Volume 3 ▪ Issue 2 ▪ May 2011

Editor-in-Chief
Dr. Min Song

INTERNATIONAL JOURNAL OF COMPUTER NETWORKS (IJCN)

ISSN : 1985-4129
Publication Frequency: 6 Issues / Year

CSC PUBLISHERS
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER NETWORKS (IJCN)

VOLUME 3, ISSUE 2, 2011

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-4129

International Journal of Computer Networks (IJCN) is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJCN Journal is a part of CSC Publishers
Computer Science Journals
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER NETWORKS (IJCN)

Book: Volume 3, Issue 2, May 2011
Publishing Date: 31-05-2011
ISSN (Online): 1985-4129

IJCN Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2011

EDITORIAL PREFACE

The International Journal of Computer Networks (IJCN) is an effective medium to interchange high quality theoretical and applied research in the field of computer networks from theoretical research to application development. This is the third issue of volume second of IJCN. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJCN emphasizes on efficient and effective image technologies, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of computer networks. Some of the important topics are ad-hoc wireless networks, congestion and flow control, cooperative networks, delay tolerant networks, mobile satellite networks, multicast and broadcast networks, multimedia networks, network architectures and protocols etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 3, 2011, IJCN appears in more focused issues. Besides normal publications, IJCN intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

IJCN give an opportunity to scientists, researchers, engineers and vendors to share the ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development. This journal is helpful for the researchers and R&D engineers, scientists all those persons who are involve in computer networks in any shape.

Highly professional scholars give their efforts, valuable time, expertise and motivation to IJCN as Editorial board members. All submissions are evaluated by the International Editorial Board. The International Editorial Board ensures that significant developments in computer networks from around the world are reflected in the IJCN publications.

IJCN editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCN. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJCN provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Computer Networks (IJCN)

TABLE OF CONTENTS

Volume 3, Issue 2, May 2011

Pages

IEEE 802.11n Based Wireless backhaul Enabled by Dual Channel IPT (DCH-IPT) Forwarding

Ehab Mahmoud Mohamed ehab@mobcom.ait.kyushu-u.ac.jp
Current: Faculty of Engineering/ Advanced Information Technology Dept/ Wireless Communication Section/Kyushu University, Japan.

Daisuke Kinoshita kinoshita@mobcom.ait.kyushu-u.ac.jp
Faculty of Engineering/ Advanced Information Technology Dept/ Wireless Communication Section/Kyushu University Motooka 744, Nishi-ku, Fukuoka-city 819-0395, Japan.

Kei Mitsunaga mitsunaga@mobcom.ait.kyushu-u.ac.jp
Faculty of Engineering/ Advanced Information Technology Dept/ Wireless Communication Section/Kyushu University Motooka 744, Nishi-ku, Fukuoka-city 819-0395, Japan.

Y.Higa higa@mobcom.ait.kyushu-u.ac.jp
Faculty of Engineering/ Advanced Information Technology Dept/ Wireless Communication Section/Kyushu University Motooka 744, Nishi-ku, Fukuoka-city 819-0395, Japan.

Hiroshi Furukawa furuhiro@ait.kyushu-u.ac.jp
Faculty of Engineering/ Advanced Information Technology Dept/ Wireless Communication Section/Kyushu University Motooka 744, Nishi-ku, Fukuoka-city 819-0395, Japan

Abstract

Wireless backhaul has received much attention as an enabler of future broadband mobile communication systems because it can reduce deployment cost of pico-cells, an essential part of high capacity system. A high throughput with a minimum delay network is highly appreciated to sustain the increasing proliferation in multimedia transmissions. In this paper, we propose a backhaul network using the Multi-Input Multi-Output (MIMO) IEEE 802.11n standard in conjunction with the Dual Channel Intermittent Periodic Transmit IPT (DCH-IPT) packets forwarding protocol. By using these two techniques (IEEE 802.11n + DCH-IPT), wireless backhaul nodes can meet more demanding communication requirements such as higher throughput, lower average delay, and lower packet dropping rate than those achieved by the currently used backhaul. The current backhaul is based upon Single-Input Single-Output (SISO) IEEE 802.11a,b,g standards in conjunction with Single Channel Conventional (SCH-Conv) relaying protocol in which packets are transmitted continuously from source nodes using single channel. The proposed backhaul will accelerate introduction of picocell based mobile communication systems.

Keywords: Wireless Backhaul Networks, IEEE 802.11n, IEEE 802.11a, MIMO-OFDM, IPT forwarding.

1. INTRODUCTION

Wireless backhaul is a wireless multihop network in which base nodes are linked wirelessly [1] [2]. Wireless backhaul has received much attention as an enabler of future broadband mobile communication systems because it can reduce deployment cost of pico-cells, an essential part of high capacity system. A high throughput with a minimum delay network is highly appreciated to sustain the increasing proliferation in multimedia transmissions. In wireless backhaul, base nodes have capability of relaying packets, and a few of them called core nodes serve as gateways connecting the wireless multihop network and the outside network (i.e. the Internet) by cables. Although wireless backhaul has many attractive features over wired backhaul networks like ATM, T1 or DSL line, it is still lack for high throughput design [3]. Recently, researchers in the field of wireless multihop try to improve its performance using MIMO [4-7].

One of the advanced wireless standards driven by MIMO technology is IEEE 802.11n (Dot11n) [8]. Dot11n is an amendment, proposed by a group in IEEE802.11 committee called TGn group, over the previous OFDM based 802.11 standards (802.11a/g) (Dot11a/g) with PHY and MAC enhancements [8] [9]. Using different space time code structures, the Dot11n's MIMO-OFDM can support data rates up to 600 Mbps which is higher than IEEE 802.11a/g (SISO standard) data rate (54 Mbps at maximum). In order to take the full advantage of Dot11n, TGn also enhances its MAC layer by introducing new frame structures that can be used to aggregate multiple subframes to improve throughput [9]. In [9], the authors extensively studied the IEEE 802.11n MAC aggregation mechanisms and their effects upon the IEEE 802.11n throughput performance.

Because Dot11n is in small age, an adoption of the interface to wireless multihop is in its infancy [5] [6]. Some studies concerned about investigating more efficient MIMO-based MAC protocol than the Dot11n's one so as to be suitable for Ad-Hoc and Wireless Mesh Networks [6] [7]. Others concerned about improving the IEEE 802.11s performance, an IEEE 802.11 standard for Wireless Mesh Networks (WMN), by using Dot11n based mesh nodes [5]. To do that, the authors in [5] first utilize the variable transmission rate property of Dot11n to find the best PHY data rate related to instantaneous channel quality, then by using this data rate, they find the best MAC aggregation number that guarantees low packet dropping rate, At last, they use theses two settings (PHY rate + MAC aggregation number) to optimally allocate bandwidth to each type of traffic [5]. According to the authors' knowledge, most of the studies adopted Dot11n utilizes its MAC and PHY enhancements to improve the performance of Wireless Mesh Networks (WMN) and Ad-Hoc networks, whereas few proposals concerned about adopting the standard to wireless backhaul. Even though all the three networks can be categorized as a wireless multihop network, wireless backhaul is the only one that can accept static tree topology routing. This is because all the nodes are fixed in their positions and all uplink and downlink packets are distributed to entire backhaul network via core nodes. Hence, the static tree topology routing is the ratified topology for the predominantly fixed infrastructure networks like wireless backhaul networks, and it is chosen for the IEEE 802.11s in case of fixed Mesh Point Portal MPP (core node in case of wireless backhaul). This can be found in the IEEE 802.11s draft, and extensively explained in [10]. On the other hand, dynamic mesh routings are preferred for WMN and Ad-Hoc because of the specific structure of these networks, i.e. multi-point to multi-point connections among all neighboring nodes and dynamic changes of nodes positions. The difference between the two routings is shown in Fig. 1. In wireless backhaul with a static tree topology routing, since we can reduce the number of intersections on its routes as found in Fig. 1, each node can maintain fewer number of connecting nodes, which will contribute to reduce complexity in necessary processes relating to MIMO signal detection, such as synchronization, channel state acquisition and so on. This feature in wireless backhaul will deliver us a larger benefit of the Dot11n adoption compared with other wireless multihop networks, i.e., WM and Ad Hoc networks.

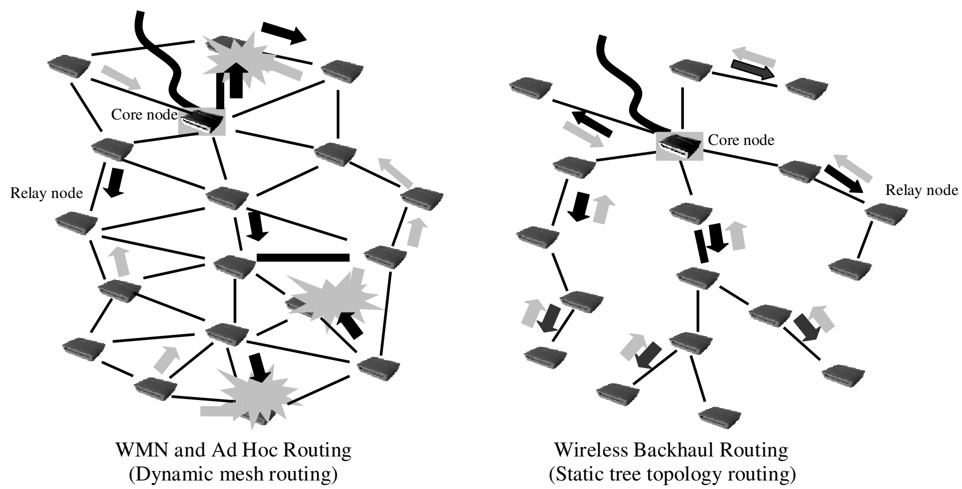


FIGURE 1: WMN and Ad Hoc Routings versus Wireless Backhaul Routing.

The application of Dot11n as MIMO wireless interface to wireless backhaul with static tree topology to enhance throughput and spectrum efficiency of its relay network was given by the authors in [11] and [12]. In this proposal, the authors compared the effectiveness of Dot11n and Dot11a based wireless backhauls under same transmission rates and bandwidth. They proved that Dot11n based wireless backhaul has higher throughput, lower average delay, and lower packet dropping rate than Dot11a based one although of the same transmission rates. In addition, they proved the robustness of Dot11n-wireless backhaul performance even in highly correlated MIMO channels environment. At the end of the research, they further improved Dot11n-wireless backhaul performance through utilizing the single channel IPT (SCH-IPT) forwarding protocol.

For further research development on the Dot11n based wireless backhaul given by the authors in [11] and [12]. And, in analogy to the IEEE 802.11 MAC enhancements to be more compatible with Dot11n high transmission rates [9], this paper concerns about studying Dot11n in conjunction with packet forwarding protocol issues for wireless backhaul, i.e., finding an appropriate packet forwarding protocol suitable for Dot11n to produce highly efficient wireless backhaul. To cope with this issue, we propose the Dual Channel relay (DCH) protocol for Dot11n transmission. Rather than Single Channel (SCH) forwarding protocol in which single channel is used for both uplink and downlink packets transmissions, our proposed Dual Channel (DCH) forwarding protocol uses two different channels for packets transmissions; one channel for uplink transmissions and the other for downlink. Therefore, fast with low interference relay is produced. Using this DCH scheme, packets are transmitted faster than using SCH because in the proposed DCH there is no waiting time between the two kinds of transmissions. Instead, simultaneous transmissions are used. In addition, this protocol mitigates the nodes congestion problem due to heavy uplink and downlink traffic. Also, using the proposed DCH forwarding protocol, two different forwarding protocols can be used for each packet type, i.e. uplink or downlink depending upon the traffic conditions [13][14], which will increase the reliability of the scheme and further fast the relay. All these DCH advantageous get it compatible with Dot11n more than SCH in which delay and interference withstand the full utilization of the high transmission rates (up to 600 Mbps) supported by Dot11n results in a low throughput performance. Dual channel strategy is used by many wireless multihop researchers in order to modify the IEEE 802.11 MAC protocol (initially designed for WLAN) to be more suitable for wireless multi-hop networks [15]-[18]. Although these Dual channel MAC protocols solve many of the IEEE 802.11 based multi-hop networks problems (hidden node, exposed node, receiver blocking, intra-flow, and inter-flow), packet transmissions from hop nodes are still one direction basis (uplink or downlink at a time). Hence, any multi-hop node can’t transmit in both directions (uplink and downlink) simultaneously. So, a big transmission delay exists in addition to the nodes congestion problem in case of heavy uplink and downlink traffic. Also, high packet dropping rate occurs due to opposite direction packets collision.

These problems are solved by our proposed DCH scheme results in a high utilization of the Dot11n high transmission rates. To prove the efficiency of the proposed DCH forwarding protocol, we compare the throughput performance between Dot11n-SCH-Conv and Dot11n-DCH-Conv in which conventional (Conv) method of relaying is used for both schemes. In contrast to the Intermittent Periodic Transmit (IPT) protocol [13], in which packets are transmitted from source nodes with a predefined transmit period, in conventional method packets are transmitted continuously from the source nodes. Also, we compare the throughput performance between Dot11a-SCH-Conv and Dot11a-DCH-Conv. The obtained results show a much better throughput performance of DCH-Conv than that achieved using SCH-Conv. In addition, simulation results ensure higher computability of DCH-Conv forwarding protocol to Dot11n (MIMO) than Dot11a (SISO) transmission. Based upon these achievements, we suggest the DCH-IPT, which is based on our previous work on wireless backhaul networks with static tree topology routing [13] [14], as a forwarding protocol suitable for Dot11n (MIMO) transmission. The combination of Dot11n and DCH-IPT produces more efficient wireless backhaul than that obtained by combining Dot11n and DCH-Conv.

The rest of the paper is organized as follows. Section 2 describes the IPT forwarding protocol. Section 3 describes the proposed DCH-IPT in more details. The simulation scenarios and performance metrics are given in section 4. Section 5 gives a comparison between SCH-Conv and DCH-Conv using Dot11a and Dot11n. A comparison between DCH-IPT and DCH-Conv under Dot11n environment is given in section 6 followed by the conclusion in section 7.

2. THE INTERMITTENT PERIODIC TRANSMISSION IPT FORWARDING

The IPT protocol is an efficient packet relay method with which a constant throughput irrespective of the node counts can be obtained [13]. Figures 2 and 3 clearly explain the difference between the conventional CSMA/CA (Conv) relaying method and IPT forwarding. In these figures, 9 nodes are linearly placed. In case of the Conv method, the source node sends packets with a random transmission period of P_Conv, i.e., as the packet appear at the source node it is immediately go through the transmission schedule, and each intermediate relay node forwards received packets from its preceding node with a random backoff period. In the case of IPT, the source node transmits packets intermittently with a certain transmission period of P_IPT and each intermediate relay node immediately forwards the received packets from the preceding node without any waiting period. In the conventional method co-transmission space, which is defined as the distance between relay nodes that transmit packets at the same time, is not fixed. In such situation, packets collisions could occur due to co-channel interference if the co-transmission space is shorter than the required frequency reuse space, as shown in Fig.2. On the other hand, in case of the IPT forwarding it can be readily understood that the co-transmission space could be controlled by the transmission period P_IPT that is given to the source node, as shown in Fig.3 in which reuse space is assumed to be 3. Reduction of the packet collisions will help to reduce re-transmissions and will consequently help to improve the system performance. In [14], the authors proposed an adaptive IPT duration in order to remove interference between co-channel relay nodes and maximize resultant throughput observed at the destination node. Figure 4 schematically shows the normalized throughput versus hop count feature of the conventional method and IPT forwarding for the systems in Fig.2 and 3. In Fig.4, constant IPT duration is applied for all slave nodes and thus the resultant throughputs are all the same [13].

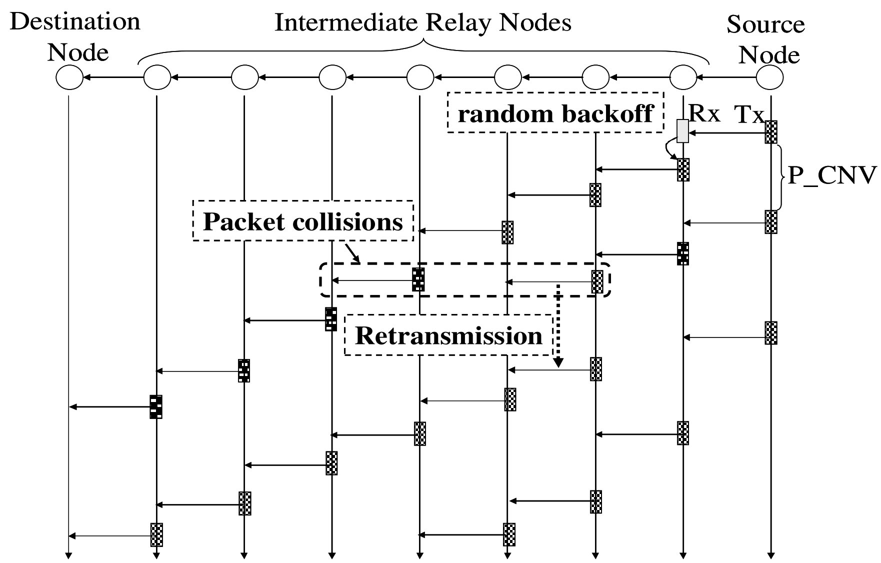


FIGURE 2: Packet relay in conventional method.

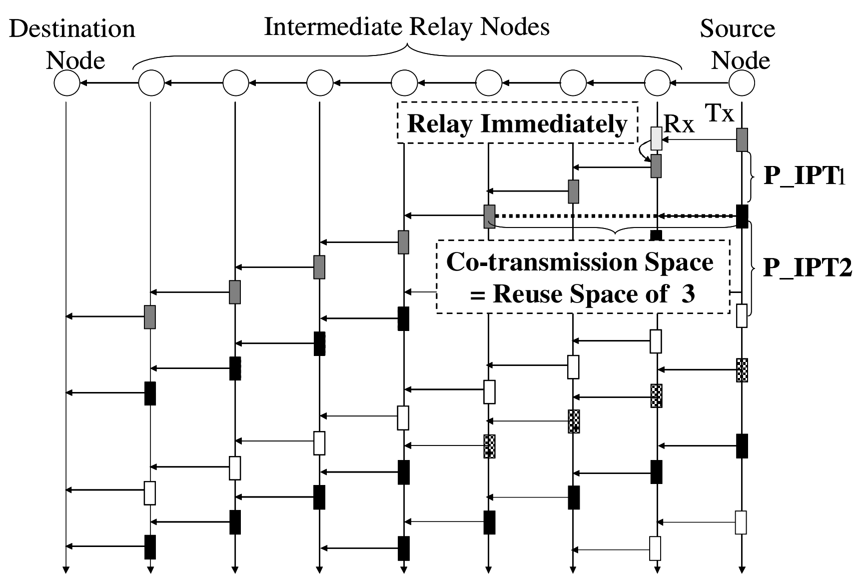


FIGURE 3: Packet relay in IPT method.

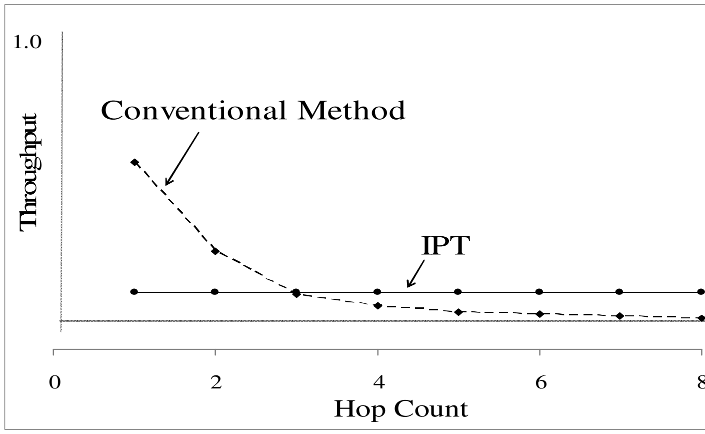


FIGURE 4: The performance comparsion between conventional method and IPT forwarding.

3. THE DUAL CHANNEL IPT (DCH-IPT) FORWARDING

As explained in the previous section, IPT forwarding is best suited for wireless backhaul networks with static tree topology routing (our main concern), where there are few number of intersections between nodes in which IPT can work efficiently. Also, the idea of IPT forwarding relies on one-way traffic either uplink or downlink transmission at a time using the same channel with predefined transmission duration of P_IPT, by which relay efficiency can be maximized. Hence, in case that downlink and uplink packets are duplexed on a single route, packets from one direction collide with ones from the other direction, which would spoil the benefit of the intermittent period transmit, and prevents the scheme from fully exploiting the high transmission data rate of Dot11n and MIMO in general. To cope with this problem and to adapt the current IPT protocol with the high data rate Dot11n-based nodes and produce high speed backhaul, we investigate the Dual-Channel IPT (DCH-IPT) relaying protocol. In this protocol, we assign two channels to transmit the relying packets simultaneously. One for uplink packets transmissions and the other for downlink packets transmissions. Figure 5 shows the DCH-IPT protocol using channel 1 for uplink packets transmissions and channel 2 for downlink packets transmissions. In this figure, we choose IPT protocol to transmit downlink packets and conventional method to transmit uplink packets. This is because downlink traffic is always heavier than uplink traffic, and IPT is more efficient than Conv method when the traffic is heavy [13]. Otherwise, when both uplink and down link traffic are heavy, IPT can be used for both kind of traffic. This flexibility will contribute in producing fast reliably relay.

Our proposed DCH-IPT can be implemented by assigning two different interfaces (radios) per relay node, one for uplink transmissions and the other for downlink transmissions, and each radio is tuned to a different channel. These two assigned channels are separated so as there will be no interference between each type of transmissions. This implementation ultimately reduces the cost of the scheme thanks to cheap Dot 11n interfaces. Although DCH-IPT uses 2 radios with double bandwidth compared to single channel IPT SCH-IPT (single radio using single channel), it contains some interesting features which get it compatible with Dot11n. By using this scheme, relay node can transmit on both channels at the same time, so we can efficiently exploit the Dot11n high transmission capability without any delay resulting from the relaying protocol. In addition, and by using this scheme, we can exploit two different relaying protocols one for each channel in relation to traffic conditions, which further improves the relay. For example, if the traffic is low, it's better to use conventional method over IPT and vice versa [13]. So by using this flexibility, we can produce a highly efficient relay. Also, packet collisions are extremely reduced through DCH-IPT. So, DCH-IPT has a much lower interference than SCH-IPT which when combined with Dot11n greatly enhances relay performance in terms of higher throughput, lower average delay and lower packet dropping rate. As a result, DCH-IPT produces a fast and efficient wireless backhaul when combined with Dot11n.

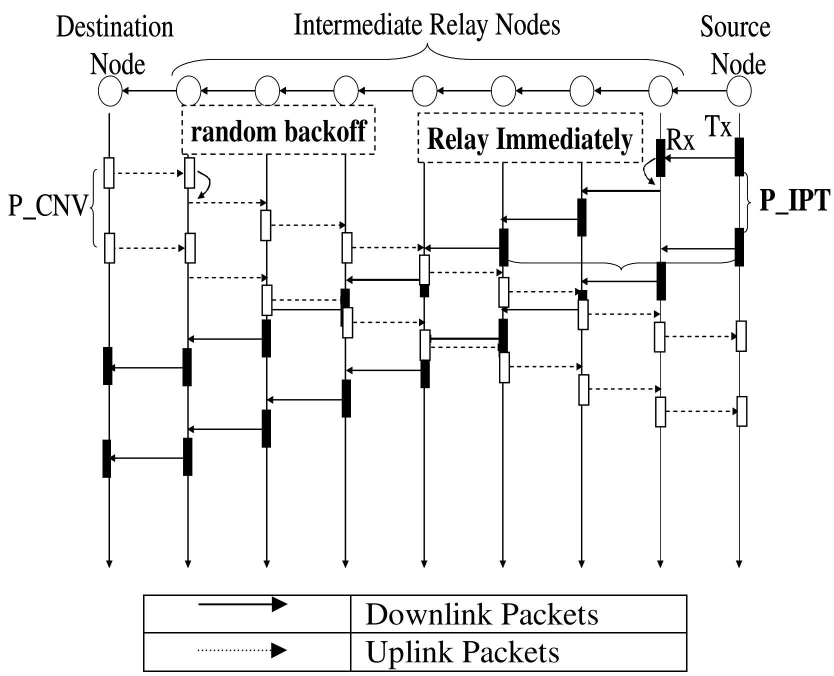


FIGURE 5: The Dual Channel IPT (DCH-IPT) forwarding.

4. SIMULATION SCENARIOS AND PERFORMANCE METRICS

To evaluate the performance of the suggested Dot11n-DCH-IPT wireless backhaul network and prove its efficiency, we evaluate PHY layer Bit Error Rate (BER) using the MATLAB program which utilized by our original network simulator to evaluate the whole network performance.

4.1 Simulation Scenarios and Parameters

4.1.1 PHY Layer

In order to prove the effectiveness of using DCH relay with Dot11n over using DCH relay with Dot11a, we compare the performance of the two PHY layers

- **IEEE 802.11n** (MIMO standard)**.**
- **IEEE 802.11a** The currently used (SISO standard)**.**

We evaluate Dot11n and Dot11a PHY performances under the following transmission rates: Dot11a 54Mbps, Dot11n 36Mbps, 48Mbps and 60 Mbps. The IEEE 802.11n 60Mbps uses high throughput specifications.

Table 1 shows Dot11n and Dot11a PHY layers simulation parameters.

Parameter	Dot11n	Dot11a
Number of Data Subcarrier	48/(108)	48
IFFT Size	64/(128)	64
Cyclic Prefix length	16	8
Pilot Subcarriers/ Symbol	4/(6)	4
QAM mapping	QPSK-16QAM/(16QAM)	64 QAM
Transmitter antennas	2	1
Receiver antennas	2	1
FEC Rate	¾, 0.5/(0.5)	¾
Raw Data Rates	36, 48 M/(60Mbps)	54 Mbps
MIMO Detector /channel equalizer	SOMLD (Soft Output Maximum Likelihood)	
Channel Estimation	Perfect	
Synchronization	Perfect	
Quasi-Static Channel model Rayleigh fading (NLOS) with exponential PDP (Power Delay Profile) indoor or outdoor (large open space) case T_{max}(MAX delay spread)=300ns, T_{rms} (RMS delay spread)=150ns		

TABLE 1: PHY layers parameters

4.1.2 MAC Layer

For MAC operation, the operation mode of CSMA/CA, the MAC standardized by IEEE 802.11, dynamically changes in between the Basic and RTS/CTS modes depending on message transmit method and IPT activation: when IPT forwarding is carried out, the Basic mode is applied otherwise the RTS/CTS mode is chosen. When the simulator is turned to evaluate the Dot11n 60Mbps performance, a 4-packet A-MSDU MAC aggregation is used [9]. Otherwise in the Dot11n 36 and 48 Mbps non high-throughput modes, no packet aggregation is used.

4.1.3 Traffic Model

Downlink traffic directed to terminals that stay under base nodes is all generated at a core node and forwarded to each base node. Uplink traffic caused by terminals is gathered at the base node in which the terminals stay and forwarded to a core node. The Poisson process is employed as a traffic model. The number of data packets per session is randomly determined by the log-normal distribution, the mean of which is 20 for downlink and 3 for uplink. The ratio of the total offered load of downlink to uplink is 10:1 [19].

4.1.1 Packet forwarding methods

In order to prove the efficiency of DCH-Conv wireless backhaul over SCH-Conv, and also, in order to prove the efficiency of DCH-IPT over DCH-Conv, we compare the following relaying methods with the same forwarding path shown in Fig 4.

- **SCH-Conventional method (SCH-Conv)-** packets are transmitted continuously, using single channel for both uplink and downlink transmissions, with minimum path loss with RTS/CTS MAC mode for all transport sessions.
- **DCH-Conventional method (DCH-Conv)-** packets are transmitted continuously, using two channels one for uplink and the other for downlink transmissions, with minimum path loss and RTS/CTS MAC mode for all transport sessions.
- **Dual Channel IPT (DCH-IPT) protocol-** in this scheme, we use IPT protocol with basic mode MAC for downlink packets, and conventional method with RTS/CTS MAC for uplink

packets. We use the conventional method for uplink packets because the conventional method is better than IPT when the traffic is low [20].

4.1.5 Evaluation Site

We chose the floor of our department building as a test site Fig 6. In order to handle a complex interference situation as correctly as possible, we use a simple deterministic radio propagation model such as a path loss coefficient of 2 dB until 5m and 3.5 dB beyond this distance [21], 12 dB penetration loss of the wall [22]. 23 base nodes are placed on the floor and a core node is placed on stairs area of the floor Fig 6. A forwarding path is formed in advance and fixed during simulation Fig 6.

Spectrum assigned to the wireless repeater network is assumed to be different from one assigned to wireless communication links between mobile terminals and base station (access network), so interference between access network and repeater network can be excluded.

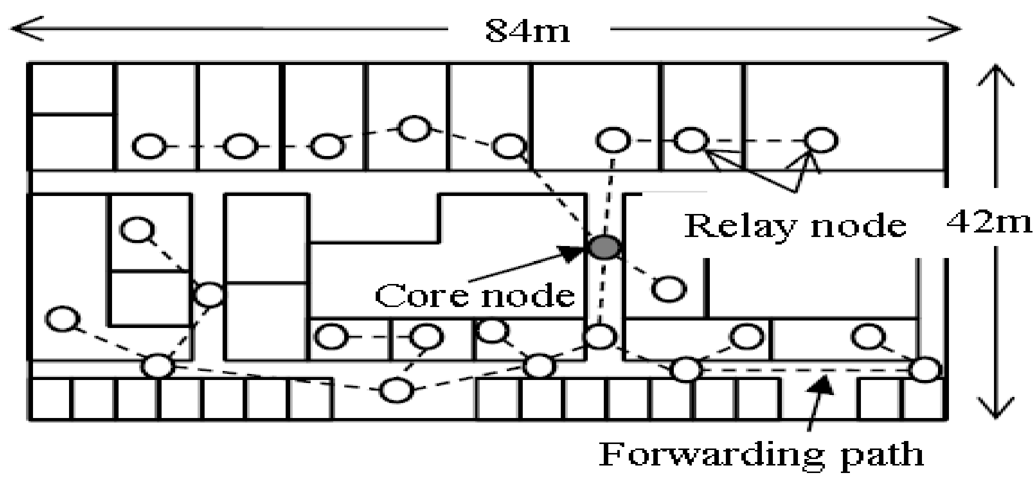


FIGURE 6: Floor plan and node layout for evaluations.

4.2 Performance Metrics

4.2.1 PHY Layer Simulator

We evaluate the BER (Bit Error Rate) performance of each tested PHY layer, i.e., Dot11a 54, Dot11n 36, 48, and 60 Mbps.

4.2.2 System Level Simulator

Three performance metrics are used: Aggregated end-to-end throughput, Average delay, and packet loss rate.

a. ***Aggregated end-to-end throughput*** is defined as sum of throughputs for all sessions each of which successfully delivered to a destination.
b. ***Average delay*** is defined as an average time period from the instant when a packet occurs at a source node to the instant when the destination node completes reception of the packet.
c. ***Packet loss rate*** is defined as follows:
   Packet loss rate [%] $= ND*100 / (NS + ND)$.
   Where *NS* denotes the number of packets received successfully by destination nodes.
   ND denotes the number of discarded packets due to exceeding a retry limit

Each simulation is carried out for 200 sec; this simulation period has been ensured to achieve a good convergence. Also, we assume UDP traffic.

5. COMPARISON BETWEEN SCH-CONV AND DCH-CONV USING DOT11A AND DOT11N

Mont Carlo simulations are carried out for evaluating the comparison between SCH-Conv and DCH-Conv based wireless backhauls.

5.1 PHY Layer Simulator

Figure 7 shows the BER performance of the compared PHY layers (Dot11a 54 Mbps and the high-throughput Dot11n 60 Mbps). Form this figure; we notice the extremely enhanced BER performance of Dot11n over Dot11a counterpart although of nearly same transmission rate. This Dot11n better BER performance comes from using multiple antennas at both transmitter and receiver (MIMO) which is not the case for Dot11a (SISO). By using MIMO, Dot11n uses lower MCS (Modulation Coding Scheme) than Dot11a to obtain the same transmission rate under the same bandwidth. For example, in order to obtain a transmission rate of 36 Mbps for both Dot11a and Dot11n under the same bandwidth of 20 MHz, Dot11a uses 16-QAM with FEC=3/4, but 2x2 MIMO Dot11n uses QPSK with FEC=3/4. This Dot11n's MCS reduction resulting from using MIMO greatly enhances its BER performance.

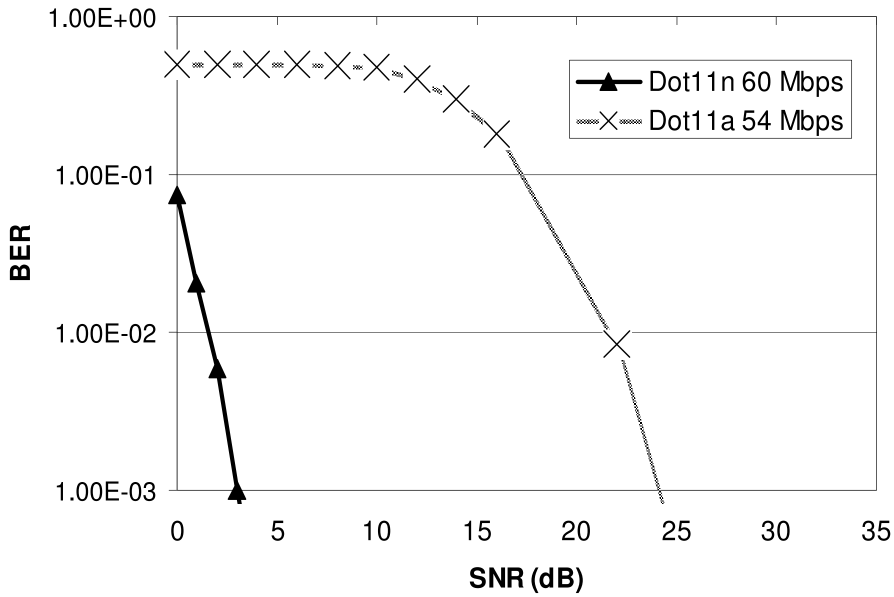


FIGURE 7: BER performances of Dot11n and Dot11a.

5.2 System Level Simulator

Using the BER performances of Dot11a 54 Mbps, Dot11n 60 Mbps evaluated in section 5.1, we compare their system level performances using SCH-conventional and DCH-conventional methods of relaying. Figure 8 shows the aggregated end to end throughput resulting from this simulation.

First, we notice that the throughput performance of Dot11n 60-SCH-Conv is about 50% larger than Dot 11a 54-SCH-Conv although Dot11n 60 transmission is only 6Mbps higher than Dot11a 54 transmission. This Dot11n higher performance comes from its better BER performance as explained in section 5.1. Better BER performance means that for a certain required PER (Packet Error Rate) performance value, Dot11n operates at a lower SINR (Signal to Interference Ratio) value than Dot11a, which gives Dot11n robust characteristics against interference, i.e., Dot11n has a higher tolerance to interference than Dot11a. This important Dot11n phenomenon has a great impact on system performance in which interference causes a significant degradation in performance. In addition, the MAC Aggregation process used in Dot11n, which is not Dot11a case, causes a lower network delay than that occurred using Dot11a especially for high Dot11n rates like 90, 180Mbps, etc. These facts are reflected on the system level performance in terms of higher throughput than that achieved by Dot11a based one as revealed by this figure. These simulations support the idea of enhancing wireless backhaul performance using Dot11n-based nodes previously stated by the authors [12].

In addition, from this figure, we can notice the improved performance of the Dot11a 54-DCH-Conv / Dot11n 60 DCH-Conv over the Dot11a 54-SCH-Conv / Dot 11n 60-SCH-Conv respectively. This improved DCH-Conv performance comes from using two channels relaying protocol, one for uplink and the other for downlink packets transmissions, which fasts the relay compared to the SCH-Conv relay, in which uplink or downlink packets can be transmitted at a time. Also, by using Dual Channel technique, their will be less interference occurred between uplink and downlink packets which contributes in enhancing the wireless backhaul performance.

Furthermore, it is clear that the achieved throughput improvement using Dot11n-DCH-Conv is about twice the achieved improvement using Dot11a-DCH-Conv. This is because Dot11n has a better PER (Packet Error Rate) than Dot11a in addition to the MAC Aggregation scheme used in Dot11n. These features require a low interference environment to work efficiently. Dual Channel protocol has lower interference than Single Channel one, where uplink and downlink packets collisions greatly affect Single Channel performance. So, the combination of Dot11n and Dual Channel protocol gives us a much better achieved performance improvement than that achieved by combining Dot11a and Dual Channel protocol.

In conclusion, Dot11n (MIMO+MAC Aggregation) is a promising wireless interface in enhancing the wireless backhaul performance which requires a good relaying protocol that can fully exploit this powerful interface. Dual Channel relaying protocol is the convenient relaying protocol which when cooperates with Dot11n extremely enhances wireless backhaul performance.

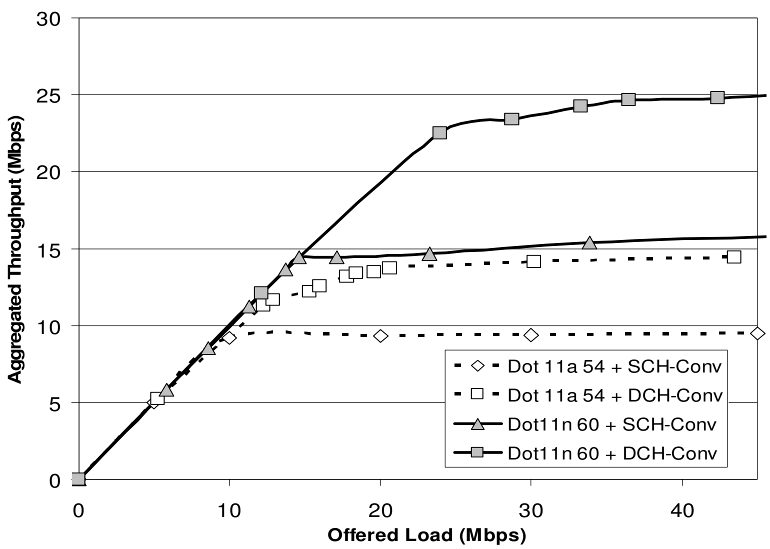


FIGURE 8: Comparison between SCH-Conv and DCH-Conv under Dot11a and Dot 11n.

6. COMPARISON BETWEEN DCH-IPT AND DCH-CONV USING DOT11N

In these simulations, by using the BER of the non high-throughput Dot11n 36 and 48Mbps shown in Fig 9, we compare the system level performances of theses PHY layers with DCH-Conv and DCH-IPT relaying protocols. Figures 10-12 show the aggregated end-to-end throughput, Average Delay, and Packet Loss rate of this comparison. The simulation results ensure better performance of the DCH-IPT based network over DCH-Conv based one. About 20 % improvements in throughput and average delay result from using DCH-IPT relaying. In addition, the packet dropping rate is extremely enhanced, about 10^{-3} reduction in packet loss rate. This enhanced DCH-IPT performance comes from the interference rejection resulting from the intermittent packets transmissions introduced by the IPT, which solves the hidden terminals problems and greatly enhances relay performance. Figures 13-15 compare the proposed Dot11n

(MIMO)-DCH-IPT wireless backhaul with the currently used Dot11a (SISO)-SCH-Conv one using the 48Mbps PHY layer. About 40 % improvements are obtained in the throughput and Average delay using the proposed Dot11n-DCH-IPT backhaul. Also, very large improvement occurred in packet dropping rate results from using the high inference tolerances Dot11n and DCH-IPT schemes. These results verify the effectiveness of the Dot11n-DCH-IPT based wireless backhaul as a key enabler for the next wireless mobile communication generation.

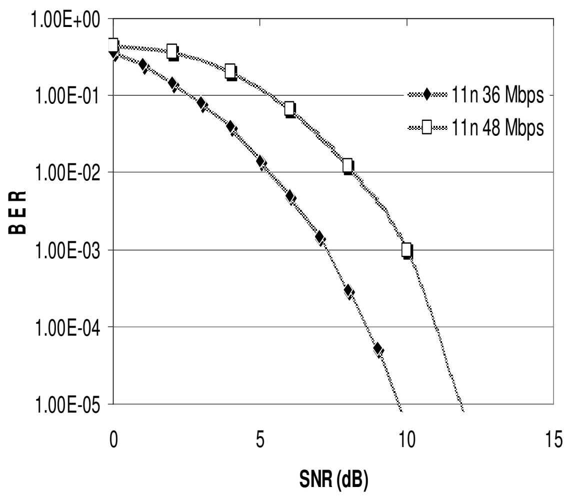


FIGURE 9: BER performances of Dot11n 36 and 48 Mbps.

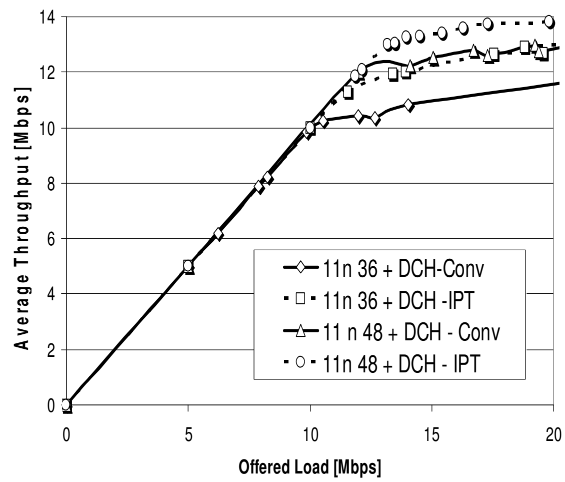
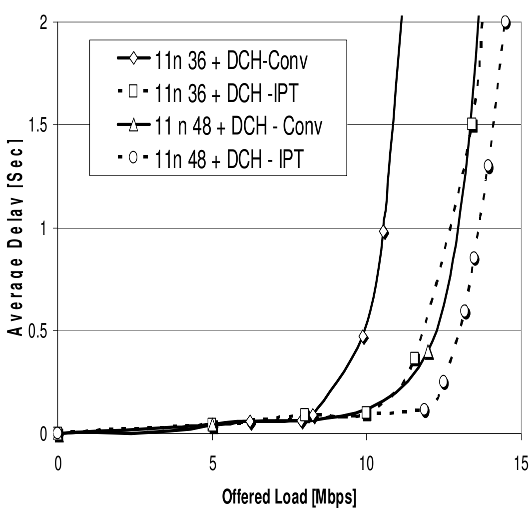


FIGURE 10: Aggregated end-to-end



FIGUER 11: Average Delay.

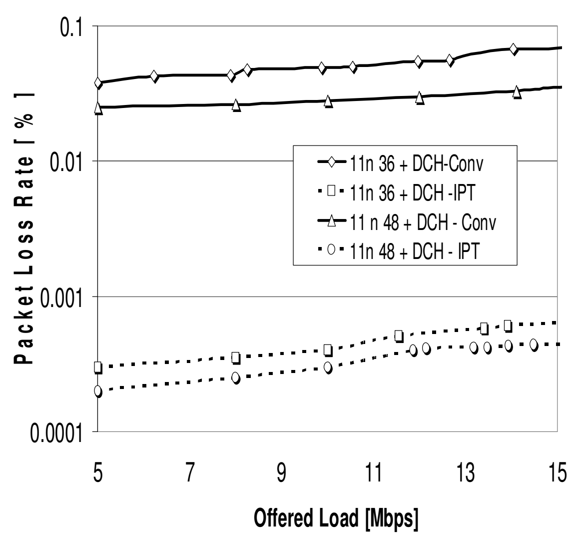


FIGURE 12: Packet Loss Rate.

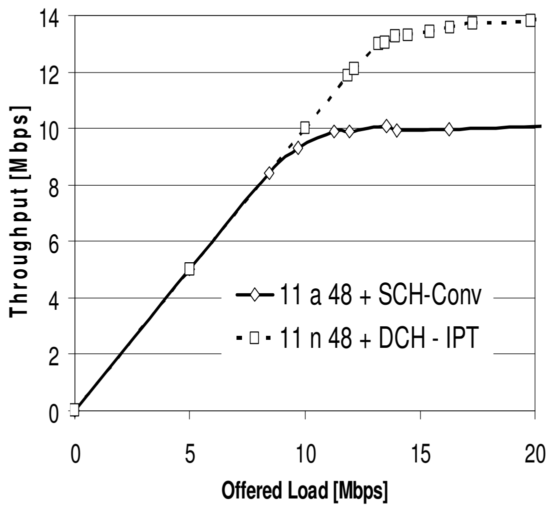
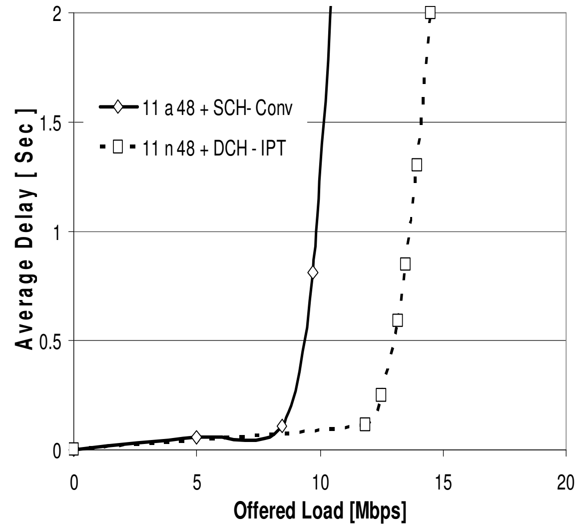


FIGURE 13: Aggregated end-to-end throughput.



FIGUER 14: Average Delay.

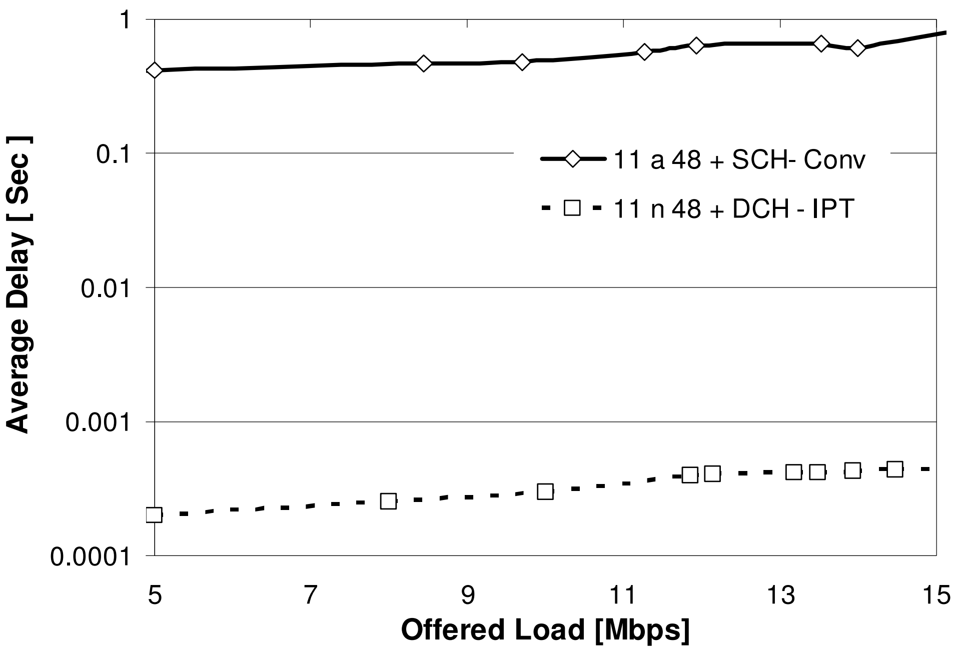


FIGURE 15: Packet Loss Rate.

1. CONSLUSION & FUTURE WORK

The IEEE 802.11n MIMO-OFDM standard is a promising technique for the next generation wireless backhaul networks. The application of this standard as a wireless interface in these networks needs an efficient relaying protocol to exploit its high speed capabilities. We showed that Dual Channel (DCH) relaying protocol is an efficient relaying protocol suitable for Dot11n-based backhaul, and we proved that the achieved throughput improvement of Dot11n-DCH is more than those achieved by Dot11a-DCH with conventional method of relaying. In addition, and in order to further improve this Dot11n based wireless backhaul, we adopt the IPT forwarding protocol to be compatible with the Dot11n standard by introducing DCH-IPT. We proved the effectiveness of the Dot11n-DCH-IPT based backhaul over the Dot11n-DCH-Conv based one, and over the currently used Dot11a-SCH-Conv one.

For further investigations, we will try to optimize the spectral efficiency of DCH-IPT and produce an optimized relying protocol for MIMO (Dot11n) transmissions.

7. REFERENCES

[1] S. Romasszko and C. Blondia, “Cross layer PHY-MAC protocol for wireless static and mobile Ad Hoc networks,” In Eurasip Journal on Advances in Signal Processing 2009, art. No. 278041, 2009.

[2] R. Pabst et al, “Relay based deployment concepts for wireless and mobile broadband radio,” In IEEE Communication Magazine, Vol. 42, No.9, pp.80-89, 2004.

[3] G. Narlikar, G. Wilfong, and L. Zhang, “Designing Multihop Wireless Backhaul Networks with Delay Guarantees,” In proc INFOCOM 2006. 25th IEEE International Conference on Computer Communications, pp.1-12, 2006.

[4] Mohinder Jankiraman, “Space-Time Codes and MIMO Systems,” Book, Cambridge university press.

[5] A. Ashtaiwi and H. Hassanein, “Utilizing IEEE 802.11n to enhance QoS support in wireless mesh networks,” In Local Computer Networks LCN 2008. 33rd IEEE Conference, pp. 689-696, 2008.

[6] A. Ashtaiwi and H. Hassanein, “Rate Splitting MIMO-based MAC Protocol,” In Proc Local Computer Networks, 2007. LCN 2007.

[7] K. Sundaresan, R. Sivakumar, M. A. Ingram, and T. Chang. “Medium access control in ad hoc networks with MIMO links: optimization considerations and algorithms,” In IEEE Transc. Mobile Computing, Vol.3, No.4, pp. 350-365, 2004.

[8] T. K. Paul and T. Ogunfunmi, “Wireless LAN Comes of Age: Understanding the IEEE 802.11n Amendment,” In IEEE Circuits and Systems Magazine, Vol.8, No.1, pp. 28 -54, 2008.

[9] D. Skordoulis, Q. Ni, H. H. Chen, A. P. Stephens, C. Liu, and A. Jamalipour, “IEEE 802.11n MAC frame aggregation mechanisms for next-generation high-throughput WLANs,” In IEEE Wireless Communications Magazine, Vol. 15, No. 1, pp. 40–47, 2008.

[10] J. D. Camp and E. W. Kinghtly, “The IEEE 802.11s Extended Service Set Mesh Networking Standard,” In the IEEE Communication Magazine, Vol. 46, No. 8, pp. 120-126, 2008.

[11] E. M. Mohamed, D. Kinoshita, K. Mitsunaga, Y. Higa and H. Furukawa, “An Efficient Wireless Backhaul Utilizing MIMO Transmission and IPT Forwarding,” International Journal of Computer Networks, IJCN, Vol. 2, No. 1, pp. 34-46, 2010.

[12] E. M. Mohamed, D. Kinoshita, K. Mitsunaga, Y. Higa and H. Furukawa, “MIMO Wireless Backhaul Enabled by IPT Forwarding,” In proc 17th ICT 2010, pp 233-240, 2010.

[13] H. Furukawa, “Hop Counts Independent Throughput Realization by a New Wireless Multihop Relay,” In Proc. VTC’04-Fall, 2004.

[14] Y. Higa and H. Furukawa, “Time Interval Adjustment Protocol for the New Wirless multihop Relay with Intermittent Periodic Transmit” In IEICE Transc., B-5-180, 2004.

[15] H. Zhai, J. Wang, and Y. Fang, “DUCHA: A New Dual-channel MAC Protocol for Multihop Ad Hoc Networks,” In IEEE Transc. on Wireless communications, Vo1. 5, No. 11, pp 3224 – 3233, 2006.

[16] S. Wu, Y. Tseng, and J. Sheu, “ Intelligent medium access for mobile ad hoc networks with busy tones and power control,” In IEE J. Select. Areas Commun., vol. 18, No.9, pp. 1647-1657, 2000.

[17] Z. J. Hass and J. Deng, “ Dual busy tone multiple access (DBTMA)-A multiple access control for ad hoc networks,” in IEEE Transc. Commun., Vol. 50, pp. 975-985, 2002.

[18] Z. J. Hass and J. Deng, “ Dual busy tone multiple access (DBTMA): performance results,” In Proc. IEEE WCNC, Vol. 3, pp. 1328-1332, 1999.

[19] K. Fukuda, K. Cho, and H. Esaki, “The Impact of Residential Broadband Traffic on Japanese ISP Backbones,” In Computer Communication Review, Vol. 35, pp. 15-22, 2005.

[20] H. Furukawa, Y. Higa, R.-G. Cheng and J.-S. Yang, “Proposed text for frame forwarding on 802.11s,” In IEEE802.11-05/770r0, Jun. 2005.

[21] IEEE 802 11-03/161r2, TGn Indoor MIMO WLAN Channel Models, 2004.

[22] ITU-R WP3K, “Draft revision of recommendation ITU-R P.1238; Propagation data and prediction models for the planning of indoor radio communication systems and radio local area networks in the frequency range 900MHz to 100GHz,” ITU-R Document 3/53, Mar. 1999.

Performance Model of Key Points At the IPTV Networks

Snezana Tilovska *snezana.tilovska@telekom.mk*
Makedonski Telekom
Kuzman Josifovski Pitu bb, Stip, 2000, Republic of Macedonia

Aristotel Tentov *toto@feit.ukim.edu.mk*
Computer Science Department
Faculty of Electrical Engineering and Information Technologies
University "St. Kiril and Metodij"
Rugjer Boshkovikj II bb,Skopje, 1000, Republic of Macedonia

Abstract

In this paper we propose a new analytical model for modeling of the key points at the IPTV networks. This model uses Gamma distribution with Intergroup Characteristics for modeling self similar nature of processes in key points of IPTV network.
Enclosed Gamma Distribution results are compared with results from real measurements. Calculated discrepancies confirm that enclosed analytical model is optimal estimation model for process modeling of the key points at the IPTV network.
The used methodology for real-time analyses of the key points at the IPTV Network is very important for achieving IPTV service best performance.

Keywords: IPTV Network, Measurements, Monitoring, Analytical model, Real-time measurements.

1. INTRODUCTION

This research analyses measurement parameters for a key points of IPTV network over the period March 2008 - June 2010. The data includes information about the stream outage distribution, delay factor and stream loss in key points of overall IPTV Network defined in hours and minutes per each day in a sample defined time periods.

Categories that included equipment failure and human error are not analyzed into the research. In addition, information about the total number of customers served by the affected IPTV network, as well as total service population and service population density of the state affected in each point, also was not direcly analyzed into the research.

The resulting database included information about 2.010 values over this period. The database was used to carry out analyses for stream outage distribution, delay factor and stream loss in key points of overall IPTV Network per given time interval.

The research encloses mathematical model of Gamma Distribution with Intergroup Characteristics and networking model for typical IPTV Network.
We perform comparison analyses of the values of stream outage distribution, delay factor and stream loss in key points of overall IPTV Network.

With comparison analyses we found 2.32% disrcapances for delay factor into the Back Up Pop Router, 5.83% for stream loss pparameter into the Head-End eqipment, 3.56% for stream outage distribution parameter into the Head-End Eqipment, 5.02% for first iteration measurement stream loss parameter into Core Router, 3.67% for second iteration measurement stream loss parameter into Core Router.

Also, the enclosed mathematical model is compared with other models that are user in similar resrches into the field. There were not founded any similarity with other already developed models in this field.

2. IPTV ARHITECTURE

A typical IPTV network is consisted with functional blocks (see figure 1 defined below), [16], [17]:

- National head-end: Where most of the IPTV channels enter the network from national broadcasters
- Core network: Usually an IP/MPLS network transporting traffic to the access network
- Access network: Distributes the IPTV streams to the DSLAMs
- Regional head-end: Where local content is added to the network
- Customer premises: Where the IPTV stream is terminated and viewed.

2.1 Factors that Affect Service into the Network

Main factors that can be measured into the IPTV network and have important influence on the quality of the services into the network are:

(A) Encoding and Compression, The quality of the video being distributed across the network can be affected right at the source; i.e., at the video head-end. The encoding and compression process usually creates a trade-off between the quality of the video and the desired compression level, [17].

(B) Jitter, Defined as a short-term variation in the packet arrival time, typically caused by network or server congestion. If the Ethernet frames arrive at the STB at a rate that is slower or faster, as determined by the network conditions, buffering is required to help smooth out the variations. Based on the size of the buffer, there are delivery conditions that can make the buffer overflow or underflow, which results in a degradation of the perceived video.

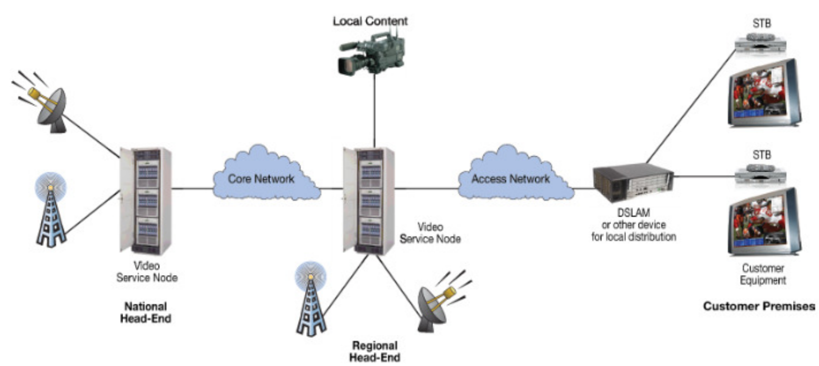


FIGURE 1: General IPTV Network Topology

(C) Limited Bandwidth, as core IP infrastructure is usually based on optical networks with a low level of congestion, bandwidth limitations (and the total amount of video-stream data that can be sent) is limited mostly by the access network or the customer's home network supported rate. When traffic levels hit the maximum bandwidth available, packets are discarded, leading to video quality degradation.

(D) Packet Loss, Loss of IP packets may occur for multiple reasons—bandwidth limitations, network congestion, failed links and transmission errors. Packet loss usually presents a busty behavior, commonly related to periods of network congestion.

2.2 Quality of Experience (QoE)

Due to the structure of Ethernet and IP networks, the quality of the video/audio traffic is primarily influenced by network jitter and packet loss. With the type of video encoding that is used in MPEG or other similar compression algorithms, the actual impact to the user perception depends

on the packet type that is lost in the network. In MPEG-2, the transported packets that are used to form an image are divided into I-frames, P-frames and B-frames. In simple terms, I-frames contain a complete image, while P-frames and B-frames contain predicted information from the other frames.
Figure 2 provides a sample of the relationships between the various types of frames included in a group of picture (GOP). As shown, I-frames are independent and provide input to support the other frames; this means that an error in the I-frames will have more repercussions to the image being viewed than losing P-frames or B-frames, [17].

2.3 Key Points into IPTV Networks

There are four major locations that need to be monitored for IPTV service delivery into IPTV network. Generally demarcation points in each of the four areas must be defined considering the different groups and units within a service provider's organization that will handle issues based on where they are found.
First Network Monitoring point is "Head-End" monitoring point. The first monitoring location is between the traditional cable head-end, and the network interface.

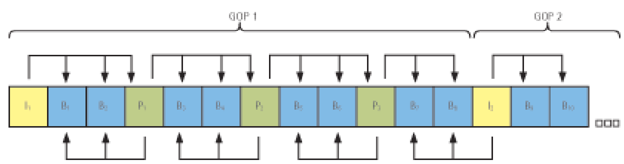


FIGURE 2: Typical group of picture (GOP) relationship in MPEG.

The second monitoring point is network transport monitoring. This would typically be monitored at the VHO/Regional center egress from the transport network and represents issues that can occur in the transport of video from the SHO to the VHO.
Third Monitoring Point is "Last Mile" monitoring point. The third monitoring point is the VDSL or FTTH line. Some T&M can be done at the egress from the VHO or at a remote DSL cabinet with the use of a permanent or longer term leave behind test device. In addition to IPTV measurements with a measurement probe in the DSLAM qualification of the copper lines can be completed without having a technician "onsite".
The last demark being the actual Set-Top-Box, residing behind the Residential Gateway, and connected through a variety of LAN technologies (Ethernet, MOCA, HPNA, etc.).

2.4 Measurements With Probes

There are several recommended ways for measurements into the key points, [18]:

- Option One: In-line equivalent passive monitoring - Passive monitoring is the safest method for adding a probe because the probe cannot impact service. If a probe is required to be in-line then an in-line equivalent can be setup using two mirror ports on a router.
  Advantages: Provides an equivalent setup to in-line, but assures passive behavior meaning the probe will never impact the service. Disadvantages: Requires two ports on a router/switch. Depending on the router/switch mirror ports can be service impacting.
- Option Two: Passive Monitoring, Potentially Measurement Impacting - The second option is to mirror both ingress and egress of a line under test to a single mirror port, likely on two different VLANs using distinct VLAN IDs.
  Advantages: Provides all line data as if in-line, and assures passive behavior meaning the probe will never impact the service. Disadvantages: Because you are mirroring a full duplex line onto a single direction of a mirror port line, there is a potential for congestion on the mirror port that could impact measurement.
  From a probe perspective, this doesn't look exactly like an in-line setup so the probe must support this configuration (i.e. support measurement VLAN's) or the measurement

of concern must not be directional in nature (for example, getting a “program ID” wouldn’t matter).

- Option Three: Active, but no uncast monitoring - Sometimes passive monitoring is not possible. For example if you would like to test the response of a D-Server to a channel change request you would need to join the service. Any multicast can be joined to provide semi-passive monitoring of multicast streams. In this case you can setup a probe to be a part of the service (e.g. in the VPLS domain of the service similar to how an end-user would be attached if directly Ethernet connected).

3. INTRODUCING THE MODEL OF SELF-SIMILAR PROCESSES FOR PERFORMANCE MODELING OF KEY POINTS OF THE IPTV NETWORK

In this study Self-Similar process with Gamma Distribution is introduced as a new approach for analyzes of key points of IPTV Network.
Self-similar processes are types of stochastic processes that exhibit the phenomenon of self-similarity, [15]. A self-similar phenomenon behaves the same when viewed at different degrees of magnification, or different scales on a dimension (space or time). Self-similar processes can sometimes be described using heavy-tailed distributions, also known as long-tailed distributions. Example of such processes includes traffic processes such as packet inter-arrival times and burst lengths. Self-similar processes can exhibit long-range dependency.

The design of robust and reliable networks and network services has become an increasingly challenging task in today's world. To achieve this goal, understanding the characteristics of traffic plays, performance models, bottlenecks into the networks, throughput, specific functions of the packets that are transmitted over the network, key points into the network is more and more critical role. Empirical studies of measured traffic traces have led to the wide recognition of self-similarity in network traffic.

3.1 Fractals in Self Similar Processes

B. Mandelbrot introduced the term ‘fractal’ for geometrical objects: lines, surfaces and spatial bodies having a strongly irregular form, [1]. These objects can possess the property of self similarity. The term ‘fractal’ comes from the Latin word fractals and can be translated as fractional or broken. The fractional object has an infinite length, which essentially singles it out on the traditional Euclidean geometry background. As the fractal has the self-similar property it is more or less uniformly arranged in a wide scale range; i.e. there is a characteristic similarity of the fractal when considered for different resolutions. In the ideal case self-similarity leads to the fractional object being invariant when the scale is changed. When self-similar traffic is mentioned, it will be assumed that its time realizations are fractals.

3.2 Self Similarity Processes in Telecommunication Networks

Modern investigations show that self-similarity can occur as a result of combining separate ON/ OFF sources, which can be strongly changeable (i.e. ON and OFF periods have DHT and infinite variances, e.g. Pareto distribution). In other words, the superposition of the ON/OFF sources exhibiting the infinite variance syndrome results in self-similar combined (aggregated) network traffic, tending to fractional Brownian motion. Moreover, the research of various traffic sources shows that the highly changeable ON/OFF behavior is the typical characteristic for the client–server architecture, [1].

3.3 Characteristics of Video Traffic

Video signals can be analyzed as a sequence of continuous pictures or as e sequence of ‘frames’. Each fixed picture should be presented by the coding algorithm in the digital form and should be compressed to decrease the bandwidth. Normally, in order to decrease the bandwidth, typical transmission of video signals cover the transmission of the full initial frame and then transmission of the difference frames. This method of transmission is called inter frame coding.

Since the ongoing frames differ little from each other (for the motion is continuous) to avoid transmission errors it is possible to ensure the periodic transmission of the full frame, where varying the frames no longer depends on the previous frames.
For these and some other reasons video traffic differs from broadband data traffic and, therefore, the models and the conclusions revealed for video data cannot be used for other traffic types, [1].
Every frame has MPEG coding. The MPEG standard uses three modes for frame coding. They are called intra frame (I), predicted (P) and interpolating (B). The I frame is the JPEG (Joint Photography Experts Group) coding of a separate frame (i.e. without using time redundancy).
During coding the coding template is usually given, i.e. an accurate sequence determining the time moments of the full frame arrival. This template is referred to as the GOP (group of pictures) and appears to be self-sufficient for decoding the frame sequence.
From the video sequence correlation function can be seen that MPEG coding introduces strict periodicity. To avoid this periodicity MPEG data can be grouped into blocks of 12 frames, called the GOP.

Since the frames inside this template do not differ very much from each other (only the difference between them is transmitted), this leads to the existence of an essential correlation of their sizes. During transmission of the next full frame the correlation between them practically disappears. For this reason the video traffic differs considerably from the usual traffic of the telecommunication network. Therefore, the conclusions and the models obtained for the usual network traffic cannot be applied for the analysis and modeling of video traffic.

3.4 Model for Intergroup Characteristics

The intergroup traffic character can be described by the first-order and second-order statistical characteristics of I frame processes, [1]. Gamma distribution is a good approximation of the I process:

$$F_{X_I}(r) = \frac{r^{m_I - 1}}{\Gamma(m_I) l_I^{m_I}} e^{-r/l_I}, \forall r > 0$$

Where m_I is the shape parameter and l is the scaling coefficient. They are related by the m value and the variance σ^2_I of the I frame trace using the following relations:

$$m_I = \frac{\sigma_I^2}{\mu_I^2} \quad \text{And} \quad l_I = \frac{\sigma_I^2}{\mu_I}$$

The I frame trace has self-similar properties and can be characterized by the SRD parameter I_I, LRD parameter H_I and the ‘boundary parameter’ K_I. Thus the autocorrelation function calculated as has the form

$$R_{X_I X_I}(m) = \begin{cases} e^{-\lambda_I m}, m \le K_I \\ L m^{-\beta_I}, m > K_I \end{cases}$$

where $\beta_I = 2 - 2H_I$. The same procedure can be used to describe P and B frame distributions.

Can be shown that the ACF of I processes for analyzed sequences has two different characteristics: the self-similar character (long-range dependence), described by the Hurst exponent H_I, and exponential decay similar to the function $e^{-\lambda_I x}$ over short time intervals.
Two regions are divided by coefficient K_I characterizing the boundary. For example, in the case of the cartoon the exponent H_{Ij}=0.873, λj=0.891 and coefficient K_{Ij} = 30 frames. The same character is typical for the correlation functions of B and P frames.
Taking into account that the gamma distribution is fully characterized by the mean and the variance, it is necessary to analyze the statistical characteristics of the processes reflecting the B and P frame size distributions for the given I frame sizes.

This model reflects two main statistical characteristics of the real video sequence: the quotient distribution with the heavy tail and the long-range part of the autocorrelation function.

4. EXISTING MODELS FOR ANALYZES OF IPTV NETWORK PARAMETERS

Into research area can be found other approaches for performing analyzes for different topics of IPTV Networks, IP Networks, User Behavior, WWW Traffic, VBR Performance Analyses, etc. [9], [10], [11], [12], [13], [14], [19], [20], [21]. Analyzes for performance parameters into key points of IPTV Network with Self-Similar Process with Gamma Distribution can be found into the research area in the World for the first time.

Previously was described the managing of the IPTV service (performance monitoring data such as device usage and error logs, user activity logs, detail network alarms and customer care tickets) with design of Giza. In this research was presented the first characterization study of faults and performance impairments in the infrastructure of a large IPTV service provider in North America. Analysis are performed from the spanned routers in the backbone to set top boxes (STB) and residential gateways (RGs) in home networks, including hardware and software crashes to video quality impairments. This research was covered on University of Texas and AT&T Lab, [3].

Another approach was research for modeling of channel popularity dynamic in large IPTV systems with Zipf-like model for distribution. In that research, was analyzed and modeled channel popularity based on user channel access data in a nation-wide commercial IPTV system. Into the document was found that the channel popularity is highly skewed and can be well captured by a Zipf-like distribution, [4].

User Behavior for CUTV service via STB devices is modeled with non stationery Poisson process with Catching Algorithm. This paper shows that transport capacity requirements of the network supporting interactive television services risks growing enormously, if these services continue to gain in popularity. Caches deployed in strategic places in the network supported by good caching algorithms can alleviate this large increase in traffic volume for CUTV services where content is still of interest to many viewers, but which viewers do no longer watch it simultaneously. This research was prepared by Alcatel-Lucent Bell Lab and Ghent University, [2].

Performance bounds for peer-assisted live streaming for IPTV network are analyzed with distributed multi tree configuration with bottleneck removal algorithm by Princeton University. In this research, were studied three performance metrics: minimum server load, maximum supported rate, and minimum tree depth, under three cases: unconstrained peer selection, single peer selection, and constrained peer selection. The analysis on the performance bounds also suggest the tradeoffs between tree depth, server load, and degree bound, [5].

Real time monitoring of video quality (model for mapping of packet losses from which can be analyzed loss distortions) in IP networks is analyzed with GE model. The goal of the research was to devise a lightweight solution that would allow real-time, large-scale monitoring of video quality. First contribution was the development of a loss-distortion model that accounts for the impact of various network-dependent and application-specific factors on the quality of the decoded video. Second contribution was in using this model to define a relative video quality metric, rPSNR, that can be evaluated without parsing or decoding the transmitted video bit streams, and without knowledge of video characteristics - thereby significantly reducing complexity while still providing reasonably accurate video quality estimates. The robustness and accuracy of the rPSNR-based method were demonstrated through a broad range of simulations and experiments, [6].

At the end, the model for VBR traffic with Gamma distribution can be found into the research prepared by Moscow State technical University of Service and Moscow Power Engineering Institute. This model reflects two main statistical characteristics of the real video sequence: the quotient distribution with the heavy tail and the long-range part of the autocorrelation function.

Since this model does not approximate the SRD accurately enough, it is suitable for modeling the set of a large number of video sources. In this case the quotient distributions are close to Gaussian and the singular correlation effects are random over the short time spaces, [1].
Implementation of social features over regular STB is described in the material prepared in University of Aveiro. Developing a Social TV application for a commercial IPTV platform regarding the STB hardware and software limitations compelled the team for alternative technical implementations and mishaps were topic on this research. These limitations along with the usual constraints of designing for television also led to different experiments concerning the information layout and interaction patterns, resulting in some directions to other iTV developments over regular browser based set-top boxes, [1].

The self similarity nature of WWW traffic was described into the paper prepared by Boston University. The self similar processes are implemented over the WWW traffic via transferred files and over silent time periods, [8].

5. SELF-SIMILARITY WITH GAMMA DISTRIBUTION FOR REAL TIME TRAFFIC INTO KEY POINTS OF IPTV NETWORK EXPERIENCES

5.1 Measurement Configuration

This research analyses stream outage distribution, delay factor and stream loss at the key points of overall IPTV Network into the period March 2008 - June 2010. The real measurements database includes around 2010 values about the stream outage distribution, delay factor and stream loss at the key points of overall IPTV Network defined in hours and minutes per each day in a year for defined research time period interval of almost two years.
The range value per each figure parameter comes from about 260 up to 4.250 values metric for stream outage distribution parameter, from about 1.080 up to 1.800.000 value metric for stream loss and from about 190 up to 520 values metric for delay factor.
The simplified subject configuration of IPTV network is shown in Figure 3.
The telecommunication network (TN) covers a considerable terrain and combines a large number of various protocols of the link layer. The figure shows only the generalized network structure. There are terrestrial and satellite communication channels inside the network.
In the network configuration shown in Figure 3 the device (example: Cisco Catalyst 3750) is the network 'kernel'.

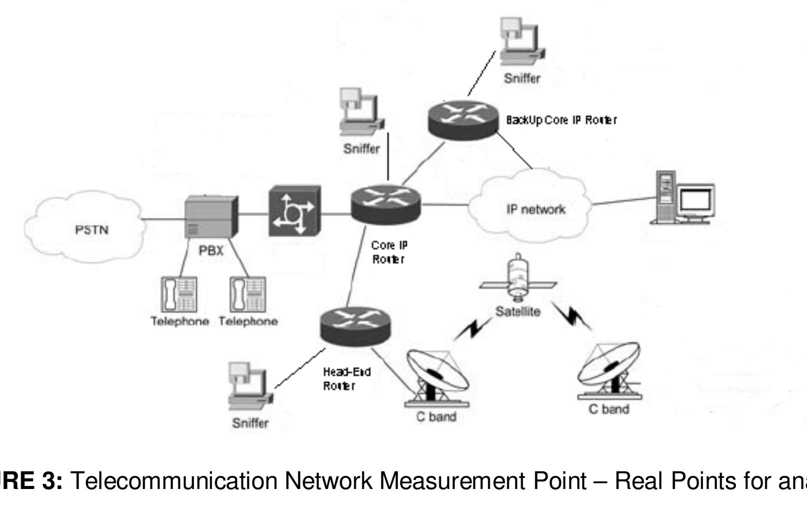


FIGURE 3: Telecommunication Network Measurement Point – Real Points for analyses.

The key network analyses performed during this study are performed on this key point:

- Measurements on key points in IPTV Network. For realization of real time measurements, the free port of network "kernel' equipment, Head-End equipment and Backup of network 'kernel' equipment are connected through the fiber optic interface to the personal computer (PC). The PC has the measurement software and the scheduler, which initiated

the chosen traffic measurement program every 15 minutes. At the link layer the connection was realized over the fast Ethernet protocol. The chosen device port was configured as the SPAN port. As measurement software, the sniffer software catches all packets falling into the interface. As a result about 90 files were recorded.
Since the sniffer software caught total traffic passing through the Head-End Equipment, Core Router and the Backup Core Router Equipment the recorded (by the IP addresses known in advance) log-files were subjected to filtering in order to allocate the measurements for IPTV packets.

- Key Measurement Parameters. Stream Outage distribution as number in packets is the total number of packets that are out during the15-minute time period. Delay Factor is defined as a number in milliseconds that is the maximum value of delay factor before setting an alarm condition. Stream loss as a number in packets is the total number of media packets lost for 15-minute inspection period.

These measurement parameters are recorded into about 90 files by measurement software and systemized, analyzed and compared with the results from the mathematical model.
During the research the frames for connection which made the TCP operate at the transport layer were not taken into consideration.

5.2 Measurement Practice

Into the real time measurements the following traffic characteristics were registered:

1. Total traffic. This traffic, measured in bytes and packets, demonstrates the packet number (or the total packet size in bytes) for which the MAC address of the appropriate device or the MAC address of the destination was present in the field of the source's MAC address.
2. Input traffic. This traffic, classified during the analysis as incoming into the router and used for the measurement, was fixed in the packet form as well as in the total volume in bytes or packets.
3. Output traffic. This is the traffic that is classified at the analysis as outgoing from the router and used for measurement. This parameter is also fixed in packets and in the total volume in bytes or packets.

Shape parameter	Log-log correlation	Variance Time	R/S Statistics
$\alpha = 2.0$	H = 0.592	H = 0.544	H = 0.672
$\alpha = 1.7$	H = 0.758	H = 0.794	H = 0.684
$\alpha = 1.4$	H = 0.7996	H = 0.8198	H = 0.811
$\alpha = 1.1$	H = 0.828	H = 0.745	H = 0.884

TABLE 1: Hurst exponent estimate (example).

The SPAN session is opened on Ethernet port of the routers. The terminal with the loaded sniffer software is connected to this port. Therefore, during the measurement only the traffic of the ports involved in the SPAN session will be analyzed.
The following scenario was configured:

- Step1. All the traffic considered for IPTV Stream into the Head-End equipment, Core Router and Backup Core Router was analyzed.
- Step 2. The considered traffic is measured via SPAN session.
- Step 3. The parameters: stream outage, delay factor and stream loss are measured.
- Step 4. The measurement results were defined into Documents.
- Step 5. The enclosed curve from the real measurement was analyzed for catching the best effort mathematical model.
- Step 6. Enclosed mathematical model with Gamma distribution is defined with specific parameter values related with real measurements.
- Step 7. The values of the real measurements are related with results from the mathematical model.

- Step 8. Performing the comparison analyzes calculation of the discrepancies between the results from the analytical model and results from the real measurements.

5.3 Assumed Self Similar Model

It is assumed that the self-similar model as mathematical model can explain the behavior of the real IPTV traffic into the key points of the IPTV network.
With analyzes of different types of mathematical functions that can be found into the theory of self -similar processes is assumed that the Gamma distribution with intergroup characteristic can be used as a mathematical model for performance analyzes of key points of the IPTV network.
The Intergroup characteristic is described with first order and second order statistical characteristic of I frame processes.

$$F_{X_I}(r) = \frac{r^{m_I - 1}}{\Gamma(m_I) l_I^{m_I}} e^{-r/l_I}, \forall r > 0$$

For the assumed model in this research we define the parameter values as follows: H = 0.873, λ=0.891, L = 100 and k = 30.
The parameter β is calculated with formula β = 2-2*H, and receive value β = 0.25.
The shape parameter and the scaling coefficient are related by the μ value and σ^2 of the I frame process by equations:

$$m_I = \frac{\sigma_I^2}{\mu_I^2} \quad \text{and} \quad l_I = \frac{\sigma_I^2}{\mu_I}$$

So, they receive values in range from 0.1 to 2.5.

5.4 Comparison Analyzes

Next, we show through comparative evaluation that considered self-similar process as mathematical model defined with Gamma Distribution with Intergroup characteristic with first order and second order statistical characteristic of I frame processes is important for discovering of the stream outage distribution, delay factor and stream loss parameters in key points of overall IPTV Network. A qualitative comparison of all this technique is provided in this Section.
To compare Analytical model with Gamma Distribution and RMSOD (Real Measurements for Stream Outage Distribution), RMDF (Real Measurements for Delay Factor) and RMSL (Real Measurements for Stream Loss), we use two years database aggregated at RMSOD, RMDF and RMSL with resolution measurements per hour for each day in a year for sample time period.

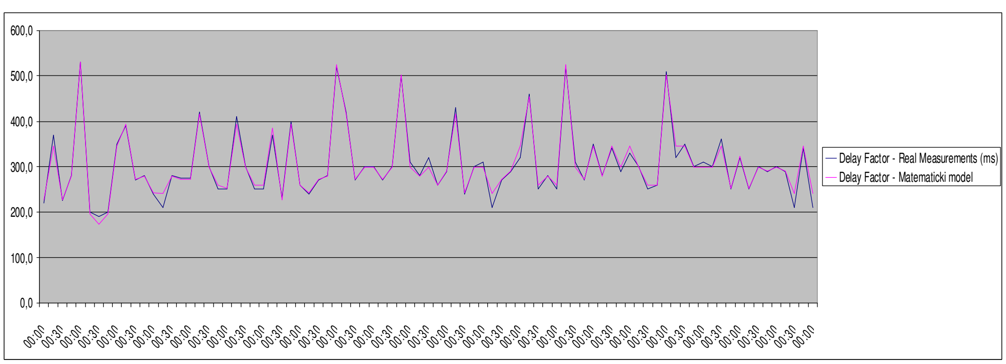


FIGURE 4: Comparison Analyzes for delay factor for Back Up Core Router into IPTV Network.

We describe our experiences in applying Analytical model for Gamma Distribution on the RMSOD, RMDF and RMSL database and we demonstrate how we can apply the suite of techniques in Gamma Distribution to receive the value for stream outage distribution, delay factor and stream loss parameters into the overall IPTV network. In this research, we consider the values from the RMSOD, RMDF and RMSL database for the related sample period as our input series. These are the direct measures reflecting the stream outage distribution, delay factor and stream loss parameters of overall IPTV network performance impairments. We apply Analytical model with Gamma Distribution and founding the differences. We find that Analytical Model with Gamma Distribution particularly useful in defining the value of stream outage distribution, delay factor and stream loss parameters of IPTV network for every time period. This is an important conclusion for IPTV Network detailed diagnosis that would exhaust cost effective operation into IPTV networks.

- 1st class of measurements. We perform the correlation analysis for delay factor data collected over time period of 21 hours with measurement intervals of 15 minutes. The correlation time window is set on 15 minutes. We observe strong correlations between values of analytical model and values for delay factor in the Backup Pop up router into the IPTV network in defined time period. We have also validated the discovered dependencies of delay factor into the Backup Pop up router of the IPTV network with analyzed model. We found discrepancy level for this time period of about 2.32% (Figure 4).
- 2nd class of measurements. We perform the correlation analysis for delay factor data collected over time period of 7 days with measurement intervals of 24 hours. The correlation time window is set on 24 hours. We observe strong correlations between values of analytical model and values for stream loss parameter from Head-End equipment into the IPTV network in defined time period. We have also validated the discovered dependencies of stream loss parameter from Head-End equipment into the IPTV network with analyzed model. We found discrepancy level for this time period of about 5.83% (Figure 5).

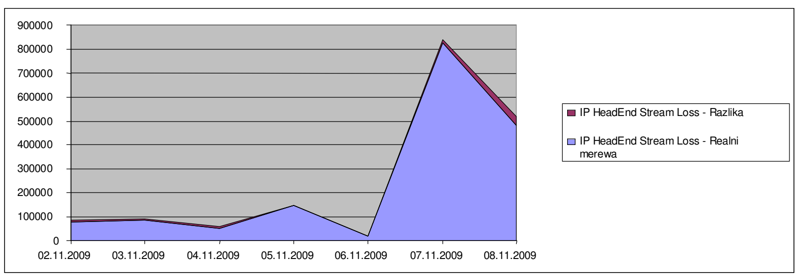


FIGURE 5: Comparison Analyzes for stream loss for Head End Equipment into IPTV Network.

- 3rd class of measurements. We perform the correlation analysis for stream outage distribution data collected over time period of 7 days with measurement intervals of 24 hours. The correlation time window is set on 24 hours. We observe strong correlations between values of analytical model and values for stream outage distribution parameter from Head-End equipment into the IPTV network in defined time period. We have also validated the discovered dependencies of stream outage distribution parameter from Head-End equipment into the IPTV network with analyzed model. We found discrepancy level for this time period of about 3.56% (Figure 6).

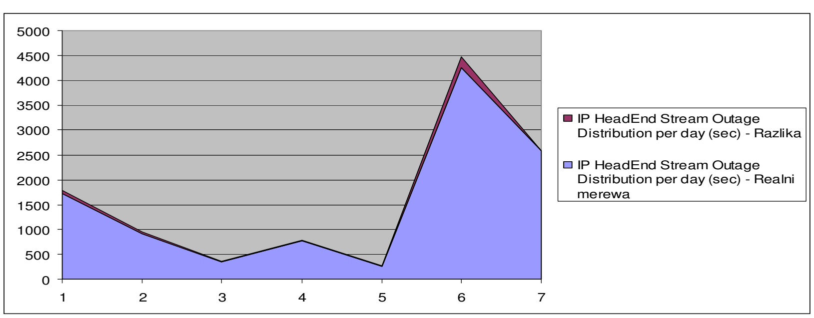


FIGURE 6: Comparison Analyzes for stream outage distribution for Head End Equipment into IPTV Network.

- 4th class of measurements. We perform the correlation analysis for stream loss data collected over first time period of 11 days with measurement intervals of 24 hours. The correlation time window is set on 24 hours. We observe strong correlations between values of analytical model and values for stream loss parameter from Core Router equipment into the IPTV network in defined time period. We have also validated the discovered dependencies of stream loss parameter from Core Router equipment into the IPTV network with analyzed model. We found discrepancy level for this time period of about 5.02% (Figure 7).
- 5th class of measurements. We perform the correlation analysis for stream loss data collected over second time period of 11 days with measurement intervals of 24 hours. The correlation time window is set on 24 hours. We observe strong correlations between values of analytical model and values for stream loss parameter from Core Router equipment into the IPTV network in defined time period. We have also validated the discovered dependencies of stream loss parameter from Core Router equipment into the IPTV network with analyzed model. We found discrepancy level for this time period of about 3.67% (Figure 8).

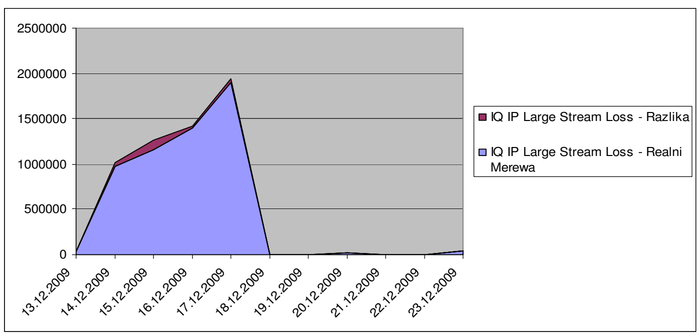


FIGURE 7: Comparison Analyzes for stream loss for Core Router into IPTV Network.

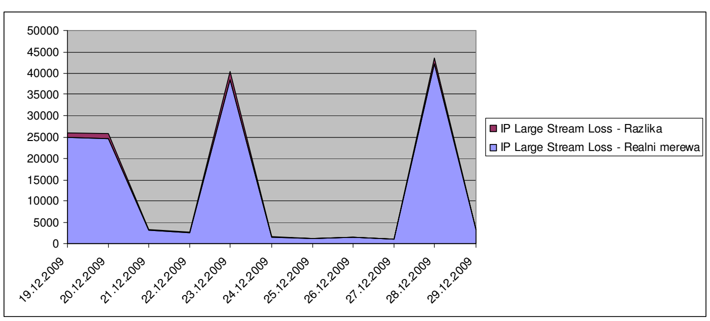


FIGURE 8: Comparison Analyzes for stream loss for Core Router into IPTV Network.

6. CONCLUSION

The first characterized research for outage distribution parameter, stream loss parameters and delay factor parameter at the key points of IPTV network is finished and described in this paper.
Research consist strict specification of covered model with details for defined parameters and used functions based on fractal shapes represented by self similar processes and description of used network IPTV topology.

An achieved result confirms that covered research model can be used for performing real time analyses at the key points of the IPTV networks.
The simplicity of the covered model and the accuracy of the enclosed results confirms that the described research have extra high level of quality and can be used as the base in further analyses, studies and researches in the field.

The covered model can be implemented for performing analyzes of the stream at the CPE devices into the IPTV network, analyzes related with handling of the HD stream into the IPTV network, analyzes of IPTV stream into the 4G networks and other.

The research has impact on better automation of IPTV operations in order to achieve better detection and troubleshooting performance.

7. REFERENCES

Books

[1] O. I. Sheluhin, S. M. Smolskiy, A. V. Osin. (2007). Self-Similar Processes in Telecommunications. (1st edition). [On-line]. Available: www.wiley.com [Nov 07, 2010].

Journal

[2] D. D. Vleeschauwer, Z. Avramova, S. Wittevrongel, H. Brunneel. (2009. Jun.). “Transport Capacity for a Catch-up Television Service.” Euro/TV’09. Available: http://portal.acm.org [Oct. 14, 2009].

[3] A. Mahimkar, Z. Ge, A. Shaikh, J. Wang, J. Yates, Y. Zhang, Q. Zhao. (2009. Aug.). “Towards Automated Performance Diagnosis in a Large IPTV Network.” Sigcomm’09. Available: http://citeseerx.ist.psu.edu/viewdoc [Oct. 30, 2009].

[4] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, J. Xu. (2009. Jun.). “Modeling Channel Popularity Dynamics in a Large IPTV.” Sigmetrics. Available: http://portal.acm.org [Nov. 02, 2009].

[5] S. Liu, R. Zhang-Shen, W. Jiang, J. Rexford, M. Chiang. (2008. Jun.). “Performance Bounds for Peer-Assisted Live Streaming.” Sigmetrics. Available: http://www.princeton.edu [Nov. 10, 2009].

[6] S. Tao. (2008. Oct.). “Real-Time Monitoring of Video Quality in IP Networks.” IEEE/ACM Transactions on networking, vol. 16, no.5. Available: http://portal.acm.org org [Nov. 17, 2009].

[7] J. F. Abreu, P. Almeida, R. Pinto, V. Nobre. (2009. Jun.). “Implementation of Social Features Over Regular IPTV STB.” Euro/TV’09. Available: http://portal.acm.org [Oct. 03, 2009].

[8] M. E. Crovella, A. Bestavros. (1997. Dec.). “Self Similirity in World Wide Web Traffic: Evidence and Possible Causes.” IEEE/ACM Transactions on networking, vol. 5, no.6. Available: http://portal.acm.org [Nov. 23, 2009].

[9] B. Erman, E. P. Matthews. (2008), “Analyzes and realization of IPTV Service Quality.” Bell Labs Technical Journal 12(4), 195–212 (2008). Available: http://www.interscience.wiley.com [Apr. 01, 2010].

[10] N. Degrande, D. D. Vleeschauwer, K. Laevens. (2008). “Protecting IPTV against Packet Loss: Techniques and Trade-Offs.” Bell Labs Technical Journal 13(1), 35–52 (2008). Available: http://www.interscience.wiley.com [Apr. 01, 2010].

[11] K. Sridhar, G. Damm, H. C. Cankaya. (2008). “End-to-End Diagnostics in IPTV Architectures.” Bell Labs Technical Journal 13(1), 29–34 (2008). Available: http://www.interscience.wiley.com [Apr. 01, 2010].

[12] M. Verhoeyen, D. D. Vleeschauwer, D. Robinson. (2008). "Content Storage Architectures for Boosted IPTV Service." Bell Labs Technical Journal 13(3), 29–44 (2008). Available: http://www.interscience.wiley.com [Apr. 01, 2010].

[13] B. Krogfoss, L. Sofman, A. Agrawal. (2008). "Caching Architectures and Optimization Strategies for IPTV Networks." Bell Labs Technical Journal 13(3), 13–28 (2008). Available: http://www.interscience.wiley.com [Apr. 01, 2010].

World Wide Web

[14] Multichannel. "Cable's IPTV Future." Internet: www.multichannel.com, Sep. 2009 [Mar. 31, 2010].

[15] Wikipedia. "Self Similar Process." Internet: www.en.wikipedia.org/wiki/Self-similar_process, Jan. 2010 [Feb. 23, 2010].

[16] Wikipedia. "IPTV." Internet: www.en.wikipedia.org/wiki/IPTV, Oct. 2008 [Apr. 04, 2010].

[17] Exfo. "IPTV-Technology Overview." Internet: www.images.exfo.com/www/Products/FTB-400/Index.swf, Jan. 2006 [Feb. 11, 2010].

[18] Alcatel-Lucent. "IPTV test and measurement best practices." Internet: www.bitpipe.com, May. 2007 [Feb. 19, 2010].

[19] J. Graves. "Innovation: The Key to Long-Term IPTV Success." Internet: www.findarticles.com, Oct. 2009 [Mar. 15, 2010].

[20] Ericsson. "Ericsson IPTV Solutions – enabling a new kind of IPTV." Internet: www.ericsson.com, Jul. 2009 [Mar. 28, 2010].

[21] NetUp. "NetUP's IPTV Complex." Internet: http://www.netup.tv, 2009 [Apr. 01, 2010].

Energy Consumption in Ad Hoc Network With Agents Minimizing the Number of Hops and Maintaining Connectivity of Mobile Terminals Which Move from One to the Others

Kohei Arai *arai@is.saga-u.ac.jp*
Department of Information Science, Faculty of Science and Engineering
Saga University Saga,
8400021, Japan

Lipur Sugiyanta *lipurs@gmail.com*
Department of Electrical Engineering,
Faculty of Engineering
State University of Jakarta
Jakarta, 13220, Indonesia

Abstract

Wireless mobile ad-hoc network (MANET) is a special kind of network, where all of the nodes move in time. Node is intended to help relaying packets of neighboring nodes using multi-hop routing mechanism in order to solve problem of dead communication. MANET which engages broadcasting and contains multiple hops becomes increasingly vulnerable to problems such as mobile node's energy degradation, routing problem and rapid increasing of overhead packets. This paper provides an extensive study of energy consumption in the MANET that consists of two network areas with the presence agents. Agents will minimize number of hops and its affect in linearity with the delay. As nodes grow, either in data transmission services or coverage of node's communication or more agents stand in overlapped locations, the intensive data exchange and topology construction to adapt the network are becoming an important issue. As a result, agents are needed to support these process automation, high-level connectivity, and intelligent service. We evaluate the agents' performance and network energy consumption for supporting MANET that divided into two domain/network areas. The proposed agents provides service for packets transmission between networks; e.g. determine appropriate relay nodes dynamically, maintain the transmission between networks through another nodes, share the topology knowledge among agents, and route packets between source and final destination that are unable to communicate directly. The achievement on research with this approach is conducted via simulation study. A similar network without agents is presented to derive such referential bounds by using appropriate functions of network agents. The proposed algorithm is confirmed with composite simulation results.

Keywords: Energy, Agents, Multi-hops, Connectivity

1. INTRODUCTION

MANETs is a multi-hop wireless network in which nodes can communicate with each other without support of any existing infrastructure. This network is fully autonomous and free to move anywhere (in the area or across areas) any time. Node is referred to a mobile device which equipped with built-in wireless communications devices attached and has capability similar to autonomous router. The nodes can be located in or on airplanes, ships, cars, or on people as part of personal handheld devices, and there may be multiple hosts among them. Each node is autonomous. The system may operate in isolation, or have gateways to a fixed network. In the future operational mode, multiple coverage of the network is expected to operate as global "mobile network" connecting to legacy "fixed network".

At each time and every node's positions, a wireless connectivity in the form of a random, single-hop, multi-hop path may exist between nodes and areas. This topology may change as the nodes move or adjust their parameters. Among networks, Wireless (Ad-Hoc) Mesh Network has several characteristics; i.e. dynamic topologies, bandwidth-constrained, energy-constrained operation, and limited physical security. These characteristics create a set of underlying assumptions and performance considerations for protocol design which extend beyond static topology of the fixed network. The design should reacts efficiently to topological changes and traffic demands while maintaining effective routing in a mobile networking context.

In this paper we propose a mobile agent - based routing protocol for delivering packets in clustered networks whose performance increases with the increasing traffic in the network due to high degree of cooperation among both the nodes and agents. The agents were used to deliver packets where they acted as a messenger that will migrate packets from a source in certain clustered network area to a final destination which is located in different network area. Thus when there are a number of sources to send messages to a final destination simultaneously; a group of parallel redundant traffic with similar responsibility will be generated. This traffic will eventually consume the energy, the bandwidth, and other crucial resources of the ad hoc wireless network. The novelty of this work is to analyze the effect of cooperating agents at the border of clustered networks to the overall network energy consumption and compare it with the similar network which has all the pairs of source-final destination carrying packets for the transmission travel across different clustered networks. Delay performance of those networks, however, will be key issues. This problem is caused by complex network interactions that complicate the multi-hop routing mechanisms. The agents, which are relieved of their responsibilities in the process, will be slept if possible and thus reducing the traffic load heavily [7].

In this paper, we analyze various formation options of agents and nodes and their potential overheads and impacts on efficiency are evaluated via simulation study. The remainder of this paper is organized as follows: Section 2 gives the previous research related with our model. Section 3 discusses the detail design of the simulation model, its notations, and assumptions. Simulation algorithm that suits mobile environment is presented in Section 4. A performance evaluation of networks and its comparison to network without implementing agents at the similar network topology are presented in Section 5 and Section 6, respectively. Section 7 concludes the paper.

2. PREVIOUS WORKS

All nodes in MANET rely on batteries or other exhaustible energy modules for their energy. For this network, the most important system design criterion for optimization is energy conservation. Thus one critical design issue for future Wireless Ad-Hoc Network is the development of efficiently energy consumption that suits communication architectures, protocols and services of network enabled wireless devices. Energy conservation means to maximize the operational lifetime of a node, thus, enhancing the overall user experience [1][18].

Previous researches for energy conservation of MANET are focus on transmission energy control and dynamic turning off active nodes in network. Adjusting the energy for transmission reduce the energy consumption significantly and increase lifetime of the network. Our previous propose framework concentrated on energy aware broadcasting technique for wireless ad hoc networks. The framework uses a packet forwarding technique where neighbor nodes can be elected to be relay on behalf of source-destination path with the goal of optimized the overall energy consumption to deliver packets in the network, while maintaining the connectivity among nodes. Transmission to a distant node may consume a higher amount of energy in comparison to transmission to a node in closer range and more energy will be used for sending packets than receiving or processing packets. In addition, transmission of larger packets may consume a higher amount of energy in comparison to transmission of smaller packets. The framework is based on the principle that adding additional relay nodes with appropriate energy and routing metric between source and final destination nodes significantly reduces the energy consumption necessary to deliver packets in Wireless Ad-Hoc Network while maintaining connectivity among nodes [1]. Yu Wang introduced the method to control energy over network layers in MANET [2]. Energy management based on cross-layer design make network more robust and adaptive. At physical and data link layers, the method triggers nodes to transfer between energy-save mode and active mode, while at network layer, a routing protocol is equipped with new defined joint function, which can realize hop-by-hop and end-to-end energy control. Other researcher, Cartigny proposed an algorithm that requires local information, i.e. full knowledge of network its distance to all neighboring nodes and distances between its neighboring nodes. Distances can be measured by using signal strength, time delay or more sophisticated techniques like microwave distance [3]. Nodes adjust their transmission energy so as to achieve the minimum energy consumption according the local information. Ramanathan and ElBatt remarked method of adjusting energy for delivering packets implement with considering levels to achieve a desired degree of connectivity in the network [4][5]. Bergamo etc. submitted a routing algorithm based on distributed energy control, which provide optimum transmit energy while maintaining limited degradation in throughput and delay [6]. Each node in MANET estimates the energy necessary to reach its neighbors, and this estimation is used both for tuning the transmit energy and as the link cost for minimum energy routing. Simulations results confirm that this method can save the network's energy.

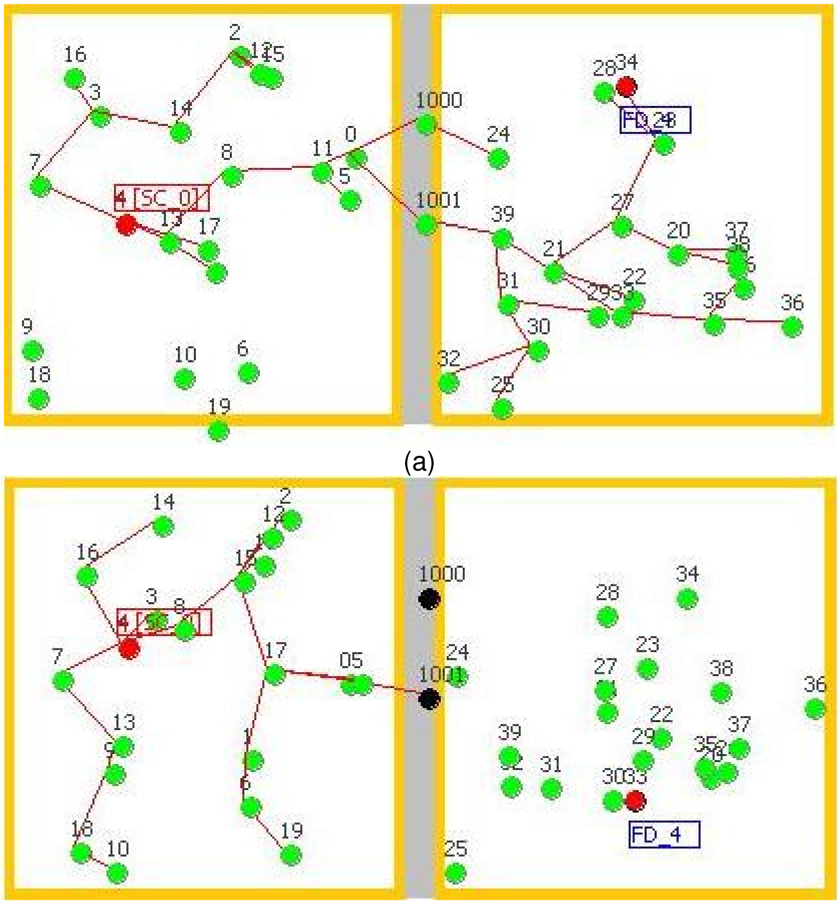
Most of those researches did well in one network area. Taking into consideration all the challenges mentioned above, we deal with energy control over nodes that are distributed into several network areas. The main contribution of this paper is that we propose agents that collect local network information. If agents are positioned in the overlapped area between networks, each source node in certain network area only needs to get its agents' information whenever it require to contact other node located in other network area from that is where full knowledge of network is required to make decision.

3. SIMULATION MODEL, NOTATIONS, AND ASSUMPTION

As foundation for this mobile environment, the core algorithm is developed from static mode (e.g., sensor networks). The enhancement algorithm for serving mobility then detailed in support of topology development, topology maintenance, and routing maintenance.

The model is initiated from broadcast mechanism and propagated through node-to-node based routing metrics approach. Each source injects single big packet which fragmented into multiple packets in the network, which traverse through the network until those reach the final destination. Packets are queued at each node in its path where it waits for an opportunity to be transmitted. This model is not only applicable in direct communication (one hop transmission) but it can also work in multi-hop transmission. In this situation, when the source and final destination nodes are located outside its clustered network area, source node is capable to discover multiple hop route lead to agents thus maintaining the connectivity required in comparison to standard flooding based ad hoc routing designs.

It is square of Cartesian model area with 200x200 areas and one overlapped section. We consider the case where all nodes in the network are similar, i.e., assuming a homogeneous infrastructure. Inside areas, nodes are deployed uniformly, distributed at random position in the both areas. This deployment produces a connected topology under some assumptions; sometimes a completely connected topology is built and sometimes topology is not fully connected. Simulation build a large connected component quickly using a communications radius considerably smaller than the radius needed to have the entire network connected. Agents are located in such place to facilitate communication between wireless areas and minimize the number of hops to achieve optimum throughput of mesh clients which communicate each other. Agents have static position located in the overlapped between two network areas.



(a)

(b)

FIGURE 1: Development of network topology within two areas.
(a) With Agents involved in delivery of packets.
(b) Without Agents involved in delivery of packets.

This mode of messenger can be made clear from the Figure 1. At Figure 1(a), if the source (node 4 mark with red color) coming from certain clustered network area and having destinations (e.g. node 33) located at different (adjacent) clustered network area, the paths are able to meet at a common Agents (node 1000 or node 1001) simultaneously then only any one of the agents will be sufficient to carry all the messages to the proper destination. Thus our routing scheme utilizes the agents capability where a collection of independent request transmission come together for the purpose of cooperative task behaviors and maintaining these connectivity among pairs of sources - destinations. All these tasks essentially work directly through node - agent communication. The entire algorithm works on the fact that agents need to know the existence of each nodes at each clustered areas. On the other side, at Figure 1(b), predefined source (node 4) start the topology development protocol by sending (broadcast) an initial Hello Message. With receive-transmit subsequent routine, the process continues to all reachable nodes. Not every node will be selected to be part of the tree, and those which were not selected will keep silent (in the propagation of packet). Without any knowledge of destination topology (located in other network area), then the packets must travel to each vertex to reach the destination node. In both pictures, if there are more than one source nodes starts to transmit packets simultaneously, then several trees may be built in parallel.

The Model

Simulation consists of multi clustered network environment of homogeneous nodes that communicate with each other using the broadcast services of IEEE 802.11. There are nodes with different roles simulated in this simulation, namely initiator node/source node, receiver node, sender node, destination node, and final destination node. Initiator node/source node is node that initiates transmission of packet. Packet can be either route discovery or data transmission. Like other nodes, initiator is always moving with random direction, speed, and distance. At the time it is moving, initiator node is always sensing its neighbor to maintain connectivity. Receiver node is node that can be reached by source/sender node. Nodes are defined as neighbors if it located within its distance radius range. At initial time, node senses its neighbors before packet data is required to be transmitted. Coverage neighbor nodes always receive packets that are broadcasted from sender. Destination node is selected receiver node in multi hop transmission that should relay packets to the next receiver node. Final destination node is node that became the finish destination of packets.

The layered concept of networking was developed to accommodate changes in local layer protocol mechanism. Each layer is responsible for a different function of the network. It will pass information up and down to the next subsequent layer as data is processed. Among the seven layers in the OSI reference model, the link layer, network layer, and transport layer are 3 main layers of network. The framework is configured in those layers. Genuine packets are initiated at Protocol layer, and then delivered sequentially to next layer as assumed that fragmented packets to be randomly distributed. Simulation models each layer owned with finite buffers. Limited buffer makes packets are queued up according to the drop tail queuing principle. When a node has packets to transmit, they are queued up provide the queue contains less than K elements ($K \geq 1$). To increase the randomization of the simulation process, simulation introduces some delay on some common processes in the network, like message transmission delay, processing delay, time out, etc. This behavior will result that at each instance of a simulation would produce different results. The packets exchanged between sender and receiver is of a fixed rate transmission λ based on a Poisson distribution. Nodes that have packet queued are able to transmit it out using in each available bi-directional link channel.

Our work extends the chance of contacting of the agents that arranged at the fixed cluster from nodes which are distributed randomly within the network. Agents navigating through the network for delivering messages must understand these clustered nodes whenever they are initiating a new route request and thereby increasing the degree of spatial coordination (agents must be on the same place). The temporal coordination has been enhanced with the introduction of a short waiting delay offered to each nodes/agents by the clustered overheads packets. This node – agent coordination will reduce the number of hops and waiting time in spite of further increase the overheads packets hang around with agents and head to highly the agent-chasing problem. The place hosted for the clustered agents can be called as the overlapped area within the network and the detainment period by the agents can be called agent periods.

This overlapped area actually offers a temporary space to be used by all agents for sharing network knowledge, and exchanging messages. Thus when an agent comes a fresh it can exploit all other agents who are currently experiencing their agent periods. Here the traveling agents are allowed to carry the information of already visited clusters along with them. The idea behind this is to capture and share the partial network information present with roaming agents. The integration of all such partial information at a common overlapped area helps cooperative tasks like taking the decision for next destination, suitable exchange of messages between agents, getting up-to-date knowledge of the network and reducing unnecessary redundant overhead packets used to visit nodes.

Thus the model of coordination clustered network area where the autonomous agents will be able to deliver messages within a large network with the cooperative communication between them at suitable overlapped area is necessary. There is need of knowing the routes proactively or reactively where part of the network capacity is used for exchanging chunks of routing table data.

We built network simulator to evaluate this proposed algorithm. The simulator supports physical, link and routing layers for single/multi hop ad-hoc networks. We assume that IEEE 802.11 Distributed Coordination Function (DCF) or MAC protocol which uses Channel Sense Multiple

Access with Collision Avoidance (CSMA/CA) already deployed. Successfully received packet by receiver's interface is packet whose SNR is above a certain minimum value otherwise the packets cannot be distinguished from background noise/interference. Packets are transmitting through physical layer in accordance with Poisson distribution. Communication between two nodes in IEEE 802.11 uses TCP signaling before the actual data transmission takes place. Simulation simulates this with random hearing to link's condition. If link allow packets to be sent, then sender executes some packets already queued. To execute preferential event in sequentially distributed events, we used a simple approach that consists of applying a different time-event execution by means of the triggering event sequences action. The lower and upper bound of the queuing interval are set such that they do not interfere with predefined timers used by the other events for layers and modification events.

Nodes

Nodes are equipped with antenna module installed as capable of dynamically adjusting the transmission energy used to communicate with other nodes. Industrial standard of antenna module supports a management for controlling this energy consumption. The energy consumption required to transmit packet between nodes A and B is similar to that energy required between nodes B and A if and only if the distance and the size of packet are same. The coverage distance range of the nodes is a perfect symmetric unit disk (omni-directional). If $d_{x,y} \leq r_x \rightarrow x$ and y can see each other. This assumption may be acceptable in the condition that interference in both directions is similar in space and time; which is not always the case. Usually interference-free Media Access Control (MAC) protocol such as Channel Sense Multiple Access (CSMA) may exist. In addition, wireless link channel is assumed to have no physical noise; i.e., the errors in packet reception due to fading and other external interferences are not considered as a serious problem. Packets from sender to receiver will be transmitted as long as the bandwidth capacity is sufficient and the received signal to noise ratio (SNR) is above a certain minimum value. Thus every packet successfully received is acknowledged at the link layer and de-encapsulate at the higher layer. Each node is capable of measuring the received SNR by analyzing overheard packet. A constant bit error rate (BER) is defined for the whole network. Whenever a packet is going to be sent, a random number is generated and compared to the packet's CRC. If the random number is greater, the message is received, otherwise it is lost. The default value for the BER is 0, which means there is no packet loss due to physical link error.

Energy is power kept in each node. [13] was assumed that the radio dissipates E_{elec} = 50 nJ/bit to run the transmitter or receiver circuitry and ε_{amp} = 100 pJ/bit/m^2 for the transmit amplifier. The radio model is shown in the Figure 2 below.

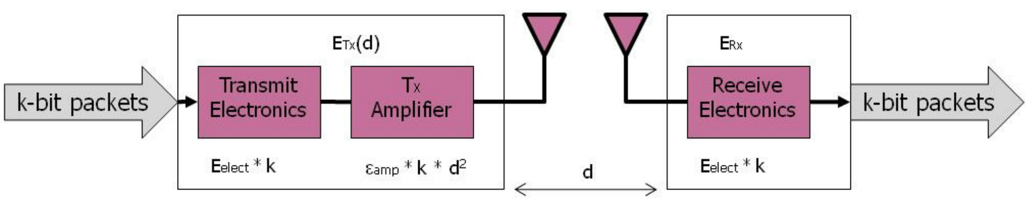


FIGURE 2: The radio model [13].

Thus, to transmit a k-bit message a distance d using this radio model, the radio expends:

$$E_{Tx}(k,d) = E_{Tx-elect}(k) + E_{Tx-amp}(k,d) \quad (1)$$

$$E_{Tx}(k,d) = E_{elect} * k + \epsilon_{amp} * k * d^2$$

and to receive this message, the radio expends:

$$E_{Rx}(k) = E_{Rx-elect}(k) \quad (2)$$

$$E_{Rx}(k) = E_{elect} * k$$

Let Emin_i is the minimum energy ratio of node I at which a node can still receive, process, and transmit packets. Node j finds out the energy level of neighbor node I through analyzing of received reply packet from node I as it responded the previous transmitted Hello packets. The computation of Emin_i is done through two-step propagations. The use of two-steps propagation model is to simulate interactive propagation in the operation of the protocol in dynamic environment. As a future research, the appropriate propagation model that best matches to this environment should replace the simple two-steps model presented here [10][15][16]. The two-steps propagation model is appropriate for outdoor environments where a line of sight communication existed between the transmitter and receiver nodes and when the antennas are omni-directional. The two-steps propagation model assumes there are two main signal components. The first component is the signal traveling on the line of sight to reached neighbors along with its reply from neighbors and the second component is a confirmation packet transmitted to selected neighbors.

The energy behaviors of node are defined as follow:

- During the idle time, a node does not spend energy. Even though this assumption has been proven untrue because being idle might be as costly as receiving data, this is still an assumption that can be done in most experiments, since the most important factor is the overhead in terms of message exchange and its associated cost.
- The nodes are assumed to have one radio for general communication. The main radio is used in all operations when the node is in active mode, and to send and receive control packets. When this radio is turned off, then no messages will be received and no energy will be used.
- Energy distribution among nodes can either be constant value, normally distributed, Poissonly distributed, or uniformly distributed.

Agents

In this paper we introduce mobile agents to hop around the networks for delivering packets .The agents are allowed to meet with other agents at fixed overlapped places as has been mentioned. The cooperation among agents will mutually benefit each other by cooperating in delivering packets. In this current flexible and decentralized framework any autonomous node can send message to any other node at any instant within the network by just issuing a mobile agent. The agent then communicates with other agents to determine the proper one to carries the message to the corresponding destination node. The selected agent then becomes responsible for delivering the message to proper destination. Analog to the real life, these agents actually play the role of messengers and the selected agent play the role of post offices in the ad hoc wireless scenario. Such cooperation among agents scheme has been explicitly designed to reduce the agent traffic in the network. The unnecessary redundant node visits made by the agents to reach destination node has been avoided by sharing and merging with other agents. These agents take the responsibility of providing communication services and improvement of overall traffic in the network.

While delivering messages, an agent will maintain the path records of all the visited nodes in both clustered networks and its corresponding topology. Carrying this network information provide coordination and share the updated network knowledge with other agents. This information field carries topology tree along with numeric values of this membership list (of nodes) collected from each clustered network. In few cases the length of the information list carried by an agent gets longer due to the course of journey made from agent to reach the proper destination. If this happened, the list has been restricted using the hop count limit in order to avoid a huge series of data to be carried along by an agent and subsequent nodes. No packet is allowed to travel forward further whenever its hop count has got exhausted and it is compelled to move back to its originating source node/agent. Thus whenever an agent has finished its forward journey it will eventually follow the same path back to the source node.

The objective of the navigation procedure is to minimize the hops between the agent at overlapped location (current location of the node where the agent is residing) and the source and destination node's location. This criterion would enable an agent to select a neighbor of its current location and take out the packets to the destination nodes. If there is no neighbor available at that instant of time satisfying the above-mentioned criterion, the agent waits for a pre specified amount of time (randomly) and tries to communicate with other agents (any agents can be reached) to get its knowledge. Such contacted agents will respond the request. Through intensive communication among agents, the best agent can be selected to take responsibility for delivering received packets.

In the simulation, when a source node within the ad hoc network wants to send some packets, it immediately senses whether an agent is needed. Each such agent attaches with itself a topology bag to accommodate the request (certain) destination node. This bag is of a given capacity, which can be made full or can be made empty. The source node after initiating the agent puts the packets to (appropriate) agent. The agents are able to exchange these packets with other agents on having suitable coordination with them, which will have the best route path with minimum hops to reach destination node. These agents will be inactive automatically when there is no more packets need to be delivered.

Because of high degree of mobility, the topology will change and it is assumed that the agent will eventually succeed to migrate [30][31]. Whenever an agent wants to leave its current location for delivering packets to some unknown networks it will collect the topology list information from the nodes and will try to reach for a boundary node through which it may get an exit point. The node lying at the boundary will have neighbors from two or more different agents and can act as gateway nodes to other clustered network area. Thus if an agent can reach such a cluster area boundary, it can start visiting a fresh (other) agents. As the location of the agents can be made available from any node of that same clustered network area it can easily track the new nodes there, which has been compulsory. Though the order of cluster visits take place in a random manner still the redundancy in the path visit has been avoided by maintain the path visit list (using BFS function). The agents are free to roam among the overlapped area within the network in a random manner. This agents' capability will be presented on the next paper.

4. Performance Evaluation

In this section, an evaluation of the framework is discussed and followed by a number of performance issues associated with network traffic and network energy consumption. The following evaluation and subsequent charts are obtained from simulation with 40 nodes. Nodes are randomly positioned in both network areas. There are two scenarios with simulation in order to show the effectiveness of agents, i.e. simulation with agents and simulation without agents. Same initial topologies are created for both scenarios. The simulation then follows its autonomous movement during simulation. To simulate agents, 4 nodes are positioned in the overlap area. Nodes are moving during 10 steps of simulation with speed of 20 km/h. The source is set to transmit packet with length of 10Kbyte to reach final destination, either located in same network area or different network area. The following charts show the energy consumption, span of network and average transmitted packets as parameters change. The energy of network is defined as the period from the beginning of simulation to the end of simulation when all nodes exhaust their energy.

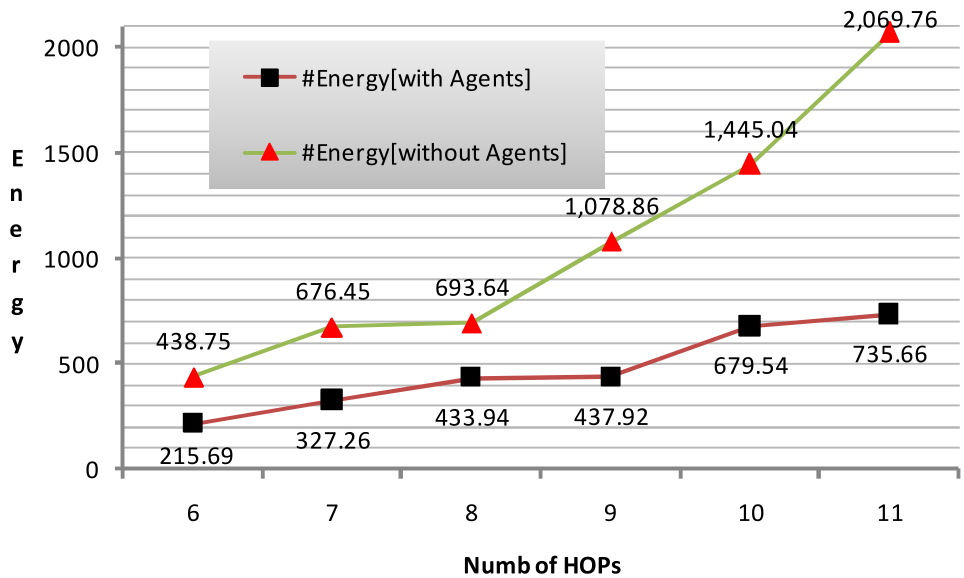


FIGURE 3: The comparison of average energy consumption (in nJ/bit) on nodes in the network using agents and without agents in each successful data transmission with certain number of hops between source and final destination (with pair of source and final destination nodes are located in different network area).

In this set of simulations, we evaluate the performance of schemes with and without agents' mode. CBR is used to send packets and random way-point is adopted as is and the one hop routing discovery is applied. The data packets can be persistently transmitted along the existing route and more energy will be saved. Figure 3 shows the relationships between the network energy consumption and number of hops, for a 10Kbyte file being transferred between source and final destination nodes.

As is shown in the Figure 3, with the performance of both routing scheme stay the same, network achieve low energy consumption as well as longer span-time of network because agents can dynamically limits the propagated packets and consider the residual energy of node in routing setup phase. While the mobility of nodes is high, the advantage of the agents based routing becomes more remarkably. The actual gain in this scheme, expected by the use of agents, ranges around 37% for 64% saved energy. Simulation methods for energy conservation are focus on broadcast energy control and dynamic turning off active nodes in network. Controlling the broadcast energy allows to significantly reduce the energy consumption for data transmission and increase lifetime of the network. Another proposed protocol based on energy management where each node requires the knowledge of its distances to all neighbor nodes, throughput metric, and direction metric is being prepared. Nodes adjust their transmission energy so as to achieve the minimum energy consumption according the neighborhood information.

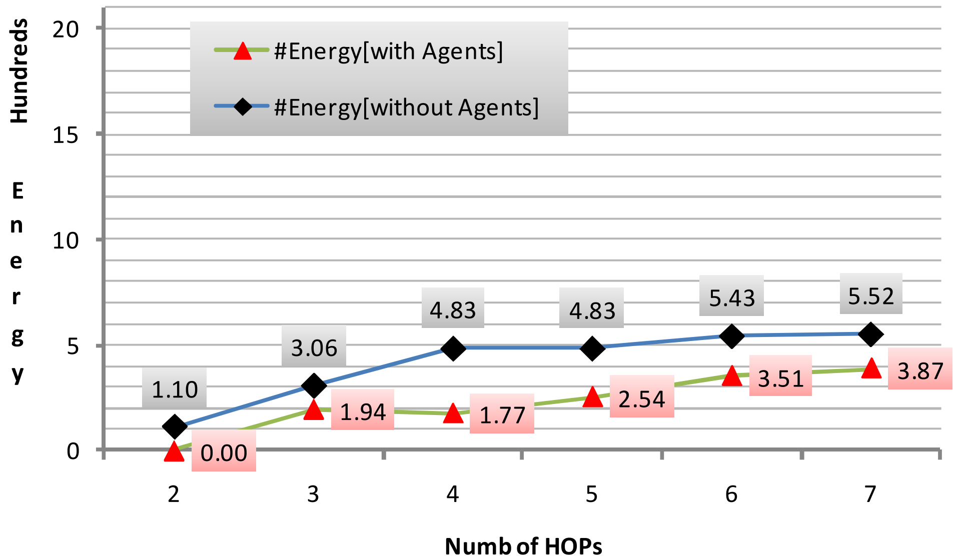


FIGURE 4: The comparison of average energy consumption (in nJ/bit) on nodes in the network using agents and without agents in each successful data transmission with certain number of hops between source and final destination that both located in the same network area.

In the simulation, where source and final destination nodes are located in the same network area, agents also give important results. Communication between source and final destination nodes is set to complete if data packets are received by correct final destination node and it corresponding respond packet already arrived at source node. As the traffic load and number of hops increases, high data collision results in the more energy consumption as well as end-to-end delay. While the scheme with agents' mode gains energy efficiency because of dynamically reducing the number of waking nodes. Again, network with agents' mode can achieve nearly 50% energy conservation. They prevent packets to propagate to other network area. With such agent's capability, the dissipated energy for network to broadcast packets is significantly reduced. It can be shown at Figure 4. Almost for different number of hops, the network with agents has lower dissipated energy at the same hops compared with the other without agents. In obtaining this result, the two similar networks are simulated with same initial topology.

The following chart shows the energy consumption, redundant packets of network and successfully transmission as the number of source nodes changes. Simulation is taken with 40 nodes and constant speed of network at 20 km/h. High mobility of nodes will result in more frequent changing of network topology. Logically, the greater the pause time of nodes (the instant time of node to have static position) is, the more stable network topology is and the routing discovery and maintenance will be simple. The data packets can be persistently transmitted along the existing route and more energy will be saved.

Figure 5 and Figure 6 show the dissipated energy consumption for redundant packets as the number of source nodes of network changes. Redundant packets are received packets by active nodes and then relayed to next hop in order to reach final destination, but these packets are not arrived at correct destination or final destination nodes. These packets are existed during wireless network with broadcast mechanism. The behavior of packets is illustrated either pair of source and final destination nodes located in the same network area or different network area. Each of that location determines different results. We can observe that the energy of network is proportional to the number of source nodes and their neighborhood in network. With the source nodes increasing, the energy of network grows gradually. Since the shortest path based routing scheme uses the maximal transmission energy to send data packets, the increasing number of nodes will not extend the survival time of network. As mentioned in section 3, the topology based routing adds some packets to gain knowledge of nodes' neighborhood, which increases the size of control packets. As is shown in Figure 5, in the case of network with 40 nodes, the network without agents' scheme consumes more energy because the routing control message guide packets to propagate with more active nodes and brings additional energy consumption.

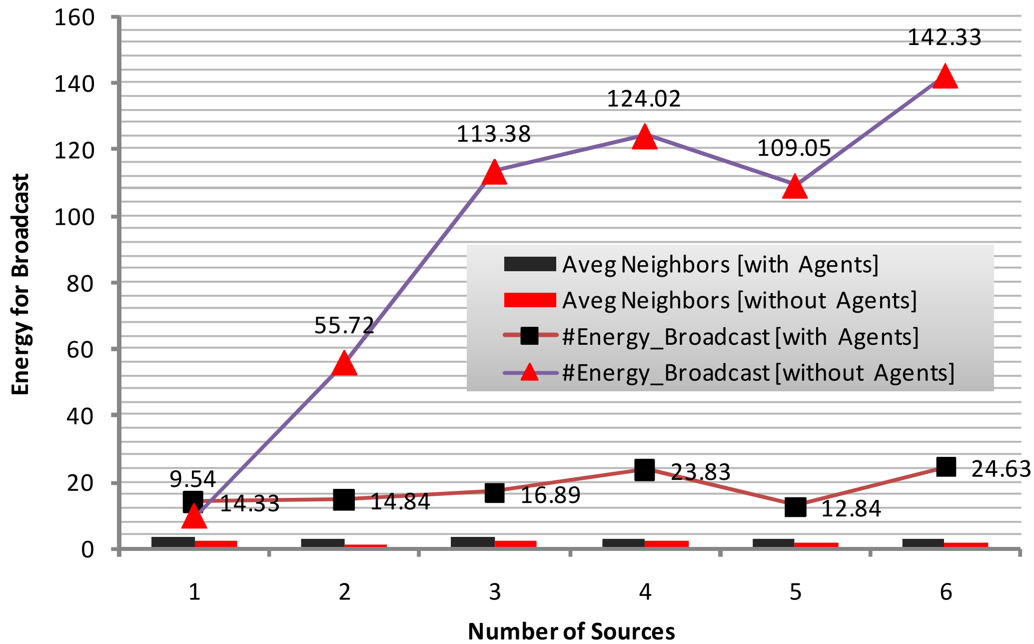


FIGURE 5: The comparison of total Energy dissipated during broadcast of redundant packets for nodes in between the network with agents and without agents (with pair of source and final destination nodes are located in different network area).

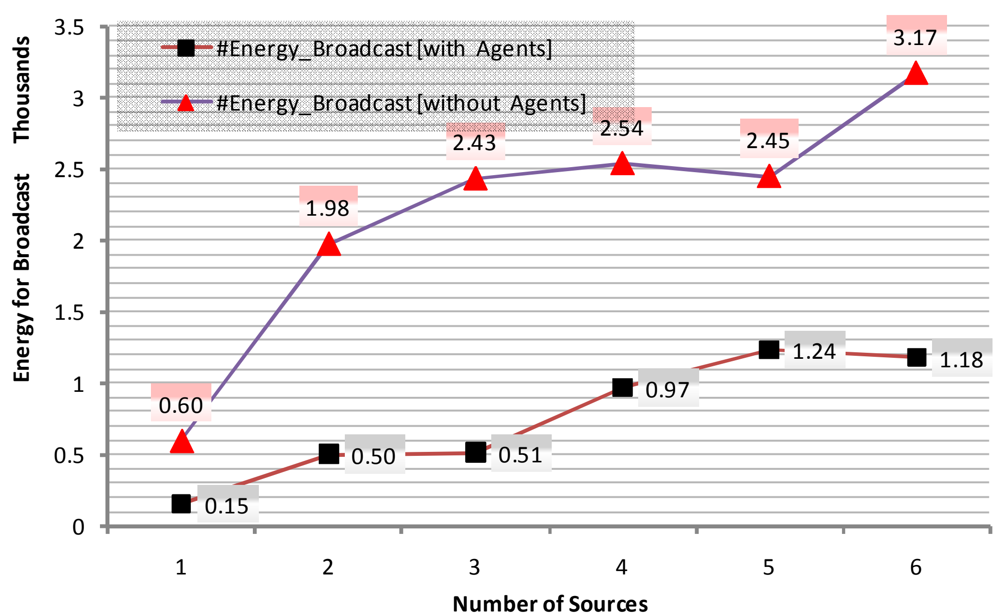


FIGURE 6: The comparison of total Energy dissipated during broadcast of redundant packets for nodes in between the network with agents and without agents (with pair of source and final destination nodes are located in same network area).

First, Figure 5 is considering the situation where the pair of source and final destination is located in different network area. The chart with red marker value shown in Figure 5 is 10% lower than real value for #Energy_Broadcast [without Agents] to make different view with the small value of other with Agents. The chart has slope when number of sources is 5 because the average neighbor in that situation is come down for about 30% compared with the previous. Agents have affected the whole network in reducing broadcast flooding of unnecessary packets. Agents will not propagate packets to next relay nodes if final destination cannot be reached. If such situation happened, agents inform the source to stop the broadcast. They segment the network to make effective the data transmission. Second, even if the source and final destination nodes located in the same network area, agents still play an important role, as shown in Figure 6.

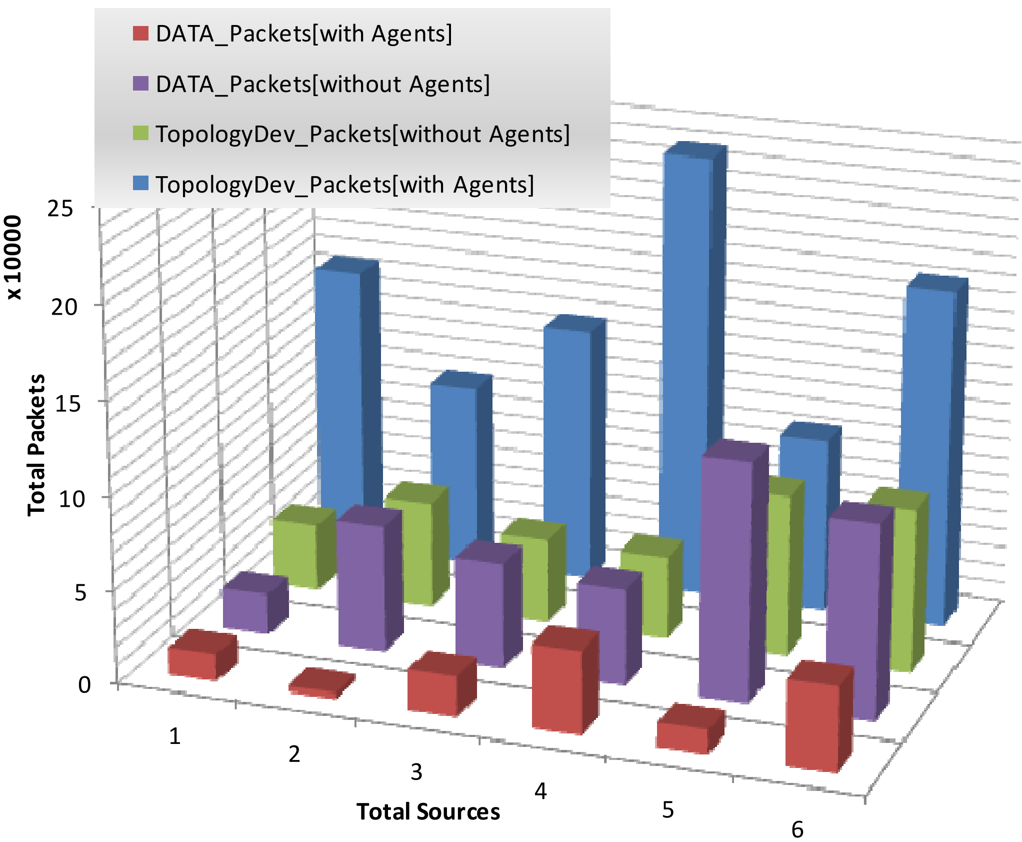


FIGURE 7: Total redundant packets for different number of sources of simulation with agents and without agents (with pair of source and final destination nodes are located in different network area).

These total redundant packets in the network with agents and without agents are illustrated in Figure 7. The previous energy chart in Figure 5 and Figure 6 are made clearer with the total redundant packets propagated during simulation. Although the network with agents is flooded with overhead packets, it experiences fewer redundant data packets (packet with data contained inside). Overhead packets in network with agents are required to prepare agents with knowledge about network topology at its both sides. Depending on nodes' position and neighborhood, agents may or may not able to have route paths to all nodes. If source cannot contact agents to reach final destination, it gives up sending packets. With the same situation, if final destination node cannot be reached by agents, immediately agents notify the source. Such this agent's behavior, it gives significant effect because most data packets have size longer. If nodes receive and relay redundant packets, they will consume energy which actually could be saved for other execution (useless).

5. DISCUSSION

In this research, we focus on the impact of energy models on the performance of data packets transmission in the large MANET that segmented into two domain areas. Significant academic and industrial research that led to the development of a variety of MANET protocols, and also the development of platforms and architectures for reliable communication provided by the MANET assumed that reliable communication mostly be provided in the single uniform network. Research approached in the field, like SANDMAN [34] and DEAPSpace [33], is done by grouping nodes with similar mobility patterns into clusters in one network area; where in each cluster/group, one of the nodes (called cluster head) stays awake permanently and answers discovery requests. However, by splitting the network into independent layer 3 domains, it has been shown that the small domains allow energy dissipated, average energy consumption, and flooding of overhead packets to operate limitary on fewer nodes, with cross-domain interaction only through overlapped area nodes. This division has several key benefits. First, it reduces overall protocol overhead. Second, it made network life time longer. Nodes have salient feature of energy-constrained devices. The battery of node is depleted by: (i) computational processing and (ii) transmission/reception of signal to maintain the signal-to-noise ratio above a certain threshold. Although the energy consumption by computations can be further reduced with new developments in low power devices, the energy consumption by communications cannot be overcome. Therefore, partitioning network into smaller domain is essential to develop efficient networking algorithms and protocols that are optimized for energy consumption. As a consequence, when partitioning a big network into domains, there are some engineering rules that need to be taken into account. For example, the partitions should minimize the expected traffic among different areas so as to not overload the overlapped area; there must be at least one path between every pair of nodes in certain area and the nodes in overlapped area, and at least one path must exist from overlapped area to nodes that belong to other area for such transmission exchange-area to be successfully take place; there must be at least one node that existed in the overlapped area to act as an relay exchange-area domain. Such nodes are called agents. Agents ensure that no single area suffers adversely from a disproportionately large volume of overhead and data packets.

As in line with our previous research results [1], this proposed framework implemented broadcast oriented scheme to construct and maintain such network topology. Compared with topology control oriented scheme, this framework emphasize the broadcast process from a given source node by means of the minimum-energy broadcast tree. It has condition that the source can reach every node of the network. While other researcher, e.g. [32], describes a localized protocol where each node requires only the knowledge of its distance to all neighboring nodes and distances between its neighboring nodes (or, alternatively, geographic position of itself and its neighboring nodes), while our proposed framework optimized the broadcast mechanism by means of energy level and distance metrics. We compare the performance of the network life time that consist of two network areas either with agents or without agents. The performance of both schemes degrades while number of source nodes increase. Especially, the energy consumption of the network without agents grows drastically under high traffic, which results in poor survival time of network. When the traffic load is high, more packets need to be transmitted. In the network with agents, the propagation mechanism is analyzed before relaying packets so as to effectively save energy. Furthermore, since the residual energy of nodes is also considered in such routing function, it can achieve load balance and extend life-span of network. It needs to propose the optimal transmitting energy of the nodes in the routing setup stage. This mechanism will be covered in the next research.

6. CONSLUSION & FUTURE WORK

In this paper, we present an energy management model which can effectively reduce energy consumption in MANET. Mobile nodes in such model transfer packets between pairs of source and final destination nodes in multiple network area modes. Simulation is intensively conducted and it triggered by communication events execution. Base on the agents' existence, we evaluate routing protocol with joint function considering both transmission energy and agents' connectivity. This framework is analyzed and compared its performance to similar network without existence of agents. Simulation experiments show that the proposed agents with energy control mode can achieve higher performance and extend the life-span of network. In the future research, we will try to effectively reduce the end to-end delay with agents' capability, evaluate the agents with higher mobility of nodes, and proposed optimal transmitting energy in the routing setup.

7. REFERENCES

[1] Arai, Lipur S., "Energy Behavior in Ad Hoc Network Minimizing the Number of Hops and Maintaining Connectivity of Mobile Terminals Which Move from One to the Others", International Journal of Computer Networks (IJCN), Volume(2): Issue (6), 2011.

[2] Wang Yu, "Study on Energy Conservation in MANET", Journal of Networks, Vol. 5, No. 6, June 2010.

[3] J. Cartigny, D. Simplot and I. Stojmenovie, "Localized Minimum-energy Broadcasting in Ad HOC Networks," Proceedings of the IEEE Infocom 2003, IEEE, San Francisco, 2003, pp. 2210-2217.

[4] R. Ramanathan and R. Rosales-Hain, "Topology Control of Multihop Wireless Networks using Transmit Power Adjustment," Procedings of IEEE INFOCOM 2000, March 2000.

[5] T.A. ElBatt, S.V. Krishnamurthy, D. Connors, and S. Dao, "Power Management for Throughput Enhancement in Wireless Ad-Hoc networks," Procedings of IEEE ICC'00, New Orleans, LA, June 2000.

[6] P. Bergamo, "Distributed power control for energy efficient routing in ad hoc networks", Wireless Networks, 2004, pp. 29 – 42.

[7] Benjie Chen, Kyle Jamieson, Hari Balakrishnan, Robert Morris, Span. "*An energy-efficient coordination algorithm for topology maintenance in Ad Hoc wireless networks*". Proceedings of the 7th annual international conference on Mobile computing and networking, Rome, Italy, July 2001, p.85-96.

[8] Chang, J., and Tassiulas, L. "*Energy Conserving Routing in Wireless Ad Hoc Networks*". In Proceedings of IEEE INFOCOM, Tel Aviv, Israel, 2000.

[9] Chang-Woo Ahn, Sang-Hwa Chung, Tae-Hun Kim, Su-Young Kang. *"A Node-Disjoint Multipath Routing Protocol Based on AODV in Mobile Adhoc Networks"*. In: Proceeding of Seventh International Conference of Information Technology ITNG2010: 828-833, April 2010.

[10] F. Ingelrest, D. Simplot-Ryl. "*Localized broadcast incremental power protocol for wireless ad hoc networks*". Proceedings of the 10th IEEE Symposium on Computers and Communications, Cartagena, Spain, 2005, pp. 28–33.

[11] Feeney, L., and Nilsson, M. "*Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment*". In Proceedings of IEEE INFOCOM, Anchorage, AK, 2001.

[12]Fenner, W. "*Internet Group Management Protocol*", Version 2, Nov 1997. RFC-2236.

[13] Heinzelman, W., Chandrakasan, A., Balakrishnan, H. *"Energy-efficient communication protocol for wireless microsensor networks"*. In: Proceedings of the 33rd International Conference on System Sciences (HICSS): 1–10, 2000.

[14] IEEE Std. 802.11-1999, Part11. *"Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications"*. IEEE Std. 802.11, (1999).

[15] J. Wu and F. Dai. "*Broadcasting In Ad Hoc Networks Based On Self-Pruning*". Int'l J. Foundations of Computer Science, vol. 14, no. 2, Apr. 2003, pp. 201-221.

[16] J. Wu and F. Dai. "*Efficient broadcasting with guaranteed coverage in mobile ad hoc networks*". IEEE Trans. Mobile Comput. 4 (3) (2005), pp. 1–12.

[17] J. Wu and F. Dai. "*Mobility-sensitive topology control in mobile ad hoc networks*". IEEE Trans. Parallel Distributed Syst. 17 (6) (2006), pp. 522–535.

[18] Javier G., Andrew T. C., Mahmoud N., Chatschik B. *"Conserving Transmission Power in Wireless Ad Hoc Networks"*. Network Protocols Ninth International Conference on ICNP: 24-34, Nov 2001.

[19] Julien Cartigny, David Simplot , Ivan Stojmenovic. "*Localized Minimum-energy broadcasting in ad-hoc networks*". Twenty-Second Annual Conference of the IEEE Computer and Communication society, INFOCOM 2003.

[20] Masato, Tsuru. *"Simulation-based Evaluation of TCP Performance on Wireless Networks"*. Journal of the Japan Society for Simulation Technology: 67-73, 2009.

[21] Moreno M.T., C. Steven, E. Felix Schmidt, Hartenstein H. *"IEEE 802.11-based one-hop broadcast communications: understanding transmission success and failure under different radio propagation environments"*. In: Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems, SESSION: Modeling and performance evaluation II (physical and link layer), ISBN:1-59593-477-4: 68 – 77, 2006.

[22] N. Bansal, Z. Liu. "*Capacity, delay and mobility in wireless ad-hoc networks*". INFOCOM, San Franciso, USA, 2003, pp. 1553–1563.

[23] Nozomu, Nemoto. *"Consideration and Evaluation of Wireless Mesh Network"*. Nomura Research Institute (NRI) Pacific Advanced Technologies Eng.: 70-85, 2006.

[24] Prasanthi. S, Sang-Hwa Chung. *"An Efficient Algorithm for the Performance of TCP over Multi-hop Wireless Mesh Networks"*. In: Proceeding of Seventh International Conference of Information Technology ITNG2010: 816-821, April 2010.

[25] Robin Kravets, P. Krishnan. "*Power management techniques for mobile communication*". Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking, p.157-168, Dallas, Texas, United States, October 25-30, 1998.

[26] Rodoplu, V., and Meng, T. H. "*Minimum Energy Mobile Wireless Networks*". In Proceedings of the IEEE International Conference on Communications (ICC), Atlanta, GA, vol. 3, pp. 1633-1639, June 1998.

[27] T. Camp, J. Boleng and V. Davies. "*A survey of mobility models for ad hoc network research*" Wireless Commun. Mobile Computing, special issue on mobile ad hoc networking: research, trends and applications 2 (5) (2002), pp. 483–502.

[28] Y.C. Tseng, S.Y. Ni, Y.S. Chen, and J.P. Sheu. *"The broadcast storm problem in a mobile ad hoc network"*. Wireless Networks, 8(2/3):153–167, Mar-May 2002.

[29] Ya Xu , John Heidemann , Deborah Estrin. "*Geography-informed energy conservation for Ad Hoc routing*". Proceedings of the 7th annual international conference on Mobile computing and networking, Rome, Italy, p.70-84, July 2001.

[30] Romit RoyChoudhury, Krishna Paul, and S. Bandyopadhyay, "MARP: A Multi-Agent Routing Protocol for Mobile Wireless Ad Hoc Networks" Autonomous Agents and Multi-Agent Systems (Kluwer), 8 (1): 47-68.

[31] Romit RoyChouldhury, S. Bandyopadhyay and Krishna Paul, "A Mobile Agent Based Mechanism to Discover Geographical Positions of Nodes in Ad Hoc Wireless Networks", Accepted in the 6th Asia-Pacific Conference on Communications (APCC2000) , Seoul, Korea, 30 Oct. - 2 Nov. 2000

[32] F. Ingelrest, D. Simplot-Ryl. "Localized broadcast incremental power protocol for wireless ad hoc networks". Proceedings of the 10th IEEE Symposium on Computers and Communications, Cartagena, Spain, 2005, pp. 28–33.

[33] M. Nidd, "Service Discovery in DEAPspace", IEEE Personal Communications, August 2001, pp. 39-45.

[34] G. Schiele, C. Becker and K. Rothermel, "Energy-Efficient Cluster-based Service Discovery for Ubiquitous Computing," Proc. ACM SIGOPS European Workshop, Leuven, Belgium, Sept. 2004.

A Critical Note On Ring Flushing

Syed Muhammad Atif syed.muhammad.atif@gmail.com
M.S Computer Networks
Department of Computer System Engineering
Usman Institute of Technology
Karachi, Pakistan.

Abstract

Rapid Spanning Tree Protocol enable switches to flood incoming frame with broadcast or unknown unicast destination address even in switched Ethernet networks having redundant links. It also allows switches to secretly learn location of connected devices in such networks. However some of those learnt location may become stale if a topology change is detected by RSTP and need to be flushed by switches in the network. It is found that standard address flushing technique of RSTP flushes too many addresses from large number of switches after a topology change. As a result there is a sudden massive increase in flooding traffic which may cause network-wide congestion, frame delay and frame loss. Recently a new address flushing technique named as Ring Flushing was proposed for RSTP that flushes addresses from small number of selective ports of very selective switches and so dramatically reduces the amount of flooding traffic after a topology change. However, numbers of flaws are identified in the current implementation of this newly proposed technique. This paper will not critically discuss the flaws in the current implementation of Ring Flushing but will also propose there simple yet effective solutions.

Keywords: Ring Flushing, Frame flooding, Network Scalability, RSTP compatibility.

1. INTRODUCTION

The current standard spanning tree protocol – Rapid Spanning Tree Protocol – is an indispensable management protocol for switched Ethernet networks. It constructs, in a distributive manner, an overlying logical tree spanning over all the switches in the underlying physical topology. The protocol is very essential for switch for flooding unknown broadcast or unicast frame and for learning addresses by inspecting incoming frame because of two reasons. First, RSTP prevents flooded broadcast or unknown unicast frame from persisting forever in the network. Second, RSTP make it feasible for a switch to learn location of devices in the network by inspecting incoming frames by ensuring that there always exist one and only one path from a switch to a particular device at any particular instant of time. Learnt locations of network connected devices are temporarily cached into a table called forwarding table.

Rapid Spanning Tree Protocol [1] is specifically designed to converge switched Ethernet network after topology change. It is basically the extension of its ancestor protocol Spanning Tree Protocol [2] (STP), first proposed by Perlmen in [3]. The enhancements and extensions incorporated in RSTP were first proposed by Mick Seaman in [4],[5],[6] and [7]. RSTP, due to their modifications, can converge within 1-3s [8] sharply in contrast to its ancestor STP which may take as much as 50s [9].

To compute the spanning tree, RSTP assigns a unique Identifier to each switch in the network. Each port of a switch also has an identifier unique within the scope of switch. All switches in the network elects a Root Switch and then try to compute and maintain the shortest path to that Root Switch thus creating logical spanning tree for the physical network. Switches blocks all ports, i.e. does not allow ports to transmit or receive data frames, expect the port necessary for itself to get access to the Root Switch through the shortest path (the root port) or ports required to provide the shortest path to neighboring switches (designated ports). Both the root port and designated ports of a switch will eventually moved into forwarding state, a state in which a port is allowed to transmit and receive data frames. A port of a switch **A** is called Backup port if it is connected to a link (network segment) that designated port also belongs to the switch **A**. A port of a switch is considered as an Alternate port if it can provide an alternative path to the Root Switch in case of failure of the root port of the switch. Both Backup and Alternate ports will eventually moved into blocking state.

RSTP [1] uses Bridge Protocol Data Unit (BPDU) for communication, a message to convey information to neighboring switches. It modifies STP in four different ways. First, an RSTP switch can immediately puts its alternate port in forwarding state after failure of its root port [1][4]. Second, an RSTP switch uses a handshaking mechanism called sync to quickly moves a designated port, connected to a point-to-point link, in a forwarding state [1][5]. Third, an RSTP switch can accept even an inferior BPDU on its non-designated port if it is transmitted by a designated bridge through its designated port [1][6]. Four, a Non-Root RSTP switch may generate BPDUs in contrast to STP in which Non-Root switch are only allowed to relay BPDUs originally transmitted by the Root Switch [1].

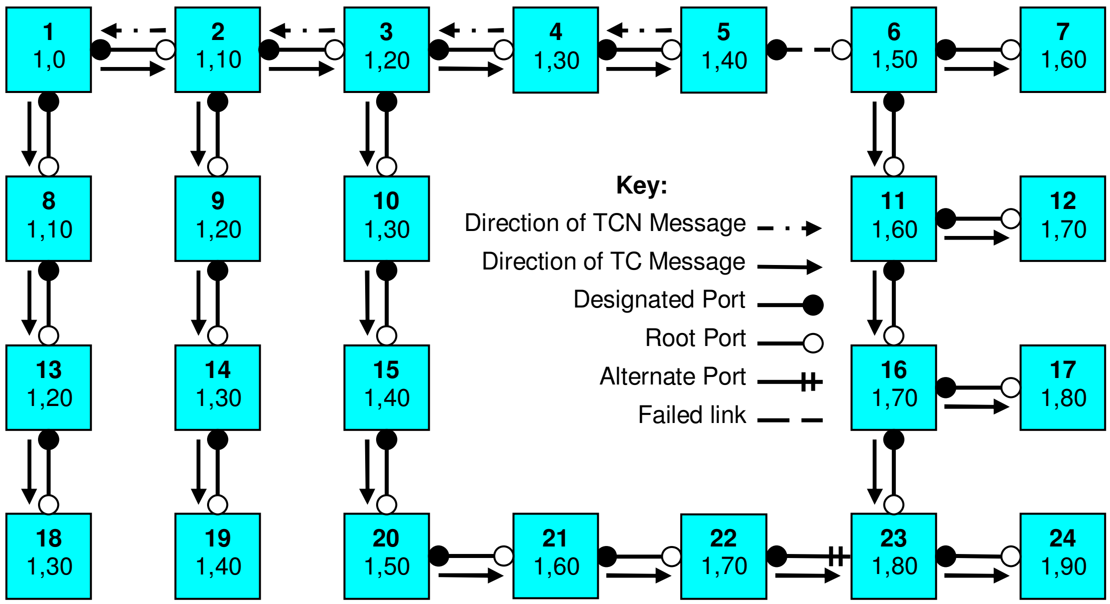
A change in physical topology of an RSTP controlled Ethernet network always urge RSTP to recomputed its spanning tree. This may result in change of position of some network connected devices with respect to some switches and thus making their cached learnt locations stale on those switches. Hence, stale cached learnt locations must be flushed from suffering switches to ensure accessibility to those devices. Unfortunately, the current standard address flushing technique used by RSTP to flush stale cached location unnecessarily flushes too many valid cached locations. Also, the scope of address flushing for this technique is network-wide. Due to this large amount of flushing of large number of switches produces sudden massive flooding traffic, due to unknown unicast frame, resulting in network congestion, frame delay and frame loss. Recently, a new address flushing technique named as Ring Flushing was proposed by Horvath et al. in [10]. This new technique does not only reduce the amount of flushed addresses but also reduces the scope of flushing to very few switches and so have a dramatic direct impact on the amount of flooding traffic. Unfortunately, current implementation of Ring Flushing has number of flaws. This paper will identify the flaws in the current implementation of Ring Flushing and will propose their simple yet effective solutions.

The rest of paper is organized as follows. Section 2 will give an overview of classical address flushing techniques used by STP and RSTP. Section 3 will describes the newly proposed Ring Flushing technique. Section 4 will identifies the flaws, along with their proposed solutions, in current implementation of Ring Flushing. Then section 5 will concludes the paper.

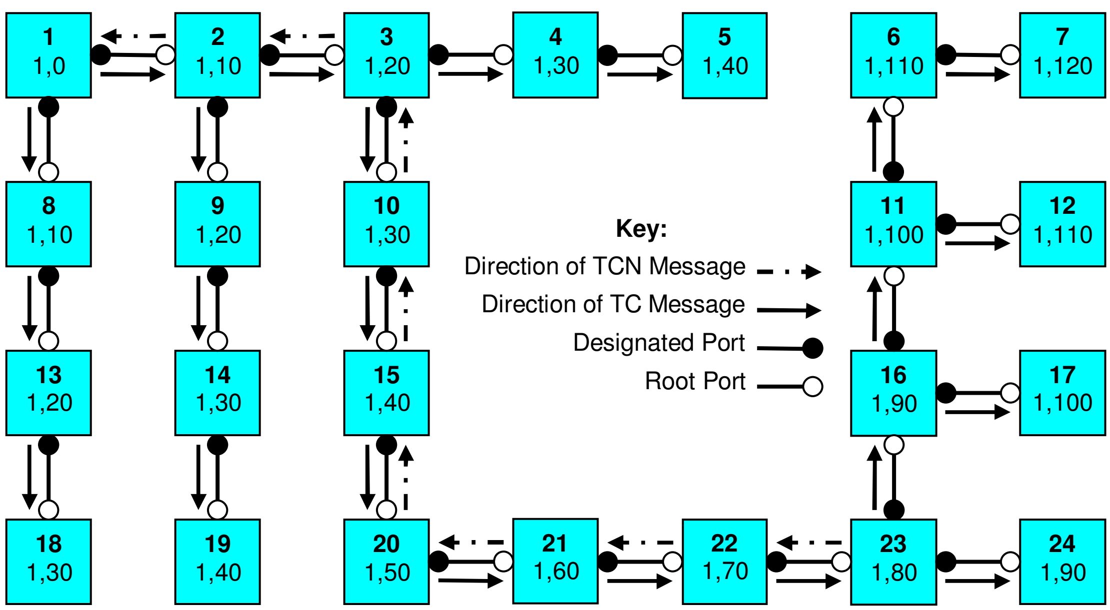
2. CLASSICAL ADDRESS FLUSHING TECHNIQUES

Both STP and RSTP force their probably suffering switches to flush probably stale (invalid) cached learnt locations of network connected devices from their filtering tables after a topology change. But they use slightly different techniques for this purpose. This section will give an overview of address flushing techniques used by STP and RSTP.

In STP both moving of a port into blocking state and moving of a port into forwarding state are marked as topology change. When a Non-Root switch detects a topology change, it transmits a Topology Change Notification BPDU on the link to which its Root Port is attached. This transmission is repeated until the switch receives an acknowledgment from the designated switch for that link. The designated switch passes the notification to, or towards, the Root using the same procedure. If the Root receives such a notification, or detects a topology change itself, it will set a Topology Change flag in all Configuration Messages transmitted for some time. This time is such that all switches will receive one or more of the Configuration Messages. While this flag is set, switches use a short value to age out cached learnt locations in the forwarding table. When the flag is reset again, switches revert to original Ageing Time.



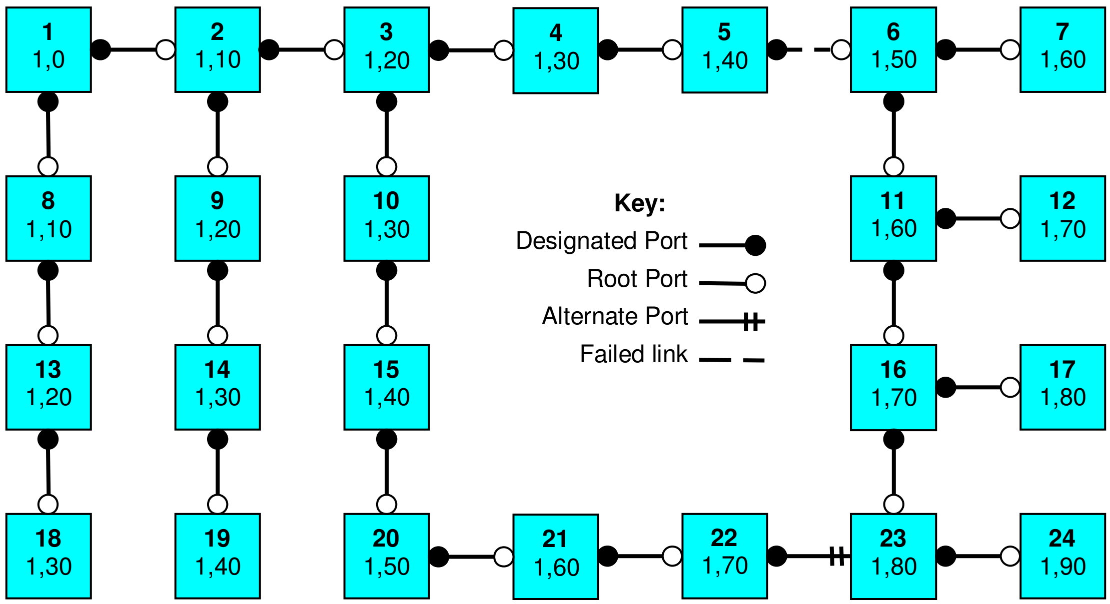
(a). Flow of TC and TCN messages when a designated port of switch **5** and the root port of switch **6** is moving into blocking state due to failure of link between switch **5** and switch **6**.



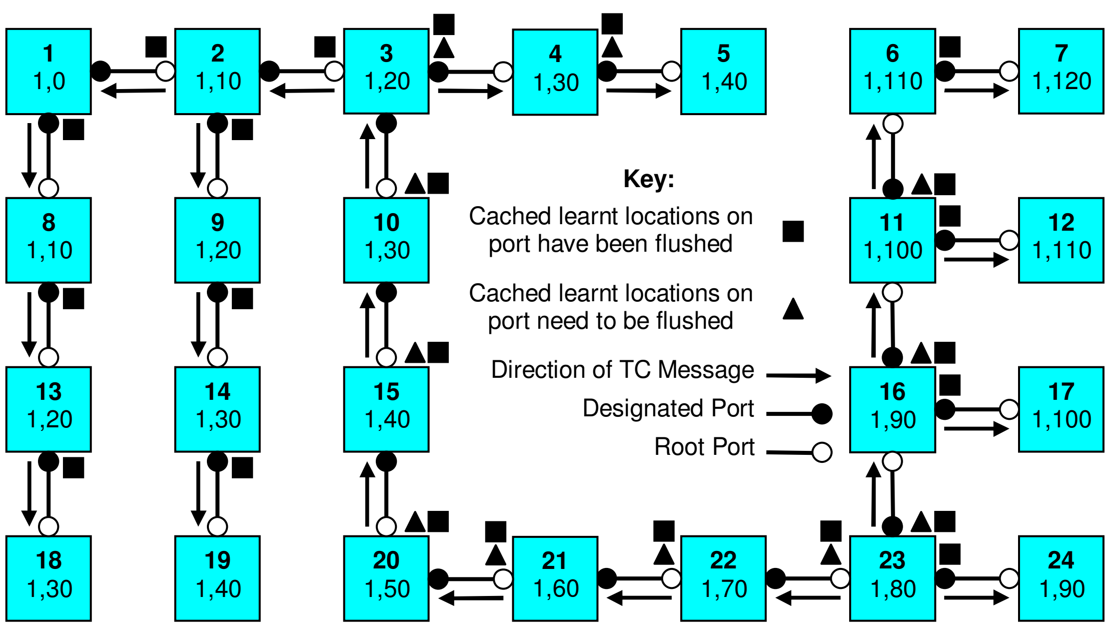
(b). Flow of TC and TCN messages when the new root port of switch **23** is moving into forwarding state to reconverge the network.

FIGURE 1: Flow of TC and TCN messages in STP.

Figure 1 is showing the flow of TCN message and TC message, Configuration BPDU such that its TC flag is set, in STP. In all figures, switches are represented by small boxes. The top number in the box is the Switch Identifier, the lower set of numbers represents the Root Switch Identifier as perceived by the switch and the cost to this Root Switch. It is assumed that all links have cost of 10.



(a). Topology at failure of link between switch **5** and switch **6**



(b). Flow of TC messages when alternate port of switch **23** become the new forwarding root port.

FIGURE 2: Flow of TC messages in RSTP

In RSTP, a topology change is recognized if a port that is previously not part of spanning tree (i.e. Disable, Alternate or Backup port) now becomes part of spanning tree (i.e. Designated or Root port in forwarding state). As such a port may relocate some network connected devices with respect to some switches in the network. So, when a switch detects a topology change it flushes the cached learnt locations in the forwarding table and transmits TC message, RST BPDU such that its TC flag is set, from all active ports (designated or root ports). A switch that receives an TC message on its active port flushes the cached learnt locations associated with all other active port and then propagates TC message through them. Figure 2 is illustrating the flow of TC message in RSTP after a topology change.

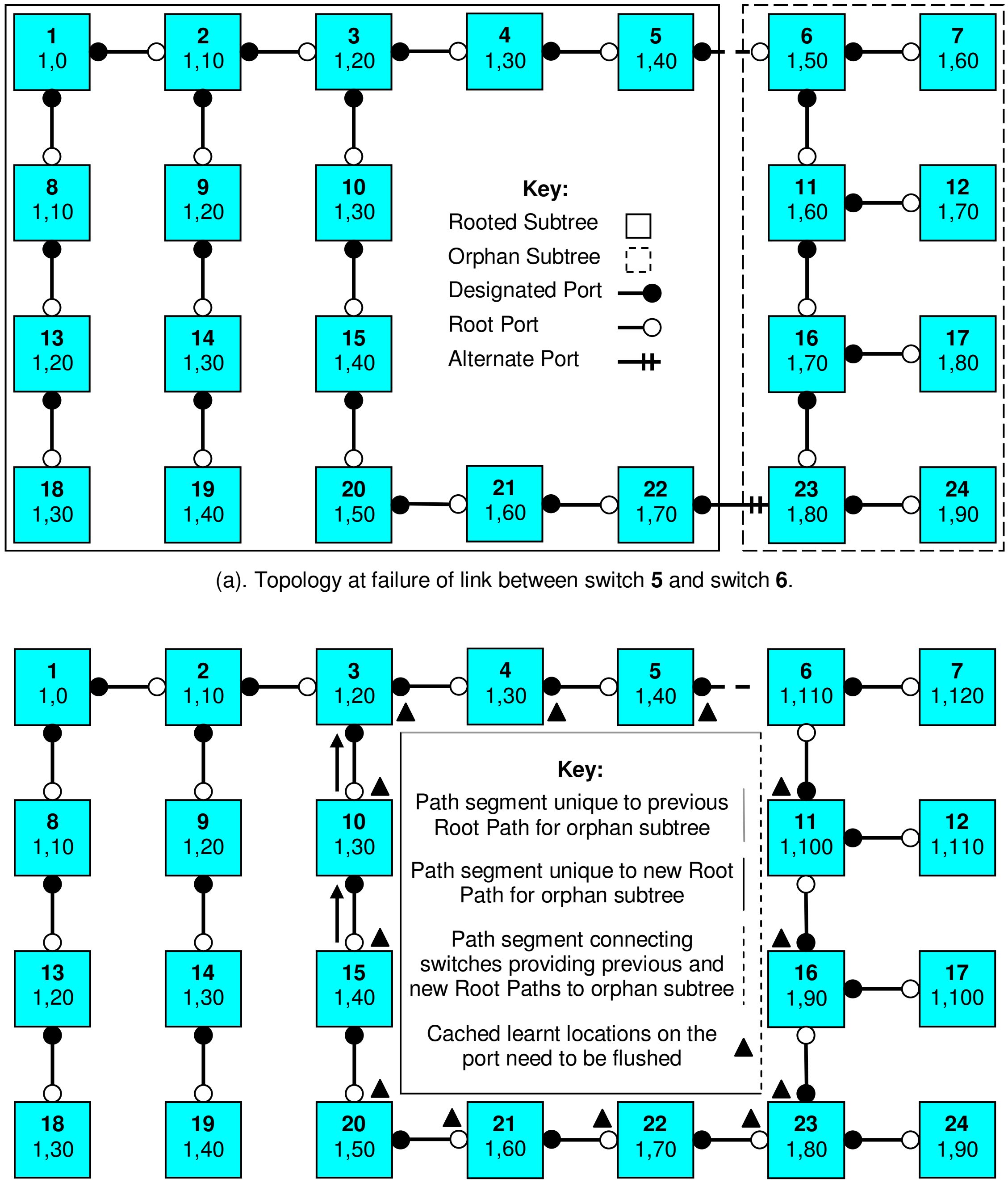
STP uses a centralized approach for address flushing i.e. a switch that detects a topology change first informs its Root Switch and then the Root Switch instructs all the switches in the network to shorten their aging time. In contrast, RSTP uses a decentralized approach i.e. a switch that detects a topology change not only informs its Root Switch but also immediately transmit RST BPDU with TC flag on all its other active ports on behalf of the Root Switch. It helps RSTP to quickly propagate the topology change information to all switches in the network. Moreover, unlike STP, RSTP immediately flushes the probably stale cached learnt locations instead of quickly aging them out.

3. RING FLUSHING

Ring Flushing is an address flushing technique that recently proposed by Horvath et al. The technique is specifically designed to reduce the amount of flooding traffic after a topology change generated due to excessive flushing of cached learnt locations in forwarding table. It is achieved not only by reducing the number of potential switches that may have stale (invalid) cached learnt locations in their forwarding tables but also by reducing the number of potentially stale cached learnt locations on those switches. This section will give an overview of ring flushing technique.

Theoretical Background

When the root port (or designated port associated with the root port) of a switch fails (or stops acting as the root port), it segregates the network into two distinct subtrees namely a *rooted subtree*, a subtree that still have the Root Switch, and an *orphan subtree*, a subtree that no longer have the previous Root Switch. The segregated subtrees can be merged again by turning an alternate port into forwarding designated port or the forwarding root port. If this happens, three path segments are very important with respect to changes locations of network connected devices. First, the path segment unique to previous Root Path for orphan subtree. Second, the path segment unique to new Root Path for orphan subtree. Third, the path segment, in the orphan subtree, between the switches providing the previous Root Path and the new Root Paths to orphan subtree respectively. All cached learnt locations associated with the designated ports along the path segment unique to previous Root Path for orphan subtree are compromised, and so they need to be flushed or removed, because they no longer providing accessibility to network connected devices in orphan subtree. Similarly, all the cached learnt locations associated with the root ports along the path segment unique to new Root Path for orphan subtree are also compromised, and so they also need to be flushed or removed, because they no longer providing accessibility to network connected devices in orphan subtree. Moreover, all cached learnt locations associated with now designated ports (previous the root ports) along the path segment, in the orphan subtree, the switches providing the previous Root Path and the new Root Paths to orphan subtree respectively are also compromised, and so they must be flushed or removed, as they no longer providing accessibility to network connected devices in rooted subtree. The above three mentioned path segments called sector 1, sector 2 and sector 3 respectively in [10]. No other port except the above mention ports requires flushing or removal of their cached learnt locations because all remaining learnt locations are still fresh (valid). If all three above mentioned path segments are present, they form a ring or cycle in the physical topology of the network and address flushing is required only in that ring or cycle. That is why this address flushing technique is named as “Ring Flushing” by Horvath et al. in [10].

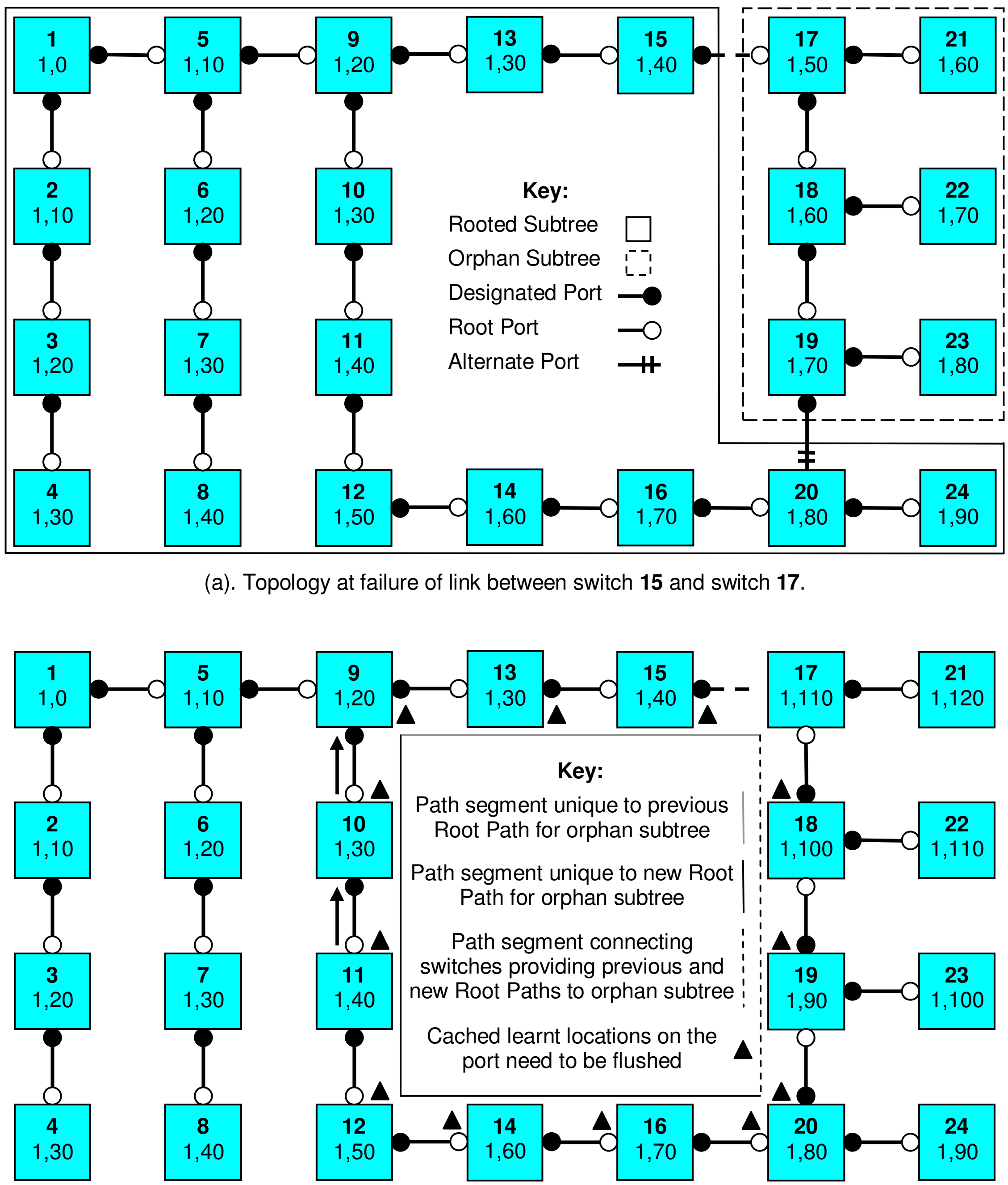


(a). Topology at failure of link between switch **5** and switch **6**.

(b). Topology when alternate port of switch **23** become new forwarding root port to reconverge the network.

FIGURE 3: Three path segments in Ring Flushing along with ports that need address flushing.

No address flushing except on the corresponding Designated port of Backup port is required, If Backup port is used to merged the two segregated subtrees by making it a forwarding designated port. This is because none of the three mentioned path segments are presents. Moreover, a Backup port transitioning to forwarding root port act as if it were previously a Designated port A newly enabled port of a switch that is becoming either the forwarding designated port or the forwarding root port also pretend as if it were previously an alternate port. Figure 3 and 4 are illustrating the ports on the three mentioned path segments that need flushing or removal of cached learnt locations when an alternate port of a switch becomes new forwarding root port or becomes a forwarding designated port respectively.



(a). Topology at failure of link between switch **15** and switch **17**.

(b). Topology when alternate port of switch **20** become forwarding designated port to reconverge the network.

FIGURE 4: Three path segments in Ring Flushing along with ports that need address flushing.

It is not necessary that all three mentioned path segments are present simultaneously in the segregated network having rooted and orphan subtrees. For example the path segment, in the orphan subtree, between the switches providing the previous and new Root Paths to orphan subtree respectively may absent if both the failed root port and the alternate port that is becoming either the forwarding designated port or the forwarding root port are present on the same switch. Similarly, path segment unique to previous Root Path of orphan subtree may absent if the two trees merging together are not part of a single network in the recent past.

Horvath Implementation of Ring Flushing

Inventers of Ring Flushing, Horvath et al., proposed an implementation for Ring Flushing in [10] using three messages of two different types. The two types of messages are an DTC message and an RTC message. A switch receiving an DTC message flushes the cached learnt locations associated with the designated port receiving the message. Whereas a switch receiving an RTC message flushes the cached learnt locations associated with its root port. Two DTC messages are generated respectively on path segment unique to previous Root Path for orphan subtree and path segment between the switches providing previous and new Root Paths to orphan subtree respectively. These two DTC messages are responsible for flushing of cached learnt locations on designated ports along the two paths. An RTC message is generated on path segment unique to new Root Path for orphan subtree and it is responsible for flushing cached learnt locations on the root ports along the path segment. Figure 4 is illustrating the flow of two required DTC messages and a required RTC message in Horvath implementation for Ring Flushing.

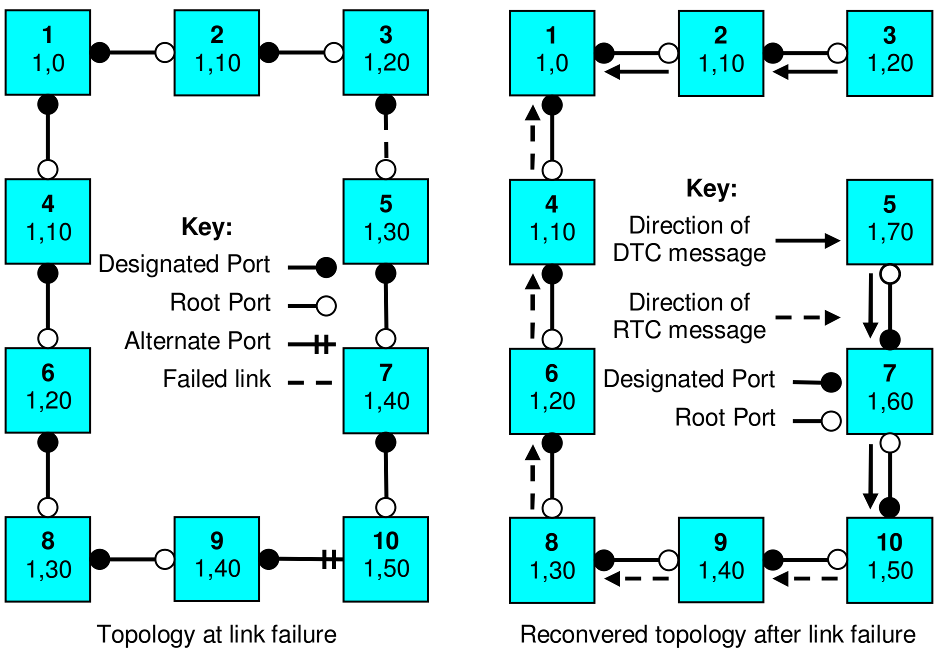


FIGURE 5: Flow of two DTC messages and an RTC message that are required in Horvath implementation for Ring Flushing

To generate and propagate the above mentioned three messages for Ring Flushing, following set of rules are prescribed in [10].

1. A port moving into blocking (discarding) state must generate an DTC message on its (new) root port.
2. A designated port receiving an DTC message must first flushes the cached learnt locations associated with it and then propagates the received DTC message on the (new) root port.
3. A port moving into forwarding state must generate an RTC message on its (new) root port.
4. A designated port receiving an RTC message must first flushes the cached learnt locations associated with the root port and then propagates the received DTC message on the (new) root port.

Above prescribed rules may be implemented by making subtle changes in the RSTP port state machine. As DTC and RTC messages generated only on the root port of switch, so they will eventually terminate at the Root switch.

Limitation of Ring Flushing

Ring Flushing is heavily dependent on generation of DTC message on path segment unique to previous Root Path for orphan subtree because this message is necessary for flushing or removal of cached learnt locations on designated ports along that path segment. The required DTC message may be generated by the root port previously connecting the orphan subtree to rooted subtree. But if the segregation of spanning tree, into orphan and rooted subtrees, is due to failure of the link connecting the two subtrees, then the DTC message will never reach to the path segment unique to previous Root Path for orphan subtree. A solution to this problem is that the designated port associated with the root port previously providing the Root Path to the orphan subtree will generate the DTC message on behalf of that root port. But a designated port can timely sense the physical failure of link connected to the root port if and only if the link is true point-to-point link. Hence, Ring Flushing technique cannot be used in a network having at least one multipoint link. Ring Flushing considers a link as point-to-point link if and only if two switch ports are connected to the link and there is no other network connected device on the link.

4. IDENTIFIED FLAWS IN HORVATH IMPLEMENTATION

Ring Flushing technique can be useful in reducing the amount of flooding traffic after a topology change as it reduces both the number of effective switch and number of cached learnt locations that need to be flushed from those switch. But, it is unfortunately find that current implementation of Ring Flushing presented in [10] and restated in section 3.2 have number of flaws. This section will point out those flaws and give their simple yet effective solutions.

Flaw 1: No Generation Of Required DTC Message In Rooted Subtree

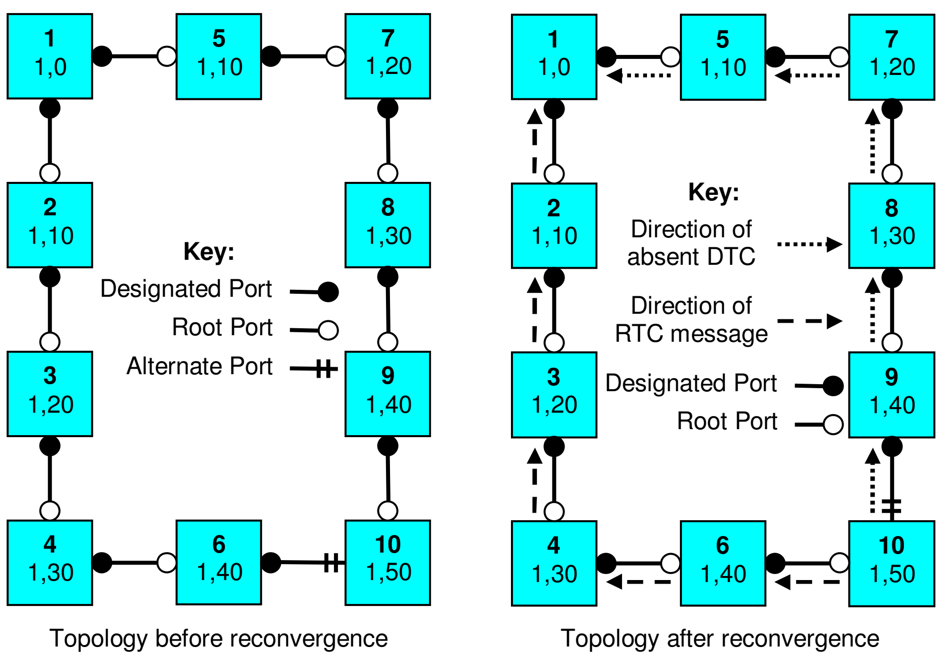


FIGURE 6: Flow of absent DTC message that is not generated by switch **10** when it receives a superior BPDU on its alternate port.

An DTC message is required to be generated on previous Root Path in Ring Flushing to flush stale cached locations associated with designated ports that are along that path segment. However, Horvath implementation of Ring Flushing is unable to generate the required DTC message if the switch silently transitions from current root port to a new root port. The reason is that in current implementation of Ring Flushing the switch transitioning to the new root port generates the DTC message only on its new root port whereas designated port associated with previous root port of the switch unable to generate the required DTC message as that designated port is not experiencing any change in its port state. A switch may silently transitions from current root port to a new root port if it receives a superior BPDU (better information) on one of its port as depicted in figure 6.

Flaw 2: No Generation Of Required DTC Message In Orphan Subtree
An DTC message is required to be generated by the switch previously providing the Root Path to orphan subtree. This message is necessary to flush stale cached locations associated with now designated ports (previous root ports) that are along the path segment, in orphan subtree, between the switches providing previous and new Root Paths to orphan subtree respectively. However, in the Horvath implementation of Ring Flushing, the responsible switch (the switch previously providing the Root Path to orphan subtree) is unable to generate the required DTC message. The reason is that in RSTP a Non-Root switch declares itself the Root Switch as soon as its root port fails and it has no alternate port (Clause 17.6 IEEE 802.1D 2004). This newly declared Root Switch of orphan subtree can later restore connectivity with the rooted subtree by transitioning one of its forwarding designated port to the root port. However, no DTC message is not generated by the switch, as illustrated in figure 7, because, in RSTP, the port does not necessarily change its state when it transitions into the root port (Clause 17.29.2 IEEE 802.1D 2004).

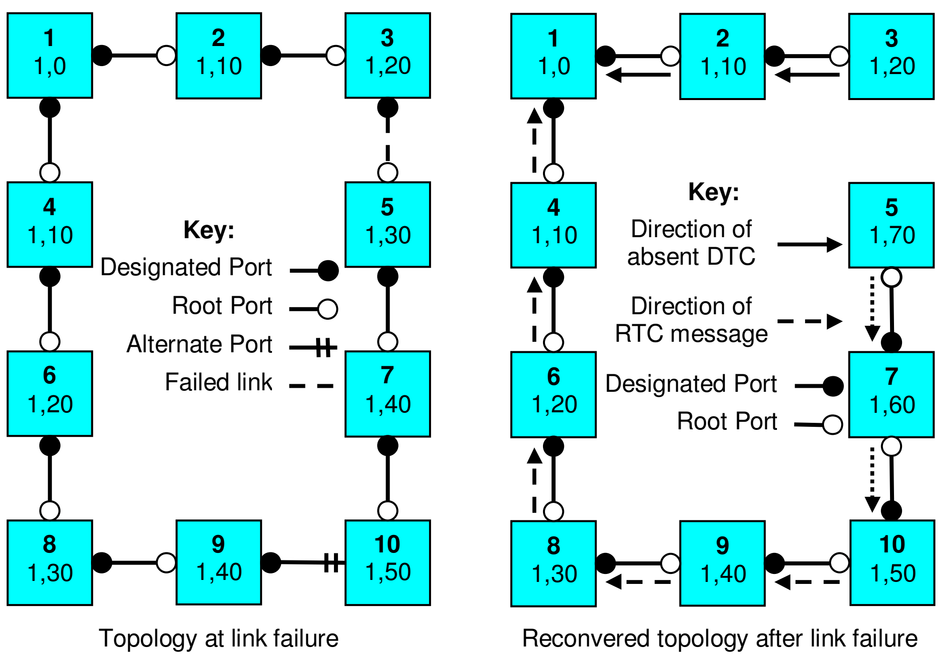


FIGURE 7: Flow of absent DTC message when link between switch **3** and switch **5** fails.

Flaw 3: Generation Of Unnecessary DTC And RTC Messages
In Horvath implementation of Ring Flushing, unnecessary DTC and RTC messages are generated by all the switches on the path segment between the switches, in orphan subtree, providing previous and new Root Paths to orphan subtree respectively. The reason is three folded. First, Horvath implementation of Ring Flushing urges a switch to generate DTC message (RTC message) on its (new) root port whenever it moves into blocking state (forwarding state).

Second, an RSTP switch has tendency to temporarily block its port to prevent loops when the port is transitioning its role from Root Port to Designated Port (Clause 17.29.3 IEEE 802.1D 2004). Third, all the switches on the path segment between the switches, in orphan subtree, providing previous and new Root Paths to orphan subtree respectively are transitioning their respective previous root ports into designated ports. As a result these designated ports are temporarily reverting back to blocking state forcing their respective switches to generate unnecessary DTC messages (see Figure 8). Later these designated ports are promoted again to forwarding state and thus subsequently forcing their respective switches to generate unnecessary RTC messages (see Figure 8).

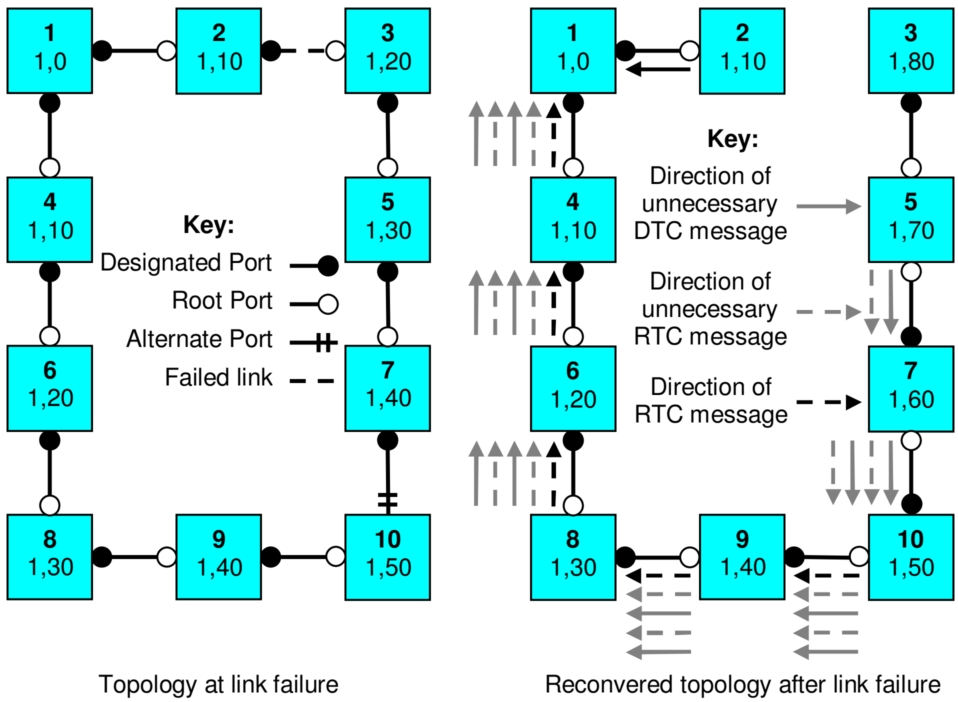


FIGURE 8: Flow of unnecessary DTC and RTC messages when link between switch **2** and switch **3** fails.

5. PROPOSED SLOUTION: ROLE BASED RULES FOR RING FLUSHING

Rules for Horvath implementation of Ring Flushing are port state based i.e. DTC and RTC messages are generated only on sensing a change in port state. The above mentioned three flaws can be removed by using new port role based rules for Ring Flushing instead of current port state based rules. Here are port role based rules for Ring Flushing.

1. A port flushes its associated cached learnt locations and generates an DTC message on the current root port when its role transitions from forwarding Designated Port to Disabled Port or Alternate Port.
2. A port flushes its associated cached learnt locations and generates an DTC message on itself if its role is transitioning from forwarding Root Port to Disabled Port or Alternate Port.
3. A port generates an RTC message on the current root port when its role transitions from Disable or Alternate to Designated or Root and the current state of port is forwarding.
4. A port flushes its associated cached learnt locations if its role is transitioning from Root Port to Designated Port.
5. A designated port that is receiving an DTC message flushes its associated cached learnt locations and propagates the received DTC message on the current root port.
6. A designated port that is receiving an RTC message flushes its associated cached learnt locations and propagates the received RTC message on the current root port.

Rule 2 ensures that the forwarding root port must inform its designated port when it becomes Alternate or Disabled to eliminate flaw 1. Rule 4 ensures that now designated ports (previous root ports) flushes their cached learnt locations without any need of DTC message and thus eliminate flaw 2. This rule also prevents generation of unnecessary DTC messages and thus partially eliminates the flaw 3. Rule 3 ensures that only the switch providing the new Root Path to orphan subtree will generates RTC message and thus prevent unnecessary generation of RTC messages and so subsequently partially eliminating flaw 3. Therefore, the new port role based rules for Ring Flushing are well protected against all three identified flaws unlike the port state based rules presented in [10].
It is not necessary for a port to generate an DTC message if it transitions its role from Designated to Disabled or Alternate but its neighboring port on the attached point-to-point link have Alternate or Backup role. This type of port will be called Passive Designated Port in this text. On the other hand, a port must generate an DTC message if it transitions its role from Designated to Disabled or Alternate and its neighboring port on the attached point-to-point link have Root role. This sort of port will be called Active Designated Port in this text. Hence, rule 1 from above six role based rules for Ring Flushing can modified as.

1. A port flushes its associated cached learnt locations and generates an DTC message on the current root port when its role transitions from forwarding Active Designated Port to Disabled Port or Alternate Port.

6. BACKWARD COMPATIBILITY WITH LEGACY SWITCHES

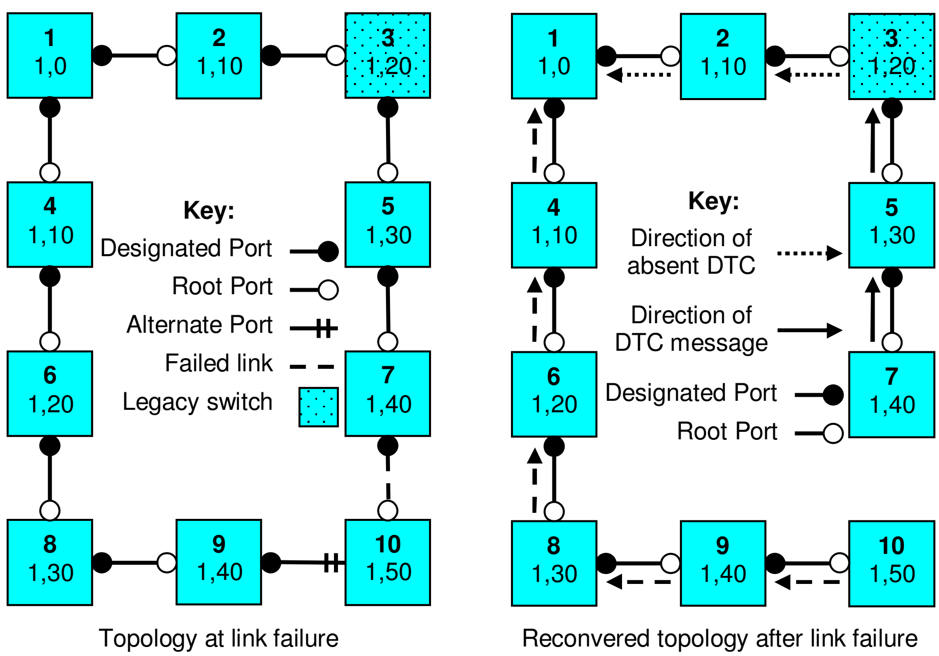


FIGURE 9: Flow of absent DTC message due to its interception by a legacy RSTP switch.

Horvath et al., inventers of Ring Flushing technique, proposed encoding of RTC message and DTC message into TC flag and unused TCAck flag of RST BPDU respectively. Moreover, they claimed that this encoding of DTC and RTC messages within RSTP BPDU ensures backward compatibility with legacy RSTP switches as no new BPDU was introduced. However, it is unfortunately not true. It is because of two reasons. First, a legacy RSTP switch receiving an DTC message on its designated port is unable to flush cached learnt locations associated with that port because set DTC flag in RST BPDU will be perceived as TCAck flag by receiving legacy RSTP switch. Second, the received DTC message encoded in RST BPDU will not be propagated further on the root port by receiving RSTP switch as TCAck flag is always clear in RST BPDU (Clause 17.21.20). As a result not only the legacy RSTP switch but also all its upstream switches, whether they legacy RSTP switches or Ring Flushing capable RSTP switches, till the Root Switch are unable to flush their designated ports along the previous Root Path for orphan subtree (See Figure 9). Hence, a more comprehensive mechanism is required to ensure backward compatibility. This section will present a mechanism that will not only provide backward compatibility with STP and RSTP switches but it will also allow multipoint links in the network.

The proposed mechanism divides each port on Backward Compatible Ring Flushing RSTP switch into two type namely Ring Flushing port and legacy port. The root port, an alternate port, a backup port and an edge designated port are all always considered as Ring Flushing port (RF port). Only non-edge designated port may or may not be an RF port. All non-edge designated ports are by default considered as legacy active designated port. A legacy active designated port connected to a point-to-point link becomes an RF active designated port when all the ports on its neighboring switch are RF ports and revert back to legacy active designated ports as soon as there is a legacy port on the neighboring switch. This information can be communicated through a newly introduced RF BPDU. The new BPDU will contain two flags specifically a “Total Ring Flushing switch” Flag and an DTC flag as depicted in Figure 10. Legacy active designated ports in “Backward Compatible Ring Flushing RSTP switch” will behave like designated ports in an RSTP switch i.e. flushes their associated cached learnt locations when an TC message is received on one of the designated ports or the root port of switch, or one of the designated ports or the root port of switch fails. While RF ports will behave like “Ring Flushing RSTP switch” i.e. flushes their associated cached learnt locations according to rules for ring flushing defined in section 5.

Total Ring Flushing Switch Flag
DTC Flag

FIGURE 10: Structure of Ring Flushing BPDU

The above mentioned backward compatibility mechanism will work because it ensures a legacy active designated port connected to a point-to-point link becomes an RF designated port if only if it can receive all the DTC messages transmitted by their downstream switches i.e. there is no legacy RSTP switch in its downstream. Moreover, there is no need of transmitting a separate RTC message because a legacy TC message, RST BPDU with TC flag, has capability to flush cached learnt locations associated with the respective root ports of switches along the new Root Path for orphan subtree. Hence, rules for “Backward Compatible Ring Flushing RSTP switch” can be summarized as:

1. The root port, an alternate port, a backup port, a disabled port and an edge designated port is always an RF port.
2. A port transitioning into Designated role is always assumed as a legacy active designated port.
3. A designated port becomes an RF active designated port when it receives an RF BPDU such that its “Total Ring Flushing Switch” flag is set.
4. A designated port becomes legacy active designated port as soon as it receives an RF BPDU such that its “Total Ring Flushing Switch” flag is clear.
5. A switch sets the “Total Ring Flushing Switch” flag of RF BPDU and transmits it on the current root port as soon as its all ports become RF ports.
6. A switch having only RF ports clears the “Total Ring Flushing Switch” flag of RF BPDU and transmits it on the current root port as soon as one of its RF ports becomes legacy port.
7. A port flushes its associated cached learnt locations and generates an DTC message on the current root port of the switch when its role transitions from forwarding RF active Designated Port to Disabled Port or Alternate Port or it become disputed designated port, a disputed designated port is a designated port in forwarding state such that its neighboring port on point-to-point is also a designated port with forwarding state.
8. A port flushes its associated cached learnt locations and generates an DTC message on itself if its role is transitioning from forwarding Root Port to Disabled Port or Alternate Port.

9. A port generates an TC message on the current root port and its all legacy active designated ports when its role transitions from Disable or Alternate to Designated or Root and the current state of port is forwarding.
10. A port flushes its associated cached learnt locations if its role is transitioning from Root Port to Designated Port.
11. A designated port that is receiving an DTC message flushes its associated cached learnt locations and propagates the received DTC message on the current root port of the switch.
12. A designated port that is receiving an TC message flushes cached learnt locations associated with the current root port and all other legacy active designated ports of the switch and propagates the received TC message on the current root port and all other legacy active designated ports of the switch.
13. The root port that is receiving an TC message flushes cached learnt locations associated with all legacy active designated ports of the switch and propagates the received TC message on all legacy active designated ports of the switch.

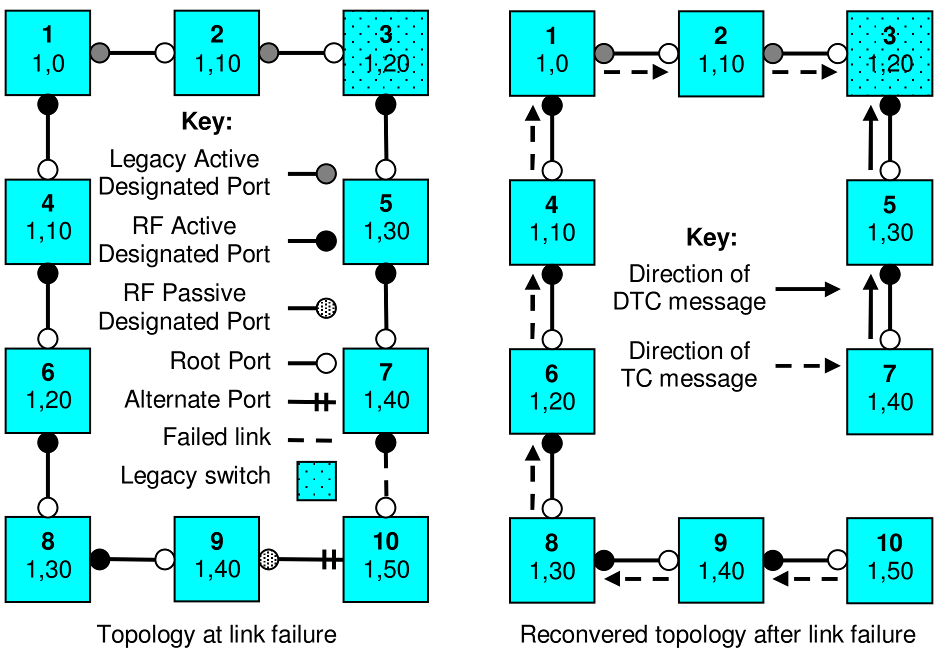


FIGURE 11: Flow of TC and DTC messages in backward compatible implementation for Ring Flushing.

Figure 11 is illustrating the flow of TC and DTC messages in backward compatible implementation of Ring Flushing. To implement the above mentioned rules, subtle change are required in Port Information state machine, Port Transmit state machine, Port Role Transitions state machine and Topology Change state machine of RSTP. A new state machine is needed to keep track that when all ports of the switch become RF port and when a port of the switch becomes legacy port. It is very important to note that the above mentioned rules will reduce to what used by RSTP switches when all the switches in the network are legacy switch and reduce to that defined in section 5 when all switches in the are “Backward Compatible Ring Flushing RSTP switches”. Moreover, there is no requirement of any extra measure for backward compatibility with legacy STP switches. Hence the proposed solution is completely backward compatible to legacy STP and RSTP switches.

7. RELATED WORKS

Although Address flushing techniques are essentially required to flush potentially stale cached learnt locations after a topology changes in a spanning tree protocol controlled switched Ethernet networks but there is very little literature available on this topic. To, the best of my knowledge, the first address flushing technique was one incorporated with STP. It is a centralized technique in which the Root Switch is responsible to instruct all switches in the network to reduce their aging time after a topology change. Later a much improved technique was proposed by Vipin Jain and Mick Seaman in [7] and it is now the standard address flushing technique for RSTP. This technique generates TC message only when a port of a switch changes its state to forwarding. Moreover, this technique instantly flushes the potentially stale cached locations instead of quick aging to improve network availability. However, the scope of effected switches after a topology change in this technique is network wide. Hence, this technique is not very much scalable. In contrast, Ring Flushing [10], an address flushing technique proposed by Horvath et al., reduces the scope effected switch to a ring in physical topology and thus not only reduces the amount flooding traffic but also enhance the network scalability. Ethernet Ring Protection (ERP) [11] is a spanning tree protocol specifically designed to improve reliability of switched Ethernet networks. Two techniques are recently proposed to specifically reduce flooding traffic in ERP. They are FDB Flip [12], proposed by Rhee et al., and selective FDB advertisement [13], proposed by Lee et al.

8. CONSLUSION

This paper critically discusses a new address flushing technique, Ring Flushing, proposed by Horvath et al. for RSTP switches. This paper mentions that although the Ring Flushing itself is a very good technique but there are three very serious flaws in the implementation of Ring Flushing proposed by Horvath et al. Due to those flaws some of the stale cached learnt locations remain persist on the switches while some fresh and valid cached learnt locations are unnecessarily flushed from the switches in the network. The paper the proposed a set of rule to successfully eliminated those mentioned flaws. In the end, it is discussed that the solution provided by the inventers of this technique for backward compatibility with RSTP switch is also not working. A new more comprehensive solution is then proposed to achieve complete backward compatibility not only with legacy RSTP switches but also with legacy STP switches. Moreover, due to the newly proposed compatibility mechanism it is now possible to use legacy multipoint link even in the network having Ring Flushing switches.

9. REFERENCES

[1] LAN/MAN Standards Committee of the IEEE Computer Society. *"IEEE Standard for Local and metropolitan area networks: Media Access Control (MAC) Bridges - 802.1D"*. 2004.

[2] LAN/MAN Standards Committee of the IEEE Computer Society. *"IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – common specifications, Part 3: Media Access Control (MAC) Bridges"* , ISO/IEC 15802-3, ANSI/IEEE Std 802.1D, 1998.

[3] R. Perlman. "*An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN"*. In the proceedings of 9th ACM Data Communications Symposium. New York, USA, 1985.

[4] M Seaman. *"High Availability Spanning Tree"*. [online] Available at: www.ieee802.org/1/files/public/docs1998/hasten7.pdf. [Accessed 21 March 2011].

[5] M. Seaman. *"Speedy Tree Protocol"*. [online] Available at: www.ieee802.org/1/files/public/docs1999/speedy_tree_protocol_10.pdf. [Accessed 21 March 2011].

[6] M. Seaman. *"Truncating Tree Timers"*. [online] Available at: www.ieee802.org/1/files/public/docs1999/truncating_tree_timing_10.pdf. [Accessed 21 March 2011].

[7] V. Jain and M. Seaman. *"Faster flushing with fewer addresses"*. [online] Available at: www.ieee802.org/1/files/public/docs1999/fast_flush_10.pdf. [Accessed 21 March 2011].

[8] M. Huynh, S. Goose and P. Mohapatra, "*Resilience technologies in Ethernet*" , Computer Networks 54(1): 57-78, 2010.

[9] Cisco Systems, Inc. *"Spanning Tree Protocol Problems and Related Design Considerations"*. [online] Available at: www.cisco.com/en/US/tech/tk389/tk621/technologies_tech_note09186a00800951ac.shtml [Accessed 21 March 2011].

[10] D. Horvath, G. Kapitany, S. Plosz, I. Moldovan and C. Lukovszki. "*Ring Flushing for Reduced Overload in Spanning Tree Protocol Controlled Ethernet Networks*". Lecture Notes in Computer Science, 5733: 11-20, 2009.

[11] ITU-T Rec. G.8032. "*Ethernet Ring Protection Switching*", ITU-T, Geneva, 2008.

[12] J.K. Rhee, J. Im and J. Ryoo. "*Ethernet Ring Protection Using Filtering Database Flip Scheme for Minimum Capacity Requirement*", ETRI Journal, 30(6):874-876, 2008.

[13] K. Lee, J. Ryoo and S. Min. "*An Ethernet Ring Protection Method to Minimize Transient Traffic by Selective FDB Advertisement*", ETRI Journal, 31(5):631-633, 2009.

A Survey on the Common Network Traffic Sources Models

Ahmed M. Mohammed *ahmed@engr.uconn.edu*
Faculty of Engineering / Electrical Department /
computer&control Section
South Vally University
Aswan, 81542, Egypt.

Adel F. Agamy *a.f.agamy@gmail.com*
Faculty of Engineering / Electrical Department /
Communications&Electronics Section
South Vally University
Aswan, 81542, Egypt.

Abstract

Selecting the appropriate traffic model can lead to successful design of computer networks. The more accurate the traffic model is the better the system quantified in terms of its performance. Successful design leads to enhancement the overall performance of the whole of network. In literature, there is innumerous traffic models proposed for understanding and analyzing the traffic characteristics of computer networks. Consequently, the study of traffic models to understand the features of the models and identify eventually the best traffic model, for a concerned environment has become a crucial and lucrative task. Good traffic modeling is also a basic requirement for accurate capacity planning. This paper provides an overview of some of the widely used network traffic models, highlighting the core features of these models and traffic characteristics. Finally we found that the N_BURST traffic model can capture the traffic characteristics of most types of computer networks.

Keywords: Traffic Modeling, Stochastic Process, Queue Theory, Chaotic Maps, Performance, Quality of Service, Poisson Process, Markova Process.

1.INTRODUCTION

Why modeling traffic? In general, traffic modeling aims to provide the computer network designer with relatively simple means to characterize traffic load on a computer network. Ideally, such means can be used to estimate performance and to enable efficient provisioning of network resources. Modeling a traffic stream emitted from source or traffic stream that represents a multiplexing of many Internet traffic streams is part of traffic modeling. It is normally reduced to finding a stochastic process that behaves like the real traffic stream from the point of view of the way it affects computer network performance. The more accurate the traffic model is the better the system quantified in terms of its performance. Traffic modeling should focus on capturing the aspects of the application which posts special demand on the system performance in the traffic model case, the long rang dependency (LRD) is the key characteristic that needs to be captured, because high burstiness resulting from LRD posts high demand on both transport and buffering capability in the system. Some network applications are real-time and some are non-real-time applications. In this paper, we are interested in real–time computer network applications. Some real time network applications are bursty and dynamically change their bandwidth demands overtime (e.g., compressed video), while others require constant bandwidth (e.g., uncompressed video). Bursty applications produce VBR (Variable Bit Rate) traffic streams, while constant applications produce CBR (Constant Bit Rate)traffic streams. In the literature there are many research in traffic model classification in [1,2] the authors divided traffic models into stationary and non stationary. Stationary traffic models can be classified into two classes: short range and long-range dependent. Traffic can be as above VBR or CBR. CBR (smooth) traffic is easy to model and predict its impact on the performance of the network. VBR (bursty) traffic models can be classify as packet level traffic models and burst (customer) level traffic model [3], where individual customers represented by individual packets rather than packet level where individual customers represent complete bursts rather than individual packets. In this paper, we classify VBR traffic models into two main categories the first category is bound (envelope) based source model, these models provide a bound or an envelope on the volume of source traffic characteristics, the bounding characterization can be deterministic or stochastic bound interval independent, bound interval dependent(BIND). The second category is unbound (exact) source model, these models characterize source behavior by describing their stochastic properties through suitable distribution functions and this category is divided into many subcategory, the last recent type of traffic model is models that use chaotic map to generate bursty traffic. We insert this type as a subcategory of the unbound category. The characteristics of chaotic map present the researchers with a method to model the non-linear nature of network traffic and make chaotic map to be a main way to generate network traffic. An overview of the progress made using chaotic maps to model individual and aggregated self-similar traffic is presented by Mondragón [6]. The organization of the rest of paper is as follows, section 2 describes traffic models and its need in design of computer network and section 3 is the conclusion.

2. TRAFFIC MODELS

The design of robust and reliable computer networks and network services is becoming increasingly difficult in today's world. The only path to achieve this goal is to develop a detailed understanding of the traffic characteristics. An accurate estimation of the computer network performance is vital. Networks, whether voice or data, are designed around many different variables. Managing the performance of computer networks involves optimizing the way networks function in an effort to maximize capacity, minimize latency and offer high reliability regardless of bandwidth available and occurrence of failures. Network performance management consists of tasks like measuring, modeling, planning and optimizing computer networks to ensure that they carry traffic with the speed, capacity and reliability that is expected by the applications using the network or required in a particular scenario. The term Quality of Service, in the field of networking, refers to control procedures that can provide a guaranteed level of performance to data flows in accordance to requests from an application/user using the network. A network that provides supports QOS usually agrees on a traffic contract with an application and reserves a finite capacity in the network nodes, based on the contract, during the session establishment phase. While the session is in progress, the computer network strives to adhere to the contract by monitoring and ensuring that the QOS guarantees are met. The reserved capacities are released subsequently after the session. There are several factors that might affect such QOS guarantees. Hence, to design a network to support QOS is not an easy task. The primary step is to once again have a clear understanding of the traffic in the network. Without a clear understanding of the traffic and the applications that might be using the network, QOS guarantees cannot be provided. Therefore, modeling of traffic becomes a crucial and necessary step. Analysis of the traffic provides information like the average load, the bandwidth requirements for different applications, and numerous other details. Traffic models enables network designers to make assumptions about the networks being designed based on past experience and also enable prediction of performance for future requirements. Traffic models are used in two fundamental ways: (1) as part of an analytical model or (2) to drive a Discrete Event Simulation (DES). Here we describe in details the most common traffic models.

2.1 BOUND (ENVELOPE) SOURCE TRAFFIC MODELS

These models provide abound or an envelope on the volume of source traffic characteristics, the bounding characterization can be deterministic or stochastic bound interval independent, bound interval dependent (BIND). As mentioned above this model can be divided into two subcategories, bound interval dependent and bound interval independent. Each subcategory can be deterministic or stochastic we describe in details here.

2.1.1 Deterministic Bound Interval Independent

Deterministic traffic modals that provide some means of bounding source peak and average bandwidth over an averaging interval such models are not only practical but they also result in an analysis that doesn't suffer from many of problems of stochastic traffic models. A traffic model for a deterministic service has several fundamental requirements. First, the model must be a worst-case characterization of the source to provide an absolute upper bound on a source's packet arrivals. Second, the model must be parameterized so that a source can efficiently specify its traffic Characterization to the network. Next, the model should characterize the traffic as accurately as possible so that the admission control algorithms do not overestimate the resources required by the connection. A worst-case representation of a traffic source may be described as follows. If the actual traffic of a connection is given by a function A such that A[τ , t+τ] denotes the traffic arrivals in the time interval [τ , t+τ], an upper bound on A can be given by a function A* if for all times τ ≥0 and all t ≥0 the following holds [38, 18].

$$A[\tau, \tau + t] \leq A^*(t). \qquad (1)$$

We refer to a function A* (t) that satisfies the property in (1) as a traffic constraint function. Note that a traffic constraint function provides a time-invariant bound on A, so that a source is bounded for every interval of length t. In practice, a source specifies its traffic characterization with a parameterized model. The parameterized deterministic traffic model defines a traffic constraint function that bounds the Source we list of most common of them here. The model should have a traffic constraint function that is as tight as possible so that the admission control algorithms do not overestimate the resources required by the connection. While, in general, a model with more parameters can achieve a more accurate traffic constraint function, the additional parameterization causes an increase in the complexity of poking the traffic model. Thus, the selection of an appropriate traffic model for a deterministic service must find a compromise between the high accuracy preferred by the admission control tests and the simplicity required for the implementation of traffic policers. The policing mechanisms must verify in real-time whether the traffic transmitted on an established connection adheres to a specified set of parameters of a deterministic traffic model. To ensure that the policing mechanisms can monitor and control traffic at high data rates, the complexity of the traffic model is limited. In [36], it was shown that a traffic model with a piece-wise linear concave traffic constraint function can be policed by a fixed number of leaky buckets. Since a leaky bucket can be implemented with a counter and a single timer [I5], concavity of the model's traffic constraint functions ensures a simple implementation of the traffic policer here we give the prosperities of the deterministic traffic model and then a short view of common deterministic traffic model.

(X_{min}, X_{av}, I, $S_{max)}$ Deterministic Traffic Model [39]

Is proposed for providing real time service over real time channel where clients declare their traffic characteristics and performance requirement at the time of channel establishment in this model. It characterizes the traffic (offered load) by the minimum packet interarrival time on the channel X_{min} the minimum value, X_{av} of the average packet interarrival time over an interval of duration I,s maximum packet size Smax and the maximum service time t in the node for the channel's packets. For the performance bounds, the source-to-destination delay bound (or bounds) for the channel's packets, and the maximum loss rate. Note that X_{av} is the average interarrival time during the channel's busiest interval of duration I. Note also that specifying X_{min} and X_{av} together with S_{max} corresponds to requesting that the network provide a certain peak bandwidth and a certain long-term average bandwidth, respectively. If the channel request is accepted, i.e., if the desired bandwidths are allocated to the channel, the client is expected to satisfy by the offered load parameters, whereas the delay bounds are to be guaranteed by the provider, i.e., by the network. A traffic constraint function

$$A^*(t) = \lfloor \tfrac{t}{I} \rfloor \cdot \frac{I \cdot s}{x_{ave}} + \min\left\{ \lceil (\tfrac{t}{I} - \lfloor \tfrac{t}{I} \rfloor) \cdot \frac{I}{x_{min}} \rceil, \frac{I}{x_{ave}} \right\} \cdot s$$

2.1.2 Stochastic Bound Traffic Models

In this model the traffic generated by a source , i ,is said to be bounded over an interval of time of length t by a discrete random variable ,R_t. If R_t is stochastically larger than a random variable ,x is said to be stochastically larger than a random variable ,y, (denoted by X ≥ st Y). If and only if prob(X>x) ≥prob(Y>x) for all x. A source can bounded by different random variable, each of which bounds the source over a different length of time as the characterization of the source. In [12], is the first bounding stochastic model. It is proposed for computing upper bounds on the distribution of individual per session performance measures such as delay and buffer occupancy for networks in which sessions may be routed over several hops. Other stochastic bounds traffic modeled are proposed in [13] [16].

2.1.3 BIND Traffic Model

In[39], a deterministic traffic model(x_{min},x_{av},I,s_{max}) is proposed ,the author in [20] proposed the extension of this deterministic model to a probabilistic model within kurose's framework[12] ,further he extended kurose's model to make bounding random variables explicit functions of the interval length (BIND)in order to better characterize the properties of the source there are two general requirements for the stochastic BIND model:

$$R_t + R_s \geq_{st} R_{t+s}$$
$$E(R_t)\backslash t \leq E(R_s)/s \quad ,\text{if } t>s$$

Where R is a random variable stochastically bounds the total number of packets that can arrive on connection during any interval and E[R] is the mean bounding rate over any interval. The first property is stochastic subadditivity. The second property requires that the mean bounding rate over smaller time intervals is greater than the mean bounding rate over large time intervals. The author in [8] gave two examples of bounding model the first with discrete random variables and the second with continuous random variable . In discrete example the author used binomial random variables to bound the number of packets that can be generated by a source in intervals of different length. by choosing different parameters for each of the family's random variables, it is possible to bound different processes with complicated distribution. For the family of binomial bounding random variables let the j source denoted by S be described by {(R_{t},j,t), t≥0} where R_t stochastically bounds the total number of packets that can arrive on connection j during any interval of length t. the binomial distribution parameters is M_t and P_t which is given by the following equations:

$$R_t = \int_0^t \left\{ r_{1,t,1}\left(\sum\nolimits_\tau = 1\right) + r_{2,t,1}\left(\sum\nolimits_\tau = 2\right) \right\} d\tau$$

$$p_t = \left\{ \begin{matrix} c\left((\lambda_{pk} - \lambda_{pk})e^{-\gamma t/I} + \lambda_{av} - \lambda_{pk}e^{-\gamma}\right) \\ \end{matrix} \right\} T \leq t \leq I ,$$

$$p_t = \left\{ \lambda_{av} / \lambda_{pk} \right\} t > I$$

Where $c = 1/(\lambda_{pk}(1-e^{-\gamma}))$ and γ ≥0 is a client specified parameter that controls how rapidly how rapidly the mean bounding rate over an interval approaches the long term average rate λ_{av} as the interval gets larger. A larger γ means that the speed with which λ_{av} is approached is faster. This is illustrated in figure 1.1 for I=133 ms .the figure show the mean bounding rate over {E(R_t)/t= M_tP_t/t} vs. interval length t for various values of γ. In the figure 1.2 we show the effect of the peak rate (burstiness) on the mean bounding rate. It shows that if two connections have the same long term average rate λ_{av} (with interval length no less than I), the mean bounding rate over any interval is greater for the connection with the higher peak rate λ_{pk} .the same property hold in the deterministic model with a fixed λ_{av} , a larger λ_{pk} means burstier traffic.

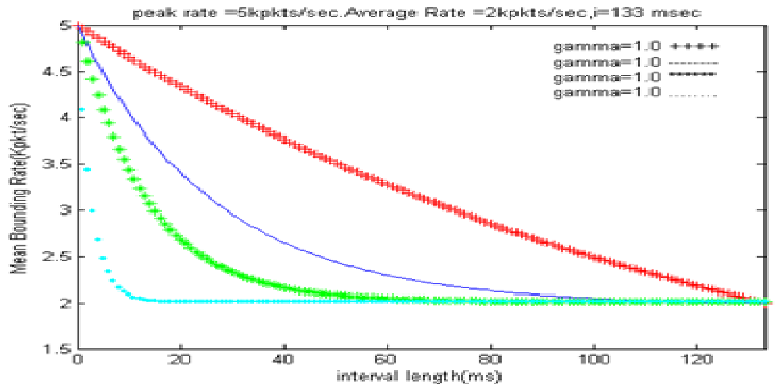


FIGURE 1.1

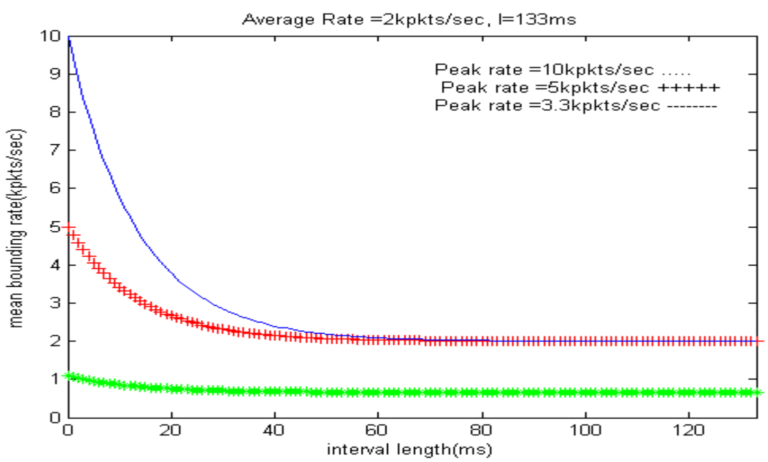


FIGURE 1.2

There are some researchers developed the BIND traffic model to overcome the multiplexing problem and low utilization. One of them extended this model with deterministic model [40] called (D-BIND) which can achieve 15-30% utilization with delay bound 30 ms. Other one introduced in [16] called (H-BIND) which achieve average utilization up to 86% , recently the author in [41] presented a(S-BIND) traffic model to on-line traffic. By using the S-BIND as input, Gamma H-BIND algorithm can achieve the maximum valid network utilization higher than the achievable network utilization under D-BIND traffic for both low-bursty and high-bursty on-line traffic, which is 50%~70 % model

2.2 Unbound (exact) Source Models

Unbound (exact) source models characterize source behavior by describing their stochastic properties through suitable distribution function and this category is divided into many subcategory, we describe each here.

2.2.1 Poisson Distribution Model

One of the most widely used and oldest traffic models is the Poisson Model. The memoryless Poisson distribution is the predominant model used for analyzing traffic in traditional telephony networks [2]. The Poisson process is characterized as a renewal process. In a Poisson process the inter-arrival times are exponentially distributed with a rate parameter µ: $p\{A_n \le t\} = 1-\exp(-\mu t)$. The Poisson distribution is appropriate if the arrivals are from a large number of independent sources, referred to as Poisson sources. The distribution has a mean and variance equal to the parameter µ. The Poisson distribution can be visualized as a limiting form of the binomial distribution, and is also used widely in queuing models. There are a number of interesting mathematical properties exhibited by Poisson processes. Primarily, superposition of independent Poisson processes results in a new Poisson process whose rate is the sum of the rates of the independent Poisson processes. Further, the independent increment property renders a Poisson process memoryless. Poisson processes are common in traffic applications scenarios that comprise of a large number of independent traffic streams. The reason behind the usage stems from Palm's Theorem which states that under suitable conditions, such large number of independent multiplexed streams approach a Poisson process as the number of processes grows, but the individual rates decrease in order to keep the aggregate rate constant. Nevertheless, it is to be noted that traffic aggregation need not always result in a Poisson process. The two primary assumptions that the Poisson model makes are:

1. The number of sources is infinite
2. The traffic arrival pattern is random.

The probability distribution function and density function of the model are given as:

$$F(t) = 1 - \exp(-\mu\, t)$$
$$f(t) = \mu \exp(-\mu\, t)$$

There are also other variations of the Poisson distributed process that are widely used. There are for example, the Homogeneous Poisson process and Non-Homogeneous Poisson process that are used to represent traffic characteristics. An interesting observation in case of Poisson models is that as the mean increases, the properties of the Poisson distribution approach those of the normal distribution.

2.2.2 Pareto Distribution Process

The Pareto distribution process produces independent and identically distributed (IID) inter-arrival times [1]. In general if X is a random variable with a Pareto distribution, then the probability that X is greater than some number x is given by

$$P(X > x) = (x / x_m)^{-k} \quad \text{For all } x \ge x_m$$

Where k is a positive parameter and x_m is the minimum possible value of X_i ,The probability distribution and the density functions are represented as

$$F(t) = 1 - (\alpha / t)^{\beta} \quad where\, \alpha, \beta \ge 0 \;\&\; t \ge \alpha$$
$$f(t) = \beta \alpha^{\beta} t^{-\beta - 1}$$

The parameters ß and α are the shape and location parameters, respectively. The Pareto distribution is applied to model self-similar arrival in packet traffic. It is also referred to as double exponential, power law distribution. Other important characteristics of the model are that the Pareto distribution has infinite variance, when ß≥ 2 and achieves infinite mean, when ß ≤ 1.

2.2.3 Markov Modulated Poisson Process

Markov models attempt to model the activities of a traffic source on a network, by a finite number of states. The accuracy of the model increases linearly with the number of states used in the model. However, the complexity of the model also increases proportionally with increasing number of states. An important aspect of the Markov model - the Markov property- states that the next (future) state depends only on the current state. In other words the probability of the next state, denoted by some random variable X_{n+1}, depends only on the current state, indicated by X_n, and not on any other state X_i, where i<n. The set of random variables referring to different states {X} is referred to as a Discrete Markov Chain. If the state transitions of the system under study happens only at integral values 0,1,2,3...n, then the Markov chain (MC) is discrete time and the random variable X follows a geometric distribution; otherwise, it is continuous time, with the random variable taking an exponential distribution. In a simple Markov traffic model, each of the state transition represents a new arrival process on the network. For modeling a continuous time system, the inter-arrival times are A Semi-Markov model is one that is obtained by allowing the time between state transitions to follow an arbitrary probability distribution. The time distribution between state transitions can also be ignored. In this model, the state transitions are then modeled as discontinuous entities with respect to time. The MC developed under such an assumption, is also referred to as an embedded or discrete Markov chain. Traffic models based on MMPP have been used to model bursty traffic. Due to its Markovian structure together with its versatility, the MMPP can capture bursty traffic statistics better than the Poisson process and still be amenable to queuing analysis. The simplest MMPP model is MMPP(2) with only four parameters: λo ,λ1 ,δo and δ1 Queuing models involving MMPP input have been analyzed in the 70s and 80s using Z-transform [26, 27, 28, 29]. Neuts developed matrix methods to analyze such queues [34]. For applications of these matrix methods for Queuing models involving MMPP and the use of MPP in traffic modeling and its related parameter fitting of MMPP the reader is referred to [30, 31, 32, 33]

2.2.4 Markov Modulated Fluid Models

Fluid flow models are conceptually simple. For instance, event simulation for an ATM multiplexer has several advantages, when fluid flow models are used for the simulation. Models other than the fluid flow models that distinguish between the cells and consider the arrival of each cell as a separate event, typically consume huge amounts of memory and CPU time for the simulation. On the contrary, a fluid flow model that characterizes the incoming cells by a finite flow rate, require comparatively less resources [1]. This is because in a fluid flow model, an event is generated only when the flow rate changes; and changes in flow rates are less frequent compared to the arrivals of cells. A fluid flow model as a consequence, utilizes lesser computing power and memory resources, compared to simulation using other models.The basic feature of a fluid model is to characterize the traffic on a network as a continuous stream of input with a finite flow/stream rate. In other words, the incoming traffic rate is represented as a stream with a finite rate. By capturing the rate changes at the input, the models analyze the different events that occur in the network. Because of the simple method of characterization of traffic, the fluid modes are analytically tractable and easier to simulate. Like any other Markov modulated process the Markov Modulated Fluid Model (MMFM), uses an underlying MC that determines the rate of the sources. At any instant, the current state of the underlying MC determines the flow rate of the inputs.

2.2.5 Autoregressive Models

The Autoregressive model is one of a group of linear prediction formulas that attempt to predict an output y_n of a system based on previous set of outputs $\{y_k\}$ where k < n and inputs x_n and $\{x_k\}$ where k < n. There exist minor changes in the way the predictions are computed based on which, Several variations of the model are developed. Basically, when the model depends only on the previous outputs of the system, it is referred to as an auto-regressive model. It is referred to as a Moving Average Model (MAM), if it depends on only the inputs to the system [1]. Finally, Autoregressive-Moving Average models are those that depend both on the inputs and the outputs, for prediction of current output.Autoregressive model of order p, denoted as AR(p), has the following form

$$Xt = R1Xt-1 + R2Xt-2 + ... + RpXt-p + Wt$$

where W_t is the white noise, R_i are real numbers and X_t are prescribed correlated random numbers. The auto-correlation function of the AR(p) process consists of damped sine waves depending on whether the roots (solutions) of the model are real or imaginary. Discrete Autoregressive Model of order p, denoted as DAR(p), generates a stationary sequence of discrete random variables with a probability distribution and with an auto-correlation structure similar to that of the Autoregressive model of order p

2.2.6 Wavelet-based Models

These models use wavelet transform function to model long-rang dependence(LRD) traffic such as traffic measured on Ethernet. Note that long-rang dependence refers to degree of dependence of samples taken at one time on those taken at an earlier time. This dependence is measured by the autocorrelation function. Long-rang dependence (LRD) traffic has its autocorrelation function slowly decrease with time .anew multi scale tool for synthesis of non Gaussian LRD traffic called multifractal wavelet model (MWM) is presented in [10]. In [10], a sequence of the hear scaling coefficients and wavelet coefficients of different scale are recursively computed. Finally synthesis traffic in the time domain is reproduced. The study shows that the correlation structure of traffic is not the only factor on queuing network, but marginal and higher order moments of traffic captured by the MWM also have a tremendous impact on the queuing traffic behavior. In [9], an estimator of the Hurst parameter H from wavelet analysis is introduced

2.2.7 Traffic Models Using Chaotic Maps

Chaotic maps are low dimensional nonlinear systems whose time evolution is described by knowledge of an initial state and a set of dynamical laws. The chaos (irregular or seemingly stochastic behavior) exhibited by such systems arises from a property known as Sensitive dependence on Initial Conditions (SIC). In [5], the author illustrates traffic characteristics that can be modeled by considering several sample maps such The Intermittency map, Piecewise Linear Maps, The Intermittency map can be heavy tailed interarrival time densities, 1/f Noise, Fractal Dimensions, here we describe Piecewise Linear Maps in details.

Piecewise Linear Maps

As the name implies, for this class of maps f(.) consists of a number of piecewise linear segments. The Bernoulli Shift is a particularly simple example, and is defined as follows:

$$Xn+1 = \begin{pmatrix} xn \setminus d \ldots\ldots\ldots\ldots 0 \prec xn \le d \\ xn - d \,/ 1 - d \ldots\ldots d \prec xn \prec 1 \end{pmatrix}$$

The associated indicator variable y_n, representing the packet generation process, is as before equal to 1 when x_n exceeds the threshold d, and is 0 otherwise. It is shown in [7] that the invariant density for this map is uniform and with every iteration, it generates a packet with probability with the arrivals forming an independent, identically distributed process. It follows that the active and passive periods are geometrically distributed.Burstier arrivals can be generated using additional piecewise linear segments. For example, it is shown in [7,8] that a three segment, two parameter map can generate a discrete analog of the Interrupted Poisson Process. More generally, it is shown in [7] that one can generate geometrically distributed dwell times in any segment by using piecewise linear segments with a uniform reinjection probability. One can then view this combination as a building block of geometrically distributed dwell times, analogous to the notion of an exponential stage in phase type processes. While it would be interesting to pursue this analogy, and investigate ways of combining piecewise linear segments, the chaotic map formulation may not offer any particular advantages in this regard. The real power of the chaotic map approach may lie in using nonlinear segments, which is illustrated next. In [9] the author use time Bernoulli Shift map to generate 100 second packet traffic by using chaotic map, Bernoulli shift map. And the characteristics of packet traffic such as packet rate, HURST exponent and Lyapunov characteristic exponent can be adjusted by the parameters of the Bernoulli shift map.

2.2.8 N- BURST TRAFFIC MODEL

The N-Burst introduced in [21][23] is a variant of the many ON/OFF models described in literature. The N-Burst arrival processes superposition of traffic streams from N independent, identical sources of ON/OFF type during its ON-time each source generates packets at rate λ_b and is quiet during its OFF time. This arrival process, with arbitrary ON and OFF time distributions (having Matrix-Exponential (ME) representations, see [19]) is analytically modeled as a Semi-Markov process of the Markov Modulated (MMPP) type. The details of this model Poisson Process can be found in [21] and [23]. When using Power-Tail Distributions for the duration of the ON periods, self-similar properties, which are critical for understanding teletraffic. Accommodating both burstiness and self-similarity in an analytic point-process model is not easy, and thus many approximations have been used by various researchers to understand buffer overflow problems and packet delay Some examples include the M/G/1 queue where the service time has infinite variance [15], continuous flow models during bursts [17, 25], and batch arrivals. The burst models are also known as ON-OFF models. For very low intra burst packet rates, the N-Burst/G/l model reduces to an M/G/l queue. For $\lambda_b \rightarrow \infty$ all packets in a burst arrive simultaneously and the model becomes a Bulk arrival, or M(X)/G/l, queue. In. the same limit, the packet-based model can be compared to a model on the burst level, an M/G/l queue where the individual customers represent complete bursts rather than individual packets. Thus the for the last mean system time describes the mean delay packet in a burst rather than the average over all packets The continuous flow model is also shown to be a limiting case of the N-Burst model by letting the number of packets in a burst n, and the router's packet service rate, v, go to infinity while holding their ratio constant. The author in [11] gave Numerical results comparing the steady-state results for Mean Packet Delay (mPD) and for Buffer Over flow Probabilities (BOP) of the different analytic models. They collectively show the critical importance of the Burstiness Parameter (the fraction of time that a source is OFF). The self similar N-burst /M/1 shows drastically changing steady-state performance for specific values of the Burstiness Parameter. The limiting models are incapable of describing the detailed structure of the performance in this transition region. In [11] the author defines the model by these parameters:

K: = the mean rate for each source (the average for ON- and OFF-times together),

λ := Overall arrival rate (packets per time unit)that generated by N-source

Where λ =KN

n_p: = Mean number of packets during a burst;

λ_p:= peak transmission rate during a burst (packets per time unit);

ON: $=n_p/\lambda_p$ =Mean ON time for a burst (time units);

OFF: = Mean OFF time between bursts (time units);

v := Mean packet service rate of router (packets per Time unit);

$\lambda_b := \lambda/n_p$= Mean burst arrival rate (bursts per time unit);

$v_b := v/n_p$ = Mean burst service rate (bursts per time unit);

$x_b := n_p/v = 1/v_b$= Mean time to service a burst (time unit);

$\rho := \lambda/v$=router utilization

In addition, the author introduce the Burstiness Parameter, b, defined as

$$b = \frac{off}{off + on} = 1 - \frac{k}{\lambda_p}$$

This parameter can be thought of as a shape parameter, since

λ, or the amount of data sent per unit time, and ρ can be held constant as b is varied over its range, [0, 1].

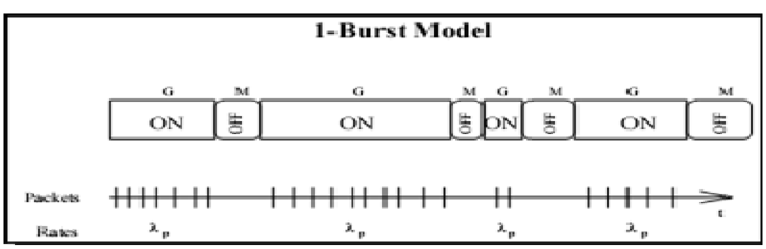


FIGURE 3: Diagram of 1-Burst traffic model

N-Burst model depends on four separate distributions, with random variables denoted by X_{SV}, X_{oF}, X_{ON} and X_{IN}, respectively. They are:

SV: Packet Service Time Distribution with mean1/v (distribution depends on packet-size distribution, service rate v depends upon router speed and size of packets)

OFF: OFF-Time distribution with mean OFF (depends on how bursts are generated, and how often);

ON: ON-Time distribution with mean ON causing a mean number of $n_p=\lambda_p$ ON a packets in a burst (e.g. . . ON-Time distribution depends on file size distribution, n_p depends on mean size of files, and on packet size);

IN: Inter-packet Time Distribution during a burst, with mean l/λ_p.

Recall that for b = 0 the SM /M /l queue reduces to the $M_\lambda/M_v/l$ queue, in which case, the mean packet delay is given by the elementary formula:

$$\text{mPD(b=0)} = \left(\frac{1/\text{v}}{1-\rho}\right)$$

Where $\rho = \lambda / \nu$ *. At the other extreme (b = 1) is the bulk arrival* $M_{\lambda b}^{[ON]}$ */ Mv /l queue. This behavior is also well known (see, [9]), and can be written as:*

$$\text{mPD(b=1)} = \left(D\frac{1/\text{v}}{1-\rho}\right)$$

Where

$$D = \left(\frac{E\left(\frac{L(L+1)}{2}\right)}{E(L)}\right)$$

3. CONSLUSION

The different traffic models each have its own advantages and disadvantages and each can be suitable for special or general type of network .consequently, the type of network under study and the traffic characteristics strictly influence the choice of the traffic model used for analysis. Traffic models that cannot capture or describe the statistical characteristics of the actual traffic on the network are to be avoided, since the choice of such models will result in under-estimation or over-estimation of network performance. There is no one single model that can be used effectively for modeling traffic in all kinds of networks. For heavy-tailed traffic, it can be shown that Poisson model under-estimates the traffic [37]. In case of high speed networks with unexpected demand on packet transfers, Pareto based traffic models are excellent candidates since the model takes into the consideration the long-term correlation in packet arrival times[1].Similarly, with Markov models, though they are mathematically tractable, they fail to fit actual traffic of high-speed networks. All this model traffic describes by a random process such Poisson process or other have the following limitation:

- model fitting some sources cannot fit the model (no analytical model then no statistical guarantee can be made)

- homogeneous traffic sources and this not suitable for various service provide by network(mean heterogeneous source)
- need per connection analysis to provide different service to different application (don't take the effect of network element)
- limiting result
- single switch analysis
- at last stochastic process traffic models is suitable to characterize the one type data or for network that support one type of traffic
- does not scale the bursty traffic properly

The bound traffic models solve a lot of problem of stochastic process traffic model. bound traffic models are a good model to achieve hard QOS in real time application on the account of resource utilization in case of internet traffic the chaotic traffic model may be a good model since it describe the traffic self similarity characteristic and chaotic characteristic in a packet level .but this model does not capture the bursty characteristic well .it shown that N-BURST traffic model is general traffic model where model traffic in a burst level rather than all model where the individual customers represent complete bursts rather than individual packets, Accommodate self similar ,bursty, and long rang dependency property for the traffic, Can describe the performance of the system under various critical points, and describe various application with it's rich parameters (λ , λ_p , v , ON , OFF ,….e.g.).as mention above we can see many traffic models as special case of N-BURST model ($b \to 0, b \to 1$).A number of factors come into play while evaluating the efficiency of a traffic model. In general, the factor that differentiates one model from the other is the ability to model various correlation patterns and marginal distributions. Traffic models should have a manageable number of parameters, and parameter estimation should be simple; and, models that are not analytically tractable are preferred only for generating traffic traces.

4. REFERENCES

[1] Abdelnaser Adas, "Traffic Models in Broadband Networks", IEEE Communications Magazine, Jul. 1997

[2] Victor S. Frost and Benjamin Melamed, "Traffic Modeling for Telecommunications Networks", IEEE Communications, Mar. 1994.

[3] LesterLipsky,Hans-PeterSchwefel,Manfred Jobmann,And Micael Greiner, "Comparison of the Analytic N-Burst Model With Some Other Approximations to Present-Day Telecommunications Traffic", in Technical Report ,Univ.of Connecticut,(2000).

[4] Erramilli and R.P. Singh, "Application of Deterministic Chaotic Maps to Characterize Broadband Traffic", 7th ITC Specialists Seminar, Morristown, 1990.

[5] Erramilli, R.P. Singh, and P. Pruthi, "Chaotic Maps as Models of Packet Traffic," The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks, Proc. of the 14th ITC, pp. 329-338, 6-10 June, 1994

[6] R. J. Mondragón, D. K. Arrowsmith, and J. M. Pitts," Chaotic maps for traffic modeling and queueing performance analysis," Performance Evaluation, vol, pp. 223-240, March 2001.

[7] A. Erramilli and R.P. Singh, "An Application of Deterministic Chaotic Maps to Model Packet Traffic", Bellcore Technical Memorandum, 1992; preprint 1993, co-author P.Pruthi.

[8] L.S. Liebovitch and T.I. Toth, ''Ion Channel Kinetics: Random or Deterministic Process?", Biophysical Journal, vol. 57, pg 317a, 1990

[9] Patrice Abry, Darry1 Veitch, "Wavelet Analysis of Long Rang Dependent Traffic " IEEE Transactions on Information Theory (1998)

[10] Vinary J. Ribeiro, Rudolf H. Riedi, Matthew S.Crouse, and Richard G. Baraniuk, “simulation of nonGausian Long-Rang-Dependent Traffic Using Wavelets”, In Proceeding SIGMETRICS ‘99(April 9,1999)

[11] LesterLipsky,Hans-Peter Schwefel,Manfred Jobmann, And Micael Greiner, “ComparisonoftheAnalyticN-Burst ModelWithSomeOtherApproximations to Present-Day Telecommunications Traffic”, in TechnicalReport,Univ.of Connecticut,(2000)

[12] J. Kurose On computing per session performance bounds in high speed multihop computer networks In ACM SigMetri92, pages128-139, Newport Rhode Island June 1992.

[13] O.Yaron ,and,M.Sidi. Calculating performance bounds in communication networks In Proceedings of IEEE IN FOCOM93, pages539-546, SanFrancisco, CA, April 1993.

[14] Dallas E. Wrege, Edward W. Knightly, Hui Zhang, and Jorg Liebeherr, " Deterministic Delay Bounds for VBR Video in Packet-Switching Networks: Fundamental Limits and Practical Trade-offs"IEEEIACM TRANSACTIONS ON NETWORKING, VOL. 4, NO. 3, JUNE 1996.

[15] Daniel Heyman,”Performance Implications of Very Large Service-Time Variances.” Proceedings of the SPIE conference on Performance and Control of Network Systems. Boston, November 1998.

[16] O .Yaron, and M .Sidi, Performance and stability of communication networks via robust exponential bounds IEEE Transaction on Networking 1(30:pages372-385, june 1993

[17] Predrag Jelenkovic and Aurel Lazar, ”Multiplexing On- Off Sources with Subexponential On Periods: Part 11.” In V. P. Wirth (eds.), ‘Teletraffic Contributions for Ramaswami and the Information Age, Vol 2A’, pp. 965-974. Elsevier Science B.V., 1997.

[18] R.Cruz, A calculus for network delay parts I and II. IEEE Transaction of Information Theory37 (1):114-141-1991.

[19] Lester Lipsky, QUEUEING THEORY: A Linear Al- gebraic Approach, MacMillan and Company, New York, 1992.

[20] Hui zhang,Edward w.knightly, providing end to end statistical performance guarntees with bounding interval dependent stochastic models In ACM SigMetri94,pages211-220,june,1994

[21] Lester Lipsky, Pierre Fiorini and Hans-Peter Schwefel, “Analytical Models of Performance in Telecommunication Systems, Based on ON-OFF Traffic Sources with Self-similar Behavior”, 7th International Conference on Telecommunication Systems Modeling and Analysis, Nashville, March 1999.

[22] Edward W. Knightly, Hui Zhang “D-BIND: an accurate traffic model for providing QoS guarantees to VBR traffic” IEEE/ACM Transactions on Networking (TON) Volume 5 , Issue 2 (April 1997) Pages: 219 - 231

[23] Hans-Peter Schwefel and Lester Lipsky, “Performance Results For Analytic Models of Traffic In Telecommunication Systems, Based on Multiple ON-OFF Sources with Self-Similar Behavior”, in P. Key and D. Smith (eds.), ‘Teletraffic Engineering in a Competitive World, Vol 3A’, pp. 55-66. Elsevier, 1999.

[24] Robert B. Cooper, Introduction to Queueing Theory, 2nd Ed., Northland, New York 1981.

[25] Mark Crovella, Virgilio Almeida, Azer Bestavros and Adriana deOliveira, “Characterizing Reference Locality in the WWW,” ACM F‘DIS, 1996

[26] U.Yechiali and P.Naor, “Queuing Problems with Heterogeneous Arrivals and Service,” Operations Research, vol. 19, pp. 722–734, 1971.

[27] U.Yechiali and P.Naor, “Queuing Problems with Heterogeneous Arrivals and Service,”Operations Research, vol. 19, pp. 722–734, 1971.

[28] M. Zukerman and I.Rubin,“Queueing performance of demand-assigned multi access communication systems under bursty traffic conditions,” Proceedings of IEEEICC’86, vol. 3, no.57.2, Toronto, Canada, June 1986,pp.1827–1832.

[29] M.Zukerman and .Rubin, “On multi channel queueing systems withfluctuating parameters,” Proceedings of IEEEINFOCOM’86, Miami, Florida, April 1986, pp. 600–608.

[30] M. F. Neuts, Structured stochastic matrices of M/G/1 type and their applications, Marcel Dekker, NY 1989.

[31] G.Latouche and V.Ramaswami, Introduction to matrix analytic methods in stochastic modeling, SIAM, 1999.

[32] H.Heffes and D.Lucantoni, “AMarkov modulated characterization of packetized voice And data traffic and related statistical multiplexer performance,” IEEE Journal on Selected Areas in Communications, vol. 4, no. 6, pp. 856–868, Sep 1986.

[33] W.Fischer and K.Meier-Hellstern, “The Markov Modulated Poisson process (MMPP)cookbook, ” Performance Evaluation, vol. 18, pp. 149–171, 1992.

[34] M.F.Neuts, Matrix-geometric solutions in stochastic models-algorithmic approach, The John Hopkin’s University Press, Baltimore (1981).

[35] P Brady.A techniques for investigating on off patterns in speech Bell System Technical Journal, 44:1-22 January1965

[36] B.Maglaris, D.Anastassiou, P .Sen, G. Karlsson, and J.D.Robbins. Performance models of statistical multiplexing in packet video communications IEEE Transaction on Communication ,36(7):834-844,july1988

[37] V. Paxson and S. Floyd, "Wide-area Traffic: The Failure of Poisson Modeling", IEEE/ACM Transactions on Networking, Jun. 1995.

[38] C. Chang, “Stability, queue length, and delay of deterministic and stochastic queueing networks”, IEEE Trans. Automat. Contr., vol. 39 no. 5, pp. 913-931, May 1994

[39] D. Ferrariand ,D. Verma “AScheme for Real Time Channel Establishment in Wide Area Networks IEEE Journal On Selected Areas in Communications, 8(3):368-379, April 1990.

[40] E. W. Knightly and H. Zhang, “D-BIND: an accurate traffic model for providing QoS guarantees to VBR traffic,” IEEE ACM Transactions on Networking, vol. 5, Issue 2, pp. 219-231, April 1997.

[41] Lie Qian , Anand Krishnamurthy , Yuke Wang , Yiyan Tang , Philippe Dauchy , and Alberto Conte “A New Traffic Model and Statistical Admission Control Algorithm for Providing QoS Guarantees to On-Line Traffic” IEEE Communications Society Globecom 2004

Tight Coupling Internetworking Between UMTS and WLAN: Challenges, Design Architectures and Simulation Analysis

Safdar Rizvi *safdarrizvi@ieee.org*
Electrical and Electronic Engineering Department
Universiti Teknologi PETRONAS
Tronoh, 31750, Malaysia

Asif Aziz *msasifbashir@hotmail.com*
Electrical and Electronic Engineering Department
Universiti Teknologi PETRONAS
Tronoh, 31750, Malaysia

N.M. Saad *naufal_saad@petronas.com.my*
Electrical and Electronic Engineering Department
Universiti Teknologi PETRONAS
Tronoh, 31750, Malaysia

Nasrullah Armi *nasrullah.armi@gmail.com*
Electrical and Electronic Engineering Department
Universiti Teknologi PETRONAS
Tronoh, 31750, Malaysia

Mohd Zuki Yusoff *mzuki_yusoff@petronas.com.my*
Electrical and Electronic Engineering Department
Universiti Teknologi PETRONAS
Tronoh, 31750, Malaysia

Abstract

To provide seamless internet connectivity anywhere at any time to the mobile users, there is a strong demand for the integration of wireless access networks for all-IP based Next Generation Networks (NGN). The Wireless Local Area Network (WLAN) is capable of providing high data rate at low cost. However, its services are limited to a small geographical area. Universal Mobile Telecommunications System (UMTS) networks provide global coverage, however, cost is high and the provided data rate do not fulfill the requirements of bandwidth intensive applications. By integrating these two promising technologies; UMTS and WLAN several benefits can be achieved, i.e., load balancing, extension of coverage area, better Quality of Service (QoS), improved security features, etc. Therefore, the integration of these two technologies can provide ubiquitous connectivity and high data rate at low cost to wireless clients. In this paper different integration mechanisms of UMTS and WLAN are investigated. More precisely, an integrated mechanism for the integration of UMTS and WLAN based on two different variations of tight coupling, i.e., interconnecting WLAN with Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN) is designed and analyzed. The simulated results reveal that the GGSN-WLAN integration performance is better than the SGSN-WLAN integration for all the applied applications and measurement parameters.

Keywords: Vertical Handoff, Heterogeneous Networks, UMTS-WLAN Integration, Tight Coupling Integration, Loose Coupling Integration.

1. INTRODUCTION

Widespread acceptance and technology advancement in internet usage and mobile communication has changed the shape of the telecommunication world. At the time when the internet was launched, internet usage was considered limited to research, scientific and academic institutions. In a similar manner, mobile communications was limited because of its high cost and low coverage area. Nowadays, the World Wide Web (WWW), file transfers, and communications through the internet are deemed necessary in daily life and not only limited to the business purposes [1]. In the mid-1980s, first commercial mobile network has been launched. Since then, the communication world has been witnessing drastic changes and significant improvements of services offered by the cellular networks. The evolution from 1G to 4G networks is clearly indicating the fulfillment of customer's demands and better enhanced services provided by the cellular networks. These facts represent a clear success story of cellular communication [2].

Network	Coverage	Data Rates	Cost
Satellite	World	Max. 144kb/s	High
GSM/GPRS	Approx. 35 km	9.6 kb/s up to 144kb/s	High
IEEE 802.16a	Approx. 30 Km	Max. 70 Mb/s	Medium
IEEE 802.20	Approx. 20 km	1-9 Mb/s	High
UMTS	20 km	Up to 2 Mb/s	High
HIPERLAN2	70 up to 300m	25 Mb/s	Low
IEEE 802.11a	50 up to 300m	54 Mb/s	Low
IEEE 802.11b	50 up to 300m	11 Mb/s	Low
Bluetooth	10m	Max. 700 kb/s	Low

TABLE 1: Diversity in existing and emerging wireless technologies [3].

At present, there are number of wireless access networks providing services to the end user. However, as illustrated in Table 1 [3], every access network has some inherent limitations and ranges in terms of data rate, coverage area, cost of services, etc. To provide services anywhere at any time, integration of wireless networks is a milestone. In the integrated unified network scenario, a user can be facilitated by the best features of each integrated wireless access network for the same data session.

The vision of 4G wireless networks is to integrate different wireless access networks to form a unified network that can provide services anywhere at any time. It is speculated that 2G, 3G, UMTS, WLAN, Worldwide Interoperability for Microwave Access (WiMAX), etc., would be operated under a unified integrated 4G network. In addition, ad hoc and wireless sensor networks would also be the part of integrated 4G networks. These heterogeneous wireless access networks will be interconnected with each other via an Internet Protocol (IP) network back bone and the internet. Such a unified network of heterogeneous wireless access networks is illustrated in Fig. 1. The main drivers of such integrated heterogeneous wireless access networks are excellent mobility supports and extensive supports of IP-based traffic [4].

Intensive research has been ongoing to integrate two most promising wireless technologies, i.e., UMTS and WLAN, to provide ubiquitous connectivity and high data rate to the wireless client. WLAN provides high data rate; however, it serves a small geographical area. IEEE 80.11b operates in license-exempt band i.e., 2.4 GHZ and provides a data rate up to 11Mbps [5]. Whereas, 802.11a and 802.11g operate at license-exempt band i.e. 5GHZ and 2.4GHZ, respectively, and provide data rate up to 54Mbps[6]. Although, WLAN coverage ranges are limited to 300 meters. On the other hand, UMTS provides global coverage; however, the data rate is limited to the maximum of 2Mpbs at high operational costs. The major limitation of UMTS networks is that they cannot serve indoor, small and densely populated areas, i.e., hotspots [7]. A WLAN network is an optimal network to provide high data rate at a very low operational cost to serve such dense and bounded regions.

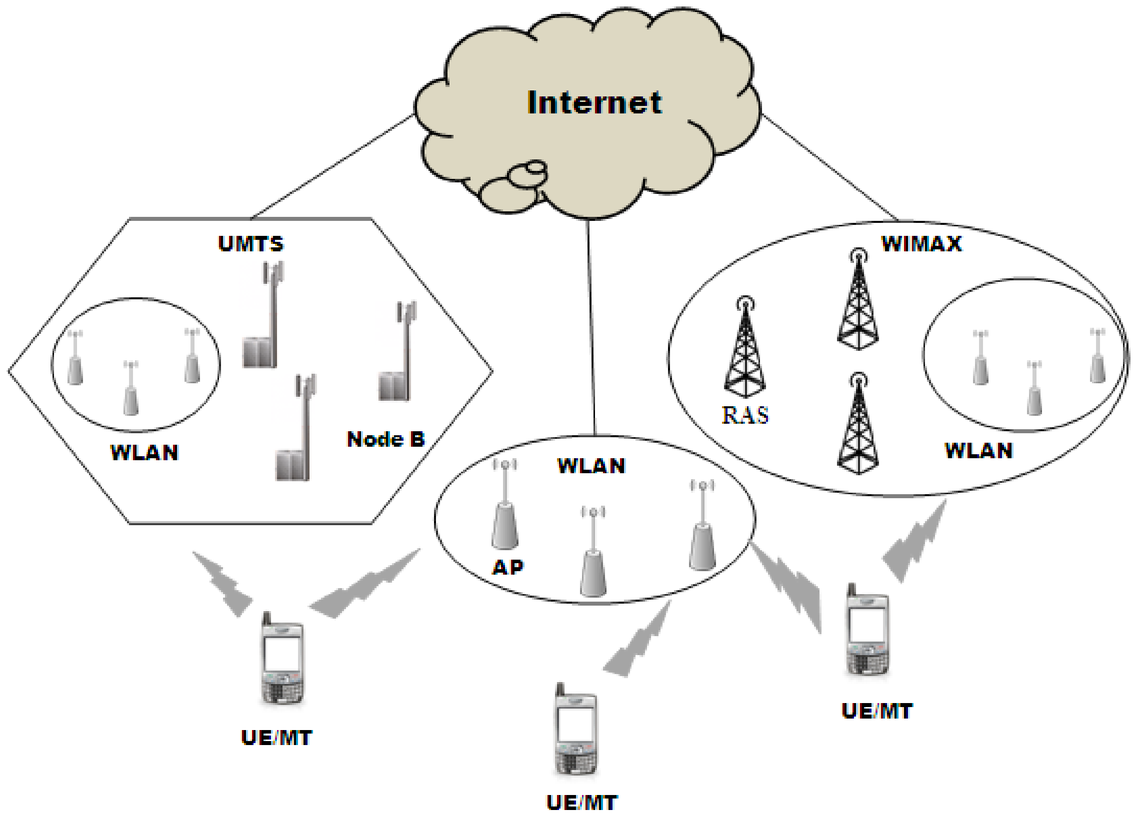


FIGURE 1: Integration of heterogeneous wireless access networks.

Therefore, the integration of UMTS and WLAN networks offers the wireless client best services experience at hotspot areas where bandwidth intensive applications are more demanding. For example, if a user is moving from the UMTS network towards a hotspot region that is being served by the WLAN network, then the same running session can be facilitated with the high data rate and low cost WLAN network. In contrast, if the user is moving away from the hotspot then the overlaid WLAN network data session can be facilitated by the UMTS. Consequently, discontinuing the running session on the Mobile Terminal (MT), due to an “out of coverage” constraint, can be eliminated by the integrated overlaid wireless network.

For the integration of UMTS/WLAN network, six scenarios have been proposed [8] [9] [10]:
I. Simplest form of integration: Common Customer Care and Billing.
II. 3G-Based Charging and Access Control.
III. Access to 3G Packet Switched (PS) services.
IV. Access to 3G PS services with the Service Continuity maintenance.
V. Access to 3G PS services with Seamless Service Continuity.
VI. Access to 3G Circuit Switched (CS) services with Seamless mobility.
These scenarios play a major role in defining the way of coupling or the integration technique. Broadly, coupling schemes are divided into three classes: open coupling, loose coupling and tight coupling.

This paper mainly focuses on the tight coupling integration scheme of UMTS and WLAN. Both variations of the tight coupling architecture, GGSN-WLAN and SGSN-WLAN, have been analyzed with respect to different services and parameter values. The remainder of the paper is organized as follows. Section 2 describes the technical challenges and issues related to the integration of the wireless heterogeneous networks. Section 3 illustrates the advantages of the 4G heterogeneous wireless environment from users’ and network operators’ point of view, and the advantages of the integration of UMTS-WLAN networks. Section 4 discusses the related works and associated contribution for the integration of wireless networks. In Section 5, we demonstrate the different approaches for the integration of wireless networks and their advantages and disadvantages. Simulated network architecture and results obtained are represented in Section 6. Finally, conclusion drawn from the paper and future work are discussed in Section 7.

2. TECHNICAL CHALLENGES

The integration of UMTS and WLAN networks brings a lot of technical challenges and problems. UMTS and WLAN are two different technologies that came into existence to serve different goals and to fulfill different service demands [11]; the implemented protocols, algorithms, data rate, authentication mechanisms, handoff mechanism, coverage ranges, etc., are dissimilar. In the same fashion, an integrated user terminal mechanism that equips the user terminal with multiple wireless access networks is required. To support seamless mobility among heterogeneous wireless networks, sophisticated mechanisms and vertical handoff decision algorithms are required.

In vertical handoffs (VHO), i.e., handoffs between radio access networks which are representing different technologies, an additional delay is occurred to disconnect from the current serving radio access network, and to connect to the target radio access network. Therefore, correct time to initiate a VHO request and selection of the best available target network among the range of available network are crucial. Moreover, minimizing the vertical handoff delay is an important factor to avoid packet loss and degradation of services during VHO.

Vertical handoff can be categorized as downward VHO (handoff towards high data rate network with smaller coverage area, from a low data rate network with larger coverage area, e.g., from UMTS to WLAN) and upward VHO (handoff towards low data rate network with larger coverage area, from a high data rate network with smaller coverage area, e.g., from WLAN to UMTS). Downward VHO are basically opportunity based VHO; i.e., User Equipment (UE) performs VHO to the target network, e.g., from UMTS to WLAN network, even when the current network is still available. Such VHO is performed to achieve high data rate and low cost services. Handoff timing is not often crucial in downward VHO. In upward VHO, UE is currently served with the high data rate and low cost network; it needs to perform VHO because it is moving away from the current network coverage area, e.g., from WLAN to UMTS network. Hence, in such type of VHO timing is crucial. An early handoff results in unnecessarily high cost and low data rate services of the target network, whereas a late VHO results in packet loss and degradation of service [12]. Downward and upward VHOs are illustrated in Fig. 2.

Similarly, a unified protocol stack is required to make networks compatible among one another. Seamless mobility, QoS, Authentication Authorization and Accounting (AAA), vertical handoffs etc., in such heterogeneous environment are still under intense research. Before answering these issues the robust integration of these promising technologies would never exist.

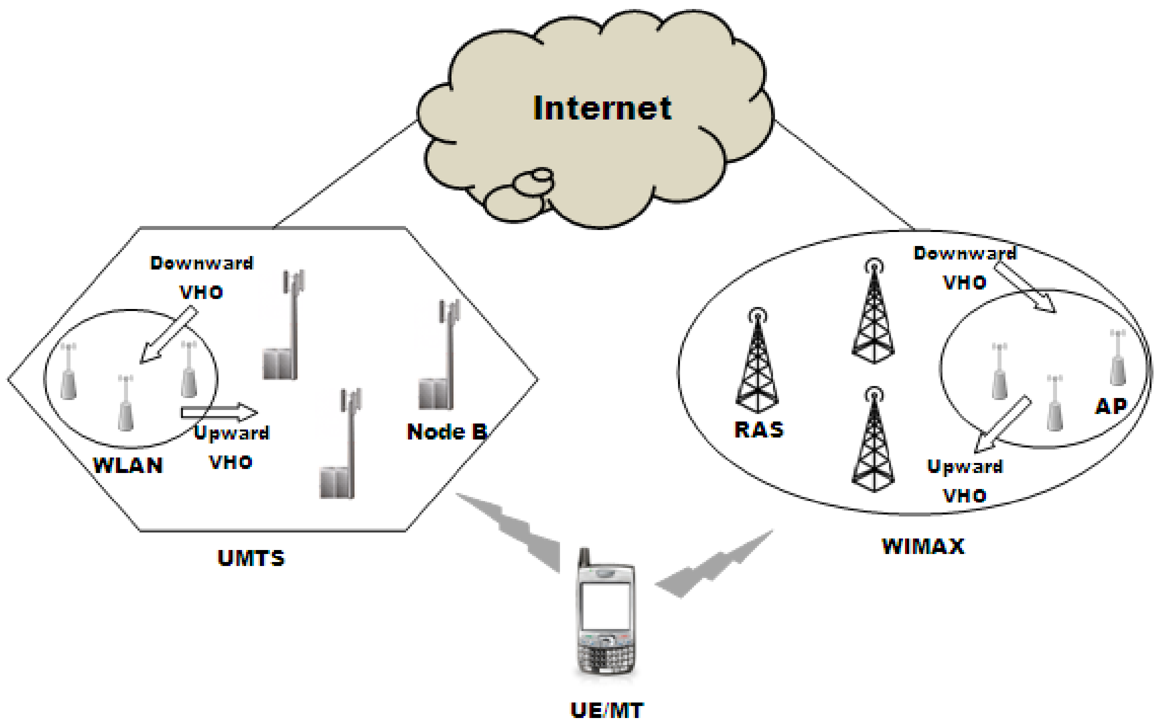


FIGURE 2: Downward and upward VHO in wireless heterogeneous networking environment.

3. ADVANTAGES OF UMTS AND WLAN INTEGRATION

The 4G networks will provide several benefits to the user and service providers [13, 14].From the users' point of view; they will have an opportunity to select a network among all the integrated networks. Therefore, the issues such as being out of coverage and limited capacity offered by some specific network will no longer exist. Moreover, 4G networks will allow them to seamlessly connect to the network that is providing a large number of available resources. Therefore, the concept of "anywhere, at any time, and always connected with the best network" can be provided to the user. From network operators' point of view, such integrated wireless heterogeneous networks will provide efficient utilization of available network resources of each wireless network. Furthermore, already deployed networks can be reused to provide anywhere at any time and always-on services, global coverage, low cost for the running session of mobile terminal and best services.

Moreover, the integration of UMTS-WLAN networks leads to achieve [7]:

- load balancing and avoidance of congestion. For example, in case of congestion in any specific network, user's data can be sent to multiple integrated wireless access networks. Therefore, by sending data to multiple networks, load balancing and avoidance of congestion can be achieved.
- extension in coverage area. In other words, cellular and WLAN coverage areas can be extended by the integrated UMTS-WLAN network. For example, a UMTS user can be facilitated by the WLAN in hotspot regions. Likewise, a WLAN user can be facilitated by the UMTS network when he/she moves away from WLAN coverage region.
- better QoS for the running application of the MT.
- improved security features, as the WLAN security features are not robust to provide the required networks security from the network attacks. Therefore, in an integrated UMTS-WLAN network, UMTS security features can be reused for the WLAN.
- interference avoidance and less power consumption, as the far user of the UMTS can use the WLAN as the relay network which consequently, improves network capacity.

4. RELATED WORKS

Intensive research has been conducted by the various parties to integrate 3G and WLAN networks. Tsao et al. [15], analyzed different UMTS and WLAN network integration approaches namely gateway, Mobile IP and emulator schemes. They concluded that Mobile IP is the easiest way to achieve the integration. Furthermore with the Mobile IP, networks can be deployed independently and standards are ready to use. However, the Mobile IP approach is not an appropriate solution for the real time services as the latency is too high during the handoffs. In the gateway approach, handoff latency is much lower compared to that in the Mobile IP approach. However, service and application mobility could not be supported. The emulator approach is the most difficult approach among the three applied approaches as networks are tightly coupled; however, it provides the lowest handoff latency among all of the analyzed approaches.

In [16], Apostolis proposed and discussed some 3G-WLAN integration architectures which will enable high throughput at hotspot locations for 3G subscribers. In [7], Fauzi and Mohammad proposed architecture for WLAN-UMTS integration and discussed a protocol to reserve resources for handoff.

In [17], F. Siddiqui et al. proposed and implemented architecture for a dual interface mobile terminal that switches its active data transmission, and evaluated handoffs performance in between UMTS and WLAN networks. In [18], Yu Zhou et al. designed a dual-mode mobile terminal module for the integrated UMTS-WLAN network and proposed a utility-based access selection algorithm for the load balancing in between them. Usman et al. [19], proposed a scheme for the authentication of a mobile node in a heterogeneous wireless environment. While switching its current running application from one network to another network when GRPS-WLAN networks are tightly coupled [19], MT shows its authentication certificate to the target network. This authentication certificate is given to the MT when it authenticates for the first time. This authentication certificate is used for re-authentication whenever MT performs vertical handoff.

M. Shi et.al [20], proposed an agent based scheme for a WLAN-cellular network. This scheme supports relevant authentication and event tracking for billing. Moreover, it does not require peer-to-peer roaming agreements between different wireless networks. In [21], the authors proposed an analytical mobility model for soft handoff regions. It is analyzed that the proposed model reduces call blocking and dropping probabilities when a user is moving in between the loosely coupled 3G-WLAN networks.

In contrast to the above mentioned research efforts, the prime concern of this paper is to study the effects of two tight coupling architecture variations on different applications and services. When WLAN is connected to the SGSN and when it is connected with the GGSN. These variations are termed as SGSN-WLAN (when WLAN is connected to SGSN) and GGSN-WLAN (when WLAN is connected to GGSN).

5. INTEGRATION SCHEMES

In the literature, several types of integration schemes have been proposed [5], [6], [22] [23], [24]. European Telecommunication Standards Institute (ETSI) has defined two generic approaches for the integration of UMTS and WLAN; namely, these are loose coupling and tight coupling. These two schemes differ in terms of the connecting point of WLAN with a UMTS network. Tight coupling indicates that the WLAN is directly connected to the UMTS core network, i.e., either to SGSN or GGSN. In such an internetworking scenario, WLAN appears as another access network of the UMTS core network. Signalling and data traffic traverse through UMTS network. On the contrary, loose coupling suggests an internetworking scenario in which WLAN and UMTS networks are deployed independently; as WLAN is connected to the internet, it maintains indirect connectivity to the UMTS network.

5.1 Loose Coupling

In the loose coupling inter-networking, networks are deployed and interconnected to each other independently. The WLAN is connected to the Internet Protocol (IP) network to maintain an indirect link with the UMTS network. From the UMTS network point of view, this inter-connecting point exists after the GGSN. For the mobility management, networks use the Mobile IP mechanism [25, 26].

As illustrated in Fig. 3, three promising wireless technologies, i.e., UMTS, WLAN and WiMAX, are integrated to form a unified network. To maintain interconnection, interconnecting devices such as WiMAX and WLAN gateways are required for roaming purposes. These gateways are used for the support of billing and authentication purposes. Mobile IP is used to support mobility among UMTS, WLAN and WiMAX networks [27]. Both gateways are connected to UMTS AAA server which in turn connects them to the internet. Hence, no direct connection between UMTS and WLAN/WiMAX exists. Subsequently, WiMAX and WLAN data traffic directly traverse to the internet instead of traversing through the UMTS core network. This approach enables a UMTS service provider to collect accounting records of WLAN and WiMAX and generate a combined billing statement for all integrated networks.

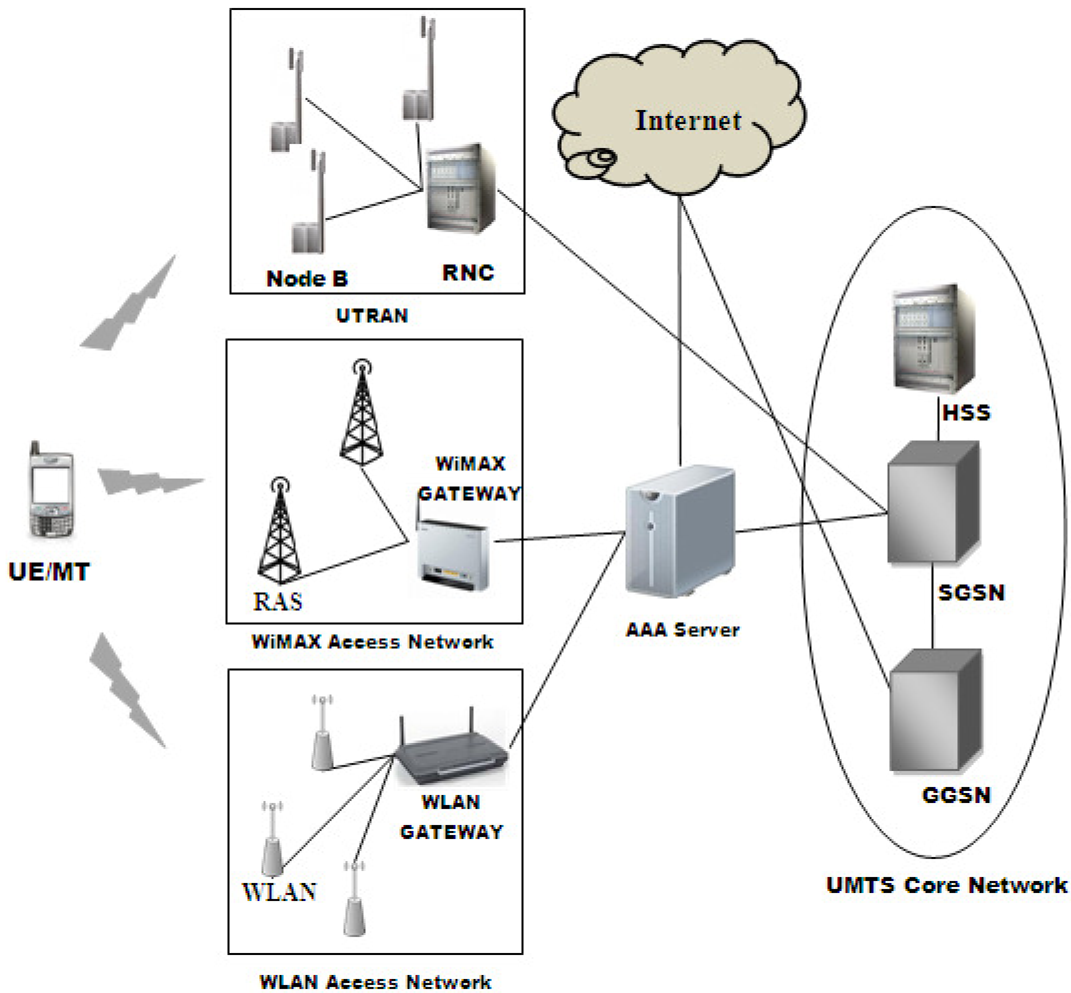


FIGURE 3: Loosely coupled integrated networks.

The main advantage of using the loose coupling scheme is that it allows the independent deployment and operation of networks. This allows the network service providers to take advantage of other provider's already deployed networks. Moreover, minimal enhancements in deployed network devices are required; hence integration does not require major investments. Subscribers can use all integrated wireless access networks while subscribing to only one service provider. For example, a UMTS subscriber can reuse its Subscriber Identity Module (SIM) or User Service Identity Module (USIM) for WLAN or WiMAX multimedia services [28].

The main disadvantage of this approach is that two networks are integrated via the internet. Therefore, signal traffic needs to traverse long paths which cause high handoff latency. As defined in [29], average handoff duration reaches 400ms; therefore, real-time services are highly affected.

5.2 Tight Coupling

In the tight coupling internetworking approach, WLAN and WiMAX are connected directly to the UMTS core network. In such an internetworking scenario, WLAN or WiMAX gateways are connected with the UMTS network in the similar manner as it is another UMTS Radio Access Network (RAN). A UMTS network deals with WLAN and WiMAX as its own RAN and finds no difference between UMTS and WLAN/WiMAX access networks, as illustrated in Fig. 4. Therefore, UMTS features such as mobility management, security, authentication, etc., can be applied to WLAN and WiMAX networks. For internetworking, each network needs to modify its services,

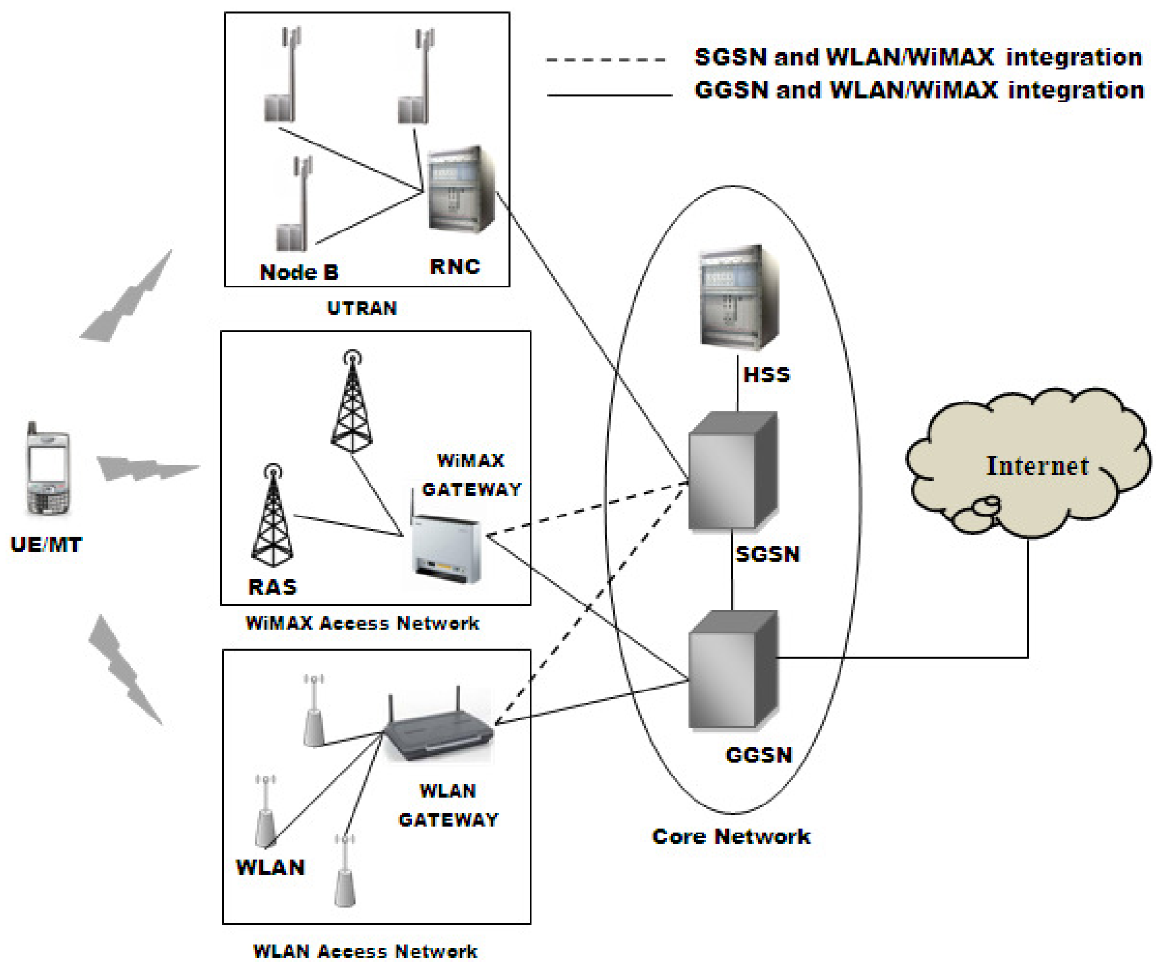


FIGURE 4: Tightly coupled integrated networks.

protocols and interfaces. WLAN or WiMAX data traffic traverse through the internet via the UMTS network. The tight coupling approach can be achieved by connecting the WLAN/WiMAX network either to UMTS' SGSN or UMTS' GGSN.

The main advantage of the tight coupling mechanism is the efficient mobility management; i.e., UMTS mobility management features are applied to the integrated networks. It offers service continuity, which includes billing and AAA services. Therefore, inter-domain seamless mobility can be achieved, while reducing handoff latency. Hence, packet loss can be reduced by minimizing the service degradation during vertical handoffs. Furthermore, UMTS core network resources, subscriber data base, billing system, authentication mechanism, etc., are reused. This reuse of resources leads to low cost deployment [25, 28].

The disadvantage of the tight coupling approach is that, as the data traffic of WLAN and WiMAX traverse via the UMTS network, they potentially create a bottle neck in the UMTS network. Moreover, this technique is considered more complicated than the loose coupling technique, because different wireless network protocol stacks need to be compatible for such type of integration. As the WLAN/WiMAX is directly connected with the UMTS network, the same operator needs to own the UMTS and WLAN/WiMAX networks. Therefore, the subscriber under the coverage of such WLAN/WiMAX networks which do not have the physical connection with the UMTS core network cannot be facilitated by the UMTS services and features.

6. SIMULATED NETWORK DESIGN

In this section, we are defining our simulation network design and parameters for the integration of UMTS and WLAN networks. Two variations of the tight coupling architecture are implemented and analyzed; in other words, one proposal is to interconnect WLAN with the GGSN of the UMTS core network, while the other mechanism is by interconnecting WLAN with the UMTS core networks' SGSN. The integrated network was designed on OPNET Modeler tool.

The network design consists of three major parts: UMTS network, WLAN network and Internet Service Provider (ISP). In this network design, only PS domain is considered and CS domain of the UMTS network is neglected for simplicity. A UMTS network is composed of Node-B, Radio Network Controller (RNC), SGSN, GGSN, UE, etc. A WLAN network is composed of WLAN AP router, which is connected either with SGSN or GGSN, and a MT. At the back of internet cloud FTP, Voice, HTTP and E-mail servers are located to provide services to the MT or UE. UE is located under the coverage of UMTS. Whereas, MT is located under the coverage of WLAN. The WLAN is located within the UMTS cell coverage area. This simulation scenario reflects a real-world scenario where WLAN is operated as a hotspot under the coverage of UMTS cell. Such hotspot serves airports, campuses, buildings, train stations, hotels, etc.

In our simulation scenario, RNC is connected with the Node-B and SGSN with the ATM OC-3 link that supports data rate up to 155.52 Mbps. The GGSN is connected to the SGSN with the PPP DS-3 bi-directional link that supports data up to 44.736 Mbps. For both GGSN-WLAN and SGSN-WLAN integration cases, WLAN AP is connected to the PPP DS-3 bi-directional link with GGSN or SGSN.

In the case of SGSN-WLAN integration [30], as illustrated in Fig. 5, WLAN does not appear as an external packet data network. Instead, it appears as another UMTS RAN. For the SGSN-WLAN integration, WLAN needs to be capable of processing UMTS messages. Therefore, some additional features in WLAN AP need to be added to make it capable of processing UMTS messages. The WLAN MT first updates its location and then establishes a packet switched signalling connection to the SGSN by using GPRS Mobility Management (GMM) attach procedure. The WLAN AP is responsible for sending these messages to the SGSN on behalf of MT. The WLAN MT is authenticated to the UMTS network after completing the GMM attach procedure.

In the case of GGSN-WLAN integration [30], as illustrated in Fig. 5, WLAN appears as an external packet data network to the UMTS network. First, UE needs to establish a Packet data protocol (PDP) context to the core network of UMTS to make itself known to the WLAN. If UE wants to communicate with the WLAN MT, then the user data are transferred between them using the encapsulation and tunneling techniques. GPRS Tunneling Protocol (GTP) is used for such encapsulation and tunneling mechanisms to send data transparently. For GGSN-WLAN integration, the WLAN does not need to process UMTS messages. Thus a simple WLAN AP is used for such type of integration.

6.1 MT Communication With ISP Servers

As illustrated in Fig. 5, WLAN access point is connected either to the SGSN or GGSN of the UMTS core network. Different types of services were used for the simulation. This includes Voice over IP (VoIP), FTP, E-mail and HTTP (web browsing). In this simulation scenario, MT communicates with the internet FTP, HTTP, Voice and E-mail servers. Table 2 represents application and measurement parameters tested for the integrated UMTS/WLAN networks.

These applications match different QoS classes of UMTS; i.e., a conversational class represents real time traffic flows such as VoIP. More precisely, PCM and GSM encoded voices have been analyzed in this paper, interactive class represents Web browsing (HTTP) and background class represents the both i.e. FTP and E-mail services.

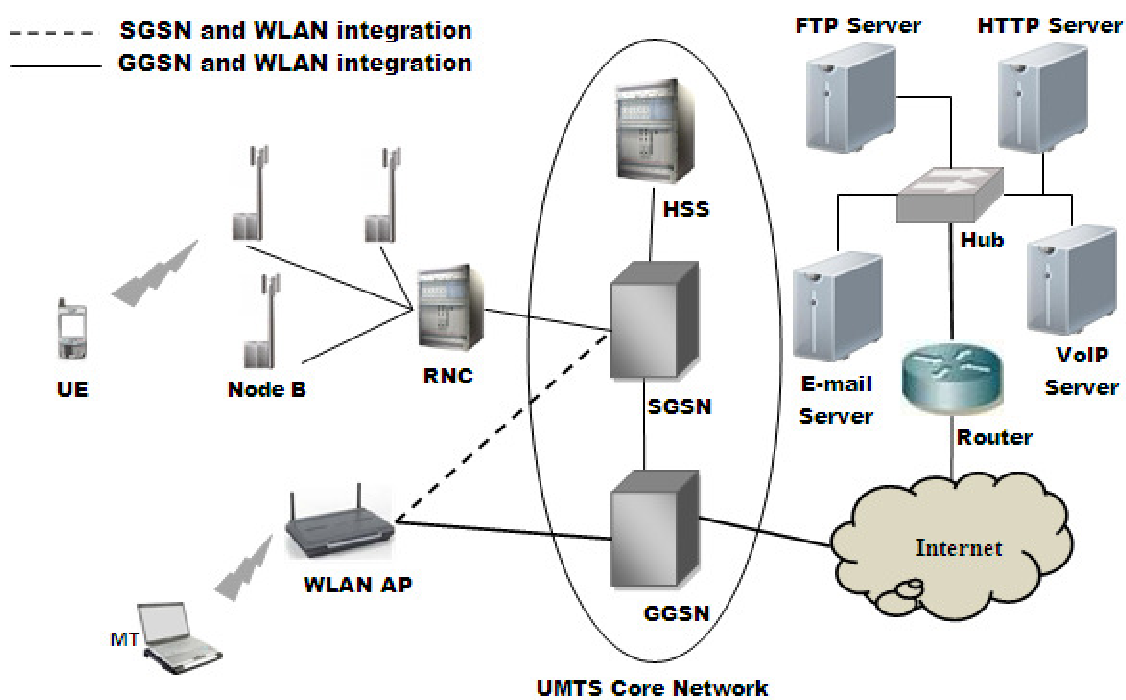


FIGURE 5: MT communicating with ISP servers in GGSN/SGSN integrated simulation scenario.

Application	QoS Class	Measurement Parameter	Size	Protocol	Figure
PCM encoded voice	Conversational	Packet Delay Variation	80 Bytes	UDP	Fig. 6
GSM-FR encoded voice	Conversational	Packet Delay Variation	33 Bytes	UDP	Fig. 7(a)
GSM-FR encoded voice	Conversational	Jitter	33 Bytes	UDP	Fig. 7(b)
E-mail	Background	Download Response time	1000 Bytes	TCP	Fig. 8(a)
E-mail	Background	Upload Response time	2000 Bytes	TCP	Fig. 8(b)
HTTP	Interactive	Downloaded Pages	500 Bytes	TCP	Fig. 9(a)
HTTP	Interactive	Object Response Time	500 Bytes	TCP	Fig. 9(b)
HTTP	Interactive	Page Response Time	500 Bytes	TCP	Fig. 9(c)
FTP	Background	Download Response Time	8000 Bytes	TCP	Fig. 10(a)
FTP	Background	Upload Response Time	8000 Bytes	TCP	Fig. 10(b)

TABLE 2: Description of the application and measurement parameters tested.

The PCM encoded voice is considered for VoIP service evaluation. Fig. 6 illustrates the simulation run times with the corresponding packet delay variation, when the user is communicating with the voice server. The voice server is located behind the internet. Throughout the simulation run time, it can be observed that the packet delay variation in the case of SGSN-WLAN integration is higher than those in the GGSN-WLAN integration.

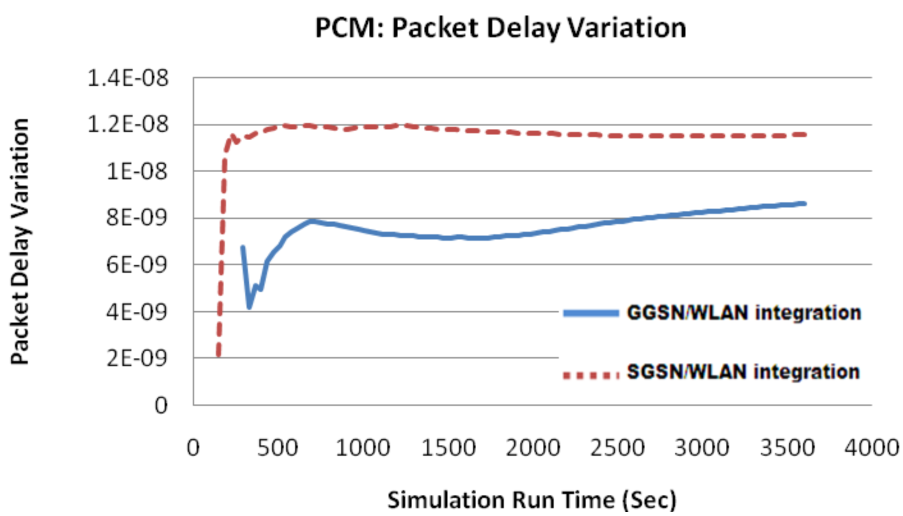


FIGURE 6: PCM packet delay variation.

Similarly for VoIP services, we consider the GSM-FR encoded voice. Fig. 7(a) and Fig. 7(b) illustrate the simulation run times with the corresponding packet delay variation and jitter, respectively. Throughout the simulation run time, it can be observed that the packet delay variation and jitter in the case of SGSN-WLAN integration is higher than those in the GGSN-WLAN integration.

Fig. 8(a) and Fig. 8(b) illustrate the simulation run times with the corresponding download and upload e-mail response time, respectively. For the diversity purpose, 1000 bytes and 2000 bytes e-mail size are used for downloading and uploading cases, respectively. It is observed that for both cases i.e. download and upload email response time, GGSN-WLAN and SGSN-WLAN integration results are initially high. However, as the simulation progresses, rapid responses can be observed. It is speculated that initially it take a longer time for the networks to negotiate with the server, and after the negotiation with the servers, email retrieval time becomes lower. The GGSN-WLAN integration represents more rapid responses in both download and upload cases compared to the responses in the SGSN-WLAN integration.

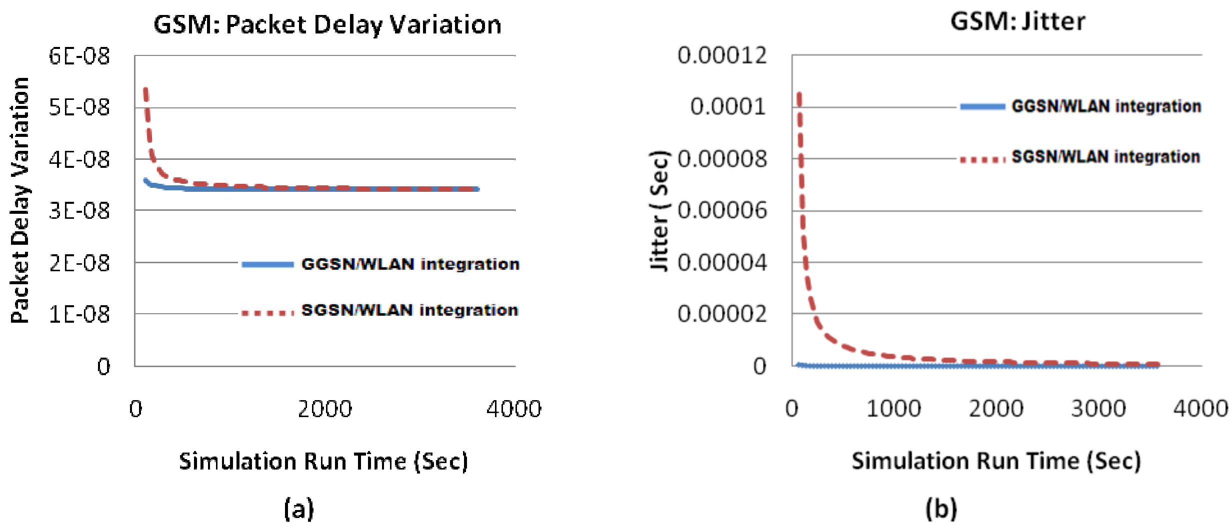


FIGURE 7: GSM encoded voice: (a) Packet delay variation, (b) Jitter.

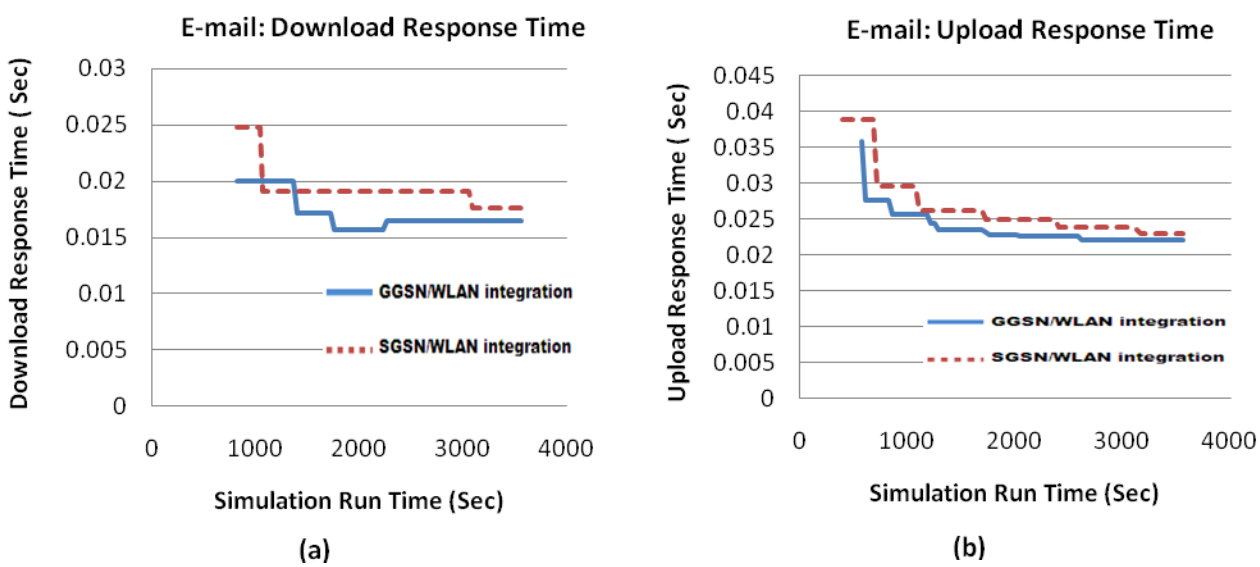


FIGURE 8: E-mail: (a) Download response time, (b) Upload response time.

For HTTP, throughout the simulation run time the GGSN-WLAN integration shows better performance than the SGSN-WLAN integration in terms of number of downloaded pages, object response time and page response time, as illustrated in Figs. 9 (a), 9 (b) and 9 (c), respectively. The page and object sizes are 500 bytes.

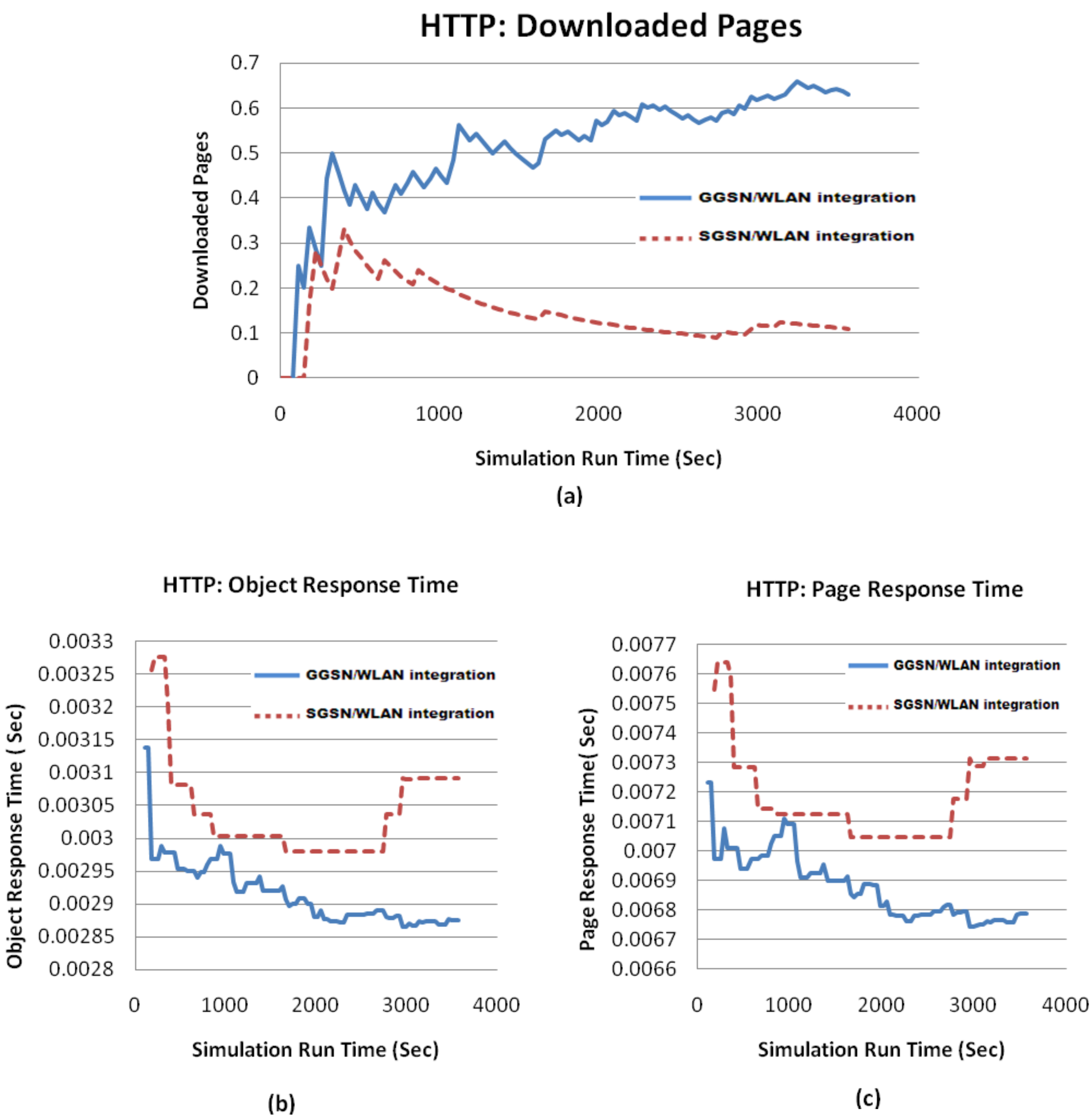


FIGURE 9: HTTP (a) Downloaded pages, (b) Object response time, (c) Page response time.

For FTP, Figs. 10(a) and 10 (b) illustrate the simulation run times the with the corresponding download and upload response time, respectively, when downloading an 8000 byte file by the MT from the FTP server located behind the internet. It can be observed that the overall performance of the GGSN-WLAN integration is better than that in the SGSN-WLAN integration, throughout the simulation run time for both cases.

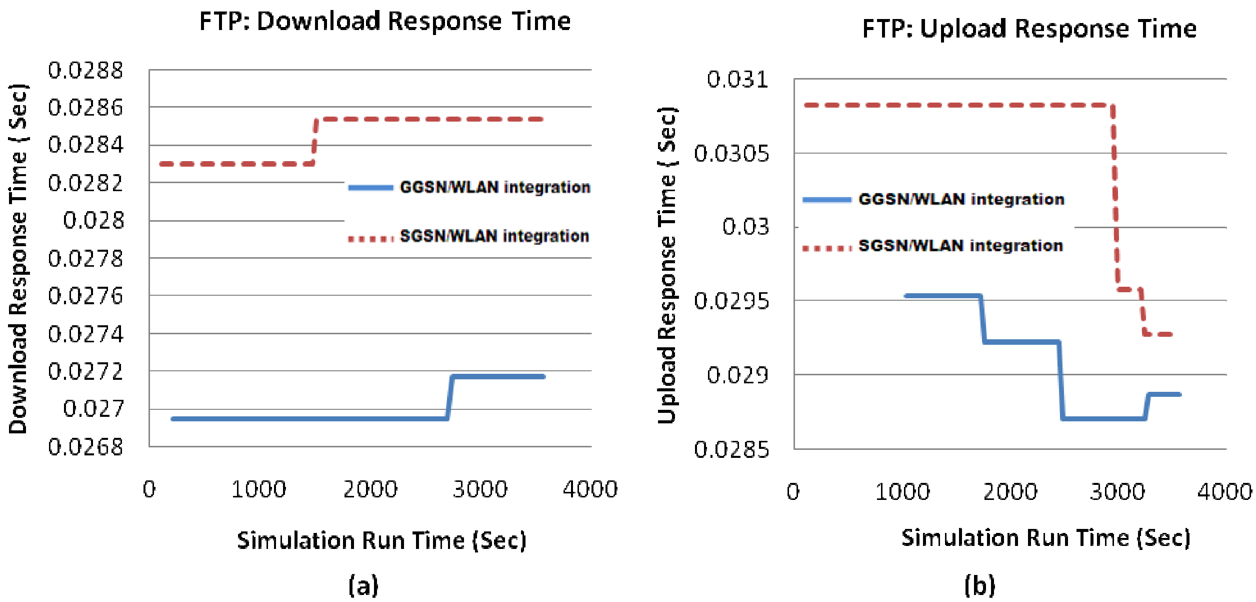


FIGURE 10: FTP: (a) Download response time, (b) Upload response time.

6.2 MT Communication With UE

As illustrated in Fig. 11, WLAN access point is either connected with the SGSN or GGSN of the UMTS core network. Different types of services were used for the simulation; these are summarized in Table 3. In this simulation scenario, MT is communicating with the UE. Therefore, it is dissimilar the simulation scenario discussed previously, in which MT is communicating with the ISP servers.

Application	QoS Class	Measurement Parameter	Size	Protocol	Figure
PCM encoded voice	Conversational	Packet Delay Variation	80 Bytes	UDP	Fig. 12 (a)
GSM-FR encoded voice	Conversational	Packet Delay Variation	33 Bytes	UDP	Fig. 12 (b)
HTTP	Interactive	Page Response Time	2000 Bytes	TCP	Fig. 13 (a)
HTTP	Interactive	Object Response Time	2000 Bytes	TCP	Fig. 13 (b)

TABLE 3: Description of the application and measurement parameters tested.

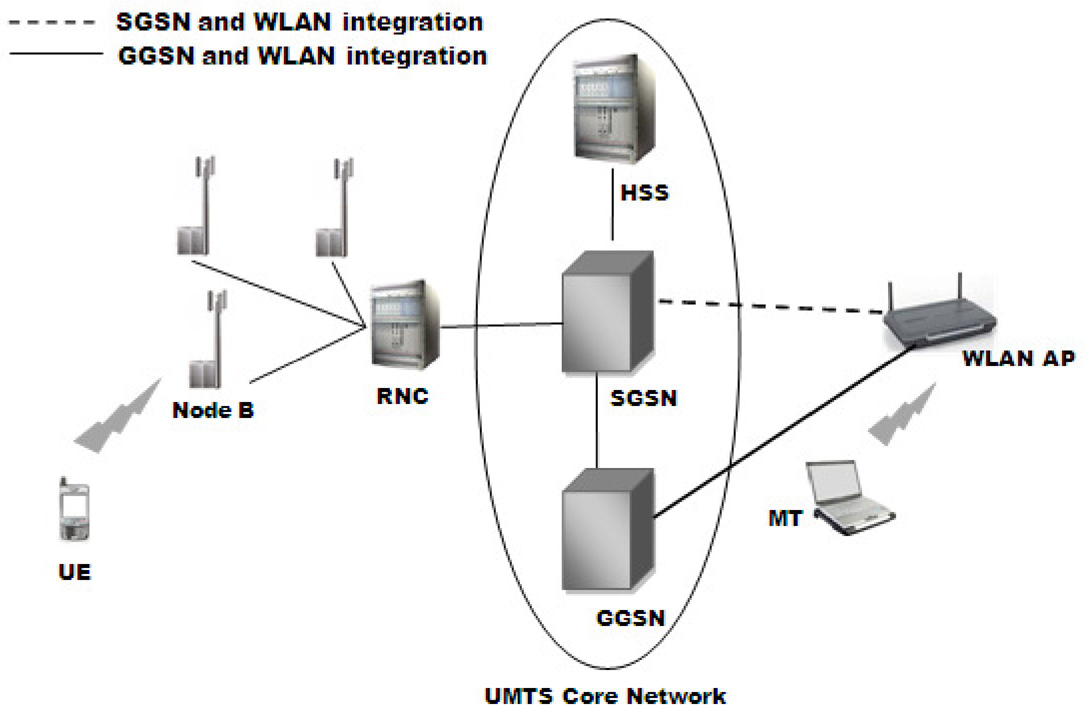


FIGURE 11: MT communicating with UE in GGSN/SGSN integrated simulation scenario.

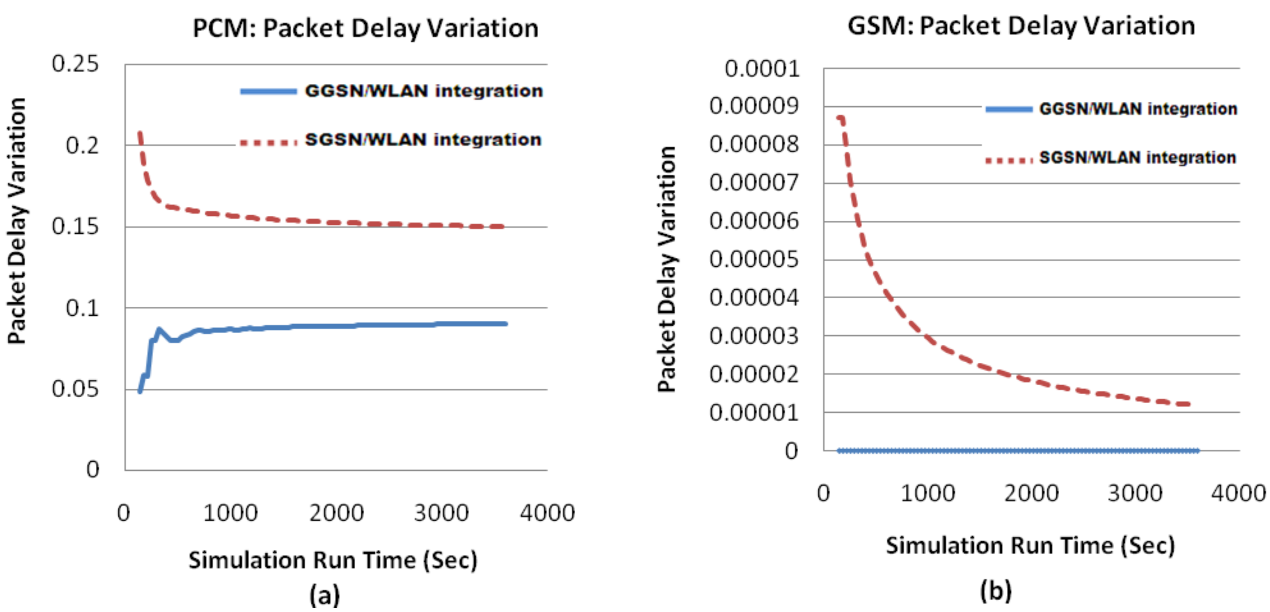


FIGURE 12: VoIP application: (a) PCM packet delay variation, (b) GSM packet delay variation.

For VoIP services, PCM encoded voice and GSM-FR encoded voice were considered, as illustrated in Figs. 12 (a) and 12 (b), respectively; the two figures represent the simulation run times with their corresponding packet delay variations. It can be observed that the packet delay variation in the case of SGSN-WLAN integration is higher than that in the GGSN-WLAN integration, throughout the simulation run time.

For HTTP, throughout the simulation run time, the GGSN-WLAN integration shows better performance in terms of page response time and object response time, compared to the SGSN-

WLAN integration, as illustrated in Figs. 13 (a) and 13 (b), respectively. The page and object sizes are 2000 bytes.

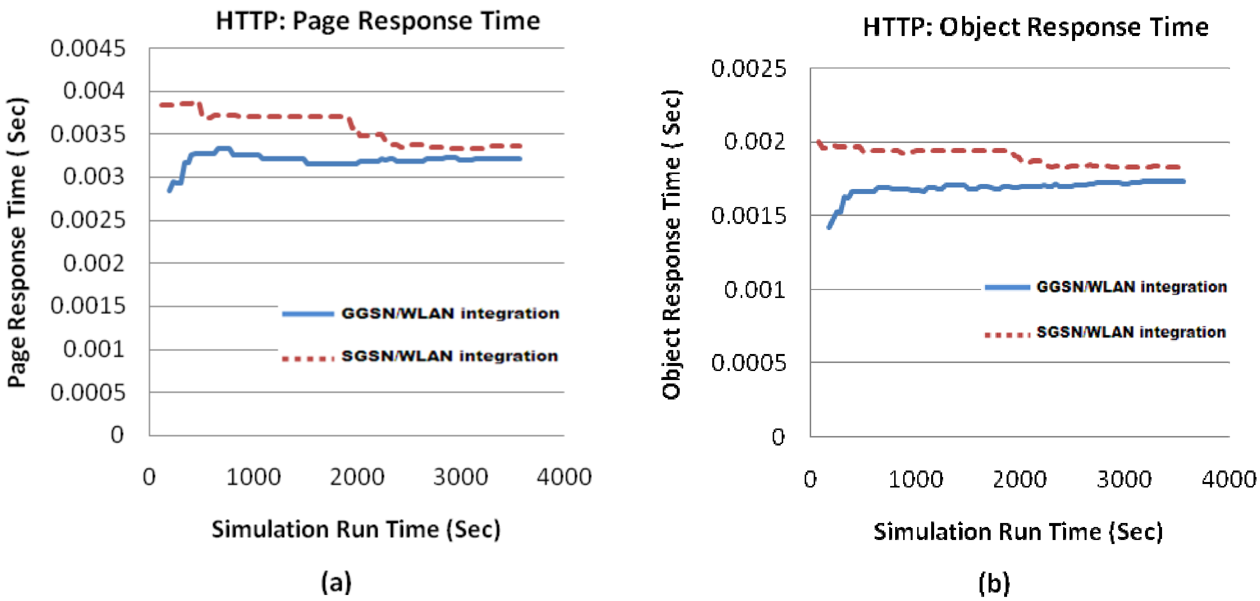


FIGURE 13: HTTP: (a) Page response time, (b) Object response time.

6.3 Discussion on Simulation Results

From the simulation results, it is apparent that the integration of the WLAN to the GGSN is of better-quality compared to the integration of WLAN to the SGSN, for all the applied applications. In the case of VoIP using either GSM or PCM encoded voice; the packet delay variation and jitter are higher in SGSN-WLAN integration compared to those in GGSN-WLAN. Similarly, in the case of E-mail, HTTP and FTP, for all applied measurement parameters, the performance of UMTS and WLAN integration is far better in the GGSN-WLAN integration than in the SGSN-WLAN integration.

In the case of SGSN-WLAN integration, this difference in the performance is due to the additional processing time required by the WLAN access point, to process the UMTS messages and establish compatibility with the UMTS network. The WLAN AP is performing dual tasks, i.e., at one time operating as a RAN to manage the compatibility to the UMTS network, and on the same communication session, operating as an AP for the MT. Therefore, it requires longer time to establish the session and has to process more messages for communication. For the authentication of MT to the UMTS network, it has to perform a GMM attach procedure with the UMTS network. On the other hand, for the GGSN-WLAN integration, WLAN AP is a simple IEEE 802.11b access point which requires no additional capabilities for processing the UMTS messages. The straightforwardness of the GGSN-WLAN integration, leads to the lower latency and low processing requirement for the communication, as no additional task needs to be performed.

Consequently, in the case of GGSN-WLAN integration, improvement in the performance for all applied applications and their parameters are achieved. For VoIP using either GSM or PCM encoded voice; improvement in the packet delay variation and jitter is achieved. For E-mail, download and upload response time is reduced. For HTTP, number of downloaded pages is noticeably high. Whereas, object and page response time is significantly reduced. For FTP, download and upload response time is considerably low etc. Therefore, the simplicity of communication without extra processing and additional tasks leads to the performance enhancements.

7. CONCLUSION AND FUTURE WORK

In this paper, we have reviewed several internetworking techniques for the integration of UMTS and WLAN networks. The inherent differences between WLAN and UMTS bring a lot of technical challenges that need to be resolved for the integrated Next Generation Wireless Networks. These differences are in terms of protocols, algorithms, data rate, authentication mechanism, handoff mechanism, coverage ranges, etc. To achieve an integrated heterogeneous wireless network, several techniques have been proposed in the recent literature. Moreover, we have comprehensively investigated two different techniques of tight coupling schemes for the integration of the UMTS and WLAN, namely GGSN-WLAN and SGSN-WLAN.

The network model has been designed on OPNET modeler tool. For simplicity, only the PS-domain of the UMTS core network is considered in the designed model, while neglecting the CS-domain. Comprehensive results have been obtained by designing two different simulation scenarios, i.e., MT is communicating with the ISP server and MT is communicating with the UE. For both cases, our results demonstrate that the performance of the integrated UMTS and WLAN is far better in the case of GGSN-WLAN than that of SGSN-WLAN internetworking, for all of the applied applications and measurement parameters. This is because the WLAN AP needs to have some additional capabilities to process UMTS messages; the SGSN-WLAN integration requires more processing and latency for the communication. Whereas for GGSN-WLAN, a simple IEEE 802.11b WLAN AP is required; therefore, GGSN-WLAN requires no additional tasks for communication. Our future research focuses on the evaluation and optimizations of the vertical handoff decision algorithms and the maintenance of a seamless mobility, when the user is moving across the heterogeneous wireless networks.

8. REFERENCES

[1] T.L. Mann. “*A Network System Level Simulator for Investigating the Interworking of Wireless LAN and 3G Mobile Systems*”. Master’s Thesis, Virginia Polytechnic Institute and State University, 17 April, 2003

[2] S. Kasera and N. Narang. ”*3G networks Architecture, Protocols and procedures*”, Tata McGraw-Hill Publishing Company Limited, pp.3 (2008)

[3] P. Vidales, G. Mapp, F. Stanjano, J. Crowcroft, and C. J. Bernardos. “*A Practical approach for 4G systems: Deployment of overlay networks*”. IEEE proceedings of the First International Conference on Testbeds and Research Infrastructures for the DEvelopment of NeTworks and COMmunities (TRIDENTCOM’05). 2005

[4] D. O’Mahony and L. Doyle. “*Beyond 3G: 4th Generation IP-based mobile networks*”. In Wireless IP and Building the Mobile Intemet. Artech House, Norwood, 2002

[5] W. Song, H. Jiang, W. Zhuang, X. Shen. "*Resource management for QoS support in cellular/WLAN interworking*". *IEEE Network*, 19(5): 12-18, 2005

[6] W. Song, W. Zhuang, A. Saleh. “*Interworking of 3G cellular networks and wireless LANs*”. International Journal of Wireless and Mobile Computing, 2(4): 237-247, 2007

[7] F. Zarai, N. Boudriga, M.S. Obaidat. “*WLAN-UMTS Integration: Architecture, Seamless Handoff, and Simulation Analysis*”**.** SIMULATION, 82(6): 413-424, 2006

[8] 3GPP TR 22.934 6.2.0. Feasibility study on 3gpp system to wlan Interworking (release 6). Technical specification, 3rd Genration Partnership Project, September 2002

[9] A.K. Salkintzis. "*Interworking techniques and architectures for WLAN/3G integration toward 4G mobile data networks*". IEEE Wireless Communications, 11(3): 50-61, 2004

[10] I. Maarouf. "*A Comparative Analysis of Different Integration Approaches between UMTS and WLAN Networks*", http://www.kfupm.edu.sa/ocwtest/user071\COE54001/FINAL_PAPER_COE587.pdf

[11] M.Y. Gerla, R. Kapoor, T. Kwon, and A. Zanella. "*UMTS TDD: Asolution for internetworking Bluetooth piconets in indoor environments*". Proceedings of IEEE Symposium on Computers and Communications (ISCC'02). Taormina, Italy, 2002

[12] H. Petander, E. Perera, A. Seneviratne. "*Multicasting with Selective Delivery: A SafetyNet for Vertical Handoffs*". Wireless Personal Communications, 43: 945- 958, 2007

[13] A. H. Zahran, B. Liang, A. Saleh. "*Signal Threshold Adaptation for Vertical Handoff in Heterogeneous Wireless Networks*". Mobile Networks and Applications, 11: 625-640, 2006

[14] A. Hecker, H. Labiod, G. Pujolle, H. Afifi, A. Serhrouchni, P. Urien. "*A New Access Control Solution for a Multi-Provider Wireless Environment*". Telecommunication Systems, 29(2): 131-152, 2005

[15] S.Tsao and C.C. Lin. "*Design and evaluation of UMTS-WLAN interworking strategies*". Proceedings of Vehicular Technology Conference. 2002

[16] A.K. Salkintzis. "*WLAN/3G interworking architectures for next generation hybrid data networks*". IEEE International Conference on Communications. 2004

[17] F. Siddiqui, S. Zeadally, and S. Fowler. "*A Novel Architecture for Roaming between 3G and Wireless LANs*".1st International Conference on Multimedia Services Access Networks, MSAN '05. 2005

[18] Y. Zhou, Y. Rong, H.A. Choi, J.H. Kim, J. K. Sohn, and H. In-Choi. "A *Dual-mode Mobile Station Module for WLAN/UMTS Interworking Systems*", CiteSeerX, 2009

[19] M. Usman, W. Noshairwan, M. Gilani, A. Irshad and E. Irshad."*Seamless vertical handoff using Authentication Certificate in GPRS-WLAN tightly coupled integrated networks*". 4th International Conference on emerging Technologies, ICET. 2008

[20] S. Minghui, H. Rutagemwa, X.Shen, J.W. Mark, A. Saleh. "*A Service-Agent-Based Roaming Architecture for WLAN/Cellular Integrated Networks*". IEEE Transactions on Vehicular Technology, 56(5): 3168-3181, 2007

[21] R. B. Ali, S. Pierre. "*On the impact of soft vertical handoff on optimal voice admission control in PCF-based WLANs loosely coupled to 3G networks*". IEEE Transactions on Wireless Communications, 8(3):1356-1365, 2009

[22] P. Vidales, G. Mapp, F. Stajano, J. Crowcroft and C.J Bernardos. "*A practical approach for 4G systems: deployment of overlay networks*". First International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, Tridentcom.2005

[23] M. A Amara et.al. "*Performance of UMTS/WLAN Integration at Hot- Spot Locations Using OPNET*". Department of Computer Engineering, King Fahd University of Petroleum and Minerals.2008

[24] R. Samarasinghe, V. Friderikos, and A.H. Aghvami. "*Analysis of Intersystem Handover: UMTS FDD & WLAN*". Centre for Telecommunications Research, King's College London. http://www.ee.ucl.ac.uk/lcs/previous/LCS2003/61.pdf

[25] L. Lampropoulos, N. Passas, A. Kaloxylos, L. Merakos. "*Handover Management Architectures in Integrated WLAN/Cellular Networks*". IEEE Communications Surveys and Tutorials, 7: 30-47, 2005

[26] L. Lampropoulos, N. Passas, A. Kaloxylos, L. Merakos. "*A Flexible UMTS/WLAN Architecture for Improved Network Performance*". Wireless Personal Communications, 43: 889-906, 2007

[27] S. Park, J. Yu, and J. T. Ihm. "*A Performance Evaluation of Vertical Handoff Scheme between MobileWiMax and Cellular Networks*". In Proc. of 16th International Conference on Computer Communications and Networks. 2007

[28] A. K. Salkintzis, C. Fors, R. Pazhyannur. "*WLAN-GPRS Integration for NG Mobile Data Netwoks*", IEEE Wireless Communications, 9: 112-124, 2002

[29] J. Y. Song, H. J. Lee, S. H. Lee, S. W. Lee, D. H. Cho. "*Hybrid Coupling Schemes for UMTS and Wireless LAN Interworking*". AEU International Journal on Electronics and Communications, 61: 329-336, 2007

[30] F. Siddiqui, S. Zeadally, E. Yaprak, "*Design and architectures for 3G and IEEE 802.11 WLAN Integration*", ICN 2005, LNCS 3421, 1047–1054, 2005

Wideband Sensing for Cognitive Radio Systems in Heterogeneous Next Generation Networks

Ahmed Barnawi *ambarnawi@kau.edu.sa*
Faculty of Computing and IT
King Abdulaziz University
Jeddah, PO Box 80221, Saudi Arabia

Abstract

Mobile Next Generation Network (MNGN) is characterized as heterogeneous network where varieties of access technologies are meant to coexist. Decisions on choosing an air interface that meets a particular need at a particular time will be shifted from the network's side to (a more intelligent) user's side. On top of that network operators and regularities have come to the realization that assigned spectrum bands are not utilized as they should be. Cognitive radio stands out as a candidate technology to address many emerging issues in MNGN such as capacity, quality of service and spectral efficiency. As a transmission strategy, cognitive radio systems depend greatly on sensing the radio environment. In this paper, we present a novel approach for interference characterization in cognitive radio networks based on wideband chirp signal. The results presented show that improved sensing accuracy is maintained at tolerable system complexity. Remittance

Keywords: Cognitive Radio, Spectrum Sensing, MNGN, Interference Characterization.

1. INTRODUCTION

Mobile Next Generation Network is perceived as a collection of unconventional concepts for network management rather than a universally standardized technology. It is widely accepted by all players that no single mobile wireless technology will prevail on the expense of the others in the foreseen future. Thus the novelty in future networks is in facilitating the coexistence of different technologies not in waging competition between them.

For many years to come we will have 2nd, 3rd and 4th mobile generations along with WIMAX, WIFI and Bluetooth in operation. The heterogeneous nature of next generation networks have pushed the technology frontiers toward decentralization

For network operators, the distributed and localized decision making which will feature in next generation networks will bring dramatic reduction in operational cost (OPEX) and improve the performance [1]. A key technology to bring this into realization is cognitive radio systems.

Cognitive Radio (CR) is defined as a system that senses its operational electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify system operation, such as maximize throughput, mitigate interference, facilitate interoperability, access secondary markets [2]. Hence, one main aspect of cognitive radio is related to autonomously exploiting locally unused spectrum to provide new paths to spectrum access [3].

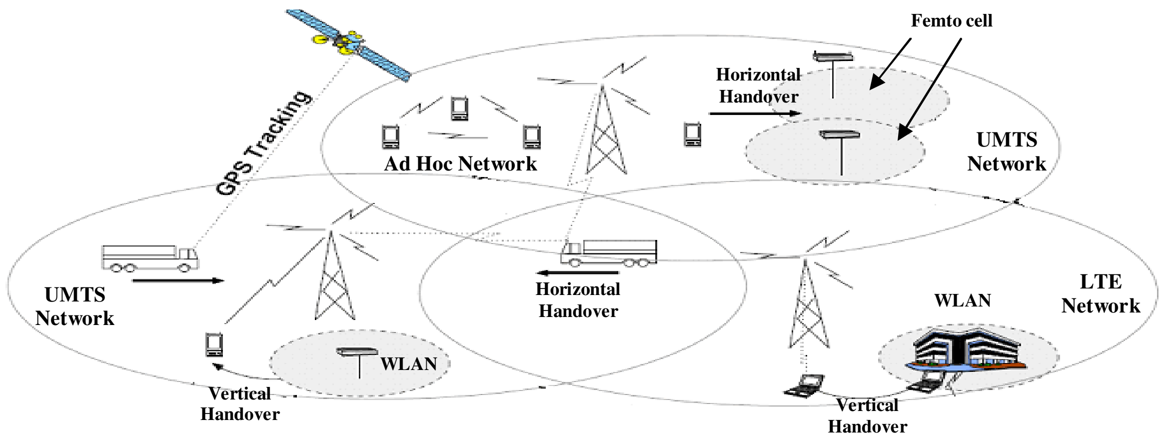


FIGURE 1: Mobile Next Generation Heterogeneous Network

The success of this transmission strategy relays on the quality and quantity of cognition possessed by CR user. This cognition is obtained through rigorous sensing of the radio channel and ability to characterize the interference. Based on the sensing functionality, CR users are ought to adapt their transmission accordingly in a manner not to harm the transmission of primary (incumbent) users of the channel.

The problem of spectrum sensing is a typical tradeoff problem where the accuracy and system simplicity are inversely related. The most known sensing techniques used are match filtering, energy detection, and cyclostationary features detection [4][5]. Match filtering is the technique with optimal detection, however due to system requirements it is practically difficult to implement [4]. Though at a lower level of implementation difficulty the performance of cyclostationary features detection is near optimal, system complexity is not trivial [4]. Energy detection is the least complex and most inaccurate among the three methods [5].

Wideband spectral sensing is a challenging aspect in cognitive radio sensing [3][7][8][9]. In [10][11][12], sequential sensing was introduced where a wideband radio channel is sensed using tunable narrowband bandpss filter at the RF front-end to sense one narrow frequency band at a time. In [8][9], multiband joint detection approach for wideband spectrum sensing was proposed where a bank of narrowband subchannels is used to concurrently sense a wideband rather than considering a single band at a time. The complexity of the two techniques (sequential or concurrent) is considered none travail especially as the bandwidth sensed grows. The accuracy is also compromised as they suboptimal detection is utilized.

In this paper we address the issues of wideband sensing. We introduce novel sensing techniques to be used for interference characterization for cognitive radios in heterogeneous networks. These techniques use of the chirp signal in an infrastructural cognitive radio networks.

Chirp signal is a wideband signal that has “interesting” cross correlation characteristics in time and frequency domains. The use of chirp signal is shown to ease system complexity and improve quality of sensing and therefore offer enhanced cognition at cognitive radio user in heterogeneous networks.

This paper is organized as follows: In II, we set up the scene for heterogeneous next generation network. In III, we discuss how cognitive radio can be utilized in heterogeneous environment. In V, we explain and analyze our sensing methodology. In VI, we present the simulation models. In VII, the results are presented and discussed. In VIII, we conclude.

These instructions are for authors of submitting the research papers to the International Journal of Computer Science and Security (IJCSS). IJCSS is seeking research papers, technical reports, dissertation, letter etc for these interdisciplinary areas. The goal of the IJCSS is to publish the most recent results in the development of information technology.

These instructions are for authors of submitting the research papers to the International Journal of Computer Science and Security (IJCSS). IJCSS is seeking research papers, technical reports, dissertation, letter etc for these interdisciplinary areas. The goal of the IJCSS is to publish the most recent results in the development of information technology. These instructions are for authors of submitting the research papers to the International Journal of Computer Science and Security (IJCSS). IJCSS is seeking research papers, technical reports, dissertation, letter etc for these interdisciplinary areas. The goal of the IJCSS is to publish the most recent results in the development of information technology. These instructions are for authors of submitting the research papers to the International Journal of Computer Science and Security (IJCSS). IJCSS is seeking research papers, technical reports, dissertation, letter etc for these interdisciplinary areas. The goal of the IJCSS is to publish the most recent results in the development of information technology.

2. MOBILE NEXT GENERATION NETWORK

Mobile Next Generation Network (MNGN) is a heterogeneous network. Figure 1 demonstrates the “operational theater” for heterogeneous network in which different access technologies are arranged in different topologies. And being accessed by a range of user' devices that enable variety of applications. The management of such scene constitutes a challenge for both operators and regulators as resources being depleted. To overcome those challenges innovative solutions are coming into place.

Some of challenges related to our work in heterogeneous network could be summarized as follows [1];

1. Increased demand on capacity and bandwidth
2. Offering differentiated services and application
3. Controlling operational cost

As for the capacity question, increasing spectral efficiency is more viable solution than boosting operators' spectrum that is a result of scarcity of available frequency bands [13]. Classical solutions to improve spectral efficiency include; (1) the usage of Multi Input Multi Output systems (MIMO), (2) adaptive Modulation and Coding (AMC) and (3) Deployment of smaller cells (Macro/Micro cells). Even though the overall spectral efficiency improves dramatically, system complexity and hence the operational cost increases significantly.

2.1 Innovative Solutions

Innovative solutions are designed to be effective in terms of cost and operational management. We will be focusing in here on four of them, i.e. (1) the use of femto cells, (2) vertical handoff, (3)self organization, and (4) Direct Spectrum Access (DSA)

2.1.1 Femto Cells

The deployment of smaller cell i.e. femto cells is becoming an attractive solution. Network operators are projecting that deployment of smaller cells in indoor spaces such as houses, enterprise building and public spaces will dramatically improve special coverage and off-load traffic from macro/micro-cells. The operational cost for this strategy is also minimized having the mini cells getting connected to the backhaul network wirelessly [19].

2.1.2 Vertical handoff

In a heterogeneous network where different access technologies overlap with one another, vertical handoff (handover) become a necessity to guarantee certain level of quality of service for mobile users. Vertical handoff will become a feature of MNGN especially in scenarios where connectivity is established on bases of content awareness. Thus for example a mobile user with urge to download a multimedia file will have to handoff from 3.5G connection to WiFi connection. The overall outcome of such technique is offloading traffic from bandwidth limited network [14].

2.1.3 Dynamic Spectrum Access

Dynamic spectrum access is a technology aimed at enabling cellular networks to exploit gaps and opportunities for accessing the radio channel when no activities are monitored. This technology can dramatically increase the spectral efficiency at low level of complicity as it won't impose any further requirement on existing networks operating using the current regulations of spectrum licensing, this in turn will boost current infrastructure and investments.

2.1.4 Self Organization

In this innovative solution, the problem of network capacity is addressed from the prospective of the mobile part of the system rather than the fixed part. In another word, group of users will organize their transmissions (among themselves) in an ad hoc fashion of communication. They would probably also use each other as intelligent relays to take their messages to farther destinations. Capitalizing on improved network coverage at low complexity, self organization is one important candidate to overcome the capacity demands for MNGN [15].

3. COGNITIVE RADIO NETWORKS

Cognitive technology is the underlying technology behind the solutions proposed to address capacity and performance improvement in MNGN. A network that enables self organization, DSA, handoffs between access technologies or handovers between micro/macro/femto cells will defiantly require this technology [1]. In such a system, possession of local cognition (via rigorous sensing) determines the vital decisions to be made by users in heterogeneous wireless networks.

3.1 Cognitive Network Architecture

Cognitive Radio Network can be deployed in different methods such as infrastructural and distributed architectures, to serve licensed and unlicensed applications. In this work we are concerned with infrastructural architecture. Figure 2 shows the system architecture. The system is hybrid and contains two networks; a primary radio network and a cognitive "adaptive" radio network [16]. The two networks are not physically connected however they meant to coexist.

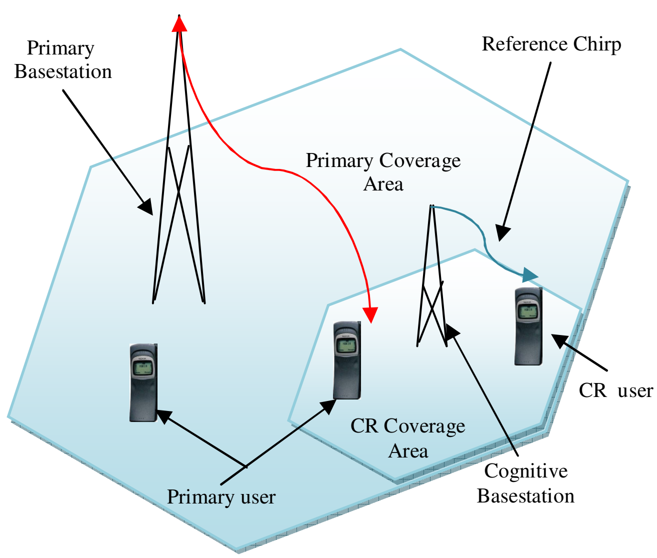


FIGURE 2: System Architecture

3.1.1 Primary Radio network
Primary radio network is the essential part of the system. It consists of a primary base station serving primary “licensed” users over the primary coverage area. The primary base station performs normal functions of a cellular base station.

3.1.2 Cognitive Radio network
The cognitive radio network is the adaptive part of the system. It consists of cognitive radio base station (CR-BS) serving cognitive radio users (CR-user). The cognitive radio coverage area overlaps with primary coverage area. A CR-user is ought to behave in opportunistic manner where it only can transmit after sensing the availability of the radio resources thus guarantee no excessive interference occurs at a primary user’s receiver.

4. Wideband Sensing Methodology
The proposed wideband cognitive radio sensing techniques is designed for infrastructural cognitive radio networks. A reference chirp signal that is transmitted over the coverage area of cognitive radio “femto” cell is used in this process [19]. The idea is to utilize resolution characteristics of chirp signal (in time and frequency) while removing excessive noise interference in the reception process.

4.1 Chirp Signal [17]
Wideband chirp signal is a result of linear frequency modulation of digital signal. The instantaneous frequency of the chirp signal increases or decreases linearly with time, Figure 3a shows a chirp signal. The bandwidth of a chirp signal, F, extends from the starting frequency sweep, f1, to the final frequency sweep f2. With proper choice for processing gain i.e. FT product, where T is the bit period, the spectrum of chirp signal has a distinctive near square shape, Figure 3b.

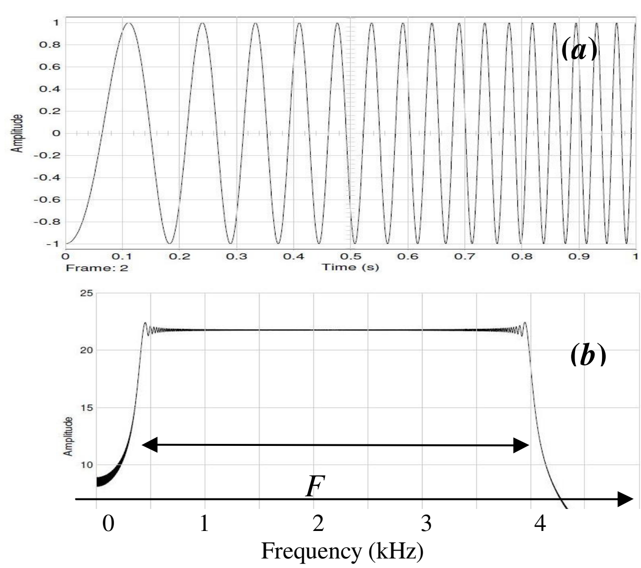


FIGURE 3: (a) Chirp signal, (b) Chirp spectrum

Chirp signal has very interesting correlation characteristics that gave it multi use in different applications [20]. In our methodology we are interested in two characteristics which will be helpful for sensing both frequency and time related behavior of primary user.
As for frequency sensing, spectral resolution in the presence of white noise is sought. The spectral resolution is obtained by cross-correlating the chirp signal with locally generated copy of itself (i.e. matched filtering). The result of this is optimal reception of chirp signal where excessive noise components are removed, Figure 7b.

As for temporal (time related) sensing the resolution sought is in the time domain. This resolution is obtained by correlating the received chirp with a locally generated conjugate of itself. The result of this operation is removal of noise components and resolution in time domain, Figure 13a.

4.2 Frequency Sensing Methodology

The novel frequency sensing methodology can be summarized as follows:

CR-BS broadcast low powered reference chirp signal with a bandwidth covering the sensed frequency spectrum.

After traveling over the radio channel and interfering with primary users' transmissions and noise, the reference chirp signal is then received at the CR-user using a chirp signal matched filter.

Fast Fourier Transform (FFT) is then applied to the output of the match filter. Figure 7a shows the spectrum of the received reference signal. As it is shown, the utilized (interfering) frequencies appears as spikes (peaks) rising above the flat floor of the received chirp signal spectrum.

The output from the previous step is fed into a decision circuit stage where a threshold value is set to decide the minimum amplitude of utilized frequencies.

4.2.1 Decision Circuit

Decision circuit is an algorithm implemented in software to detect the peaks representing primary users' frequencies. This algorithm belongs to search algorithms family and could be implemented either using sequential or binary search. The sequential search algorithm will sweep across the magnitude values of FFT samples and sequentially comparing them to a threshold value (which determines the existence of the tone.) However, the binary search algorithm will first sort the values of the FFT samples then it will discard the samples below the threshold value. Either algorithm should return the frequency values at which the FTT magnitude values exceed the threshold.

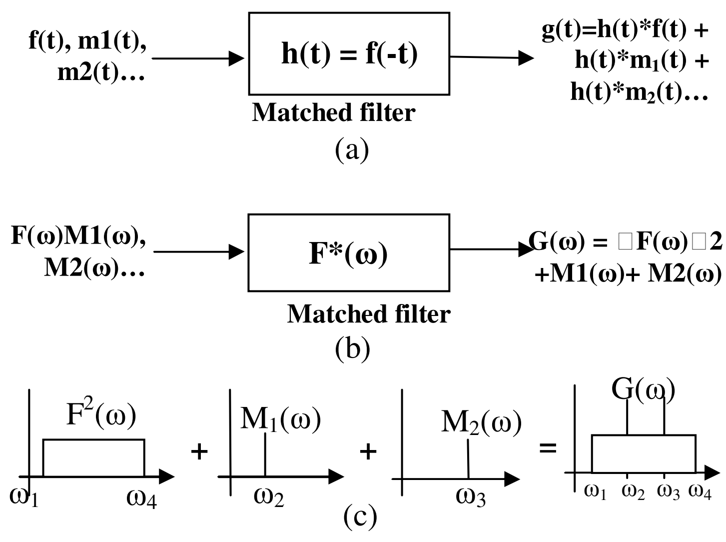


FIGURE 4: Analysis

4.2.2 Mathematical Modeling

Mathematical modeling can show a proof the concept for our methodology. Figure 4a shows the block diagram of the system. The inputs to the chirp signal matched filter is the reference chirp signal f(t), and the interfering carrier signals (tones) of primary users m1(t), m2(t)…

For simplicity, we will address the problem in the frequency domain as time convolution is transferred into frequency multiplication, Figure 4b. Therefore:

$$G(\omega) = F2(\omega)+ F(\omega)\, M1(\omega) + F(\omega)M2(\omega) \qquad (1)$$

Since the frequency representation of a sine wave is the unit impulse function in the frequency domain shifted to its corresponding frequency, (2) can be further simplified as:

$$G(\omega) = F2(\omega) + M1(\omega) + M2(\omega) \qquad (2)$$

Assuming that the spectrum of chirp signal is a square wave function, Figure 3c visualizes $G(\omega)$.

4.2.3 Temporal (time related) Sensing Methodology
Temporal sensing for primary user's time related behavior could be summarized as follows:
CR-BS broadcasts a low power reference chirp signal with a bandwidth covering the sensed spectrum.
After traveling over the radio channel and interfering with primary users' transmissions and noise, the reference chirp signal is correlated with a locally generated conjugate of the reference chirp signal. Figure 13b shows the received signal. As it is shown, the presence of the tone is sensed as soon as the flat top starts to change.
Finally, the output of chirp signal correlation is fed to delay estimation circuit to estimate the delay referenced to the starting moment of tone's reception.

4.2.4 Delay Estimation Circuit
Delay estimation circuit is simply a timer that starts counting the tone delay referenced to the starting time of the chirp signal reception. The timer is re-set as soon as the flat top of received chirp signal has begun to deform. The deformation corresponds to the moment a primary user starts to transmit. To sense this moment, received samples must be compared against a threshold value. The threshold value should be set just above the flat top of the received waveform.

5. Simulation Model
Simulations models using Matlab are constructed to draw initial conclusions on the sensing methodology.

5.2 Frequency Sensing Simulation Model
Figure 5 shows a block diagram for the proposed algorithm implemented using Matlab™. The reference chirp signal is received at the CR-receiver after interfering with primary user's signals in AWGN channel. The Chirp signal is firstly received by chirp signal matched filter. Then FFT is applied and the output of the FFT passes to the decision circuit to decide whether primary users interfering with the referenced chirp signal.

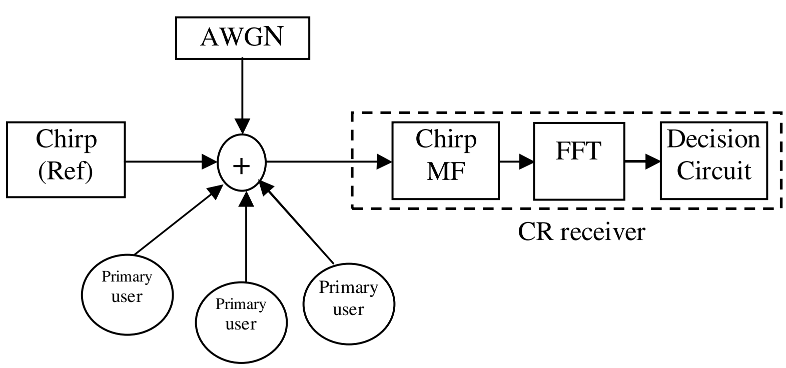


FIGURE 5: simulation model

5.3 Temporal Sensing Simulation Model
Figure 6 shows a block diagram for the simulation model that was implemented using Matlab™. The reference chirp signal is received at the CR-receiver after interfering with primary user's signals in AWGN channel. The Chirp signal is firstly received and cross-correlated with a locally

generated conjugate of the reference chirp signal. Then the output passes to a delay estimation circuit to estimate the moment when primary user's transmission took place.

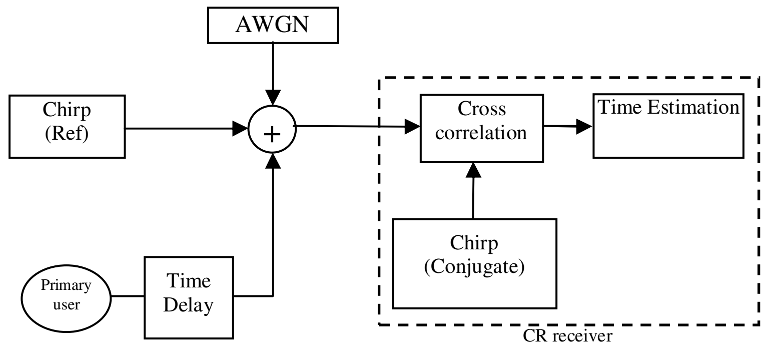


FIGURE 6: simulation model

6. RESULTS AND DISCUSSIONS

6.1 Frequency Sensing Performance Evaluation

Figure 7 shows spectrum of received chirp signal. The "interesting" characteristic or received chirp signal is obvious where a near flat floor extends over the bandwidth i.e. 3600 Hz. For chirp signal period of 1 s, the processing gain is 3600. It is shown that two distinctive peaks occur at frequencies corresponding to primary users' carrier frequencies at 500 Hz and 1800 Hz. Figure 7a shows the scenario where Signal to Interference plus Noise Ratio (SINR) for the interfering (primary users) carriers is 20 dB. Figure 7b shows the scenario where SINR is -5 dB. It is obvious that as SINR decreases, noise floor rises toward the peak value.

In order to quantify the performance of the system we define d which measures the distance (in dB) between the peak of the carrier's spike and the flat floor of chirp signal spectrum. In Figure 8, we plot the carrier's SINR versus d, normalized to value of d at SINR = 10 dB. The value of d = 0 dB signifies that the spike is no longer distinguishable from the noise and therefore probability of false alarm is very likely. The level of the noisy floor should determine the threshold value for the decision circuit. It is obvious that as SINR decreases d decrees. For our setup, it is shown that d = 0 at SINR= -25 dB which is extremely a low SINR value.

From above discussion it is becoming very important for a cognitive user to be able to set the threshold value for its decision circuit. This process, which depends on the level of Signal to Noise Ratio (SNR) of received reference chip signal, should be dynamic as the value of SNR changes upon many factors in the mobile environment. In order to address this problem, CR-user must be capable of estimating the SNR of received reference signal. This estimation could be further fine tuned by CR-user to minimize probability of false detection.

The results in Figure 8 assume that primary users' transmissions are in synch with the reference chirp signal. The performance is expected to degrade as this assumption is getting relaxed which is probably more a realistic assumption. Figure 9 illustrates the delay scenario between reference chirp signal and primary user's carrier.

To investigate the effect of a more realistic scenario where both networks (primary and cognitive) are not synchronized, Figure 10 shows how d changes with respect to the delay between primary user's signal and reference chirp signal. The results shows SINR in two cases, SINR = 10 dB and -10dB. As expected, system performance degrades as delay increases. For example, if the user's signal is delayed by 0.5Tc s, Tc is the chirp signal bit period, d drops by 25% in both SINR scenarios. Nonetheless, the system is showing tolerant to delay up to 0.25Tc s in both SINR cases. Also from Figure 10, we can deduce that below -10 dB the performance will greatly degrades especially if synchronization is not maintained.

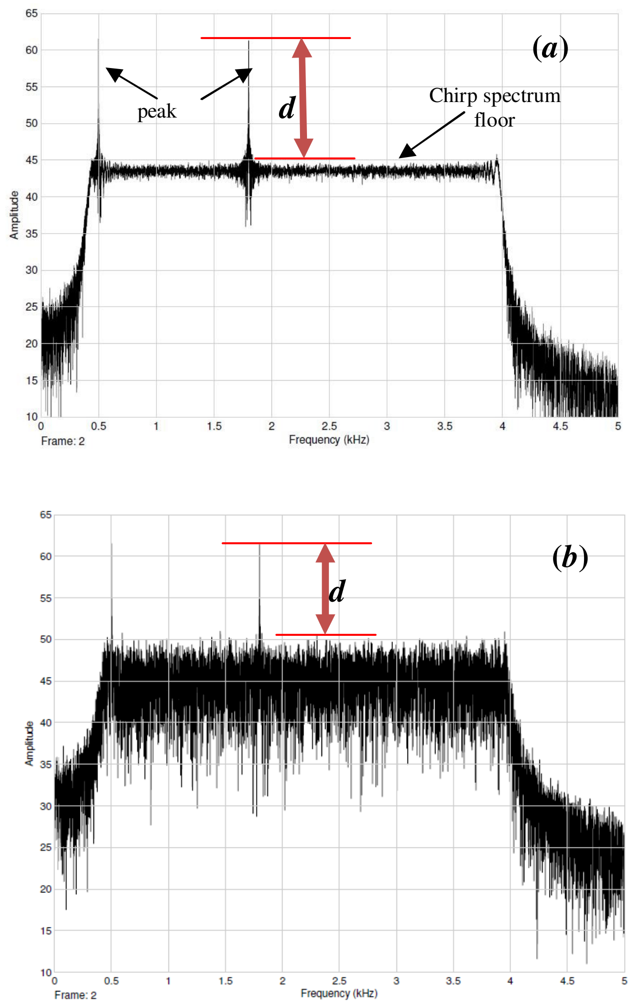


FIGURE 7: received chirp spectrum (a) SINR = 20 dB, (b) SINR = -5dB

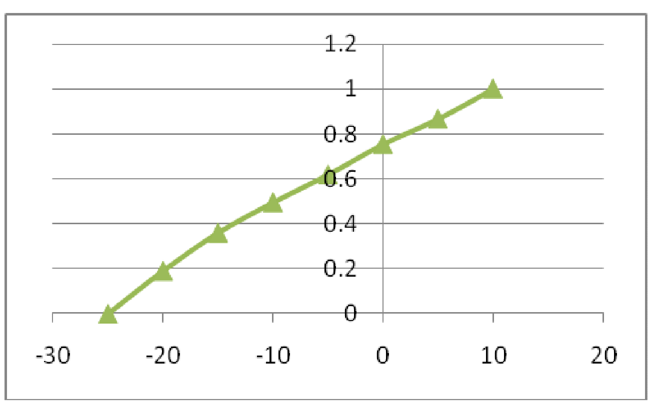


FIGURE 8: SINR versus normalized ***d***.

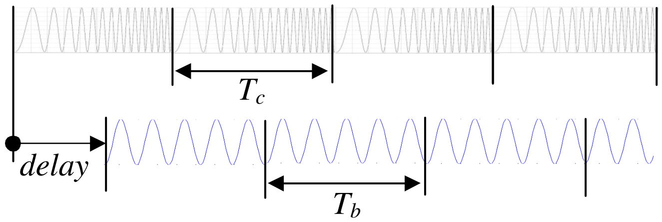


FIGURE 9: delay between reference chirp signal and received primary signal

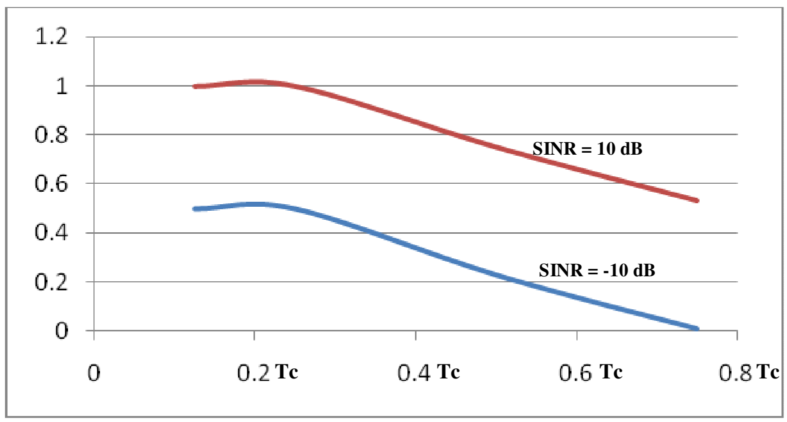


FIGURE 10: delay versus normalized d.

It must be noted that the results shown in Figure 10 simulate the scenario of one chirp signal with bit period Tc and a delayed carrier with bit duration of Tb. Also it must be noted that in this simulation Tb = Tc. The obvious conclusion to be drawn from this is that d is dependent on the overlapping duration between the chirp duration (Tc) and the bit duration (Tb). Therefore, even in worst case scenarios when the overlapping between the first chirp duration and the carrier bit is not enough, the consecutive overlapping should be enough to make a better decision. This may lead to the conclusion that observation interval must be at least twice as much as chirp signal period to avoid this drawback.

From this discussion, we can determine other important criteria necessary for the success of the sensing strategy which is the relation between Tb and Tc since it is evident that if Tb is not long enough to produce enough overlapping duration, the sensing strategy will fail. To drive the relation between Tb and Tc, we need to observe the relationship between the delay and d. in Figure 10, it is shown that in worst case scenario i.e. SINR = -10 dB, d has values greater than 0 if the delay is less than 0.75Tc. This is interpreted as that the minimum overlapping duration should be greater than one forth the chirp signal duration. On the other hand, if we want to estimate the duration of overlapping which is just enough to give optimal performance, we need to observe the performance between delays equal to 0 to 0.25Tc. It is shown that between delays equal to 0 to 0.25Tc, d remains unchanged regardless to the SINR. This is interpreted as that it is enough for Tb to last 0.75Tc and still being received optimally. To summarize this point, we imply that 0.25Tc <Tb< 0.75Tb or 1.33Tb < Tc < 4Tb.

Another important suggestion to be concluded from observing the relationship between Tb and Tc is possibility for sensing signals with spread spectrum and frequency hopping. However this point is out of scope in this paper, we argue that since we are capable of sensing tones with signaling rate higher than the signaling rate of the reference signal, spread spectrum sensing seems possible.

6.1.1 Compassion Between Spectrum Sensing Techniques

It is always difficult to compare results of different systems tested on different simulation testbeds. Therefore in this subsection we are concerned with subjective comparison. Future work should investigate this point thoroughly. And hence, the above discussion entitles us to highlight how our strategy addresses the main shortcoming of other known sensing strategies i.e. energy detection, matched filtering and cyclostationary feature detection.

As for energy detection technique, the main problem is setting the threshold value in presence of background noise. This problem is alleviated by the virtue of flat floor of chirp signal matched filter output which maintains good resolution. Here our sensing technique outperforms energy detection in two ways; firstly setting the threshold will be easier as it depends on estimating the SNR of received reference signal, and secondary, wideband resolution is very important as in case we are dealing with the 4G modulation OFDMA where subcarriers frequencies are set adjacent to one another. Figure 11a shows the spectrum of a tone received using our strategy of sensing and Figure 11b shows a tone received based on energy detection. It is obvious that the resolution in frequency domain is maintained better in the latter case. As a result of, improved measurement accuracy is insured. This should be obvious in dispersive mobile channel where frequency dispersion become a problem, in [18] this issue has been addressed for OFDMA.

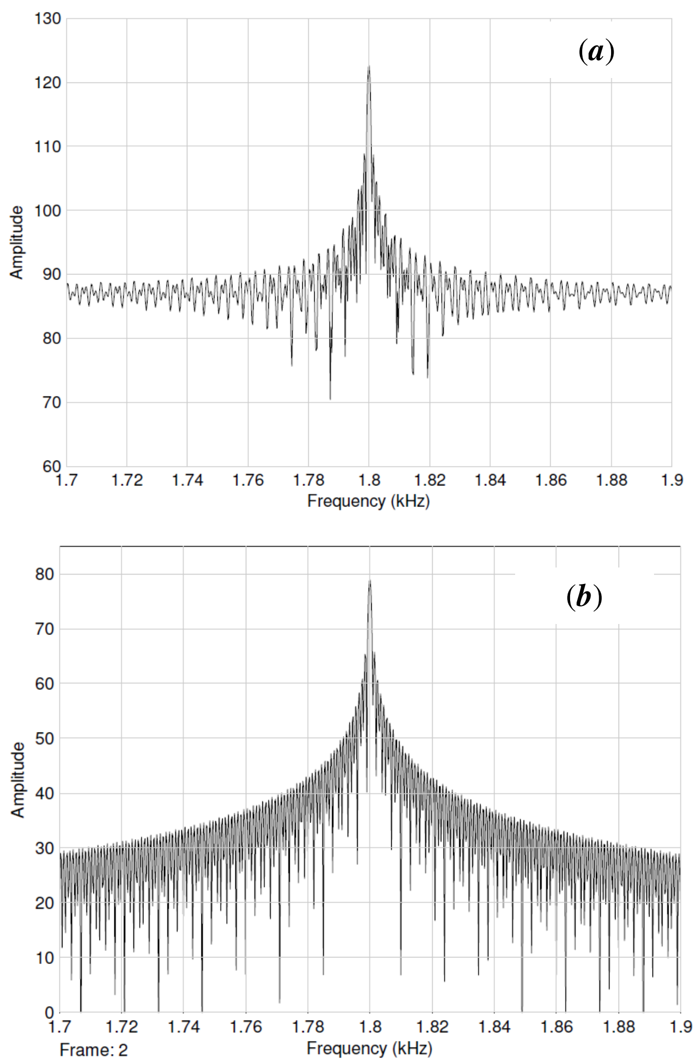


FIGURE 11: A simulation shown a tone received using (a) chip sensing and (b) energy detection

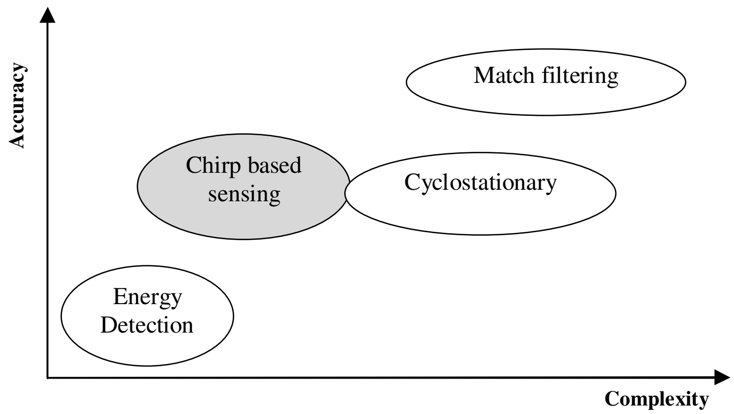


FIGURE 12: Sensing methods compression

As for match filtering, the main problem is that this method requires priori knowledge of primary user signal for optimal detection, in our strategy this is not the case. Although the accuracy of our strategy is not expected to match the optimal solution, the relief of complexity to sense wideband spectrum is considered an advantage. Finally, in cyclostationary feature detection, the main problem is increasing complexity however it is shown that the complexity of our system is moderated. Figure 12 shows a schematic of how we compare our sensing technique to the other techniques.

6.2 Temporal Sensing Performance Evaluation

Figure 13a and 3b show the output of received signal after cross-correlation with the conjugate of referenced chirp signal in time domain without and with an interfering tone respectively. As it is shown in Figure 13b, the presence of the tone is sensed as soon as the flat top of the cross-correlation's output starts to change. We denote Td as the time at which primary user started to transmit. Td is referenced to the beginning of chirp signal's reception.

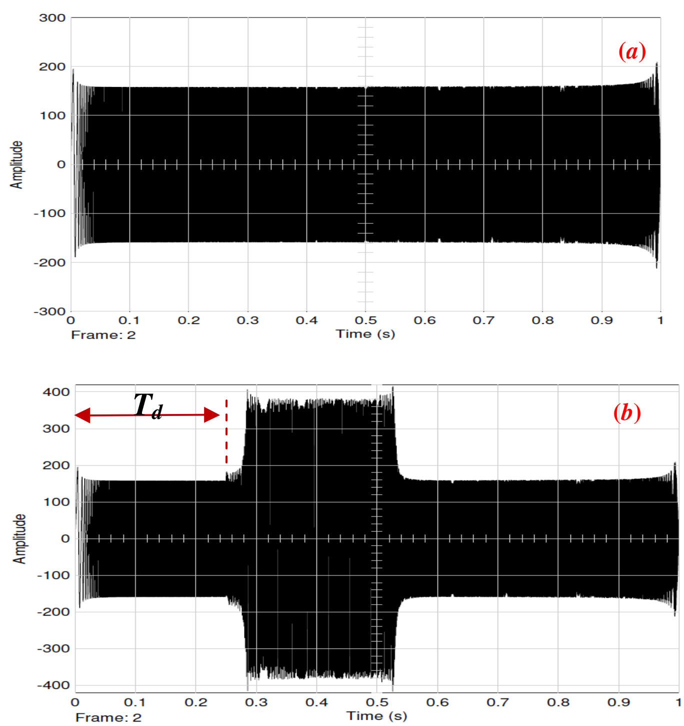


FIGURE 13: Output after correlation (a) without the presence of primary tone and (b) with the presence of primary tone.

In order to evaluate the performance, Figure 14 shows the delay estimation error probability Pd versus SNIR. It is shown that as SNIR decreases, error probability increases as the marking of the tone presence become difficult to recognize from background noise.
Further performance improvement is possible have we applied coherent “optimal” detection of the tone. The requirement for such improvement is a prior knowledge of the carrier frequency. This knowledge can easily be obtained from spectrum sensing based on chirp signal as we have shown above. Figure 14 shows the improvement using optimal detection.

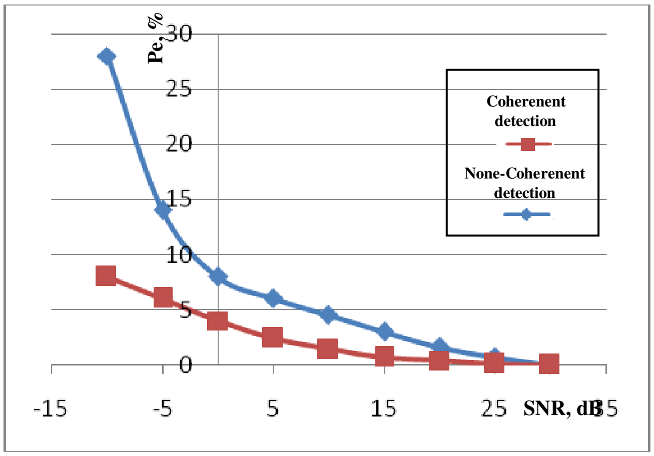


FIGURE 14: SNR vs probability of error

Another aspect to be investigated is the case of multi carrier reception. An example of the superposition of carriers resulted from this scenario is shown in Figure 15. To resolve this ambiguity, interference suppression technique should be used. This technique can be accomplished using band pass filtering to filter out the tones of interest (one at a time) having had knowledge of their frequencies.

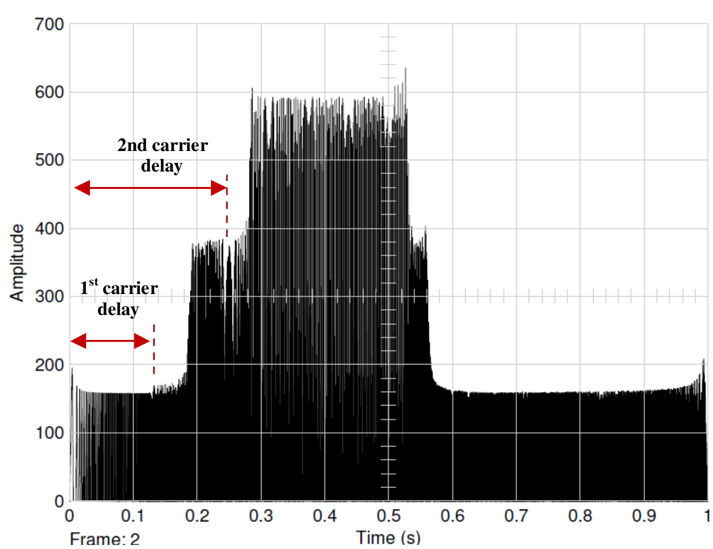


FIGURE 15: Superposition of two received tones

6.3 Sensing Based on Chirp Signal System Overview

Based on above discussion we can put together a design for intelligent sensing strategy. Figure 16 shows a block diagram of the system. Both modules for time and frequency sensing will interact to resolve ambiguity either in time or frequency estimation. Information from both modules is used to map interfering carriers along with the time it accesses the channel and their relative power measurements. Addition algorithms will be developed to obtain further "value added" information such as the set of subcarriers for OFDMA air interface, temporal behavior of users or unutilized timeslots for CDMA/TDD. In future work, hardware implementation based on Software Defined Radio (SDR) platform will be developed and tested.

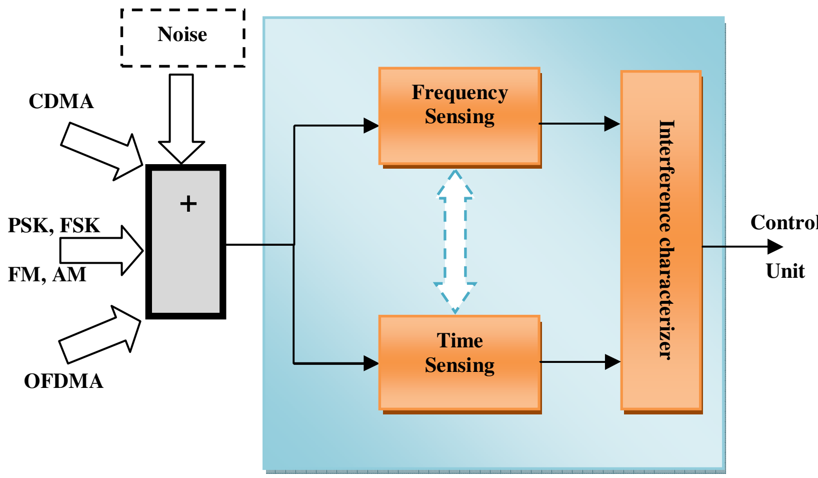


FIGURE 16: System block diagram

7. CONCLUSIONS

Our novel methods for sensing in cognitive radio environment significantly enhance spectral and temporal sensing at moderate complexity. We have evaluated the performance against different parameters and created different related arguments. Future work aiming to design an SDR-based system for interference characterization in heterogeneous future networks will benefit from these findings.

8. REFERENCES

[1] M. M. Buddhikot, "*Cognitive Radio, DSA and Self-X: Towards Next Transformation in Cellular Networks.*" In Proceedings of 2010 IEEE Symposium on New Frontiers in Dynamic Spectrum, 2010

[2] Federal Communications Commission, *"Notice of proposed rulemaking and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies,"* ET Docket No. 03-108, February 2005.

[3] T. Yucek. and H. Arslan, *"A survey of spectrum sensing algorithms for cognitive radio applications.",* IEEE journal for Communications Surveys & Tutorials, 11 (1), 2009

[4] D. Cabric, S.M. Mishra, R. W. Brodersen, *"Implementation Issues in Spectrum Sensing for Cognitive Radios."* In Proceedings of the 38th Annual Asilomar Conference on Signals, Systems and Computers, 2004

[5] *W. A. Gardner, C. M. Spooner, "Signal Interception: Performance advantages of Cyclic-feature detectors", IEEE Transaction on Communications, vol. 40, January 1992.*

[6] T. Yucek and H. Arslan, *"Spectrum Characterization for Opportunistic Cognitive Radio Systems* In Proceedings of IEEE MILCOM 2006, 2006.

[7] S. Hussain and X. Fernando, *"Spectrum Sensing in Cognitive Radio Networks: Up-to-date Techniques and Future Challenges,"* In Proceedings of the IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH), Toronto, Canada, 2009.

[8] Z. Quan, S Cui, V. Poor, *"Wideband Spectrum Sensing in Cognitive Radio Networks."* In Proceedings of IEEE International Conference on Communications, Beijing, China, 2008.

[9] X. Chen L and Zhao J Li, *"A Modified Spectrum Sensing Methods for Widband Cognitive Radio Based on Compressive Sensing."* In Proceedings of the 4th International Conference on Networking in China, Xian, China, 2009.

[10] A. Sahai and D. Cabric, *"A tutorial on spectrum sensing: Fundamental limits and practical challenges,"* in Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Baltimore, MD, 2005.

[11] Z. Tian and G. B. Giannakis, *"A wavelet approach to wideband spectrum sensing for cognitive radios."* In Proceedings of the 1st Int. Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), Mykonos Island, Greece, 2006.

[12] Y. Hur, J. Park, W. Woo, K. Lim, C.-H. Lee, H. S. Kim, and J. Laskar*, "A wideband analog multi-resolution spectrum sensing technique for cognitive radio systems."* In the Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), Island of Kos, Greece, 2006

[13] *"FCC September Commission Meeting Presentation,"* September 29,(2009).

[14] H. J. Wang, R. H. Katz, and J. Giese, *"Policy-enabled handoffs across heterogeneous wireless network*s." In Proceedings of IEEE Workshop on Mobile Computing Systems and Applications, 1999.

[15] A. Barnawi, "*Adaptive Technologies for Hybrid Ad hoc/Cellular Network Architecture,"* to appear in International Journal on Ad hoc and Ubiquitous Computing (IJAHUC), May 2011.

[16] *K*-C Chen and R. Prasad , "*Cognitive Radio Networks*" John Wiley & Sons Ltd, pp. 183-186, (2009)

[18] M. R. Winkley, *"Chirp signals for communications."* In the Proceedings of the IEEE WESCON conference, 1962.

[19] A. Ben Salem, M. Siala and H. Boujemaa, *"Performance Comparsion of OFDMA and OFDM/OQAM Systems Operating in Highly Time and Frequency Dispersive Radio Mobile Channels."* in Proceedings of the IEEE International Conference on Electronics, Circuit and Systems, Gammarth, 2005

[20] R. J. Haines, "*Cognitive Pilot Channels for Femto-cell Deployment."* in Proceedings of the 7th International Symposium on Wireless Communication Systems (ISWCS), York, UK, 2010

[21] L. Huang, Y. Lu and W. Liu, *"Using Chirp Signal for Accurate RFID Positioning."* In Proceedings of the International Conference on Communications, Circuit and Systems (ICCCAS), Chengdu, China, 2010

VEGAS: Better Performance Than Other TCP Congestion Control Algorithms on MANETs

M.Jehan *jehanjoy@yahoo.co.in*
Associate Prof., Department Of Computer Science,
D.J.Academy for Managerial Excellence,
Bharathiar University,
Coimbatore, 641 032, India.

Dr.G.Radhamani *radhamanig@hotmail.comProf.*
& Director, Department of Computer Science,
Dr.G.R.Damodaran College of Science,
Bharathiar University,
Coimbatore, 641 014, India.

T.Kalakumari *kalakumari_t@yahoo.co.in*
Assistant Prof., Department of Commerce & IT,
Sri Krishna Arts and Science College,
Bharathiar University,
Coimbatore, 641 008, India.

Abstract

The wireless communication TCP/IP protocol is an important role in developing communication systems and which provides better and reliable communication capabilities in almost all kinds of networking environment. The wireless networking technology and the new kind of requirements in communication systems need some extensions to the original design of TCP for on coming technology development. In this paper we have analyzed six TCP Congestion Control Algorithms and their performance on Mobile Ad-hoc Networks (MANET). More specifically, we describe the performance behavior of BIC, Cubic, TCP Compound, Vegas, Reno and Westwood congestion control algorithms. The evaluation is simulated through Network Simulator (NS2) and the performance of these congestion control algorithms is analyzed with suitable metrics.

Keywords: TCP Congestion Control Algorithms, MANET, BIC, Cubic, Compound, Vegas, Reno, Westwood.

1. INTRODUCTION

Ad-hoc networks are self-organizing wireless networks, in which all end nodes act as routers. This network improves the efficiency, range of fixed or mobile internet access and enables totally with new applications. A Mobile Ad hoc Networks (MANET) consists of a set of mobile hosts within communication range and exchange the data among themselves without using any pre-existing infrastructure. MANET nodes are typically distinguished by their limited power, processing and memory resources as well as high degree of mobility. In such networks, the wireless mobile nodes may dynamically enter the network and leave the network. Due to the limited transmission range of wireless network nodes, multiple hops are usually needed for a node to exchange information with any other node in the network.

MANET has potential use in a wide variety of disparate situations. Such situations include moving battlefield communications to disposable sensors which are dropped from high altitude and dispersed on the ground for hazardous materials detection. Civilian applications include simple scenarios such as people at a conference in a hotel with their laptops comprise a temporary

MANET to more complicated scenarios like highly mobile vehicles on the highway which form an ad-hoc network in order to provide vehicular traffic management.

In this paper, we have evaluated the Control Window (cwnd), Round Trip Delay Time (rtt) and the Throughput using the six algorithms on the performance of TCP in the wireless communication.

2. BACKGROUND WORK

2.1 Transmission Control Protocol and Congestion Control

TCP is one of the core protocols of the internet protocol family. TCP operates at a higher level, concerned only with the two end systems. In particular, TCP provides reliable, ordered delivery of a stream of bytes from a program on one computer to another program on other computer. Among its other management tasks, TCP controls segment size, flow control, the rate at which data is exchanged, and the network traffic congestion.

For the reliable packets delivery, TCP can support the mechanisms of flow and congestion control. Due to the unconstrained movement of the mobile nodes, TCP is unable to notice network congestion or link down to activate related controls on the MANET [2]. The standard congestion control mechanism of the TCP is not able to handle the special properties of a shared wireless multi-hop channel well. In particular, the frequent changes of the network topology and the shared nature of the wireless channel pose significant challenges [4].

TCP provides reliable end-to-end delivery of data over wired networks, several recent studies have indicated that TCP performance degrades significantly in MANET [6] [7] [8]. In [13], TCP-F is proposed to overcome the TCP false reaction towards route failures in MANETs. In [14] the simulation shows that the route change results in link disconnections, which reduces TCP throughput.

Vegas TCP was the first attempt to depart from the loss-driven paradigm of the TCP by introducing a mechanism of congestion detection before packet losses [9]. Westwood TCP is a new congestion control algorithm that is based on end-to-end bandwidth estimate [15].

Using TCP more computers are interconnected to increase data transaction between users rapidly. The MIMD and PIPD protocols developed and provides better throughput for the wireless networks [12], [16] and [17].

So, this study on the existing TCP congestion control algorithms and its performance on MANET will be very useful to design new algorithms exclusively for mobile wireless communication scenarios.

2.2 The Congestion Control Algorithms Under Evaluation

In the past years, after the invention of TCP, there is numerous congestion control algorithms discovered for different purposes. Each of them has unique characteristics [3]. But in this work we have selected six congestion control algorithms only because of these are the default algorithms in most of the open source and commercial operating systems.

2.2.1 Binary Increase Congestion Control (BIC)

BIC-TCP (Binary Increase Control-TCP) incorporated binary search increase in the protocol. Binary search increase provides reliable feedback on any network congestion and lost packets, allowing BIC-TCP to aggressively increase its transmission speed toward the maximum allowed by the high-speed network.

2.2.2 CUBIC

Binary Increase congestion Control for TCP v2.0 is called as CUBIC and it is a default TCP algorithm in Linux. The protocol modifies the linear window growth function of existing TCP standards to be a CUBIC function in order to improve the scalability of TCP over fast and long distance networks. CUBIC is a less aggressive and more systematic derivative of BIC, in which the window is a CUBIC function of time since the last congestion event, with the inflection point set to the window prior to the event.

2.2.3 COMPOUND TCP (C-TCP)

Compound TCP is a Microsoft implementation of TCP which maintains two different congestion windows simultaneously, with the goal of achieving good performance on, long fat networks (LFNs) while not impairing fairness. It has been widely deployed with Microsoft Windows Vista and Windows Server 2008 and has been ported to older Microsoft Windows versions as well as Linux.

2.2.4 TCP VEGAS

Until the mid 1990s, all TCPs set timeouts and measured round-trip delays were based upon only the last transmitted packet in the transmit buffer. In TCP Vegas, timeouts were set and round-trip delays were measured for every packet in the transmit buffer. In addition, TCP Vegas uses additive increases in the congestion window.

2.2.5 TCP NEW RENO (Reno)

TCP New Reno is the most commonly implemented algorithm; SACK support is very common and is an extension to Reno/New Reno. Starting with 2.6.8 the Linux kernel switched the default implementation from Reno to BIC.

2.2.6 TCP WESTWOOD+ (westwood)

TCP Westwood+ is based on end-to-end bandwidth estimation to set congestion window and slow start threshold after a congestion episode, that is, after three duplicate acknowledgments or a timeout. The bandwidth is estimated by properly low-pass filtering the rate of returning acknowledgment packets. The rationale of this strategy is simple: in contrast with TCP Reno, which blindly halves the congestion window after three duplicate ACKs, TCP Westwood+ adaptively sets a slow start threshold and a congestion window which takes into account the bandwidth used at the time congestion is experienced. The Algorithm significantly increases throughput over wireless links and fairness compared to TCP Reno/New Reno in wired networks.
A detailed description of TCP Westwood (TCPW) is shown in [15].

3. THE SIMULATION

The proposed simulations has been successfully implemented and evaluated using NS-2 simulator on a computer with Intel Core 2 Duo CPU (T6400 processor @ 2.00 GHz) 2 GB of RAM. A random wireless mobile ad hoc network topology was used for these experiments.

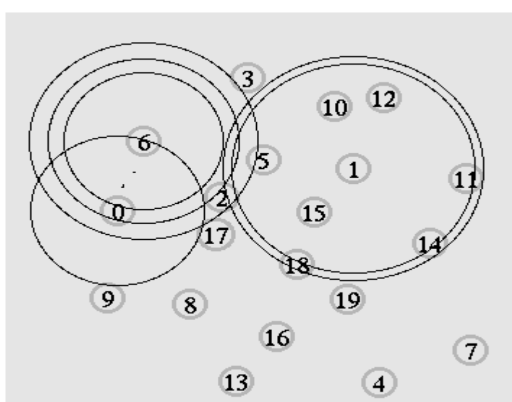


FIGURE 1: The MANET Scenario.

Some of the important parameters of the Ad-hoc Network simulation are:

Number of Nodes	20
Number of Sending Nodes	1
Topography	x=500 y=500
Mobility	0 or 20m/s
Mobility Start Time	20th Sec
Routing Protocol	AODV
Mac Type	802.11
Queue	DropTail / PriQueue
Queue Size	50
The Traffic Application	FTP
TCP Packet Size	1448
TCP Initial Window Size	30000

As far as the different parameters of congestion algorithm are concerned, all default parameters of TCP-Linux have been used in all our simulations. For simplicity and clarity of outputs, we used only one TCP flow during evaluating the algorithms.

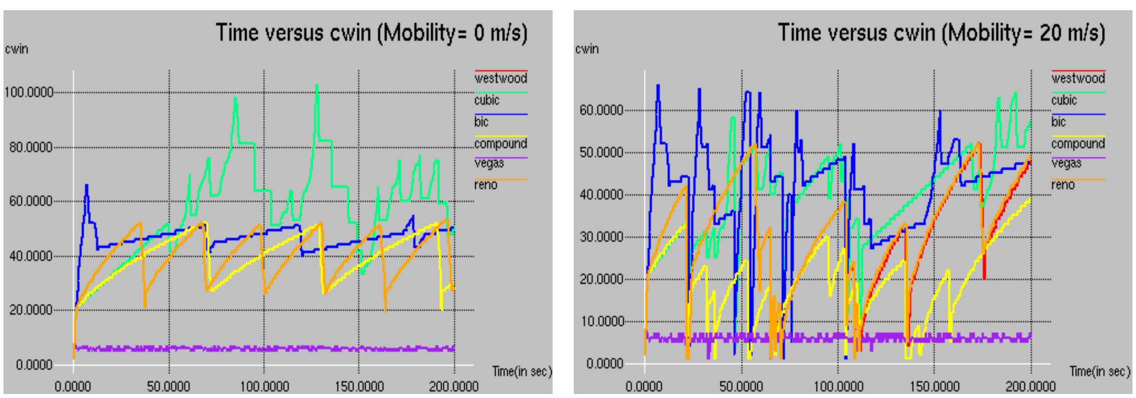
4. RESULT AND DISCUSSION

4.1 Simulation Results in MANETs

In this section, we carried out the simulation results of congestion control window, Round Trip Delay Time and Throughput in the Wireless Ad hoc Network. This simulation has been run for 200 seconds.

4.1.1 Control Window in MANETs

In the experimental network, we have used to perform evaluation of congestion control window comparison between above mentioned six algorithms as shown in the simulation result.



(a) (b)

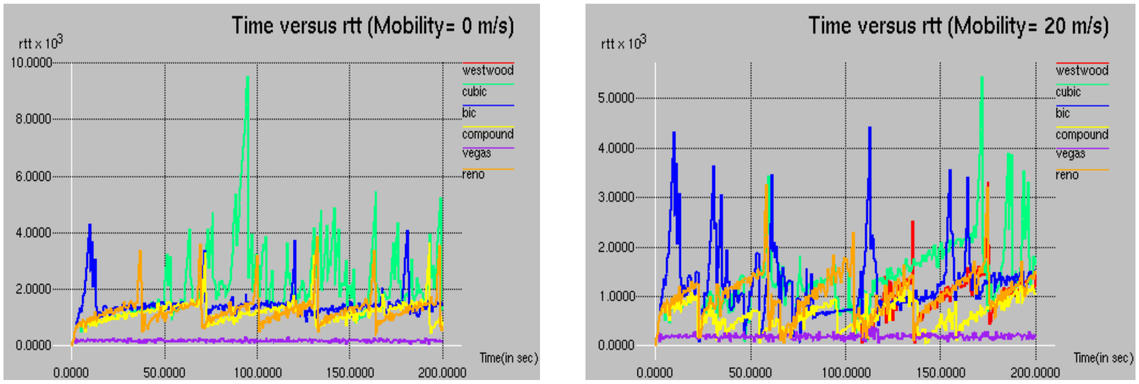
FIGURE 2: The cwnd Dynamics on Ad hoc Network – With and Without Mobility.

Figure 2 shows the, congestion control window increasing and decreasing in the all algorithms without any sequence except TCP Vegas.

In the both Figure 2 (a) and 2 (b), the exponential window size increase, linear increase and drop-off occurs irregularly during the simulation. In this Mobile Ad-hoc Networks scenario the TCP Vegas giving good result than other algorithms from starting to end of the evaluation.

4.1.2 Round Trip Delay Time in MANETs

The Round Trip Delay Time estimation of congestion control algorithms on wireless ad hoc network as shown in the Figure 3 (a) and 3 (b). This simulation result shows the TCP Vegas performance is better than other algorithms from the group of algorithms. As per the simulation setup, all nodes started to move randomly after 20 m/s.



(a) (b)

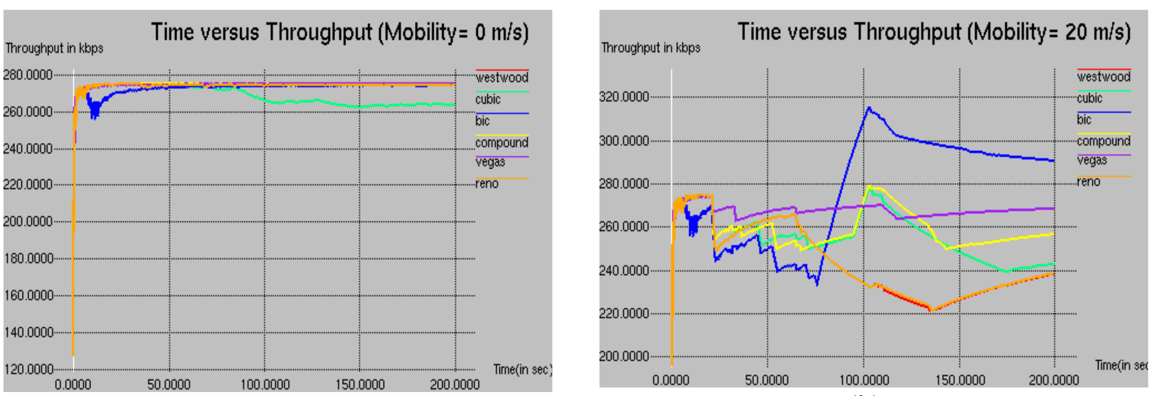
FIGURE 3: The rtt Dynamics on Ad hoc Network – With and Without Mobility.

As per the simulation result, the Figure 3 (a) and 3 (b) shows the TCP Vegas giving better result than other algorithms in the node mobility and non mobility environment.

4.1.3 Throughput Over Time in MANETs

The amount of data transferred from sender to receiver is processed in specified time duration. Data transfer rates for nodes and networks are measured in terms of throughput.

In this simulation, the throughput is the number of packets arriving at the sink per ms/second. Here we have find out the instant throughput over time during nodes movement and non movement.



(a) (b)

FIGURE 4: Throughput – without and with Mobility.

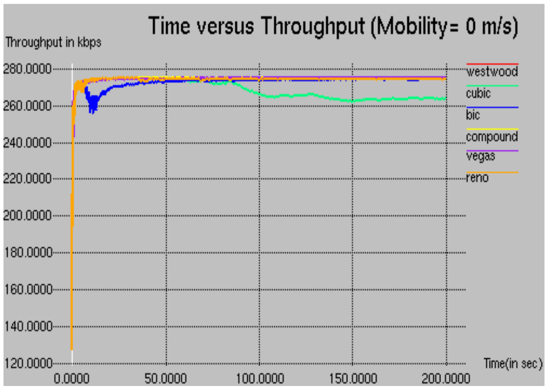
As shown in the Figure 4 (a), algorithms in the scenario without mobility, there was not much difference in provided throughput. But in the case of mobile scenario, as shown in the Figure 4 (b), the algorithm Vegas provided better throughput over time. The algorithm BIC provided best throughput after 75 seconds only but it provided poor results during initial phase of the communication.

4.2 Discussion of the Closer Analysis of Throughput Over Time

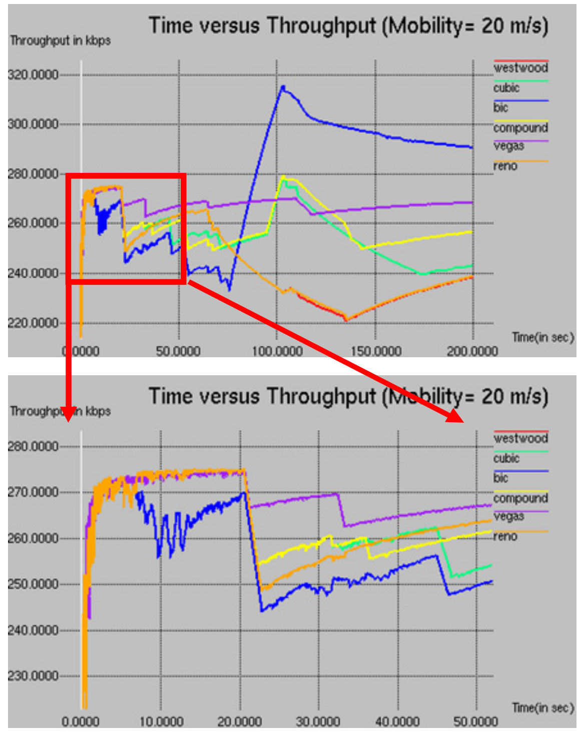
The following graphs illustrate the closer analysis of the congestion control algorithms in terms of throughput in the whole time duration.

4.2.1 Throughput Without Mobility

Figure 5 (a) shows the throughput over time in the case of mobility=0ms or without mobility. After 50 seconds, mostly all algorithms provided equal performance except CUBIC. During initial stage, even TCP BIC also very low throughput; but over time, it started to perform equal to other algorithms except CUBIC.



(a)



(b)

FIGURE 5: Throughput -without Mobility and with mobility closer analysis up to 50 seconds.

4.2.2 Throughput With Mobility

The Figure 5 (b) shows the throughput over time in the case of mobility=20ms. In this wireless network, the nodes started to move after 20 seconds. So up to 20 seconds, the graph is almost similar to after 50 seconds; all the algorithms provided equal performance except CUBIC. During the initial stage, TCP BIC also provided very low throughput; but all over the time TCP BIC performance is excellent than other algorithms from the group.

If we carefully observe the two sections (up to 50 seconds and 100 to 200 seconds) of Figure 5 (b), we can say that Vegas is the only algorithm which tried to provide almost constant and better throughput from starting to end of the simulation time. TCP BIC has started to provide better throughput after 75 seconds only. But it provided very poor throughput in the initial stage.

So as a final verdict, we have selected Vegas as the best performer in mobile ad hoc network scenario. As per the arrived results, we can say the algorithm Vegas can be used for short time TCP communication applications as well as long time TCP communication applications.

5. CONCLUSION & FUTURE WORK

We have successfully evaluated six congestion control algorithms using NS2 simulation tool in the Mobile Ad hoc Networks. The results are more significant and comparable. We have appraised the performance of these congestion control algorithms in very ideal condition without any cross traffic and any additional flows. In this small MANET scenario, the algorithm BIC provided good throughput after 75 seconds but algorithm Vegas provided stable and excellent throughput almost all over on the whole run time. So we move towards to the wrapping up that the algorithm Vegas will be the suitable algorithm for small and dynamic mobile ad hoc network scenario. Except Vegas, all other assessed algorithms provided very poor throughput during initial stage of the communication (less than 50 Seconds). So we conclude TCP Vegas will be the best algorithm from the list.

In this work, we have preferred six algorithms for evaluation, because they are default algorithms in several standard operating systems. But we have planned to do another evaluation based on the types of algorithms namely, Slow Start, Congestion Avoidance, Fast Retransmit and Fast Recovery. In our future work, we will select few algorithms from each of these four categories and will evaluate their performance in MANET scenario. There are other varying network parameters and metrics that the authors are working on the same. Based on the results, we can extend the further enhancement towards specific application on MANETs.

6. REFERENCES

[1] D. X. Wei and P. Cao, "*NS-2 TCP- Linux: An NS-2 TCP Implementation with Congestion Control Algorithms from Linux*", proceedings of ValueTool'06 -- Workshop of NS-2, Oct, 2006.

[2] Shin-Jer Yang , Yung-Chieh Lin, Soochow University, Taipei, Taiwan , "*Tuning Rules in TCP Congestion Control on the Mobile Ad Hoc Networks*", Proceedings of the 20th International Conference on Advanced Information Networking and Applications, Volume 01 , 2006.

[3] Jehan.M, G.Radhamani, T.Kalakumari, "*A survey on congestion control algorithms in wired and wireless networks*", Proceedings of the International conference on mathematical computing and management (ICMCM 2010), Kerala, India, June 2010.

[4] Christian Lochert, Björn Scheuermann, Martin Mauve, "*A survey on congestion control for mobile ad hoc networks*", Wireless Communications & Mobile Computing, Volume 7 , Issue 5, pp. 655 – 676, June 2007, ISSN:1530-8669.

[5] M.Allman, V.Paxson, and W.Stevens. *TCP Congestion Control. RFC2581 (Proposed Standard), Apr.1999. Updated by RFC3390*.

[6] L. A. Grieco, S. Mascolo, "*Performance evaluation and comparison of Westwood+, Vegas and New Reno TCP congestion control*", ACM Computer Communication Review, April 2004.

[7] I. Chlamtac, M. Conti, and J. Liu, "*Mobile ad hoc networking: imperatives and challenges*", Ad Hoc Networks Journal, vol.1, no. 1, pp. 13-64, Jul, 2003.

[8] A. Al Hanbali, E. Altman, P. Nain, "*A Survey of TCP over Mobile Ad Hoc Networks*", Research Report no. 5182, INRIA Sophia Antipolis research unit, May 2004.

[9] Brakmo, L. S., O'Malley, S.W., and Peterson, L.,"*TCP Vegas: End-to end congestion avoidance on a global Internet*", IEEE Journal on Selected Areas in Communications (JSAC), 13(8), pp. 1465-1480 (1995).

[10] Mascolo, S. "*Congestion control in high-speed communication networks*", Automatica, Special Issue on Control Methods for Communication Networks, Vol. 35, no. 12, pp. 1921-1935 Dec, 1999.

[11] Sally Floyd and Kevin Fall, "*Promoting the use of end-to end Congestion control in the Internet*," IEEE/ACM Transactions on Networking, vol. 7(4), pp. 458–472, Aug, 1999.

[12] Chandrasekaran M. and Wahida Banu R.S.D., "*Interaction Between Polynomial Congestion Control algorithms Queue Management schemes in wired TCP Networks*," International Journal of Soft Computing, Vol. 1, No. 2, pp.83-90, 2006.

[13] K. Chandran, S. Raghunathan, S. Venkatesan, and R. Prakash,, "*A feedback based scheme for improving TCP performance in Ad-Hoc wireless networks*", Proceedings of the International Conference on Distributed Computing Systems (ICDCS 98), Amsterdam, Netherlands, May 1998.

[14] Foez ahmed, Sateesh Kumar Pradhan, Nayeema Islam , and Sumon Kumar Debnath, "*Performance Evaluation of TCP over Mobile Ad-hoc Networks*" in (IJCSIS) International Journal of Computer Science and Information Security ,Vol. 7, No. 1, 2010.

[15] Mascolo, S., Casetti, C., Gerla, M., Sanadidi, M., Wang. R. "*TCP Westwood: End-to-End Bandwidth Estimation for Efficient Transport over Wired and Wireless Networks*", In the Proceedings of ACM Mobicom 2001, (Rome, Italy, July 2001).

[16] Chandrasekaran M, Kalpana M and Wahida Banu R.S.D., "*Congestion Control using Polynomial Window size adjustment Algorithms for wired and wireless networks*", In the Proceedings of International Conference on Network –ICN06 conducted at Mauritius, Apr, 2006.

[17] Chandrasekaran M, , Kalpana M and Wahida Banu R.S.D., "*Interaction between MIMD-Poly & PIPD-Poly Algorithms and other TCP Variants in Multiple Bottleneck TCP Networks*", In the Proceedings of IEEE Conference WOCN2006, Bangalore, Apr, 2006.

[18] Tomoya Hatano, Hiroshi Shigeno, Ken-ichi Okada, "*TCP-friendly Congestion Control for High Speed Network*", International Symposium on Applications and the Internet-SAINT'07, pp.10, 2007.

[19] David X. Wei, Cheng Jin, Steven H. Low, and Sanjay Hedge., "*Fast TCP: Motivation, Architecture, Algorithms, Performance*", IEEE/ACM transactions on networking, 2006.

[20] K. Satyanarayan Reddy and Lokanatha C. Reddy., "*A survey on congestion control mechanisms in high speed networks*", IJCSNS-International Journal of Computer Science and Network Security, vol. 8, no. 1, pp. 187 – 195, 2008.

[21] Colin Perkins and Ladan Gharal., "*Rtp and the datagram congestion control protocol*", In Proceedings of IEEE International Conference on Multimedia and Expo, Toronto, Canada, July 2006.

[22] Van Jacobson., "*Congestion Avoidance and Control*", Computer Communications Review, Volume 18 number 4, pp. 314-329, Aug, 1988.

[23] Van Jacobson., "*Modified TCP Congestion Control Avoidance Algorithm*", end-2-end-interest mailing list, pp.1-14 April 30, 1990.

[24] A Linux TCP implementation for NS-2. URL: http://www.cs.caltech.edu/weixl/ns2.html

[25] A mini – tutorial for NS-2 TCP-Linux. URL: http://www.cs.caltech.edu/weixl/ns2.html

INSTRUCTIONS TO CONTRIBUTORS

The International Journal of Computer Networks (IJCN) is an archival, bimonthly journal committed to the timely publications of peer-reviewed and original papers that advance the state-of-the-art and practical applications of computer networks. It provides a publication vehicle for complete coverage of all topics of interest to network professionals and brings to its readers the latest and most important findings in computer networks.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCN.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 3, 2011, IJCN appears in more focused issues. Besides normal publications, IJCN intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

IJCN LIST OF TOPICS

The realm of International Journal of Computer Networks (IJCN) extends, but not limited, to the following:

- Algorithms, Systems and Applications
- ATM Networks
- Cellular Networks
- Congestion and Flow Control
- Delay Tolerant Networks
- Information Theory
- Metropolitan Area Networks
- Mobile Computing
- Multicast and Broadcast Networks
- Network Architectures and Protocols
- Network Modeling and Performance Analysis Network
- Network Security and Privacy
- Optical Networks
- Personal Area Networks
- Telecommunication Networks
- Ubiquitous Computing
- Wide Area Networks
- Wireless Mesh Networks
- Ad-hoc Wireless Networks
- Body Sensor Networks
- Cognitive Radio Networks
- Cooperative Networks
- Fault Tolerant Networks
- Local Area Networks
- MIMO Networks
- Mobile Satellite Networks
- Multimedia Networks
- Network Coding
- Network Operation and Management
- Network Services and Applications
- Peer-to-Peer Networks
- Switching and Routing
- Trust Worth Computing
- Web-based Services
- Wireless Local Area Networks
- Wireless Sensor Networks

CALL FOR PAPERS

Volume: 3 - **Issue:** 4 - July 2011

i. Paper Submission: July 31, 2011 **ii. Author Notification:** September 01, 2011

iii. Issue Publication: September / October 2011

CONTACT INFORMATION

Computer Science Journals Sdn BhD
M-3-19, Plaza Damas Sri Hartamas
50480, Kuala Lumpur MALAYSIA

Phone: 006 03 6207 1607
006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2011
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org