MASSEY UNIVERSITY
COLLEGE OF SCIENCES

Declaration Confirming Content of Digital Version of Thesis

I confirm that the content of the digital version of this thesis

Title: Performance Evaluation of Multihop Wireless Network

is the final amended version following the examination process and is identical to this hard bound paper copy.

Student's Name: Liang, Shuai (Lynn)

Student's Signature:

Date: 2010-09-20

Performance Evaluation of Multihop Wireless Network

A thesis presented in fulfillment of the requirements for

Master of Engineering Degree

In

Electronics and Communications Engineering

School of Engineering and Advanced Technology
Massey University at Albany
New Zealand

Liang, Shuai

2010

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude and appreciation to my supervisor Dr. Mohammad Rashid for his guidance and support throughout my two-year study at Massey University. His enthusiastic supervision, professional suggestions and in-depth review of the thesis have extremely helpful in completing my research.

Furthermore, I take this opportunity to thank Dmitri Roukin who helped me with simulation environment configuration and troubleshooting in the early stage of my research. Sincere thanks go to Gill Sanders and her family for all the help and encouragement they offered during my stay at Albany.

My deepest gratitude goes to my beloved family for their loving considerations and great confidence in me all through these years. Thank you so much, Dad, for all your valuable time spent on listening to me and supporting me working out the problems all the way long.

ABSTRACT

In recent years, there has been an upsurge of interest in wireless broadband access networks in both industry and academia. This study aims at evaluating the performance of wireless access networks implemented in the multihop mesh architecture based on IEEE802.11 standards.

An implementation model is defined with the objectives to assess the impact of the variation of several network parameters including the number of mesh access points (MAPs) and stations (STAs), supported profiles, etc. A detailed analysis of the results gathered from 168 simulation runs in OPNET Modeler reveals that the number of MAPs in each extended service set (ESS) could be configured up to 4, the number of STAs associated to each MAP could be up to 8. On the other hand, the EDCA mechanism for QoS support from IEEE802.11e has been considered in the evaluation on both STAs and MAPs. The results show that enabling EDCA mechanism improves the global multihop network performance significantly in the scenarios with more streaming service (more bandwidth demanding) and more real-time applications (more delay stringent and more uplink bandwidth required).

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AC	Access Category
ACK	Acknowledgment
AIFS	Arbitration Interframe Space
AIFSN	Arbitration Interframe Space Number
AP	Access Point
BSS	Basic Service Set
BSSID	Basic Service Set Identification
CA	Collision Avoidance
CCA	Clear Channel Assessment
CD	Collision Detection
CRC	Cyclic Redundancy Code
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
CW	Contention Window
DA	Destination Address
DCF	Distributed Coordination Function
DIFS	Distributed (Coordination Function) Interframe Space
DS	Distribution System
DSSS	Direct Sequence Spread Spectrum
EDCA	Enhanced Distributed Channel Access
EDCF	Enhanced Distributed Channel Function
EIFS	Extended Interframe Space
ERP	Extended Rate PHY
ESS	Extended Service Set
FCS	Frame Check Sequence
FHSS	Frequency Hopping Spread Spectrum

FTP	File Transfer Protocol
HCCA	HCF Controlled Channel Access
HCF	Hybrid Coordination Function
HTTP	Hyper Text Transfer Protocol
IBSS	Independent BSS
IFS	Interframe Space
IP	Internet Protocol
ISM	Industrial, Scientific and Medical frequency band
LAN	Local Area Network
LLC	Logical Link Control
MAC	Medium Access Control
MAP	Mesh Access Point
MIMO	Multiple Input Multiple Output
MPDU	MAC Protocol Data Unit
MSDU	MAC Service Data Unit
MWN	Multihop Wireless Network
NAV	Network Allocation Vector
NLOS	Non-line-of-sight
NRTM	Non Real-Time Maximum
NRTC	Non Real-Time Centric
NRT	Non Real-Time
OFDM	Orthogonal Frequency Division Multiplexing
PBCC	Packet Binary Convolutional Coding
PCF	Point Coordination Function
PHY	Physical Layer
PINC	Pairwise Intersession Network Coding
PLC	Power Line Communications
PLCP	Physical Layer Convergence Procedure

PPDU	PLCP Protocol Data Unit
PSDU	PLCP Service Data Unit
QoS	Quality of Service
RT	Real-Time
RTC	Real-Time Centric
RTM	Real-Time Maximum
RA	Receiver Address or Receiving Station Address
RTS	Request to Send
SA	Source Address
SIFS	Short Interframe Space
STA	Station
TA	Transmitter Address or Transmitting Station Address
TCP	Transmission Control Protocol
TXOP	Transmission Opportunity
TGs	802.11s Task Group
UP	User Priority
VoIP	Voice over Internet Protocol
WDS	Wireless Distribution System
WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WM	Wireless Medium

CHAPTER 1 INTRODUCTION

Chapter 1 commences with an introduction to the background information including a brief history of wireless access technology. The advantage of multihop wireless networks in some particular situations is presented along with some technical limitations in this field. A short overview of the work is then provided, presenting the aims and methodology of the work. At the end, the chapter provides an outline of the thesis.

1.1 Background

As a senior futurist Thomas Frey predicted in 2006, “The world of wires has already begun its long descent into oblivion as wireless technology improves to the point where wires become obsolete” (Frey, 2006). Although fibre optics has overwhelming advantages in bandwidth and transmission loss, the cost in installation and maintenance, especially for rural areas, limit their applications in access network services. Furthermore, the increasing demand for having the ability to communicate wherever and whenever has lead to an inevitable trend in wireless access.

Regarding the current state of wireless technology applications, wireless connectivity became virtually universal, to a great extent, because of the efforts of the Wireless Fidelity Alliance (WiFi) certification program, which has been the single most important factor for being able to use wireless adapters or embedded chips universally. Today, personal computers, PDAs and telephones, almost every form of consumer electronics without exception, support WiFi connectivity getting access to the Internet via wireless access networks. As a commonly-used wireless access mode, WiFi technology which is built on IEEE802.11 standards has been developed for the last

two decades.

Since first released in 1997, the IEEE802.11 standards have addressed the Physical (PHY) layer and Media Access Control (MAC) layer standards separately. The original PHY standard provides data rates of 1-2 Mbps and three fundamentally different mechanisms of operation, namely, Infrared, 2.4 GHz Frequency Hopping Spread Spectrum (FHSS), and 2.4 GHz Direct Sequence Spread Spectrum (DSSS). Since wireless stations do not have the capability of detecting collisions, the MAC employs an access method that made every effort to avoid collisions, which is known as Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). In Chapter 3, we will take an in-depth investigation on how the PHY and MAC standard support the operation of wireless networks.

Following the deployment of the standards, several amendments were introduced, trying to compensate for all the drawbacks it had, compared to wired networks: less secure, less reliable, or less throughput. These amendments included the definition of the new physical layer specifications (IEEE802.11a/b/g) to enable higher data rate and throughputs (IEEE802.11n), to increase security (IEEE802.11i), to enable its usage in countries or areas with different spectrum regulation (IEEE802.11d/h/j), or to form a unified network of mesh topology composed of multiple WLAN cells (IEEE802.11s), as listed in Table 1.1

Table 1.1 IEEE802.11 Standards(IEEE LAN/MAN Standards Committee, 2010)

standard	Specification
802.11	The original standard: 1/2 Mbps, 2.4GHz RF and IR standard (1997)
802.11a	New physical layer, 54 Mbps, 5GHz (1999)
802.11b	New physical layer, 11 Mbps,2.4GHz (1999)
802.11d	Deals with issues related to regulatory differences in various countries (2001)
802.11g	New physical layer, 54 Mbps, 2.4GHz, compatible with 802.11b (2003)
802.11h	Spectrum managed 802.11a (for Europe) (2004)
802.11i	Provides a stronger encryption than WEP and other security enhancements (2004)
802.11j	Extensions for Japan (2004)
802.11e	New Media Access Control layer, to enable Quality of Service Support (2005)
802.11k	Defines the information that should be provided to higher layers, in order to facilitate the management and maintenance of a WLAN (2008)
802.11n	Enhancements for higher throughput (2009)
802.11s	Mesh networking (in process)

Most of the currently deployed wireless networks operate in infrastructure mode, which rely upon wireless links between wired infrastructure devices (access points) and end user (mobile stations), such as the cellular mobile networks (GSM, CDMA, etc.). However, cost-effective deployment of infrastructure-based solutions is desired in order to meet economic feasibility criteria when emerging requirement for getting access to Internet in a tolerable data rate. In the circumstances, it could be appealing for interconnect access points (APs) via wireless, instead of any physical wired connection to the core network.

The demand and constraints on currently deployed wireless networks outlined above

lead to a multihop mesh architecture, where APs become Mesh Access Points (MAPs) able to deliver traffic from source to destination by means of multihop relaying. In this architecture, some MAPs might operate as a portal or gateway to allow access to the Internet, as depicted in Figure 1.1.

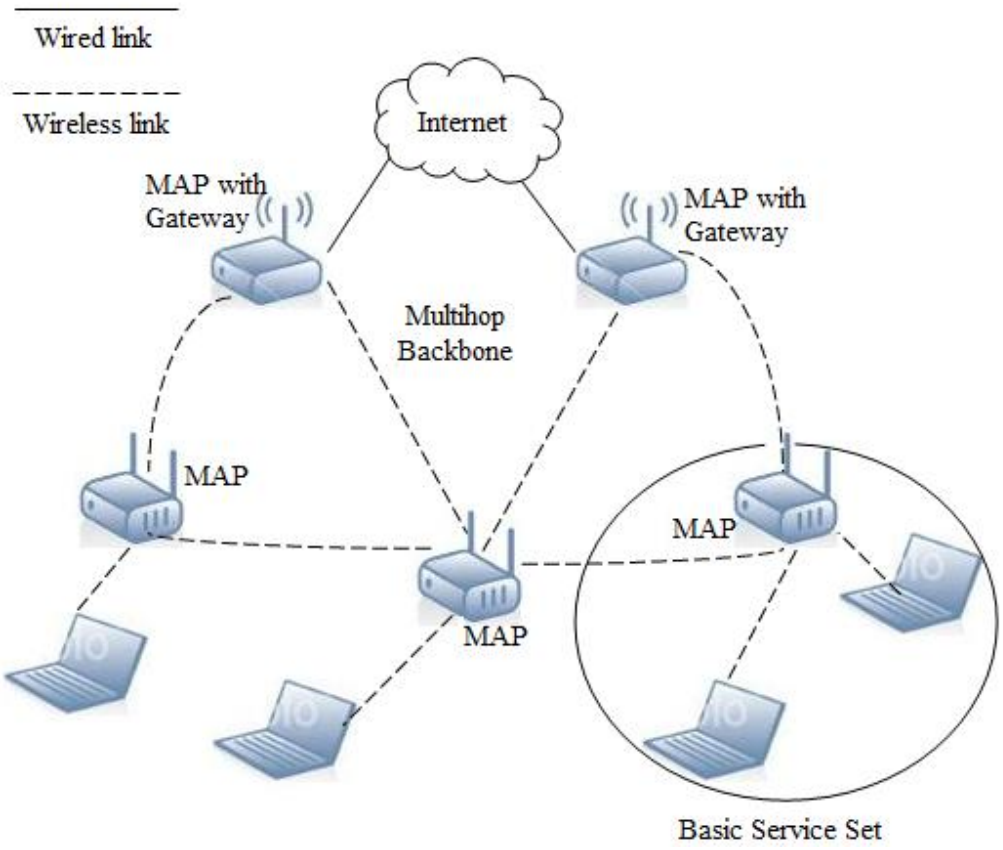


Figure 1.1 An Example of Multihop Wireless Network

The above architecture inherits the characteristics of a multihop wireless network (MWN), which can extend the network coverage using variable wireless access technology. The literature, which contains significant research in the field of MWN, will be reviewed in Chapter 2.

1.2 Objectives of the Study

MWN is, more than a promising technology, now considered as a fundamental

instrument to enable a ubiquitous wireless Internet. Nevertheless, there is still a number of challenging research topics at all protocol layers level that need to be addressed, in order to take advantage of all mesh network potentialities. Among them is the identification of the relationship between network capacity and other factors, such as: network architecture, network topology, traffic pattern, network node density, number of channels used for each node, transmission power level and node mobility (Akyildiz, Wang, & Wang, 2005).

Although the increasing importance of mesh networks and the great number of foreseen applications has driven them to become a hot topic in wireless communications research worldwide, there are still many challenges to be tackled, both on the theoretical and practical sides, for understanding performance limits and devising design principles for infrastructure-based MWNs, for both narrowband and wideband application. Especially, the combination of wireless forwarding and routing protocols allows the establishment of all wireless end-to-end routes among communicating devices placed far away from each other, which could not exist if only standard IEEE802.11 networks were used. Taking this observation into account, the present study aims at evaluating the impact of several parameters into multihop wireless networks capacity and performance in the scope of IEEE802.11 standards.

1.3 Research Approach

The approach used for achieving the above objectives can be divided into four stages.

Firstly, an in-depth review of the current literature published in the area of MWNs was carried out. This was significant for identifying the trends and issues, and defining the scope of the study. Besides, the study on wireless access technology based on IEEE802.11 was necessary for the operational part of the present study. It could be considered as the theoretical foundation to the following stages.

Secondly, a model of MWN was developed followed by scenarios configurations and several degrees of variations. Since the goal was merely to study the performance of IEEE802.11 based MWNs, the networks provided data rate up to 11 Mbps using DSSS with 2.4 GHz range for the reference scenario. To see how the real-time demand of the applications would affect the network performance, six combinations of commonly-used applications (Data Access, Email, File Transfer, File Print, Web Browsing, Voice over IP Call and Video Conferencing) were imported into the initial model. On the other hand, the variations of the mobiles number and system parameters were taken into consideration in the model.

Thirdly, the study was led to a scenario-based simulation, implemented using the methodology of Discrete Event Simulation (DES), which managed a system as a chronological sequence of event. Limited by the topic, this thesis would not take deep investigation in developing simulator based on this method, but directly chose an appropriate simulation tool OPNET. The basics of OPNET is presented later in Chapter 4.

The fourth stage of the study was to make comparisons and analysis of the results gathered from the simulations in the previous stage. The global performance of the scenarios in terms of throughput, average delay and dropped data was analyzed in groups. From the comparison of various scenarios, we were able to know that how the number and distribution of users, real-time applications and EDCA mechanism affect the global capacity of a MWN.

1.4 Thesis Outline

The first chapter introduces IEEE802.11 standard and multihop wireless network architecture. The study objectives and approach are also discussed in this chapter.

Chapter 2 provides an in-depth discussion on the current state of Multihop Wireless Networks. The literatures on the characteristics, open issues and research trends are reviewed in this field. Related works on performance evaluation of MWNs are then presented at the end of this chapter.

In Chapter 3, wireless technology based on IEEE 802.11 standards is introduced, following by a description and classification of services. Furthermore, the physical layer specifications and medium access control relevant to the study are discussed.

Chapter 4 starts from a comparison of simulators in communication networks. The remaining of the chapter is dedicated to the description of the used simulation tool, OPNET Modeler, and to the detailed presentation of the simulation model and its implementation.

In Chapter 5, the simulation scenarios are described based on the reference scenario. The variations of user number, profiles and EDCA parameters are introduced in detail.

Chapter 6 presents the analysis of results obtained from the OPNET modeler simulations. The global throughput, delay and dropped data gathered from 168 scenario simulations are compared respectively.

Finally, Chapter 7 presents the conclusion and discussion on the potential directions for future study in the area of the topic.

CHAPTER 2

REVIEW OF LITERATURE ON MULTIHOP WIRELESS NETWORKS

This chapter provides a literature review of previous research and industry work undertaken in the field of multihop wireless networks, following by an in-depth discussion on open issues, research trends and outcomes.

2.1 Introduction to Multihop Wireless Networks

As various wireless networks evolve into a new development era to provide low-cost, ubiquitous broadband Internet access, Multihop Wireless Networking (MWN) has emerged as a promising technology. In MWNs, nodes are comprised of mesh routers and mesh clients (Akyildiz, Wang, & Wang, 2005), which operate not only as hosts but also as routers, forwarding packets on behalf of other nodes that may not be within direct wireless transmission range of their destinations.

Compared to the connection between mobile clients and access points in a traditional wireless LAN, a MWN consists of access routers communicating with each other wirelessly, potentially over multiple hops, where a small fraction of those access router are wired to the Internet and serve as Internet gateways for the rest of the network (Marina & Das, 2005).

The architecture of a typical MWN is similar to the example illustrated in Figure 1.1. The network is formed by a collection of wireless access points (routers, Internet gateways, etc.) and terminal users (static or mobile stations). It depends on some form of wireless communication between the nodes, which are in accordance with the

IEEE802.11 standard in the present study. As mentioned before, other technologies like Zigbee, ultra-wideband (UWB) and WiMAX have also been utilized to form WMNs respectively or hybridizably (Wang, 2008). Because MWNs based on WiFi technology have been the most commercially successful to date, the most common sense reflects that multihop wireless mesh networks are implemented over a wireless LAN based on IEEE802.11 standards.

The primary objective to develop MWNs is to extend the coverage range of current wireless networks without sacrificing the channel capacity. On the other hand, providing non-line-of-sight (NLOS) connectivity among the users without direct line-of-sight (LOS) links is another purpose of building MWNs. Due to these features, the characteristics of MWNs are summarized as below:

- Multi-hopping

  Compared to other wireless networks, multihop wireless links collectively exhibit graceful drop behavior: as the offered load increases, the link contention drop probability also increases, but saturates eventually (Fu, Zerfos, Luo, Lu, Zhang, & Cerla, 2003).

  In 2004, Qiu, et al. offered an impelling support for multihop approach by optimizing the placement of access points. They developed algorithms to make informed placement decisions based on neighborhood layouts, user demands, and wireless link characteristics (Qiu, Chandra, Jain, & Mahdian, 2004).

  Multi-hopping is indispensable for MWNs, as it can highly increase throughput without sacrificing effective radio range via shorter link distances (Fu, Zerfos, Luo, Lu, Zhang, & Cerla, 2003), less interference between the nodes (Jein, Padhye, Padmanabhan, & Qiu, 2005), and more efficient frequency reuse (Ramchandran, Belding, Almeroth, & Buddhikot, 2006).

- Support for ad hoc networking, and capability of self-forming, self-healing, and self-organization

  MWNs can efficiently enhance network performance, largely because of flexible network architecture, easy deployment and configuration (Marina & Das, 2005), fault tolerance (Qiu, Chandra, Jain, & Mahdian, 2004), and mesh connectivity (Garetto, Shi, & Knightly, 2005). Also due to the features above, MWNs have low upfront investment requirement, and the network can grow gradually as demanded (Wang, 2008).

- Mobility

  Access points (i.e. routers) usually have minimal mobility, while terminal users can be static, nomadic or mobile nodes. This offers another option for the integration of data communication networks (i.e. Internet) and telecommunication system (i.e. CDMA).

- Dependence of power-consumption constraints on the type of nodes

  Because of the mobility, access points usually do not have strict constraints on power consumption. In contrast, users may require power efficient protocols. Sensor network designers need the ability to obtain accurate and dependable power consumption figures to tune their applications before deployment in real environments. Apart from aggregate power consumption over time, the pattern of power load is important to consider, as this affects the ability of the power source to deliver adequate energy over time (Shnayder, Hempstead, Chen, Allen, & Welsh, 2004).

- Compatibility and interoperability with existing wireless networks

  As MWNs are developed from existing wireless technologies, which are widely

deployed in cellular networks and data communications, MWNs inevitably need to be inter-operable with other wireless networks such as UMTS, ZigBee, etc.

2.2 Open Issues and Research Trends

Despite the fact that there is a vast range (or flavors) of multihop wireless networks that have been studied, there exist many open issues. This section attempts to summarize some of these issues, which may be considered as potential research themes in the further study.

2.2.1 Protocols for Network Management

Many management functions are needed to maintain the appropriate operation of MWNs (Akyildiz, Wang, & Wang, 2005).

- Mobility management

  There are two components in mobility management: location management and handoff (also referred to as handover in the literature) management (Yu, Wong, Song, Leung, & Chan, 2010). In MWNs, location management enables the network to deliver massages to mobile users by tracking their locations between connections, while handoff management maintains service continuity by enabling a mobile terminal to keep connected when it moves form one access point region to another.

  Although the mobility management schemes developed for other networks including cellular, ad hoc and mobile IP networks could be useful for MWNs, these schemes cannot perform well for MWNs due to the specific features of MWNs. These schemes assume that the fixed networking paths are stable and the hierarchies can be deterministically created following the network layout at deployment. In such networks, MAPs can be placed at the root nodes in the tree

hierarchy and sub-domains for the MAPs are formed with the children under the root nodes. These assumptions do not hold true for WMNs which usually have an unplanned graph topology and where the wireless links are unstable and exhibit varying latencies. In 2008, Wu et al. presented that Hierarchical mobility management scheme HMIP is applicable on MWNs and is able to give a close to optimal packet loss rate and handoff latency.

Location service is a desired feature in MWNs. A novel mobility management mechanism Wireless Mesh Mobility Management (WMM) for MWNs was proposed in 2008 (Huang, Lin, & Gan, 2008). WMM adopts the location cache approach, where the mobile nodes cache the IP address of user's serving MAP (known as user's location information) while routing the data for the user.

Handoff management is closely related to multiple layers of network protocols. The development of multi-layer handoff management schemes as shown in (Yu, Wong, Song, Leung, & Chan, 2010) is a challenging topic.

- Power management

  Usually MAPs do not have a constraint on power consumption as individual nodes. However, power management appears to be significant in the scope of a MWN, as power management schemes provide every level of energy savings at each layer of network protocol stack (Klues, 2006).

  The various techniques are used to conserve energy form the application layer all the way down to the physical layer. At the application layer, it is worth to exploring techniques that specifically deal with reducing the power consumed while running the common applications such as database operations and video processing (Jones, Sivalingam, Agraval, & Chen, 2001). The TCP-Probing

(Tsaoussidis & Badr, 2000) and Wave and Wait Protocols (Zhang & Tsaoussidis, 2001) have been developed to guarantee end-to-end data delivery with high throughput and low power consumption. The techniques existing at the network layer are concerned with performing power efficient routing through a multihop network either backbone based, topology control based, or a hybrid of both (Karl, 2003). The schemes at data link layer are used to reduce the number of packet errors at a receiving node. In most cases, lots of power is wasted listening on the radio channel while there is nothing there to receive, and the MAC layer power saving is deployed to avoid it. Generally, proper hardware design techniques at physical layer allow one to decrease the level of parasitic leak currents in devices (Jacome & Catthoor, 2003).

Ongoing research has been providing innovative solutions including power aware routing, sleep scheduling, energy harvesting, etc. Although the protocols are similar in terms of principle, the power savings achieved using each of them varies from system to system and application to application. One technique is not better than the other in this sense, so efforts are being made to define exactly when each type should be used.

- Network monitoring

  From a network management perspective, monitoring offers several benefits to the robust operation of wireless networks, including (i) providing statistics to pinpoint the sources of network failure, (ii) verifying if the strategies adopted by protocols, applications or middleware perform well, and (iii) locating potential bottlenecks so as to redimension the network (Sailhan, et al., 2007).

  The monitoring of MWNs, however is challenging. The rapid pace of development of wireless technologies brings a large amount of proprietary

solutions with little standardization. Consequently, operating and maintaining a set of such devices becomes cumbersome and scales poorly with increasing network size. To deal with the above limitations, a scalable framework VISUM was presented (Ho, Ramachandran, Almeroth, & Belding-Royer, 2004). Using a modular architecture, VISUM collects MAC layer information from wireless network infrastructure components to provide real time views of network status. This implementation developed in JAVA has been successfully tested in a real-time environment for a month.

Due to the inherent uncertainty in the wireless medium, network administrators require a comprehensive set of data and metrics to deal with them. Besides, the method of storing accumulated information must be considered carefully, allowing efficient data retrieval. To this end, a multi-tiered monitoring framework MeshMon was proposed (Raghavendra, Acharya, Beling, & Almeroth, 2009). MeshMon dynamically controls the granularity of data collection based on observed events in the network, thereby achieving significant bandwidth savings and enabling real-time automated management. The evaluation of MeshMon on a real testbed shows that it can diagnose a majority (87%) of network faults with a 66% savings in bandwidth required for network monitoring.

A number of monitoring protocols have been designed and implemented, however, the extensibility of these schemes needs to be achieved for the purposes of collecting information form newly developed devices. In addition, how to quickly determine network topology is also an open issue.

2.2.2 Security

Due to the unique characteristics of MWNs, they are highly vulnerable to security attacks compared to wired networks. Designing a robust security mechanism for

MWNs is a challenging task. The security can be provided in various layers of protocol stack. Current security approaches may be effective against a particular attack in a specific protocol layer, but they lack a comprehensive mechanism to prevent or counter attacks in different protocol layers. The following issues pose difficulty in providing security in MWNs.

- Shared radio channel

  Because the wireless links between the nodes in MWNs are broadcast in nature, a malicious node could easily obtain data being transmitted if it is placed in the transmission range of MAPs. An attacker may paralyze nodes in its neighborhood by sending CTS frames periodically, setting the "Duration" field of each frame to at least the interval between such frames (Hu & Rerrig, 2005).

- Lack of association

  In MWNs, the MAPs form a fixed mesh topology for the mobile users. Hence, the clients can join and leave the network at any time via MAPs. If no proper authentication mechanism is provided for association of nodes with MWNs, an intruder would be able to join the network quite easily and carry out attacks may sneak into the network by misusing cryptographic primitives (Borisov, Goldberg, & Wagner, 2001).

- Limited resource availability

  Normally, the mesh clients are limited in resources such as bandwidth, battery power, computational power, etc. Therefore, it is difficult to implement complex cryptography-based mechanisms at the client nodes. As MAPs are resource-rich in terms of battery power and computational power, security mechanisms can be implemented at MAPs. Due to wireless connectivity between MAPs, they also have bandwidth constraints. Consequently, the communication overhead incurred by the security mechanism should be minimal (Zhang, Zheng, & Hu, 2008).

The key management is one of the most important tasks for network security (Akyildiz, Wang, & Wang, 2005). However, the key management for MWNs becomes much more difficult, because there is no central authority, trusted third party or server to manage security keys. To enhance security, two strategies also need to be adopted: either to embed security mechanism into network protocols such as secure routing and MAC protocols; or to develop security monitoring and response systems to detect attacks, monitor service disruption and respond quickly to attacks. However security attacks in a network may come simultaneously from different protocol layers. Thus, a multi-protocol layer security scheme is desired for network protocols. How to design and implement a practical security monitoring system, including cross-layer secure network protocols and various intrusion detection algorithms, is a challenging research topic.

2.2.3 Cross-Layer Design

There are many studies that discuss performance improvement of MWNs in recent years from the single-layer point of view. Because of the direct coupling among different layer, the traditional layered design is not sufficient for MWNs (Zhang & Zhang, 2008). All the controls in different layers potentially have mutual impact, and it is necessary to consider all the controls across different layers jointly to optimize the overall performance. Such interactions demand a cross-layer design among these layers. Figure 2.1 illustrates the cross-layer framework and the potential interaction among layers. In this design, interdependencies between layers are characterized and exploited by adapting to information exchanged between layers and building the appropriate amount of robustness into each layer. For instance, scheduling and channel management in the MAC layer can avoid links experiencing deep fades and resending messages.

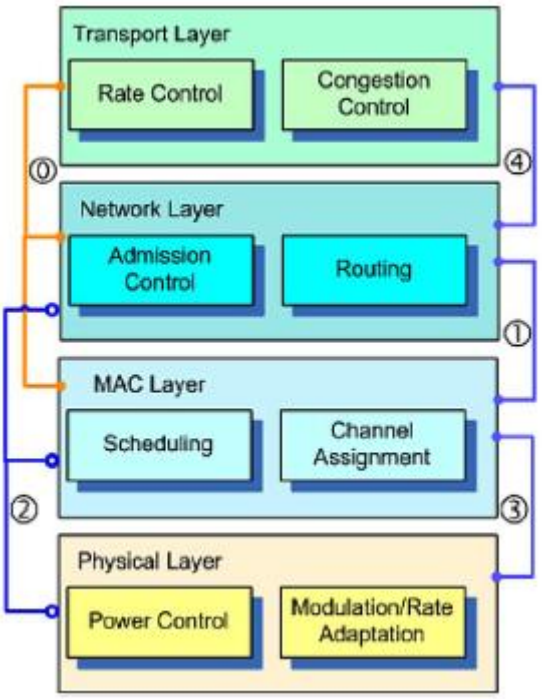


Figure 2.1 Cross-Layer Framework and interaction among layers (Zhang & Zhang, 2008)

Although it is reasonable to believe that cross-layer optimization will continue to be one of the most important tasks in protocol design for MWNs, critical issues must be considered for cross-layer design. The method of optimization decomposition, which was proposed in (Akyildiz & Wang, Cross-Layer Design in Wireless Mesh Networks, 2008), took fully consideration of risks due to loss of protocol-layer abstraction, incompatibility with existing protocols, unforeseen impact on the future design of the network, and difficulty in maintenance and management.

Unlike building around the existing routing concept which focuses on cross-layer design that considers the corresponding network utility maximization problem, network coding allows intermediate nodes to perform coding operations in additions to pure packet forwarding. In (Khreishah, Wang, & Shhroff, 2009), they proposed a jointly optimal coding, scheduling, and rate-control scheme based on pairwise intersession network coding (PINC). Numerical experiments had proven that in a MWN, the throughput advantage of PINC could be achieved without sacrificing the

stability conditions. Moreover, PINC had minimal impact on the optimal rate-control and scheduling as the only new component necessary for scheduling PINC traffic is the balance update performed at the receivers.

To date, cross-layer design of MWNs lacks standard on framework and methodology. To further improve the viability of cross-layer design schemes, standardizing the framework is necessary.

2.3 Performance Evaluations

Considering the flexibility of a MWN, the designers of the networks have to take much attention on network performance trade-offs. Despite of the impact of various factors of a MWN, a lot of works have been done in the last decade on performance evaluations, such as throughput, transmission delay, fairness, etc.

Quite a few performance evaluations have taken place on testbeds or straightly gone with a project implementation. Dated back to December 2006, Technology for All (TFA) and Rice University have partnered to develop and test a MWN for low-income communities in Houston. In this project, a hardware platform EPIA VE10000 Mini-ITX Board has been used in the evaluation. On systems running Linux kernel with LocustWorld open-source networking software, the platform met expectation of 1 Mb/sec for commercial connection (Camp, Knightly, & Reed, 2006).

In contrast to network-layer statistics in previous work, Cheng, et al. (2008) evaluated the system performance of application layer video quality in a MWN testbed and provided insights in potential problems and solutions for supporting video streaming. The paper analyzed impacts on video performance, and found that impact of interference was the major source of quality degradation. In the experiment, the tradeoff between video coding rate and the achieved video quality has been studied,

and the conclusion was again that interference in mutihop scenatios significantly worsened the video performance (Cheng, Mohapatra, Lee, & Banerjee, 2008). For evaluating impact of the rate adaptation algorithms on MWNs performance, indoor and outdoor testbeds have been built separately (Ancillotti, Bruno, & Conti, 2008).

Hamidian, et al. (2009) created a testbed for MWNs using the Mesh Connectivity Layer (MCL), which is part of the Microsoft Mesh Toolkit. Three groups of outcomes from experiments on scenarios with FTP/TCP session, VoIP streams and video streams were presented (Hamidian, Palazzi, Chong, Nanarro, Korner, & Gerla, 2009). In addition, another major finding from this paper was that the testbed could be used to evaluate complex and realistic scenarios with more complex interactive applications, such as online games, cooperative multimedia management, and augmented reality services.

In the same year of 2009, FloorNet, a 802.11-based MWN testbed, was presented by Serrano, et al (Serrano, Bernardos, Oliva, Banchs, Soto, & Zink, 2009). The unique characteristics of the false floor constituted strong support for the deployment of MWN testbeds. But it was noticed that nonideal behavior of off-the-shelf hardware existed in i) the impact of the entity generating traffic in the measurements, and ii) the strong interference between non-overlapping channels.

Most of the performance evaluations on MWNs have been based on simulation. As a free open-source simulator, NS-2 became one of the most popular tools in networks performance evaluation. In (Fu, Zerfos, Luo, Lu, Zhang, & Cerla, 2003), the analysis based on NS-2 simulations showed the relationship between throughput and specific topology and flow settings. Using a conflict graph, Jain, et al. (2005) modeled wireless interference in MWNs to evaluate the impact of it on optimal throughput with the assistance of NS-2. Nevertheless, the assumption of packet transmissions in

their simulations is unrealistic (Jein, Padhye, Padmanabhan, & Qiu, 2005). We will present the drawbacks of using NS-2 on performance evaluation in the next chapter.

On the other hand, the performance evaluation of WMNs has been developing in various directions. Some researches focused on different communication protocols rather than IEEE802.11, such as (Han, Jia, & Lin, 2006). But Gambiroza et, al. had objective of studying fairness in MWNs. To implement their target, they performed an extensive set of simulation experiments and then developed and studied a distributed MAC layer fairness algorithm which aimed to achieve the fairness of the reference model without modification to TCP (Gambiroza, Sadeghi, & Knightly, 2004).

Apart from the simulation environment used as above, Ernst and Denko proposed the other two approaches “Performance Metrics and Simulation Parameters” and “Analysis of the Experimental Results” in their paper, which focused on fair scheduling algorithms for MWNs (Ernst & Denko, 2010).

CHAPTER 3

AN OVERVIEW OF WIRELESS NETWORK STANDARDS

Chapter 3 provides a technical overview of IEEE802.11 based wireless networks, particularly of physical layer protocols and medium access control protocols considered appropriate for the present study for identification of parameters and network characteristics.

3.1 Introduction

Although the multihop mesh architecture could be implemented with all kinds of wireless access technologies, like Zigbee, UWB and WiMAX, only IEEE802.11 standards are deployed in the present study, for the reason mentioned in the objective (see Chapter 1) and the literature review (see Chapter 2).

Wireless networks based on IEEE802.11 standards have fundamental characteristics that make them significantly different from traditional wired LANs. The PHYs used in IEEE802.11 are fundamentally different from wired media. Thus IEEE 802.11 PHYs use a medium that has neither absolute nor readily observable boundaries outside of which STAs with conformant PHY transceivers are known to be unable to receive network frames. Although they are unprotected from other signals that may be sharing the medium which is significantly less reliable than wired PHYs, wireless PHYs have time-varying and asymmetric propagation properties and dynamic topologies (IEEE LAN/MAN Standards Committee, 2007).

3.1.1 IEEE802.11 LAN Topology

The concept of service set is the basis of the different types of wireless LAN topologies. A service set is a grouping of devices that access the network by broadcasting a signal across a wireless radio frequency (RF) carrier. The basic service set (BSS) is the basic building block of an IEEE 802.11 LAN. Figure 3.1 shows two BSSs, each of which has two stations (STAs) that are members of the BSS.

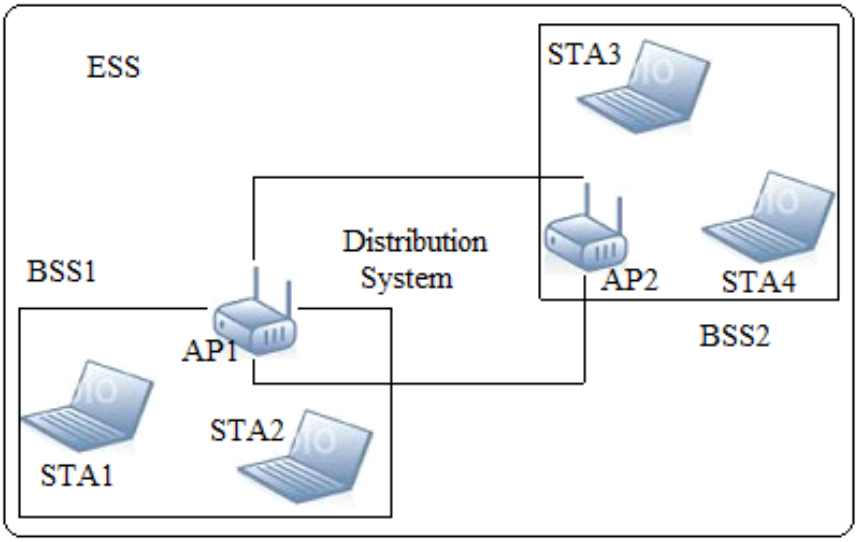


Figure 3.1 IEEE802.11 LAN topology

In a BSS, the service set consists of two entities: the STA and the AP. There can be several stations that communicate with one another via the AP, which acts as a relay station. An AP can also function as a bridge to the outside world, providing a connection to some kind of backbone Distribution System (DS).

The independent BSS (IBSS) is the most basic type of IEEE 802.11 LAN. A minimum IEEE 802.11 LAN may consist of only two STAs. The role of an AP does not exist in an IBSS. Stations communicate directly with one another without the use of an intermediate. This self-contained network is a simple peer-to-peer WLAN,

which is also referred to as an ad-hoc network. Typically, an IBSS is small and only lasts enough time until the communication being performed is completed.

The extended service set (ESS) is the most generic topology for a WLAN, consisting of two or more BSSs that are interconnected by a DS. Figure 3.1 shows a simple representation of an ESS, where it is possible to identify a collection of BSSs, grouped via a DS. In this case, if STA4, located in BSS2, wants to send a frame to STA1, it has to send it first to AP2, which acts as the AP of BSS2, being responsible for forwarding the frame to AP1 (the AP of BSS1). Finally, AP1 is able to deliver the frame to its final destination. It is important to note that this process is performed at the MAC level, thus, the ESS appears as a single logical unit to the Logic Link Control (LLC) one. This way, the frame that is exchanged according to the described process between MAC users is known as the MAC Service Data Unit (MSDU). Moreover, the MSDU delivery from the MAC to the upper layer constitutes the basic service of an IEEE802.11 LAN.

3.1.2 IEEE802.11 Station Services

From the above simple description of a frame traversing an ESS, the need of the IEEE802.11 standard to define a set of complementary services for the basic MSDU delivery becomes evident, which are listed in Table 3.1. The provider column indicates who is responsible for the service. Station services are provided among all stations, therefore, being implemented in every 802.11 station, including APs. Distribution services are available among BSSs, by the DS, being implemented only in APs or in another special-purpose device attached to the DS. The first five services are used to support MSDU delivery, which is discussed in section 3.3, while the last three are used to control IEEE 802.11 LAN access and confidentiality.

Table 3.1 IEEE802.11 Services (IEEE LAN/MAN Standards Committee, 1999)

Service	Provider
Distribution	Distribution system
Integration	Distribution system
Association	Distribution system
Reassociation	Distribution system
Disassociation	Distribution system
Authentication	Station
Deauthentication	Station
Privacy	Station

Distribution is the primary service used by stations to send MAC frames to another station located in a different BSS within the same ESS. In the example of Figure 3.1, AP2 uses the distribution service in order to send a frame to STA1. In the case of stations exchanging a frame that are located in the same BSS, the distribution service goes through the single AP of that BSS. The other service that is responsible for the distribution of messages within a DS is integration, which enables transfer of frames between a station on an IEEE 802.11 LAN and another on an IEEE 802.x LAN that is physically connected to the DS.

For a correct operation of the services that are responsible for the transfer of MSDUs among MAC users, some kind of information about the location of the various stations within an ESS is necessary. This requirement is fulfilled by the association, reassociation and disassociation services. The association service establishes an initial association between a station and an AP, by which the AP is able to register the identity and address of the station. The AP can then communicate this information to other APs within the ESS, to facilitate routing and delivery of frames. Association is usually preceded by a probe process that is used by a station to select the most

adequate AP to associate with. Concerning the mobility of stations, when a station moves from a BSS to another, the established association must be transferred to another AP using the reassociation service. The end of an existing association, because a station is either leaving the ESS or shutting down, must be notified using the disassociation service.

In order to provide a minimum level of security, three services are provided: authentication, deauthentication and privacy. Before the association process is accomplished, the station that wants to communicate with another one needs to prove its identity using the authentication service. The standard does not mandate any particular authentication scheme, which can range from a simple handshaking to a public key encryption scheme. The reverse process, when an existing authentication is to be terminated, is performed by the deauthentication service. Another security mechanism, used to prevent messages from being read by a casual eavesdropper, is the privacy service. This service consists of an optional encryption mechanism that takes the content of a data frame and passes it through an encryption algorithm, in both the sending and the receiving stations.

3.2 IEEE802.11 Physical Layer

In an IEEE802.11 LAN, an underlying physical layer is defined to support all the functions of the MAC layer. Besides this primary function, it is also responsible for other secondary ones, such as assessing the state of the wireless medium and reporting it to the MAC.

3.2.1 The Various Physical Layers

The basic function of the 802.11 PHY layer is to provide wireless transmission mechanisms for the MAC layer. As described previous, the PHY layer comprises two sublayers: the PLCP and the PMD. While the former is responsible for mapping

MPDUs frames, coming from the upper MAC layer, onto an appropriate frame format, the latter provides adequate methods for transmitting and receiving user data through a wireless medium. The PLCP sublayer can also be seen as an interface between MAC and PMD, defining a set of primitives that enable communication between the two adjacent layers. These primitives provide the interface for transfer of data between the MAC and the PMD. Moreover, on transmission, there are primitives that enable the MAC to tell PMD when to initiate transmission. On reception, PLCP primitives indicate the start of an incoming transmission from another station to the MAC.

The IEEE 802.11 original standard has defined the MAC layer and three PHY layer specifications, which are based on the following methods:

- Infrared at 1 Mbps and 2 Mbps, operating at wavelength between 850~950 nm.
- Frequency Hopping Spread Spectrum (FHSS) also operating in the 2.4 GHz ISM band, at the same data rates. This technique uses 1 MHz channels and splits the available bandwidth into 79 non-overlapping channels.
- Direct Sequence Spread Spectrum (DSSS) operating in the 2.4 GHz Industrial, Scientific and Medical (ISM) band, at data rates of 1 Mbps and 2 Mbps. DSSS WLANs use 22 MHz channels that allow three non-overlapping channels in the 2.4 to 2.483 GHz range.

To overcome some limitations of the original PHY layer, several PHY specifications have been standardized: IEEE 802.11a, 802.11b and 802.11g.

3.2.2 IEEE802.11a

IEEE 802.11a operates in the 5 GHz frequency band (IEEE LAN/MAN Standards Committee, 1999). It defines a set of 20 MHz channels within the Universal

Networking Information Infrastructure (UNII) band, which is divided into three parts: the UNII-1 band (5.15 to 5.25 GHz), intended for indoor use; the UNII-2 (5.25 to 5.35 GHz), to be used for either indoor or outdoor; and the UNII-3 (5.725 to 5.825 GHz), exclusively for outdoor use. The standard provides mandatory data rates up to 24 Mbps (6, 9, 12, 18 and 24 Mbps) and some optional rates up to 54 Mbps (36, 48 and 54 Mbps). 802.11a does not use a spread spectrum scheme; instead, it uses Orthogonal Frequency Division Multiplexing (OFDM), which uses multiple subcarriers for sending bits on each one.

The PLCP sublayer provides all the framing and signalling needed for PMD operations. It takes the MPDUs coming from the upper MAC layer and adds some additional information, forming a PLCP PDU (PPDU). Figure 3.2 illustrates the PPDU frame format.

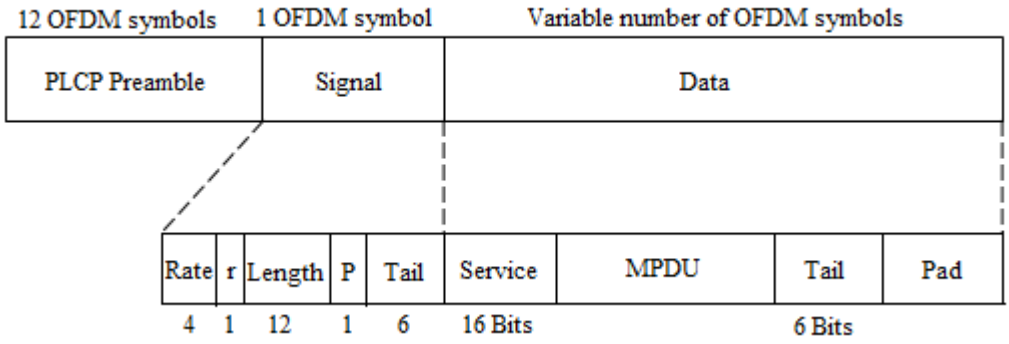


Figure 3.2 IEEE802.11a PPDU

The PLCP preamble enables the receiver to acquire an incoming signal and to synchronize the demodulator. The preamble and signal fields are transmitted at the minimum data rate, i.e., 6 Mbps. Subsequently, follows the transmission of the data field at the rate specified in the rate subfield. Prior to transmission, the data field has to pass through a scrambling process. Besides the MPDU, the data field consists of the following subfields:

- Service: used to synchronize the scrambler at the receiver.
- Tail: used for encoding purposes.
- Pad: used to provide the remaining bits to make the data field a multiple of the number of bits in an OFDM symbol.

Together with OFDM, the PMD sublayer supports several modulation and coding alternatives. Each carrier is divided in up to 48 subcarriers that are modulated using BPSK, QPSK, 16-QAM or 64-QAM. Table 3.2 shows the correspondence between available data rates and used modulation.

Table 3.2 IEEE802.11a Data Rates (IEEE LAN/MAN Standards Committee, 1999)

Data Rate (Mbps)	Modulation
6	BPSK
9	BPSK
12	QPSK
18	QPSK
24	16-QAM
36	16-QAM
48	64-QAM
54	64-QAM

3.2.3 IEEE802.11b

IEEE 802.11b is an extension of the original DSSS scheme, providing additional data rates of 5.5 and 11 Mbps in the same ISM band (IEEE LAN/MAN Standards Committee, 1999). Channel bandwidth is also 22 MHz, providing the same 3 non-overlapping channels. A chipping code, or pseudonoise sequence, is the basis of DSSS, which is used to spread the data rate of the signal. The original 802.11 DSSS

uses an 11-chip Barker sequence.

The preamble field represented in Figure 3.3 has two subfields: the sync one, to synchronise the demodulator, and the Start-of-Frame Delimiter (SFD). The header follows the preamble, consisting of the following subfields:

- Signal: provides the data rate at which the MPDU field is transmitted.
- Service: indicates which encoder is used, among other functions.
- Length: indicates the number of microseconds that are necessary to transmit the MPDU field.
- CRC: error detection code.

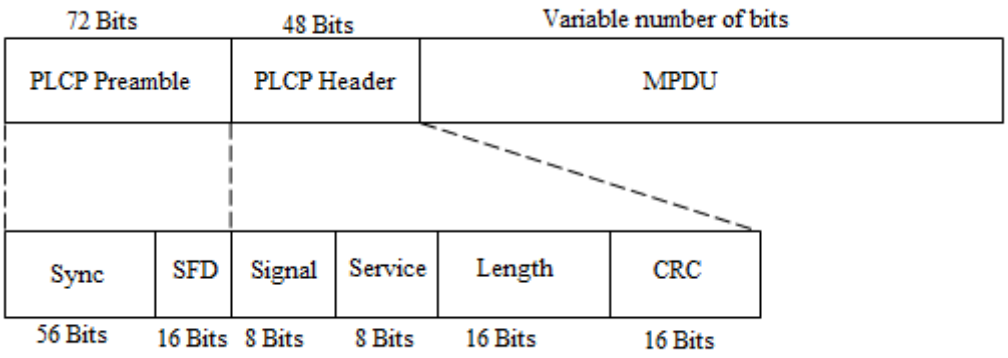


Figure 3.3 IEEE802.11b PPDU (IEEE LAN/MAN Standards Committee, 1999)

To achieve the lower data rates, the PMD 802.11b sublayer employs the same techniques as the original standard, which is DSSS using an 11-chip Barker sequence with DBPSK modulation, for 1 Mbps, and with DQPSK modulation for 2 Mbps. To achieve the higher data rates with the same chipping rate and using the same bandwidth, a more complex modulation scheme is needed. The mandatory scheme is a Complementary Code Keying (CCK) one, which takes 8 bits at a 1.375 MHz rate. Six of these bits are mapped onto one of 64 codes sequences. The output of the mapping and the remaining two bits are applied to the input of a DQPSK modulator. Additionally to CCK, the standard also provides an optional modulation scheme,

named Packet Binary Convolutional Coding (PBCC), which provides a more efficient transmission at the cost of increased computation requirements.

3.2.4 IEEE802.11g

The IEEE 802.11g standard introduces an Extended Rate Physical (ERP) layer to support higher data rates in the 2.4 GHz ISM band (IEEE LAN/MAN Standards Committee, 2003). The standard guarantees interoperability with the older 802.11b system, providing the same modulation and framing schemes for the lower data rates of 1, 2, 5.5 and 11 Mbps. Additionally, the 802.11g also provides a wide range of data rates: 6, 12 and 24 Mbps that are mandatory, and 9, 18, 36, 48 and 54 Mbps that are optional. To support this additional data rates, a new modulation scheme is defined, referred to as ERP-OFDM. It is also possible to optionally use a DSSS-OFDM scheme to support the same data rates and an ERP-PBCC scheme to support 22 and 33 Mbps. Table 3.3 summarizes some of the options in terms of data rates and modulation schemes.

Table 3.3 IEEE802.11g Options (IEEE LAN/MAN Standards Committee, 2003)

Data Rate (Mbps)	Modulation Scheme	Data Rate (Mbps)	Modulation Scheme
1	DSSS	18	ERP-OFDM
2	DSSS	22	ERP-PBCC
5.5	CCK or PBCC	24	ERP-OFDM
6	ERP-OFDM	33	ERP-PBCC
9	ERP-OFDM	36	ERP-OFDM
11	CCK or PBCC	48	ERP-OFDM
12	ERP-OFDM	54	ERP-OFDM

Five PPDU formats are provided, differing in the way the preamble is defined; three are mandatory preambles (a long, a short, and an ERP-OFDM one), the remaining two, which are optional, being the long and short DSSS-OFDM preambles.

3.3 IEEE802.11 Medium Access Control

3.3.1 MAC Data Services

MAC data services provide peer LLC entities with the ability to exchange MSDUs. To support this service, the local MAC uses the underlying PHY-level services to transport an MSDU to a peer MAC entity, where it will be delivered to the peer LLC. This frame exchange between MAC entities requires a mechanism to access the common medium in a WLAN.

The IEEE 802.11 MAC defines two transmission modes for data packets: the Distributed Coordination Function (DCF) based on CSMA/CA and the contention-free Point Coordination Function (PCF) where the Access Point controls all transmissions based on a polling mechanism. The DCF and PCF modes are time multiplexed in a superframe, which is formed by a PCF contention-free period (CFP) followed by a DCF contention period (CP), positioned at regular intervals, as shown in Figure 3.4.

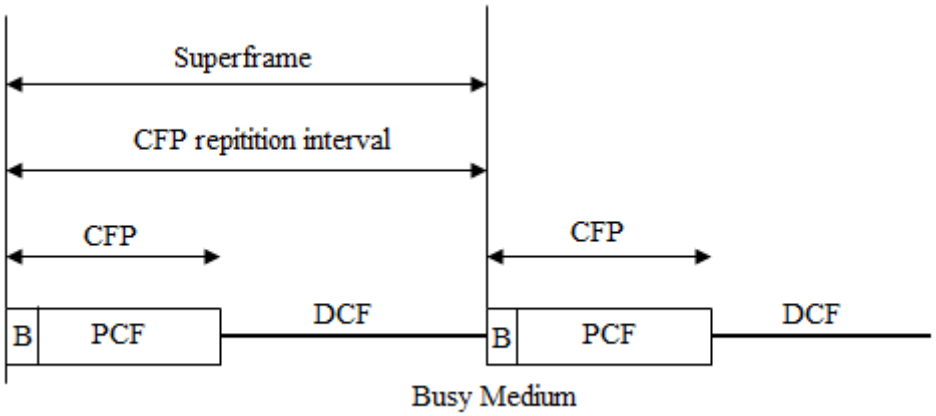


Figure 3.4 Beacons and Contention Free Periods

The AP transmits beacon frames periodically in order to deliver management information to terminals. The boundaries between CFPs and CPs are marked by

beacons carrying the Delivery Traffic Indication Message (DTIM). Terminals can use the information present in the beacons in order to associate with the AP, which is performed during the CP. This association is mandatory if the terminal needs to have its transmissions scheduled by the PCF, which is usually required for QoS sensitive data.

Packet priorities are implemented defining three different length Interframe Spaces (IFSs):

- SIFS (Short IFS): is the shortest IFS, used for all immediate response actions, as the ASK transmission.
- PIFS (Point coordination function IFS): is a middle-length IFS, after whose interval expires, any PCF mode frames can be transmitted.
- DIFS (Distributed coordination function IFS): is the longest IFS, used in the DCF operation as a minimum delay for frames contending the medium according to the CSMA back off mechanism.

The DCF mode is based on a CSMA/CA mechanism. The access control scheme is shown in Figure 3.5.

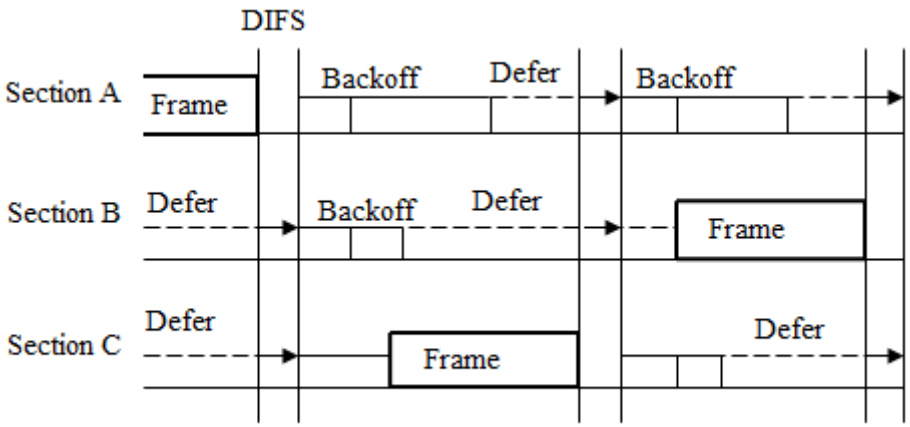


Figure 3.5 Backoff Mechanism in DCF

A terminal that intends to transmit and senses the channel busy waits for the end of the ongoing transmission, then waits for a time period of DIFS length, and then randomly selects a time slot within the backoff window. The backoff length is calculated as follows:

$$Backoff\_length = Random(0, CW) \times aSlotTime \quad (1)$$

The slot duration, $aSlotTime$, depends on the network round-trip propagation delay. The number of backoff slots is derived from a uniform distribution over the interval $(0, \mathrm{CW})$, where the contention window (CW) parameter ranges from a minimum value of aCWmin up to a maximum value aCWmax. Initially, the CW parameter is set to aCWmin and can be increased up to 255.

If no other terminal starts transmitting before the intended slot is reached, the transmission of a fragment with maximum size of a Fragmentation Threshold is started. Collisions can only occur in the case that two terminals have selected the same slot. For each unsuccessful transmission the contention window is updated as follows:

$$\mathrm{CW} = 2^{2+\mathrm{i}} - 1 \quad (2)$$

Where i is the number of transmission attempts.

If another terminal has selected an earlier slot, transmission is deferred and its backoff counter is frozen. Then, the terminal waits for the channel to become idle and then waits for the backoff slots remaining from the previous competition. After the successful transmission of the first fragment of a MSDU, the remaining fragments are transmitted sequentially separated by a SIFS interval. Transmission ends when all fragments of the MSDU are transmitted or the maximum dwell time (aMediumOccupancyLimit) expires.

In order to guarantee undisturbed transmission even if hidden terminals are present,

an RTS/CTS mechanism is used. When this mechanism is applied, the contention winner does not transmit the data immediately. Instead it sends an RTS frame to which the receiver answers with a CTS frame. This guarantees, that all terminals in the range of both the sender and the receiver, know that a packet will be transmitted, remaining silent during the entire transmission. Only then the sender transmits the data frames. While the two extra messages present additional overhead, the mechanism is particularly useful in the case of large data frames because the RTS and CTS frames are short.

The PCF mode is based on a polling mechanism controlled by the AP as depicted in Figure 3.6.

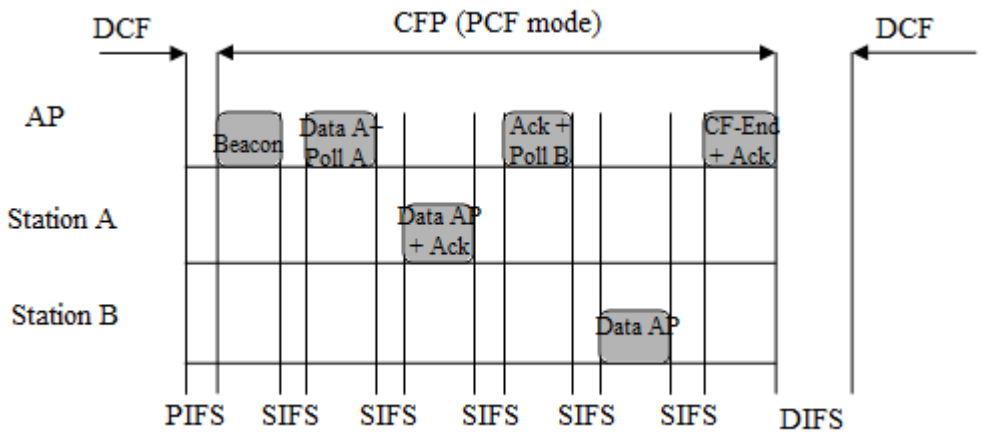


Figure 3.6 Polling Mechanism in PCF

During the CFP, the AP polls the terminals registered in its polling list and allows them undisturbed contention-free access to the medium. As already said, in order to become registered in the polling list, the terminals have to associate with the AP during the CP. The maximum duration of a CFP is given by the 802.11 Management Information Base (MIB) variable aCFPMaxDuration, while the frequency is given by the variable aCFPRate.

During the CFP a frame can be a composite of control and data information. The following combinations are allowed for PCF frames: DATA, CF-ACK, CF-POLL, DATA+CFACK, DATA+CF_POLL, DATA+CF-ACK+CF-POLL and CF-ACK+CF-POLL. Only the AP has the capability to issue frames with CF-POLL. The terminals can answer with DATA, CF-ACK, DATA+CF-ACK or a NULL frame (the latter is sent when there is no data to transmit and no pending acknowledgement). Each DATA frame cannot be longer that the maximum size aFragmentationThreshold. If the terminal does not answer a polling request within an interval of PIFS, the AP concludes that an uplink frame was lost and may decide to poll the same terminal again.

DCF is the most appropriate in a context where there is no network infrastructure and users must form temporary ad hoc networks to communicate directly between each other. On the other hand, if the objective is to offer a permanent network infrastructure to provide access to the Intranet/Internet with guaranteed QoS bounds, PCF is the best choice.

It can be easily noticed that legacy DCF cannot fulfill the QoS requirements of multimedia applications such as telephony and videoconference, as it does not include prioritization mechanisms. IEEE 802.11 Task Group E elected Virtual DCF (VDCF) as the Enhanced Distributed Coordination Access (EDCA) mechanism to be incorporated in the IEEE 802.11e standard, (IEEE LAN/MAN Standards Committee, 2005). EDCF introduces a prioritization enhancement based on different Access Categories (ACs). One or more User Priorities (UPs) can be assigned to each AC. In this case, each UP within each AC has a different queue, different IFS (Arbitration IFS – AIFS) and contention window parameters. Each AC contends for medium access with only one CSMA instance using the parameters that belong to its lowest UP. This corresponds to the priority of the AC as a whole.

The AIFS length of UP *i* is set according to the following formula:

$$\mathrm{AIFS_i} = \mathrm{SIFS} + \mathrm{aAIFS_i} \times \mathrm{aSlotTime} \quad (3)$$

The default value for aAIFS$_i$ is 2 slots, which makes *AIFS$_i$* equal to DIFS time as in the legacy DCF. A terminal having several ACs maintains a separate backoff timer for each of those ACs, with each backoff timer independently counting down. The backoff of each AC *j* is chosen according to a uniform distribution over [0, *CW$_i$*], where *i* is the lowest UP of the AC and *CW$_i$* is the corresponding contention window:

$$\mathrm{Backoff_j} = \mathrm{Random}(0, \mathrm{CW_i}) \times \mathrm{aSlotTime} \quad (4)$$

CW$_i$ is an integer within the range aCWmin$_i$ and, optionally, aCWmax$_i$. Upon collision it is updated using the same generator function as for legacy DCF. When the backoff timer of an AC counts down to zero, the terminal transmits a frame from the queue with highest priority and initiates a transmission opportunity (TXOP), which is a bounded duration time interval in which the station may transmit a sequence of SIFS-separated DATA frame exchanges. Internal conflict between local ACs occurs when the corresponding backoff timers expire at the same time. In that case, the ESTA transmits a frame from the AC of highest priority, and then resets all expired backoff timers. During the TXOP, the terminal can send a burst of DATA frames separated by SIFS in the same way already explained for legacy DCF. The TXOP ends when there are no more frames to be transmitted or when the TXOP maximum duration expires. The default TXOP maximum duration is given by the MIB variable dot11DefaultCPTXOPLimit, but the TXOP limit can be modified the AP in beacons or association response frames.

Besides enhancing DCF, IEEE 802.11e also specifies a new mode of operation named Hybrid Coordination Function (HCF) (IEEE LAN/MAN Standards Committee, 2005). HCF is based on a polling mechanism similar to legacy PCF, but it allows the HC to start contention-free Controlled Access Periods (CAPs) at any time during a CP, after

the medium remains idle for at least a PIFS interval (Figure 3.7). This more flexible contention-free mechanism renders PCF useless, although IEEE 802.11e terminals are still allowed to support PCF.

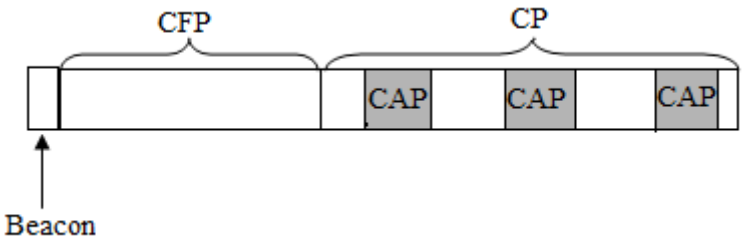


Figure 3.7 Generation of CAPs during the CP

A new set of frames is defined, which is similar to the legacy PCF frame set but with a QoS attribute added: QoS NULL, QoS DATA, QoS CF-ACK, QoS CF-POLL, QoS DATA+CF-ACK, QoS CF-ACK+CF-POLL, QoS DATA+CF_POLL and QoS DATA+CF-ACK+CF-POLL and. A CAP is a sequence of TXOPs initiated by the HC with the transmission of a QoS Data frame or QoS CFPOLL frame.

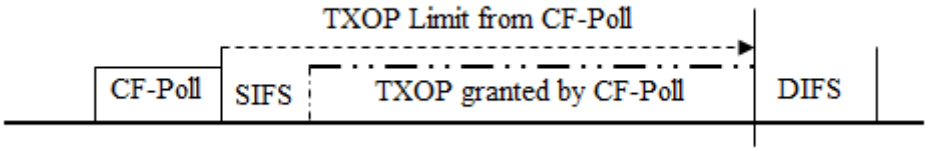


Figure 3.8 Transimission opportunity in HCF

A TXOP ends when at least once one of the following conditions is met:

- Transmission of a QoS DATA frame with the non-final (NF) flag set to 0, which means that there are no further frames queued for transmission.
- Expiration of the TXOP duration implicitly given by the default MIB variable dot11DefaultCPTXOPLimit or explicitly set by the HC in beacons, association response frames or QoS CF-POLL frames.
- The polled station allows the wireless medium to remain idle for PIFS.

The CAP ends when the wireless medium remains idle for a DIFS interval.

As already seen, in legacy IEEE 802.11 the multiplexing of DCF and PCF in a superframe is fixed. The MIB variables aCFPRate and aCFPMaxDuration clearly define the length and frequency of the CFPs and, by extension, the proportion between the CP and the CFP. In HCF this proportion must also be specified in order to limit polling, as contention is still needed for important management tasks (e.g. association of terminals to the AP). In HCF, the rate and proportion of CAPs are given by the MIB variables dot11CAPRate and dot11CAPMax. The meaning of dot11CAPRate is very different from the corresponding PCF variable aCFPRate and specifies the fraction of the CP that can be used for CAPs, expressed in units of microseconds per 64 microseconds (e.g. a dot11CAPRate value of 32 means that at most one half of the CP can be spent with CAP). On the other hand, the variable dot11CAPMax specifies the maximum duration of a CAP. Together, the two HCF MIB variables define a token bucket of time whose state is given by a CAP timer. The CAP timer is initialised to zero and counts upwards at a rate defined by dot11CAPRate, until it reaches the maximum value of dot11CAPMax. At any time, the AP can deduct from the CAP timer a number of units equal or less than its current value and start a CAP whose duration in microseconds corresponds to the number of deducted units.

HCF also defines Reservation Request (RR) frames that can be used by the stations to request TXOPs to the HC. Additionally, Controlled Contention (CC) frames can be used by the HC to initiate a controlled contention interval (CCI) in which the stations contend for transmission using the short RR frames, with actual data transmission being done after contention by the winning station. The HCF specification will also include appropriate signalling to negotiate QoS parameters for specific data streams (IEEE LAN/MAN Standards Committee, 2005).

3.3.2 MAC Frame Formats

The format of the MAC frames is specified in this section. A STA shall be able properly to construct a subset of the frames specified in this clause for transmission and to decode a (potentially different) subset of the frames specified in this clause upon validation following reception. The particular subset of these frames that a STA constructs and decodes is determined by the functions supported by that particular STA (IEEE LAN/MAN Standards Committee, 2007). All STAs shall be able to validate every received frame using the frame check sequence (FCS) and to interpret certain fields from the MAC headers of all frames.

Each frame consists of the following basic components:

① A MAC header, which comprises frame control, duration, address, and sequence control information, and for QoS data frames, QoS control information;

② A variable length frame body, which contains information specific to the frame type and subtype;

③ A FCS, which contains an IEEE 32-bit CRC.

The MAC frame format comprises a set of fields that occur in a fixed order in all frames. Figure 3.9 represents all frame types derived from the general IEEE802.11 frame format.

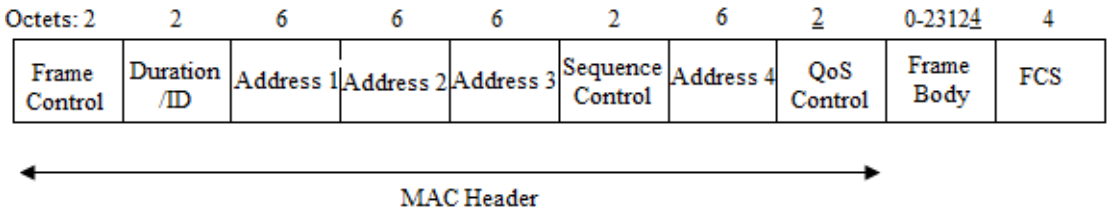


Figure 3.9 MAC frame Format

The first three fields (Frame Control, Duration/ID, and Address 1) and the last field (FCS) constitute the minimal frame format and are present in all frames, including

reserved types and subtypes. The fields Address 2, Address 3, Sequence Control, Address 4, QoS Control, and Frame Body are present only in certain frame types and subtypes. The Frame Body field is of variable size. The maximum frame body size is determined by the maximum MSDU size (2304 octets) plus any overhead from security encapsulation.

A description of all the fields and subfields that compose the frame header is given in what follows:

- Frame control: contains all the information that the MAC requires to correctly interpret all the subsequent fields. As shown in Figure 3.10, it is made up of several subfields.

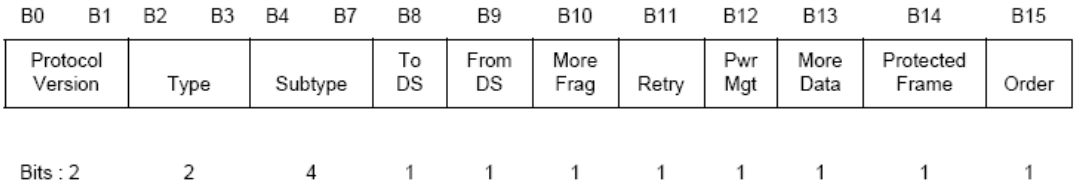


Figure 3.10 Frame Control Field

  - Protocol version: specifies the version of the MAC protocol used to construct the frame. To date, this subfield has only one valid value, since there is only one version.
  - Type and subtypes: identify the function of the frame and which other MAC header fields are present in the frame.
  - To DS and From DS: indicate whether the frame is destined to the DS or comes from the DS, respectively. When both subfields have a value of 0, the frame is directed from one station to another, in the same IBSS. On the other hand, both subfields set to 1 indicate that an IEEE 802.11 WLAN is being used as the DS.
  - More fragments: indicates whether a frame is the last fragment of a larger frame or not.

- ✧ Retry: allows a receiver to realize if the frame is being retransmitted.
- ✧ Power management: used to announce the power management state of a station. If the subfield is set to 1, the station enters in power save mode when the frame exchange is completed. Frames from an AP always have a value of 0.
- ✧ More data: a station receiving a frame with the more data subfield set to 1 is notified that there is at least one more data frame buffered at the AP.
- ✧ WEP: when set to 1, it indicates that the frame body has been encrypted using the Wired Equivalent Privacy (WEP) algorithm.
- ✧ Order: indicates if the frame was provided to the MAC with a request for strictly ordered service.

◆ Duration/ID: the information contained in this field varies according to the state of the station that is accessing the medium. Table 3.4 shows the different possible values for the duration/ID field.

Table 3.4 Values for the Duration/ID Field

Duration/ID Field			State of the station
Bit 15	Bit 14	Bit 13 - 0	
0	0-32767		DCF operation contains the duration of frame exchange (in microseconds), allowing other stations to update their NAVs.
1	0	0	PCF operation.
1	0	1-1683	Reserved.
1	1	0	Reserved.
1	1	1-2007	PS mode used in PS-poll frames to indicate the station AID.
1	1	2008-16 383	Reserved.

- Address: each of these addresses contain one of the following subfields, depending on the to DS and from DS subfields in the frame header, as shown in Table 3.5.

Table 3.5 Information Contained in the Different Address Fields

To DS	From DS	Address 1	Address 2	Address 3	Address 4	Notes
0	0	RA=DA	SA	BSSID	-	Frame exchange within an IBSS
0	1	RA=DA	BSSID	SA	-	Frame from an AP
1	0	RA=BSSID	SA	DA	-	Frame to an AP
1	1	RA	TA	DA	SA	IEEE802.11 WLAN as DS

  - BSS Identifier (BSSID): represents a unique identifier assigned to each BSS within an ESS.
  - Transmitter Address (TA): MAC address of the station that transmits the frame.
  - Receiver Address (RA): MAC address of the station to which the frame is sent over the wireless medium.
  - Source Address (SA): MAC address of the station that originates the frame.
  - Destination Address (DA): MAC address of the final destination of the frame.
- Sequence control: sequence or fragment number of a frame.
- QoS Control: identifies the TC or TS to which the frame belongs and various other QoS-related information about the frame that varies by frame type and subtype, as illustrated in Table 3.6.
- Frame body: this field carries the payload, or MSDU, of a frame delivered by upper layers. There are several frames with an empty frame body. These are control frames, management frames and the null data frame.
- Frame check sequence (FCS): this field contains a 32-bit cyclic redundancy check

(CRC) value calculated over all fields in the MAC header and frame body.

Table 3.6 QoS Control Field

Applicable Frame (sub) types	Bits 0-3	Bit 4	Bits 5-6	Bit 7	Bits 8-15
QoS (+)CF-Poll frames sent by HC	TID	EOSP	Ack Policy	Reserved	TXOP Limit
QoS Data, QoS Null, and QoS Data+CF-Ack frames sent by HC	TID	EOSP	Ack Policy	Reserved	AP PS Buffer State
QoS data frames sent by non-AP STAs	TID	0	-	Reserved	TXOP Duration Requested
	TID	1	-	Reserved	Queue Size

CHAPTER 4

WIRELESS NETWORK SIMULATION TOOLS

In order to introduce the simulation on wireless networks, this chapter begins with pointing out the importance of simulation. Then, a comparison of simulation tools that are already available is presented. The remaining of the chapter focuses on OPNET Modeler, which is used for the study.

4.1 The Need for Simulation

The construction of a real test bed for any proposal network is always a costly or even impossible task, if factors like mobility, reaching distance, etc. are taken into account (Schilling, 2005). Additionally, most measurements are time-consuming and unrepeatable, therefore simulations on network performance are required to bypass the problem. The principle is that if a network can be modeled, we can change its features and then analyze the corresponding results.

However, network simulators just try to model the real networks in a specific way, they cannot perfectly model all the details of the networks (Heidemann, Bulusu, Elson, & Intanagonwiwat, 2001). The goal of a simulator is to give a meaningful insight into the network in order to gain realistic and adaptable results.

In context with networks, especially wireless ones, simulators are usually used for testing the network capacity and efficiency under specific scenarios before installing a new network.

4.2 Type of Simulators

Network simulators can be categorized into several types based on some criteria such as technology, processing method, commerciality, etc. A short list of the current network simulators include OPNET, NS-2, NS3, GloMoSim, OMNeT++, SSFNet, NetSim, QualNet. Limited by the topic, we do not intend to cover all the available network simulators, but focus on some typical ones and introduce others briefly.

① Methods of Simulations

Currently there are two typical methods of simulation: analytical and discrete event simulation (WANsim, 2009). The former uses mathematical models to represent a network topology, defining the scenarios, specifying the nodes on the network, the links between those nodes, but may sacrifice accuracy. The latter produce predictions in the network at a low level (packet-by-packet), which makes them more accuracy but slow to generate results. The common approach is to combine both methodologies in one simulator, like OPNET, NS, etc.

② Commerciality of Simulators

Commercial network simulators are those would only provide the source code of the software or some affiliated packages to the users who pay to get the license or pay to order some particular packages to meet their own specific usage requirements. OPNET is one of the commercial network simulators, with the advantage that they generally have complete and up-to-date documentations and they can be consistently maintained by professionals from the companies. On the other hand, the open source network simulator is defeated in this aspect that people working on the documentation are mostly from all different organizations. This problem can be serious when the different versions come with many new features and it will become difficult to trace and understand the previous codes

without appropriate documentations.

However, the open source network simulators have the advantage of offering everything open to any individual or organization can contribute to it and amend it. Moreover, the interfaces are also open for further improvement. They are so flexible that can reflect the most recent technologies in a faster way than commercial network simulators. On the contrary, we notice that some strong points for commercial network simulators are the deficiency of the open source ones. Short of enough systematic and complete documentations and lack of version control supports can lead to some unexpected troubles that is somehow waste of time and also limit the applicability and life-time of the open source network simulators.

4.3 A Brief Comparison

Several comparisons from difference points of view have been done between network simulators. Afterwards some comparisons on wireless network packets of some popular simulators are resumed.

In the paper of Lucio, et al. (2003) OPNET modeler and NS-2 have been compared to a real testbed. In order to test the network performance with different types of traffic, CBR(Constant Bit Rate) data traffic and an FTP session are generated for both real and simulated results.

According to the results from this paper, NS-2 performed better than OPNET modeler using the default setting, in terms of accuracy of bandwidth estimation for the pure CBR traffic. However, when network load getting higher, NS-2 behaved differently to the testbed, meanwhile the OPNET modeler gave more accurate results overall.

As for FTP session results, NS-2 FTP simulation model only indicated a general

transfer rate rather than replicating the actual network flow (ISI, 2009). Modeler performed closely to the testbed results, and similar results appeared in the case of when two types of data traffic are generated together. On the other hand, both NS-2 and OPNET modeler proved to be fast, requiring small amount of time to obtain the results.

Regarding simulation speed, NS-2 has a “small suite”, so several modifications and extra care has to be taken to manage memory allocation and CPU time for large-scale networks. Nevertheless, OPNET modeler has a “heavy suite” (large software overhead) which provides diverse statistics modules at different levels.

In another paper by D. Cavin, Y. Sasson and A. Schuper (2002), it was presented a set of measures collected during the simulation of the flooding algorithm on OPNET modeler, NS-2 and GloMosim (Cavin, Sasson, & Schiper, 2002). The simulations are based on the scenario: wireless ad-hoc network of 50 mobile nodes, uniformly placed on a terrain of 1km×1km. All nodes runs the IEEE802.11 MAC protocol in the ad-hoc mode (or peer-to-peer).

From the result of effective transmission range, big differences could be found between the simulators. The resolving success rate in percent of OPNET modeler is sensitive to the power range in meters, but NS-2 and GloMosim have low dependence between these two factors.

The results also show that the success rate of every single simulator does not change much when the mobility of the nodes is increased. But the results among simulators differ significantly. The success rate of OPNET model and GloMosim are more than twice as high compared to the one of NS-2.

As for overhead produced during the simulation process, OPNET modeler produces a lot more duplicates than NS-2. However, for the time needed by each simulator to flood a message through the entire network, the results indicate that OPNET modeler needs about 10 times more than NS-2 (Cavin, Sasson, & Schiper, 2002).

Considering all the features of wireless network simulators available, OPNET modeler is the first choice for modeling and simulating the wireless network scenario in this study. Although there are indeed some limitation of this tool, OPNET modeler's performance in the significant features overwhelms the others.

4.4 OPNET Modeler Basics

OPNET Modeler provides a comprehensive development environment supporting the modeling of communication networks and distributed systems. It is traditionally used for performance measures and behavioral analysis of a proposed system. This section provides an overview of OPNET Modeler's structure and capabilities.

4.4.1 Modeler Architecture

OPNET models are structured hierarchically, in a manner that parallels real network systems. Specialized editors address issues at different levels of the hierarchy. This provides an intuitive modeling environment and also permits re-use of lower level models.

Network, Node, Process, and External System modeling environments are referred to as the modeling domains of OPNET Modeler because they span all the hierarchical levels of a model (OPNET Technologies,Inc., 2009), as shown in Figure 4.1.

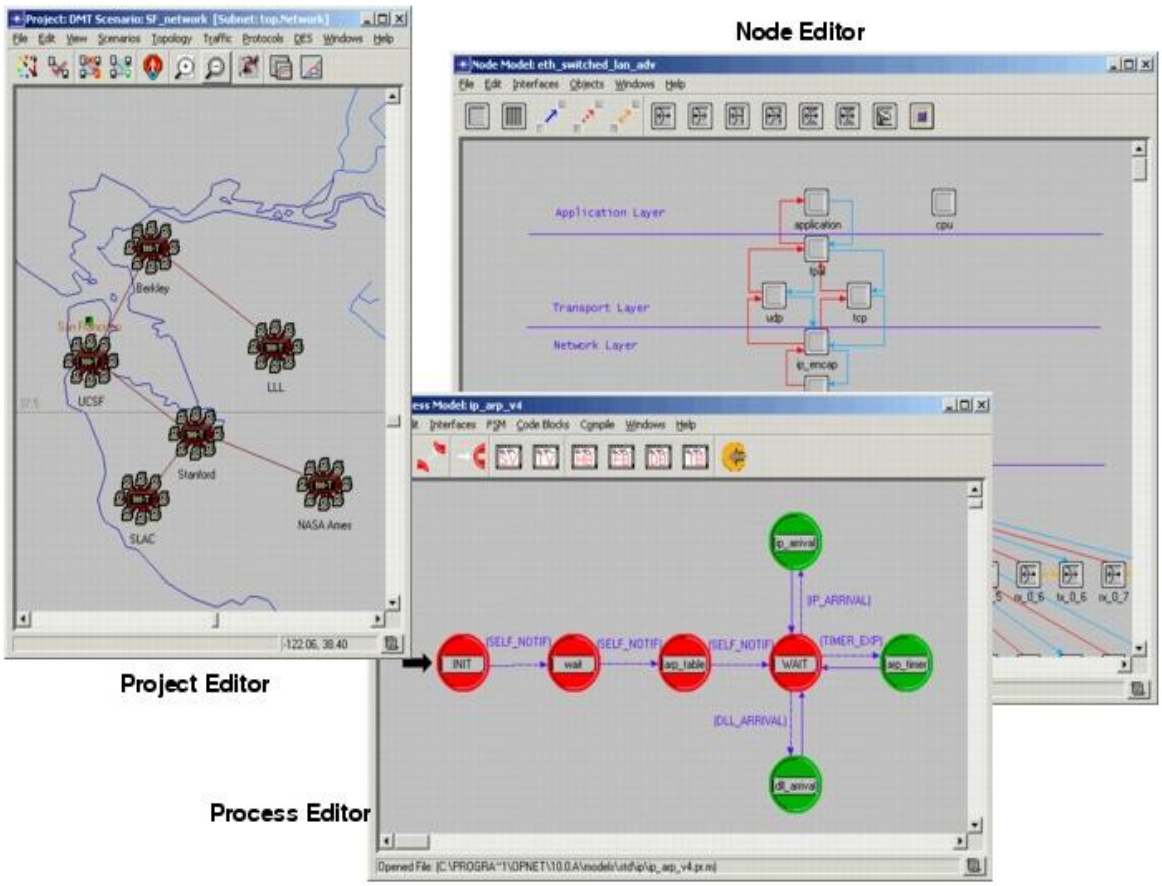
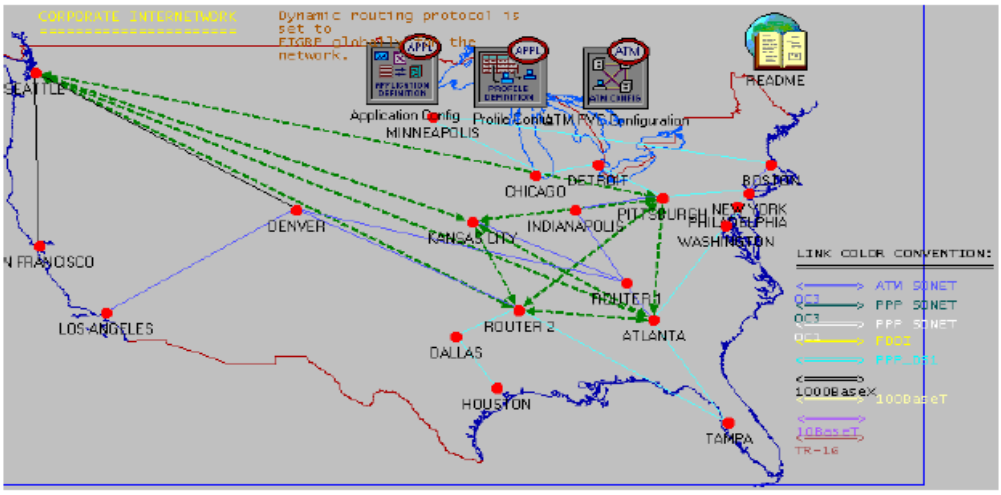


Figure 4.1 Graphical Editors for Network, Node and Process Models

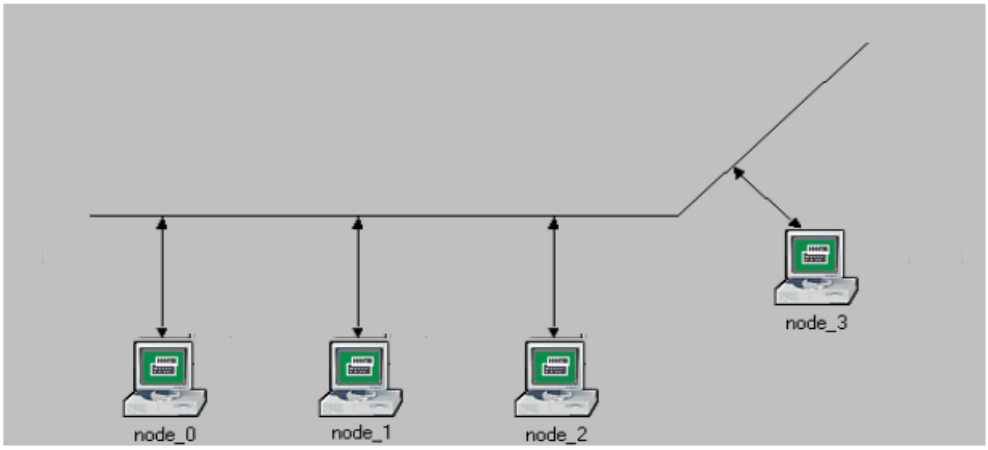
The Network Domain at the top level is to define the topology of a communication network. The communicating entities are called nodes and the specific capabilities of each node are defined by designating their model. Network models consist of nodes and links that can be deployed within a geographical context.

Most nodes require the ability to communicate with other nodes to do their function in a network model. Several different types of communication link architectures are provided to interconnect nodes that communicate with each other. OPNET Modeler provides simplex (unidirectional) and duplex (bidirectional) point-to-point links to connect nodes in pairs and a bus link to allow broadcast communication for arbitrarily large sets of fixed nodes. The Wireless functionality adds the capability for fixed, satellite, and mobile nodes to communicate with each other via radio links. While bus

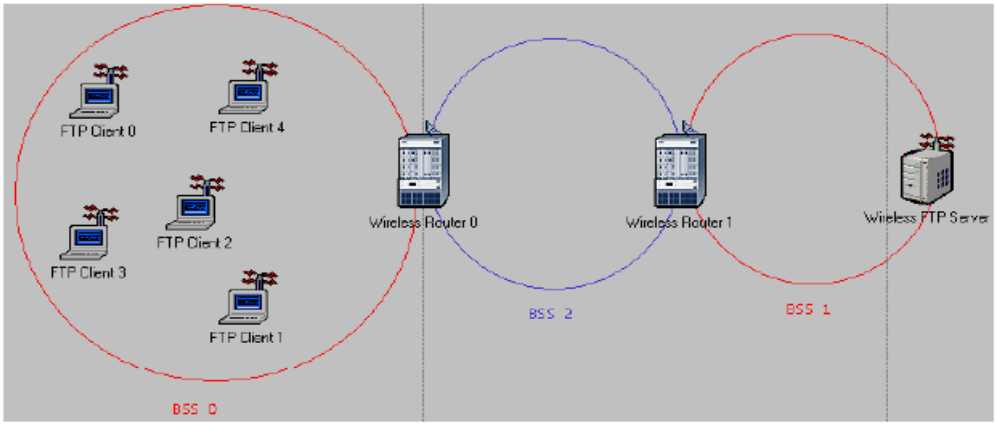
and point-to-point links are modeled as explicit objects that you must create, radio links are dynamically evaluated based on characteristics of the communicating nodes. Figure 4.2 shows some typical network model diagrams involving each of the supported link types.



Point-to-Point Links



Bus Links



Radio Links

Figure 4.2 Network Models with Point-to-Point, Bus and Radio Links

Fixed, mobile, and satellite subnetwork objects provide hierarchy in the network model and are used to break down complexity into multiple levels, as illustrated in Figure 4.3. Subnets can contain various combinations of nodes, links, and other subnets, and can be nested to any depth, (OPNET Technologies,Inc., 2009).

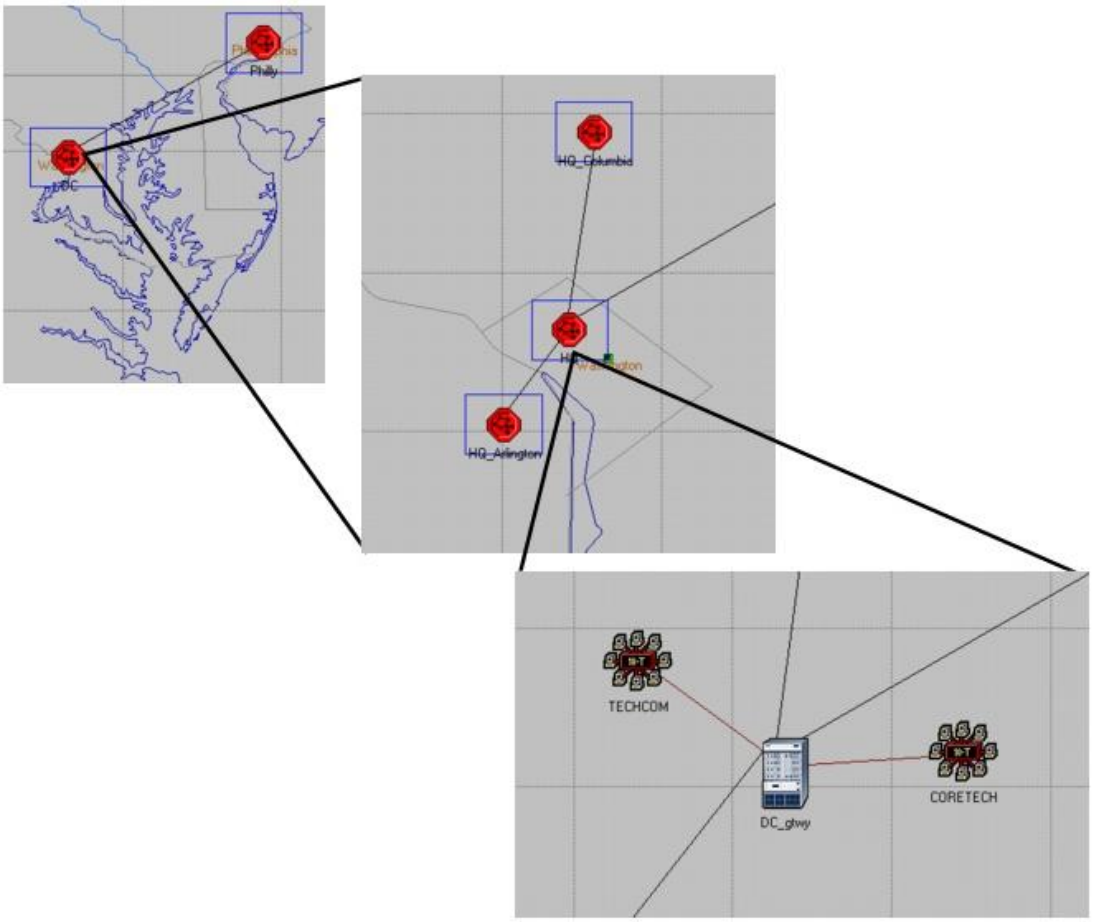


Figure 4.3 A Hierarchical Network with Two Levels of Subnetworking

The second level of the Modeler Architecture is the node domain and the associated node editor. It provides for the modeling of communication devices that can be deployed and interconnected at the network level. In OPNET Modeler terms, these devices are called nodes, and in the real world they may correspond to various types of computing and communicating equipment such as routers, bridges, workstations, terminals, mainframe computers, file servers, fast packet switches, satellites, and so

on.

Node models are developed in the Node Editor and are expressed in terms of smaller building blocks called modules. Some modules offer capability that is substantially predefined and can only be configured through a set of built-in parameters. These include various transmitters and receivers allowing a node to be attached to communication links in the network domain. Other modules, called processors, queues, and external systems, are highly programmable, their behavior being prescribed by an assigned process model. Process models are developed using the Process Editor.

A node model can consist of any number of modules of different types. Three types of connections are provided to support interaction between modules. These are called packet streams, statistic wires, and logical associations, (OPNET Technologies,Inc., 2009). Packet streams allow formatted messages called packets to be conveyed from one module to another. Statistic wires convey simple numeric signals or control information between modules, and are typically used for monitoring the performance or state of another. If needed, both packet streams and statistic wires have parameters that may be set to configure aspects of their behavior. Logical associations identify a binding between modules. Currently, they are allowed only between transmitters and receivers to indicate that they should be used as a pair when attaching the node to a link in the Network Domain. A typical node model that includes the three types of connections is shown in Figure 4.4.

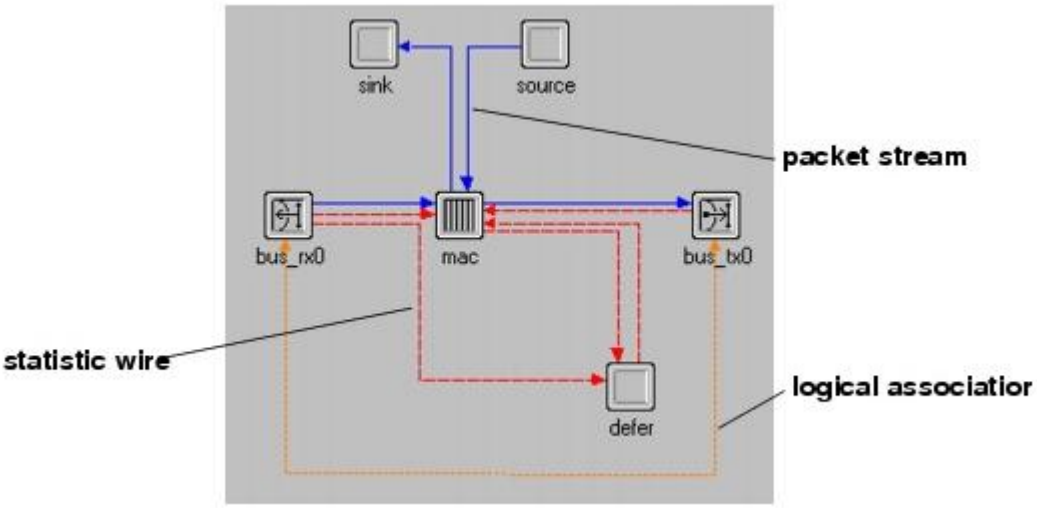


Figure 4.4 Node Model Employing Packet Streams, Statistic Wires and Logical Association

The last level in the Modeler architecture is the process domain, Figure 5. As indicated in the description of the Node Domain, queue and processor modules are user-programmable elements that are key elements of communication nodes. The tasks that these modules execute are called processes. Because it has a set of instructions and maintains state memory, a process is similar to an executing software program. Processes in OPNET Modeler are based on process models that are defined in the Process Editor.

The relationship between process model and process is similar to the relationship between a program and a particular session of that program running as a task (in fact, the term process is used in many operating systems as well). Just as nodes created in the Project Editor are instances of node models defined with the Node Editor, each process that executes in a queue, processor, or esys module is an instance of a particular process model.

The process modeling paradigm of OPNET Modeler supports the concepts of process groups. A process group consists of multiple processes that execute within the same

processor or queue. When a simulation begins, each module has only one process, termed the root process. This process can later create new processes which can in turn create others as well, etc. When a process creates another one, it is termed the new process' *parent*; the new process is called the child of the process that created it. Processes that are created during the simulation are referred to as dynamic processes.

OPNET Modeler places no limits on the number of processes that may be created in a particular processor or queue. Processes may be created and destroyed based on dynamic conditions that are analyzed by the logic of the executing processes. This paradigm provides a very natural framework for modeling many common systems. In particular, multitasking operating systems where the root process represents the operating system itself and the dynamically created processes correspond to new tasks; and multi-context protocols where the root process represents a session manager, for example, and each new session that is requested is modeled by creating a new process of the correct type.

Only one process can be executing at any time. A process is considered to be executing when it is progressing through new instructions that are part of its process model. When a process begins execution it is said to be invoked. A process that is currently executing can invoke another process in its process group to cause it to begin executing. When this happens, the invoking process is temporarily suspended until the invoked process blocks. A process blocks by indicating that it has completed its processing for its current invocation. After the invoked process has blocked, the invoking process resumes execution where it had left off, in a manner similar to the procedure-call mechanism in a programming language such as C.

Processes in OPNET Modeler are designed to respond to interrupts and/or invocations. Interrupts are events that are directed at a process and that may require it to take some

action. They may be generated by sources external to a process group, by other members of a process group, or by a process for itself. Interrupts typically correspond to events such as messages arriving, timers expiring, resources being released, or state changes in other modules, (OPNET Technologies,Inc., 2010). After a process has been invoked due to an interrupt, it may invoke other processes in the group and these may in turn invoke other processes, etc. An interrupt's processing is completed when the first process that was invoked blocks.

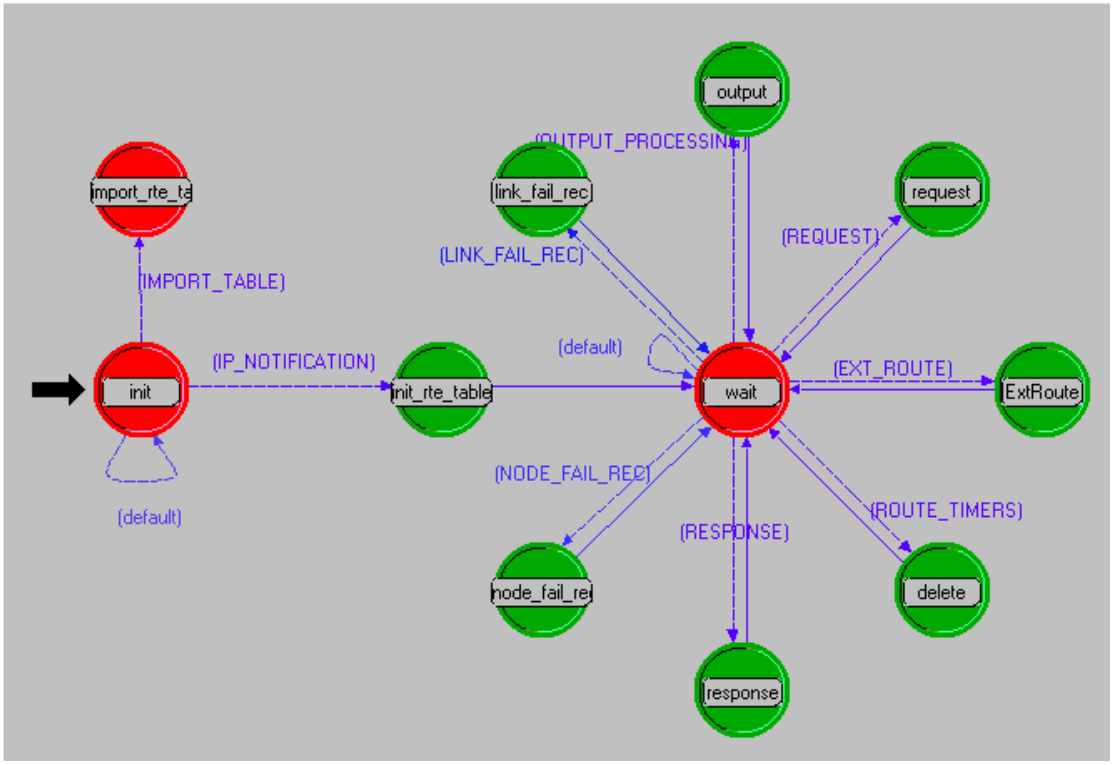


Figure 4.5 State Transition Diagram in the Process Editor

The Process Editor expresses process models in a language called Proto-C, which is specifically designed to support development of protocols and algorithms. Proto-C is based on a combination of state transition diagrams (STDs), a library of high-level commands known as Kernel Procedures, and the general facilities of the C or C++ programming language. A process model's STD defines a set of primary modes or *states* that the process can enter and, for each state, the conditions that would cause

the process to move to another state. The condition needed for a particular change in state to occur and the associated destination state is called a transition.

The three domains briefly mentioned represent the basis of the Modeler simulation tool, but there are also some other editors available to offer specific capabilities to address the diverse issues encountered in networks and distributed systems. Just to name a few, there is External System Editor to develop external system definitions, the Link Model Editor to create, edit and view link models, or the ICI Editor that are used to communicate control information between processes.

4.4.2 Discrete Event Simulations

As mentioned before, all the nodes in a given network being simulated are represented by a system model, which evolves through a sequence of states generated by discrete event simulations (DES) as a function of time, based on the specifications of the behavior of model components and of their interactions (Raczynski, 2006). To the extent that model specifications are accurate, this evolution is representative of the way in which the actual system functions over time. The notion of time in a simulation is not directly related to the actual time that it takes to run a simulation (as measured by a wall-clock or the computer's own clock), but simply a variable maintained by the simulation program. This variable is more precisely referred to as simulation time to clearly distinguish it from real (wall-clock) time.

In the discrete-event modeling approach, the progression of the model over simulation time is decomposed into individual points where change can take place. The OPNET Modeler term for each such point in time is an event (OPNET Technologies,Inc., 2010). Each event represents a need for the model to possibly effect some change in its state or to make some decision. Certain events may result in no change in the system's state due to the fact that the actions that would cause change

are specified conditionally (that is, they depend on the current state of the system and possibly on properties of the event itself). Some common actions that are modeled as events are listed below,

- receipt of a message (a packet) or a command by a process
- expiration of a timer (for example, while waiting for an acknowledgment)
- indication of availability of a resource
- indication of completion or partial completion of a task by a resource
- change in a statistic that is monitored by a process
- start or end of a packet's transmission/reception on a link
- generation of a new message, command, or task by an application process
- failure or recovery from failure of a device

Every time that a new event occurs, it is said to be executed by the simulation. Simulation time increases monotonically as new events are executed. Because OPNET Modeler supports modeling of distributed systems, it must allow multiple events to occur simultaneously in terms of simulation time, and to affect different components of the system. Therefore, it is possible for any number of distinct events to occur at the exact same simulation time.

The simulation times at which events are executed are expressed using the maximum-precision real number supported by the host computer. There is no requirement that events be spaced regularly in simulation time; in fact, it is common for the density of events to vary significantly as a function of time while a simulation is progressing, as shown in Figure 4.6:

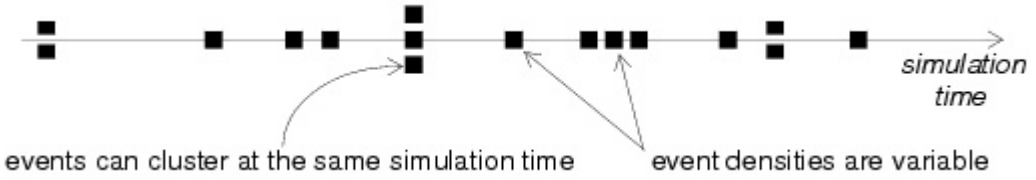


Figure 4.6 Typical Simulation Timeline (OPNET Technologies,Inc., 2010)

Because discrete event simulations allow multiple events to occur simultaneously, they must also impose the constraint that all events have a duration equal to zero. If this were not the case, then it would implicitly be possible to observe a different simulation time at the beginning and at the end of an event. However, the subsequent event may be specified to begin at the same simulation time, which would force time to regress (Banks, 1998).

Therefore, simulation time is not allowed to progress during an event, but only between events. In fact, this time model states that simulation time is always equal to the time at which the current event began. Therefore, simulation time can be viewed as a variable that "jumps" to track the time specified for each new event.

A simulation in progress can be viewed both as a generator and a consumer of events. New events are generated by the components of the simulation model to schedule new future activities, based on their underlying behavioral models. Events are consumed as simulation time progresses and specified event times are overtaken.

Discrete event simulations manage events with an event list. The purpose of maintaining an event list is to make sure that all events are executed in the correct time order. Each event has an associated time at which it is specified to occur. The requesting of new events, which requires specification of the simulation time at which they must execute, is called event scheduling. Events are said to be scheduled at particular times. Events may be scheduled for any future time, or for the current time,

but never for past times, because the simulation time variable can only progress forward. It is common for new events to be scheduled for the current simulation time because this corresponds to a chain of events that cause each other, using mechanisms that are modeled as having a delay of zero. For example, transfers of packets via packet streams, or values via statistic wires, are both often performed with no delay.

The simulation event list (see Figure 4.7) maintains all of the events in time-order so that the next event may be readily determined when the current one has completed execution. The earliest remaining event is referred to as the head of the event list, and the latest event is called the tail. During the execution of each event, new events may be scheduled, including events for the current simulation time. Therefore during execution of an event, the head or tail of the event list may be replaced with a new event.

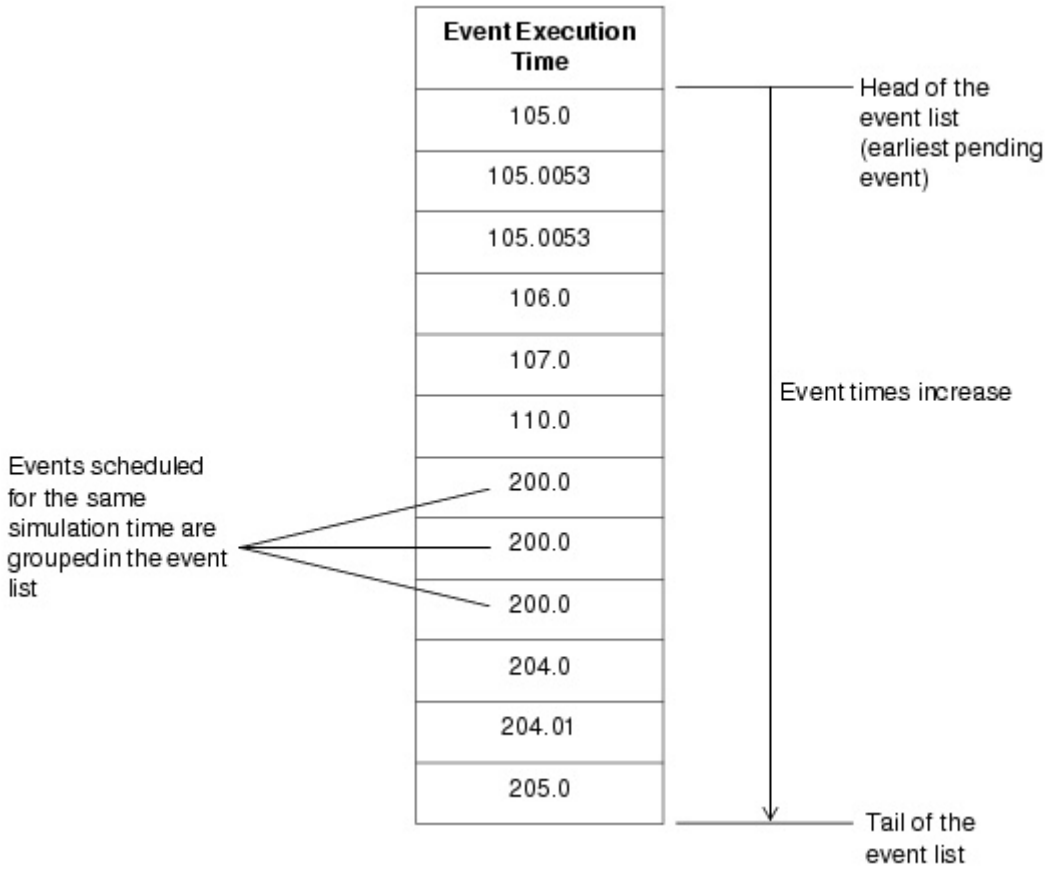


Figure 4.7 Simulation Event List (OPNET Technologies,Inc., 2010)

The term "event list" only refers to the conceptual structure of the list. While the events behave as if they are organized in the form of a sequential list, they in fact are stored in proprietary data structures that support very efficient insertion and extraction of events while maintaining the required ordering. The "list-like" abstraction is maintained for all user-visible interfaces to the event storage. In particular, the Event package of Kernel Procedures, used to scan the event list and query the properties of events, hides the underlying organization from the user.

During a simulation, the event list may continually grow and shrink as new events are scheduled and previously scheduled events are executed or canceled (Banks, 1998). Each simulation will have its own particular pattern of growth for the event list, depending on the activities that are modeled. Because all future events that are known are scheduled and placed in the event list, a simulation need only continue if the list contains new events to execute. Therefore, if the list becomes empty as a result of executing the last remaining event, a simulation will terminate as that event completes, unless it has in turn scheduled at least one new event. An empty event list will cause a simulation to end regardless of the simulation time elapsed so far.

At the beginning of simulation, the event list must receive at least one event so that execution may begin. The initially scheduled events "bootstrap" the simulation by in turn causing other events. A special type of event called "begin simulation" may be scheduled automatically for processes by enabling the "begsim intrpt" attribute of processor and queue modules within the node domain. Also processor modules automatically choose a time at which to issue their first packet, based on their assigned distributions. These are the types of events that usually appear in the event list when it is examined at the start of simulation.

The size of the event list is continuously varying, as new events are generated/deleted and consumed by the simulation itself. These events are more than just a simple indication regarding the simulation activity and its associated simulation time. In fact, each event is a complex entity and a wide range of events types exist, representing different situations within a simulation. Each event by itself has a number of attributes describing how it should be executed, including, among others, the time, the identification, or the type of the process that will receive the event. These attributes are enough to differentiate and characterize the event. As a given process model receives an event to process, it uses these attributes to know which actions it must perform.

In order to simulate the behavior of a given network and obtain performance results, traffic must be added. All the traffic being generated, transmitted, and received within the network is represented by events. OPNET allows the addition of traffic in the network mainly in two different ways: either manually, by setting the attributes from the various applications/packet generators, or by importing traffic data from external files or programs (OPNET Technologies,Inc., 2010).

OPNET has multiple options for importing traffic. This allows using real traffic in the network to obtain better and more realistic results for scenarios. Regarding the manual setup of traffic, two different types can be modeled: explicit and background. "Explicit traffic" is packet-by-packet traffic, in which the simulation models each packet-related event (packet created, packet queued, packet transmitted, etc.) that occurs during the simulation. Explicit traffic modeling provides the most accurate results because it models all protocol effects. However, this also results in longer simulations and higher memory usage, because the simulation allocates memory for each individual packet.

"Background traffic" is analytically modeled traffic that affects the performance of explicit traffic by introducing additional delays. Unlike explicit traffic, background traffic can affect not only discrete event simulations, but also flow analyzes. Discrete event simulations that include background traffic use the hybrid simulation model. This model includes the effects of background traffic to calculate queue build-ups on intermediate devices and delays based on the queue length, at any time during in the simulation (Zhu, Yang, Aweya, Oullette, & Montuno, 2002). Because each packet that produces traffic on the network is not explicitly modeled, using background traffic can speed up simulations considerably.

There are three general methods for the generation of background traffic, (OPNET Technologies,Inc., 2009):

- Traffic Flows—a traffic flow describes an end-to-end flow of traffic from a source to one or more destination nodes. Traffic flows can be created manually, using traffic flow objects. Traffic flows can also be imported from external files. With a license for eXpress Data Import functionality, traffic can be imported from programs like Netflow Collector and NetScout nGenius.
- Baseline Loads—this type of traffic (also called "static background utilization") represents traffic as a background load on a link, node, or connection. Unlike a traffic flow, which can span multiple links and nodes, a traffic load is "static" and applies to one object. You can also convert existing link loads to traffic flows, which allows flow analyzes to account for these loads.
- Application demands—you can use application demands to represent background traffic flowing between two nodes. Besides background traffic, application demands can also be configured to purely discrete (explicit) traffic or a combination of the two (hybrid traffic).

4.4.3 Wireless LAN Model Suite

The wireless LAN model suite includes the features of the IEEE 802.11, 802.11b, 802.11a, 802.11g, and 802.11e standards that allow users to analyze wireless network implementations using discrete event simulation.

The model supports 4 types of network configurations, which are Ad-hoc Network, Infrastructure BSS, Extended Service Set and Wireless Backbone, as shown in Figures 4.8-11.

Figure 4.8 Ad-hoc Network

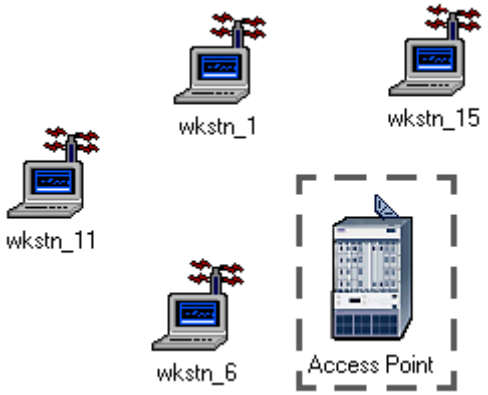


Figure 4.9 Infrastructure BSS

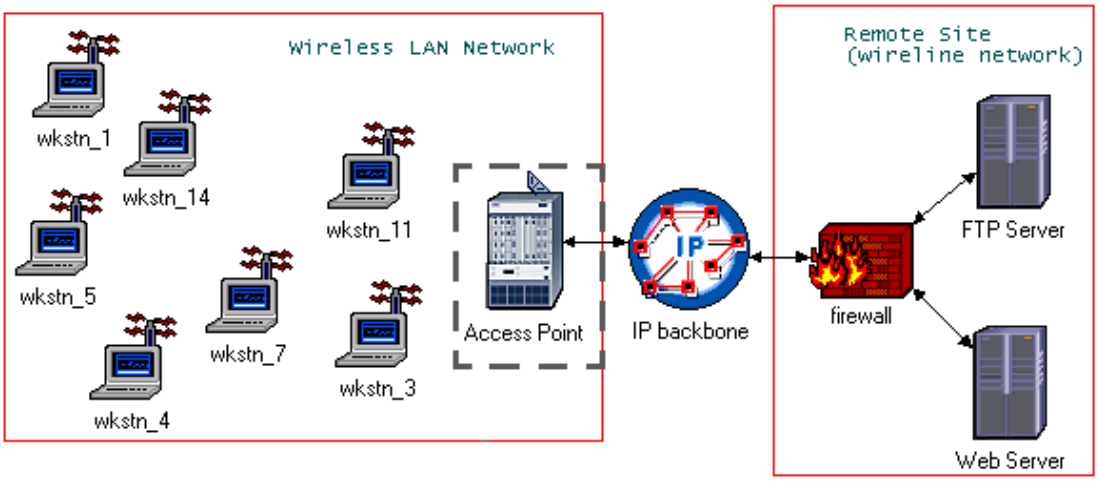


Figure 4.10 Extended Service Set

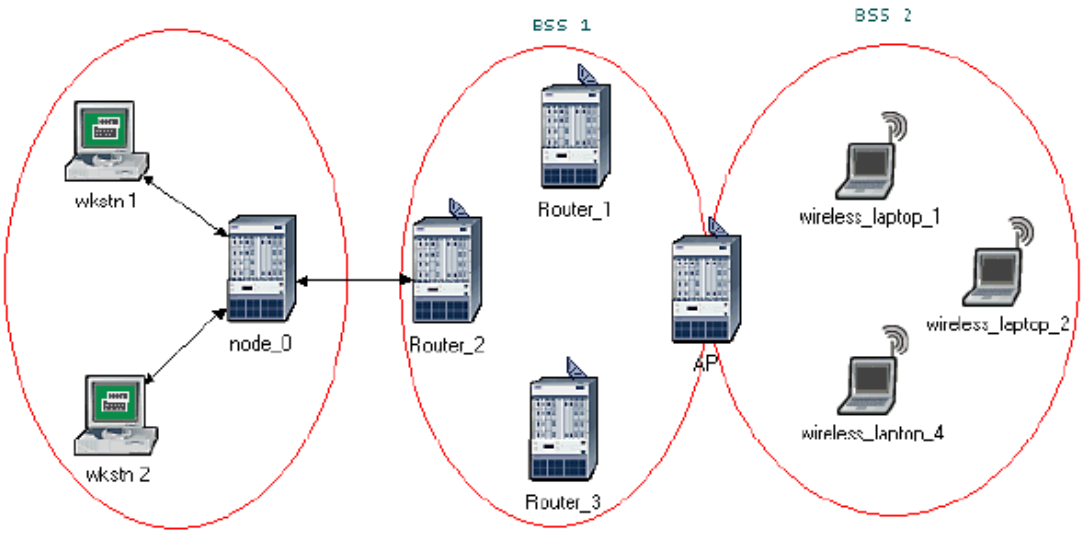


Figure 4.11 Wireless Backbone

In an ad-hoc network of several stations, the workstations can have peer-to-peer connections with other stations in the BSS, but communication is limited to within the BSS, while the workstations of Infrastructure BSS communicate with each other through the Access Point. Similarly, workstations in Extended Service Set (ESS) communicate with each other and with nodes outside their LAN through the AP. In wireless backbone architecture, you can set a BSS identifier for each router interface.

Wireless models rely on a broadcast medium and wireless nodes and subnets can move during a simulation, so there are additional things to consider when creating wireless networks.

Once the WLAN network topology is built, you can introduce traffic into the model. You can do this through the use of standard applications (such as VoIP or E-mail), through raw packet generation, or with traffic demands. Application support is included in the WLAN workstation and server nodes.

Standard network applications are implemented in a two-tier architecture in which the client issues a request and a server or client receives the request and returns a response (OPNET Technologies,Inc., 2009). This request-response exchange typically

happens within one "conversation" between the client and the server, or between a client and another client. In this chapter, we use the term "conversation" to represent a sequence of activity between a client and a server within the context of a given application. A conversation includes a pattern of data exchanges, typically defined in a statistical manner that repeats over time.

4.4.4 Results Collection

Modeler has several different mechanisms that allow the simulator user to have flexibility in the way results are obtained and shown. There are four output types:

- Output vectors: a given variable can vary as the simulation time increases, and output vectors are used to record this behavior. This way, each output vector is composed of a series of values and associated times. Multiple independent output vectors can be collected simultaneously. As an example, Figure 4.12 shows an example of a trace obtained from a multi-simulation vector output.

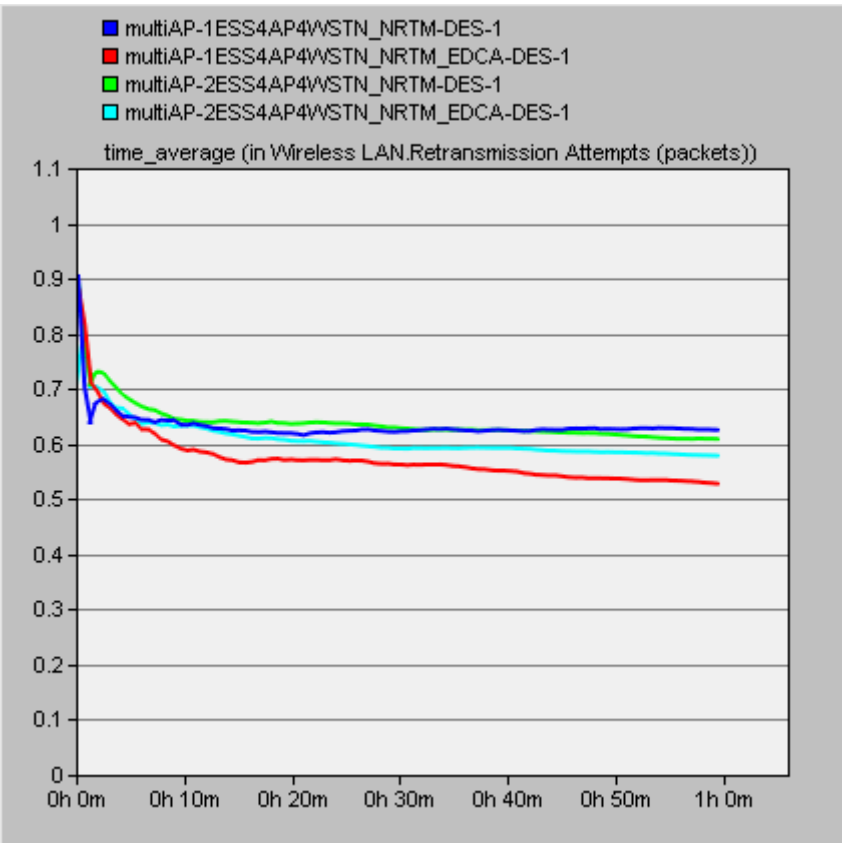


Figure 4.12 Example of Vectors Data Result Panel

- Output scalars: certain metrics of interest do not vary over time; instead, they have one value that is representative of some kind of system performance, as averages or standard deviations. Each scalar is usually recorded only once for a given simulation, but scalars from multiple simulations can be combined to analyze the dependency on a given simulation input. Figure 4.13 is the visualization of a scalar collected during a simulation run. Here, network load standard deviation continuously modified over a simulation to evaluate the behavior of global delay.

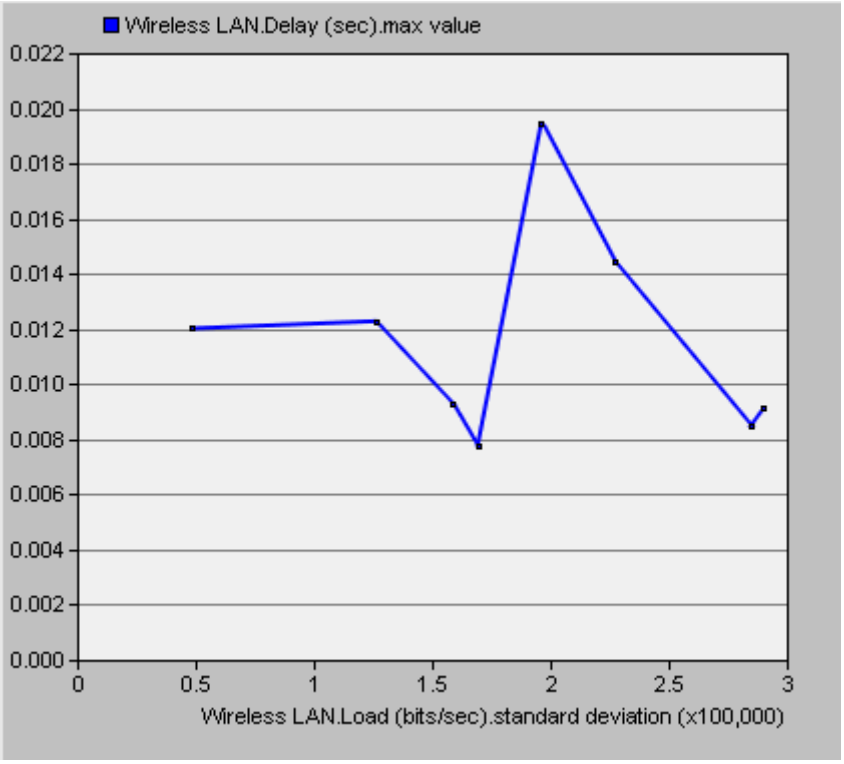


Figure 4.13 Example of a Scalar Data Result Panel

- Animations: this type of output is not a numerical statistic. This is a method to visualize system behavior and interactions among system components. Simulations can generate animations as they run in order to be shown after the end of simulation. With these animations, it is possible to see, e.g., packets

flowing across the nodes of the network, the movement of the nodes, or state transitions within a given process.

- Proprietary reports: a given user can include user-defined processes or link models, and define its own outputs and reports to be generated by Modeler. This can be done by using the general functionalities of the C/C++ programming language.

CHAPTER 5 SIMULATION MODEL OF THE MULTIHOP WIRELESS NETWORK

This chapter begins with the modelling of a Multihop Wireless Network based on IEEE802.11 standards. The representational simulation scenarios will be presented by defining the parameters of the reference (base) scenario, such as services, their definition and usage profiles, the number of mobiles (including access points and users). Then using the reference scenarios, variations are performed on the number of MAPs and STAs, usage profiles and system parameters.

5.1 Reference Scenario Definition

Before studying the impact of the variation of scenarios on network performance in a systematic manner, the definition of the simulation scenarios is detailed, starting with defining the base/reference scenario, identifying the number of users, describing the applications and the usage profiles. Then, the varying scenarios are described, identifying the changes made on the base scenario. Three major situations are considered in the definition of scenarios: variation of mobiles number, variation of applications and usage profiles, with and without using EDCA mechanism (enabled and disabled). The first two situation aims at analyzing the effect of users' parameters, while the latter aims at testifying the advantages and disadvantages of using EDCA in the network.

5.1.1 Layout of the Scenario

The reference scenario used in the simulations is based on the Multihop Mesh Architecture. The Internet-connected gateway is implemented by a WLAN Ethernet router. As for the dual-mode relay gateway, the OPNET has a built-in node model functioned the same way. The user nodes either mobile or fixed or nomadic are

presented by workstations. The overview of the reference scenario is shown in Figure 5.1 and Figure 5.2.

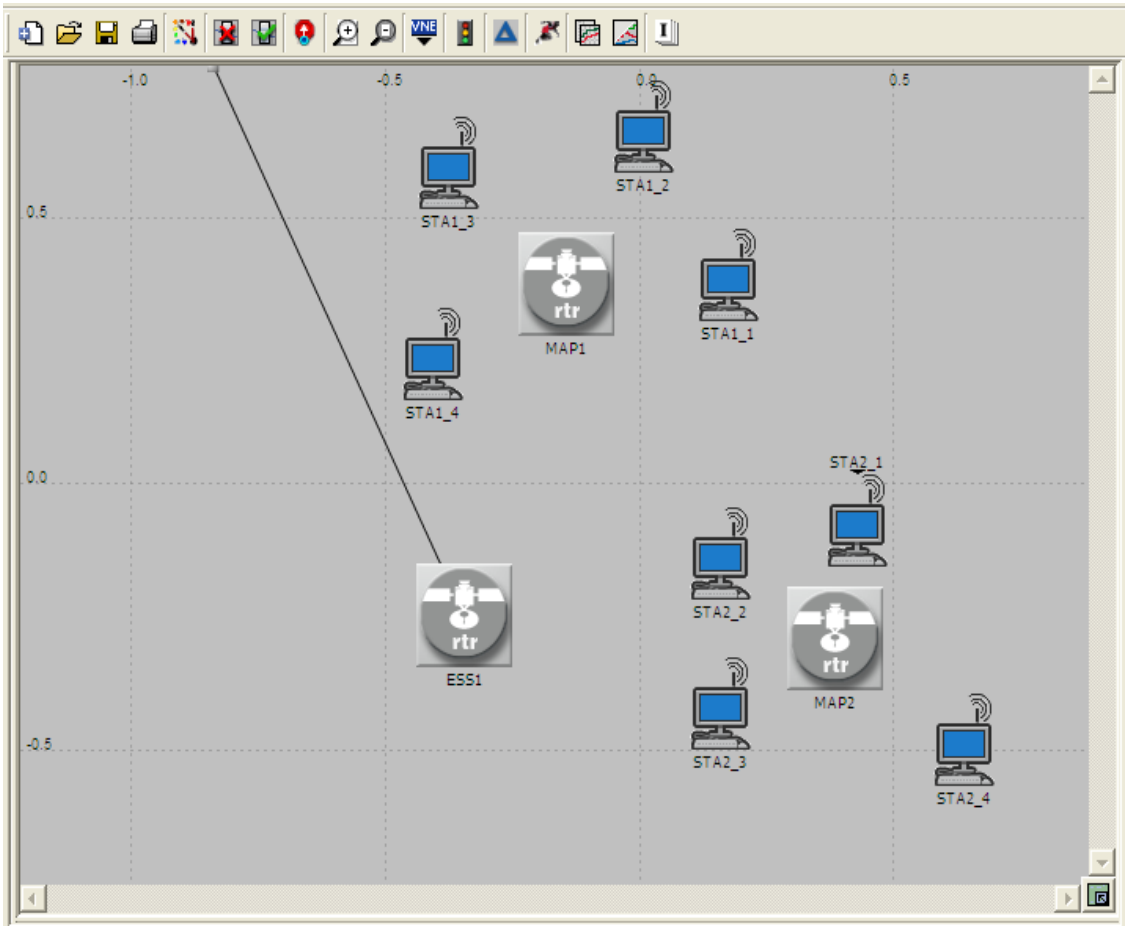


Figure 5.1 The ESS Configuration

Beyond the ESS, Ethernet router is connected to IP cloud (Internet) with PPP_DS3 link which has a data rate of 44.736Mbps. The combination of the server and gateway along with IP cloud provide the traffic from the Internet.

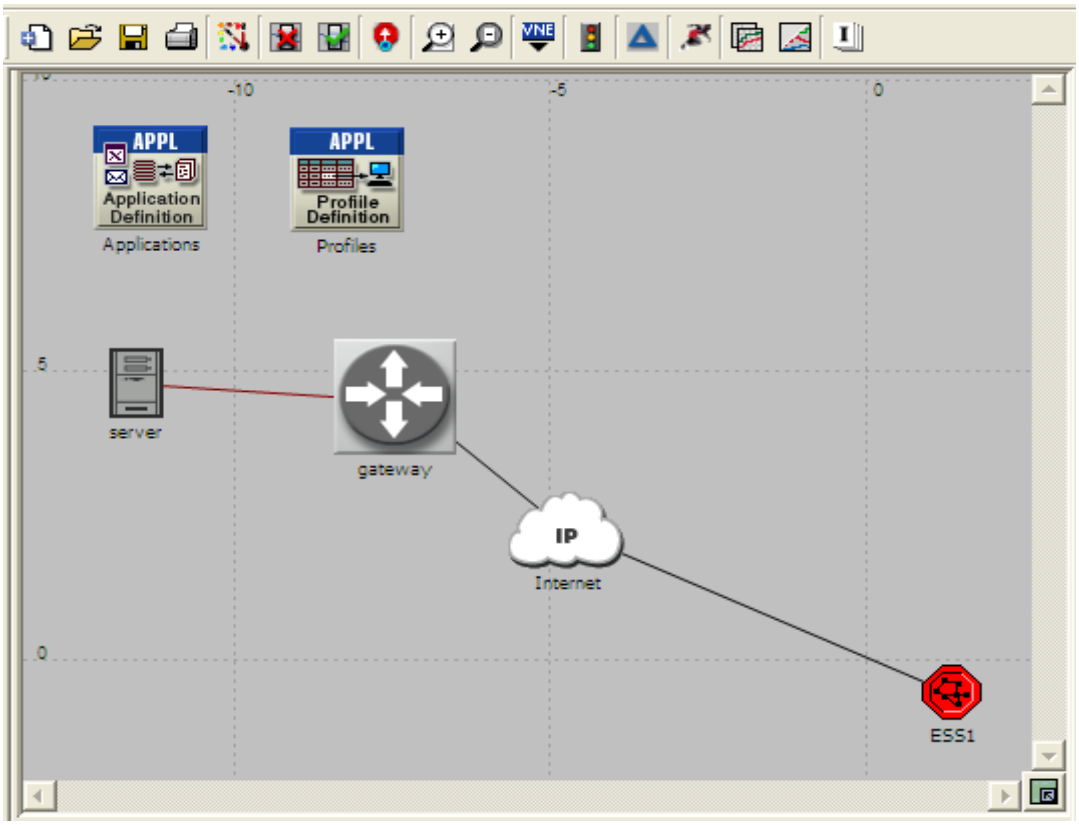


Figure 5.2 The Reference Scenario

5.1.2 Application Definition

In order to characterise the traffic load delivered to the network, the application parameters, including traffic and QoS parameters, are defined in the Application Definition object. Considering the common applications on Internet, there are 7 applications defined in the reference scenarios:

1. Database Access

Attribute	Value
Transaction Mix (Queries/Total Transactions)	50%
Transaction Interarrival Time (seconds)	exponential (12)
Transaction Size (bytes)	constant (32768)
Symbolic Server Name	Database Server
Type of Service	Best Effort (0)
RSVP Parameters	None
Back-End Custom Application	Not Used

Figure 5.3 Data Access Definition

There are two categories of database operations are taken in to account: Database Entry and Database Query. A database entry results in a fixed amount of data being written into the database. A database query results in the client issuing a query, and the server responding with some data. The default transport protocol for the database application is TCP.

- ✧ 50% Transaction Mix represents the database queries equal to the database updates, in terms of transactions.
- ✧ The Transaction Interarrival Time is time between transactions.
- ✧ The average size of an entry or a response to a query, Transaction Size is set as a constant 32,768 bytes.
- ✧ Type of Service represents QoS parameter for assigning priority to this application's traffic.

2. Email

Attribute	Value
Send Interarrival Time (seconds)	exponential (1200)
Send Group Size	constant (3)
Receive Interarrival Time (seconds)	exponential (1200)
Receive Group Size	constant (3)
E-Mail Size (bytes)	constant (1000)
Symbolic Server Name	Email Server
Type of Service	Best Effort (0)
RSVP Parameters	None
Back-End Custom Application	Not Used

Figure 5.4 Email Definition

The default transport protocol used in the email application model is TCP, i.e., messages are sent and received using TCP. Modern email packages use a combination of SMTP (Simple Mail Transfer Protocol) and POP (Post Office Protocol). Both SMTP and POP use TCP as the underlying transport. SMTP transfers an email from the client to the mail server.

- ✧ Send/Receive Interarrival Time is respectively time between e-mails sent/received from the client/server to the server/client.
- ✧ The numbers of e-mail messages grouped before transmission and reception are all 3.
- ✧ The average size of an e-mail message is 1KB.

3. File Transfer

Attribute	Value
Command Mix (Get/Total)	50%
Inter-Request Time (seconds)	exponential (1200)
File Size (bytes)	constant (2000)
Symbolic Server Name	FTP Server
Type of Service	Best Effort (0)
RSVP Parameters	None
Back-End Custom Application	Not Used

Figure 5.5 File Transfer Definition

File transfers between a client and a server are supported by FTP application. FTP has two basic commands for transferring a file: "get"(download) and "put"(upload). The "get" command triggers the transfer of a file from a remote server. The "put" command sends a file to a remote server. For connection-oriented transport protocols such as TCP, the model opens a new transport connection for each file transfer. The model does not include separate channels for data and control traffic—the data and control messages for a file transfer use the same TCP connection.

- ✧ Like the Transaction Mix of database applications, the Command Mix is ratio of “get” commands to the total number of commands.
- ✧ Inter-Request Time is time between subsequent file requests.
- ✧ File Size here is the maximum size of a file being transferred not the average size like the previous two.

4. File Print

Attribute	Value
Print Interarrival Time (seconds)	exponential (360)
File Size (bytes)	normal (30000, 9000000)
Symbolic Printer Name	Printer
Type of Service	Best Effort (0)

Figure 5.6 File Print Definition

A print application allows the user to initiate print jobs. TCP is the default transport protocol used for this application. Each print job creates a new TCP connection with the printer.

5. Web Browsing

The HTTP application models Web browsing. The user downloads a page from a remote server. The page contains text and graphic information (also referred to as "inline objects"). TCP is the default transport protocol for HTTP. Each HTTP page request may result in opening multiple TCP connections for transferring the contents of the inline objects embedded in the page. The number of concurrent TCP sessions is determined by the application configuration.

	Object Size (bytes)	Number of Objects (objects per page)	Location	Back-End Custom Application
constant (1000)	constant (1000)	constant (1)	HTTP Server	Not Used
Medium Image	Medium Image	constant (5)	HTTP Server	Not Used

Attribute	Value
HTTP Specification	HTTP 1.1
Page Interarrival Time (seconds)	exponential (60)
Page Properties	(...)
Server Selection	(...)
RSVP Parameters	None
Type of Service	Best Effort (0)

Attribute	Value
Initial Repeat Probability	Browse
Pages Per Server	exponential (10)

Figure 5.7 Web Browsing Definition

Http Specification is included the name of the supported HTTP version, maximum number of simultaneous TCP connections that HTTP can spawn, maximum idle time after which a connection is torn down, number of pipelined requests, and request size.

- ✧ Page Interarrival Time is time between subsequent pages that a user browses.
- ✧ Page Properties setup average object size, number of objects, and locations.
- ✧ Server Selection provides Initial Repeat Probability and Page per Server.

6. Voice over IP Call (PCM Quality)

The application set provides a voice application, which enables two clients to establish a virtual channel over which they can communicate using digitally encoded voice signals. UDP is the default transport protocol used for this application. The voice data arrives in spurts that are followed by a silence period. Encoding schemes can be specified for the voice-to-packet translation. Internally, the voice packets are sent over real-time protocol (RTP) streams. No special configuration is needed for RTP.

Attribute	Value
Silence Length (seconds)	default
Talk Spurt Length (seconds)	default
Symbolic Destination Name	Voice Destination
Encoder Scheme	G.711 (silence)
Voice Frames per Packet	1
Type of Service	Interactive Voice (6)
RSVP Parameters	None
Traffic Mix (%)	All Discrete
Signaling	None
Compression Delay (seconds)	0.02
Decompression Delay (seconds)	0.02
Conversation Environment	(...)

Figure 5.8 VoIP Call Definition

- ✧ Silence Length is for the incoming and outgoing calls along with the associated distributions. The default values are all exp(0.65).

- ✧ Talk Spurt Length, on the other hand, represents length of a talk spurt for the incoming and outgoing calls along with the associated distributions, default values as exp(0.352).
- ✧ Encoder Schemes, in effect at the client, are managed through the top-level attribute "Voice Encoder Schemes".
- ✧ Voice Frames per Packet show the number of voice frames that can be sent in a single packet.
- ✧ Traffic Mix is the proportion of traffic that should be modelled analytically instead of discretely. Higher amounts of analytic traffic decrease simulation run times but may limit the statistics that can be collected.
- ✧ Signaling specifies the method used for establishing and tearing down a voice call.
- ✧ Compression and Decompression Delays are the delay in compressing and decompression a voice packet, respectively.
- ✧ Conversation Environment assigns incoming and outgoing conversation environment.

7. Video Conferencing

A video conferencing application lets users transfer streaming video frames across the network. VCR quality video standard is using here.

Attribute	Value
Frame Interarrival Time Information	30 frames/sec
Frame Size Information (bytes)	352X240 pixels
Symbolic Destination Name	Video Destination
Type of Service	Best Effort (0)
RSVP Parameters	None
Traffic Mix (%)	All Discrete

Figure 5.9 Video Conferencing Definition

- ✧ Frame Interarrival Time Information includes Incoming and Outgoing Stream Interarrival Times, which respectively means time between frames generated within a video conferencing session from the destination and source. 30 frames/sec is when incoming and outgoing stream interarrival times are all 0.0333 seconds.
- ✧ Frame Size Information determines the incoming and outgoing stream frame size. 352*240 pixels converts to 253440 bytes.

5.1.3 Profiles Definition

In order to fully characterize an application, usage profiles must be also defined to describe how applications behave through time, *e.g.*, when they start, their repetition, or their duration. Although several applications can be present in a profile, the option was to configure each of the applications has its own usage profile defined, but all these values can be changed.

Some of the parameters, like the application start time or the repetition of the application profile itself, were defined equally in all profiles. Others, like the applications duration or repetition in the profile, were defined accordingly in each application, as shown in Table 5.1.

The 7 different profiles, running on the reference scenario and the varying scenarios, assign the activity patterns of each application used over a period of time.

Operation Mode defines how applications will start. When set to “Simultaneous”, the applications can start all at the same time.

Start Time (seconds) defines when during the simulation the profile session will start.

Table 5.1 Profile Definition

	Database Access	Email	File Transfer	File Print	Web Browsing	VoIP	Video Conference
Operation Mode	Simultaneous						
Start Time	Uniform(100,110)						
Start Time Offset	Uniform(5,10)						
Duration	End of Profile	End of Profile	End of Profile	End of Profile	End of Profile	Uniform (100,140)	Uniform (100,140)
Inter-repetition Time	-	-	-	-	-	Exponential (600)	Exponential (900)
Number of Repetition	-	-	-	-	-	Unlimited	Unlimited
Repetition Pattern	Serial						

Start Time Offset (seconds) has two interpretations based on the value specified for the "Operation Mode". As the Operation Mode is set to “Simultaneous”, this offset refers to the offset of the first instance of each application (defined in the profile), from the start of the profile.

The duration values for the Database Access, Email, File Transfer and File Print applications were set to the “End of Profile”, because, in these applications, the repetition of the applications is defined in the application definition itself as inter-request/interarrival time, so another repetition time does not needed to be defined within the application profile. Repetition Pattern specifies the pattern in which this application repeats serially or concurrently.

5.2 Scenarios Variations

5.2.1 Mobiles Variations

Taking the base scenario, some more scenarios were defined by varying number of

users. As every user connects to one MAP or involves in one BSS, and every MAP connects to one gateway or involves in one ESS, 3 kinds of scheme can be considered when the user number increases.

① Build more gateways and keep every ESS (corresponding to a gateway) configuration the same, as shown in Figure 5.10. Table 5.2 lists every variation in this scheme.

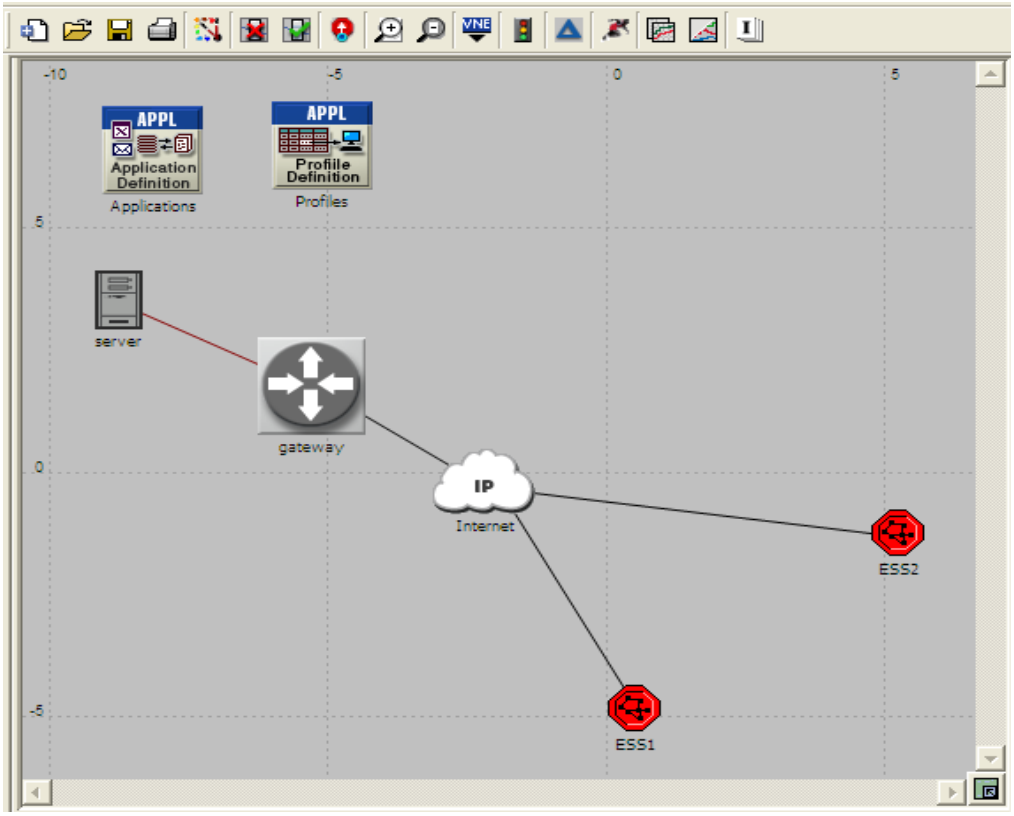


Figure 5.10 An Example of Multiple ESSs Scenarios

Table 5.2 Mobiles Variations (Scheme 1)

Scenarios	Number of ESS	Number of MAPs per ESS	Number of STAs per MAP	Total Number of STAs
1ESS2MAPs4STAs	1	2	4	8
2ESS2MAPs4STAs	2	2	4	16
4ESS2MAPs4STAs	4	2	4	32
6ESS2MAPs4STAs	6	2	4	48
8ESS2MAPs4STAs	8	2	4	64

② Increase the number of BSS (MAPs) in one ESS and keep every BSS configuration the same. Figure 5.11 shows the configuration of Scenario 1ESS4MAPs4STAs as an example, while Table 5.3 lists all the possibility of this scheme of variation.

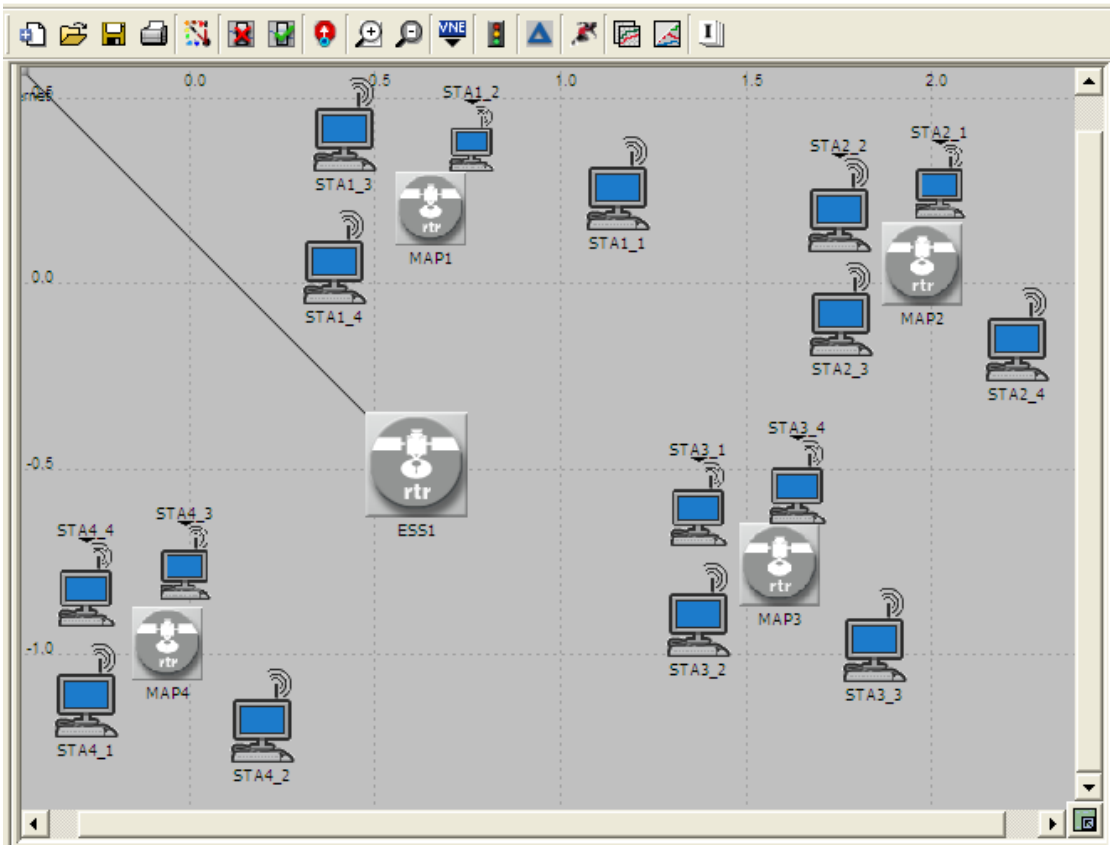


Figure 5.11 A Scenario with 4 MAPs per ESS

Table 5.3 Mobiles Variations (Scheme 2)

Scenarios	Number of ESS	Number of MAPs per ESS	Number of STAs per MAP	Total Number of STAs
1ESS4MAPs4STAs	1	4	4	16
1ESS8MAPs4STAs	1	8	4	32
2ESS4MAPs4STAs	2	4	4	32
2ESS8MAPs4STAs	2	8	4	64

③ Simply add the new coming users to existing MAPs. In Figure 5.12, a scenario with more STAs getting access to each MAP is shown.

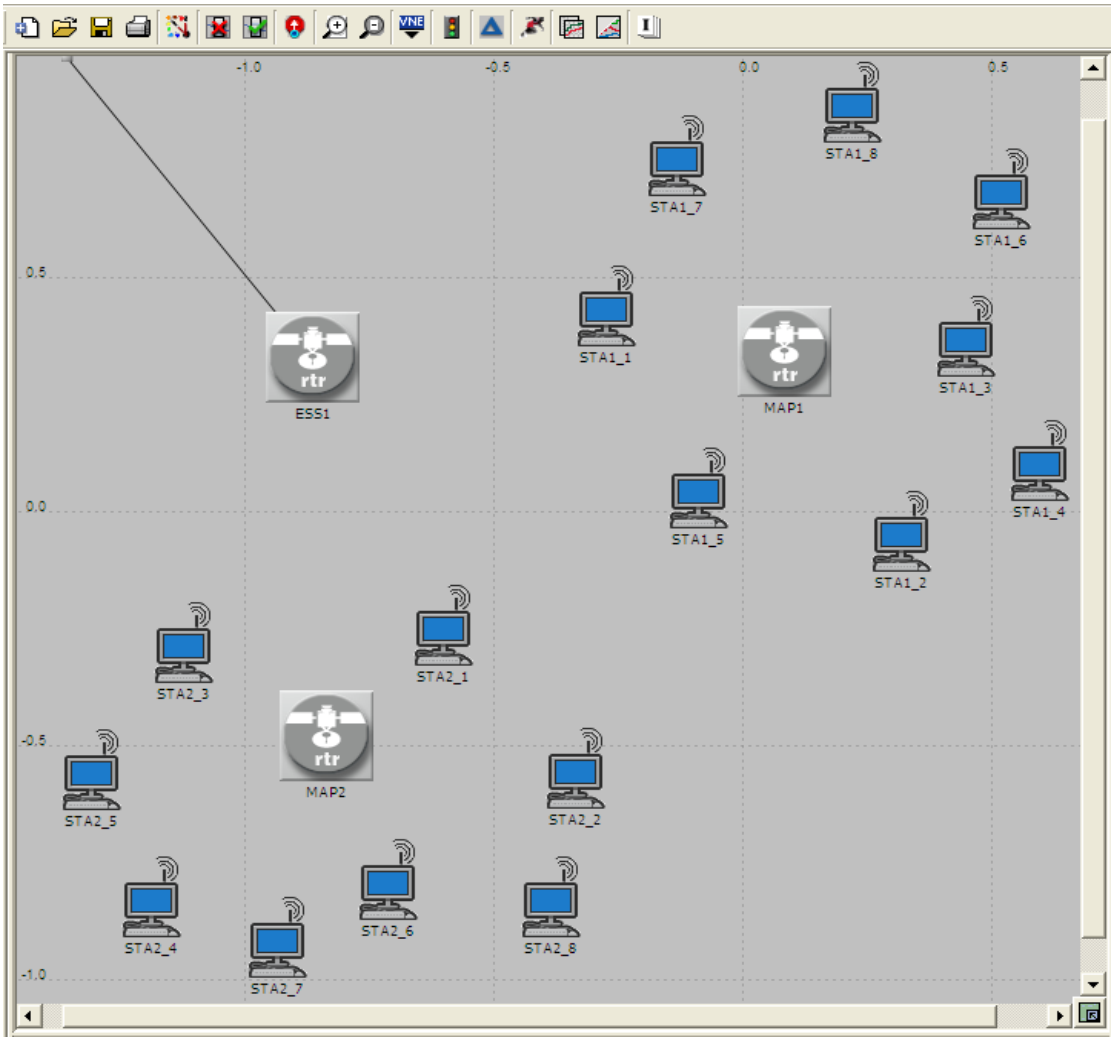


Figure 5.12 Scenario with 8 STAs per MAP

Table 5.4 Mobiles Variations (Scheme 3)

Scenarios	Number of ESS	Number of MAPs per ESS	Number of STAs per MAP	Total Number of STAs
1ESS2MAPs8STAs	1	2	8	16
1ESS2MAPs16STAs	1	2	16	32
2ESS2MAPs8STAs	2	2	8	32
2ESS2MAPs16STAs	2	2	16	64
2ESS4MAPs8STAs	2	4	8	64

5.2.2 Profiles Variations

In order to evaluate the impact of varying profiles/services supported/used in the network, all the 7 applications configured in the last section need to be divided into two groups in terms of real-time demanding. Database Access, Email, File Transfer, File Print, Web Browsing are non real-time applications; VoIP and Video Conference are real-time ones.

As the profiles variations can be assigned in Applications>Application: Supported Profiles of models' attributes, the relative weights of individual applications are varied in the following 6 scenario types:

- NRTM: Non Real-Time Maximum
- NRTC: Non Real-Time Centric
- NRT: Non Real-Time
- RT: Real-Time
- RTC: Real-Time Centric
- RTM: Real-Time Maximum

Table 5.5 shows the distribution of the 6 profiles.

Table 5.5 Profiles Distribution

	Database Access	Email	File Transfer	File Print	Web Browsing	VoIP	Video Conference
NRTM	√	√	√	√	√	-	-
NRTC	-	√	-	√	√	-	-
NRT	-	√	-	√	√	√	-
RT	-	√	-	-	√	√	-
RTC	-	-	-	-	√	√	-
RTM	-	-	-	-	-	√	√

√ supported applications - non supported applications

5.2.3 EDCA Parameters

As Section 3.3 presented, IEEE802.11e standard proposed a new mechanism EDCA in order to provide QoS guarantees to applications. The setting of these parameters is shown as Figure 5.13.

Attribute	Value
Large Packet Processing	Drop
⊞ PCF Parameters	Disabled
⊟ HCF Parameters	(...)
Status	Supported
⊟ EDCA Parameters	(...)
⊟ Access Category Parameters	(...)
⊟ Voice	(...)
CWmin	(PHY CWmin + 1) / 4 - 1
CWmax	(PHY CWmin + 1) / 2 - 1
AIFSN	2
⊞ TXOP Limits	Default
⊞ Video	Default
⊞ Best Effort	Default
⊟ Background	(...)
CWmin	PHY CWmin
CWmax	PHY CWmax
AIFSN	7
⊞ TXOP Limits	Default
⊞ Traffic Category Parameters (8 R...	Default
Block ACK Capability	Supported
⊞ AP Specific Parameters	Default

Figure 5.13 EDCA Parameters Setting

Based on the three major variations of scenarios in this simulation model presented above, the rest of this thesis would analyze the performances of these scenarios according to the results from simulations.

CHAPTER 6 ANALYSIS OF SIMULATION RESULTS

The results collected from simulations with OPNET modeler are presented and analyzed in this chapter. The performance of various scenarios are compared and analyzed in terms of throughput, average delay, and data drop to measure the QoS of Multihoop Wireless Network.

6.1 An Overview

Before starting the analysis, it is important to underline a limitation of the IEEE802.11 model suite used for simulation in OPNET. The transmission data rate used by WLAN nodes is static through the entire simulation time, in other words, the model does not implement the rate adaptation feature specified in the standard, which is 11 Mbps. This limitation, together with the use of the free-space propagation model, presents a limitation in the model implementation using OPNET.

In most cases, a large number of system variables and metrics are available for collection in OPNET, which offer two types of user-defined statistics: local and global, in general. Local statistics are ideal for reporting activity that is private to a particular node in the system model. However, in the case of network performance evaluation, we are more interested in obtaining quantitative information about the system as a whole. Therefore, the present study works on the data collected from global statistics.

Several scenarios were defined in the previous chapter for a varying number of users in the network, as well as for various usage profiles, the existence of QoS (EDCA mechanism enabled/disabled) being considered in parallel. To analyze the QoS of the network, we would mainly focus on global average delay and throughput. Meanwhile,

the data drop, retransmission attempts, the standard deviations and so on are considered significant in some cases.

As for the simulation duration, the objective is to make one hour simulations for each scenario, as it is considered that one hour is a large enough period to have meaningful and stable results in the network. It is noted that the simulation time set in simulations is not equal to the CPU run time or real clock time as explained in Section 4.2.2.

As expected the most demanding scenarios, in terms of resources usage, are the ones with more streaming service (more bandwidth demanding), and the ones with more real-time applications (more delay stringent and more uplink bandwidth required). Looking, for instance, at the global average throughput and global average delay, these most resource demanding scenarios can be identified.

6.2 Global Average Throughput

From Global Average Throughput of the Scheme 1 scenarios with mobiles variations shown in Figure 6.1, we can see the clear increase with the number of ESS, with similar results in all of the scenarios, but with smaller variations when real-time profiles are used or served. On the other hand, the network load is very high in most of the scenarios, with the maximum observed average global throughput of 44Mbps approximately, at the cost of building 7 more base stations which is represented in the form of WLAN gateways. To keep the high throughput of the network when the users increase, we can also take some other schemes into account.

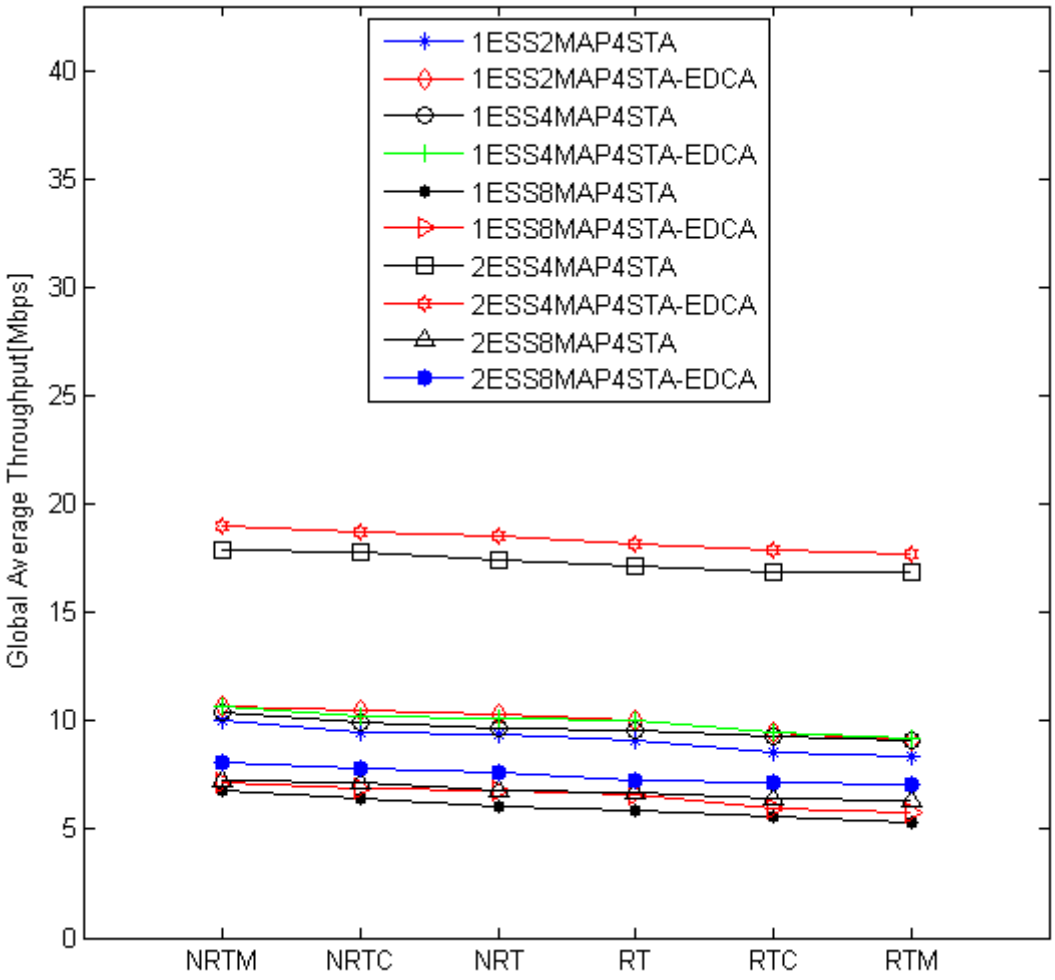


Figure 6.1 Global Average Throughput of Scheme 1 Scenarios

From the results of the scenarios of Scheme 2 shown in Figure 6.2 and Figure 6.3, we can see that doubling the number of MAPs does not make remarkable difference to the throughput. But the average throughput drops sharply from above 10Mbps to below 7Mbps, when the number of MAPs is doubled to 8 per ESS. Likewise, for 2ESS situations, 8 MAPs per ESS brings significant drop to the global average throughput. This indicates that the scheme of 8 MAPs per ESS is not a good choice for the network QoS in terms of global throughput. For this reason the scenarios with up to 8 MAPs per ESS would not be analyzed further.

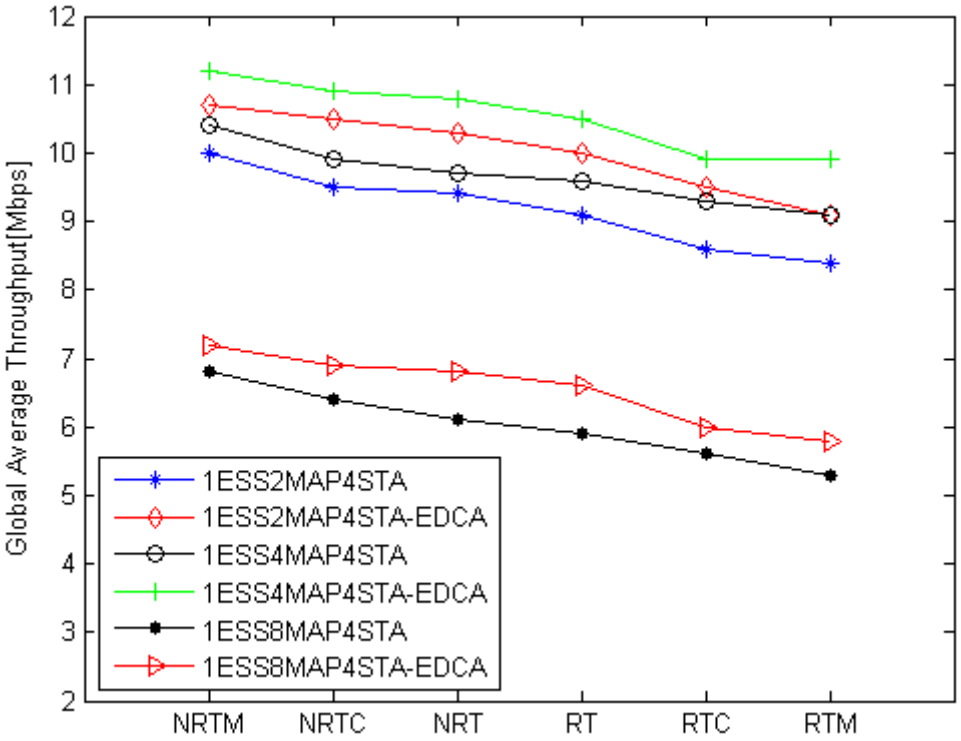


Figure 6.2 Global Average Throughput of Scenarios

with variable MAP, 1 ESS and 4 STAs

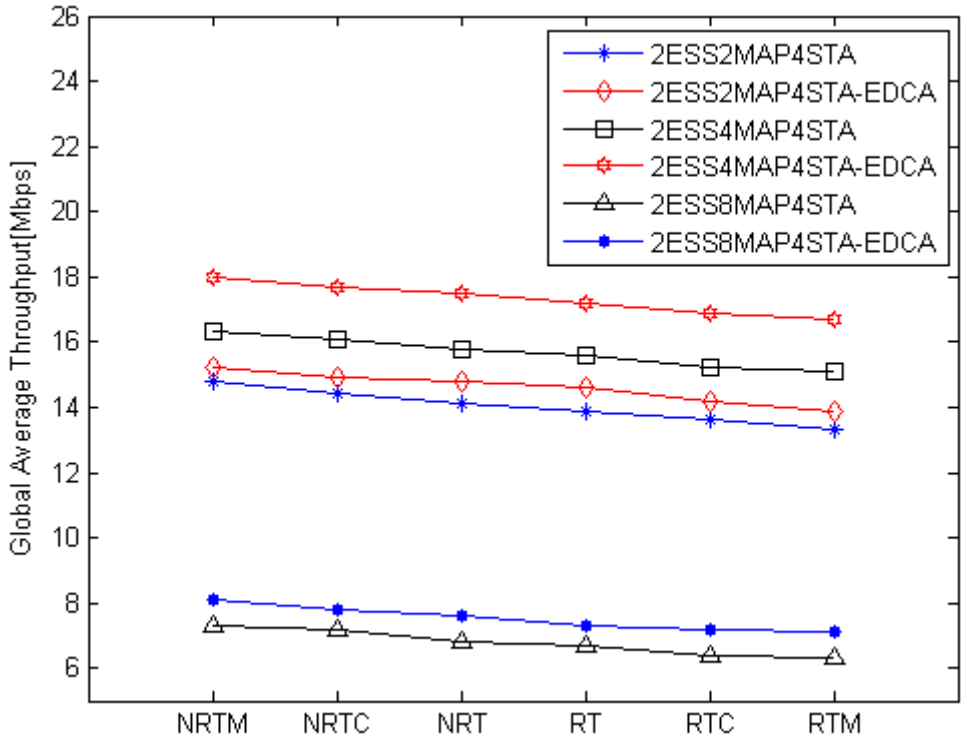


Figure 6.3 Global Average Throughput of Scenarios

with variable MAPs, 2 ESSs and 4 STAs

As for the third scheme of scenarios, when the number of workstations increases, the global throughput drops only by several hundred Kbps, which differs slightly compared to the variations in the second scheme, shown in Figure 6.4. Notice Figure 6.5, when the total number of workstations increases up to 64, 2EES4MAP8STA has the highest value of global throughput, without concern of building 8 base stations.

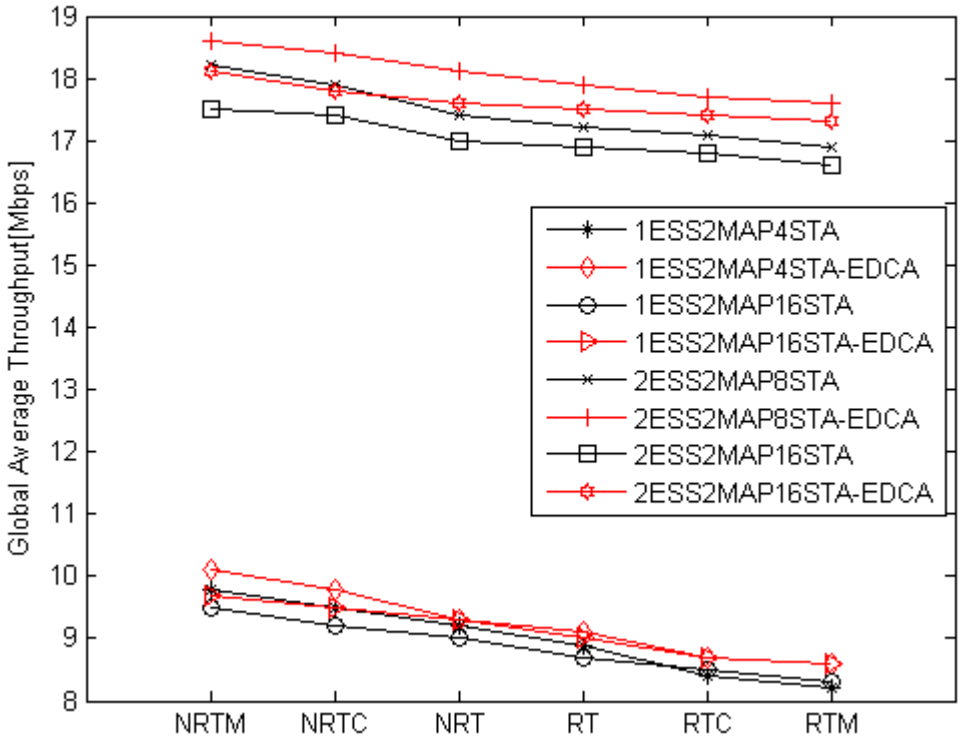


Figure 6.4 Global Average Throughput of Scenarios

with variable ESSs and STAs, 2 MAPs

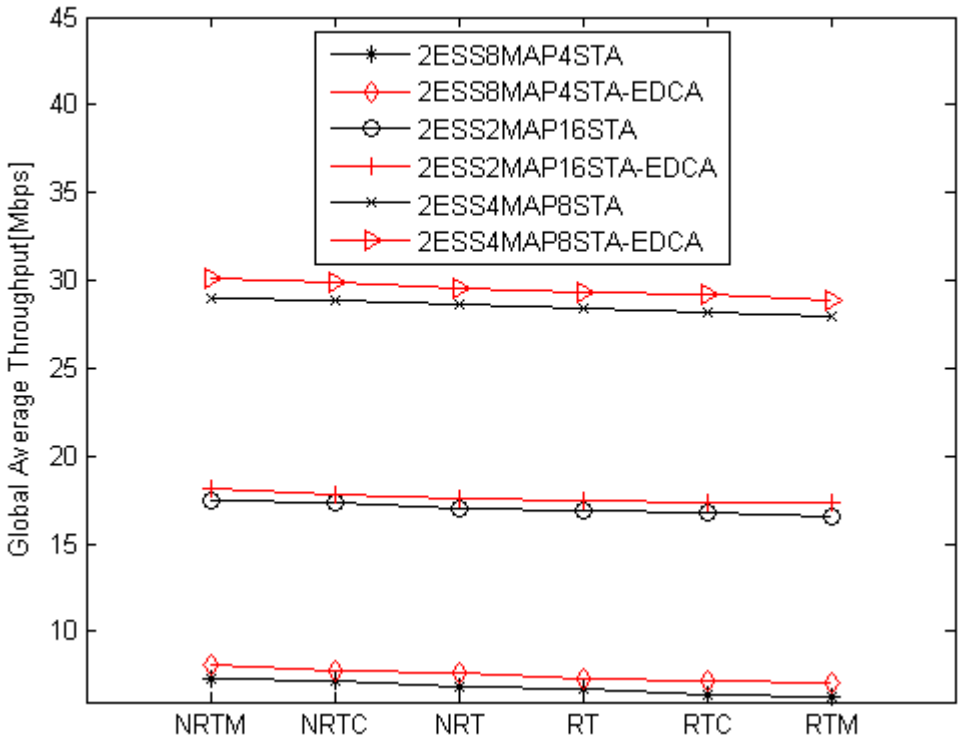


Figure 6.5 Global Average Throughput of Scenarios

with variable MAPs and STAs, 2 ESSs

However, it is noticed that no clear distinction between the EDCA enabled/disabled scenarios occurs, even when increasing the number of workstations. Although this might lead you to think that there are no throughput gains when enabling the EDCA mechanism, the overall tendency in all scenarios shows that EDCA mechanisms can, in fact, improve the throughput of the network, at least in the majority of the situations.

6.3 Global Average Delay

As global average delay shows end-to-end delay for all the packets received at the MAC layer of all the WLAN stations, it plays a very important role in evaluating performance of the network.

For the first scheme of scenarios with EDCA disable, we can see the trend of global delay from Figure 6.6. Except for the scenarios with ESS number no more than 4, the delay values go over 100ms even reach almost 150ms in the situations of using real-time maximum. The delay values as such high could lead to severe performance degradation, which is the last thing we would let it happen. Another phenomenon from this figure is that the more real-time profiles are used, the more delay it brings to the network.

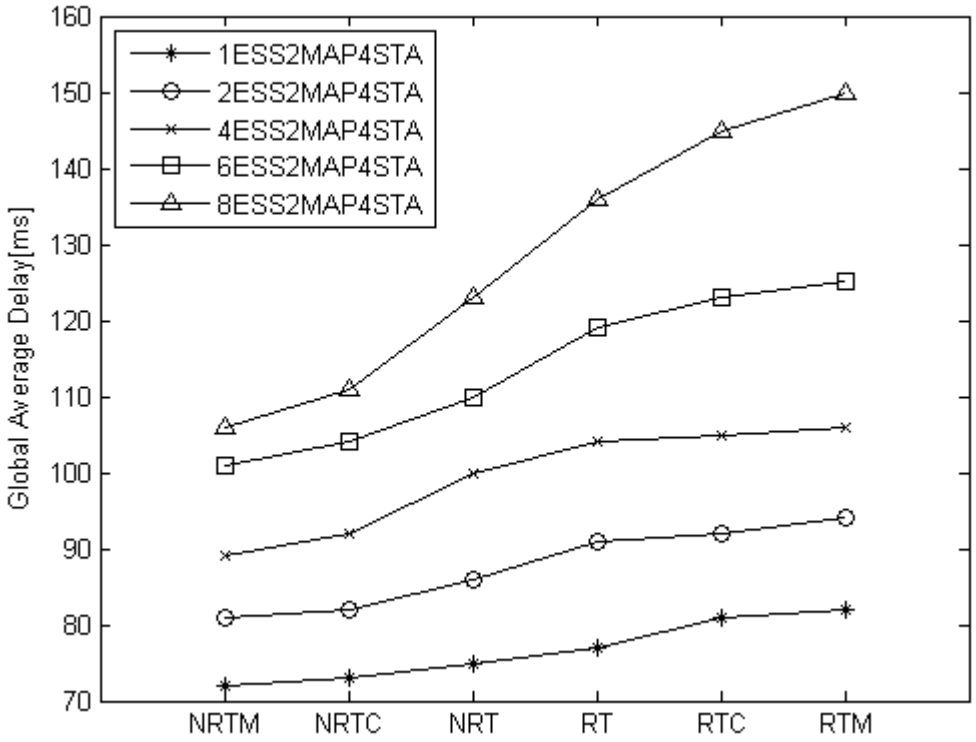


Figure 6.6 Global Average Delay of Scheme 1 Scenarios without EDCA

However, the delay value obtains significant reduction in every scenario with EDCA enabled, as Figure 6.7 shows. The delay reductions of the RTM scenarios are higher than the others with non real-time applications. This trend indicates that EDCA mechanism improve the network QoS very well especially when there are more applications in higher priority and demanding.

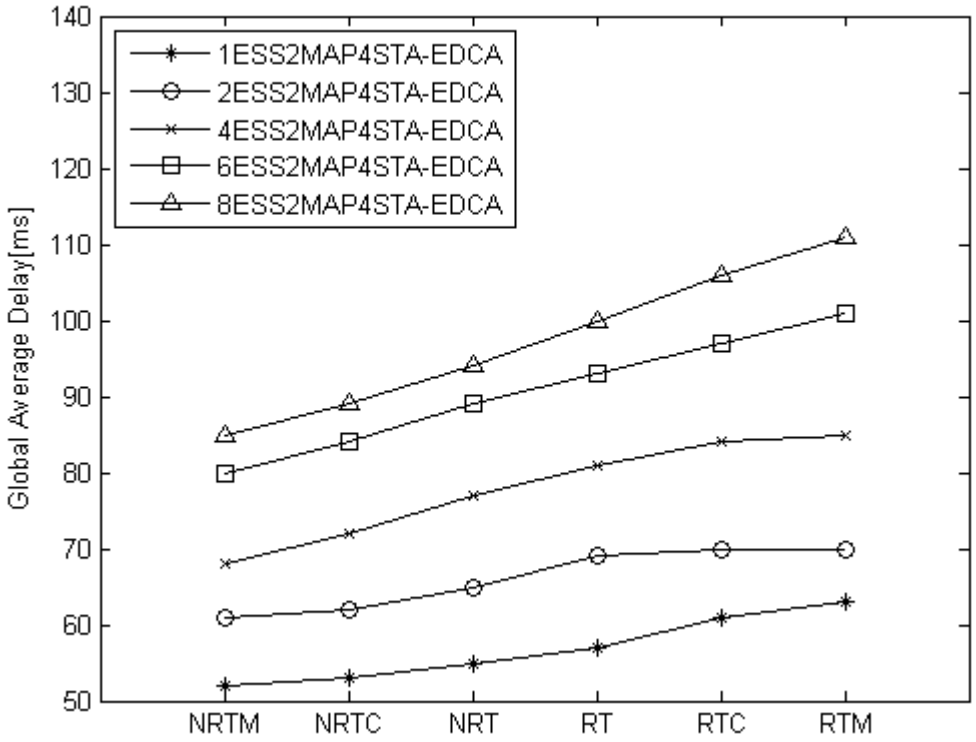


Figure 6.7 Global Average Delay of Scheme 1 Scenarios with EDCA

Again in Figure 6.7, we notice the remarkably high performance degradation for the scenarios with 6 and 8 ESS. The delay values without EDCA are already very large, and on top of that, the delay reduction in the RTM scenario is much smaller comparing to the scenarios with less ESS. Therefore, we come to a conclusion that 4 is the maximum number of ESS that the network can sustain, while still being able to maintain acceptable behavior. For this reason only the scenarios with 1 and 2 ESS are further analyzed. But a fact should be pointed out that the maximum number of ESS depends on the scenario that is being considered. Any change in the parameters of the scenario would cause totally different results.

On the other hand, a very large variation appears in the global average delay of the second group of scenarios when doubling the number of MAPs from 2 to 8, as shown in Figure 6.8 and Figure 6.9.

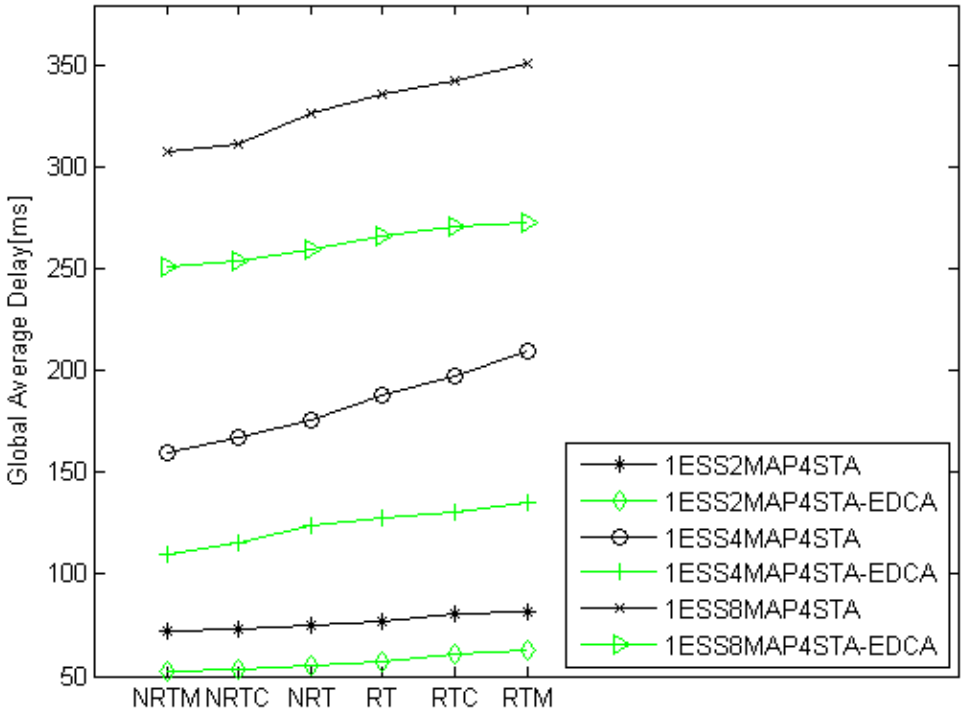


Figure 6.8 Global Average Delay of Scenarios

with variable MAPs, 1 ESS and 4 STAs

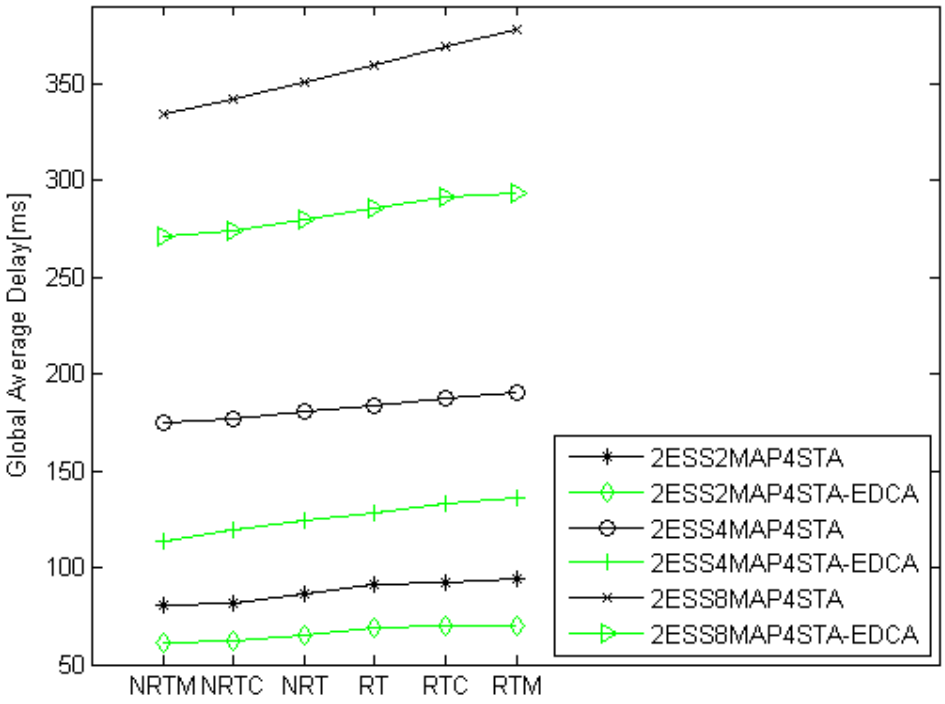


Figure 6.9 Global Average Delay of Scenarios

with variable MAPs, 2 ESSs and 4 STAs

Notice that, EDCA mechanism helps the scenarios reduce delay dramatically, especially when the real-time applications are highly demanded. Even though the remarkable improvement has been brought by EDCA, the global average delays of scenarios with 8 MAPs per ESS are out of range. However, EDCA decreases delay for scenarios with 4 MAPs per ESS to a reasonable level.

Like the impact of increasing workstations number in a BSS on the global throughput, the delays only increase a little bit when the number of workstations doubles to 8, as despicted Figure 6.10 and Figure 6.11. However, when the number doubles to 16, the situation turns up to be extremely unacceptable, regardless of the improvement from EDCA.

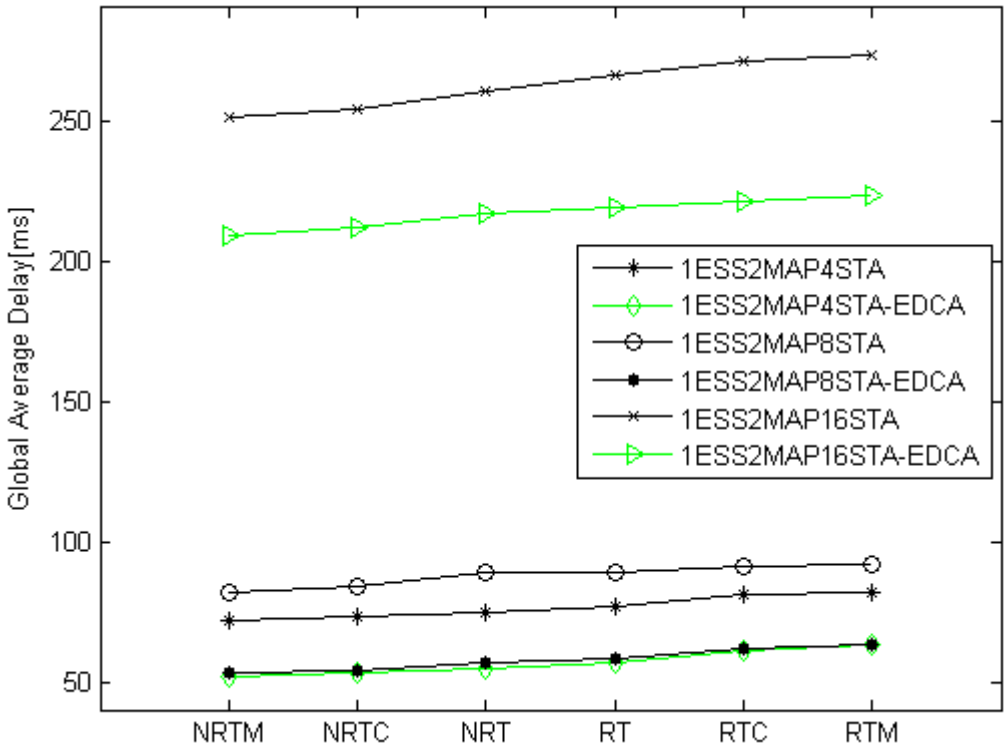


Figure 6.10 Global Average Delay of Scenarios with variable STAs, 1 ESS and 2 MAPs

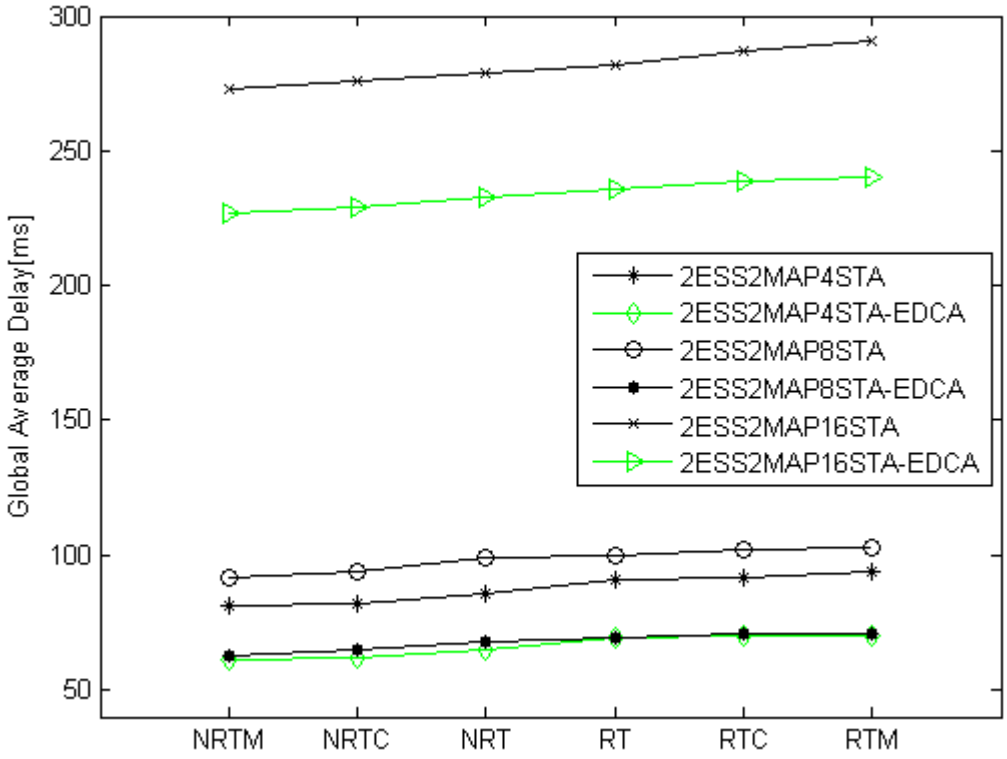


Figure 6.11 Global Average Delay of Scenarios

with variable STAs, 2 ESSs and 2 MAPs

6.4 Dropped Data

To analyze the global performance of the network, we must consider not only the throughput and delay, but also the number of retransmission attempts and the dropped data, in order to check if eventual performance gains in the throughput or delay caused by EDCA mechanisms are not causing undesired effects. Dropped data (by either buffer overflow or retry threshold exceeded) leads to performance losses, thus, increasing the number of retransmission attempts.

In Figure 6.12, dropped data values of the scenarios except for the ones with 8 MAPs/ESS and 16 STAs/MAP are presented. In the scenarios with fewer total number of workstations, the values are minor in all situations, but when the total number increases, a rise in dropped data, especially for the most resource demanding scenarios, appears. Again, it can be seen that the EDCA enabled scenarios have much

better performance with values always below 1 Kbps, while the EDCA disable scenarios almost reach 12 Kbps. It may be noticed that the dropped data, similar to the global throughput and global delay, just take MAC layer issues into account, i.e., eventual retransmissions made/caused by the higher layers, e.g., application layer retransmissions are not considered in these values.

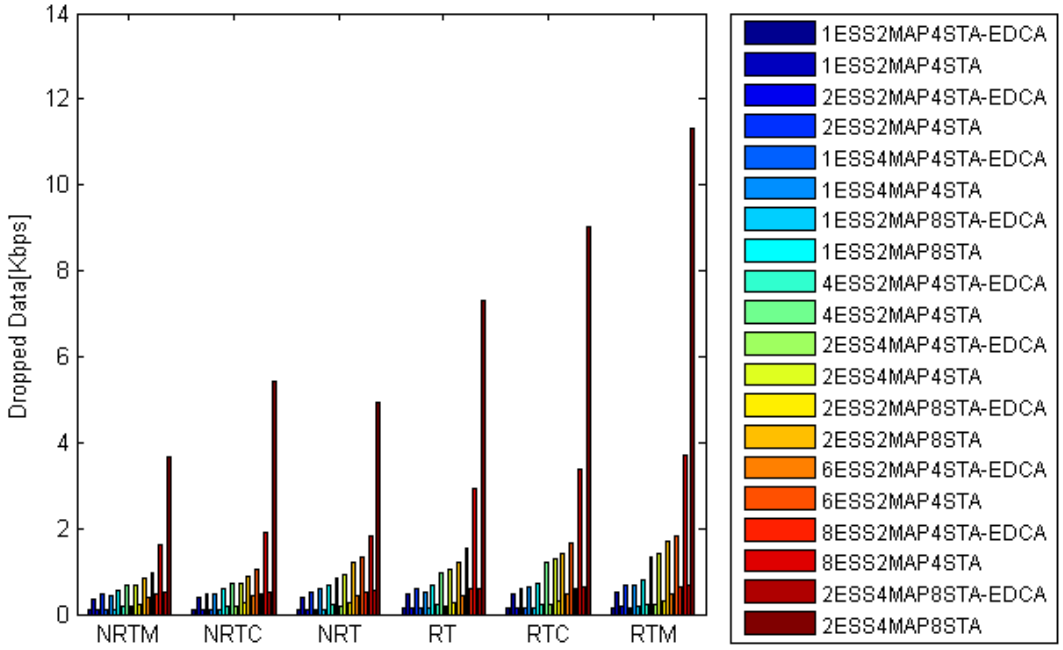


Figure 6.12 Dropped Data by Scenario

From the results shown so far, it may be concluded that EDCA mechanism can improve the overall performance of the network, enabling smaller global delays, while still being able to achieve higher maximum throughputs, in the majority of scenarios. EDCA mechanism aim at improving the priority of real and near real-time applications, hence, as expected, larger gains are observed in real-time centric/maximum scenarios. These improvements are more tangible when the network load increases, where better results are achieved. From the analysis of three different schemes for devices configuration, we noticed that in order to keep the QoS of network stable, we can increase the number of ESS for the newly joining users, but

keep the number of MAPs per ESS no more than 4 and number of workstations per MAP no more than 8.

CHAPTER 7

CONCLUSIONS AND FUTURE STUDY

This chapter is to go through the works done to fulfill the objective explained at the beginning of the thesis, draw the conclusions from the finding in the study, and eventually point out future directions based on the present study.

7.1 Conclusions

In the context of the emerging demand on Multihop Wireless Networks investigation, a number of studies could be found in literature addressing several important issues, such as network management, security and cross-layer design. An overview on them was provided, investigating some remarkable related studies, along with a description of some commercial projects providing multihop networking solutions. In parallel with these research efforts, a lot of works had been done in the last decade on performance evaluations.

Although every wireless technology had claimed its advantage in certain circumstances, most consumer electronics supported WiFi technology built on the IEEE802.11 standards. In general, the basic principle of IEEE802.11 standards, as well as network topology and station services, was also studied in the first stage of the work. To provide the background for wireless network technology, IEEE802.11 PHY and MAC protocols had been described in detail.

Considering all the features of wireless network simulators available, OPNET modeler was chosen for simulating the MWNs performance in this study. After the introduction of modeler architecture, the principle of discrete event simulation used for network modeling and simulation was explained.

To study the impact of several parameters on MWNs performance, a specific implementation model was then proposed, defined and developed with several degrees of variations, including the number of mobiles, variations of profiles and QoS support (EDCA mechanism). The number of mobiles also varied in three schemes. In the first scheme, the number of ESS increased from 1 to 8; the second scheme changed the number of MAPs per ESS from 2 to 8; the number of STAs per MAP in the third scheme varied from 4 to 16. In parallel, the variations of profiles had been deployed in every scenario with different mobiles number. The 6 profiles were combinations of non real-time applications (Database Access, Email, File Transfer, File Print and Web Browsing) and real-time ones (VoIP and Video Conference). Another condition, which has been considered in this study, was EDCA mechanism introduced as the IEEE802.11e standard.

A comprehensive evaluation of performance of MWNs under the IEEE802.11 standard was conducted based on simulation in OPNET modeler. Global average throughput, global average delay and dropped data were collected from 168 simulation runs for the system performance analysis.

Taking an overview of the results, we are able to conclude that the global average throughput and delay are largely depended on the configuration of devices, especially when the system had high demand on real time applications. For keeping a stable system performance, the maximum number of MAPs per ESS is 4 while the number of workstations per MAP should be no more than 8. This finding could be used for optimizing MWNs configuration.

In general, EDCA mechanism brought more remarkable improvement in global

average delay, compared to the minor difference in global average throughput. Another useful conclusion obtained from the results was that EDCA mechanism could effectively increase global average throughput, reduce global average delay and minimize retransmission probability for the most resource demanding systems.

7.2 Future Study

Apart from the usage of network configuration mentioned before, future works can also be carried out in some other directions.

With regards to improving the MWN architecture, the data collected from the simulations can be used as inputs to more complex studies, such as interworking heterogeneous wireless networks, to implement widespread deployment of wireless networking domains. The MAPs could be designed for supporting not only WiFi devices under standards of IEEE802.11, but also some other protocols such as WiMAX, Zigbee, etc.

In the case of heterogeneous MWNs design, considerations may certainly be taken on protocols for network management and security issues. The previous mobility and power management schemes developed for other types of networks are similar in terms of principle, so they could be used for MWNs in a certain way. However, the network management schemes achieved using each of them varies from application to application. Therefore, efforts must be made to determine the protocol of network management for each particular MWN. As we can see from this study, there is no central authority or trusted server to manage security issues. How to detect attacks, monitor service disruption and respond rapidly to attacks will be a challenging topic for future practical projects.

REFERENCES

[1] Akyildiz, I. F., & Wang, X. (2008, 3). Cross-Layer Design in Wireless Mesh Networks. *IEEE Transactions on Vehicular Technology* , pp. 1061-1076.

[2] Akyildiz, I., Wang, X., & Wang, W. (2005). Wireless mesh networks: a survey. *Computer Networks* , pp. 445-487.

[3] Ancillotti, E., Bruno, R., & Conti, M. (2008). *Experimentation and Performance Evaluation of Rate Adaptation Algorighms in Wireless Mesh Networks.* Vancouver,BC,Canada: ACM.

[4] Banks, J. (1998). Discrete-Event Simulation of Computer and Communication Systems. In J. Banks, *Handbook of Simulation* (pp. 659-676). New York: Wiley.

[5] Borisov, N., Goldberg, I., & Wagner, D. (2001). Intercepting mobile communications: the insecurity of 802.11. *International Conference on Mobile Computing and Networking*, (pp. 180-189). Rome,Italy.

[6] Breslau, L., & Shenker, S. (1998). Best-effort versus reservations: A simple comparative analysis. *ACM SIGCOMM Conference* (pp. 3-16). Vancouver: ACM.

[7] Camp, J., Knightly, E., & Reed, W. (2006). Developing and Deploying Multihop Wireless Networks for Low-Income Communities. *Journal of Urban Technology* , pp. 129-137.

[8] Cavin, D., Sasson, Y., & Schiper, A. (2002). On the Accuracy of MANET Simulators. *POMC'02* (pp. 38-43). Toulouse: ACM.

[9] Cheng, X., Mohapatra, P., Lee, S.-J., & Banerjee, S. (2008). Performance Evaluation of Video Streaming in Multihop Wireless Mesh Networks. *International Workshop on Network and Operating System Support for Digital Audio and Vide* (pp. 57-62). Braunschweig, Germany: ACM.

[10] Ernst, J. B., & Denko, M. K. (2010, 2). The Design and Evaluation of Fair

Scheduling in Wireless Mesh Networks. *Journal of Computer and System Sciences* .

[11] Forouzan, B. (2001). *Data Communications and Networking.* Boston: McGraw-Hill.

[12] Frey, T. (2006, 12). *Driving Forces.* Retrieved 12 2009, from FuturistSpeaker: http://www.futuristspeaker.com/2006/12/driving-forces/

[13] Fu, Z., Zerfos, P., Luo, H., Lu, S., Zhang, L., & Cerla, M. (2003). *The Impact of Multihop Wireless Channel on TCP Throughput and Loss.* Los Angeles.

[14] Gambiroza, V., Sadeghi, B., & Knightly, E. W. (2004). End-to-End Performance and Fairness in Multihop Wireless Backhaul Networks. *MobiCom'04* (pp. 287-301). Philadelphia, Pennsylvania,USA: ACM.

[15] Garetto, M., Shi, J., & Knightly, E. (2005). Modeling media access in embedded two-flow topologies of multi-hop wireless networks. *International Conference on Mobile Computing and Networking* (pp. 200-214). New York: ACM.

[16] Hamidian, A., Palazzi, C. E., Chong, T. Y., Nanarro, J. M., Korner, U., & Gerla, M. (2009). *Deployment and Evaluation of a Wireless Mesh Network.*

[17] Han, B., Jia, W., & Lin, L. (2006, 10). Performance Evaluation of Scheduling in IEEE802.16 Based Wireless Mesh Networks. *Computer Communications* , pp. 782-792.

[18] Heidemann, J., Bulusu, N., Elson, J., & Intanagonwiwat, C. (2001). Effects of Detail in Wireless Network Simulation. *SCS Communication Networks and Distribued Systems Modeling and Simulation Conference.* USC/Information Sciences Institute.

[19] Heidemann, J., Mills, K., & Kumar, S. (2001). Expanding Confidence in Network Simulation. *IEEE Communications Magazine* , 58-63.

[20] Ho, C. C., Ramachandran, K. N., Almeroth, K. C., & Belding-Royer, E. M. (2004). A Scalable Framework for Wireless Network Monitoring. *Proceedings of the 2nd ACM international workshop*, (pp. 93-101). Philadelphia.

[21] Hu, Y.-C., & Rerrig, A. (2005). Ariadne: A Secure On-Demand Routing

Protocol for Ad Hoc Networks. *Wireless Networks* , pp. 21-38.

[22] Huang, D.-W., Lin, P., & Gan, C.-H. (2008, 5). Design and Performance Study for a Mobility Management Mechanism (WMM) Using Location Cache for Wireless Mesh Networks. *IEEE Transactions on Mocile Computing* , pp. 546-556.

[23] IEEE LAN/MAN Standards Committee. (1999). IEEE Std 802.11a-1999 High-speed Physical Layer in the 5 GHz Band. NY: IEEE.

[24] IEEE LAN/MAN Standards Committee. (1999). *IEEE Std 802.11b-1999 Higher-Speed Physical Layer Extension in the 2.4 GHz Band.* NY: IEEE.

[25] IEEE LAN/MAN Standards Committee. (2001). *IEEE Std 802.11d-2001 Amendment 3: Specification for operation in additional regulatory domains.* NY: IEEE.

[26] IEEE LAN/MAN Standards Committee. (2005). *IEEE Std 802.11e™-2005 Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements.* NY: IEEE.

[27] IEEE LAN/MAN Standards Committee. (2003). *IEEE Std 802.11g™-2003 Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band.* NY: IEEE.

[28] IEEE LAN/MAN Standards Committee. (2003). *IEEE Std 802.11h™-2003 Amendment 5: Spectrum and Transmit Power Management Extensions in the 5 GHz band in Europe.* NY: IEEE.

[29] IEEE LAN/MAN Standards Committee. (2004). *IEEE Std 802.11i™-2004 Amendment 6: Medium Access Control (MAC) Security Enhancements.* NY: IEEE.

[30] IEEE LAN/MAN Standards Committee. (2004). *IEEE Std 802.11j™-2004 Amendment 7: 4.9 GHz–5 GHz Operation in Japan.* NY: IEEE.

[31] IEEE LAN/MAN Standards Committee. (2010, 2 26). *Official IEEE 802.11 Working Group Project Timelines*. Retrieved 2 26, 2010, from IEEE 802.11 Local Area Networks: http://www.ieee802.org/11/Reports/802.11_Timelines.htm

[32] IEEE LAN/MAN Standards Committee. (2007). *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.* New York: IEEE.

ISI. (2009). *The Network Simulator NS-2:Tips and Statistical Data for Running Large Simulations in NS.* Retrieved from http://www.isi.edu/nsnam/ns/ns-lagesim.html

[33] Jacome, M., & Catthoor, F. (2003, 8). Special issue on power-aware embedded computing. *ACM Transactions on Embedded Computing Systems* , pp. 251-254.

[34] Jein, K., Padhye, J., Padmanabhan, V., & Qiu, L. (2005). Impact of Interference on Multi-Hop Wireless Network Performance. *Wireless Networks* , 471-487.

[35] Jones, C., Sivalingam, K. M., Agraval, P., & Chen, J. C. (2001, 7). A Survey of Energy Efficient Network Protocols for Wireless Networks. *Wireless Networks* , pp. 343-358.

[36] Karl, H. (2003). *An Overview of Energy-Efficiency Techniques for Mobile Communication Systems.* Berlin: Technical University Berlin.

[37] Khreishah, A., Wang, C.-C., & Shhroff, N. B. (2009, 6). Cross-Layer Optimization for Wireless Multihop Networks with Pairwise Intersession Network Coding. *IEEE Journal on Selected Areas in Communications* , pp. 606-621.

[38] Klues, K. (2006). *Power Management.* Citeseer.

[39] Liu, Y., Hoshyar, R., Yang, X., & Tafazolli, R. (2006, 11). Integrated Radio Resource Allocation for Multihop Cellular Networks With Fixed Relay Stations. *IEEE Journal* , pp. 2137-2146.

[40] Marina, M., & Das, S. (2005). A Topology Control Approach for Utilizing Multiple Channels in Multi-Radio Wireless Mesh Networks. *2nd International Conference on Broadband*, (pp. 381-390).

[41] NS2. (2009). *NS2 Official Website*. Retrieved from http://www.isi.edu/nsnam/ns/

[42] NS3. (2009). *NS3 Official Website*. Retrieved from http://www.nsnam.isi.edu/nsnam/index.php/Main_Page

[43] OMNeT. (2009). *OMNeT++ Official Website*. Retrieved from

http://www.omnetpp.org/

[44] OPNET Technologies,Inc. (2010). Retrieved from OPNET Technologies: http://www.opnet.com/

[45] OPNET Technologies,Inc. (2009). *OPNET Modeler 15.0 Documentation.* Retrieved from http://www.opnet.com/products/modeler/home.html

[46] Oyman, O., Laneman, N., & Sandhu, S. (2006, 12 29). Multihop Relaying for Broadband Wireless Mesh Networks: From Theory to Practice. *IEEE Comm.Mag* .

[47] Pabst, R., Walke, B., & Schultz, D. (2004). Relay-Based Deployment Concepts for Wireless and Mobile Broadband Radio. *Wireless World Research Forum* (pp. 80-89). IEEE.

[48] Pawlikowski, K., Jeong, H.-D., & Lee, J.-S. (2002). On the Credibility of Simulatin Studies of Telecommunication networks. *IEEE Communications Magazine* , 132-139.

[49] QiuLili, ChandraRanveer, JainKamal, & MahdianMohammad. (2004). Optimizing the Placement of Integration Points in Multi-hop Wireless Networks. IEEE ICNP.

[50] Raczynski, S. (2006). Discrete and Combined Simulation. In S. Raczynski, *Modeling and Simulation : the computer science of illusion* (pp. 67-110). Hoboken, NJ: Wiley.

[51] Raghavendra, R., Acharya, P., Beling, E. M., & Almeroth, K. C. (2009). MeshMon: a multi-tiered framework for wireless mesh networkmonitoring. *Proceedings of the 2009 MobiHoc S3 workshop*, (pp. 45-48). New Orleans, Louisiana, USA .

[52] Ramchandran, K., Belding, E., Almeroth, K., & Buddhikot, M. (2006). *Interference-Aware Chnannel Assignment in Multi-Radio Wireless Mesh Networks.* IEEE ICNP.

[53] Sailhan, F., Fallon, l., Quinn, K., Farrell, P., Collins, S., Parker, D., et al. (2007). *Wireless Mesh Network Monitoring: Design, Implementation and Experiments.*

IEEE DANMS.

[54] Schilling, B. (2005, 1 31). Qualitative Comparison of Network Simulation Tools. Stuttgart, Germany.

[55] Shnayder, V., Hempstead, M., Chen, B.-r., Allen, G. W., & Welsh, M. (2004). *Simulation the Power Consumption of Large-Scale Sensor Network Applications.* ACM.

[56] Stallings, W. (2005). *Wireless Communications & Networks.* Upper Saddle River, NJ: Pearson Prentice Hall.

[57] Tsaoussidis, V., & Badr, H. (2000). TCP-Probing: Towards an Error Control Schema with Energy and Throughput Performance Gains. *ICNP'00.* Osaka, Japan : Citeseer.

[58] Wang, X. (2008). Wireless Mesh Networks. *Journal of Telemedicin and Telecare* , pp. 401-403.

[59] WANsim. (2009). *WAN simulators and emulators*. Retrieved from http://www.wan-sim.net/

[60] Yu, R., Wong, V., Song, J.-H., Leung, V., & Chan, H. (2010, 1 8). Next Generation Moibility Management:An Introduction.

[61] Zhang, C., & Tsaoussidis, V. (2001). TCP Real: Improving Real-time Capabilities of TCP over Heterogeneous Networks. *Proceedings of the 11th international workshop on Network and operating systems support for digital audio and video* (pp. 189 - 198). New York: Port Jefferson.

[62] Zhang, Q., & Zhang, Y.-Q. (2008, 1). Cross-Layer Design for QoS Support in Multihop Wireless Networks. *Proceedings of The IEEE* , pp. 64-76.

[63] Zhang, Y., Zheng, J., & Hu, H. (2008). Security in Wireless Mesh Networks.

[64] Zhao, R., Walke, B., & Hiertz, G. (2006, 11). An Efficient IEEE 802.11 ESS Mesh Network Supporting Quality-of-Service. *IEEE Journal on Selected Areas in Communications* , pp. 2005-2017.

[65] Zhu, C., Yang, O. W., Aweya, J., Oullette, M., & Montuno, D. (2002). A

Comparison of Active Queue Management Algorithms Using the OPNET Modeler. *IEEE Communications Magazine* , 158-167.