Performance Analysis of Differentiated *QoS* *MAC* in Wireless Local Area Networks (WLANs)

by

Kenan Xu

A thesis submitted to the

Department of Electrical and Computer Engineering

in conformity with the requirement for

the degree of Master of Science

Queen's University

Kingston, Ontario Canada

September 2003

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisisitons et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-86202-X
Our file Notre référence
ISBN: 0-612-86202-X

Canadä

Abstract

With wide acceptance and initial successful deployment, Wireless *LANs* (*WLANs*) have drawn great interest from both industry and academia. Quality of Service (*QoS)* is one of the most active research topics, since it plays a key role in enabling the integration of various multimedia services into a single *WLAN* domain. The current *IEEE* 802.11 distributed *MAC* component known as the Distributed Coordination Function (*DCF*), coordinates the multiple access procedure among active wireless nodes. *DCF*, however, does not have *QoS* provisions. Therefore, in recent years, several mechanisms have been proposed to enhance *DCF* with *QoS* capability. Among others, Enhanced *DCF* (*EDCF*), an extension to *DCF* defined in *IEEE* 802.11e, has attracted growing interest. *EDCF* enhances *DCF* in terms of providing differentiated *QoS* with the introduction of prioritized Virtual Stations (*VSs*) while maintaining the characteristics of simplicity and robustness to failure of the legacy *DCF*. Moreover, *EDCF* is fully compliant with *DCF* functions, so that nodes running the legacy *DCF* and nodes running the advanced *EDCF* can work concurrently in the same domain.

This thesis introduces a multi-dimensional Markov model to analyze the performance of the *IEEE* 802.11e *EDCF MAC* protocol. Based on this model, we present extensive performance evaluation in terms of throughout, throughput ratio between flows, and access delay of flows of distinct priorities under the *RTS/CTS* mode. We also provide quantitative analysis of the impact of prioritized parameters, i.e., Arbitration InterFrame Space (*AIFS*) and Contention Window (*CW*) on *QoS* differentiation. The accuracy of the proposed model is

verified by means of comparing the numerical results obtained from both the analytical model and simulation. Our research can be used as a guideline for the prediction of how flows belonging to a certain Traffic Category (*TC*) perform with their *TC*-specific parameters, as well as for the design of *QoS*-support *WLANs* with *EDCF* and the tuning of the parameters to achieve the desirable differentiated *QoS* objectives.

Acknowledgments

I would like to begin by thanking Professor Hossam Hassanein, my thesis supervisor and mentor for the past 2 years. Hossam has been a great advisor and guider, providing me with support, encouragement, and an endless source of good ideas. He is always broad-minded to novel opinions and provides students with a free and open academic environment with a lot of fun. I appreciate him for giving me such a wonderful opportunity to struggle for a higher level of life.

Also, I would like to express my appreciation to Professor Glen Takahara of the department of Mathematics and Statistics, Queen's University. It was through Professor Takahara that I have learned much of the fundamental mathematical tools and advanced derivation skills, which are crucial to my research work.

I feel grateful to Quanhong Wang, my colleague in the Telecommunication Research Lab (*TRL*). We had great collaboration on a lot of research work. Besides, I had the privilege of interacting with other wonderful, bright, and talented people, including but not limited to Abdelhamid Taha, Ahmed Safwat, Alex Oliver, Hongyan Du, Mohammed Al-Riyami, Nidal Nasser, Tiantong You and Yinghong Fan. They have taught me much, and their advice, feedback and friendship have made my Master's experience both more educational and more fun. It has been a great pleasure to work with all of you.

The financial support provided by Communications and Information Technology Ontario (CITO) and Queen's University is greatly appreciated.

Finally, I would like to thank my family, especially my sister Hong Xu. She has given me a lot of support and care since the very early time. She has always been there for myself and the family so that I can concentrate on my research. I am indebted to my parents, for everything they have given to me. They taught me the value of knowledge, the joy of love, the appreciation of life and the basic merits of a person. I am sorry that my father, Ling Xu, cannot witness my development today, which largely stems from his early education and care. My mother, Xingkai Chen, has sacrificed a lot to support me, while providing constant encouragement and love. My family has been an inspiration to me in all that they do.

Contents

List of Tables

List of Figures

List of Acronyms

AIFS	Arbitration IFS
AP	Access Point
ATM	Asynchronous Transfer Mode
BB	Black Burst
CDMA	Code Division Multiple Access
CFP	Contention Free Period
CHOICE	A network project executed by UCSD
CP	Contention Period
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
CW	Contention Window
DBTMA	Dual Busy Tone Multiple Access
DCF	Distributed Coordination Function
DIFS	DCF IFS
DR-TDMA	Dynamic Reservation TDMA
DTMP	Disposable Token MAC Protocol
EDCF	Enhanced DCF
EY-NPMA	Elimination Yield-Non-Preemptive Priority Multiple Access
FDMA	Frequency Division Multiple Access
HCF	Hybrid Coordination Function
HIPERLAN	High Performance LAN
HOL	Head Of Line
HP	High Priority
IEEE	Institute of Electrical and Electronics Engineers
IFS	InterFrame Space
ISM	Industrial, Scientific and Medical
IT	Idle Time
LANs	Local Area Networks
LP	Low Priority
MAC	Medium Access Control
MAT	Matlab

MIMO	Multiple Input Multiple Output
NAV	Network Allocation Vector
OFDM	Orthogonal Frequency Domain Multiplexing
PC	PCF Coordinator
PCF	Point Coordination Function
PEN	Prototype Embedded Network
PHY	Physical
PIFS	PCF IFS
PL	PayLoad
QoS	Quality of Service
RTS	Request To Send
SIFS	Short IFS
SIM	Simulation
SISO	Single Input Single Output
STT	SuccessTransTime
TC	Traffic Category
TDMA	Time Division Multiple Access
VSs	Virtual Stations
WEP	Wired Equivalent Privacy
Wi-Fi	Wireless Fidelity
WLANs	Wireless LANs
WSN	Wireless Sensor Network

Chapter 1

Introduction

Wireless communication networks using *ISM* free-of-charge radio frequencies, widely known as wireless Local Area Networks (*WLANs*) or Wireless Fidelity (Wi-Fi) networks, have experienced explosive growth in interest from both industry and academia in recent years.

1.1 Characteristics of *WLANs*

A number of common characteristics distinguish *WLANs* from other wireless communication technologies.

First of all, *WLANs* use the Industrial, Scientific and Medical (*ISM*) frequency band, which is free of charge. *WLANs* hence provide a significant advantage over 3G cellular networks that require frequency licensing.

Secondly, *WLANs* typically adopt distributed contention-based *MAC* protocols, such as the Distributed Coordination Function (*DCF*) in the *IEEE* 802.11 standard, and Elimination Yield-Non-Preemptive Priority Multiple Access (*EY-NPMA*) in the *HIPERLAN2* standard, to coordinate multiple access among nodes. Such contention-based wireless *MAC* provides simplicity, robustness and good performance under light to medium traffic load [1] as opposed to more sophisticated wireless multiple access paradigms, such as *FDMA*, *TDMA*, *CDMA*, polling, dynamic reservation, etc. Moreover, distributed *MAC* can work efficiently when a centralized controller/scheduler is not available. Our focus in this thesis is on an enhanced version of the *DCF* function, known as Enhanced Distributed Coordination Function (*EDCF*), which is a key component of the *IEEE* 802e standard.

The transmission radius of *ISM* radio is relatively small compared to cellular networks; hence the use of *ISM* is usually limited to Local Area Networks (*LANs*). Note that Wireless *LANs* (*WLANs*) are not limited to being the last hop of the end-to-end communication. Within an Ad Hoc network architecture, which has no centralized coordination, it is feasible to implement multi-hop peer-to-peer wireless communications with wireless ad hoc routing support [2,3].

High bandwidth is another important aspect of *WLANs*. A widely commercialized standard, *IEEE* 802.11b operating in the 2.4 GHz *ISM* band, achieves data rates of up to 11Mb/s. Newer standards, such as *IEEE* 802.11a operating at *5GHz*, *IEEE* 802.11g operating at 2.*4GHz*, and European *HIPERLAN*/2 operating at the 5 *GHz* band, will support multiple transmission modes with the Orthogonal Frequency Domain Multiplexing (*OFDM*) and provide raw data rates of up to 54 Mb/s.

Furthermore, wireless communication networks using *ISM* can be potentially deployed with either of two types of network architecture, namely, infrastructure-Based and Ad Hoc, as illustrated in Figure 1-1. An Infrastructure-Based network refers to a network that is connected by wireless link with a centralized coordinator, i.e., Access Point (*AP*), coordinating the operation of the network and ensuring connectivity, authentication, security and control. The *AP* can use the mixed mechanisms of *CSMA/CA*-based contention and polling to coordinate multiple access of the network channel. The *AP* is a close counterpart of the base station in cellular networks. By contrast, a wireless Ad Hoc network has no centralized coordinator. It is a collection of two or more devices equipped with wireless communications and networking capability. Nodes can communicate directly with other nodes that are within their radio range, or intermediate nodes are used to relay or forward the packets from the source to the destination if they are out of the transmission range of each other. Typically, all nodes equipped with routing intelligence form the peer-to-peer communication environment. Multi-hop wireless networks extend the coverage of infrastructure-based networks and compensate for the short range of the radio.

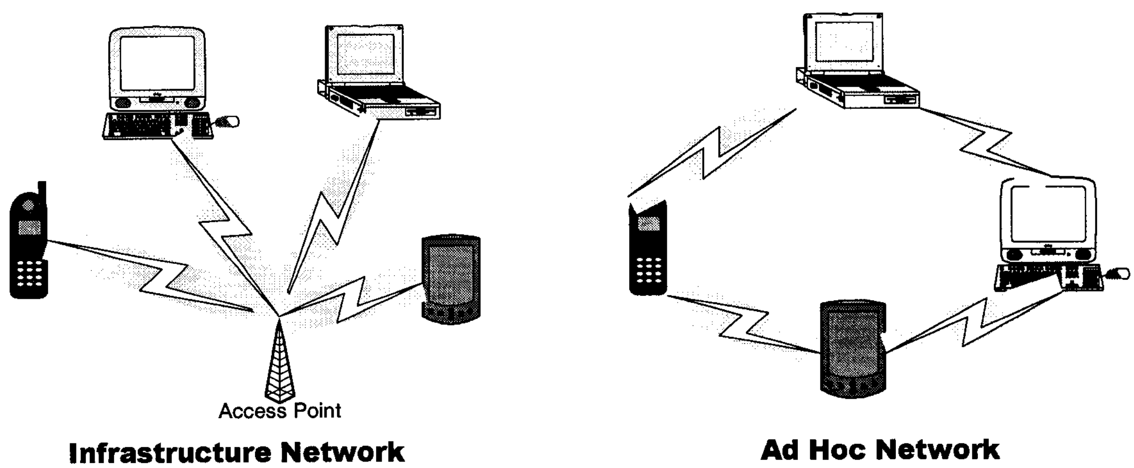


Figure 1-1: Wireless *LAN* architectures.

1.2 Applications

WLANs and their variants provide a wide spectrum of applications. Some of these are summarized below.

1.2.1 Wireless Internet Access

WLANs can be used as a platform for access to the public Internet. "Hot Spots", which are simply wireless LAN access points, have been installed in Malls, airports, coffee stores, campuses, hotels, trade shows, etc. Also, some offices and conference rooms are using *WLAN* as an alternative for the traditional Ethernet. Hot spots are challenged with billing, security and power efficiency issues.

1.2.2 Wireless Multimedia Service

With the proliferation of multimedia service in the Internet, users will not be satisfied with the plain best-effort data service. Wireless transport of multimedia data is a very challenging task due to the acute contradiction between the real-time, high-bandwidth requirement of the applications and scarce wireless transmission resources. Among others, advanced coding and modulation methods, multimedia techniques such as compression, *QoS* mechanisms, including *QoS* routing, *QoS MAC*, admission control and *QoS* adaptation are under intensive investigation. Wireless multimedia applications also intensify the demand for power-aware *WLAN* architectures.

1.2.3 Home Automation and Device Communications

WLANs and their variants are also ideal candidates for home automation and device communications. Short-range lower-power wireless nodes enable the connection of traditional appliances. Such applications are under investigation in several projects, such as *PEN*, *CHOICE* etc. [4-7].

1.2.4 Temporary Communication Platform for Military and Rescue Missions

This is another unprecedented application of *ISM* networks. Typically, under such a situation, infrastructure-based communication facilities are not available or are destroyed. An Ad Hoc Network needs little effort to deploy and mobile nodes can form a peer-to-peer network very quickly. [8]

1.2.5 Wireless Sensor Networks

A Wireless Sensor Network (*WSN*) is considered as a variant of *WLANs*. In short, a *WSN* is a collection of sensors equipped by low-power wireless transceivers. *WSN* nodes normally execute three functions, namely, sensing the environment, processing signals and data computing, and wireless communications. Compared to traditional sensor devices, *WSNs* have many advantages. For example, a *WSN* extends the range of sensing. Moreover, a *WSN* enables distributed control, thus the system becomes more robust. An overlay of sensing range among nodes provides fault-tolerance and therefore improved accuracy. The greatest concern of *WSN* is power consumption and latency [9-11].

1.3 Quality of Service (*QoS*) in *WLANs*

Although applications of and research on *WLANs* have achieved certain accomplishments, such as large-scale field deployment and a series of *IEEE* 802.11 standards and other proposals on the way, they are still drawing increasing attention and interest on some hot research topics , including energy efficiency, security, and Quality of Service (*QoS*).

The power consumption of wireless communication nodes determines the lifetime of connectivity and suitability for certain applications. Typically, mobile devices are powered by battery and so have limited power supplies. Among others, the network interface is a major source of power usage. For instance, a typical example from a Toshiba 410 *CDT* mobile computer demonstrates that 18% of the battery power consumption is due to the wireless interface [11]. Research indicates that a significant amount of energy can be saved by incorporating low-power strategies into the design of network architectures and protocols at all network layers [11-23].

Network Security is one of the major concerns in the commercialization of the *IEEE* 802.11 standard. The current *IEEE* 802.11 series of standards, for instance 802.11b, define Wired Equivalent Privacy (*WEP*) for three basic security objectives, i.e. Authentication, Confidentiality and Integrity [24-26]. However, recent research has identified some vulnerabilities with respect to *IEEE* 802.11 security, such as the static shared *WEP* keys, short cryptographic key, poor packet integrity and so on [27-29]. The *IEEE* 802.11 task group has suggested that users should not depend on the link-layer security mechanism. Instead, users should implement their own security policy at the network layer and above. However, they have promised security improvements in a new standard supplement, *IEEE* 802.11i. Securing Ad Hoc networks is a more challenging task. The absence of a central controller negatively affects key management, intrusion detection and other security operations. For a study of ad hoc network security, please refer to reference [30].

Quality of Service (*QoS*) is the major interest of this thesis. The main motivation of *QoS* is the integration of data and multimedia communications over the wireless channel. High-bandwidth and well-designed resource management mechanisms and protocols are two key issues to transmit multimedia data over *ISM* wireless networks.

On the Physical (*PHY*) layer, Orthogonal Frequency Domain Multiplexing (*OFDM*) enabling eight different PHY rates, ranging from 6 Mbps to 54 Mbps, has been adopted in the *IEEE* 802.11a standard (see Table 1-1). *OFDM* is a multi-carrier modulation scheme, where the input stream is split into several sub-streams. Each sub-stream is carried by one orthogonal frequency carrier and is then combined with other sub-streams at the receiver. Besides the obvious improvement of transmission rates over *IEEE* 802.11b, *OFDM* also enables *PHY* rate selection depending on the wireless channel condition between the transmitter and receiver. When coupled with dynamic packet fragmentation, this leads to link adaptation [31]. With link adaptation, the system can achieve optimized bandwidth utilization under all circumstances.

Table 1-1: Eight transmission rates and parameters of IEEE802.11a *OFDM.*

Data rate (Mbits/s)	Modulation	Coding rate	Coded bits per subcarrier (NBPSC)	Coded bits per OFDM symbol (NCBPS)	Data bits per OFDM symbol (NDBPS)
6	BPSK	1/2	1	48	24
9	BPSK	3/4	1	48	36
12	QPSK	1/2	2	96	48
18	QPSK	3/4	2	96	72
24	16-QAM	1/2	4	192	96
36	16-QAM	3/4	4	192	144
48	64-QAM	2/3	6	288	192
54	64-QAM	3/4	6	288	216

Another significant advancement in the *PHY* layer is the development of Multiple-Input-Multiple-Output (*MIMO*). As depicted in Figure 1-2, *MIMO* denotes a wireless link that is composed of M antennas on the transmitter end and N antennas on the receiver end, where M and N are any finite integer numbers. The essential benefit of *MIMO* systems is the ability to turn multi-path propagation, usually a pitfall of Single-Input-Single-Output (*SISO*) wireless transmission, into an advantage for increasing the user's data rate.

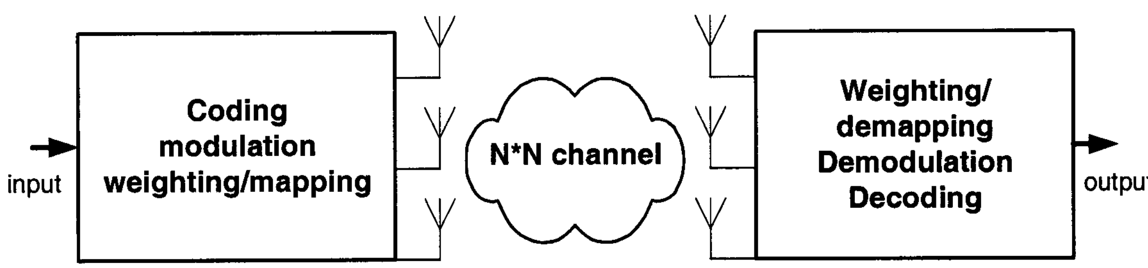


Figure 1-2: *MIMO* wireless transmission system. The transmitter and receiver are equipped with multiple antennas.

Other techniques, such as smart-antenna and multi-user diversity have positive impact on data transmission efficiency. For detailed information, please refer to [32-34].

At the *MAC* layer, the medium access control protocol has to eliminate collisions as much as possible in order to achieve high system throughput [1,14]. However, the overall throughput may still not be good enough in the context of *QoS*. To meet the service requirements of the mixed traffic of data, voice and video, the *MAC* has to differentiate service quality with enhanced strategies. The Enhanced Distributed Coordination Function (*EDCF*) of *IEEE* 802.11e is a cornerstone of those efforts. According to the *IEEE* 802.11e standard, each *EDCF* node can have up to eight network interface queues. Each queue, also known as a Virtual Station (*VS*), operates independently in conformity with *CSMA/CA* mechanisms. The *QoS* differentiation objective is achieved by differentiated run-time parameters of different *VSs*. In addition, a certain mapping policy has to be set up between the application and the virtual station. The performance of the *EDCF* function is the focus of this research.

Packet scheduling is another approach to *QoS* deployment, especially in Ad Hoc networks. Packet scheduling was originally studied in the context of wired networks, under several variations of fair queuing. Such research aims to design and analyze mechanisms to implement *QoS*-based weighted fair queuing paradigms for instantaneous fairness and bounded delays [35]. In the mid 90's, researchers shifted their focus to scheduling issues in wireless cellular network [36, 37]. The wireless environment imposes some unique problems on wireless fair queuing algorithms. Such problems include dynamically varying wireless channel capacity, location-dependent and bursty errors, channel contention, shortage of global channel state and limited processing power and battery power.

Packet scheduling in Ad Hoc networks is a very challenging task. In addition to the problems commonly inherent in all wireless environments, as outlined above, the lack of a centralized scheduling entity implies that packet scheduling has to be implemented in a distributed manner [38-40].

QoS routing is the major approach to *QoS* at the networking layer. Connectivity and shortest hop are usually the concerns of traditional ad hoc routing protocols. *QoS* routing takes other metrics, e.g., bandwidth and delay, into account, in order to measure the quality of a path. A routing path of non-*QoS* routing paradigms may not meet the quality requirement in the context of *QoS* routing. Detailed information can be found in [41].

1.4 Contributions of this Thesis

We study the *QoS* performance of the *IEEE* 802.11e *EDCF* in single-hop *WLANs*. The *IEEE* 802.11e *EDCF* originally appeared in 2001 as a draft standard [42], which defines two multiple access control mechanisms, *EDCF* and *HCF* (Hybrid Coordination Function). *EDCF* is a distributed multiple access function using the *CSMA/CA* contention resolution method. *HCF* is built over *EDCF*, using both random access mechanisms provided by *EDCF* and polling mechanisms.

Our research sets up a Markov model to analyze the transmission performance of individual prioritized virtual stations, which operate using the *EDCF* mechanism, where priority is achieved via a set of differentiated parameters for individual virtual stations. Numerical results are obtained to show how well the *QoS* differentiation can be achieved with differentiated parameters. Based on these results, we obtain guidelines for the prediction of how the prioritized *VSs* perform with specific parameters, as well as how to tune the parameters to achieve the desired *QoS* objectives. We argue that *EDCF* is an effective *QoS*

MAC mechanism to provide *QoS* differentiation. A simulation model also validates our analytical model.

Our research is an initial effort on mathematical analysis of *EDCF* performance. To the best of our knowledge, no research taking the mathematical analysis approach to study *QoS* support in *EDCF* has been reported in the literature.

1.5 Thesis Organization

In the next chapter, we will first discuss and analyze all kinds of challenges in wireless communications. This will be followed by a taxonomy of wireless *LAN MAC* protocols and *MAC* protocol performance metrics. We continue with a detailed discussion of the *IEEE* 802.11 *DCF* and *EDCF*. Related work to enhance *DCF* for *QoS* deployment is also summarized. Chapter 3 introduces the detailed development of the mathematical model of *EDCF*. Chapter 4 describes the simulation model, for the purpose of mathematical model validation. Then, we show and discuss numerical results based on the analytical model. The thesis concludes with a summary and a discussion of future research directions in Chapter 5.

Chapter 2

Medium Access Control (*MAC*) and *QoS*

In *WLANs,* Medium Access Control (*MAC*) mechanisms are responsible for coordinating access to a multiple access channel in either a static or dynamic manner. Wireless *MAC* algorithms were originally proposed in 1970 in the well known *ALOHA* protocol [43]. Afterwards, many algorithms have been designed for wireless networks.

Wireless communications use radio as the transmission carrier. The propagation of electromagnetic waves is severely influenced by the surrounding environment and is limited by three major factors, namely, reflection, diffraction, and scattering. The variation of the environment due to nodal mobility and changes of surroundings, combined with sources of interference, leads to a complex and time-varying communication environment. To cope with these challenging radio propagation characteristics, many wireless *MAC* protocols have been proposed and implemented in all kinds of wireless network topologies and environments. In this chapter, we will briefly review the challenges facing wireless communication and wireless *MAC* protocols.

2.1 Challenges in Wireless Communication

Compared with *MAC* mechanisms in wired networks, designing wireless *MAC* has to take some unique characteristics, existing in the context of wireless communications, into consideration, such as node mobility, large-scale and small-scale path loss, location-dependent carrier sensing, burst error channel and so on. Those characteristics, in turn, give birth to some notorious problems.

2.1.1 Node Mobility

In most cases, mobile nodes may move around independently and unpredictably during the operation of the system. This random movement triggers important events in the communications system, including a handoff operation from one cell to another, link breakage and routing failure, and time varying signal quality. Handoff occurs when a node moves to another cell while it is involved in an ongoing session in the original cell. If the distance between two parties of an ongoing session becomes larger than the radius of the transmission due to the movement of either node, the receiving node would fail to receive the signal, which we refer to as link breakage. In multi-hop wireless networks, link breakage can further cause routing failure [44, 45] if the victim link is along the working routing path. Node movement also causes variation of signal strength due to the effect of large-scale path loss and small-scale fading effects.

2.1.2 Large-Scale Path Loss

Radio signals propagation is regulated by three factors: reflection, diffraction and scattering. In general, the strength of the signal arrived through the multi-path fading channel decreases as the distance between the transmitter and receiver increases. The large-scale propagation model is used to predict the mean signal strength for an arbitrary transmitter-receiver separation distance [46, chapter 3]. However, it is critical to realize that the practical measurements of signal strength over a large number of measurement locations which have the same transmitter-receiver separation, are different due to different levels of clutter on the propagation path. This effect is known as shadowing.

2.1.3 Small-Scale Fading and Time Varying Radio Link

Small-scale fading is used to describe the rapid fluctuation of the amplitude of a radio signal over a short period of time or travel distance due to small-scale multipath propagation. In a nutshell, fading is caused by interference between two or more copies of the transmitted signal, which arrive at the receiver at slightly different times. These waves, called multipath waves, combine at the receiver antenna to give a resultant signal, which can vary widely in amplitude and phase, depending on the distribution of the intensity of disturbance and relative length of the propagation path. In addition to small-scale fading, some other factors, including node mobility (as above) and variation of environment, also contribute to the time-varying properties of a radio link.

2.1.4 Burst Error Channel

Errors can be attributed to several factors, such as node mobility, time varying radio link, etc. Co-channel interference, adjacent channel interference and background noise also contribute to high bit error rates in wireless networks. Co-channel interference refers to the radio cross talk caused by other cells using the same frequency. Adjacent channel interference normally refers to the interference caused by energy leakage from an on-going transmission using an adjunct frequency in the same geographic area. Typically, wireless channels may have higher bit error rates than wired channel by a factor of 3 orders of magnitude.

2.1.5 Network Topology

As mentioned earlier, *WLANs* typically form one of two types of topologies, infrastructure-based or frastructure-less, i.e. Ad Hoc networks. Here, infrastructure refers to the centralized control entity, viz. the access point. The centralized controller plays a vital role in the network, providing admission control, packet scheduling, synchronization, power control, security, billing and some other functions. In terms of MAC mechanisms, a central controller may work as the coordinator and arbitrator of media access opportunities. In contrast, Ad Hoc networks are formed by a collection of wireless nodes without the presence of the centralized controller. The lack of the centralized control poses a challenge for deploying those functions that are provided by the central controller in infrastructure-based networks.

2.1.6 Location-Dependent Carrier Sensing

In practice, regulated by three mechanisms, the signal strength received by the receiver can be modeled by a large-scale path loss model and a small-scale fading and multi-path model. Therefore, it is critical to note the fact that the wireless transmission must be detected at the receiver (not the sender) and so the detection is location dependent. In turn, the location-dependent carrier sensing causes three problems:

- **Hidden Terminal problem.** A hidden terminal refers to a node that is within the range of the intended destination but out of range of the sender, so that the hidden terminal is not aware of the on-going transmission. As illustrated in Figure 2-1, Node B can communicate with A and C and A and C cannot hear each other. When A transmits to B, C cannot detect the transmission using the carrier sense mechanism. If C transmits to D, a collision will occur at B. In such a case, C is a hidden terminal to A.
- **Exposed Terminal problem.** Exposed nodes are complementary to hidden terminals. An exposed terminal is a node that is within the range of the sender but out of range

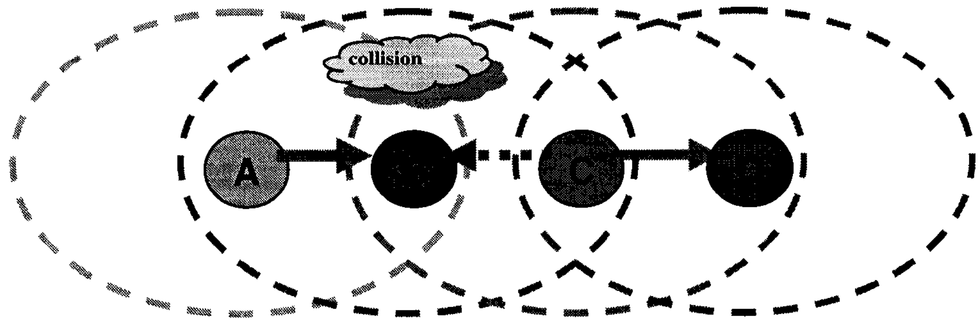


Figure 2-1: Hidden terminal problem.

of the destination, so that the exposed terminal will be improperly precluded from sending in order to avoid collision. As illustrated in Figure 2-2, Node C can communicate with B and D and B and A can hear each other. When C transmits to D, B detects the transmission using the carrier sense mechanism and postpones transmission to A, even though such transmission will not cause a collision.

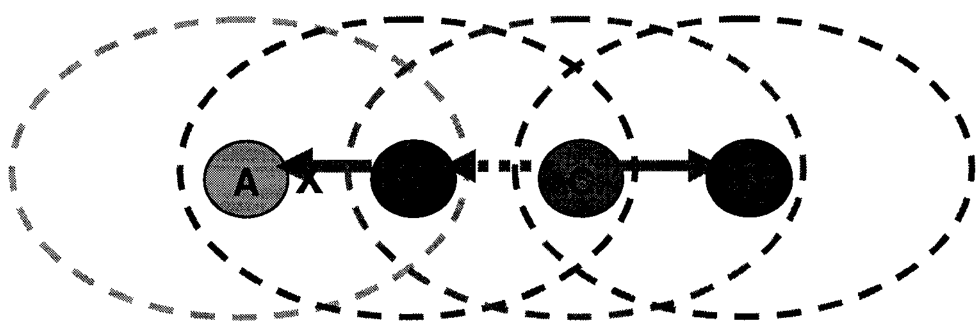


Figure 2-2: Exposed terminal problem

- **Far-near problem.** The far-near problem, also known as the capture effect, occurs when a receiver can clearly receive a transmission from one of two simultaneous transmissions, both within its range. In Figure 2-3, nodes A and C transmit simultaneously to B, the signal strength received from C is much higher than that from A, and C's transmission can be decoded without errors in the presence of transmission from A.

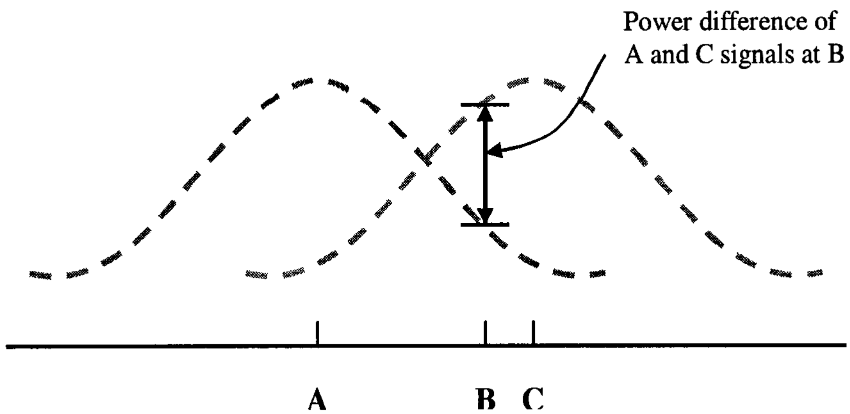


Figure 2-3: Far-near problem.

2.1.7 Half-duplex Operation

When a wireless node is transmitting data, a large fraction of the signal energy leaks into the receive path. The phenomenon is referred to as self-interference. The leakage signal typically has much higher power than the received signal, therefore it is difficult to have the node send and transmit data at the same time unless a perfect filter-circuit is available.

2.2 Wireless *MAC* Classification

Figure 2-4 shows a classification of wireless *MAC* protocols. As mentioned above, wireless networks can be divided into two types in terms of the network topology, namely, infrastructure-based and infrastructure-less, or Ad Hoc network. Accordingly, two types of network *MAC* approach, i.e., centralized *MAC* protocols and distributed *MAC* protocols, are designed for each of the network topologies. In addition, hybrid *MAC* protocols, which combine the merits of both distributed and centralized mechanisms, are the third type of *MAC* approach.

2.2.1 Centralized *MAC* Protocols

Centralized *MAC* protocols include controlled multiplexing algorithms, such as Code Division Multiple Access (*CDMA*), Time Division Multiple Access (*TDMA*) and Frequency

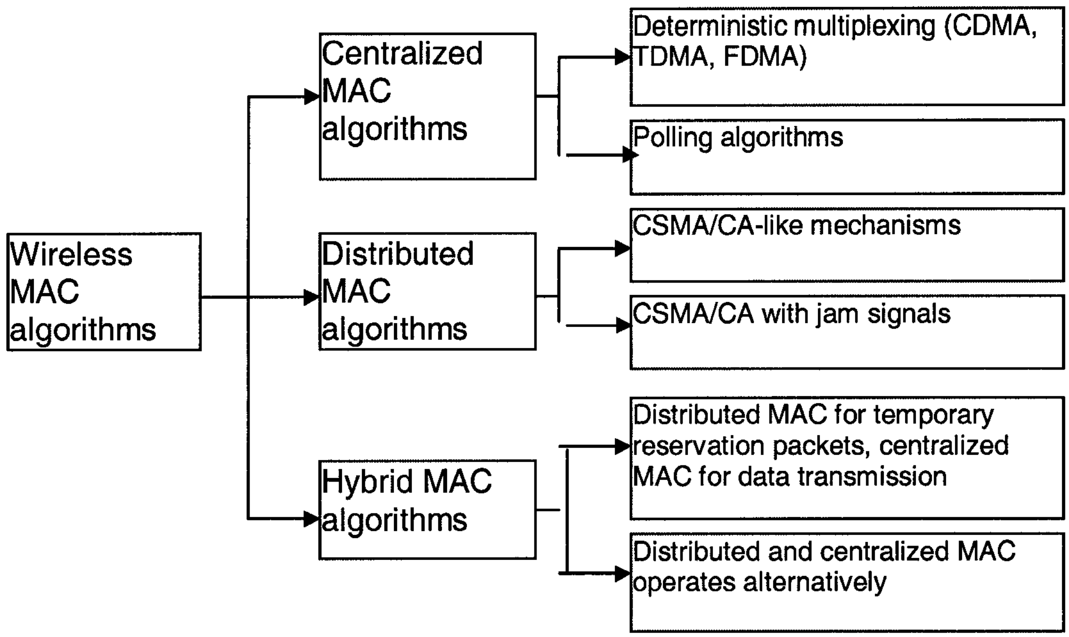


Figure 2-4: Classification of wireless *MAC* protocols.

Division Multiple Access (*FDMA*) and polling algorithms, such as Disposable Token MAC Protocol (*DTMP*) [47]. Deterministic multiplexing algorithms may adopt some dynamic reservation mechanisms to achieve flexibility and to be adaptive to the variation of traffic load.

2.2.2 Distributed *MAC* Protocols

Distributed *MAC* protocols are the major interest of this thesis. They usually provide random multiplexing access to operating nodes. Most protocols in this category adopt the principles of carrier sensing and collision avoidance.

Carrier sensing eliminates most transmission collisions by deferring the transmission when the channel is sensed busy. However, in many cases, including multi-hop wireless networks, location dependence of carrier sensing results in hidden and exposed terminal problems, which have a great effect on algorithm efficiency. Two types of *MAC*-layer mechanism are proposed to relieve such problems. In the Distributed Coordination Function

(*DCF*), the exchange of *RTS-CTS* control packets claims the transmission space for subsequent data exchange, thereby eliminating the hidden terminal transmission when data are in the air. A detailed discussion of *DCF* will follow later in Section 2.4. Another proposal, using separate control and data channels, known as Dual Busy Tone Multiple Access (*DBTMA*) access protocol, solve both the hidden terminal and exposed terminal problems by indicating the transmission or receiving status explicitly [48, 49]. As shown in Figure 2-5, when node A is transmitting data to node B using the data channel, it sends a transmitting Busy Tone (*BT*) on the control channel. Node E in our example, which senses the transmitting *BT* can initiate a data exchange without disturbing the on-going data transmission because it is outside the receiving BT area. Thus the exposed terminal problem is solved. At the same time, node B sends a receiving *BT* on the control channel while receiving the data. Node C, which senses the receiving *BT*, can not initiate a data exchange but may terminate a data transmission to avoid hidden terminal problem. Node G, which senses both the receiving *BT* and transmitting BT, will not engage in any data exchange. More details are available in [48,49].

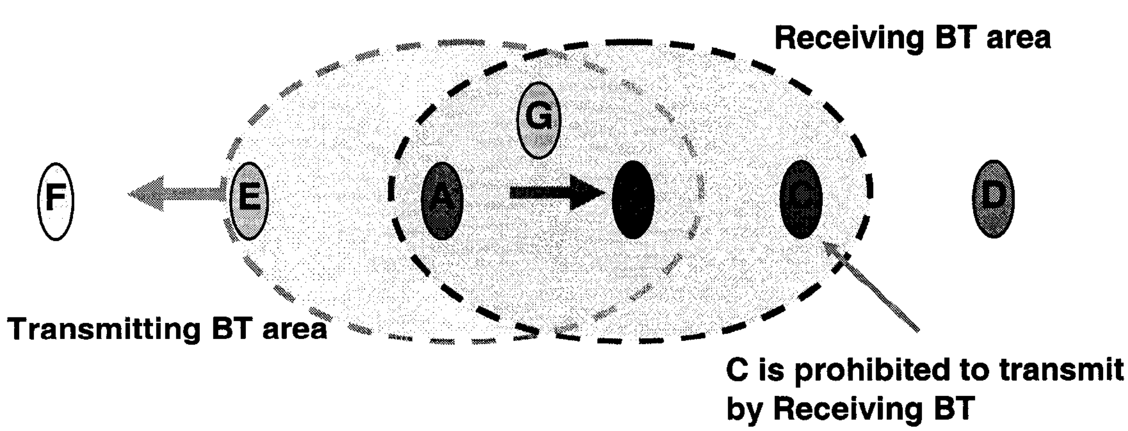


Figure 2-5: *DBTMA* solution to location-dependent carrier sensing problems

Other distributed *MAC* protocols, including *EY-NPMA* [50,51] and *BB* tone [52,53], use jamming signals. *EY-NPMA* stands for Elimination Yield–Non-preemptive Priority Multiple Access. It is the channel access protocol used in the *HIPERLAN* system being developed in Europe. A node will trigger a three-phase competition mechanism when the channel is sensed busy during the initial period corresponding to the time it takes to transmit 1700bits, otherwise the node will transmit immediately. The three phases are the prioritization phase, contention phase (including elimination sub-phase and yield sub-phase), and transmission phase, as illustrated in Figure 2-6. In the priority phase, the node is sensing the medium and will quit the competition if the medium becomes busy. The length of the priority phase reflects the priority of a packet, i.e., the higher priority one packet has, the shorter the priority phase. After sensing the medium to be free for the priority phase, the node enters into the contention phase. In the elimination sub-phase, each node transmits for a random number of slots. The node switches into yield sub-phase at the end of a random period if the node senses the channel is idle, otherwise it quits its transmission competition. During the yield sub-phase, the node keeps listening to the channel for a random number of slots. If no transmission is detected during this time, the node starts and completes its data transmission.

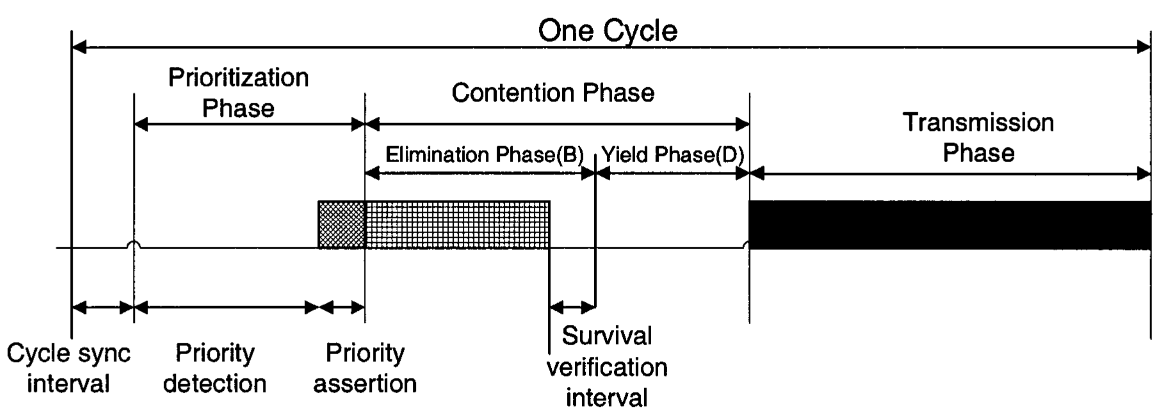


Figure 2-6: Access phase in EY-NPMA

BB is short for "Black Burst", which was proposed to support prioritized data transmission in ad hoc networks. According to [52,53], after sensing the channel idle for a preset fixed period of time, the operating nodes start transmitting jamming signals on the channel for a time period proportional to the packet priority. At the end of transmission of the jamming signal, the node turns over and starts sensing the channel. If the channel is clear for the sensing period of time, it will send its real-time packet; otherwise it will quit the competition.

2.2.3 Hybrid *MAC* Protocols

Combining the controllability of centralized protocols and flexibility of distributed protocols, hybrid protocols provide better performance at the cost of complexity. There are many different ways of combination. Two approaches are discussed below.

Dynamic Reservation *TDMA* (*DR-TDMA*) [54] is a *MAC* protocol for wireless *ATM* (Asynchronous Transfer Mode) networks. Figure 2-7 illustrates the frame structure of the *DR-TDMA* protocol. In this protocol, a distributed *MAC* mechanism regulates the transmission of temporary control messages for reservation purposes and centralized *MAC* mechanisms are in charge of data transmission. Specifically, *DR-TDMA* uses a framed pseudo-Bayesian priority Aloha protocol, which is a distributed mechanism, to regulate the access of contention slots for reservation requests. On the other hand, the base station allocates data transmission resource in *TDMA* mode corresponding to reservation requests.

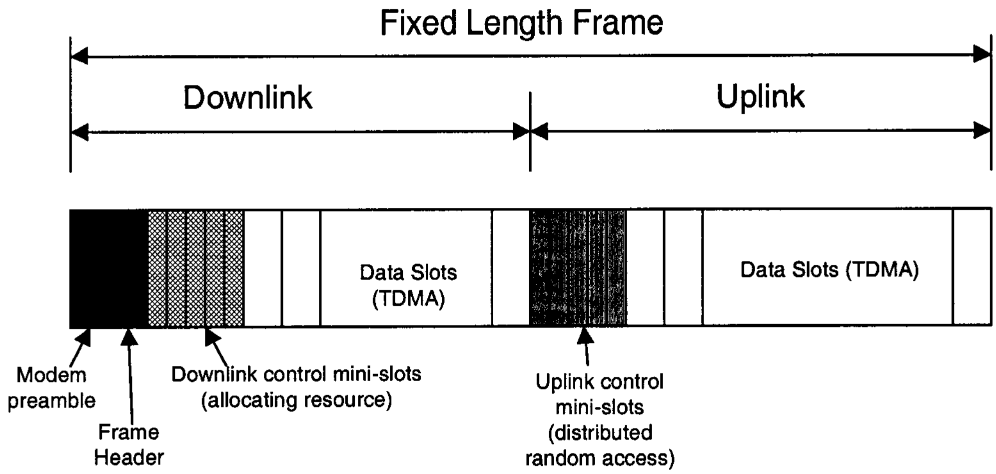


Figure 2-7: DR-TDMA frame structure.

Another example is the Point Coordination Function (*PCF*) in *IEEE* 802.11. *PCF* use a centralized polling mechanism and the *DCF* mechanism alternately in the different stages of operation. As depicted in Figure 2-8, the *PCF* access procedure consists of two periods, the Contention Free Period (*CFP*) and the Contention Period (*CP*). During the *CFP*, the PCF Coordinator (*PC*) schedules the transmission of packets using a polling mechanism, while other stations keep silent by setting their Network Allocation Vector (*NAV*) to be the length of the *CFP*. Thus most contention is avoided completely in a single-hop network. During the *CP*, the channel access obeys the rules of the *CSMA/CA*-based *DCF* function.

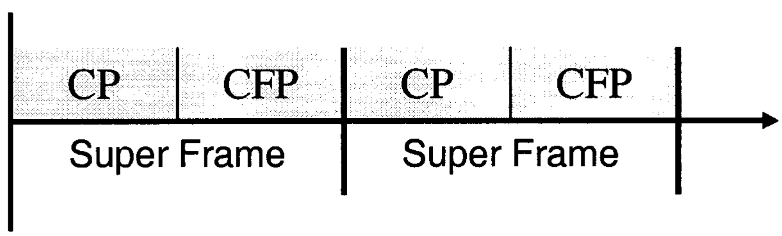


Figure 2-8: *PCF* superframe structure: controlled by distributed and centralized *MAC* protocols alternately.

2.3 Wireless MAC Performance Metrics

It is important to understand the metrics used to evaluate the *MAC* protocols. Following is a brief discussion of these metrics.

- **Throughput.** Throughput is the fraction of the channel capacity used for successful data transmission. A desired *MAC* protocol should provide high throughput to maximize the resource utilization.
- **Delay.** Delay is defined as the average time spent by a packet in the *MAC* queue, i.e., from the instant it is queued till its transmission is completed. Besides channel capacity and *MAC* contention, traffic load is the most important external factor affecting the delay.
- **Fairness.** A *MAC* protocol is fair if it does not exhibit preference to any single node or flow, referred to respectively as node-based fairness and flow-based fairness, when multiple nodes or flows are trying to access the channel. When multimedia traffic is considered, fairness is defined as being able to distribute bandwidth (delivery with delay constraint) in proportion to the requirements of the flows.
- ***QoS* Support.** *QoS* support is critical to deploy multimedia applications over wireless networks. *QoS* is achieved using prioritized access and scheduling. Protocols, such as Enhanced *DCF* (*EDCF*) in 802.11e, Black Burst (*BB*) and *EY-NPMA*, enable prioritized *MAC* access. Prioritized *MAC* access is of great interest in this thesis.
- **Stability.** Stability refers to the ability to maintain high efficiency when the offered load is close to or greater than the maximum capacity of the system. Such ability is critical to a successful *MAC* protocol in the real world.

- **Power Efficiency.** As mentioned earlier, power efficiency is a critical property of network protocols for wireless networks since the network nodes are often battery powered.
- ***TCP* Efficiency.** A *TCP* session generates two traffic flows each session, one for data packets and one for *ACK* packets. Some random access *MAC* protocols, such as *DCF*, have degraded performance as the number of sessions in the system increase [55]. Therefore, it is desirable to piggyback *TCP ACK* packets on *MAC* control messages, so that the overall performance is improved since one traffic flow is eliminated.

2.4 *IEEE* 802.11 Distributed Coordination Function (*DCF*)

Distributed *MAC* mechanisms have experienced a great proliferation since *ALOHA* was proposed in early 70's. Among others, *DCF* is a *CSMA/CA*-like protocol and has gained increased attention due to the growing popularity of Wireless *LANs*. *DCF* was first standardized by the *IEEE* 802.11 working group in 1997 [56].

In *DCF* there are three InterFrame Space (*IFS*) intervals specified in the *IEEE* 802.11 standard: short *IFS* (*SIFS*), Point Coordination Function *IFS* (*PIFS*), and *DCF-IFS* (*DIFS*). Figure 2-9 illustrates the relationship between them.

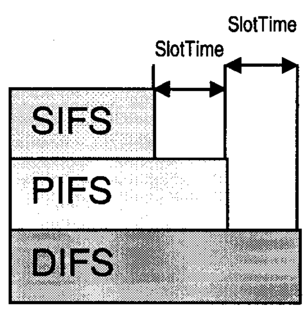


Figure 2-9: InterFrame Space (*IFS*) specified by the *IEEE802*.11 standard.

The *SIFS* is the shortest InterFrame spacing and is used for transmission of an *ACK* frame or a *CTS* frame in *DCF*. The *PIFS* is used by a node operating as an access coordinator in *PCF* mode to allocate priority access to the medium. It is one slot time longer than the *SIFS*. Nodes operating in *DCF* mode transmit data packets using the *DIFS*. It defines the fixed length of time that nodes have to wait before a random countdown procedure in order to access the channel eventually.

The *IEEE* 802.11 *DCF* works as a listen-before-talk scheme, and operates in either basic access mode or *RTS-CTS* mode.

A. Basic Access Mode

As depicted in Figure 2-10, the basic access mode is a two-way handshaking mechanism between the transmitter and the receiver of data transmission using data packets and *ACK* packets.

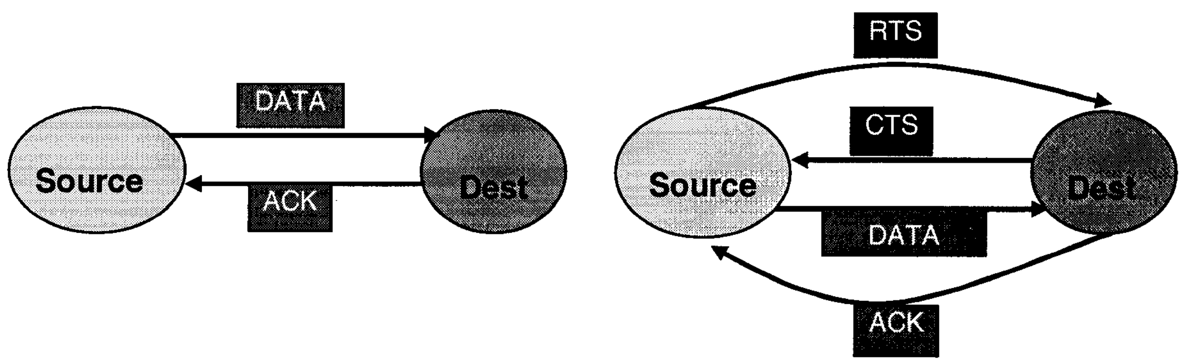


Figure 2-10: Two-way handshaking and Four-way handshaking.

In basic access mode, the station starts its transmission immediately if the medium is sensed to be idle for an interval of length *DIFS*. If the medium becomes busy during the sensing period, the station defers until an idle period of *DIFS* length is detected and then further counts down a random back-off period before transmission. The back-off period, which is an integral number of slot times, referred to as the backoff value, is uniformly chosen in the range (0, *CW*-1), where *CW* stands for a runtime parameter called the Contention Window. The backoff value decrements every slot time when the channel is sensed idle until it reaches zero. The node can start to transmit the queued packet when the backoff value counts down to zero. Upon the proper reception of one data packet, the receiver should respond with an *ACK* message after a *SIFS* time period. The *DATA-ACK* exchange is referred to as two-way handshaking. If the channel becomes busy before the backoff reaches zero, for example, due to the transmission of other nodes, the backoff procedure stops and the backoff value is frozen. The node will wait for another idle period of *DIFS* length, and resume the backoff procedure with the remaining backoff value.

If two or more nodes transmit at the same time, the transmissions are destined to collide with each other. When collisions take place, the nodes involved will timeout waiting for *ACK* packets and execute the recovery procedure.

The recovery procedure first invokes a Binary Exponential Backoff (BEB) algorithm to compute a new *CW* value. According to the BEB algorithm, *CW* is initially set to *CWmin*, a parameter that may be agreed upon among all operating nodes. *CW* is doubled every time a collision happens (implied by *ACK* timeout in *IEEE* 802.11) until it reaches the parameter *CWmax*. After *CW* reaches *CWmax*, *CW* will remain unchanged for further retransmissions. *CW* is reset to *CWmin* when finally the retransmission succeeds. Figure 2-11 illustrates the BEB algorithm.

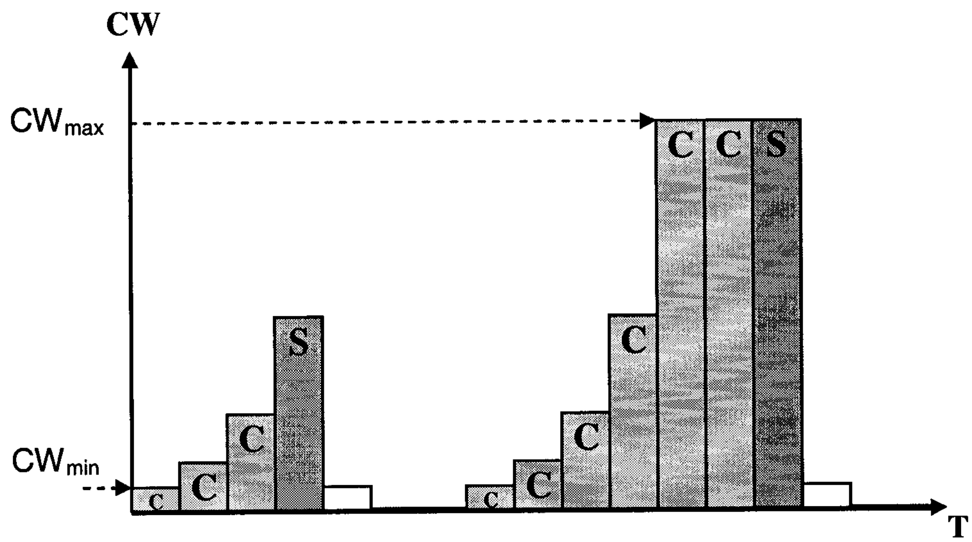


Figure 2-11: *BEB* algorithms: first packet is transmitted successfully after retrying three times; the second packet is transmitted successfully after retrying six times

After generating the new *CW* using the *BEB* algorithm, the new backoff value is drawn from (0, *CW*-1) using the same algorithm. The same backoff procedure applies afterwards. The *IEEE* 802.11 standard also introduces two other parameters, dot11ShortRetryLimit and dot11LongRetryLimit, to respectively regulate the time of retransmission efforts in order to avoid unnecessary retries and delays. Please refer to page 78 of [24] for details.

The *DCF* basic access mechanism exchanges data using the two-way handshaking; hence the data is transmitted right after the node finishes the backoff procedure. However, since two or more nodes may happen to finish the random backoff procedure at the same time, a collision may occur and it would last for more than the length of the whole packet (packet length plus timeout) due to the lack of detection ability in wireless nodes. So a collision will cause a large amount of resource wastage. Moreover, the basic access mechanism makes no effort to deal with the "hidden terminal" problem, which results in a large fraction of collision cases. The *RTS-CTS* mechanism addressed both of the two challenges by "acquiring the space" before transmitting.

B. RTS-CTS Access Mode

The *RTS-CTS* access mode executes the same carrier sensing and random backoff procedure as in the basic mode. It also uses the same *BEB* algorithm to regulate the adjustment of the contention window. It enhances the basic access mode in two respects.

Firstly, the *RTS-CTS* access mode enhances the basic mode two-way handshaking with *RTS-CTS-DATA-ACK* four-way handshaking (Figure 2-10). When the node finishes the pre-sending carrier sensing and random countdown, instead of transmitting the data packet immediately, the node sends out a Request-To-Send (*RTS*) control packet to the intended receiver. Upon the correct reception of the *RTS* message in an idle state, the receiver responds to the sender with a Clear-To-Send (*CTS*) control packet. After receiving the *CTS* message, the sender then starts transmitting the real data packet, and then expects the *ACK* packet as the confirmation of a successful data exchange.

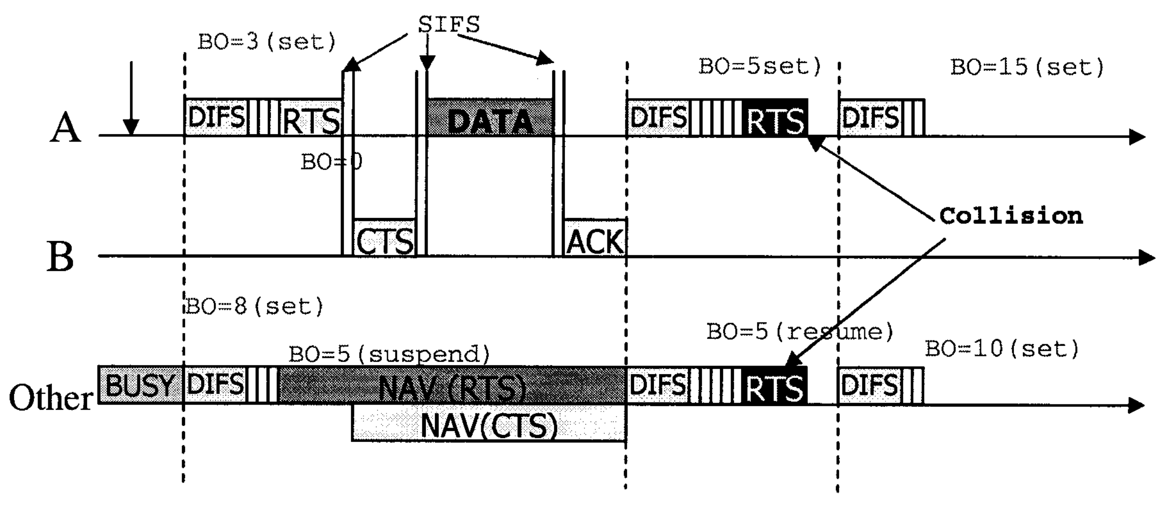


Figure 2-12: An example of virtual carrier sense and operation of *DCF* in *RTS/CTS* mode. The first transmission is successful from node A to B. Other nodes are prohibited to transmit by virtual sense.

Four-way handshaking decreases the chance of collision significantly since the *RTS* packet has much smaller length than the data packet. Even though collisions may sometimes occur, the resource waste is minimized since collisions last only for the length of the *RTS* packet. Moreover, the *RTS-CTS* exchange can solve the hidden terminal problem. The nodes which receive the *RTS* or *CTS*, but are not the intended receiver, are prevented from transmitting until the end of the current transmission. So the nodes within the radio range of the receiver would not start transmissions that would cause collision at the receiver of the first transmission. But the *RTS* message intensifies the exposed terminal problem.

Another important feature introduced by *RTS-CTS* is the virtual carrier-sense mechanism in addition to the physical carrier sense mechanism. They are both used to determine the state of the medium. When either function indicates a busy medium, the medium is considered busy. The virtual sense mechanism can save much energy in the following manner. The node employs a variable, known as the Network Allocation Vector (*NAV*), to maintain a prediction of future traffic on the medium based on duration information announced in *RTS/CTS* frames prior to the actual exchange of data. When a node other than the intended receiver receives an *RTS* or *CTS* packet, the node will set its *NAV* equal to the value in the duration field of the received frames if the latter is larger than the node's current *NAV* value (Figure 2-12). The *NAV* value counts down at a uniform rate regardless of the network status. When *NAV* reaches zero, the virtual carrier-sense indication is that the medium is idle; when nonzero, the indication is busy.

DCF provides an efficient random access mechanism for wireless *LANs*. However, *QoS* is not considered into the original design. All packets, of any user or any application are treated in exactly the same manner, following the same transmission procedure and using the same control parameters. So *DCF* does not support the integration of multimedia flows. Some researches [57-73] have been executed to enhance the *DCF* function with *QoS* ability.

Among others, *EDCF* is a novel *QoS*-based *MAC* mechanism in the process of standardization [42].

2.5 *IEEE* 802.11 Enhanced Coordination Function (*EDCF*)

EDCF is a *QoS*-enabled distributed *MAC* mechanism derived from the *DCF* function. Under *EDCF*, each node operates up to eight prioritized queues in order to provide differentiated *QoS* to the different flows of traffic. These independent queues, as illustrated in Figure 2-13, also called virtual stations in the *IEEE* 802.11E standard, execute the same random access procedure as a *DCF* node, but with differentiated, Traffic Category (*TC*) specific parameters. The *TC*-specific parameters will lead to the differentiated *QoS* of different data flows.

In *EDCF*, there are four *TC*-specific parameters. The first is the Arbitration Interframe Space (*AIFS*). *AIFS* is the counterpart of the *DIFS* in *DCF*, but the specific value is a function of the traffic category. *EDCF* defines the value of the *AIFS* as *AIFS*=*DIFS*+k*slot_time, where $k \in [0,8]$. Also, *EDCF* defines *CWmin*, *CWmax* and Persistent Factor (*PF*) to regulate the backoff and retransmission procedure. The contention window is originally set to *CWmin*. Upon the occurrence of packet collision, *CW* is bounded as (in practice, we use the low bound throughput our experiments)

$$CW[TC] \geq ((CW[TC]+1)*PF[TC]-1). \qquad (2.1)$$

Enlargement of the contention window is limited by *CWmax*. So if the new *CW* calculated using equation (2.1) is greater than *CWmax*, the new *CW* would be *CWmax*. When *PF* equals 1, *CW* remains constant, and when it is 2, the backoff procedure is almost the same as the *BEB* algorithm in *DCF*.

In *EDCF*, packets are delivered through multiple virtual station instances within one node, and each virtual station instance is parameterized with *TC*-specific parameters. Each virtual station within the node contends for a transmission opportunity (*TXOP*) and independently starts a backoff after detecting the channel being idle for an Arbitration Interframe Space (*AIFS*). After waiting for the *AIFS*, each virtual station sets a backoff counter to a random number drawn from the interval [1, *CW*+1]. If the channel is determined busy before the counter reaches zero, the backoff has to wait for the channel being idle for an *AIFS* again, before continuing to count down the backoff value. A big difference from the legacy *DCF* is that when the channel is determined idle for the period of an *AIFS*, the backoff counter is reduced by one beginning from the last slot interval of the *AIFS* period. Note that with the legacy *DCF*, the backoff counter decrements beginning from the first time slot interval after the *DIFS* period. Upon a transmission collision, *CW* is enlarged as stated above and a new backoff value is generated randomly from [1, *CW*+1] again. If the counters of two or more parallel virtual stations in a single station reach zero at the same time, a scheduler inside the station avoids the virtual collision. The scheduler grants the *TXOP* to the virtual station with highest priority, out of the virtual stations that virtually collided within the station, as illustrated in Figure 2-13. Collisions may occur between transmissions by different physical nodes.

EDCF provides prioritized channel access by running the separate virtual stations in one physical station. Each virtual station follows the same random access procedure but with differentiated parameters. With well-designed mapping mechanisms from application data flows to *MAC* layer traffic categories, integration of data traffic of different *QoS* characteristics is feasible.

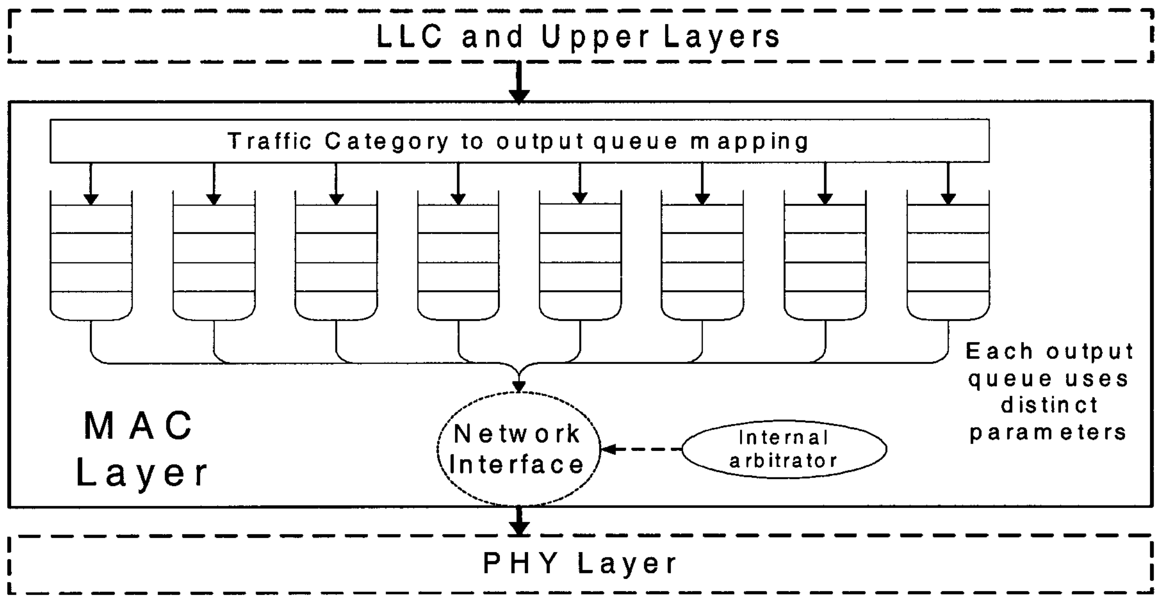


Figure 2-13: The virtual structure of *EDCF* nodes.

2.6 Related Work on *QoS MAC*

In addition to *EDCF* specified by the *IEEE* 802.11 Task Group E, there are other efforts made to enhance *CSMA/CA*-like access schemes to provide *QoS*. Some researches [52, 53,57-71] have revealed that, with a *CSMA/CA*-like access scheme, packet-level *QoS* can be achieved by optimizing the nodal parameters and improving contention window control algorithms. These proposals can be categorized based on different aspects, including network topology, *QoS* objectives, *QoS* metrics, and *QoS* leverages. A summary of existing *QoS* MAC protocols is in Table 2-1.

Table 2-1: Distributed QoS MAC proposals

Ref.	Methods	*QoS* leverages				Network Status Awareness	*QoS* Objective	Metrics	Other Factors
		Persistence	Back Off	Inter Frame Space	Packet Length				
57	Simulation		Yes			Monitoring	Fairness	Bandwidth	Per flow *QoS*
58	Math Model and Simulation		Yes			Monitoring	Weighted Fairness	Bandwidth	Per flow *QoS*; Channel utilization is considered as well
59, 60, 61	Simulation	Yes	Yes			Information Exchange	Weighted fairness	Bandwidth	Per flow *QoS*; Two methods of computing persistence factor
62	Simulation		Yes			Monitoring	Weighted Fairness	Bandwidth	Per flow *QoS*;
63	Math Model and Simulation		Yes			Information Exchange	Weighted Fairness	Bandwidth	Per flow or per node *QoS*; Assume no mobility
64, 65	Math Model and Simulation		Yes			Monitoring	Service Differentiation	Bandwidth Delay	Integration of multimedia traffic (TCP, voice, CBR); admission control is used
66	Simulation and Math Model		Yes	Yes	Yes		Service Differentiation	Bandwidth Delay Drop rate	Throughput, delay drop rate (TCP, UDP)
67	Simulation		Yes				Service Differentiation	Bandwidth	UDP(CBR) traffic
68	Simulation		Yes				Service Differentiation	Delay Throughput	Integration of multimedia traffic (voice, data); *QoS* routing; reservation
52, 53	Simulation						Service Differentiation	Delay Throughput	Black Burst (BB) gives priority to real-time traffic by blocking data traffic using
69	Simulation and Math Model						Service Differentiation	Delay, Bandwidth	Dynamic reservation; similar to D-TDMA
70	Simulation		Yes			Information Exchange	Weighted Fairness	Bandwidth	SCFQ Scheduling in distributed way; exchange transmission tag information, per-flow fairness
71	Simulation		Yes			Information Exchange	Weighted Fairness	Bandwidth	Scheduling in a distributed way; exchange fairness label; per-flow fairness

MACAW [57] is a *QoS*-based MAC protocol providing per-flow fairness. The Multiplicative Increase/Linear Decrease (*MILD*) contention window algorithm, "backoff copy", and per-flow contention window paradigm were proposed to provide better fair access to packet streams across multiple cells. As the name indicates, the *MILD* algorithm increases the contention window by a multiplicative factor upon a collision, and decreases the contention window by one upon a successful transmission. In order to implement "backoff copy", every packet contains the current value of the contention window. When a station hears (monitors) a packet, it copies that value as its own contention window. The "backoff copy" mechanism enables information sharing about the network status in terms of the contention condition. In addition, the station maintains one contention window instance for each flow, which is declared critical to realize the per-flow fairness objective.

In [58], efficient channel utilization and weighted fairness were achieved by dynamically selecting the proper contention window to reflect the relative weights among data traffic flows and the number of stations contending for the wireless medium. Therefore, the window size has to be calculated based on the network traffic condition, which is monitored and tracked down along with the operation.

In [59-61], Balanced Media Access Methods (*BMAM*) was proposed to provide per-flow fair transmission. *BMAM* is similar to the *DCF* function in that they both use listen-before-talk schemes and both execute random backoff algorithms. However, *BMAM* introduced the persistence factor, p, as the leverage of fairness, which is similar to p-persistent *CSMA*. In *BMAM*, stations send packets with probability p, after the backoff period, or back off again with probability 1-p using the same contention window value, where p can be calculated dynamically using either a contention-based or time based media access method, in order to achieve the fairness objective.

Reference [62] tried to achieve per-flow fair medium access by replacing the *BEB* backoff algorithm in *DCF*. In the proposed scheme, each station will estimate its share and other stations' share of the channel and then adjust the contention window accordingly. The other stations' share of the channel is calculated based on the constant monitoring on the network traffic. This work proposed the "fairness index" to measure the fairness of flows.

In [63], a general analytical framework was introduced to model system wide per-flow fairness via the specification of per-flow utility functions. Then the fairness model could be translated into a corresponding contention resolution algorithm. Using this translation, the backoff algorithm for achieving proportional fairness was derived.

The work in [64] and [65] proposed a set of mechanisms that form a fully distributed wireless differentiated services network. The mechanisms include a distributed radio resource monitoring mechanism, service quality estimation mechanisms, distributed traffic and admission control mechanisms, and of interest to us, a distributed, differentiated services-capable *MAC*. This *MAC* is a modification of *DCF* in *IEEE* 802.11 and the data transmission follows the same listen-before-talk scheme. However, best effort and premium traffic streams use different contention window bounds so that premium traffic tends to wait less than the best effort traffic before sending. The work in [68] takes a similar approach to enabling *QoS* at the *MAC* layer.

In reference [66], three *DCF* factors, namely, contention window algorithms, interframe spaces and packet lengths were studied as the leverage of service differentiation. Delay and bandwidth of prioritized flows were simulated and compared when using one of three differentiation mechanisms. In the first case each priority level has a different backoff increment function. In the second each priority level is assigned a different DIFS length. The last case is when each priority level has a different maximum frame length, beyond which the frame has to be segmented and transmitted multiple times. Similarly, reference [67] proposed and evaluated different contention window algorithms for the prioritized flows

As we have mentioned in Section 2.2, the authors of [52] and [53] proposed Black Burst (*BB*) mechanism to provide service differentiation between two traffic classes. The real-time traffic senses idle in the shorter fixed length period *PIFS*, before sending a pulse of energy, or black burst signals, so that the best effort data traffic is blocked whenever there is real-time traffic. Thus, the real-time traffic always has priority over the best-effort traffic.

The protocol in [69] was named Group Allocation Multiple Access (*GAMA*). It is similar to Dynamic Reservation *TDMA* (*DR-TDMA*) [54] for wireless *ATM* networks. *GAMA* divides time into dynamically sized "cycles". Each cycle consists of a contention period of up to a maximum duration and a group transmission period during which one or more stations transmit data packets without collisions. *GAMA* places a limit on the length of the group transmission period, which allows *GAMA* to bound the interval between occurrences of a transmission period. A station with data to send competes for the right to be added to the transmission group by successfully completing an *RTS/CTS* message exchange during the contention period. Once a station is a member of the transmission group, it is able to transmit a collision-free data packet during each cycle.

The work in [70] and [71] intended to achieve weighted fairness of flows with distributed control facilitated by state information exchange. The scheme introduced in [70] tried to mimic the centralized Self-Clocked Fair Queueing algorithm (*SCFQ*) using the piggy-backed timer information and choosing backoff intervals proportional to normalized packet sizes by flow weights, i.e., packet size/weight. Paper [71] chooses a *CW* size proportional to the normalized rate of each flow.

In terms of the performance of QoS support under *EDCF*, we are in the very early stages of research. Indeed, very few research efforts have been reported in the literature [72-75]. The work in [72] provides a brief illustration of the differentiated *QoS* effect of the *EDCF* with simulation results. Reference [73] presented an evaluation of IEEE802.11e in more realistic scenarios. Both the Enhanced Distributed Coordination Function (*EDCF*) and Hybrid Coordination Function (*HCF*) are simulated and compared under multimedia traffic loads. References [74, 75] presented an adaptive service differentiation scheme called Adaptive *EDCF*. This scheme is derived from *EDCF*, by incorporating a dynamic contention window algorithm, which takes into account both application requirements and network conditions. The performances of the *EDCF* and *AEDCF* algorithms were compared with prioritized traffic using simulation. However, and to the best of our knowledge, no research taking the mathematical analysis approach to study *QoS* support in *EDCF* has been reported in the literature.

EDCF operation is a complex and highly-dynamic process and the *QoS* performance is affected by a number of factors. Intuitively, the smaller the *AIFS* is, the shorter the waiting time exists before the virtual station accesses to the wireless medium. Thus the virtual station obtains higher priority when contending for the wireless media. And this is the same as what the smaller contention window will do. But the degree of differentiation depends not only on the length of the *AIFSs*, but also relies on the combination of the *AIFS* length and contention window size. In this thesis, we consider the combined impact of *AIFSs* and contention window related parameters and provide quantitative analysis of *QoS* supported by *EDCF* in terms of saturation throughput (*ST*)[1] and access delay by means of a multi-dimensional Markov model. We derive formulas to compute the system *ST, ST* ratios and access delay of different flow*s* in terms of *AIFS* and contention window parameters. We validate our theoretical analysis by comparing with simulation results. These lead to the accurate evaluation of differentiation *QoS* provided by the *IEEE* 802.11e *EDCF* as well as the analysis of the impacts of the prioritized parameters on *QoS* differentiation. Based on these results, we obtain guidelines for the prediction of how the prioritized *VSs* perform with specific parameters, as well as how to tune the parameters to achieve the desired *QoS* objectives.

[1] Saturation throughput is defined as the limit reached by the system throughput as the offered load increases till the maximum load that the system can carry in stable conditions. This concept is also as interpreted in [76].

Chapter 3

Mathematical Modeling of *EDCF*

This chapter introduces a mathematical model to study the performance of the *EDCF* function in terms of service differentiation in a single-hop wireless *LAN*. Based on the mathematical model, formulas for the throughput and access delay of virtual stations are derived under traffic saturation conditions. The model validation and numerical results analysis will follow in the next chapter.

A discrete time Markov chain is used to model the *EDCF* channel access mechanism. Before introducing the model, it is necessary to overview some of the concepts related to Markov chains.

3.1 Markov Chain, Transient States and Time to Absorption

We denote the individual Markov chain state with lowercase letter i, j, k, l, etc, and the Markov chain state space as S. For any state i, Let f_i denote the probability that, starting in state i, the process will ever reenter state i. State i is said to be recurrent if $f_i = 1$ and transient if $f_i < 1$. Moreover, for any recurrent state j, if the one-step transition probability P_{jl}=0 for all $l \neq j, l \in S$, j is said to be an absorbing state.

Given a Markov chain, let G denote the set of transient states, and G^C denote the set of absorbing states (after mapping each recurrent class into one absorbing state); hence the state space $S = G \cup G^C$. Also, suppose there are p absorbing states and q transient states, so the size of state space, $|S|$, equals p+q. If we arrange all absorbing states to be located before the transient states when composing the transition matrix, we can write the transition matrix as $P = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$, where I is a p*p identity matrix indicating the transition probabilities from absorbing states to absorbing states; R is a q*p matrix indicating the transition probabilities from transient states to absorbing states; Q is a q*q matrix indicating the transition probabilities from transient states to transient states; 0 is a p*q null matrix indicating the transition probabilities from absorbing states to transient states, which are 0 according to the definition of transient state and absorbing state [77, 78].

We have $P^n = \begin{bmatrix} I & 0 \\ R_n & Q^n \end{bmatrix}$, where $R_n = (I + Q + Q^2 + ... + Q^{n-1})R$. P^n denotes the n-step transition matrix of the Markov chain and R_n denotes the n-step transition probabilities from transient states to absorbing states.

Let $W = I + Q + Q^2 + ...$, i.e. $W = (I - Q)^{-1}$, if $|W| < \infty$. If $i, j \in G$, $(Q^n)_{ij} = P(X_n = j \mid X_0 = i)$.

where X_0 denote the initial state of the Markov chain and X_n is the state at time n (or after n transitions).

Therefore,

$$W_{ij} = \sum_{n=0}^{\infty} (Q^n)_{ij} = \sum_{n=0}^{\infty} P(X_n = j \mid X_0 = i) = \sum_{n=0}^{\infty} E[I\{X_n = j \mid X_0 = i\}] = E[\sum_{n=0}^{\infty} I\{X_n = j \mid X_0 = i\}]$$

So,

$W_{ij} = E[number\ of\ times\ j\ is\ visited | X_0 = i]$

Summing over j,

$$\sum_{j \in G} W_{ij} = Expected\ value\ of\ total\ time\ spent\ in\ G\ starting\ in\ state\ i.$$

In other words,

$$\sum_{j \in G} W_{ij} = Expected\ time\ to\ absorption\ given\ X_0 = i.$$

Let $i \in G, l \in G^C$, $and f_{il}(n) = P(1st\ time\ to\ l\ is\ time\ n \mid X_0 = i)$, Then

$$f_{il}(n) = P(1st\ time\ to\ l\ is\ at\ time\ n \mid X_0 = i)$$

$$= \sum_{j \in G} P(1st\ time\ to\ l\ is\ at\ time\ n \mid X_{n-1} = j) \times P(X_{n-1} = j \mid X_0 = i)$$

$$= \sum_{j \in G} R_{jl} \cdot (Q^{n-1})_{ij} = (Q^{n-1} \cdot R)_{il}$$

Let $f_{il} = P(ever\, visit\, state\, l \mid X_0 = i)$. Then

$$f_{il} = \sum_{n=1}^{\infty} f_{il}(n) = \sum_{n=1}^{\infty} (Q^{n-1} \cdot R)_{il} .$$

Let F be the matrix $((f_{il}))_{i \in G, l \in G^C}$, so

$$F = \sum_{n=1}^{\infty} Q^{n-1} R = (I + Q + Q^2 + \cdots)R = WR = (I - Q)^{-1} \cdot R,$$

where f_{il} is called the reaching probability and means the probability that the Markov chain is absorbed into state l starting from transient state i.

3.2 Model Assumptions

A wireless *LAN* executing the *EDCF* function is a complex system. The performance is influenced by many factors, as stated in Chapter 2. It is very difficult to take every factor into account at the same time. We make a few necessary assumptions in order to simplify the modeling of the system without losing the major operational characteristics of the system.

Firstly, in this thesis, we study the *EDCF* performance in a single hop wireless *LAN*. We assume the network is fully connected, i.e., every node in the system can hear other nodes directly. In this case, hidden and exposed terminals won't arise.

Secondly, we assume ideal channel conditions. We ignore packet loss due to interference of all kinds and also the capture effect. Thus only packet loss due to collisions is considered. Packet collision takes place when more than one node transmits at the same time.

Thirdly, we analyze *EDCF* performance when the system operates under saturation conditions, i.e., each virtual stations always has a packet available for transmission and the transmission queue of each station is always nonempty. While this is not always the case in practice, we should note that the "Saturation Throughput" is a fundamental performance metric defined as the limit reached by the system throughput as the offered load increases, and represents the maximum load that the system can carry in stable conditions. Similar to other random access schemes, such as Carrier Sense Multiple Access (*CSMA*), the *DCF* function exhibits a non-monotonicity in term of throughput change when the traffic load increase to a certain degree. Reference [76] has illustrated and discussed this characteristic of *DCF*, (see Figure 3-1). As the offered load increases, the throughput grows up to a maximum value, referred to as the maximum throughput. However, further increase of the offered load leads to a decrease in the system throughput. This results in the practical impossibility to operate *DCF* function at its maximum throughput for a "long" period of time. Saturation throughput indicates the throughput lower bound when the network is running under high load. As well, the access delay derived under the saturation condition gives an upper bound for average packet service time. The access delay refers to the time interval between a packet becoming Head of Line (*HOL*) and its being transmitted successfully, less the time used for the successful data exchange procedure. The bound can be used to predict the performance of a WLAN system.

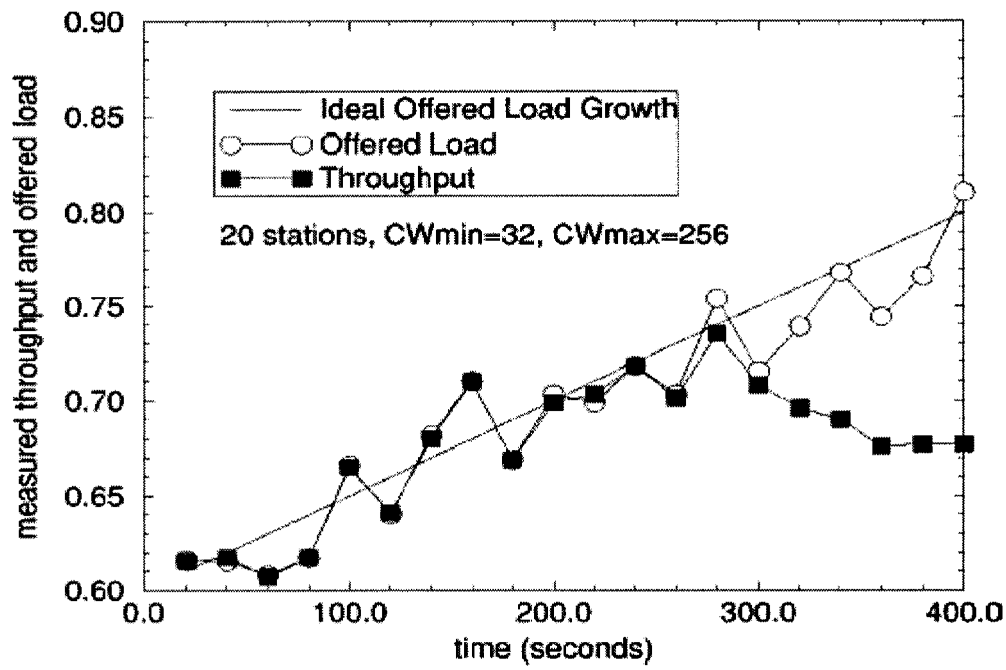


Figure 3-1: Instability of the *DCF* function.

The next assumption is regarding the setting of parameter values. To evaluate the effect of differentiated *AIFS* lengths, we set the persistence factor to one for all virtual stations, which means the contention window value remains constant when collision happens. We justify this assumption by two results; (a) one is a valid value for the persistence factor according to *IEEE* 802.11 standards [24]; and (b) the persistence factor takes effect only when collision occurs, and the contention window is reset to *CWmin* when one transmission succeed. We argue collision is a relatively rare occurrence during the operation of the system when the nodes use the *EDCF* function, which will be illustrated when we discuss the numerical results. We will discuss further how this assumption affects the results in Chapter 4.

Our last assumption is that the network consists of a finite number of nodes and each node operates only one virtual station of a certain priority. In other words, we ignore the effect of virtual arbitration on who is going to send when two virtual stations finish countdowns at the same time.

3.3 Mathematical Modeling

Now, we are ready to model *EDCF* in a single-hop *WLAN*. We focus on the differentiated *QoS* provided by the *IEEE* 802.11e *EDCF* under a traffic overload condition. Indeed, the major contribution of this thesis is the analytical evaluation of system saturation throughput, overall throughput and access delay of individual flows, and the accurate calculation of throughput ratio among them. The analytical results describe how prioritized flows share the system resources. Moreover, throughput ratios can help us to quantitatively understand to what degree High-Priority (*HP*) flows can have advantage over Low-Priority (*LP*) ones when contending for the wireless medium. Therefore, we gain insight into how well 802.11e *EDCF* can actually support differentiated *QoS* using virtual stations with different *AIFS* lengths and *CW* related parameters.

Our analysis consists of two parts. First, we will present a m-dimensional Markov chain model, defining the state space as well as the transition matrix. Then, according to the theorem of transient states and time to absorption [77, 78], we derive expressions for throughput of individual flows, throughput ratios among different flows, and access delay as functions of *AIFS* lengths and the *CW* size.

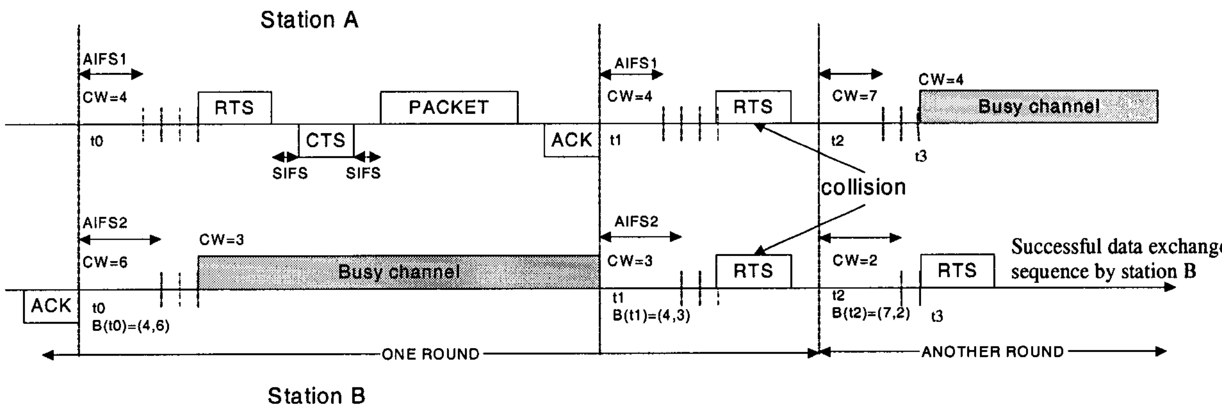


Figure 3-2: Example of the operation of the *EDCF* function

3.3.1 Multi-Dimensional Markov Chain Model

As shown in Figure 3-2, in *EDCF*, the channel would be in one of three states: idle, successful transmission, and collision (depicted in Figure 3-3). Here, the system is in the idle state when all nodes either have no packet to transmit or are executing the sensing and backoff procedure before the data transmission. The system is in the successful transmission state when one and only one node is transmitting a packet without collision. Specifically, its duration denotes the time interval from sending the *RTS* frame to receiving the *ACK* frame in *RTS/CTS* mode or the time interval from sending the *DATA* frame to receiving the *ACK* frame in basic transmission mode. The system is in the collision state if more than one node tries to exchange packets at the same time. In particular, its duration denotes the time interval from sending the *RTS* frame to the *CTS* timeout at the sender in *RTS/CTS* mode or the time from sending the *DATA* frame to the *ACK* timeout in basic transmission mode. These 3 states interleave each other on the time axis. The system throughput is the percentage of successful transmitted payload length (normalized by the transmission rate) over the total time.

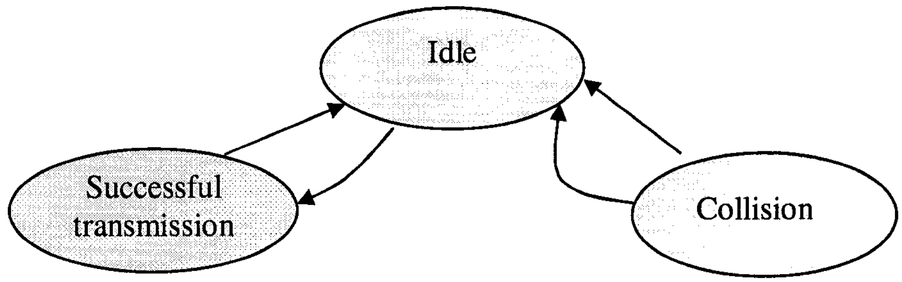


Figure 3-3: *EDCF* states and transitions.

Consider a fixed number of m flows ($f_1, f_2, \ldots, f_m$), each of which corresponds to one virtual station equipped with a set of TC-specific parameters, i.e., $AIFS_i$, $CW_{\min_i}$, $CW_{\max_i}$, and PF_i for $i = 1,2,\ldots,m$. The absolute difference between any two *AIFSs* is a non-negative integer multiple of a slot time σ.

In our derivation, a discrete and slotted time scale is adopted. Let t_n denote the time ending the n^{th} transmission attempt of the system. As illustrated in Figure 3-2, t_0, t_1, $t_2 \ldots$ represent such time points. The transmission between t_0 and t_1 is a successful attempt, and the one between t_1 and t_2 is a failed attempt due to collision. We setup the stochastic model in that the state of the process at time t_n is a vector of backoff counters *of m* virtual stations $B(n) = (b_n^{(1)}, b_n^{(2)}, \ldots, b_n^{(m)})$, where $b_n^{(i)}$ is the value of the backoff counter of the i^{th} virtual station at time t_n. The transmission could be either a successful one or a collision, hence the duration of two consecutive transmissions may be different. The state of the system changes as a transmission attempt occurs, and at the end of a transmission attempt, new backoff counter(s) are randomly generated for next transmission due to the overload condition. Since the state at time t_{n+1} only relies on the state at time t_n (it becomes clearer as the analysis proceeds), we can model this process as an m-dimensional discrete time Markov chain. The state space of this Markov chain is all possible combinations of $(b_n^{(1)}, b_n^{(2)}, \ldots, b_n^{(m)})$, where $b_n^{(i)}$ is any integer between $[1,\ CW_i + 1]$, $i = 1,2,\ldots,m$.

Next, we analyze and derive the one-step transition matrix P of the model. It will become clear that the state at time t_{n+1} is only based on the state at time t_n, so that this stochastic process is a valid a Markov chain.

From time t_n, each node starts or resumes the carrier sensing and backoff procedure in order to initiate a packet exchange. Some nodes will finish their backoff procedure earlier than others and proceed with data transmission. No matter whether the attempt succeeds or fails, in the end the system reaches the time point t_{n+1}. In other words, during the time interval $[t_n,\ t_{n+1}]$, there is at least one virtual station, VS, say virtual station *j*, VS_j, whose backoff counter reaches zero and triggers its transmission attempt, i.e., VS_j satisfies $(b_n^{(j)} + AIFS_j) = \min_{i \in (1,2,..m)} (b_n^{(i)} + AIFS_i)$. (Note there could be more than one virtual station satisfying the above minimum condition). To ease the description of the model, we denote *s,* $s \in (1,2,\ldots,m)$ as the number of virtual stations who execute the $(n+1)^{st}$ transmission attempt in $[t_n,\ t_{n+1}]$. Then $VS_{j_1}, VS_{j_2} \ldots VS_{j_s}$ are those *s* *VSs* that have to reset their backoff counters by randomly drawing an integer between $[1,\ CW_{j_r}+1]$, ($r = 1,2,\ldots s$) at time t_{n+1}. Moreover, with the different values of *s*, we can distinguish the result of transmission attempts into 3 mutually exclusive and exhaustive groups:

Group 1: successful transmission	when $s = 1$
Group 2: partial collision	when $1 < s < m$
Group 3: full collision	when $s = m$

The state of this Markov chain at time t_{n+1} becomes $B(n+1) = (b_{n+1}^{(1)}, b_{n+1}^{(2)} \ldots, b_{n+1}^{(m)})$, where $b_{n+1}^{(j_r)}$ is any integer equally likely chosen from $[1,\ CW_{j_r}+1]$, $r = 1,2,\ldots s$.

Those VSs whose backoff counters did not count down to zero within $[t_n,\ t_{n+1}]$ will have their values at t_{n+1} as:

$$b_{n+1}^{(l)}(l \neq j_1, j_2, \ldots, j_s) = \begin{cases} b_n^{(l)} & \text{if } AIFS_l \geq AIFS_{j_r} + b_n^{(j_r)} \\ b_n^{(l)} + AIFS_l - (AIFS_{j_r} + b_n^{(j_r)}) & AIFS_l < AIFS_{j_r} + b_n^{(j_r)} \end{cases}$$

(Note that $AIFS_{j_r} + b_n^{(j_r)}$ has the same value for r = 1, 2, ...s).

The first case is illustrated in Figure 3-4a, where the virtual station l uses an *AIFS* whose length is greater than or equal to the *AIFS* of virtual station j_r, $r = 1,2,\ldots s$, plus the backoff value, for at least one r, i.e., virtual station l is still sensing its *AIFS* when virtual station j_r starts transmission, so that virtual station l has its backoff value unchanged.

The second case is illustrated in Figure 3-4b, where the virtual station l uses an *AIFS* whose length is less than the *AIFS* of virtual station j_r, $r = 1,2,\ldots s$, plus the backoff value, i.e., virtual station l is in its count down procedure when virtual station j_r starts transmission, so that virtual station l will have its backoff value be $b_n^{(l)} + AIFS_l - (AIFS_{j_r} + b_n^{(j_r)})$ at time t_{n+1}.

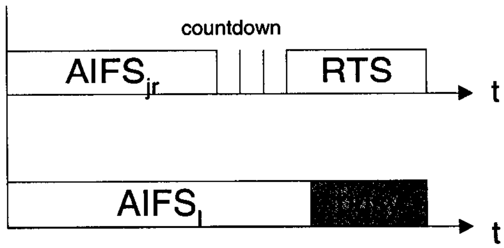


Figure 3-4a: Case one of countdown procedure.

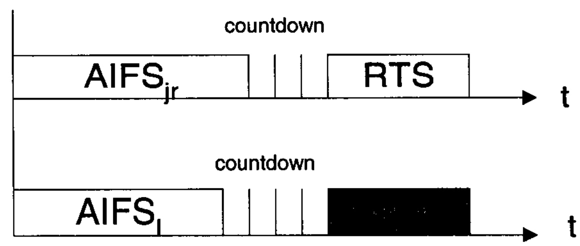


Figure 3-4b: Case two of countdown procedure.

The one-step transition probability $P_{B(n) \rightarrow B(n+1)}$ is:

$[(CW_{j_1}+1)\cdot(CW_{j_2}+1)\cdots(CW_{j_s}+1)]^{-1}$

Figure 3-5 illustrates the one-step transition diagram for successful and collision transmission situation.

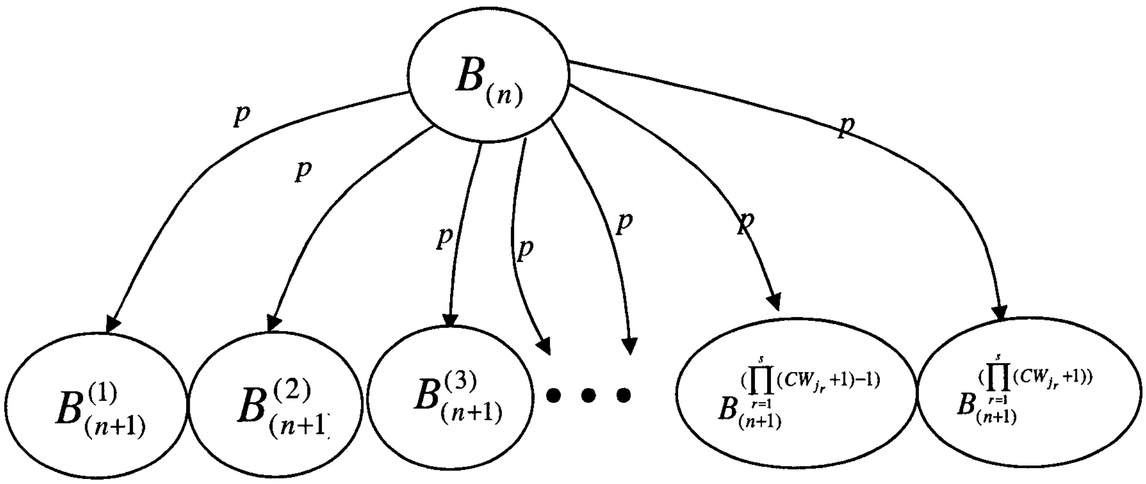


Figure 3-5: State transition diagram with collision. B(n) could transit to $[(CW(_{j1})+1)(CW(_{j2})+1)\ldots(CW(_{js})+1)]$ possible states, each with probability $p=1/[(CW(_{j1})+1)(CW(_{j2})+1)\ldots(CW(_{js})+1)]$.

3.3.2 Performance Analysis

The model can be regarded as a regenerative process. Starting from any full collision state, the system will eventually visit another full collision state; after that, the system will start over a probabilistic replica of the whole process, which ends with another full collision. We define the time interval between two full collision states as a round. Therefore, full collision states of this model are regenerative states and they restart probabilistically identical rounds. The theory of regenerative processes indicates the statistical characteristic in one round are identical to the long-run properties of the system [77, section 7.5]. The following analysis and derivations are based on one round of operation.

Specifically, we define the start of each round as leaving a regenerative state, and ends at the end of the following regenerative state. In other words, a round randomly starts from any valid Markov state after the last full collision. In order to compute the *throughput* and *access delay*, we adopt the theorem of *transient state and time to absorption* to our model.

In any round of operation, the regenerative state, i.e., the full-collision state $B(n)$ with the property $b_n^{(1)} + AIFS_1 = b_n^{(2)} + AIFS_2 = \cdots = b_n^{(m)} + AIFS_m$ can be taken as an absorbing state, in which case, all *VSs* are going to send their packets simultaneously within $[t_n, t_{n+1}]$ and reset their backoff counters at t_{n+1} consequently. The other states, including both successful and partial collision states, are transient states. Using the theorem on *time to absorption*, we can accurately calculate how many times on average the system will go through each transient state (either successful transmission or partial collision) before being absorbed into an absorbing state (full collision) starting from the beginning of a round. Note that the system operation will go on with another round, instead of being absorbed in full collision states.

A. Computation Procedure and some Notations

1. Compose one-step transition matrix P of the Markov chain; organize its layout so that the absorbing states are ordered before all transient states. The matrix P would have the form P=$\begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$, where I is an identity matrix, R consists of transition probabilities from transient states to absorbing states, and Q represents transition probabilities from transient states to transient states.

2. Calculate the absorption matrix $W = (I - Q)^{-1}$. The element of W, say w_{jk}, represents the average number of times state k is visited starting from state j, where j and k are both transient states.

3. Calculate the reaching probability matrix $F = W \cdot R$. The element of F, say F_{ji} represent the probability that the system ends up in absorbing state i starting from transient state j.

Denote G to be the set of transient states of the Markov chain corresponding to the matrix Q, and G^C the set of absorbing states. The transient states of the Markov chain can be further divided into m+1 subsets according to the transmission results (partial collision or successful) and the transmitter of the successful transmission. Let $G_f, G_1, G_2, ..., G_m$ denote those sets of states, where G_f is the set of states that results in a partial collision attempt, and G_i is the set of states that leads to a successful transmission from VS_i, $i \in (1,2,...,m)$.

B. System Throughput

We will calculate the *normalized throughput*, defined as the fraction of time when the channel is used to successfully transmit effective payload bits. To do so, we assume the system can equally likely start to operate from any one of its valid states. Let the total number of states be $|P|$, i.e., the number of rows of P. The total throughput of the system, S, can be expressed as,

$$S = \frac{E[PL]}{E[T_{round}]} = \frac{E[\textit{successfully transmitted payload}]}{E[\textit{SuccessTransTime}] + E[\textit{CollisionTime}] + E[\textit{IdleTime}]} \tag{3.1}$$

where $E[PL]$ is the average normalized (by channel rate) successfully transmitted payload (*effective payload*) during one round. It comprises successfully transmitted payload by all virtual stations. Let $E[PL_i]$ denote payload transmitted by virtual station VS_i. $E[PL]$ can be expressed as:

$$E[PL] = \sum_{i \in (1,2,...,m)} E[PL_i] \tag{3.2}$$

We notice that there could be more than one transient state (states belonging to set G_i) leading to a successful transmission by virtual station VS_i, therefore the effective payload transmitted by VS_i is the sum of transmissions from all states of G_i. While W_{jk}, ($j \in G$, $k \in G_i$) is the average time that the system stays in state k (leading to a successful transmission by VS_i) starting from any state j. Here we argue one system round starts from any valid state of the whole state space equally likely, which corresponds to the fact that the backoff counter value of any VS is chosen randomly based on its own CW. So, we can express the effective payload contributed by VS_i as:

$$\begin{aligned}
E[PL_i] &= E[PL_i \mid \textit{starting from absorbing states}] \cdot P\{\textit{starting from absorbing states}\} + \\
&\quad E[PL_i \mid \textit{starting from transient states}] \cdot P\{\textit{starting from transient states}\} \\
&= 0 + \left(E[P_i] \cdot \sum_{j \in G, k \in G_i} \sum \frac{W_{jk}}{|W|} \right) \frac{|W|}{|P|} \\
&= \frac{1}{|P|} \cdot E[P_i] \cdot \sum_{j \in G, k \in G_i} \sum W_{jk}
\end{aligned} \tag{3.3}$$

where $E[P_i]$ is the normalized average length of packets transmitted by virtual station VS_i ; $|W|$ is the number of transient states.

In (3.1), $E[SuccessTransTime]$ and $E[CollisionTime]$ refer to the channel busy time due to successful transmission and collision respectively. In the *IEEE* 802.11 specification, two access mechanisms are described, namely, the basic access mechanism and the *RTS-CTS* exchange access mechanism. The two access mechanisms would have different transmission time and collision time. In our research, we will only discuss the *RTS-CTS* exchange mechanism, which provides a better performance in the presence of high contention [76].

With *RTS-CTS* exchange, as shown in Figure 3-6, we can obtain the channel busy time needed for a successful transmission by the i^{th} virtual station (T_s^i) and the channel busy time due to collision (T_c)as:

$$
\begin{aligned}
T_s^i &= RTS + SIFS + \delta + CTS + SIFS + \delta + H + E[P_i] + SIFS + \delta + ACK + \delta \\
T_c &= RTS + PIFS + \delta
\end{aligned}
\tag{3. 4}
$$

where $H = PHY_{hdr} + MAC_{hdr}$ is the packet header, *PIFS* is the timeout for *CTS* packet, and δ is the propagation delay.

Similar to the calculation of the effective payload, the time spent on successful transmission consists of all successful transmission by all *VSs*, which can be expressed as (we shorten SuccessTransTime as *STT*):

$$
\begin{aligned}
E[STT] &= E[STT \mid starting\ from\ absorbing\ states] \cdot P\{starting\ from\ absorbing\ states\} \\
&\quad + E[STT_i \mid starting\ from\ transient\ states] \cdot P\{starting\ from\ transient\ states\} \\
&= 0 + \sum_{i\in(1,2,\dots m)} \sum_{j\in G, k\in G_i}\sum \left(\frac{W_{jk}}{|W|}\cdot T_s^i\right)\cdot\frac{|W|}{|P|} \\
&= \frac{1}{|P|}\cdot \sum_{i\in(1,2,\dots m)} T_s^i \cdot (\sum_{j\in G, k\in G_i}\sum W_{jk})
\end{aligned}
\tag{3.5}
$$

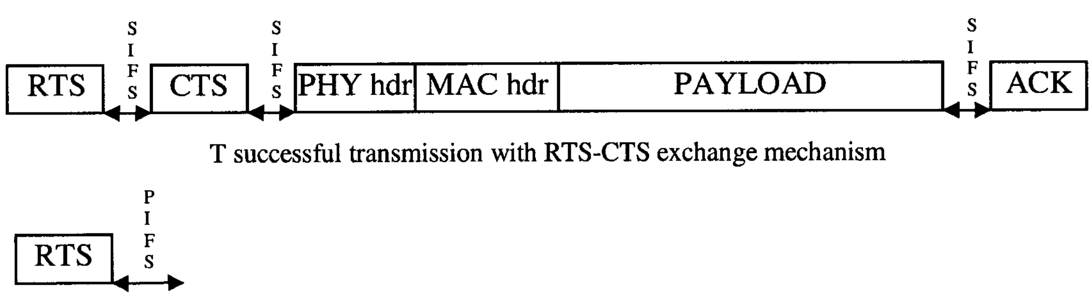


Figure 3-6: Computation of the length of transmission.

The average channel busy time due to collision consists of two parts. The first part is attributed to partial collisions before absorption, and the average length equals the product of the time length of each collision and the average number of partial collision in each round. The second part is the time length of collision attributed to the recurrent state (a full collision). The average channel busy time due to collisions during one round can be expressed as:

$$
\begin{aligned}
E[CollisionTime] &= E[\textit{length of all partial collisions}] + E[\textit{length of all full collision}] \\
&= E[\textit{number of partial collisions}] \cdot E[\textit{length of collision}] \\
&\quad + E[\textit{number of full collisions}] \cdot E[\textit{length of collision}] \\
&= \left(\sum_{j \in G, k \in G_f} \sum \frac{W_{jk}}{|W|} \cdot \frac{|W|}{|P|} \right) \cdot T_c + 1 \cdot T_c \\
&= \left(\frac{1}{|P|} \cdot \sum_{j \in G, k \in G_f} \sum W_{jk} + 1 \right) \cdot T_c \qquad (3.6)
\end{aligned}
$$

The channel will be idle while all virtual stations are sensing and waiting for a free channel. Therefore, all the three types of transmission: successful, partial collision and full collision contribute to the idle time. Given any state at time t_n, $B(n) = (b_n^{(1)}, b_n^{(2)}, \ldots, b_n^{(m)})$, let $j_1, j_2, \ldots, j_s, 1 \le s \le m$ be the virtual stations having their backoff counters count down to zero and sending packets simultaneously within $[t_n, t_{n+1}]$. Let I_n denote the total idle time before the transmission within $[t_n, t_{n+1}]$. Then I_n consists of two consecutive parts, namely, Arbitration InterFrame Space ($AIFS$) and the idle time due to the backoff counter counting down, so we can calculate it as $I_n = AIFS_{j_r} + b_n^{j_r}$, $r \in (1,2,\ldots s)$. The average idle time can be divided into two parts (we shorten IdleTime as IT):

$$E[IT] = \sum_{i \in G^c} E[IT \mid starting\ from\ state\ i] \cdot P\{starting\ from\ state\ i\}$$
$$+ \sum_{j \in G} E[IT \mid starting\ from\ state\ j] \cdot P\{starting\ from\ state\ j\}$$
$$= \frac{1}{|P|} \cdot \sum_{l \in G^c} I_l + \left[\sum_{j \in G, k \in G} \sum (W_{jk} \cdot I_k) + \sum_{j \in G} \sum_{l \in G^c} (F_{jl} \cdot I_l) \right] \cdot \frac{1}{|P|} \tag{3.7}$$

Plugging (3.2) - (3.7) into Eq. (3.1) and multiplying both the denominator and numerator by $|P|$, we get the system throughput as:

$$S = \frac{\sum_{i \in (1,2,\dots m)} \sum_{j \in G, k \in G_i} \sum (W_{jk} \cdot E[P_i])}{SP + CP + IP} \tag{3.8}$$

where

$$SP = \sum_{i \in (1,2,\dots,m)} \sum_{j \in G} \sum_{k \in G_i} (W_{jk} \cdot T_s^i), \tag{3.9}$$

$$CP = \sum_{j \in G} \sum_{k \in G_f} (W_{jk} \cdot T_c) + |P| \cdot T_c \text{ and} \tag{3.10}$$

$$IP = \sum_{j \in G} \sum_{k \in G} (W_{jk} \cdot I_k) + \sum_{j \in G} \sum_{l \in G^C} (F_{jl} \cdot I_l) + \sum_{l \in G^C} I_l \tag{3.11}$$

C. Throughput Ratios Between Flows

Now we can calculate the throughput ratios among the different flows. The effective payload sent by the i^{th} virtual station, $i \in (1,2,\dots,m)$, equals the product of the average number of successful transmissions and the average normalized packet length. So the throughput of the i^{th} virtual station can be expressed by

$$S_i = \frac{E[PL_i]}{E[T_{round}]} = \frac{\frac{1}{|P|} \cdot E[P_i] \cdot \sum_{j \in G, k \in G_i} \sum W_{jk}}{E[T_{round}]} \tag{3.12}$$

Therefore, the throughput ratios $S_1 : S_2 : \dots : S_m$ among virtual stations can be calculated as:

$$S_1 : S_2 : \dots : S_m = (E[P_1] \cdot \sum_{j \in G, k \in G_1} \sum W_{jk}) : (E[P_2] \cdot \sum_{j \in G, k \in G_2} \sum W_{jk}) : \dots : (E[P_m] \cdot \sum_{j \in G, k \in G_m} \sum W_{jk}) \tag{3.13}$$

and the throughput ratios between any two virtual stations , *u, v*, can be expressed as

$$S_u : S_v = (E[P_u] \cdot \sum_{j \in G,\, k \in G_u} \sum W_{jk}) : (E[P_v] \cdot \sum_{j \in G,\, k \in G_v} \sum W_{jk}) \tag{3.14}$$

D. Access Delay

We define the access delay as the time between a packet becoming Head of Line (*HOL*) and the starting time of its successful transmission attempt, which is confirmed by an *ACK* control frame (Figure 3-2, interval between t0 and t3, for a packet in virtual station B). Sometimes, the access delay is a more critical measurement to evaluate *QoS* in wireless networks [79]. Based on the analytical model, we can compute the average access delay for packets of individual flows. In particular, in one round of *EDCF* operation, the average access delay of the packets of the i^{th} flow, T_Q^i, is the total waiting time of the i^{th} flow divided by the total number of successful transmissions by the same flow. The total waiting time of the i^{th} flow, which includes all the time spent for sensing the channel idle, retransmission time due to collisions, as well as the successful transmission time by other flows, can also be interpreted as the average T_{round} less the total time for successful transmissions of the i^{th} *VS* within one round.

Following the derivation of (3.3) – (3.5) and (3.8) - (3.11), we obtain the average access delay of the i^{th} virtual station, T_Q^i:

$$E[T_Q^i] = \frac{E[T_{round}] - E[T_{success}^i]}{E[N_s^i]} = \frac{|P| \cdot E[T_{round}] - T_s^i \cdot \sum_{j \in G, k \in G_i} \sum W_{jk}}{\sum_{j \in G, k \in G_i} \sum W_{jk}} \tag{3.15}$$

Where E[$T^i_{success}$] is the average successful transmission time for the i^{th} virtual station and E[N^i_s] is the average number of successful transmission by the i^{th} virtual station during one round.

Chapter 4

Analytical Model Validation and Performance Evaluation

We will evaluate the *EDCF QoS* performance under different scenarios, based on the numerical results generated by the analytical model. The metrics we use in the study include throughput of the system, throughput and access delay of individual flows, and throughput ratio among flows. The throughput and access delay of individual flows indicate how the flows were served distinctively as a result of *TC*-specific prioritized parameters, while the throughput ratio reflects quantitatively the *QoS* differentiation, i.e., to what degree the *HP* flow has an advantage over *LP* ones.

4.1 Simulation Model and Analytical Model Validation

In order to validate the proposed analytical model, we compare it with a simulation model. Our simulations of the *EDCF WLAN* system are written in C++, while numerical results for the analytical model are obtained using *MATLAB*, which is a widely used mathematics software package and provides a wide spectrum of mathematics computation of wide spectrum. In our simulations, all virtual stations operate independently in the *RTS/CTS* mode of *EDCF* with a saturated traffic load. We try to have the simulation closely conform to the specifications in the *IEEE* 802.11 and 802.11E standards. However, we ignore the *PHY* layer, i.e., we assume an ideal and constant transmission environment. A summary of the constant parameter values used in both our analytical model and simulations are given in Table 4-1 (refer to [1], 15.3.3 *DS PHY* characteristics). For simplicity, we set the same constant payload size for the packets from all flows.

For the sake of convenience, we use a non-negative integer k to represent the *AIFS* length, noted as *AIFS=k*, which means the length of the *AIFS* is the length of a *DIFS* plus k slot times (σ). According to the 802.11E draft, k is chosen in the range [0,8].

Figure 4-1 and 4-2 show the experiment results in terms of throughput and throughput ratio of the analytical model (depicted by *MAT*, which applies to the following tables as well) versus simulation (I use *MAT* and *SIM* to denote analytical and simulation results respectively). Table 4-2 gives the numerical comparison of the two approaches. Figure 4-3 and 4-4 show the experiment results in terms of the access delay from the analytical model versus simulation. Table 4-3 gives the numerical results for access delay using the analytical model and simulation. The experiment scenario consists of two differentiated flows. The high priority flow has its *AIFS* fixed at 0 and the low priority flow has its *AIFS* increment from 0 to 8. Both flows use the same contention window size of 7. Simulation results in all scenarios have a 95% confidence interval lower than ± 0.012 error. The comparison illustrates that our analytical model is accurate.

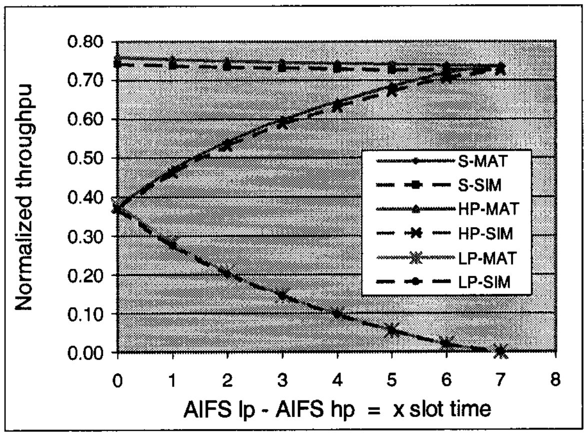
Moreover, all the results reported in next sections were also compared with those obtained by simulations, with negligible differences. However, we will proceed with our performance evaluation using only results from the analytical model.

Table 4-1: Constant parameters in analytical model and simulations.

Packet payload size (bits)	8196
PHY header (bits)	192
MAC header (bits)	272
RTS (bits)	160
CTS (bits)	112
ACK (bits)	112
WM transmission rate (bits/sec)	11M
Propagation delay (μs)	1
SIFS (μs)	10
Slot time σ (μs)	20
PIFS = SIFS + σ (μs)	30
DIFS = SIFS + 2σ (μs)	50

Table 4-2: Throughput comparison of analytical and simulation results.

AIFSlp-AIFShp	S		HP throughput		LP throughput		ratio	
	S-MAT	S-SIM	HP-MAT	HP-SIM	LP-MAT	LP-SIM	HP: LP MAT	HP : LP, SIM
0	0.759	0.740	0.379	0.371	0.379	0.369	1.000	1.004
1	0.753	0.735	0.471	0.460	0.283	0.275	1.665	1.669
2	0.749	0.731	0.542	0.530	0.207	0.201	2.626	2.634
3	0.745	0.729	0.598	0.585	0.147	0.144	4.071	4.058
4	0.742	0.726	0.644	0.630	0.099	0.096	6.526	6.561
5	0.739	0.724	0.684	0.670	0.055	0.054	12.393	12.365
6	0.737	0.723	0.717	0.704	0.020	0.020	35.352	35.644
7	0.735	0.724	0.735	0.724	0.000	0.000	inf	inf



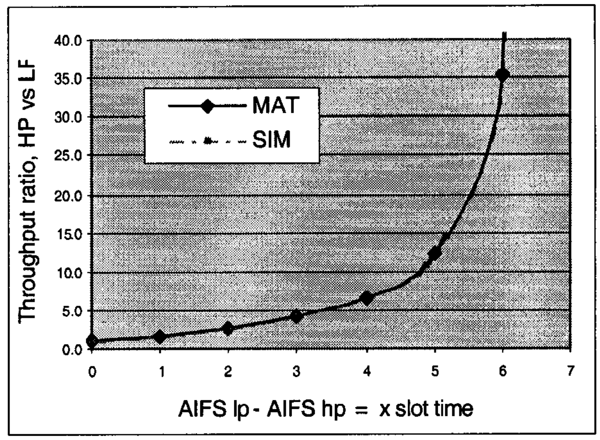
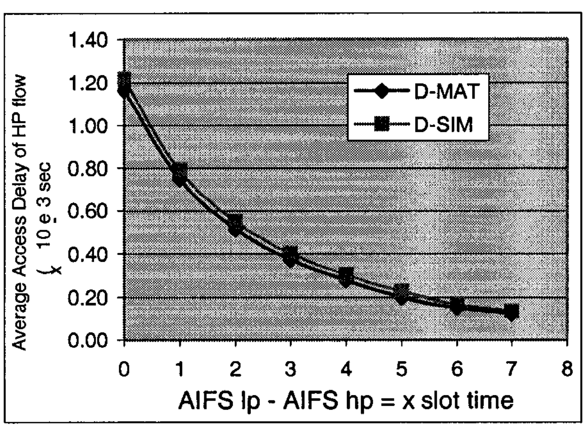


Figure 4-1: Flow throughputs vs *AIFS* difference.

Figure 4-2: Throughput ratio vs *AIFS* difference.

Table 4-3: Delay comparison of analytical and simulation results.

$AIFS_{lp}$-$AIFS_{hp}$	Delay of HP flow (10e-3 sec)		Delay of LP flow (10e-3 sec)	
	D-MAT	D-SIM	D-MAT	D-SIM
0	1.1632	1.2147	1.1632	1.2233
1	0.7446	0.7867	1.9022	1.9783
2	0.5146	0.5500	2.9701	3.0748
3	0.3733	0.4036	4.5767	4.7264
4	0.2770	0.3046	7.3096	7.5341
5	0.2018	0.2261	13.8435	14.2716
6	0.1497	0.1685	39.4228	40.4882
7	0.1230	0.1348	Inf	inf



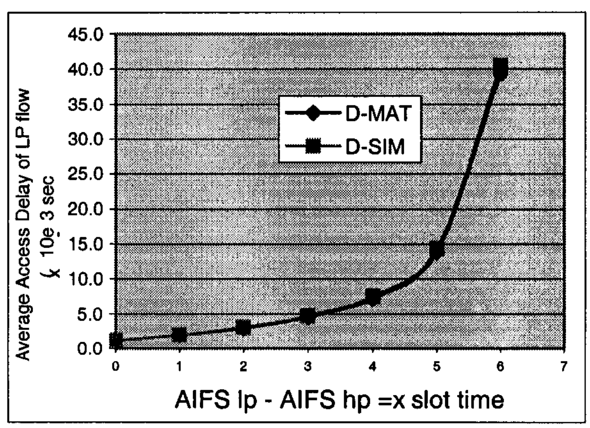


Figure 4-3: Access delay of *HP* flow vs *AIFS* difference.

Figure 4-4: Access delay of *LP* flow vs *AIFS* difference.

4.2 Experiment Design

We designed three groups of experiments as follows.

4.2.1 Experiment One – Differentiated *AIFSs*

The first group of experiments is to investigate the role of differentiated *AIFSs* on *QoS* differentiation when every virtual station uses the same contention window value, i.e., *AIFSs* are the sole differentiated parameters in the experiments. The experiment scenario consists of two independent flows, which belong to two different traffic categories. The two traffic categories use the same contention window value of 7. During the experiments, the *AIFS* of the high priority traffic category is fixed at $AIFS_{HP}$ = 0, and the *AIFS* of the low priority traffic category, $AIFS_{LP}$, increments from 0 until 8.

4.2.2 Experiment Two – Differentiated *AIFSs* and Contention Windows

The second group of experiments is to explore the compound impact of differentiated *AIFSs* and contention windows. While using the first scenario as stated in the first experiment, we design another two scenarios. The second scenario consists of two flows and each flow has the same parameters as the above experiments, except that both flows use *CW*=15; still the *HP* flow uses an *AIFS* equal to 0 and the *LP* flow increases its*AIFS* length by one slot time each time, starting from 0. The third scenario also consists of two prioritized flows. However, the *HP* flow has $CW_{HP} = 7$, the *LP* flow has $CW_{LP} = 15$. The *LP* flow increases its *AIFS* length by one slot time each time and the *HP* flow uses the fixed *AIFS* of 0, i.e., both *AIFSs* and *CWs* are differentiated.

4.2.3 Experiment Three – Traffic Loading Change

The third experiment is to explore the effect of a traffic load change on the performance of different traffic category flows. The observation is based on three experiment scenarios.

The first scenario is the same as the third scenario in the second experiment. In the second scenario, we introduced another *LP* flow identical to the original *LP* flow in the first scenario, i.e., one *HP* flow and two *LP* flows. In the third scenario, we introduced another *HP* flow based on scenario one, i.e., two *HP* flows and one *LP* flow. We will compare the performance of those differentiated flows in terms of throughput and access delay, so that we can explore the effect of traffic change.

4.3 Experiment Results and Performance Evaluation

4.3.1 The Effect of *AIFSs* on *QoS*

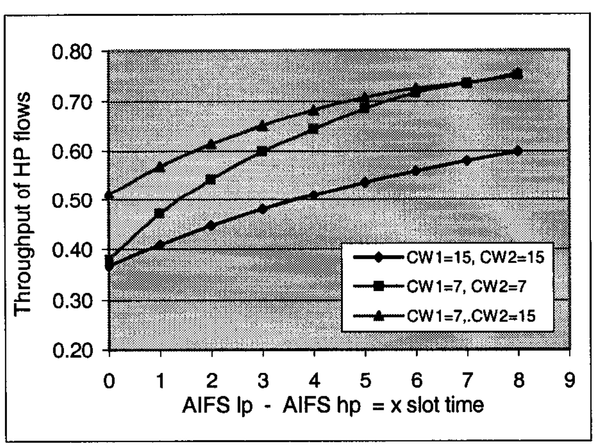
For the first experiment, we collected and illustrated the simulated and analytical numerical results in Tables 4-2 and 4-3, and Figures 4-1, 4-2, 4-3 and 4-4.

As expected, the two flows have approximately the same share of bandwidth and experience the same average access delay when the two flows use the same parameters. As the *AIFS* of the *LP* flow increases, the *HP* flow achieves more and more bandwidth and less access delay; on the other hand, the *LP* flow shares less bandwidth and experience longer delay. The degree of incline on the bandwidth of the *HP* flow decreases as the *AIFS* of the *LP* flow increases, correspondingly, the degree of decline on the access delay of the *HP* flow decreases as the *AIFS* of the *LP* increases. At the same time, the degree of decline on the bandwidth of the *LP* flow decreases as the *AIFS* of the *LP* flow increases; consequently, the degree of incline on the access delay of the *LP* flow increases as *AIFS* of *LP* flow increases. The throughput ratio of *HP* flow vs *LP* flow increase significantly as the *AIFS* of the *LP* flow increase. The degree of incline increases dramatically as the *AIFS* of the *LP* flow increases since the bandwidth of the *HP* flow climbs up while the bandwidth of *LP* flow drops down, i.e., a twofold effect. In summary, the differentiation caused by the *AIFS* difference between the two flows is not linear and the *AIFS* difference has a significant effect on differentiation.

4.3.2 The Compound Effects of *AIFS* and *CW* on *QoS*

Figures 4-5 and 4-6 illustrate throughput of the *HP* flow and the *LP* flow of the three scenarios respectively. Figure 4-7 illustrates the throughput ratio of the *HP* flow versus the *LP* flow. Figure 4-8 and 4-9 illustrate the access delay of the *HP* flow and *LP* flow of the three sets of experiments, respectively.

We have the following observations. First of all, as the first experiment reveals, the *AIFS* difference results in *QoS* differentiation. Moreover, the contention window size difference will intensify the *QoS* differentiation in term of both throughput and access delay, as Figure 4-7 indicates that a difference in both the contention window and the *AIFS* leads to the highest throughput ratio. Actually, the combination of differentiated *AIFSs* and differentiated contention window size introduce more compound and significant effect on the degree of differentiation. We also find that when the differentiated flows use the same contention window size, the smaller the contention windows are the more significant influence the *AIFS* difference will have on the *QoS* differentiation, as scenario 1 has higher *QoS* differentiation than scenario 2.



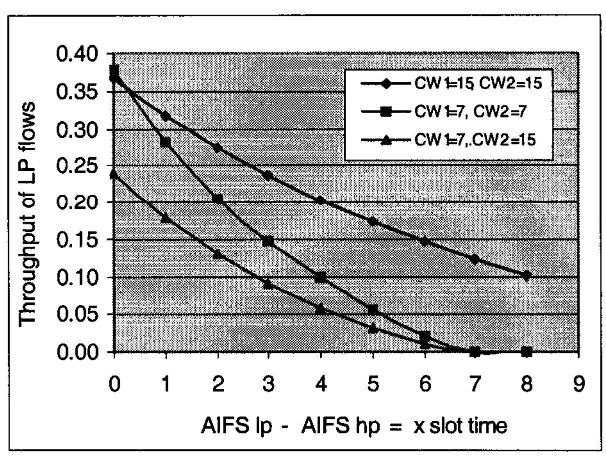


Figure 4-5: Throughput of *HP* flow vs *AIFS* difference. Figure 4-6: Throughput of *LP* flow vs *AIFS* difference.

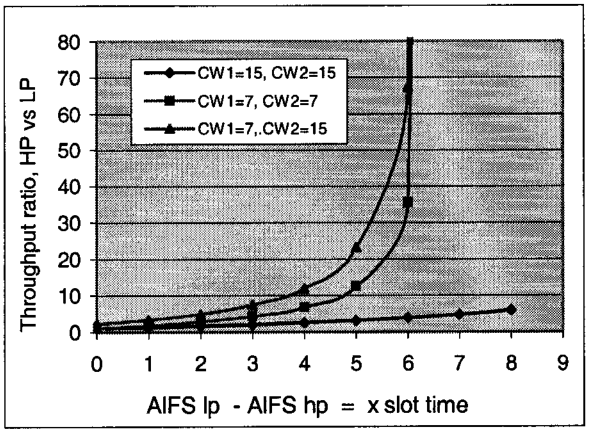


Figure 4-7: Throughput ratio vs *AIFS* difference.

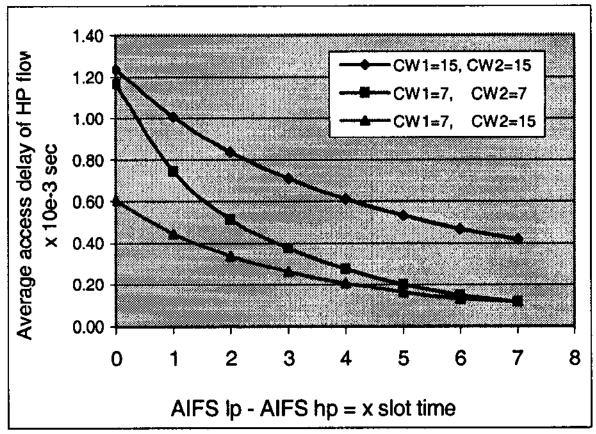


Figure 4-8: Access delay of *HP* flow vs *AIFS* difference.

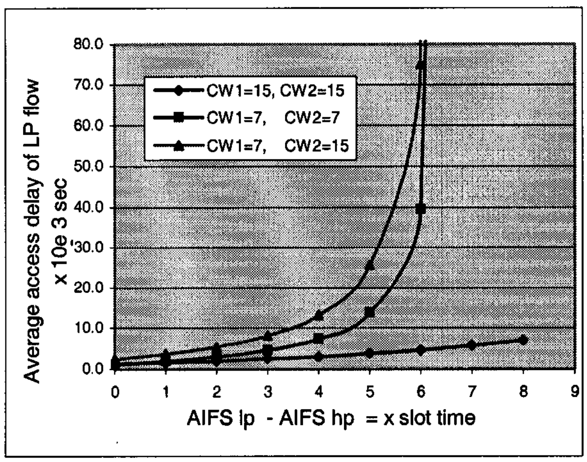


Figure 4-9: Access delay of *LP* flow vs *AIFS* difference.

Moreover, we argue that it is the *AIFS* difference, rather than the absolute *AIFS* values, that affects the degree of *QoS* differentiation. As displayed in Figure 4-7, the increment of throughput ratio accelerates with the growth of the *AIFS* difference nonlinearly. The larger the *AIFS* difference, the faster the ratio increases, and in the worst case, the *LP* flow may not receive access to the wireless medium at all. To further examine the above argument, we proceed with the fourth scenario, in which the difference of *AIFSs* between the *HP* and the *LP* priority flows remains unchanged, i.e., $AIFS_{LP} - AIFS_{HP} = 2$, $AIFS_{HP}$ is set to 0,1,...6

and $AIFS_{LP}$ is set to 2,3,...8 consecutively, while CW_{HP} is 7 and CW_{LP} is 15. As expected, the throughput ratios (Figure 4-10) remain unchanged, and the access delay (Figure 4-11) increases slightly, though the absolute *AIFS* values of both *TCs* increase synchronously.

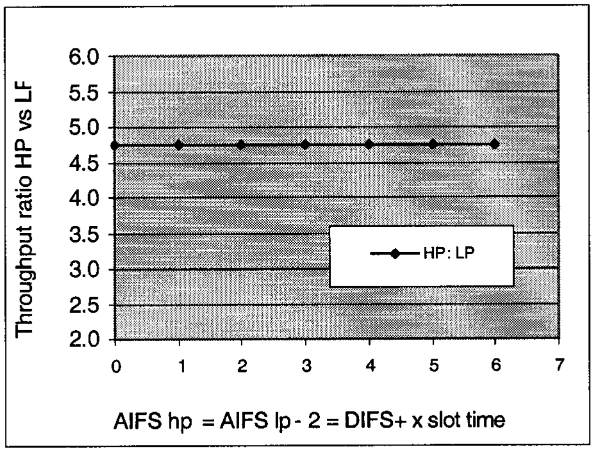


Figure 4-10: Throughput ratio vs *AIFS* difference.

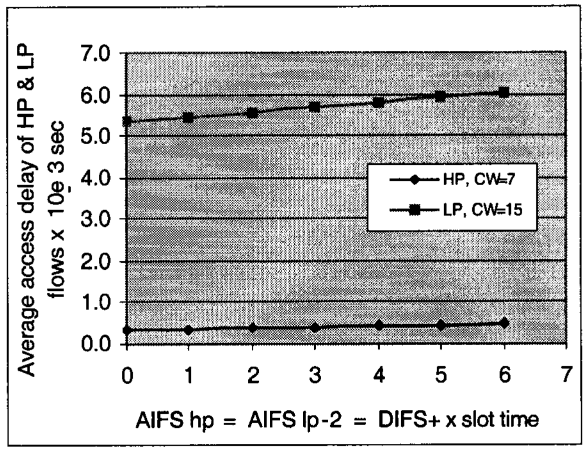


Figure 4-11: Access delay vs *AIFS* difference.

4.3.3 Effect of Traffic Loading on *QoS*

Here we evaluate the performance when the traffic loads change. Figures 4-12, 4-13, 4-14, 4-15 and 4-16 respectively show the results of the throughput of an *HP* flow, an *LP* flow, the throughput ratio between an *HP* and an *LP* flow, and access delay of an *HP* flow and an *LP* flow, for the cases of: (a) one *HP* flow and one *LP* flow, (b) one *HP* flow and two *LP* flows, and (c) two *HP* flows and one *LP* flow. Here, *HP* flows and *LP* flows use the same parameters as those in scenario 3 of the second experiment. We plot the throughput of only one *HP* flow or one *LP* flow when there are two such flows, since the analytical model derives the same numerical results for all flows of the same class. From these results we can see that the *HP* flows can dominate the wireless medium and are less affected by *LP* flows, as one more *LP* flow won't change the performance line too much compared to the performance line for the first scenario (one *HP* flow and one *LP* flow). However, the *HP* flow imposes great influence on *QoS* differentiation. It not only affects *LP* flow service, but also *HP* flow service. A new *HP* flow in scenario three can almost grab half the bandwidth from the old *HP* flow or cause the access delay to increase by several times. We also find that a small difference in the *AIFS* (one slot time) between the two flows can result in a considerably large difference in their throughput and access delay. For example, as shown in Figure 4-14, the ratio is 4.258 in the two-*HP*-one-*LP* scenario, which implies that the HP flows can take much more advantages than LP flow*s*. Similar pheromones can be found in Figures 4-15 and 4-16 in term of access delay.

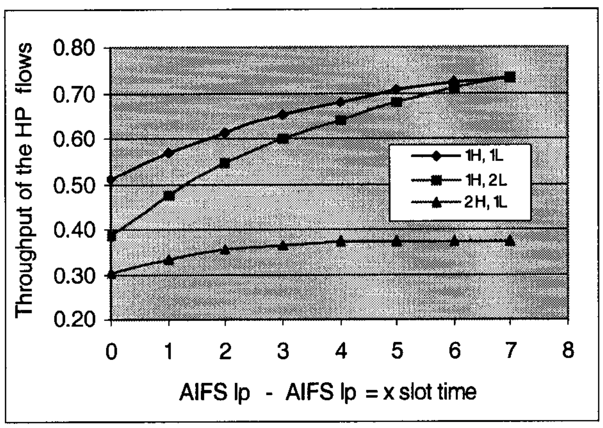


Figure 4-12: *HP* flow throughput vs. traffic load change

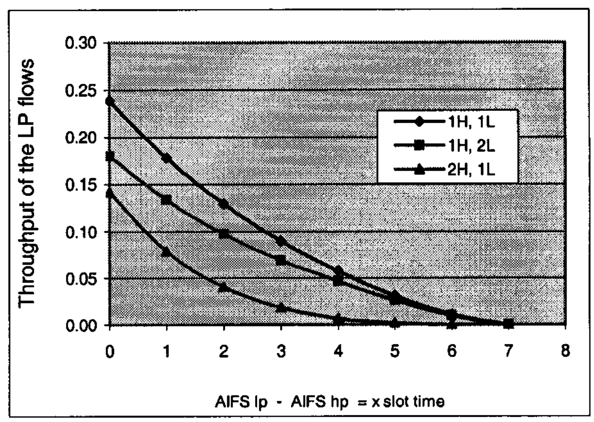


Figure 4-13: *LP* flow throughput vs. traffic load change.

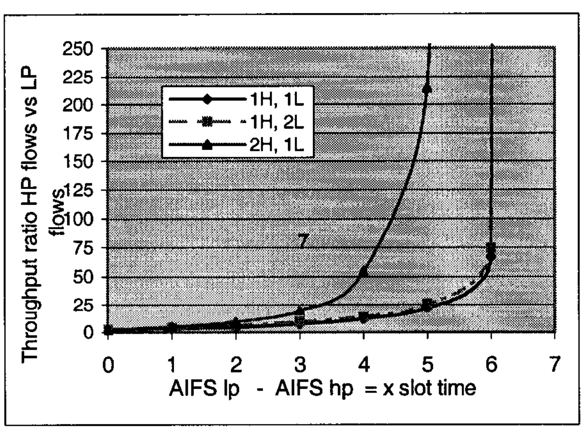


Figure 4-14: Flow throughput ratio vs. traffic load change.

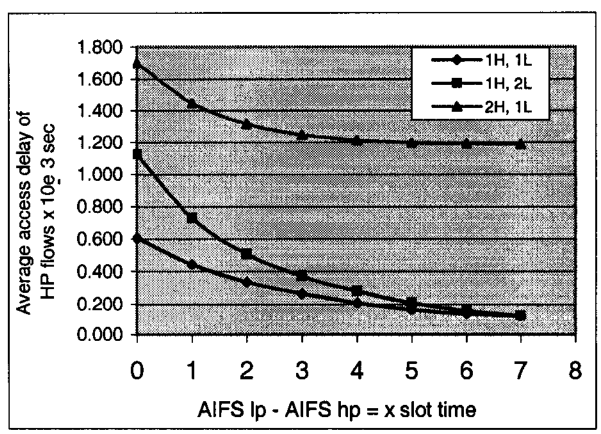


Figure 4-15: Access delay of *HP* flow vs traffic load change.

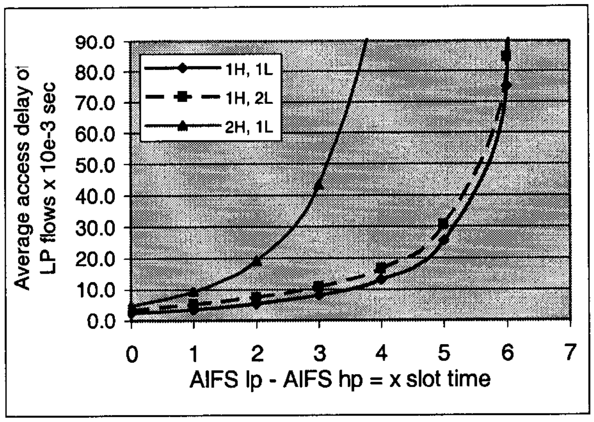


Figure 4-16: Access delay of *LP* flow vs. traffic load change.

4.4 Additional Discussion

In order to advocate the argument that collisions are relatively infrequent occurrences compared to successful transmissions, we further investigate the experimental results based on the analytical model. Table 4-4 summarizes the results of the average number of collisions and successful transmissions during one round of operation, in two system scenarios.

The first scenario has two *HP* flows and one *LP* flow. Both *HP* flows use the parameters $AIFS_{hp}$=0 and CW_{hp}=7 while the *LP* flow uses the parameters $AIFS_{lp}$=3 and CW_{lp}=15. The ratio of collision over the total transmission attempts is 0.139.

The second scenario consists of one *HP* flow and two *LP* flows. The *HP* flow and the *LP* flows use the same parameters as the first scenario. The ratio of collision over the total transmission attempts is 0.094.

We set *PF* to one throughout all experiments, which may not be desirable in the real-world system. However, since the collisions are infrequent during the operation of system, our experimental results should be able to represent the cases for all valid *PF* values. Furthermore, a *PF* value greater than one might help to reduce the happening of collisions, so the throughput results using *PF* equal to one would be an upper bound value in the corresponding experimental cases.

Table 4-4: A summary of collisions and transmission attempts

	Scenario ONE -- Two HP and one LP	Scenario TWO -- One HP and Two LP
Transmission attempts in one round	196.9	194.9
Collisions in one round	27.3	16.8
Ratio of collisions over total transmission	0.139	0.094

Chapter 5

Concluding Remarks

In this thesis, we developed an accurate multi-dimensional Markov model to analyze the performance of the *IEEE* 802.11e Enhanced Distributed Coordination Function (*EDCF*) MAC protocol. In particular, the model was used to evaluate the ability of the *IEEE* 802.11e *EDCF* to provide differentiated *QoS*.

We derived formulations of saturation throughput, throughput ratio of flows and access delay as functions of the prioritized parameters. These were used to conduct quantitative analysis of the impact of parameters such as *AIFS* and *CW* on the performance of prioritized flows.

Numerical results show that *EDCF* can well support differentiated *QoS*, and it provides significant advantage to higher priority flows. We have shown that *AIFS* has a significant impact on the *TC* priority. Regardless of other parameters, the smaller the *AIFS* is, the higher the priority of a flow, and the shorter time the flow has to wait before transmitting. This in turn translates into higher bandwidth share of the *HP* flow. Hence, we conclude that the *IEEE* 802.11 *EDCF* function can effectively provide *QoS* differentiation.

To the best of our knowledge, no research taking the mathematical analysis approach to study *QoS* support in *EDCF* has been reported in the literature. The research method and results of this thesis should provide useful hints for understanding the characteristics of distributed *QoS MAC* protocol and further research can be better planned and designed based on our effort.

Future research may include the following.

(1) Alternative mathematical solutions may simplify the calculation. The current model has the disadvantage of lack of scalability; the state space increases exponentially with the increase number of flows and the contention window size. As the state space increases, it becomes hard to solve the W matrix using the currently available computing facilities. A possible solution to scalability problem is to modify the mathematical model, so that the increase of state space is within a manageable range.

(2) In this thesis, we have explored the interaction among priority flows and their *QoS* parameters. The analytical model and the numerical result can be used for the purpose of designing the distributed scheduling algorithm, which can dynamically adjust the *QoS* parameters of each virtual stations according to fluctuations of network conditions [38, 39].

(3) It has been shown that the *EDCF* provides significant Medium Access Control (*MAC*) *QoS* differentiation in a distributed manner. However, *QoS MAC* itself is not enough to provide (soft) *QoS* guarantees for delay-sensitive and bandwidth-intensive multimedia applications. Due to the fluctuation of network traffic and the dynamic status of network resource usage, a framework, which consists of admission control, rate control and traffic shaping, *QoS MAC (EDCF)* and packet scheduling should be devised.

Reference

[1] L. Kleinrock and F.A. Tobagi, "Packet Switching in Radio Channels: Part I-Carrier Sense Multiple Access Modes and Their Throughput-Delay Characteristics", IEEE Transactions on Communications, Vol. Com-23, No. 12, December 1975, pp. 1400-1416.

[2] S. Corson and J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC2501, 1999.

[3] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", RFC3561, 2003.

[4] G. Girling, J. Li, K. Wa, P. Osborn, and R. Stefanova, "The PEN Low Power Protocol Stack", IEEE ICCCN 2000, pp. 282-288, Las Vegas, Nevada, USA.

[5] The CHOICE Network Project, http://ramp.ucsd.edu/pawn/Choice.html, September 1999.

[6] J. Weatherall and A. Jones, "Ubiquitous Networks and Their Application", IEEE Wireless Communications, February 2002, pp.18-29.

[7] P. Bahl, W. Russell, Y.M. Wang, A. Balachandran, G. M. Voelker, and A. Miu, "PAWNs: Satisfying the Need of Ubiquitous Secure Connectivity and Location Services", IEEE Wireless Communications, February 2002, pp. 40-48.

[8] C-K Toh, "Ad Hoc Mobile Wireless Networks, Protocols and Systems", Prentice Hall PTR, c2002.

[9] C. Schurgers, V. Tsiatsis, S. Ganeriwal, and M. Srivastava, "Optimizing Sensor Networks in the Energy-Lantecy-Density" Design Space, IEEE Transactions on Mobile Computing, Vol. 1, No. 1, January-March 2002, pp. 70-80.

[10] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", IEEE Transactions on Wireless Communications, Vol. 1, No. 4, October 2002, pp. 660-670.

[11] C.E. Jones, K.M. Sivalingam, P. Agrawal, and J.C. Chen, "A survey of Energy Efficient Network Protocols for Wireless Networks", ACM/Baltzer Journal on Wireless Networks, Vol. 7, 2001, pp. 343-358.

[12] A. Chandrakasan and R. W. Brodersen, "Low Power Digital CMOS Design", Kluwer Academic, Norwell, USA, 1995.

[13] N.S. Fahmy, T.D. Todd, and V. Kezys, “Ad Hoc Networks with Smart Antennas Using IEEE802.11-Based Protocols”, Vol. 5, IEEE ICC 2002, Vol. 5, pp. 3144-3148, New York City, New York, USA.

[14] T. You, C.H. Yeh, and H. Hassanein, “A New Class of Collision Prevention MAC Protocols for Wireless Ad Hoc Networks”, IEEE ICC 2003, Vol. 2, pp. 1135-1140, Alaska, USA.

[15] Y.C. Tseng, C.S. Hsu, and T.Y. Hsieh, “Power-Saving Protocols for IEEE 802.11 Based Multi-Hop Ad Hoc Networks”, IEEE INFOCOM 2002, Vol. 1, pp. 200-209, New York City, New York, USA.

[16] E.S. Jung and N.H. Vaidya, “An Energy Efficient MAC Protocol for Wireless LANs”, IEEE INFOCOM 2002, Vol. 3, pp. 1756-1764, New York City, New York, USA.

[17] P. Lettieri, C. Fragouli, and M.B. Srivastava, “Low Power Error Control for Wireless Links”, ACM/IEEE MOBICOM 1997, pp. 139-150, Budapest, Hungary.

[18] P. Lettieri and M.B. Srivastava, “Adaptive Frame Length Control for Improving Wireless Link Throughput, Range, and Energy Efficiency”, IEEE INFOCOM 1998, Vol. 2, pp. 564-571, San Francisco, California, USA.

[19] B. Narendran, J. Sienichi, S. Yajnik, and P. Agrawal, “Evaluation of an Adaptive Power and Error Control Algorithm for Wireless Systems”, IEEE Communications Conference Montreal, Vol. 1, pp. 349-355, Canada 1997.

[20] A. Safwat, H. Hassanein, and H. Mouftah, “Power-Aware Fair Infrastructure Formation for Wireless Mobile Ad Hoc Communications”, IEEE GLOBECOM 2001, Vol. 5, pp. 2832 -2836, San Antonio, Texas, USA.

[21] W. Yu and J. Lee, “DSR-based Energy-aware Routing Protocols in Ad Hoc Networks”, ICWN 2002, Las Vegas, Nevada, USA.

[22] K. Woo, C. Yu, D. Lee, and H.Y. Youn, “Non-Blocking, Localized Routing Algorithm for Balanced Energy Consumption in Mobile Ad Hoc Networks”, ACM/IEEE MASCOTS 2001, pp. 117-124, Cincinnati, Ohio, USA.

[23] M. Zorzi, and R.R. Rao, “Energy Efficiency of TCP in a Local Wireless Environment”, Mobile Networks and Applications 6, 265-278, 2001, Kluwer Academic Publishers.

[24] IEEE 802.11 WG, “Reference number ISO/IEC 8802-11:1999(E) IEEE Std 802.11, 1999 Edition. International Standard [for] Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks-Specific Requirements- Part 11: Wireless LAN Medium Access Control (MAC) AND Physical Layer(PHY) Specifications”, 1999.

[25] T. Karygiannis and L. Owens, "Wireless Network Security 802.11, Bluetooth and Handheld Devices", NIST special publication 800-48.

[26] Proxim Wireless Network, "Wireless Network Security", http://www.proxim.com/learn/library/whitepapers/index.html, 2003.

[27] W.A. Arbaugh, N. Shankar, and Y.C.J. Wan, "Your 802.11 Wireless Network Has No Clothes", IEEE Wireless Communications, December 2002, pp. 44-51.

[28] N. Borisov, I. Goldberg, and D. Wagner, "Intercepting Mobile Communications: The Insecurity of 802.11", ACM MOBICOM 2001, pp. 180-189, Rome, Italy.

[29] N. Borisov, et., al, "Security of the WEP Algorithm", http://www.isaac.cs.berkeley.edu/isaac/wep-faq.html, 2001.

[30] L. Zhou and Z.J. Hass, "Securing Ad Hoc Networks", IEEE Network, November/December 1999.

[31] D. Qiao and S. Chio, "Goodput Enhancement of IEEE 802.11a Wireless LAN via Link Adaptation", IEEE ICC2001, Vol. 7, pp. 1995-2000, Helsinki, Finland.

[32] A. Mercado and K.J.R. Liu, "Adaptive QoS for Wireless Multimedia Networks Using Power Control and Smart Antennas", IEEE Transaction on Vehicular Technology, Vol. 51, No. 5, September 2002, pp. 1223-1233.

[33] S.H. Marikar and L.A. Dasilva, "Resource Management in Third Generation Mobile Communication Systems Employing Smart Antennas", IEEE INFOCOM 2002, Vol. 3, pp. 1461 -1469, New York City, New York, USA.

[34] D. Wu and R. Negi, "Utilizing Multiuser Diversity for Efficient Support of Quality of Service over a Fading Channel," IEEE ICC 2003, Vol. 3, pp. 2202 -2207, Alaska, USA.

[35] A.K. Parekh and R. G. Gallagher, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: the Single-Node Case," IEEE/ACM Transaction on Networking, vol. 1, 1993, pp. 344-357.

[36] S. Lu, V. Bharghavan, and R. Sirkant, "Fair Scheduling in Wireless Packet Networks," IEEE/ACM Transactions on Networking, Vol. 7, No.4, August 1999, pp. 473-489.

[37] H. Fattah and C. Leung, "An Overview of Scheduling Algorithms in Wireless Multimedia Networks", IEEE Wireless Communications, October 2002, pp. 76-83.

[38] N.H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in a Wireless LAN", IEEE MOBICOM 2000, pp. 167-178, Boston, Massachusetts, USA.

[39] A. Banchs and X. Perez, "Distributed Weighted Fair Queuing in 802.11 Wireless LAN", IEEE ICC 2002, Vol. 5, pp. 3121-3127, New York City, New York, USA.

[40] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly, "Ordered Packet Scheduling in Wireless Ad Hoc Networks: Mechanisms and Performance Analysis", IEEE/ACM MOBIHOC 2002, pp. 58-70, Lausanne, Switzerland.

[41] C. Zhu and M.S. Corson, "QoS Routing for Mobile Ad Hoc Networks", IEEE INFOCOM 2002, Vol. 2, pp. 958-967, New York City, New York, USA.

[42] IEEE 802.11 WG, Draft Supplement to STANDARD FOR Telecommunications and Information Exchange Between Systems – LAN/MAN Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control (MAC) Enhancement for Quality of Service(QoS)," 2001.

[43] N. Abramson, "Development of the ALOHANET," IEEE Transactions on Information Theory, Vol. IT-31, March 1985, pp. 119-123.

[44] K. Chandran, S. Raghunathan, S. Venkatesan, and R. Prakash, "A Feedback-Based Scheme for Improving TCP Performance in Ad Hoc Wireless Networks," IEEE Personal Communication, February 2001, pp. 34-39.

[45] J. Liu and S. Singh, "ATCP: TCP for Mobile Ad Hoc Networks,", IEEE Journal on Selected Area in Communications, Vol. 19, July 2001, pp. 1300-1315.

[46] T.S. Rappaport, "Wireless Communications –Principles and Practice", Prentice Hall, 1996.

[47] R. J. Haines and A. H. Aghvami, "Indoor Radio Environment Considerations in Selecting a Media Access Control Protocol for Wideband Radio Data Communications," IEEE ICC 1993, Vol. 2, pp. 990-994, Geneva, Switzerland.

[48] J. Deng and Z.J. Haas, "Dual Busy Tone Multiple Access (DBTMA): A New Medium Access Control for Packet Radio Networks", IEEE ICUPC 1998, Vol. 2, pp. 973-977, Florence, Italy.

[49] Z.J. Haas and J. Deng, "Dual Busy Tone Multiple Access (DBTMA) – Performance Evaluation", IEEE VTC 1999, Vol. 1, pp. 314-319, Houston, Texas, USA.

[50] ETSI, HIPERLAN Functional Specification, ETSI draft standard, http://www.etsi.org, July 1995.

[51] G. Anastasi, L. Lenzini, and E. Mingozzi, "Stability and Performance Analysis of HIPERLAN", IEEE INFOCOM 1998, Vol. 1, pp. 134-141, San Francisco, CA, USA.

[52] J.L. Sobrinho and A.S. Krishnakumar, "Real-Time Traffic over the IEEE802.11 Medium Access Control Layer", Bell Labs Technical Journal, 1996.

[53] J.L. Sobrinho and A.S. Krishnakumar, "Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Networks", IEEE Journal on Selected Areas in Communications, vol17, no 8, 1999, pp. 1353–1368.

[54] J.F. Frigon, V.C.M. Leung, and H. Chan, "Dynamic Reservation TDMA Protocol for Wireless ATM Netowrks", IEEE Journal on Selected Areas in Communications, Vol. 19, No. 2, Feb. 2001, pp. 370-383.

[55] J. Tourrilhes, "PiggyData: Reducing CSMA/CA Collisions for Multimedia and TCP Connections", http://www.hpl.hp.com/personal/Jean_Tourrilhes/Papers/PiggyData.pdf, 1999.

[56] IEEE 802.11 WG, "Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks-Specific Requirements- Part 11: Wireless LAN Medium Access Control (MAC) AND Physical Layer(PHY) Specifications", 1997.

[57] V. Bharghavan, "MACAW: A Media Access Protocol for Wireless LAN's", ACM SIGCOMM 1994, pp. 212-225, London, UK.

[58] D. Qiao and K.G. Shin, "Achieving Efficient Channel Utilization and Weighted Fairness for Data Communications in IEEE 802.11 WLAN under the DCF", IEEE/IFIP/ACM IWQoS 2002, pp. 227-236, Florida, USA.

[59] T. Ozugur, "Optimal MAC-Layer Fairness in 802.11 Networks", IEEE ICC 2002, Vol. 2, pp. 675-681, New York City, New York, USA.

[60] T. Ozugur, "Weighted Hierarchical Backoff Algorithm for Wireless Ad Hoc Networks", IEEE GLOBECOM 2001, Vol 5, pp. 2937-2943, San Antonio, Texas, USA.

[61] T. Ozugur, M. Naghshineh, P. Kermani, and J.A. Copeland, "Fair Media Access for Wireless LANs", IEEE GLOBECOM 1999, Vol 1b, pp. 570-579, Rio De Janeiro, Brazil.

[62] B. Bensaou, Y. Wang, and C.C. Ko, "Fair Medium Access in 802.11 Based Wireless Ad-Hoc Networks", IEEE/ACM MOBIHOC 2000, pp. 99-106, Boston, Massachusetts, USA.

[63] T. Nandagopal, T.-E. Kim, X. Gao, and V. Bharghavan, "Achieving MAC Layer Fairness in Wireless Packet Networks", ACM MOBICOM 2002, pp. 87-98, Boston, Massachusetts, USA.

[64] A. Veres, A. T.Campbell, M. Barry, and L.H. Sun, "Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control", IEEE Journal on Selected Areas in communications, Vol.19, No.3, OCT. 2001, pp. 2081-2093.

[65] M. Barry, A.T. Campbell, and A. Veres, "Distributed Control Algorithms for Service Differentiation in Wireless Packet Networks", IEEE INFOCOM 2001, Vol. 1, pp. 582-590, Anchorage, Alaska, USA.

[66] I. Aad and C. Castelluccia, "Differentiation Mechanisms for IEEE802.11", IEEE INFOCOM 2001, Vol. 1, pp. 209-218, Anchorage, Alaska, USA.

[67] S.S. Kang and M.W. Mutka, "Provisioning Service Differentiation in Ad Hoc Networks by Modification of the Backoff Algorithm", IEEE ICCCN 2001, pp. 577-580, Phoenix, Arizona, USA.

[68] H.K. Wu and P.H. Chuang, "Dynamic QoS Allocation for Multimedia Ad Hoc Wireless Networks", Mobile Networks and Application 6, 2001, pp. 377-384.

[69] A. Muir and J. Garcia-Luna-Aceves, "Supporting Real-Time Multimedia Traffic in a Wireless LAN", SPIE Multimedia Computing and Networking 1997, San Jose, California, USA.

[70] N.H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in a Wireless LAN", ACM MOBICOM 2000, pp. 167-178, Boston, Massachusetts, USA.

[71] A. Banchs and X. Perez, "Distributed Weighted Fair Queuing in 802.11 Wireless LAN", IEEE ICC 2001, Vol. 5, pp. 3121 -3127, New York City, New York, USA.

[72] S. Mangold, et, al, "IEEE802.11e Wireless LAN for Quality of Service", European Wireless, Vol. 1, Italy, February 2002.

[73] A. Grilo and M. Nunes, "Performance Evaluation of IEEE802.11e", IEEE PIMRC 2002, Vol. 1, pp. 511-517, Portugal.

[74] L. Romdhani, Q. Ni, and T. Turletti, "AEDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad –Hoc Networks", INRIA Research Report No. 4544, 2002.

[75] L. Romdhani, Q. Ni, and T. Turletti, "Adaptive EDCF: Enhanced Service Differentiation for IEEE802.11 Wireless Ad Hoc Networks", IEEE WCNC 2003, Vol. 2, pp. 1373 -1378, New Orleans, Louisiana, USA.

[76] G. Bianchi, "Performance Analysis of the IEEE802.11 Distributed Coordination Function", IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, March 2000, pp. 535-547.

[77] S.M. Ross, "Introduction to Probability Models", seventh edition, Academic press, c2000.

[78] E.P.C. Kao, "An Introduction to Stochastic Processes", Belmont, Calif., USA: Duxbury Press, c1997.

[79] G.S. Ahn, A.T. Campbell, A. Veres, and L.H. Sun, "Supporting Service Differentiation for Real-Time and Best Effort Traffic in Stateless Wireless Ad Hoc Networks (SWAN)", IEEE Trans. on mobile Computing, Vol 1, No. 3, July~Sept. 2002, pp. 192-207.