1-1-2006

Simple and effective QoS schemes supporting VoIP traffic in IEEE 802.11 network

Samson Chung Fai Chan
Ryerson University

SIMPLE AND EFFECTIVE QOS SCHEMES SUPPORTING VOIP TRAFFIC IN IEEE 802.11 NETWORK

by

Samson Chung Fai Chan

M.I.St., University of Toronto, 2003

B.A.Sc., University of Toronto, 2001

A thesis
presented to Ryerson University
in partial fulfilment of the
requirements for the degree of
Master of Applied Science
in the Program of
Computer Networks

Toronto, Ontario, Canada, 2006

I hereby declare that I am the sole author of this thesis or dissertation.

I authorize Ryerson University to lend this thesis or dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis or dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Simple and Effective QoS Schemes Supporting VoIP in IEEE 802.11 Network

Master of Applied Science, 2006, Computer Networks, Ryerson University

Abstract

This thesis presents several simple and effective QoS schemes aimed to support VoIP in the IEEE 802.11 network. The proposed schemes are called dynamic backoff adjustment (DBA), high priority DCF (HP-DCF), adaptive DCF (ADCF) and direct priority DCF (DP-DCF). All four schemes utilize the contention mechanism and are compatible with the IEEE 802.11 standard. The first three schemes use dynamic adjustment approach to improve the delay of high-priority traffic in infrastructure network and the last schemes is designed for use in ad-hoc network.

From the simulation results, the DBA scheme shows better bandwidth utilization over DCF and provides better differentiated service performance than EDCF. Whereas, the HP-DCF and AP schemes give better protection of the real-time traffic classes from the excessive best-effort traffic load in the infrastructure mode. Moreover, the DP-DCF also shows its ability in protecting the real-time traffic class in the ad-hoc environment.

Acknowledgement

This thesis is an individual project that was completed under the supervision of Professor Bobby Ma as a partial fulfilment of the Master of Applied Science degree for the program of Computer Networks at the Department of Electrical and Computer Engineering, Ryerson University.

I must express gratitude towards Prof. B. Ma, who aided in setting up this research thesis and provided insightful ideas and guidelines throughout. Without his continued support, I would not be able to complete this thesis.

I would also like to thank my dear friends: Gary Yip, Dennis Xu, and Yifan Peng for advising on the use of the OPNET software package, and express sincere gratitude to Jennifer Mak for her contribution on final editing. In addition, I want to show appreciation towards my parents for their countless support the whole time.

Abbreviation

AC	Access Category
ACK	Acknowledgment
AD	Absolute Delay Support
AIFS	Arbitrary IFS
AIFSN	Arbitrary IFS Number
AP	Access Point
AQ	Absolute QoS Support
BI	Backoff Interval
BS	Base Station
BSS	Basic Service Set
CBR	Constant Bit Rate
CPC	Centralized Polling Control
CSMA	Carrier Sense Multiple Access
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CW	Contention Window
DBA	Dynamic Backoff Adjustment
DCF	Distributed Coordination Function
DFS	Distributed Fair Scheduling
DIFS	*DCF IFS*
EDCF	Enhanced Distributed Channel Access
EDCF	Enhanced DCF
HCCA	HCF Controlled Channel Access
HCF	Hybrid Coordination Function
IBSS	Independent Basic Service Set
IFS	Interframe Space
LAN	Local Area Network
MAC	Medium Access Control
MS	Mobile Station
MSDU	MAC Service Data Unit
PCF	Point Coordination Function
PER	Packet Error Rate
PF	Persistence Factor
PIFS	PCF IFS
QoS	Quality of Service
RTS/CTS	Request to Send/Clear to Send
SIFS	Short Inter-frame Space
TX	Transmission
TXOP	Transmission Opportunity
VBR	Variable Bit Rate
VoIP	Voice over IP
WLAN	Wireless Local Area Network

Contents

List of Tables

List of Figures

1 Introduction

Over the last few years, Wireless LAN (WLAN) based on the IEEE 802.11 standard [1] has been widely deployed in enterprises, university campuses, hotspots, etc. The IEEE 802.11 standard serves well for the best–effort data traffic. With the rapid growth of multimedia applications, there is a need to support real-time traffic class such as Voice over IP (VoIP) and Video on demand. However, the requirements for Quality of Services (QoS) support such as guaranteed bandwidth and low delay jitter present a challenge to the wireless medium. Even though VoIP application can be implemented in 802.11 networks, the number of VoIP connections that can be supported is limited. The requirements for service differentiation, guaranteed latency, and better bandwidth utilization give birth to a new access mechanism. The 802.11 standard committees formed a task group TGe to address the QoS need over the wireless LANs (WLAN) and hence, IEEE 802.11e was proposed.

However, the 802.11e is still in the draft standard and the product of 802.11e still has years to come. In this thesis, we propose four effective yet simple schemes to provide better QoS support in the legacy 802.11 mechanism. All four schemes provide service differentiation among different traffic classes. The first three schemes, called Dynamic Backoff Adjustment (DBA), High Priority DCF (HP-DCF) and Adaptive DCF (ADCF), are designed for the infrastructure mode, while the last scheme, called Direct Priority DCF (DP-DCF), is designed for the ad-hoc wireless environment.

Our goal is to design simple yet efficient schemes to provide differentiated services support not found in the legacy 802.11 mechanism. Our schemes span over the infrastructure and ad-hoc environment for a wide-range of network access coverage. Since the two modes of operation are widely in use, the effectiveness of our proposed schemes will be closely connected to the real world application.

Here, we introduce the four schemes briefly. In the DBA scheme, the AP monitors the number of simultaneous VoIP connections in the network and assigns channel access parameters dynamically to the stations in the WLAN. Since the 802.11e standard uses static contention-window (CW) sizes regardless of the network condition, bandwidth would be wasted and the network resource would not be fully utilized. This argument provides an insight to our design to seek a dynamic assignment that would better utilize the available bandwidth and achieve more optimized network utilization. DBA dynamically assigns contention window sizes based on the VoIP traffic load. This approach yields a better VoIP delay performance. However, the scheme can not completely protect the VoIP traffic from the data (low-priority) traffic. Therefore, we propose another scheme, HP-DCF, to guarantee the prioritized access for the real-time traffic class. HP-DCF uses the non-overlapping window mechanism with fixed data window lower bound. This approach gives a higher degree of priority to the VoIP traffic at the expense of lower data traffic throughput. To improve the data traffic throughput of HP-DCF, we developed an adaptive approach called ADCF. ADCF uses adaptive data window lower bound to reduce the back-off delay, thus improving channel utilization, while still provides good service differentiation. Based on the success of ADCF, we developed the DP-DCF scheme to provide prioritized access to the VoIP traffic while better utilizing the remaining bandwidth allocated to the low-priority traffic class in the ad-hoc environment.

This thesis is structured as follows: First of all, a brief overview of the legacy IEEE 802.11 standard is given and the new IEEE 802.11e operation is described in detail. In Chapter 3, related works on improving real-time traffic over IEEE 802.11 WLAN are presented. In Chapter 4, QoS design issues are discussed and four new schemes, namely, DBA, HP-DCF, ADCF and DP-DCF are proposed. The four schemes are then studied through simulations in Chapter 5. The results of their performances are compared with those of the legacy IEEE 802.11 and the IEEE 802.11e standards. The final chapter concludes the thesis and points out the direction of future research.

2 Background

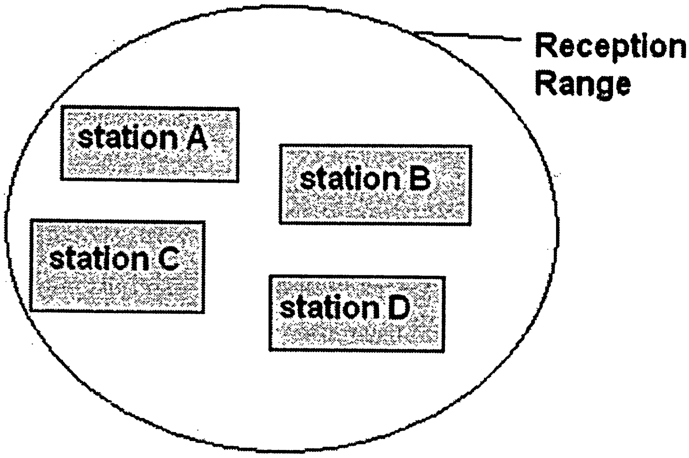
2.1 IEEE 802.11 standard

The IEEE 802.11 MAC layer provides reliable data service to the higher layer protocol and controls fair access to the shared wireless medium [13]. The idea of having a wireless MAC layer is to hide the unreliable nature of the wireless medium from upper layer protocols. Using the MAC layer allows a mobile station to roam throughout the WLAN freely and appears as stationary to the higher layer protocol above the MAC.

There are two modes of operation for the 802.11 standard: ad-hoc mode and infrastructure mode. The ad-hoc mode is a decentralized method in which stations communicate with each others in a peer-to-peer method. For the infrastructure mode, a centralized method is used in which station communication is coordinated through an access point (AP). Basic service set is a basic building block of the 802.11 wireless LAN [18]. It is of interest when we take into account of the WLAN topology as discussed in the following paragraph. Fundamentally, Basic Service Set (BSS) contains a group of number of wireless stations with or without AP.

2.1.1 Independent Basic Service Set (IBSS)

This topology is also called ad-hoc network. It consists of a number of wireless stations. Each station recognizes each other and is connected in a peer-to-peer fashion. Each station communicates directly with each other [18]. However, it requires that every station is within reception range of each other. Otherwise, it may not be able to communicate. Figure 1 shows the ad-hoc network topology.



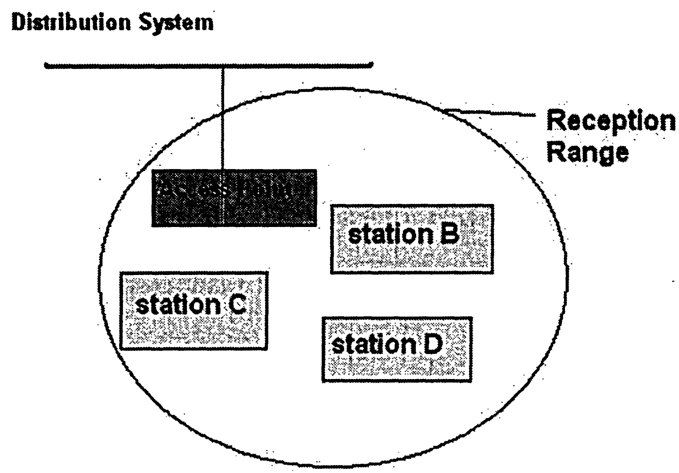
All stations need to be within range of each other

Figure 1: IBSS or Ad-hoc network

2.1.2 Infrastructure Basic Service Set

This topology consists of a number of wireless stations and an AP. The AP functions as a relay for the BSS. The sending station first transmits a frame to the AP, the AP then transfers the frame to the receiving station. All communications are relayed through the AP and stations no longer communicate directly [18]. The main purpose here is to double the reception range of the IBSS. As long as the wireless station is within reach of the AP, the distance between the wireless stations themselves has no restriction. Another advantage is that when the station is in power-saving mode, the AP may buffer frames for that particular station for later transmission [7]. However, the disadvantage is that it reduces the transmission capacity than in the case where the sender transmits frames directly to the receiver [7].

In addition, the AP is connected to a distribution system (DS). Figure 2 shows the Infrastructure BSS diagram.



Access Point acts as a relay to forward frames to/from DS and among stations

Figure 2: Infrastructure Basic Service Set

The distribution system itself may be a wired network or a special box that interconnects APs in another Basic Service Set. DS serves as the backbone of the wireless LAN for communication with another wired or wireless network. DS determines if traffic is relayed to a destination in the same BSS or forwarded to another AP through the distribution system. It can also determine if it can be forwarded to a wired network with destinations not in the extended service set (ESS). ESS is a set of infrastructure BSS interconnected by a wired network to extend the range of mobility to an arbitrary length. For example, two IBSS maybe connected through the DS to form an Extended Service Set. Figure 3 depicts the ESS diagram.

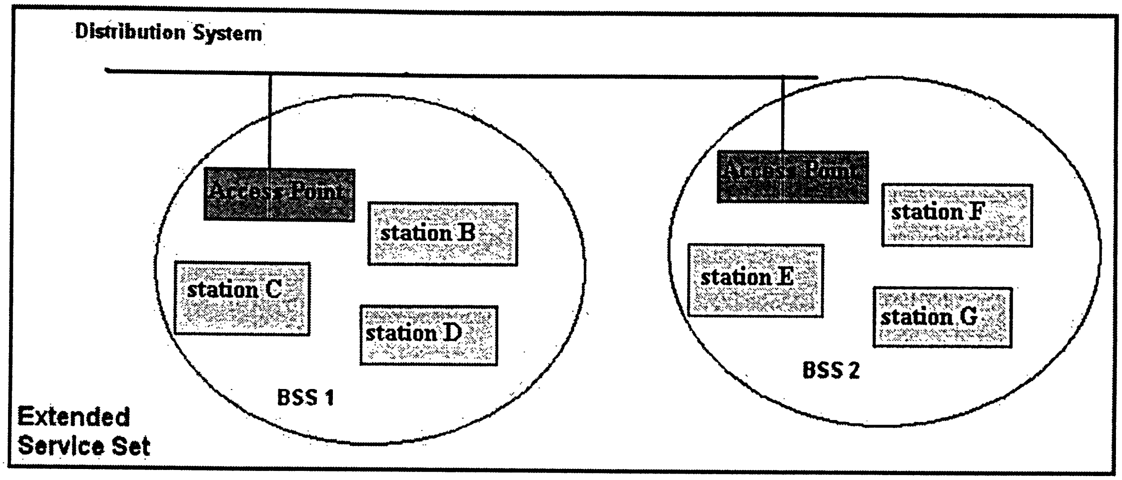


Figure 3: Extended Service Set (ESS)

To facilitate reliable data communication, IEEE 802.11 standard defines a frame exchange protocol. When a frame is received, the destination checks for the correct frame check sequence (32-bit CRC) and replies with an acknowledgment (ACK) if no error is found in the received frame. If ACK is not received or the frame check sequence is bad, re-transmission is needed. This ensures reliability of data transmission at the link layer even though the wireless medium is not as reliable as the wired medium. If the sender does not receive ACK from the destination in a timely manner, it considers the frame lost and retransmits the frame.

2.1.3 Interframe Spacing (IFS)

IEEE 802.11 relies on the concept of interframe space (IFS) to provide various priorities on channel access. Interframe space is measured in time unit. Different frame types have to wait for the durations of their respective IFS before attempting for channel access. The smaller IFS a frame uses, the higher the channel access priority it has. In IEEE 802.11 standard, several types of IFS are defined. SIFS has the shortest time among other IFSs. Frames using SIFS gain access of the channel with the highest priority over other frames that transmit using other IFSs [7]. An acknowledgement frame is an example that uses SIFS. Point Coordinate Function Interframe Space (PIFS) is the IFS used in Point Coordinate Function (PCF) mode. Stations in PCF mode can transmit frames during the contention free period after PIFS has elapsed [7]. DCF Interframe Space (DIFS) is used in Distributed Coordinate Function (DCF) mode. Stations operating in this mode are contention-based and can gain access to the wireless channel after a period longer than DIFS time [7]. Extend Interframe Spacing (EIFS) is applied only when a station attempts to retransmit a failed packet. The actual value of the interframe space depends on the physical mechanism used (e.g. DSSS or FHSS) and is not included in the diagram.

The 802.11 standard implements two channel access mechanism: distributed coordination function (DCF) and point coordination function (PCF). Details to both mechanisms are discussed below along with their limitations.

2.1.4 DCF – Distributed Coordination Function

DCF is the fundamental access method of the 802.11 MAC layer based on carrier sense multiple access with collision avoidance (CSMA/CA) protocol. CSMA/CA is similar to CSMA/CD used in 802.3. The difference is that in wireless communication, collision detection is not possible. A transmitting station cannot reliably detect collisions because the transmission signal is much stronger than the received signal. To detect collision, the cost of building such hardware transceiver is high. It consumes high power which is not economical for mobile devices with limited power. Therefore, 802.11 uses collision avoidance instead.

Carrier Sense Multiple Access makes use of both physical and virtual mechanisms. The mechanism of sensing a physical carrier in the channel depends on the medium and modulation used. Virtual carrier-sensing uses Network Allocation Vector (NAV). NAV is a timer that tells the time period the medium is reserved [7]. The station sets its NAV to reserve the time period it expects to use. Other stations count down NAV from the initial reserved value to 0. If the NAV is not zero, this tells that the medium it is busy. When NAV is zero, this means the medium is idle and the station may try to access the medium [7].

A station has to sense the medium before data transmission. Frame is transmitted after the channel has been idle for at least a DCF interframe space (DIFS) time. Figure 4 shows the details of the mechanism. If the medium is busy, access to the channel is deferred for a random backoff time measured in terms of time slots [7]. The random time slot is chosen from a range of 0 to CW_{size} -1, where CW is called Contention Window and CW_{size} is the size of the CW. The minimum and maximum of CW_{size} are denoted as CW_{min} and CW_{max}, respectively. The default value for CW_{min} is 16 and CW_{max} is 1024 slots.

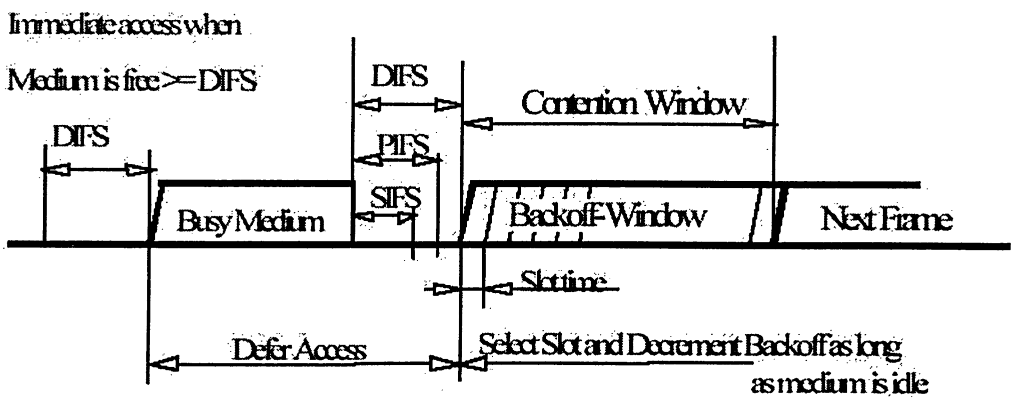


Figure 4: Backoff algorithm [1]

If the wireless medium is found to be busy during backoff, the countdown is paused and restarted until the medium is idle for a DIFS time again. When the medium is idled for at least a DIFS time again, the backoff timer decreases by one slot time and continues as long as the medium remains idle [7]. When the backoff reaches zero, the frame is sent. After a successful transmission, the CW_{size} is reset to CW_{min} and the backoff mechanism is used before sending the next frame. In case the backoff reaches zero and a collision is detected, a new backoff slot is chosen and the backoff procedure starts again. When a collision is detected (by means of not receiving acknowledgement after the data frame), CW is doubled according to $CW_i = 2^{k+i-1}$, where *i* is the number of transmission attempts (including the current one), and *k* is a constant defining the minimum contention windows, $CW_{min} = 2^k$ [12]. The CW_{size} is in the power of 2 [7]. The CW_{size} moves to the next greatest power of two whenever there is a retransmission. For instance, 1st retransmission increases the CW from 32 to 64, 2nd retransmission move the CW to 128 and so on [7]. The CW_{size} is bound by CW_{max} which is the maximum size that the contention window can grow and is physical layer-dependent. For Direct-Sequence Spread Spectrum (DSSS) physical layer, the maximum contention window size is 1024.

Thus, backoff algorithm is useful in avoiding collision and doubling CW size has the effect of reducing the likelihood of consecutive collision.

2.1.5 PCF – centrally controlled polling scheme

PCF operates only in infrastructure mode and is optional in the 802.11 standard. Notice that none of the 802.11 wireless cards and AP in the marketplace today implement this optional standard. However, the concept of PCF is important in QoS as it provides time-bound services. PCF mode is controlled by a point coordinator (PC) located inside an access point (AP).

In addition to having DCF contention period, PCF introduces a contention–free period (CFP), in which the PC polls each station in turn for frame transmission. The AP announces the start of a CFP period by sending delivery traffic indication message (DTIM) beacon frames [13]. The beacon frame contains synchronization and BSS information. (i.e. SSID, supported rates). Beacon frame format is listed in Table 1.

Table 1: Beacon Frame Body [1]

Order	Information	Notes
1	Timestamp	
2	Beacon interval	
3	Capability information	
4	SSID	
5	Supported rates	
6	FH Parameter Set	The FH Parameter Set information element is present within Beacon frames generated by STAs using frequency-hopping PHYs.
7	DS Parameter Set	The DS Parameter Set information element is present within Beacon frames generated by STAs using direct sequence PHYs.
8	CF Parameter Set	The CF Parameter Set information element is only present within Beacon frames generated by APs supporting a PCF.
9	IBSS Parameter Set	The IBSS Parameter Set information element is only present within Beacon frames generated by STAs in an IBSS.
10	TIM	The TIM information element is only present within Beacon frames generated by APs.

However, the sending of beacon frames can be delayed when the wireless medium is busy. In the beginning of the CFP, the AP gains control of the medium after sensing the medium is idle for PCF interframe space (PIFS) time at the target beacon transmission time (TBTT) [13]. TBTT is the time where AP should schedule a beacon as the next frame for transmission [1]. Time zero is defined as a TBTT where DTIM is carried in the beacon at the start of a CFP [1].

Since there is no contention in CFP, the AP schedules two-way transmission for each station by polling. Each station with CFP-enabled is polled by the AP. AP sends the

CF-Poll frame to one of the poll-able station. If the AP has a need to sent data frame need to this station, the frame is attached with the CF-Poll (DATA+CF-Poll frame). Upon receiving a poll, the station transmits its data with ACK (DATA+CF-ACK frame) or response with an ACK (CF-ACK) to indicate that nothing is to be sent. SIFS interval is used throughout the frame exchange. When one station finished its frame exchange sequence, the AP sends another CF-Poll to the next station on its poll list. A poll list is maintained in order to poll the stations but such process is implementation dependent. The actual CFP duration is announced in the beacon, and the NAV are updated accordingly for every station. The polling continues until AP has finished polling all stations in the poll list or the CFP period has expired. Then, the AP broadcasts the CF-End frame to indicate the end of CFP and all stations reset their NAVs to zero. Once the CFP is finished, CP is followed and the stations contend for access using DCF until the next DTIM beacon. The sum of the two periods CFP and CP is called a superframe. Figure 5 shows the PCF operating procedure.

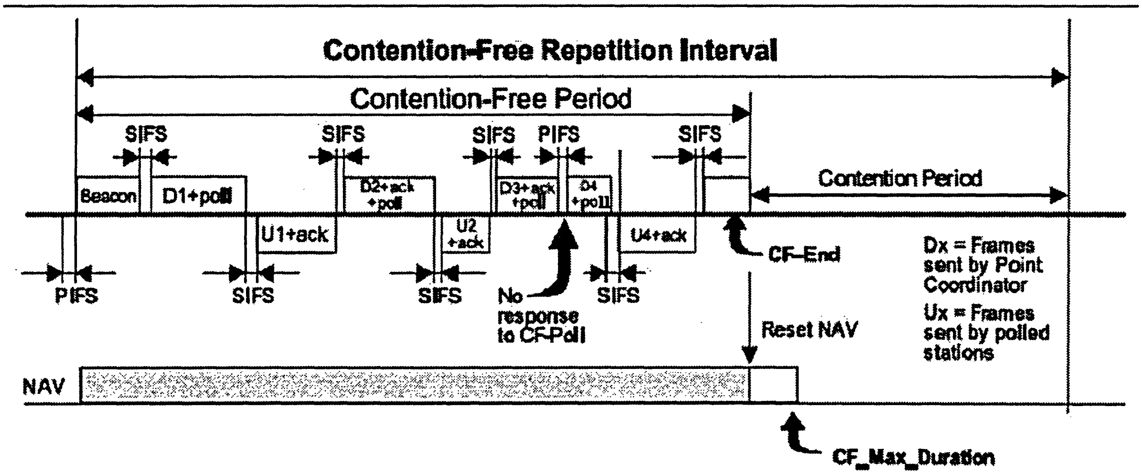


Figure 5: PCF operation [7]

There is a known problem in the implementation of CFP called Foreshortened CFP. When a station begins to send a frame at the end of the CP period, the station lengthens its transmission time carried over to the next CFP period. As a result, the actual start time of the next CFP period is delayed and has to be shortened to meet the start time of the next CP period. It is a well known issue for the 802.11 standard, and there is no solution to this problem. Many studies have addressed the problem but the standard does not include a fix. Figure 6 shows the foreshortened CFP situation.

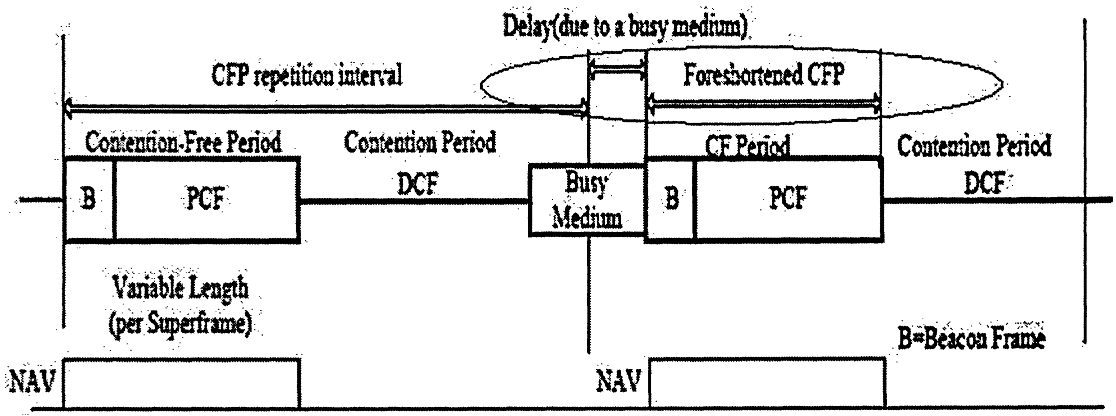


Figure 6: Foreshortened CFP [7]

2.2 The IEEE 802.11e QoS Draft

The IEEE 802.11 task group (TGe) [2] proposed a new standard, namely 802.11e, to try to provide the QoS support that is lacking in the original IEEE 802.11 standard. Two new methods were introduced: Enhanced Distribution Coordination Access (EDCF) and HCF Controlled Channel Access (HCCA). EDCF provides differentiated-services type of QoS support while HCCA offers integrated-services type of QoS support. Principally, EDCF adds different prioritizations for different traffic classes in the DCF contention-based mechanism to provide differentiated services. HCCA, on the other hand, introduces several features of QoS guarantee to the PCF polling mechanism. Polling is controlled by the hybrid co-ordinator (HC), similar to PC in legacy PCF. HC is co-located in the QoS enhanced access point (QAP).

The 802.11e defines a new superset called QoS facility. The QoS facility includes QoS enhanced station (QSTA) and access point (QAP). The QSTA is the station that supports the 802.11e QoS features; while QAP is the AP that is 802.11e QoS capable [14].

The QoS facility brings in the new coordination function named hybrid coordination function (HCF) for use in QoS enhanced basic service set (QBSS). HCF makes use of both EDCF and HCCA, which is similar to the original 802.11 access method with PCF and DCF together. HCF brings in a new feature called transmission opportunity (TXOP). A TXOP defines the time duration that a QSTA takes to transmit a number of frames per poll. This is different from the PCF as only one frame is allowed to be transmitted in each station per poll. The TXOP contains a start time and a maximum duration for that station to transmit. The concept of TXOP works in both EDCF and HCCA modes, and is named EDCF-TXOP and HCCA (polled) TXOP, respectively [14]. The relationship between HCF and PCF is shown in Figure 7 below:

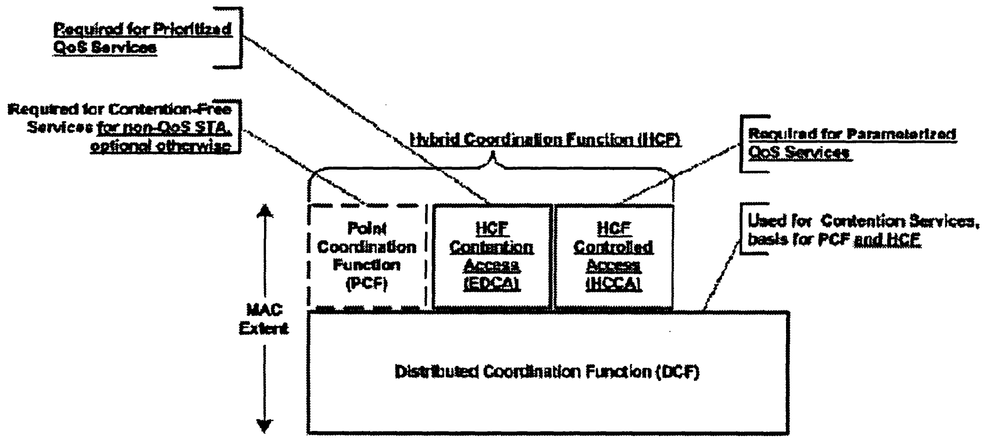


Figure 7: MAC architecture [2]

In addition, the 802.11e draft brings in many new features to provide a better QoS support. One concept worth noting is that the positive MAC-level Acknowledgment

(ACK) is optional in 802.11e [14]. It means that MAC no longer needs to reply with an ACK when the frame is received. This improves the overall MAC efficiency as VoIP/video data has strict lifetime. Block acknowledgement is another feature that allows a block of ACKs to be sent in one ACK frame. This greatly reduces the overhead due to immediate response of ACK [2]. Furthermore, the 802.11e introduces direct link protocol, which allows two stations belonging to the same AP to transmit data to each other directly without relaying traffic through the AP [2]. Likewise, 802.11e allows for piggyback, which means data can be sent piggybacked through polls and ACKs to reduce overhead [2].

Additionally, the 802.11e introduces traffic specification (TSPEC) management to match the priority assignment in higher layer QoS protocol with the MAC layer. TSPEC is discussed next.

2.2.1 Traffic Specifications (TSPEC)

The traffic Specification (TSPEC) is the traffic stream management defined in 802.11e standard, which provides management link between higher layer QoS protocol to 802.11e channel access functions and admission control. TSPEC describes the traffic characteristics and QoS requirements of a traffic stream (TS) [2]. Station requiring data to be sent must first inform the AP of its traffic class and associated TSPEC, i.e. the resource required (data rate, size of data, delay, duration, minimum PHY rate, etc.) so that AP can calculate the resource needed for that TS (i.e. bandwidth, delay, etc.) and allocate TXOP to that TS correspondingly [14]. The traffic specification element format is listed in Figure 8.

Element ID (13)	Length (55)	TS Info	Nominal MSDU Size	Maximum MSDU Size	Minimum Service Interval	Maximum Service Interval	Inactivity Interval	Suspension Interval
Octets: 1	1	3	2	2	4	4	4	4

Service Start Time	Minimum Data Rate	Mean Data Rate	Peak Data Rate	Burst Size	Delay Bound	Minimum PHY Rate	Surplus Bandwidth Allowance	Medium Time
Octets: 4	4	4	4	4	4	4	2	2

Figure 8: TSPEC - traffic specification element format [2]

Station requests for TXOP before data transmission; only a station given TXOP is allowed to transmit. If there is insufficient bandwidth, station request is deferred until enough becomes available. TSPEC is very important in guaranteeing QoS establishment and traffic congestion avoidance. TSPEC is used in both EDCF and HCCA for the purpose of admission control.

Upon successful negotiation of TSEPC between QSTA and HC, a traffic stream is created and identified by its TSID and Directions and QSTA address for use in HC. QAP then uses these parameters to filter the MSDU belonging to this TS and to deliver according to the QoS parameters already set up for the TS. The creation of TS is initiated by the non-AP QSTA at all times regardless of the direction. QSTA can transmit QoS Data frame only when TS is active [2]. TS become inactive when either QSTA/HC begins the TS deletion process or TS inactivity timer expires. When TS becomes inactive, all resources allocated for that TS are released.

In the 802.11e standard, non-AP QSTA may simultaneously maintain up to 8 traffic streams for uplink to QAP and 8 traffic streams for downlink from QAP [2]. The actual numbers of TS it support are limited by vendor implementation restrictions. As well, a HC may support 8 uplink and 8 downlink traffic streams simultaneously dependent on vendor implementation restrictions.

2.2.2 Enhanced Distribution Coordination Function (EDCF)

EDCF is an extension to DCF. EDCF includes prioritized QoS. The EDCF mechanism classifies 8 different traffic classes. Different priorities are assigned to different classes by means of varying the minimum contention window (CW_{min}) and interframe space used for data transmission. Traffic with a smaller CW results in shorter average backoff interval, thus gains a somewhat higher priority of channel access over stations with larger average CW. Further differentiation between stations with the same contention window is achieved by varying IFSs used by different traffic classes. The varying of IFS is called Arbitrary Interframe Space (AIFS). With a larger AIFS, the station will wait longer before trying to access the channel, and hence have a lower priority than other traffic with shorter AIFS. Each traffic class has its own AIFS value and CW_{min}. The value of CW could be set to zero to ensure faster access to the medium.

To support priority in IEEE 802.11D, EDCF adjusts the channel access mechanism into 8 priorities, which are further mapped into 4 Access Categories (ACs). The term AC is the same as Traffic Class (TC) in EDCF as appeared in the earlier version of IEEE draft. Table 2 shows the mapping between 802.1D and the EDCF. Figure 9 shows the EDCF using four queues for the four ACs, in contrast to DCF where only one queue is used for all traffic.

Table 2: Priority to EDCF Access Category Mapping

Priority (as in 802.1D)	Access Category (AC)	Designation
1	0	Best Effort
2	0	Best Effort
0	0	Best Effort
3	1	Video Probe
4	2	Video
5	2	Video
6	3	Voice
7	3	Voice

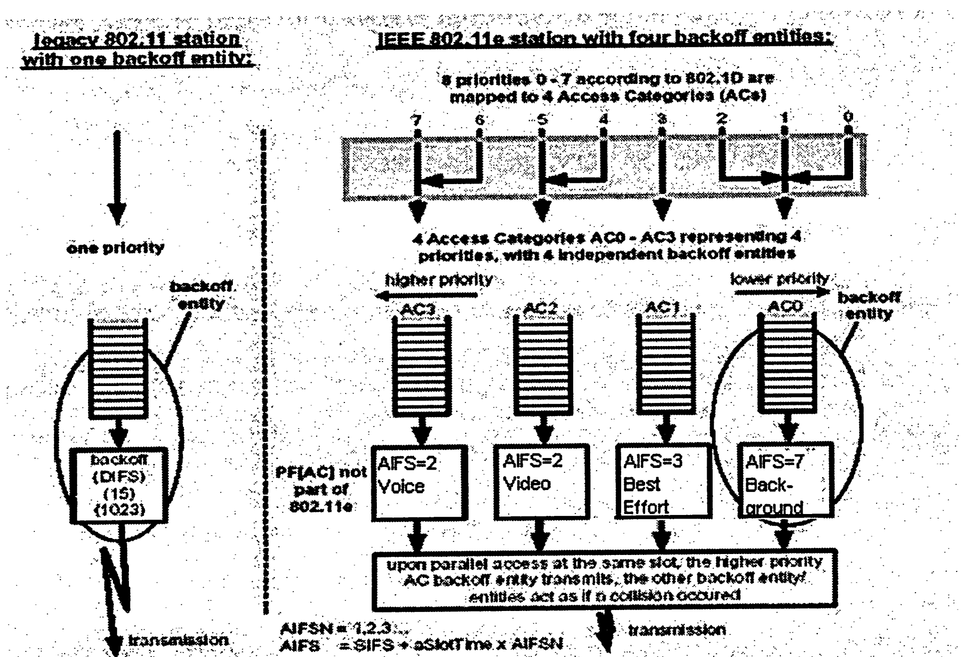


Figure 9: Access Category in EDCF compared with legacy 802.11 DCF [2]

In reality, there are only 4 different sets of parameters available for EDCF. For Voice usage, AC3 is used and has the highest priority among other ACs. For Video traffic, AC2 is used. For Best Effort traffic, AC1 is used. AC0 is the lowest priority and is used for background traffic. Each AC defines its own AIFS, CW_{min} and CW_{max} values. The default parameters are defined in 802.11e draft 13 and are listed in Table 3. Some of the terms in Table 3 are defined as follows: AIFSN[AC] is the AIFS time allocated according to that particular traffic class N; aSlotTime is defined as a single slot time duration according to the PHY layer characteristics; aSIFSTime is defined as the SIFS time duration according to the PHY layer characteristics.

Table 3: EDCF default parameter set for 802.11e Draft 13 [2]

	AC	CWmin	CWmax	AIFSN	TXOP Limit		
					For PHYs defined in clauses 15 and 18	For PHYs defined in clauses 17 and 19	Other PHYs
Background traffic	AC_BK	aCWmin	aCWmax	7	0	0	0
Best Effort traffic	AC_BE	aCWmin	aCWmax	3	0	0	0
Video traffic	AC_VI	(aCWmin+1)/2 - 1	aCWmin	2	6.016ms	3.008ms	0
Voice traffic	AC_VO	(aCWmin+1)/4 - 1	(aCWmin+1)/2 - 1	2	3.264ms	1.504ms	0

In particular, $AIFS[AC] = AIFSN[AC] \times aSlotTime + aSIFSTime$, where AIFSN[AC] is an integer greater than or equal to 2 for non-AP QSTA stations and greater than or equal to 1 for APs as defined in 802.11e draft13 [2]. In each beacon frame, the AP broadcasts the EDCF parameters set for the chosen AC. The smaller the AIFSN[AC] and CW_{min}, the greater the probability of a station winning the contention. Separate queues for each AC are kept in each station. Each queue acts as an individual backoff entity contenings for transmission [19]. When the backoff timer of the two queues completely expires at the same time, the higher AC backoff entity wins and transmits. The other backoff entity/entities responds as collision occurs as illustrated in Figure 10.

When a collision is detected, the contention window selection follows the legacy DCF binary exponential backoff procedure. In earlier draft version, the CW was expanded after a failed transmission to

$$newCW[TC] = [(oldCW[TC]) + 1) \times PF] -1 \qquad (1)$$

In equation 1, the persistence factor (PF) is another parameter for prioritization in Draft 2.0a. From draft D3.0 and so on, PF is no longer used. Regardless of different AC, PF is set to 2, which leads to a binary exponential backoff as follow:

$$newCW[TC] = [(oldCW[TC]) +1) \times 2] -1 \qquad (2)$$

If the newCW[TC] exceeds the CW_{max}[TC], the newCW[TC] remains unchanged for the remainder of any retries. If the retry limit is reached, the newCW[TC] shall be reset to CW_{min} of that TC [2].

In addition to AIFS and minimum size of CW feature, prioritization is also controlled by the TXOPLimit[TC]. Each station under EDCF contends for a transmission opportunity (TXOP). TXOP is a time-bound period that a station is allowed to initiate transmissions, defined by the start time and a maximum duration [14]. The TXOPLimit[TC] specifies the maximum duration of one TXOP as illustrated earlier in Table 3. QSTA should ensure that its transmission does not exceed the TXOP limit [2]. When the TXOP is granted in contention based access, it is called EDCF-TXOP. Likewise, HCCA (polled) TXOP is called when the TXOP is granted under HCCA. HCCA will be discussed next. Figure 10 depicts IFS relationship between SIFS, PIFS and AIFS for different AC.

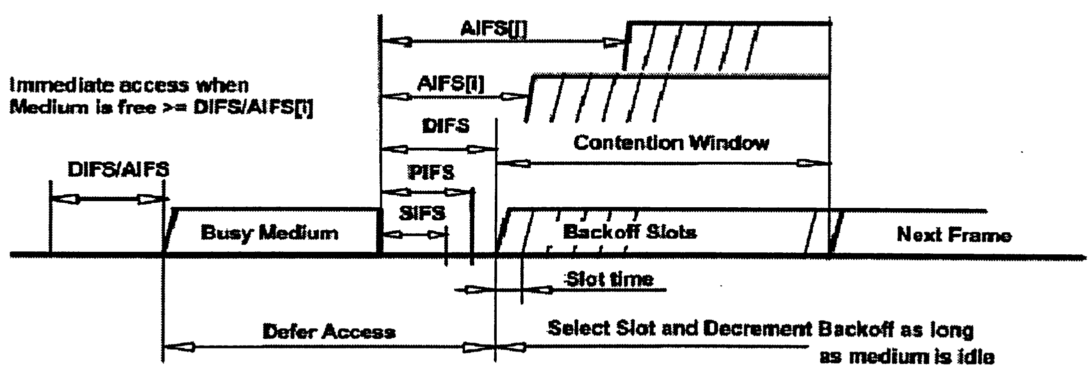


Figure 10: IFS relationship and EDCF channel access [2]

From the above diagram, it can be seen that the earliest channel access is DIFS. The minimum AIFS is equal to DIFS. Station gains immediate access to the wireless medium as long as the medium is free for at least a DIFS/AIFS[i], where i is the AC. Random backoff counter selects from interval [0, CW_{min}] according to the respective TC. As an example, the EDCF timing relationship is shown in Figure 11.

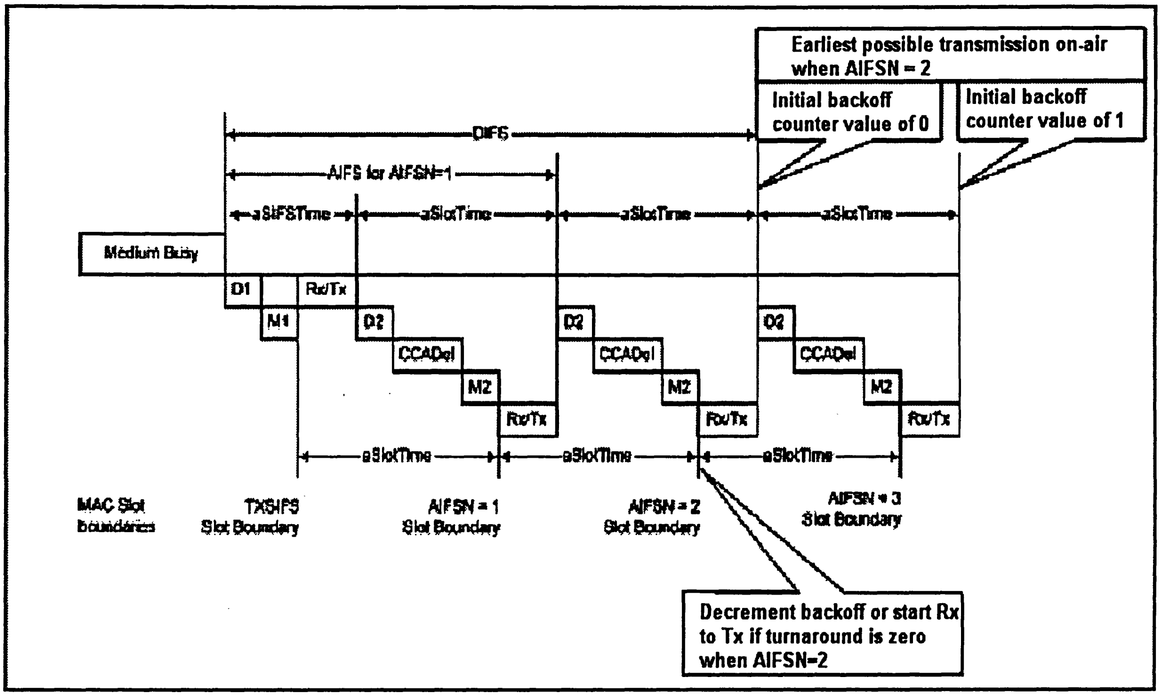


Figure 11: EDCF timing example [2]

This example demonstrates the relationship among different IFSs and slot times when busy medium is experienced (busy medium not caused by frame error) as illustrated in 802.11e draft13 [2]. For AIFSN=2, the EDCF decrease the backoff counter for the first time at interval 2x aSlotTime after the end of busy medium condition. For the example above, the initial backoff counter value is 1 when the medium became idle. Transmission starts after the backoff counter has elapsed, so the total time it waits after the medium busy ends is a*SIFSTime* + *2x aSlotTime* + aSlotTime (due to backoff elapsed) [2].

2.2.3 HCCA (HCF Controlled Channel Access)

HCCA or HCF in short, is similar to the legacy PCF polling mechanism. It is designed to support parameterized QoS. HC is a type of centralized coordinator used to control channel access; it has a higher medium access priority than non-AP stations [2]. This solves the beacon delay problem identified in PCF mode as mentioned earlier in the PCF section. To solve the foreshortening problem in PCF mode of operation, a QSTA is not allowed to transmit a frame if the transmission will not be finished before the next TBTT [21].

HC differs from the Point Coordinator (PC) used in PCF in a number of ways. One of the major differences is that it allows HCCA to start the controlled channel access at anytime during CFP and CP intervals. Another difference is that HC is capable of allocating TXOP to non-AP QSTA. HC grants a polled TXOP to non-AP QSTA with the duration specified in a QoS CF-Poll frame [2]. QSTA can transmit multiple frames within the given polled-TXOPs as long as the TXOP duration limit is not exceeded. To prevent contention during CFP due to legacy station not understanding the CF-Poll frame, all stations are required to obey the NAV rules set by the HC [2]. Figure 12 shows the superframe of HCCA and its frame exchange sequence.

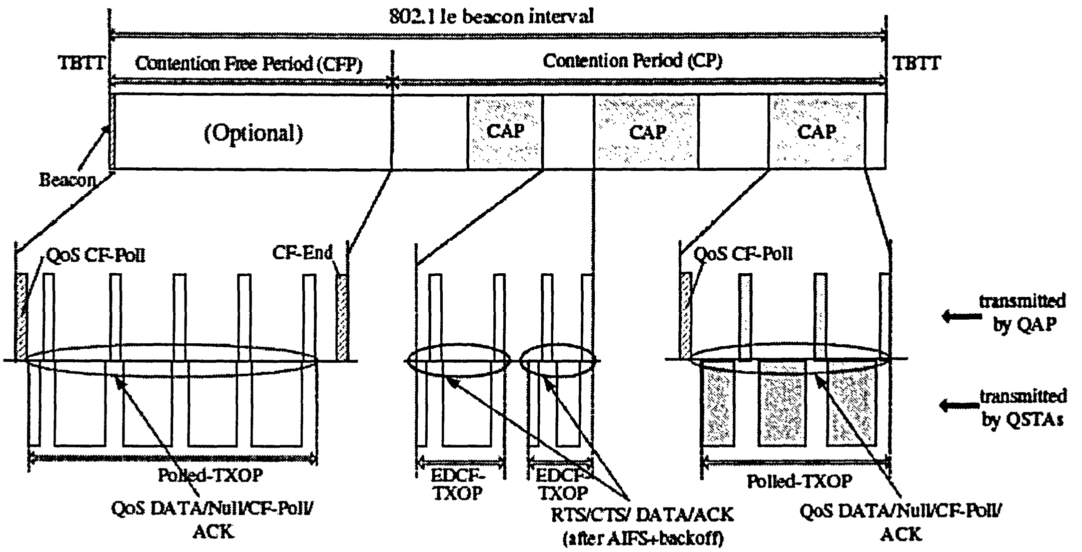


Figure 12: HCCA superframe and frame exchange sequence [21]

Like PCF, HC initiates CFP for frame delivery as long as CFP is ended with a CF-End frame. During CFP, HC provides a CF-Poll to a particular station specifying the start time and maximum duration, namely TXOP. A station is given the right to send packets when it has TXOP instance. Within a TXOP, multiple transmissions can be done as long as the time allocated for this TXOP has not expired. HC controls the time interval during which a station could transmit [5].

Other stations could not access the medium as they do not have TXOP. When a station receives a CF-Poll from the HC, they accept the TXOP and transmit accordingly within a SIFS time. If the station has nothing to send, HC will take over the medium after a PIFS interval and issue another CF-POLL [11]. A beacon frame or CF-End Frame indicates the end of the CFP [11].

In addition, HCF defines a special period called Controlled Access Period (CAP). A CAP is an interval in which multiple frames are allowed to transmit using controlled access mechanism within a contention period. During the CP, CAP starts at anytime after the medium remains idle for at least a PIFS interval [5]. This guarantees the HC granting access priority over EDCF transmission that uses DIFS and AIFS. To initiate the HCCA–TXOP, HC starts CAP to allocate polled–TXOP (CAP) to different QSTA by sending downlink QoS-frames and QoS CF-Poll frames [2]. If CAP is not used during CP, the EDCF rules are used. The polled TXOP (CAP) guarantees QoS requirements to be met by initiating CAP at any time. It is recommended in the standard that CAP should be used in Contention Period instead of CFP to reduce implementation complexity [2]. CAP is shown in Figure 13 below.

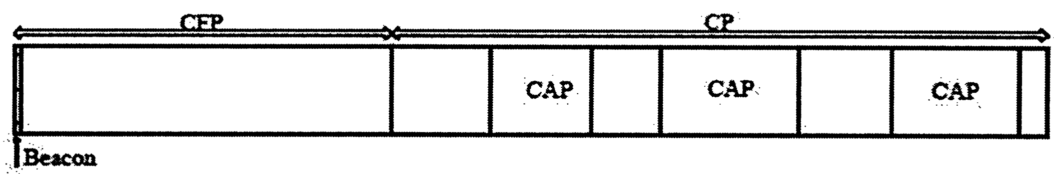


Figure 13: CAP in HCCA

3 Related Work

3.1 Analysis of the IEEE 802.11 and IEEE 802.11e standards

3.1.1 Limitation of DCF and PCF in supporting QoS:

There are many limitations in the legacy IEEE 802.11 DCF and IEEE 802.11 PCF to support QoS. Originally, the design of DCF scheme does not have QoS consideration in mind. Traffic is served in best effort using one common queue for packet buffering and transmission regardless of the traffic type. There is no differentiation among different traffic classes within a station. Even though a real-time packet wins internal competition in gaining access to the wireless channel, it has to compete with all other wireless stations that request to use the wireless medium globally. All stations contend for the medium access in the same priority, that is, real-time traffic from one station has to compete against traffic from other stations trying to access the channel. Real-time traffic needs to win both internal and external competitions in order to send a packet. This introduces unpredictable delay and jitter for delay sensitive real-time packets.

Furthermore, the CSMA/CA algorithm creates stress to AP downlink traffic flow. In infrastructure BSS, many stations transmit uplink to the AP while AP transmits downlink traffic to all stations. This one-to-many traffic pattern demands a higher bandwidth assigned to the AP. However, the IEEE 802.11 contention scheme does not support such assignment.

In contrast, PCF can provide a certain degree of support to real-time traffic. However, each transmission must be routed through an AP, which is not suitable for ad-hoc type network. PCF defines both CFP and CP in fixed positions within a superframe, next packet pending in the queue at the end of CFP must wait for the next CFP interval to transmit if that packet is not able to send in the subsequent contention period. This adds additional delay to time-sensitive packets.

Moreover, the beacon frame for PCF is transmitted if the medium has been idle after PIFS interval. When the medium is busy, the beacon will be delayed. This unpredictable beacon delay will shorten the actual start time of CFP and hence shorten the CFP. A shortened CFP means some stations may not be able to transmit its frame within a given round of CFP. During CFP, transmission time of the polled station is unknown. If the frame is fragmented, the transmission duration depends on the size of the fragmented packet. The schedule to poll the next station for the rest of CFP will be changed and is not under the control of PC. Besides, there is no management interface to integrate the QoS requirement from the higher layer (e.g. Diffserv, IntServ) to control the PCF operation.

3.1.2 Inadequacy in supporting QoS in EDCF and HCCA

Thus far, the limitation of the original IEEE 802.11 has been identified. The new IEEE 802.11e aims to resolve these QoS limitations. Yet there are still some drawbacks associated to the new standard in supporting QoS. Under EDCF scheme, it provides differentiated services and addresses the lack of priorities in original DCF scheme. However, the priorities assigned to different classes are not absolute. That is, there is still a chance that lower priority traffic from one station can access the channel sooner than the higher priority traffic from the other station.

HCCA provides QoS guarantee because of the flexible distribution of CAP within a superframe and admission control to protect traffic streams. The study shows that HCF gives lower transmission delay with an increase in superframe size and low burst traffic [5]. However, the HCCA scheme is relatively more complex than the other schemes, and the use of CAP is not mandatory in the standard. The possibility of adopting this option to the WLAN product is still under investigation.

3.2 Performance evaluation of IEEE 802.11e

3.2.1 EDCF and HCCA performance evaluation

In [17], D. He et al simulated the 802.11e EDCF and evaluated over various configurations and traffic patterns. Four Access Categories are assigned and the performance of EDCF is assessed under various traffic loads. The results found that MAC delay differentiation is achieved. The differentiation is more obvious in heavy network load scenario. Low priority traffic is starved under heavy load and high priority traffic seizes more bandwidth as compared to original 802.11 DCF scheme. It is found that the AP downlink has worse performance than the AP uplink. The cause is that downlink traffic with same AC shares one buffer in AP and AP uses the same channel access mechanism as other stations. Under heavy load conditions, high priority traffic may suffer a longer delay, thus not meeting the QoS requirement. In the concluding remarks, the authors suggested some sort of admission control and scheduling mechanism to provide guarantee service to real time traffic and to improve channel access fairness to data traffic.

In [24], S. Choi et al. evaluated the performance of EDCF and the optional contention free burst (CFB) feature against the original 802.11 DCF scheme. Contention free burst permits multiple frame exchange within a single transmission opportunity (TXOP). Simulation results found that EDCF achieved service differentiation among different classes of traffic and CFB improves the global system performance with the expense of increasing delay to other types of traffic, especially real-time traffic. From their scenarios, CFB extended transmission time for video traffic and attained higher throughput, however, it introduced more delay to voice traffic. The simulation concluded that a careful selection of the EDCF TXOP limit is needed to keep the delay of other traffic type within an acceptable range.

In another paper [25], A. Salhotra et al. evaluated the CFB and focused on analysing transmission failures during the CFB. Their analysis showed that hidden nodes or fraudulent channel conditions could possibly cause a transmission failure. The authors proposed a recovery mechanism to recover corrupted frame due to a transmission failure. The scheme proposed is called modified CFB (mCFB); it works under the two fundamental conditions:

1. If ACK is not received within Recovery Timeout (RTO) interval after the first burst frame is sent, the transmitter goes into backoff and contends for the channel again.
2. Transmitter holds on to the medium if any subsequent frames are not properly ACK. It retransmits the failed frame after the RTO expires.

Additionally, the RTO should satisfy the inequality:

$$SIFS \leq RTO - dur\ (DATA) < SIFS + 2 \times SlotTime \qquad (3)$$

The inequality makes certain that no other stations will gain access to the medium before the retransmission of the failed packet. The results from the simulation found that the proposed scheme favours high throughput applications like HDTV or TV video rather than low throughput applications like VoIP or video phone. Overall, the throughput achieved using MCFB is higher and the delay jitter is lower compared to the normal CFB. The authors recommended using mCFB mode for unidirectional downstream applications while using normal CFB for stations exchanging bidirectional traffic.

From the paper of D. Chen et al. [28] , the EDCF and HCF schemes with HDTV, VoIP, background traffic and best effort data were evaluated. The simulation results found that introducing best effort (BE) and background (BK) traffic has little effect on HDTV media access delay as HDTV traffic has higher priority over BE/BK traffic. While adding more VoIP call to the scenario, HDTV throughput dropped significantly and the BE/BK traffic could not obtain much access due to exceed of network capacity. The authors found that heavily loaded traffic streams are susceptible to network load. Hence, they proposed placing the HDTV and VoIP traffic under HCF polling and using EDCF for the other traffic. The results found that HCF improves the performance for traffic in EDCF, thus it is much better to meet the QoS requirements for heavy loaded real-time traffic. Placing heavy loaded HDTV real-time traffic under HCF polling mode is more desirable with the mix of data traffic.

3.2.2 Dynamic tuning of contention windows for traffic differentiation

Dynamic tuning of contention windows and interframe spacing has been gaining attention recently. It is known that CW and IFS dictate the time interval in accessing the wireless medium, and changing these parameters intelligently would optimize the wireless medium access. Instead of setting a range of CW statically as in the IEEE 802.11 and 802.11e standard, the CW may be assigned dynamically. The benefit of dynamic tuning is to scale the CW in achieving better utilization of the WLAN according to various traffic loads. There are two basic approaches in dynamic tuning: active online measurement based and non measurement-based. Active online measurement based is to constantly acquire the current traffic load condition and making subsequent adjustments. This measurement-based approach requires sending frames to all mobile stations explicitly to inform the changes in the parameters (e.g. CW_{min}, CW_{max}) or to notify the duration for stations' own transmission. The amount of control traffic being sent consumes a portion of transmission time. More frequent measurements require more control traffic sent, hence consume more bandwidth. Non measurement-based approach does not require the sending of control traffic. It frees up the bandwidth needed for control traffic for the use of data traffic.

In [30 - 37], various methods of dynamic tuning of contention windows are published. Fine tuning of CW is proven effective in achieving higher throughput and lower delay. However, some of these papers assume constant bit rate traffic source [36][37] or constant packet size [32]; some algorithms focused on heavy contention condition [33][32]. Many of the proposed tuning algorithms require accurate online measurement [32][34][36][37]. For example, utilizing statistical models to count for long term collision rate [34], packet drop rate [37] in order to set the CW size. Some papers also take packet length into consideration for their proposed algorithms [31][33], while other authors [32] argue that it is not necessary. Theoretically, the tuning of the backoff algorithm could be optimum if a station knows exactly the network status.

However, it is unrealistic in real applications. It is only possible to estimate the network load configuration (i.e. number of active station and packet length) and if the estimation is erroneous, the algorithm will perform badly.

Non-measurement based approaches tries to eliminate the risk posted in erroneous estimation. Instead of actively monitoring and measuring the network load condition, it uses an indirect way to tune the CW adaptively. The basic idea is if the previous transmission is successful, the CW for the next transmission would be smaller, and this process continues until the transmission is unsuccessful. Upon the detection of an unsuccessful transmission, the CW will increase accordingly. Song et al. deployed exponential increase and exponential decrease to adjust contention windows [38], while Nafaa et al. applied the linear-increase/linear-decrease model in their backoff algorithm [36].

In [22], Q.Pang et al. proposed a comparable adaptive backoff scheme called MIMLD. The mechanism used is similar to TCP congestion windows approach. After a successful transmission, CW will decrease by one. The decrement is stopped once it reaches CW_{min}. This result in shorter average waiting time spent in CW and hence reduces the overall delay. The algorithm defines a CW_{basic} as the threshold for contention. Once a collision is detected, the CW is increased to 2 x CW_{basic} if CW < CW_{basic}. If CW > CW_{basic}, it will double the CW as in the original 802.11 CW algorithm. The algorithm avoids oscillation by having a linear decrease of CW instead of multiplicative decrease. The authors evaluated its scheme for a single-hop wireless network, and found that it achieved higher throughput than the original algorithm. Although the non measurement-based schemes could achieve better performance, the scheme is not proactive. As the number of mobile stations associating/disassociating the WLAN abruptly (probable case in airport/shopping centre hotspot), this scheme may not perform as well.

3.2.3 Modifying WLAN MAC protocol to support real-time traffic

H.Liu et al. proposed a modified MAC protocol that better supports the voice traffic in the 802.11 WLAN [23]. The scheme modifies the power saving (PS) mode in IEEE 802.11 to schedule time slot for voice traffic transmission. PS mode allows AP to buffer a number of frames designated to a mobile station that is in sleep mode. Based on the PS mode traffic indication map (TIM) used for buffer frame in the original 802.11 standard, the authors defined an extension to TIM called real-time traffic indication (RTIM) to use for voice traffic bandwidth allocation. A table is maintained in the AP to include the state and Association ID (AID) of the station to keep track of the reservation. The RTIM frame format is an additional uplink time schedule field, which includes the time period between two scheduled transmissions of voice users. Voice stations transmit their frames during the scheduled time period according to the order of AID number. Priority over data traffic is achieved as voice traffic is transmitted following a beacon. RTIM interval is the multiples of TIM interval and is broadcasted to all station through Beacon frame. When a voice station has uplink traffic to send, it informs the AP by including the AIDS in the next RTIM. The AP assigns a specific time for the station to transmit at the following RTIM. It uses a mini-contention window in case the last frame transmission failed. The simulation result showed that data traffic delay increases as the number of voice users increases. The number of voice users has to be limited to keep the delay of data traffic reasonable.

3.3 Summary

Many research works have been done recently on improving the performance of the Wireless Local Area Network WLAN. The original IEEE 802.11 standard has shown deficiency in transmitting real–time traffic. The IEEE 802.11e is aimed to address this deficiency. However, there is room for improvement to the IEEE 802.11e standard as pointed out in several studies highlighted in this chapter. The primary interest is to focus on how well the WLAN can support real-time traffic coexistence with the non-real-time data traffic. The nature of voice transmission requires strict delay and jitter bound to provide acceptable quality of services. It creates quite a number of challenges to provide reliable transmission and satisfy the real-time requirements due to the unreliable nature of the WLAN. It is observed that tuning of IFS and contention window would help in service differentiation; adjusting IFS and contention window dynamically would also improve wireless medium access utilization. Whereas applying admission control would safeguard existing traffic class from new traffic stream interference. Having the literature reviews into consideration, this thesis proposes several simple yet effective mechanisms in support of VoIP traffic.

4 Proposed Scheme

4.1 General description of the proposed schemes

The main goal of this thesis is to design simple but effective schemes to support both data and VoIP traffic. To support real-time traffic over wireless network, service differentiation among different traffic classes needs to be addressed. EDCF standard has shown that service differentiation can be achieved by assigning different IFSs and CWs to different traffic classes. Our proposed schemes concentrate on the control of CWs in achieving efficient service differentiation. Essentially, we note that CW should be adaptive to the traffic load. In light traffic conditions, CW should be small to expedite the transmission. On the other hand, for heavy traffic conditions, CW should be large to reduce the probability of collision.

Four schemes are proposed to address the QoS issues. The first scheme is called Dynamic Backoff Adjustment (DBA). The DBA scheme employs a dynamic assignment of CW according to the number of active VoIP connections. By making the CWs adaptive to the VoIP traffic load, we find that the scheme provides better traffic differentiation than EDCF.

The second scheme is called High Priority DCF (HP-DCF), which is based on the DBA scheme with one difference. The difference is that the CWs of data and VoIP traffic classes are non-overlapping. More specifically, the lower bound of the data CW is fixed to the largest voice CW size. Doing so will give the highest degree of priority to VoIP traffic. The drawback to this scheme is that the data traffic may suffer an excessive delay.

The third scheme proposed is called Adaptive DCF (ADCF). ADCF also uses the concept of non-overlapping CWs. While HP-DCF uses the fixed lower bound for data CW, ADCF uses an adaptive approach to determine the lower bound. It is observed that ADCF gives a better data traffic delay performance.

The above three schemes rely on the AP to announce the desirable voice CW sizes under different VoIP traffic conditions. A modification of ADCF, called Direct Priority DCF (DP-DCF) is also proposed for the ad-hoc environment.

4.2 Proposed Scheme 1: Dynamic Backoff Adjustment (DBA)

4.2.1 DBA Design: dynamic CW selection mechanism

In this section, we present the design and the mechanism of DBA scheme. The DBA scheme aims to dynamically provide differentiated services among different traffic classes using CW differentiation. The advantage of using dynamic CW assignment is that it assigns smaller CW size to reduce the unnecessary longer backoff during the low traffic load period and hence achieves lower overall delay. Likewise, at the high traffic load period, assigning a larger CW size reduces the probability of collision under high traffic volumes and hence reduces the probability of more retransmission attempts and lower overall delay. In our study, we focus on two high priority traffic classes; one for VoIP and one for data. As a result, four CW parameters are defined in the DBA scheme, namely, $VoIP\_CW_{min}$, $VoIP\_CW_{max}$, $Data\_CW_{min}$ and $Data\_CW_{max}$. These parameters indicate the maximum and minimum contention window size for VoIP and data traffic. We define the following parameters for voice and data CWs:

Let $B_{L,V}$ be the lower bound of the CW for voice traffic

Let $B_{U,V}$ be the upper bound of the CW for voice traffic

Let $B_{L,D}$ be the lower bound of the CW for data traffic

Let $B_{U,D}$ be the upper bound of the CW for data traffic

$B_{L,V}$ and $B_{U,V}$ specify the location of the VoIP contention window. For DBA, as in EDCF, $B_{L,V}$ and $B_{U,V}$ have the following values:

$$B_{L,V} = 0 \tag{4}$$

$$B_{U,V} = \text{VoIP\_CW}_{size} - 1 \tag{5}$$

Where,

$$\text{VoIP\_CW}_{min} \leq \text{VoIP\_CW}_{size} \leq \text{VoIP\_CW}_{max} \tag{6}$$

Similarly, $B_{L,D}$ and $B_{U,D}$ specify the location of the data window and have the following values:

$$B_{L,D} = 0 \tag{7}$$

$$B_{U,D} = \text{Data\_CW}_{size} - 1 \tag{8}$$

Where,

$$\text{Data\_CW}_{min} \leq \text{Data\_CW}_{size} \leq \text{Data\_CW}_{max} \tag{9}$$

The DBA scheme basically follows the CSMA/CA scheme. The backoff is selected from 0 to minimum window size -1 after the first collision. Subsequent collisions will cause the contention window to expand according to the binary backoff algorithm. To provide service differentiation in favour of voice traffic, the binary backoff window range of the data traffic should be larger than the window range of the VoIP traffic.

Here, we also define data and voice stations as the stations attempt to send data or voice frames, respectively. A station can be a data station or voice station at any given time depending on what type of frame the station attempts to send.

Thus far, we have seen that the variables to dictate the behaviour and the performance of the scheme are VoIP_CW_{min}, VoIP_CW_{max}, Data_CW_{min} and Data_CW_{max}. The main difference for the DBA scheme from the EDCF and DCF algorithm is its ability to dynamically assign these values according to the network load. An accurate CW assignment to the appropriate traffic classes defines the success of this dynamic scheme. These values can be derived when the number of stations and their traffic characteristics (packet size, inter-arrival time, delay requirement, throughput requirement, etc.) are known to the network. Unfortunately, these characteristics could not be obtained in practice without continuous measurement. Furthermore, if the traffic load fluctuates rapidly, the measurement approach may not be very accurate.

However, if the infrastructure mode is used, AP can determine the number of voice stations in the network through the admission control process. Since voice traffic is more predictable, AP can use the information to determine the optimum CW sizes and broadcasts these values to all the stations. From our intensive simulation studies, we derived the optimum CW sizes based on the number of VoIP connections. These values are tabulated in Table 4. Simulation results have shown that the scheme gives a better service differentiation performance for the voice traffic over EDCF.

Table 4: Contention Window values recorded

Number of Voice Sessions	$VoIP\_CW_{min}$	$VoIP\_CW_{max}$	$Data\_CW_{min}$	$Data\_CW_{max}$
1	1	1	2	1024
2	4	16	16	1024
3	16	64	64	1024
4	16	64	64	1024
5	16	64	64	1024
6	16	64	64	1024
7	16	64	64	1024
8	16	128	128	1024
9	16	128	128	1024
10	16	128	128	1024
11	16	128	128	1024
12	16	128	128	1024

The table indicates that we assigned a smaller CW range for VoIP traffic and a larger CW range for data traffic. By doing so, it makes it more probable to select the random slot in favour of VoIP traffic. Simulation results have shown that the VoIP packets usually receive smaller backoff time and thus are transmitted first.

4.2.2 Shorter Variable IFS

In the EDCF scheme, it is observed that if more than one VoIP connection is going through the AP, the performance of the downlink of those connections will not be as good as the uplink. This is because AP needs to compete with other stations to access the channel while AP has more voice traffic to transmit.

In the DBA scheme, AP uses smaller IFS, called Shorter Variable IFS (SVIFS), to solve the AP unfair bandwidth competition problem for its voice traffic. The value of SVIFS is smaller than the IFS used by the VoIP traffic from the stations. Thus, VoIP traffic from the AP has a higher priority than VoIP traffic from the stations. To avoid possible monopolization of bandwidth by the AP, the maximum number of VoIP packets using SVIFS in a given period should be proportional to the number of VoIP connections through the AP. For example, if there are x number of VoIP connections going through the AP, and each connection generates a packet every 20 msec, then the maximum number of voice packet transmitted in the period of 20msec using SVIFS is x. Note that all subsequent proposed schemes in this thesis use this approach.

4.2.3 Operation of DBA

In our design, we use VoIP traffic as an example to represent the high-priority traffic and ftp traffic as an example to represent the low-priority traffic. The method described here can be extended to any high-priority and low-priority traffic types. When there is a change to the number of VoIP connections in the wireless network, the AP will announce a new $Data\_CW_{min}$, $Data\_CW_{max}$, $VoIP\_CW_{min}$ and $VoIP\_CW_{max}$ to each station in the beacon frame. A flowchart in Figure 14 shows the operation of the DBA.

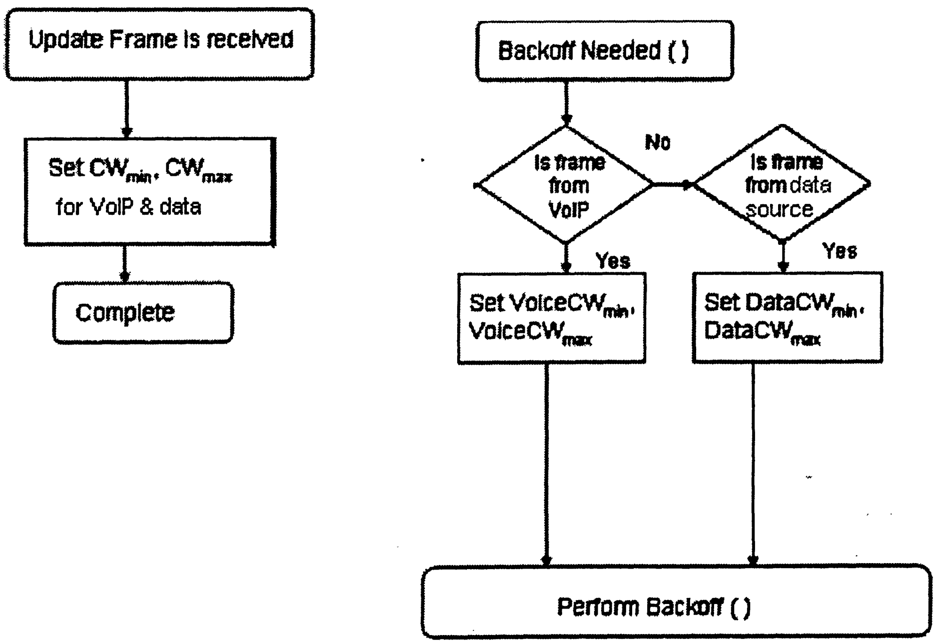


Figure 14: Flow chart of DBA

There are two stages in the DBA process implemented at the workstation. The first stage is to store the Data_CW_{min}, Data_CW_{max}, VoIP_CW_{min} and VoIP_CW_{max} values carried in the beacon frame. The second stage happens when the backoff procedure is invoked. The Backoff_Needed( ) instance would check if the frame is a VoIP frame or a data frame. For VoIP frame, VoIP_CW_{min} and VoIP_CW_{max} are used for backoff. For data frame, Data_CW_{min} and Data_CW_{max} are used.

The DBA scheme works reasonably well in the normal range of operation. However, there is one fundamental problem with this approach, i.e. when there is a collision between a data frame and a high-priority frame (VoIP) in some scenarios, the data frame could still have a smaller backoff than the voice frame, thus could still be transmitted first. Consequently, the performance of voice traffic, which is delay-sensitive, is affected significantly by the data traffic load. To protect the voice traffic from the data traffic, we propose the use of non-overlapped VoIP and data CWs. In the subsequent sections, we will discuss the schemes using this method.

4.3 Proposed Scheme 2: High-Priority DCF (HP-DCF)

4.3.1 Basic Approach

In both EDCF and DBA,

$B_{L,V} = B_{L,D} = 0$

Consequently, the CWs for data and voice are overlapping as illustrated in Figure 15.

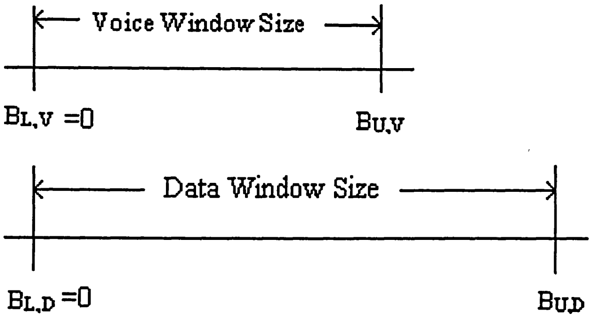


Figure 15: Window size for voice and data

The service differentiation between voice and data is derived based on the window sizes. Data traffic usually has a larger window size, and on average, has a larger backoff time. However, this kind of approach cannot completely guarantee that the data traffic has a shorter backoff. Consequently, the performance of voice traffic can still be affected significantly by the data traffic. In order to insulate the voice traffic better from the data traffic, we propose the following relationship between $B_{L,V}$ and $B_{U,V}$.

$$B_{L,D} = B_{U,V} + 1 \tag{10}$$

Thus,

$$B_{L,D} + Data\_CW_{min} - 1 \leq B_{U,D} \leq B_{L,D} + Data\_CW_{max} - 1 \tag{11}$$

In other words, the two CWs are non-overlapping as illustrated in Figure 16.

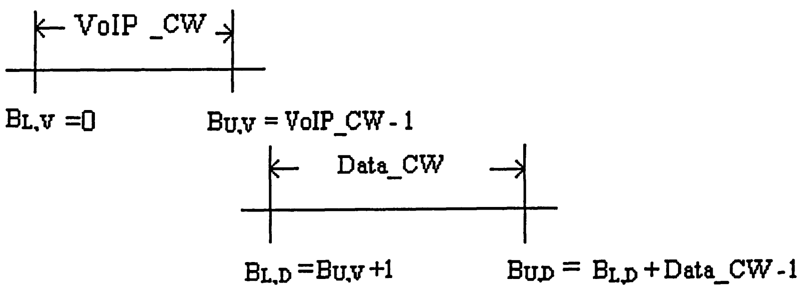


Figure 16: Non-overlapping CW

Using eq.(10), it can be guaranteed that a voice frame will have a shorter backoff than a data frame 100% of the time. This approach will provide better insulation for the voice traffic.

However, there is a problem in the implementation of eq. (10). The problem is that a data station does not know the value of $B_{U,V}$, which depends on the CW condition used by the competing voice stations. Therefore, data stations can only estimate $B_{U,V}$. Here we propose a high-priority DCF (HP-DCF) which estimates $B_{U,V}$:

$$B_{U,V} = VoIP\_CW_{max} - 1 \qquad (12)$$

With eq. (10) and (12), we can guarantee that high-priority traffic can always have a smaller backoff than lower priority traffic in the next round of competition after their collision. This leads to a better priority level assigned to the high-priority traffic.

Yet, it must be noted that a guarantee of early access to the next round of competition does not mean that VoIP frames will always access the channel earlier than the data frames. It is because VoIP frames may still compete with each other which will lead to further collisions and further backoff. The accumulated backoff of a high-priority frame may be larger than the backoff of the data frames.

The use of the non-overlapping CW mechanism introduces some penalties to the network. We can foresee that a higher delay will result for low-priority traffic due to the larger backoff time. Because of that, more bandwidth will be wasted, which will lower the overall network throughput.

4.3.2 Example

The following example illustrates the mechanism of the HP-DCF scheme. Assume there are 3 VoIP connections. Therefore, the maximum window size for voice traffic is 64. When a data station encounters the first collision, the value of $B_{L,D}$ and $B_{U,D}$ are:

$B_{L,D} = 64$

$B_{U,D} = 64 + 64 - 1 = 127$

What it means is that the backoff value for the data station is chosen randomly in the range of [64,127]. On the other hand, the backoff value chosen by the voice station is always less than 64. Consequently, voice packets are always transmitted first in the next round of competition.

This example also illustrates the problem to this approach. Specifically, the backoff of the data packet is exceedingly high. This will cause long data packet delay and lower channel utilization.

4.4 Proposed Scheme 3: Adaptive DCF (ADCF)

Based on the previous discussion, we propose another CW separation approach that gives a lower data packet delay than HP-DCF. Back in eq. (12), the lower bound of the data traffic, $B_{L,D}$, is set to a fixed value which is $VoIP\_CW_{max}$. This estimation causes excessive delay to data traffic and hence lowers the throughput. For that reason, we designed an adaptive-priority DCF scheme (ADCF). In ADCF, the value of $B_{L,D}$ changes according to the number of consecutive collisions.

Let n represents the number of consecutive collision experienced by a data station

Then,

$$B_{L,D} = 0 \quad \text{for } n=0 \quad \text{(zero collision)} \tag{13}$$

$$B_{L,D} = VoIP\_CW_{min} \times 2^{n-1} \quad (n^{th} \text{ collisions, } n > 0) \tag{14}$$

$$B_{U,D} = B_{L,D} + Data\_CW_{min} - 1 \tag{15}$$

The above equations are based on the assumption that the data and voice stations experience the same number of consecutive collisions. If the assumption holds, the voice packet will have an earlier access to the competed data packet in the next round of retransmission. The underlying assumption may not always be valid. Consequently, the voice packet will not be 100% guaranteed to win the next round of competition against the data packet. Nevertheless, the VoIP traffic should have a high probability of accessing the channel sooner than the data traffic. In addition, the average backoff time of the data traffic should be smaller than HP-DCF. Thus, we expected that this approach will yield a lower data traffic delay and higher throughput than the HP-DCF approach.

4.4.1 Example

Assume there are 3 VoIP connections. Therefore, the maximum window size for voice traffic is 64. When a data station encounters the first collision, the value of $B_{L,D}$ and $B_{U,D}$ are

$B_{L,D} = 16 \times 2^{1-1} = 16;$

$B_{U,D} = 16 + 64 - 1 = 79;$

Consequently, the backoff value chosen for the data station has a range from 16 to 79. On the other hand, the backoff value chosen by the voice station is always less than 16. Hence, voice packets are always transmitted first in the next round of competition.

In contrast to the HP-DCF, the backoff of the data packet in ADCF is usually lower than HP-DCF. After the first collision, the backoff for data is 16, which is much smaller than 64 in the HP-DCF case. So, the delay experienced by the data packets should be lower than HP-DCF.

4.5 Proposed Scheme 4: Direct Priority DCF (DP-DCF) for ad-hoc

In the previous section, we explored the methods of providing differentiated services using contention-based mechanism in the infrastructure mode. Yet, we would like to extend the approach to the ad-hoc environment. The proposed scheme is called Direct-Priority DCF (DP-DCF).

4.5.1 DP-DCF design: adaptive backoff in ad-hoc environment

The DP-DCF scheme is the modification of ADCF scheme in section 4.4. Instead of relying on AP to indicate the number of voice connections and to assign CW sizes for backoff selection, the DP-DCF scheme adjusts the contention window sizes based on the number of consecutive collisions as follows:

$$\text{VoIP\_CW}_{size} = 2^{n}, \quad n = \text{number of consecutive collisions} \tag{16}$$

$$B_{L,V} = 0 \tag{17}$$

$$B_{U,V} = \text{VoIP\_CW}_{size} - 1 = 2^{n} - 1 \leq \text{VoIP\_CW}_{max} - 1 \tag{18}$$

The CW for the data traffic is derived from the following equations:

$$B_{L,D} = 0, \quad n = 0 \tag{19}$$

$$B_{L,D} = 2^{n}, n > 0 \tag{20}$$

$$B_{U,V} = B_{L,D} + \text{Data\_CW}_{size} - 1 \tag{21}$$

Here we choose the minimum Data_CW_{size} to be 16, same value used in EDCF.

In the ad-hoc environment, stations do not know the current number of VoIP connections, the voice CW uses the concept of binary exponential backoff. Note that in this case the $VoIP\_CW_{min} = 2$. For the data CW, we use the similar concept of ADCF and let the value of $B_{L,D}$ be changed based on the number of consecutive collisions. Since VoIP_CWmin =2, eq. (14) becomes eq. (20).

To illustrate the mapping of collision counts to the CW range, Figure 17 gives a graphical representation of the CWs. The ftp CW range is totally separated from the VoIP CW range for each successive collision count.

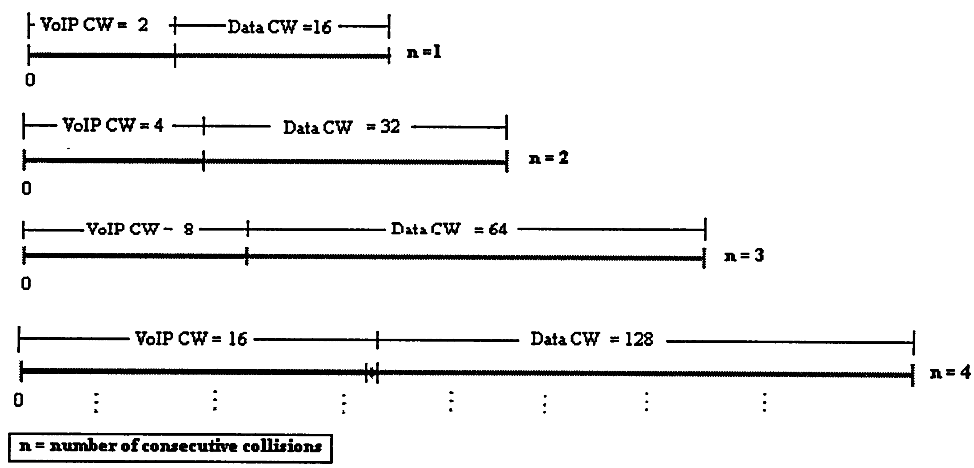


Figure 17: DP-DCF scheme graphical illustrations

4.5.3 Example for DP-DCF

When a data station encounters the first collision, the value of $B_{L,D}$ and $B_{U,D}$ are

$$B_{L,D} = 2^1 = 2$$

$$B_{U,D} = 2 + 16 - 1 = 17$$

Thus, after the first collision, the backoff value chosen for the data station is between 2 and 17. On the other hand, the backoff value chosen by the voice station is always less than 2. So, voice packets are always transmitted first in the next round of competition.

4.6 Conclusion

We have proposed four schemes that are aimed to improve the QoS support in the 802.11 standard. The first three schemes use the infrastructure mode to tackle the Differentiated Services needs for various classes of traffic. We use SIVFS for AP to address the imbalance of AP's uplink/downlink bandwidth problem. To better utilize the bandwidth, these schemes use dynamic CW assignment as opposed to the inflexible static CW assignment in EDCF. To reduce the adverse effect of low-priority traffic class gaining earlier access than the high-priority traffic class, we developed HP-DCF that uses non-overlapping CW technique to protect the high priority traffic class. Yet, the HP-DCF CW assignment for the lower classes of traffic is inflexible. Hence, we introduced an adaptive mechanism, called ADCF, to the non-overlapping CW technique to reduce delay of low-priority traffic class, and thus utilize the bandwidth resources more efficiently. To deliver better differentiated services in the ad-hoc environment, we proposed DP-DCF. It is also based on non-overlapping CW techniques and gives high-priority traffic more protection against heavy traffic injection from low-priority traffic sources.

5 Simulation and Evaluation

5.1 Simulation setup

5.1.1 Assumptions

For our analysis, several assumptions have been made:

- All wireless stations and AP are fixed and hence no mobility
- Hidden terminal problem does not exist. Every station can sense all the other's transmission and hence RTS/CTS is not used
- In our setting , since all the stations are fixed and are within a good radio reception, we do not simulate the problem of signal fading and channel error due to bad reception
- Fragmentation is ignored. In our scenario, we do not study the effect of fragmentation

5.1.2 Simulation setup parameters

OPNET is used to evaluate the performance of our proposed Model. Table 5 and Table 6 show a summary of the MAC and PHY layer parameters and packet overhead.

Table 5: MAC layer parameters

RTP Header Size	12 bytes
IP Header Size	20 bytes
UDP Header Size	8 bytes
MAC Header Size	34 bytes
Physical Header Size	192 bits (24 bytes)
VoIP Data	160 bytes

Table 6: MAC and PHY layer parameters

PHY layer specification	FHSS
Transmission Rate	1 Mbps
SIFS	10μs

All stations transmit frames at 11Mbps as in the 802.11b PHY. The simulation consists of two types of traffic: VoIP and ftp data. We use a buffer size of 2500 bytes for voice and 256000 bytes for data to simulate the effect of frame drop due to excessive delay.

To analyse the performance of the proposed scheme, EDCF is also implemented for comparison. The EDCF default parameters are shown in Table 7.

Table 7: EDCF default parameters

Type	AC	AIFS	CW_{min}	CW_{max}
Voice	3	2	3	7
Data	0	3	15	1023

Without the loss of generality, we also assume that each station can only transmit one type of traffic, either VoIP or ftp. In the simulation scenario, stations with VoIP traffic are considered high priority and require prioritized access to the wireless medium; stations with ftp traffic are considered low priority in accessing the wireless medium.

5.1.3 VoIP packet generation

In our scenario, G.711 μ-Law voice codec is used. It generates 64 kbit/s based on an 8-bit pulse coded modulation (PCM) with a sampling rate of 8000 sample/second. VoIP packet is generated every 20 msec. That is, 8 kbytes/sec x 20 msec = 160 byte voice data. The packet is carried in RTP encapsulated by UDP and IP headers. The VoIP datagram is then further encapsulated by the IEEE 802.2 Sub-Network Access protocol (SNAP) header in order to pass through the 802.11 network.

Hence, the VoIP packet size in 802.11 WLAN is:

160 byte Data + 40 RTP/UDP/IP header + 8 byte SNAP header = 208 bytes per VOIP packet

The second-layer overhead includes additional 58 bytes, which consists of 34-byte MAC and 24-byte PHY headers.

5.2 Simulation scenarios

Our simulation contains 3 ftp stations (fixed for all scenarios) and VoIP stations are added into the scenario one by one to increase the load of the system gradually until it reaches 11 VoIP stations. Each voice call is treated as a variable bit rate (VBR) connection with an uplink and a downlink uni-directional flows. VoIP station generates packets with a size of 208 bytes at an inter-arrival interval of 20msec during the talk-spurt period. Each VoIP station initiates a two-way voice call to a corresponding wired station which is connected to the Ethernet backbone through the Access Point. The QoS requirement in the Ethernet backbone connection between the wired station and the Access Point is always satisfied as the link speed is 100Mbps, which is more than sufficient for the aggregated traffic loads travelling through the backbone. Three ftp stations generate packets every 34 msec from the WLAN with a size of 1400 bytes. Although the IEEE 802.11 allows for up to 2312 bytes of data payload without fragmentation, the traffic travels through the wired LAN (in our scenario) sets a constraint to the maximum payload size of 1500 bytes [3]. So, we choose 1400 bytes as the best-effort payload size. Inter-arrival times are based on the Pareto distribution with $\alpha = 1.9$. The baseline scenario is having one VoIP station connect to the Ethernet with three ftp traffic flows. The VoIP talk spurt and silence spurt are set to 40/60 split which correspond to 1 second of talk spurt (incoming and outgoing) and 1.5 seconds of silence spurt (incoming and outgoing). Both talk spurt and silence spurt are modelled using the Pareto distribution with $\alpha = 1.9$. The scenario for 1 VoIP session is illustrated in Figure 18 and the scenario for 11 VoIP sessions is depicted in Figure 19:

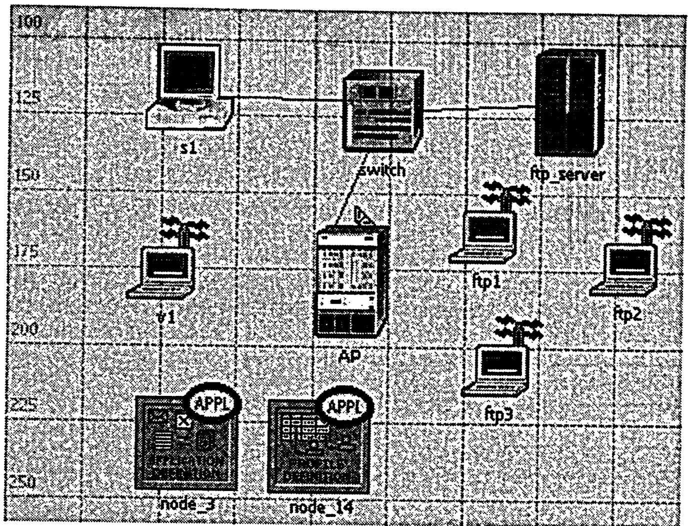


Figure 18: Simulation Topology –1VoIP session

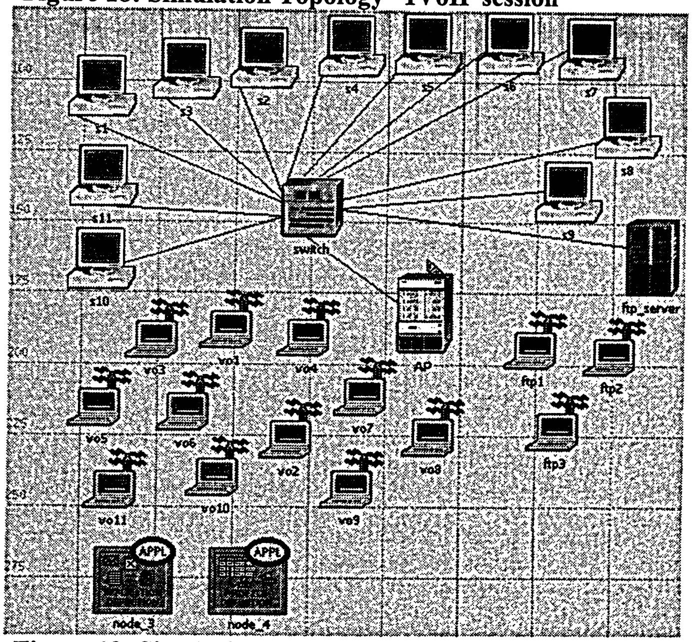


Figure 19: Simulation Topology –11 VoIP sessions

The effectiveness and efficiency of the four proposed schemes are investigated and compared with those of the legacy DCF, PCF and EDCF schemes. In particular, performances such as throughput and delay are compared. Table 8 and Table 9 summarize the simulation parameters.

Table 8: Simulation Parameters for VoIP

VoIP	
Codec	G.711
Frame Size	20msec
Start time	0.02sec
Incoming Talk Spurt	1 sec (40%) pareto (1.9)
Outgoing Talk Spurt	1 sec (40%)
Incoming Silence Spurt	1.5 sec (60%)
Outgoing Silence Spurt	1.5 sec (60%)

Table 9: Simulation Parameters for ftp

FTP	
Inter-request time	Pareto (0.016, 1.9)
File Size	1400 bytes
Start time	0.5 sec

5.3 Simulation Results and Evaluation

In this section, we present the simulation results and the analyses of the proposed schemes. We first analyse the performance of DCF with the presence of both real-time and data traffic. Then, the analysis of PCF is presented and compared with DCF to examine the ability and deficiency in supporting QoS in PCF scheme. Next, the simulation result for DBA is presented and compared with EDCF to study the effectiveness of the DBA scheme. After that, theDBA scheme is analysed and evaluated against HP-DCF and ADCF and the differences of these three schemes are discussed. Finally, the simulated results for the DP-DCF scheme in the ad-hoc environment are presented and analysed.

5.3.1 Simulation results and analysis for DCF performance

The performances of VoIP traffic and data traffic under the DCF scheme are presented in Table 10 and 11:

Table 10: Mean VoIP and FTP delay (sec) for different no. of voice connections

	Number of voice connections			
	1	3	7	11
VoIP	0.00179	0.0149	0.434	2.46
FTP	0.00234	0.01628	0.438	2.536

Table 11: Mean FTP throughput (bits/sec) for different no. of voice connections

	Number of voice connections			
	1	3	7	11
FTP /per connect	245,104	240,103	13,561	3682

From Table 10, we see that both VoIP traffic and FTP traffic exhibit similar delay as a result of no service differentiation among these two traffic classes. The VoIP delay is a bit lower because its data size is smaller than that of data traffic. The results in Table 10 also indicate that DCF cannot support large number of VoIP connections. The maximum delay requirement for VoIP traffic is 150 ms, yet the delay for VoIP traffic already reaches around 400 ms when the number of voice connections is 7.

As shown in Table 11, throughput for data traffic exhibits low value as the traffic load increases. This is due to the DCF contention based channel access mechanism, which causes large delay and heavy packet drops. Such heavy packet drops is due to the buffer overflow, even though the buffer size for data traffic queue has already been set to a large value.

5.3.2 Simulation results and analysis for PCF performance

Since DCF fails in supporting differentiated services, we tested the other IEEE 802.11 standard scheme called PCF and performed simulations to validate its ability in providing QoS. We modified the PCF scheme to better support the real-time traffic: we restricted the contention free period for use of real-time traffic only and left the contention period for the use of data and real-time traffic. Since VoIP traffic source generates a packet every 20 ms, to reduce the effect of jitter delay due to buffering periodic VoIP packet, we set the superframe size to the same value. Thus, we can better match the VoIP delay requirement by utilizing the characteristics of the PCF scheme.

5.3.2.1 Comparison of PCF and DCF throughput performance

We compare the results of DCF and PCF to demonstrate the characteristics of PCF.

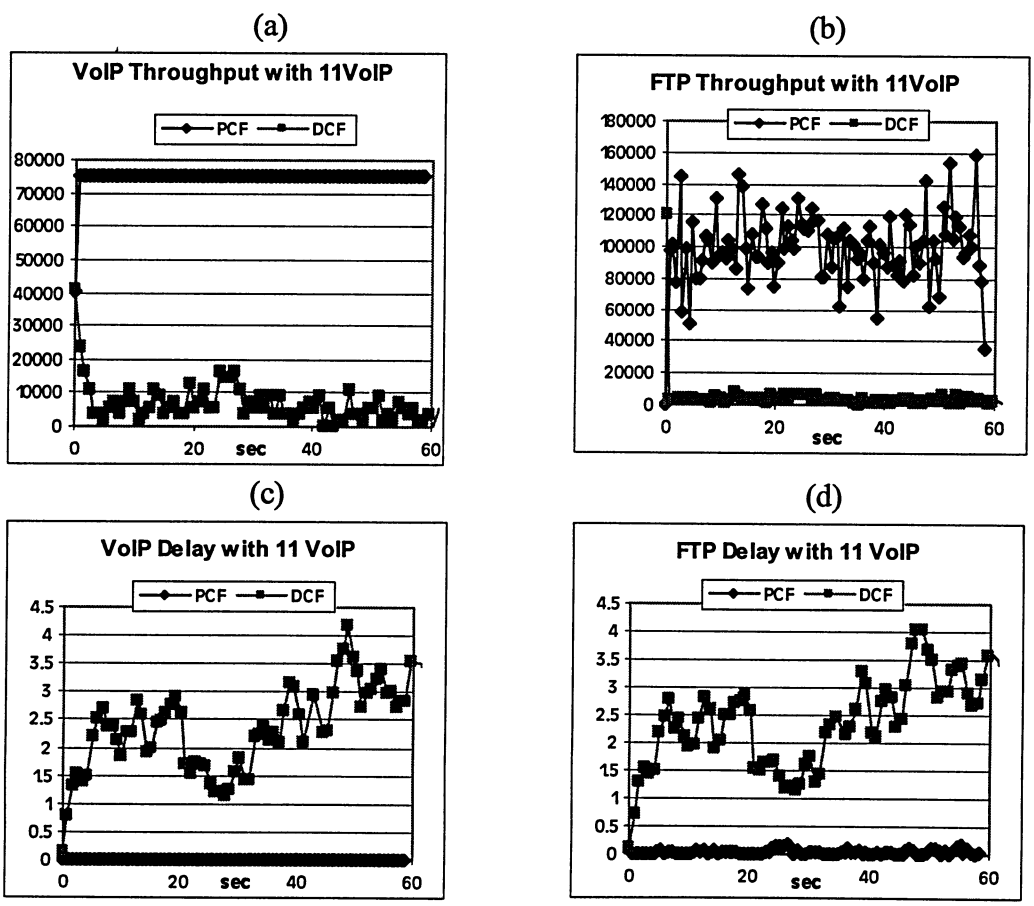


Figure 20: DCF and PCF comparison for 11 VoIP

In Figure 20a and 20b, both VoIP and FTP throughputs under DCF are low when the number of voice calls is 11. These indicate a large amount of packet drops. Under the PCF scheme, VoIP traffic streams are guaranteed for contention-free period access. As a result, the VoIP throughput is maintained at high level as shown in Figure 20a. In addition, the FTP throughput is better than the DCF as shown in Figure 20b. By studying the VoIP delay in Figure 20c, we see clearly that DCF fails to provide an acceptable VoIP delay due to no traffic differentiation in its channel access mechanism. In contrasts, the PCF scheme is able to maintain low delay for VoIP traffic. Also, in figure 20d, the PCF scheme yields a lower delay in FTP traffic than DCF.

What we have found so far is that PCF is capable in providing a high degree of QoS guarantee in supporting real-time traffic. However, PCF scheme introduces more control traffic and requires a point coordinator (support infrastructure mode only). Also, the PCF scheme itself is more complicated than contention based mechanism. Hence, we focus on the improvement over the contention-based approach. The IEEE 802.11e standard EDCF claimed that it can provide QoS support based on the differentiated services structure. We will use it as a reference scheme to investigate the performance of our proposed schemes in the next section.

5.3.3 Simulation Results and Analysis for DBA

5.3.3.1 EDCF and DBA throughput performance

Table 12: Mean VoIP delay (sec) for different no. of VoIP connections in DBA & EDCF

	Number of voice connections			
	1	3	7	11
DBA	0.000131	0.000334	0.000636	0.002073
EDCF	0.000242	0.000376	0.003897	0.005449

In Table 12, we see that the VoIP delay increases as the VoIP traffic increases for both EDCF and DBA. DBA has a lower delay than EDCF for VoIP traffic stream. Yet, both schemes yield acceptable delay values that meet the QoS requirement. Meanwhile, the DBA scheme exhibits lower VoIP delay than EDCF. We see that the VoIP delay for 7 VoIP connections is much lower than the EDCF scheme as a result of the dynamic assignments of CWs. Since EDCF deploys a small and fixed range of CW for voice traffic, the delay performance degrades as the number of VoIP connections increases. In contrast, DBA is affected in lesser degree because of using a flexible range of CW assignments.

Table 13: Mean FTP delay (sec) for different no. of VoIP connections in DBA & EDCF

	Number of voice connections			
	1	3	7	11
DBA	0.001008	0.001017	0.001955	0.003824
EDCF	0.001065	0.001081	0.006402	0.007049

The FTP traffic delay in DBA is better than EDCF as shown in Table 13. With an increasing number of VoIP connections, the gain in delay performance for DBA becomes more obvious. With 7 or more VoIP connections in DBA, ftp traffic delay is better than EDCF.

Table 14: Mean VoIP packet loss rate for different no. of VoIP connections in DBA & EDCF

	Number of voice connections				
	1	3	7	8	11
DBA /per VoIP	0	0	0.00854	0.00131	0.0319
EDCF /per VoIP	0	0	0.0089	0.00211	0.0444

The VoIP packet loss rate is shown in Table 14. The packet loss rate for both DBA and EDCF is similar. This is because both schemes deploy similar differentiated services technique for various traffic classes. Packet loss rates for both schemes are acceptable for VoIP calls. Note that increasing the number of voice connections increases the packet loss rate.

Table 15: Mean FTP Throughput (bits/sec) for different no. of VoIP connections in DBA & EDCF

	Number of voice connections			
	1	3	7	11
DBA / per connect	240,742	240,393	148,779	59,191
EDCF /per connect	240,350	219,891	223,412	68,321

In Table 15, we found that the FTP throughput is similar in DBA and EDCF in irregardless of 1 or 3 VoIP connections. While increasing the number of VoIP connections, the FTP traffic class throughput is affected. The EDCF has a better ftp data throughput when the number of voice connections is large. It is expected because the DBA scheme has a larger data CW, which leads to lower throughput.

Thus far, we see that increasing the number of VoIP connections degrades the VoIP delay and FTP traffic performance. Next, we would like to investigate the effect of increasing FTP traffic load on the delay sensitive VoIP traffic.

5.3.3.2 Delay performance for DBA scheme with various Data load

To further investigate the DBA performance with the effect of various data load conditions, we vary the data traffic load. The data is generated from the UDP sources. The following results are simulated from 11VoIP connections:

Table 16: Mean VoIP delay with 11VoIP connections under different data traffic loads with 25 data stations

VoIP delay for 11 VoIP connections with various data loads			
Data load	20%	40%	60%
DBA	0.00115	0.001385	0.0019
EDCF	0.001168	0.001487	0.002

Table 17: VoIP packet loss rate for 11VoIP connections under different data traffic loads with 25 data stations

VoIP packet loss rate for 11 VoIP connections with various data loads			
Data load	20%	40%	60%
DBA	0.076	0.0936	0.109
EDCF	0.079	0.109	0.124

The VoIP packet loss rate is high when examining Table 17. The increase in data traffic loads significantly impact on the quality of VoIP calls. The acceptable quality for a VoIP call is within 5% packet loss rate. We see that the increase in number of ftp stations causes the VoIP packet loss rate to increase to an unacceptable level. In addition, a higher data load also degrades the quality of VoIP calls significantly. The reason is that there is no complete protection to guarantee VoIP access. The DBA mechanism works well as long as the current traffic load is within the network limit. If the amount of data traffic is excessive, voice quality will suffers.

Even though the addition of data traffic has a slight impact on the mean VoIP delay as shown in Table 16, the high packet loss rate makes the voice call unbearable.

Next, we analyse the performance of HP-DCF and ADCF schemes and see if they can give a better protection to the voice traffic against the data traffic.

5.3.3.3 Performance analysis for HP-DCF and ADCF

To investigate the characteristics of the HP-DCF and ADCF schemes, we compare the simulation results of these schemes. We present scenarios and results that are significant to our analysis and discussion.

Table 18: Mean VoIP delay (sec) for different no. of VoIP connections in HP-DCF & ADCF

	Number of voice connections		
	1	7	11
HP-DCF	0.000340	0.000363	0.000450
ADCF	0.000412	0.000483	0.000892

In Table 18, we found that the VoIP delay increases as the number of voice connections increases for both HP-DCF and ADCF schemes. Yet, the HP-DCF scheme yields a lower delay than the ADCF for VoIP traffic. The result is expected since HP-DCF is designed to provide the highest degree of priority to the voice traffic while ADCF is aimed to improve the data traffic performance. In addition, there is no VoIP packet loss rate in HP-DCF. While for the ADCF, there is an extremely low packet loss rate.

Utilizing the adaptive nature of ADCF, the data traffic delay is improved as seen in Table 19.

Table 19: Mean FTP delay (sec) for different no. of VoIP connections in HP-DCF & ADCF

	Number of voice connections		
	1	7	11
HP-DCF	0.00109	0.00233	0.0032
ADCF	0.000977	0.002	0.002636

In HP-DCF, regardless of the number of voice connections, data traffic will start its initial backoff in high value, which causes the FTP delay to increase unnecessarily. As a result, the data throughput suffers as shown in Table 20.

Table 20: Mean FTP Throughput (bits/sec) for different no. of VoIP connections in HP-DCF & ADCF

	Number of voice connections		
	1	7	11
HP-DCF /per connect	251,630	1130	76
ADCF /per connect	256283	182,046	21,871

From Table 20, FTP throughput is affected negatively by the increasing number of voice calls for both HP-DCF and ADCF. Since HP-DCF uses a fixed high bound CW, there is very low throughput for HP-DCF. While utilizing ADCF, the FTP throughput is significantly improved over HP-DCF. In the case of 11VoIP connections, the FTP delays in both schemes appear to be better than the EDCF case. This is due to the fact that both HP-DCF and ADCF have lower data throughputs than EDCF. We expect that EDCF will have a better data delay performance if the throughputs of these schemes are the same. In DBA and EDCF, we see that VoIP delay is affected by the FTP traffic significantly. We will investigate to what degree the performance of VoIP traffic is affected by the introduction of various UDP data traffic loads in HP-DCF and ADCF schemes.

Table 21: Mean VoIP delay (sec) for different data traffic loads with 25 data stations in HP-DCF & ADCF

Mean VoIP delay for 11 VoIP connections with various data loads			
Data load	20%	40%	60%
HP-DCF	0.00035	0.000502	0.0007
ADCF	0.0012	0.00151	0.0018

Table 22: VoIP packet loss rate for different data traffic loads with 25 data stations in HP-DCF & ADCF

VoIP packet loss rate for 11 VoIP connections with various data loads			
Data load	20%	40%	60%
HP-DCF	0.0311	0.083	0.097
ADCF	0.044	0.087	0.103

From Table 21, we see that HP-DCF and ADCF have a very low delay under various data loads. In terms of VoIP packet loss rate, we see that both schemes exhibit similar loss rates as found in Table 22. We see that the VoIP packet loss rates for both schemes are smaller than those of the DBA and EDCF. The ADCF scheme has a slightly higher value for VoIP packet loss rate than the HP-DCF. Because the HP-DCF uses a fixed bound CW for data, therefore it lowers the throughput of the data traffic and gives a smaller delay for voice traffic. A smaller delay in voice traffic yields earlier access and hence reduces the VoIP packet loss rate. Notice that the drop rates for both HP-DCF and ADCF are lower than DBA and EDCF. The results show that VoIP traffic is better protected against the high data traffic load using HP-DCF and ADCF than EDCF and DBA.

5.3.4 Results and analysis for DP-DCF

The simulation scenario for the ad-hoc environment is different than that in the infrastructure mode. Figure 21 depicts the scenario topology.

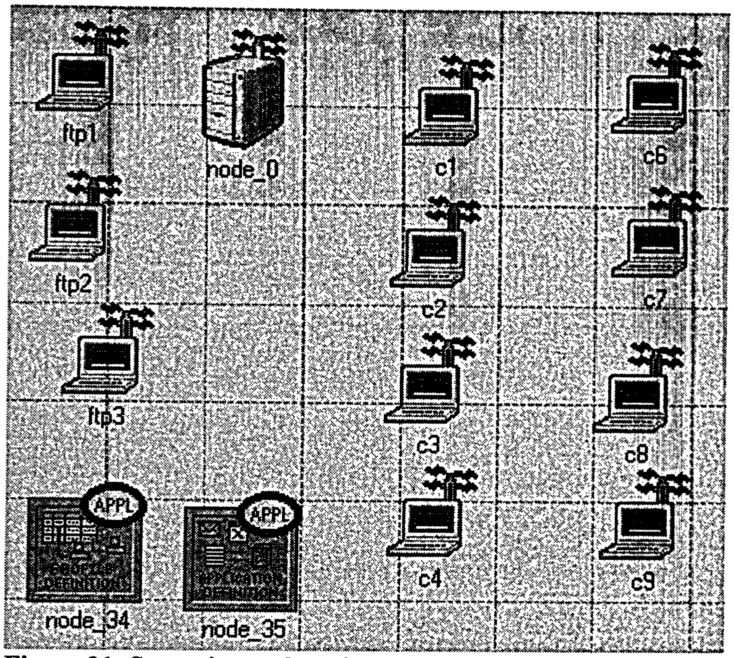


Figure 21: Scenario topology for the ad-hoc mode- 8 VoIP sessions

The differences in topology setup for ad-hoc and infrastructure mode is that there is no AP involved in the ad-hoc environment. Wireless Stations directly communicate with each others.

The results of the DP-DCF are compared with the EDCF scheme as shown below:

Table 23: Mean VoIP delay (sec) for different no. of VoIP connections in DP-DCF & EDCF

	Number of voice connections						
	2	4	5	6	7	8	12
DP-DCF	0.00031	0.00038	0.000398	0.000581	0.000431	0.00070	0.0029
EDCF	0.00032	0.00067	0.000980	0.000370	0.000661	0.00074	0.0293

Table 24: Mean FTP delay (sec) for different no. of VoIP connections in DP-DCF & EDCF

	Number of voice connections						
	2	4	5	6	7	8	12
DP-DCF	0.00147	0.00190	0.00213	0.00267	0.002826	0.0053	8.59
EDCF	0.00206	0.00263	0.00302	0.00339	0.00484	0.00831	13.53

Table 25: Mean FTP Throughput (bits/sec) for different no. of VoIP connections in DP-DCF & EDCF

	Number of voice connections						
	2	4	5	6	7	8	12
DP-DCF	243,923	256,585	246,127	266,516	256,860	269,561	242,965
EDCF	229,032	251,404	244,501	247,494	252,971	269,288	110,453

In Tables 21 and 22, we see that both VoIP delay and FTP delay in DP-DCF are obviously lower than EDCF. By increasing the number of VoIP connections, both VoIP and FTP delay increases. Figure 22 shows the dropped data packets for both VoIP and data traffic with 12 VoIP connections. We see that the VoIP loss rate for DP-DCF with 12 VoIP connections is extremely low, but it does not affect the quality of VoIP call. While for the EDCF with 12 VoIP connections, the VoIP packet loss rate seems to be higher, and may severely degrade the voice quality.

Data throughput is similar for both DP-DCF and EDCF scheme with 2 to 8 voice connections as shown in Table 23. However, in the case with 12 voice connections, the data throughput in DP-DCF outperforms the EDCF scheme significantly.

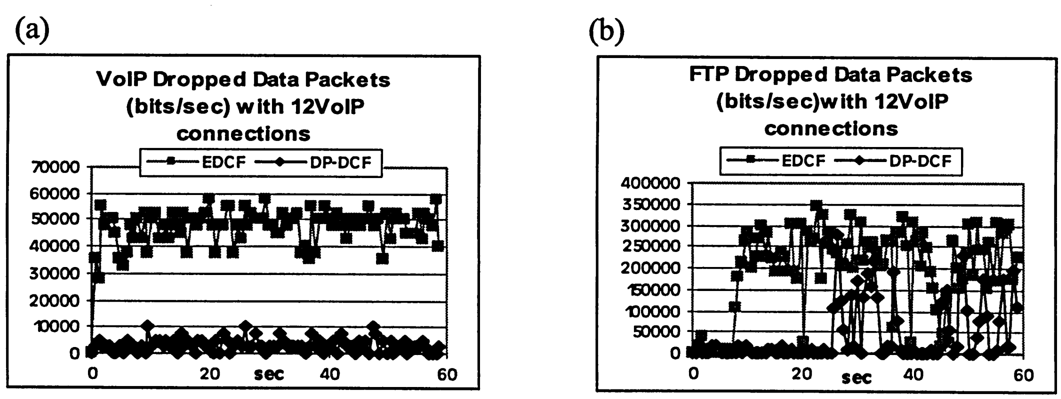


Figure 22: DP-DCF results for 12 VoIP connections with FTP present

In Table 26 and Table 27, we present the results of VoIP delay and VoIP packet loss rate with various data loads to study its performance against increasing data traffic loads.

Table 26: Mean VoIP delay (sec) for different data traffic loads with 25 data stations in DP-DCF & EDCF

VoIP delay for 11 VoIP connections with various data loads			
data load	20%	40%	60%
DP-DCF	0.011292	0.013198	0.015348
EDCF	0.4325	0.4713	0.589

Table 27: VoIP packet loss rate for different data traffic loads with 25 data stations in HP-DCF & ADCF

VoIP packet loss rate for 11 VoIP connections with various data loads			
data load	20%	40%	60%
DP-DCF	0.01649	0.0149	0.02112
EDCF	0.0678	0.0755	0.0773

The VoIP delay in DP-DCF is very low in comparison to the EDCF with increasing data loads. The EDCF yields unacceptable VoIP delay that would degrade the quality of voice call. Also, the VoIP packet loss rate in EDCF is higher than the DP-DCF. At this point, we see that DP-DCF is able to safeguard VoIP traffic against excessive data traffic injections.

We conclude that the DP-DCF scheme yields a better VoIP delay performance over EDCF while at the same time. It supports a higher data throughput than EDCF in the ad-hoc environment. Thus, this scheme is deemed suitable for the ad-hoc environment in supporting differentiated services. The findings from this scheme give leverage to further research on the priority assignment in the WLAN.

5.4 Summary

In this chapter, the performances of various proposed schemes are studied through simulations.

First, the performances of DCF and PCF are compared, and the deficiencies of DCF in QoS support are highlighted. We then compared DBA with EDCF. It is found that DBA outperforms EDCF in terms of voice delay. However, EDCF can support slightly higher data throughput. The performances of HP-DCF and ADCF are subsequently studied. Both HP-DCF and ADCF provide better insulation to VoIP traffic against data traffic. The trade-off is that these schemes give lower data throughput. Among the two schemes, ADCF supports a higher data throughput. Since both schemes provide relatively good support for VoIP traffic, it seems that ADCF is more preferable than HP-DCF.

Finally, we compared DP-DCF with EDCF in the Ad-hoc environment. We found DP-DCF to be better than EDCF in every aspect. However, more research is required to make it accommodate more traffic classes.

6 Conclusions

In this paper, we presented several novel schemes aimed at solving the lack of QoS support in the legacy 802.11 DCF mechanism in both integrated wired/wireless infrastructure and ad-hoc environment. We performed extensive simulations to study the delay and throughput performance in different ranges of contention windows. We hope that the selected parameters would contribute to further research in optimizing bandwidth utilization.

We proposed four schemes to provide differentiated services in WLAN. The first proposed DBA scheme is capable of maintaining low VoIP delay. Simulation results showing the dynamic CW assignment yield a relatively lower amount of delay than the DCF and EDCF scheme. In terms of data throughput, the proposed DBA scheme performs better than the other two schemes. DBA's dynamic assignment based on the assumption that AP knows the number of VoIP connections in the WLAN and hence assigns the appropriate CW sizes. The AP then announces this information to all stations registered in the AP.

The second proposed scheme, HP-DCF, is designed to provide a better protection to the VoIP traffic against a heavy data traffic load. It uses the non-overlapping CW concept, in which the lower bound of the data CW is always larger than the upper bound of the voice CW. The non-overlapping CWs technique gives the voice traffic a shorter backoff than the data traffic.

The third proposed scheme, ADCF, is similar to HP-DCF. Both utilize the non-overlapping CW technique and take the successive collision count to calculate the range of CW for each class of traffic when collision happens. The main difference is that ADCF deploys an adaptive technique in determining the lower bound for the data CW while HP-DCF uses a fixed lower bound. The results reveal that the adaptive scheme yields a better throughput for the best-effort traffic.

With the success of HP-DCF and ADCF, we proposed the fourth scheme, namely DP-DCF, which applies the non-overlapping CW technique in the ad-hoc environment. This scheme makes use of successive collision counts in determining CW sizes. It differs from the other three schemes as it does not require any dynamic CW announcement. Simulation results have shown that this scheme also provides a good protection to real-time traffic.

The proposed four schemes cover a broad range of WLAN scenario that addresses real-time traffic QoS requirements. We found that the four schemes are successful in yielding a better delay and throughput guarantee for real-time traffic over the legacy IEEE 802.11 standard.

In light of these findings, further research is encouraged to extend the proposed schemes in supporting QoS over WLAN. For instance, one may extend the schemes to support video traffic by adding the video traffic class. Also, one may add an admission control mechanism over the four schemes to regulate traffic injections and deliver QoS guarantee in infrastructure network. Modifications may be added to further enhance the non-overlapped CW mechanism to improve data throughput and facilitate research in ad-hoc environment.

7 References

[1] IEEE, Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification. Institute of Electrical and Electronics Engineers, Inc., 1997.

[2] IEEE, "Part 11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium, Access control (MAC) Enhancements for Quality of Service (QoS)," ANSI/IEEE Std 802.11e, Draft 13.0, Jan 2005.

[3] A. Kopsel and A. Wolisz, "Voice transmission in an IEEE 802.11 WLAN based access network," Tech. University Berlin, Germany, Proceedings of the 4th ACM international workshop on Wireless mobile multimedia, Rome, Italy, online link, *http://www.tkn.tu-berlin.de/publications/papers/wowmom01.pdf*, accessed Jan 2005.

[4] D. Gu, J. Zhang, "QoS Enhancement in 802.11 Wireless Local Area Networks," Mitsubishi Electric Research Laboratories, 2003

[5] A. Grilo, M. Nunes, "Performance Evaluation of IEEE802.11E," PIMRC, IEEE, 2002

[6] S. Mangold et al., "IEEE 802.11e Wireless LAN for Quality of Service," in Proc. European Wireless '02, Florence, Italy, February, 2002, online link, *http://www.ing.unipi.it/ew2002/proceedings/H2006.pdf*, accessed Oct 2004

[7] M. Gast, "802.11 Wireless Networks: The Definitive Guide", O'Reilly, April 2002

[8] B. Crow, I. Widjaja, J. Kim, P, Sakai, "IEEE 802.11 Wireless Local Area Networks," IEEE Communication Magazine September 1997

[9] D. Pong, T. Moors, "Fairness and Capacity Trade-off in IEEE 802.11 WLANs," School Of Electrical Engineering and Telecommunications, The University of New South Wales, Australia, online link, *http://www.ee.unsw.edu.au/~timm/pubs/04lcn/published.pdf*, accessed April 2004

[10] S. Garg, M. Kappes. "Can I add a VoIP call?" submitted for publication IEEE 2003

[11] Bermai, Inc. "QoS over 802.11 IEEE 802.11e – The Need for HCCA to Stream Video," Bermai Inc., April 2004

[12] A. Lindgren, A. Almquist, O. Schelen "Evaluation of Quality of Service Schemes for IEEE 802.11 Wireless LANs," Division of Computer Science and Networking, Department of Computer Science and Electrical Engineering, Lulea University of Technology, Sweden.

[13] S. Chung, K. Piechota, "Understanding the MAC impact of 802.11e: Part 1," EE Times, 30 October, 2003.

[14] S. Chung, K. Piechota, "Understanding the MAC impact of 802.11e: Part 2," EE Times, 30 October, 2003.

[15] M.Veeraraghavan, N. Cocker, T. Moors, "Support of voice services in IEEE802.11 wireless LANs," IEEE INFOCOM 2001

[16] P. Ansel, Q. Ni, T. Turletti "An Efficient Scheduling Scheme for IEEE 802.11e," PLANETE Project, INRIA Sophia Antipolis, FRANCE

[17] D. He, C.Q. shen, "Simulation Study of IEEE 802.11e EDCF," Institute for Infocomm Research, National University of Singapore.

[18] Intelligraphics Inc., "Introduction to IEEE 802.11," online link, *http://www.intelligraphics.com/ariticles/80211_article.html*, accessed Mar 10, 2005

[19] L. Scalia, I. Tinnirello, "Differentiation mechanisms for heterogeneous traffic integration in IEEE 802.11 networks," Universita' di Palermo, Palermo, Italy.

[20] OPNET documentation, OPNET v8.0 software, http://www.opnet.com

[21] Q. Ni, L. Romdhani, T. Turletti., "A survey of QoS Enhancements for IEEE 802.11 Wireless LAN," Journal of Wireless Communications and Mobile Computing, Wiley, 2004 Vol 4, Issue 5 pp.547-566

[22] Q. Pang et al., "Performance evaluation of an adaptive backoff scheme for WLAN", Wireless Communications and Mobile Computing, Wiley, 2004, Vol 4, pp. 867-879

[23] H. Liu, J.Wu, "A Scheme for Supporting Voice over IEEE 802.11 Wireless Local Area Network" Proceedings National Science Council ROC(A) Vol 25, No.4, 2001 pp. 259-268

[24] S. Choi et al., "IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation," Multimedia and Wireless Networking Laboratory (MWNL) School of Electrical Engineering Seoul National University, Seoul Korea

[25] A. Salhotra, R. Narasimhan, R. Kopikare, "Evaluation of Contention Free Bursting in IEEE 802.11e Wireless LANs," IEEE Communications Society WCNC 2005 pp. 107-112

[26] D. Pong, T. Moors, "Call Admission Control for IEEE 802.11 Contention Access Mechanism," School of Electrical Engineering and Telecommunications, The University of New South Wales, Australia

[27] Y. Xiao, H. Li, S. Choi, "Protection and Guarantee for Voice and Video Traffic in IEEE 802.11e Wireless LANs," in Proc. IEEE INFOCOM'04, vol. 3, Hong Kong,

Mar. 2004, pp. 2152–2162, online link, *http://www.ieee-infocom.org/2004/Papers/ 44_3.PDF*, accessed Mar12, 2005

[28] D. Chen D. Gu, J. Zhang, "Supporting Real-time Traffic with QoS in IEEE 802.11e Based Home Networks," *Consumer Communications and Networking Conference (CCNC)*, pp. 205-209, January 2004, online link, *http://www.merl.com/ reports/docs/TR2004-006.pdf*, accessed Mar 12, 2005

[29] A. Velayutham, J. Chang, "An Enhanced Alternative to the IEEE 802.11e MAC Scheme," Iowa State University, online link, *http://www.cs.iastate.edu/~vel/ research/E-802.11.pdf*, accessed Mar 12, 2005

[30] I. Tinnirello, G. Bianchi, L. Scalia, "Performance Evaluation of Differentiated Access Mechanisms Effectiveness in 802.11 Networks

[31] F. Cali, M. Conti, E. Gregori, "IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism," IEEE J. Select. Areas Commun., vol. 18, no.9, pp.1774-1786, Sep. 2000

[32] Y. Peng, H. Wu, S. Cheng, K. Long, "A new self-adapt DCF algorithm," IEEE GLOBECOM 2002, pp. 87-91

[33] L. Bononi, M. Conti, E. Gregori, "Design and Performance Evaluation of an Asymptotically Optimal Backoff Algorithm for IEEE 802.11 Wireless LANs," Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 8, pp. 8025, 2000

[34] N. Ramos, D. Panigrahi, S. Dey, "Dynamic Adaptation Policies to Improve Quality of Service of Multimedia Applications in WLAN Networks," ICST/IEEE Intl. Workshop on Broadband Wireless Multimedia, San Jose, CA, October 2004, online link, *http://www.broadnets.org/2004/workshop-papersBroadwim/broadwim2004-Paper16-ramos.pdf*, accessed April 2005

[35] H. Woesner, J. Weinmiller, J. Ebert, A. Wolisz, "Modified Backoff Algorithms for DFWMAC's Distributed Coordination Function," online link, *http://www.tkn.tu-berlin.de/publications/papers/backoff-new.pdf*, accessed May 2005

[36] A. Nafaa, A. Ksentini, A. Mehaoua, "SCW: Sliding Contention Windows For Efficient Service Differentiation in IEEE 802.11 Networks," In Proc of. IEEE WCNC 2005, The IEEE Wireless Communication and Networking Conference. New Orlean, USA 2005. online link, *http://www.prism.uvsq.fr/users/kad/ ksentini_wcnc05.pdf*, accessed May 2005

[37] A. Ksentini, A. Nafaa, A. Guéroui, M. Naimi, "Adaptive service differentiation for QoS provisioning in wireless ad-hoc networks," In Proc of. ACM PE-WASUN

2004, The 1st ACM Workshop on Performances Evaluation of Wireless Adhoc, Sensor and Ubiquitous Network, Hold in conjunction with IEEE/ACM MSWiM 2004. Venice, Italia 2004, online link, *http://www.prism.uvsq.fr/users/kad/pewasun_04.pdf*, accessed Jun 2005

[38] N. Song, B.Kwak, J.Song, L.E. Miller, "Enhancement of IEEE 802.11 distributed coordination function with exponential increase exponential decrease backoff algorithm," IEEE VTC 2003 Spring, pp. 2775-2778

[39] B. Venkatakrishnan, S. Selvakennedy, "An Enhanced HCF for IEEE 802.11e Wireless Networks," MSWiM'04 October 4-6, 2004, Venezia, Italy.