Studies on Per-Flow Fairness in IEEE 802.11 Wireless LANs

B. A. Hirantha Sithira Abeysekera

Department of Information and Communications Technology
Division of Electrical, Electronic and Information Engineering
Graduate School of Engineering
Osaka University, Japan

January 2010

Dedicated to My Beloved Parents, Wife Menaka, and to Little Saki

Preface

THIS dissertation presents studies on per-flow fairness in IEEE 802.11 wireless LANs. The contents are based on the research I carried out during my Ph.D. course at the Department of Information and Communications Technology, Graduate School of Engineering, Osaka University, Japan. The chapters in this dissertation address three key fairness issues occurred in IEEE 802.11-based wireless local area networks (LANs).

IEEE 802.11-based wireless LANs, sometimes referred to as wireless Ethernet or wireless fidelity (Wi-Fi), have become very popular and deployed widely in many areas such as airport lounges, hotels, campuses, rail-way stations, and even in private homes. The prevalence of wireless LANs is now a standard feature for laptops, personal digital assistants (PDAs), video game consoles, and mobile phones. Reasons for this explosive popularity and rapid evolving are its simplicity, convenience, mobility, and most of all, high-speed access to the Internet. Meanwhile, a wide variety of applications and services, ranging from best-effort to real-time, are running over wireless LANs, and now wireless LANs have become a part of our everyday lives.

Although wireless LANs are simple and easy to use, they have certain drawbacks as a result of changes in today's Internet. The IEEE 802.11 protocol is designed to achieve per-station fairness (or station-based fairness), so that all stations accessing the wireless channel share the wireless bandwidth fairly. This works well when all stations in a wireless LAN are identical to each other. However, achieving per-station fairness is not always reasonable, especially when some of stations in the wireless LAN behave in different ways. This dissertation discusses such three key fairness issues occurred in IEEE 802.11 wireless LANs.

Organization of the dissertation

This dissertation is organized into following chapters:

1. Introduction
2. Per-flow fairness in single-rate wireless LANs
3. Per-flow fairness in QoS-oriented wireless LANs
4. Per-flow fairness in multi-rate wireless LANs
5. Conclusions

Chapter 1 clarifies the purpose of this research by providing background. It then overviews the IEEE 802.11 MAC (medium access control) protocol, IEEE 802.11e EDCA (enhanced distributed channel access) protocol, and two fairness concepts, max-min fairness and proportional fairness, which are used in subsequent chapters.

Chapter 2 addresses a fairness issue between uplink and downlink flows in single-rate IEEE 802.11-based wireless LANs, where uplink flows dominate over downlink flows in terms of bandwidth usage. In order to ameliorate max-min fairness, in Chapter 2, we present a window control scheme by modifying the random backoff mechanism in the IEEE 802.11 MAC protocol that employs CSMA/CA (carrier sense multiple access with collision avoidance) mechanism. In our scheme, APs (access points) dynamically control their minimum contention window size $CW_{\min}$, a parameter of the random backoff mechanism in CSMA/CA, in order to adjust the ratio of the total packet rate of downlink flows to the packet rate of an uplink flow. Our scheme is evaluated through numerous simulation experiments with UDP and TCP traffic flows, and results show that our scheme is valid and works very well.

Chapter 3 describes a per-flow fairness issue in IEEE 802.11e EDCA-based wireless LANs, where both real-time and best-effort traffic flows exist. Even though the IEEE 802.11e EDCA protocol differentiates flows with different QoS (quality of services) requirements, it does not differentiate stations, i.e., it gives the same access priority to respective access categories in all stations. As a result, a bundle of flows transmitted from an access category in an AP is treated in the same way as an individual flow transmitted from the same access category in wireless terminals, and this results in unfairness between uplink and downlink best-effort flows at the best-effort access category. Chapter 3 presents a dynamic contention window control scheme to ameliorate max-min fairness among best-effort flows, while guaranteeing QoS requirements for real-time flows. In our scheme, the minimum contention window size $CW_{\min}$ of the best-effort access category at APs is first determined based on the number of best-effort flows, in such a way that this unfairness is resolved. The minimum and maximum contention window sizes (i.e., $CW_{\min}$s and $CW_{\max}$s) for real-time traffic at APs are then determined so as to guarantee QoS requirements for these traffic.

Chapter 4 considers the so-called performance anomaly issue in multi-rate IEEE 802.11 wireless LANs, where stations with the lowest data transmission rate regulate the throughput of all other stations and it is forced to be the same as the throughput of stations with the lowest data transmission rate. As a result, the total system throughput degrades badly in a multi-rate wireless LAN. To alleviate the performance anomaly, in Chapter 4, we present a dynamic contention window control mechanism that works well for UDP and TCP flows. In our scheme, flows are classified into several classes according to their data transmission rates, and at APs, downlink flows in respective classes are stored in separate buffers, as in the IEEE 802.11e EDCA protocol. Further our scheme assigns different minimum contention window sizes $CW_{\min}$s to those classes according to their data transmission rates and target packet rates. Through simulation experiments, we show the effectiveness of our scheme even when there are many downlink flows at the AP.

Chapter 5 presents the conclusions of this dissertation by summarizing all results and observations we obtained through the research. Some future works and implementation issues of our schemes are also discussed in Chapter 5.

B. A. Hirantha Sithira Abeysekera

Osaka, Japan
January 2010

Acknowledgements

First and foremost, I must sincerely express my gratitude and indebtedness to my supervisor, Professor Tetsuya Takine of Graduate School of Engineering, Osaka University. He spent a lot of time on this research. His extensive comments, kind advice, and great encouragement were extremely beneficial in improving the technical content and compiling this dissertation.

Much gratitude is also due to Professor Seiichi Sampei of Graduate School of Engineering, Osaka University, who was one my committee members. He has given me many insightful suggestions and constructive comments on this work.

I am grateful to Associate Professor Takahiro Matsuda of Graduate School of Engineering, Osaka University, who was also one of my committee members. He helped me a lot in doing this work from the beginning. Without him, my knowledge would not be enough to do this research.

I also take pleasure in thanking Professor Shozo Komaki, Professor Zenichiro Kawasaki, Professor Kenichi Kitayama, Professor Noboru Babaguchi, and Professor Kyo Inoue of Graduate School of Engineering, Osaka University, and Professor Riichiro Mizoguchi and Professor Takashi Washio of the Institute of Scientific and Industrial Research, Osaka University. Their valuable discussions and favorable comments greatly enhanced the quality of this dissertation.

I am deeply grateful to the former technical officer of Takine Laboratory, Kayoko Goto for her encouragement and creditable advice. Assistant Professor Masahiro Sasabe, Ms. Sachiko Hashimoto, and all other members of Takine Laboratory are also acknowledged for their continuous support in many ways. Without their help I could not carry out so smoothly this research. Especially I wish to thank Ph.D candidate Habibul Kabir for his many enlightening discussions and comments throughout the research.

My sincere thanks are due to Assistant Professor Kouji Hirata of Graduate School of Science and Engineering, Ehime University. He has been always willing to hear about the progress of my research and has always encouraged me. I also wish to thank Dr. Yosuke Matsushita of Panasonic Corporation for his valuable comments and suggestions in the early stage of this research.

All my friends and colleagues who have helped with providing advice and constructive comments are also appreciated. In particular, I would like to thank Dr. Thilmee Malinda

Baduge of Panasonic Corporation who enthusiastically provided me with numerous suggestions and invaluable advice throughout this research.

My special appreciation goes to "Monbukagakusho", "Ito Foundation", "International Communications Foundation", and "Tokyu Foundation". Their kind assistance and funding support greatly help me to carry out my research. This research was also supported in part by "Global COE Program". I would like to thank all committee members of "Global COE Program".

I am also grateful to all of my host families for their love and encouragement. Especially, I wish to thank Mr. and Mrs. Maezawa for their excellent support.

Last, but not least, I heartily thank my parents, parents-in-law, sister, brothers, and my loving wife, Menaka for their emotional support, understanding, encouragement, and sacrifice throughout my studies.

Contents

List of Figures

List of Tables

Abbreviations and Acronyms

ACK	Acknowledgement
AP	Access Point
BE	Best-Effort
BK	Background
BSS	Basic Service Set
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
DCF	Distributed Coordination Function
EDCA	Enhanced Distributed Channel Access
ESS	Extended Service Set
FCS	Frame Check Sequence
HCCA	HCF Controlled Channel Access
HCF	Hybrid Coordination Function
IEEE	Institute of Electrical and Electronics Engineers
IFS	Interframe Space
IP	Internet Protocol
LAN	Local Area Network
MAC	Medium Access Control
MSDU	MAC Service Data Unit
P2P	Peer-to-Peer
PCF	Point Coordnation Function
PLCP	Physical Layer Convergence Protocol
QoS	Quality of Services
RSSI	Received Signal Strength Indication
RT	Real-Time
RTS	Request To Send
RTT	Round Trip Time

TCP	Transmission Control Protocol
TXOP	Transmission Opportunity
UDP	User Datagram Protocol
VoIP	Voice over IP
WT	Wireless Terminal

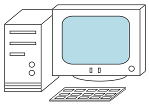
Symbols

AC_i	access category i $(i = 0, 1, 2, 3)$
$AIFS[i]$	arbitration interframe space of AC_i
$AIFSN[i]$	$AIFS[i]$ number
α	unsuccessful frame transmission probability
α_{D}	unsuccessful frame transmission probability of a downlink flow
$\alpha_{\mathrm{D},i}$	unsuccessful frame transmission probability of a class i downlink flow
α_{U}	unsuccessful frame transmission probability of an uplink flow
$\alpha_{\mathrm{U},i}$	unsuccessful frame transmission probability of a class i uplink flow
B_I	random backoff integer, $B_I = U[0, CW]$
CW	contention window
$CW_{\max}$	maximum contention window
$CW_{\max}[i]$	$CW_{\max}$ of AC_i
$CW_{\max\_\mathrm{D}}[i]$	$CW_{\max}[i]$ of downlink flows
$CW_{\min}$	minimum contention window
$CW_{\min}[i]$	$CW_{\min}$ of AC_i
$CW_{\min\_\mathrm{D}}$	$CW_{\min}$ of downlink flows
$CW^*_{\min\_\mathrm{D}}$	optimal $CW_{\min\_\mathrm{D}}$ obtained by analysis
$\widetilde{CW}^*_{\min\_\mathrm{D}}$	optimal $CW_{\min\_\mathrm{D}}$ obtained by simulation experiments
$CW_{\min\_\mathrm{D}}[i]$	$CW_{\min}[i]$ of downlink flows
$CW_{\min\_\mathrm{D},i}$	$CW_{\min}$ of class i downlink flows
$CW_{\min\_\mathrm{U}}$	$CW_{\min}$ of uplink flows
$CW_{\min\_\mathrm{U}}[i]$	$CW_{\min}[i]$ of uplink flows
$CW_{\min\_\mathrm{U},i}$	$CW_{\min}$ of class i uplink flows
D	downlink
$DIFS$	DCF interframe space
EI	efficiency index
FI	Jain's fairness index

$FI_{m,n}$	FI for m downlink flows and n uplink flows
Γ	uplink/downlink fairness index
Γ_{est}	estimated value of index Γ
Γ_{sim}	simulation results of index Γ
L	Length of a data frame
L_i	Length of a class i data frame
λ	access probability of a station
λ_{D}	access probability of a downlink flow
$\lambda_{\text{D},i}$	access probability of a class i downlink flow
λ_i	access probability of a station in class i
λ_{U}	access probability of an uplink flow
$\lambda_{\text{U},i}$	access probability of a class i uplink flow
$LongRetryLimit$	retransmission limitation of a long data frame
N	the number of flows, the number of wireless terminals
N_{D}	the number of downlink flows
$N_{\text{D},i}$	the number of class i downlink flows
$N_{\text{BE\_D}}$	the number of downlink best-effort flows
$N_{\text{BE\_U}}$	the number of uplink best-effort flows
N_{U}	the number of uplink flows
$N_{\text{U},i}$	the number of class i uplink flows
$N_{\text{VI\_D}}$	the number of downlink video flows
$N_{\text{VI\_U}}$	the number of uplink video flows
N_{VO}	the number of voice sessions
$\pi(i)$	steady state probability being in state i
$PIFS$	PCF interframe space
$P_{\text{c}}^{(\text{AP})}$	frame collision probability of the AP
$P_{\text{c}}^{(\text{WT})}$	frame collision probability of a wireless terminal
$P_{\text{loss}}^{(\text{AP})}$	packet loss probability at the AP
$P_{\text{loss}}^{(\text{Del})}$	packet loss ratio due to delay constraint
$P_{\text{loss}}^{(\text{Del\&Jit})}$	packet loss ratio due to both delay and jitter constraints
$P_{\text{loss}}^{(\text{Jit})}$	packet loss ratio due to jitter constraint
P_S	probability of a successful frame transmission
$P_{S\text{\_D}}$	P_S of a downlink flow
$P_{S\text{\_U}}$	P_S of an uplink flow
r	the number of retransmission attempts

$r_{\max}$	the number of maximum retransmission attempts
R	packet rate ratio
R^*	target packet rate ratio
R_{est}	resulting packet rate ratio, resulting throughput ratio
R_i	data transmission rate of a class i flow
$RTSThreshold$	threshold for RTS/CTS access scheme
$ShortRetryLimit$	retransmission limitation of a short data frame
$SIFS$	short interframe space
$SlotTime$	size of a time slot
$T^{(\mathrm{ACK})}$	frame transmission time of an ACK frame
$T_i^{(\mathrm{DATA})}$	frame transmission time of a class i data frame
Θ	throughput, the number of frames transmitted in a unit time
$\overline{\Theta}_{\mathrm{D}}$	average Θ of downlink flows
$\Theta_{\mathrm{D},i}$	Θ of ith downlink flow
$\overline{\Theta}_{\mathrm{U}}$	average Θ of uplink flows
$\Theta_{\mathrm{U},i}$	Θ of ith uplink flow
$TXOP[i]$	transmission opportunity of AC_i
U	uplink
Z_i	channel occupancy time of a class i flow
$Z_{\mathrm{D},i}$	channel occupancy time of a class i downlink flow
$Z_{\mathrm{U},i}$	channel occupancy time of a class i uplink flow

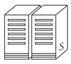
Icons



Personal Computer

Laptop

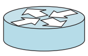
Game Console



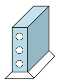
Server

Telephone

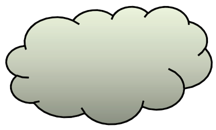
Mobile Phone



Router



Access Point



Network Cloud

Wired Link

Wireless Channel

Chapter 1

Introduction

In recent years, wireless technology has become an important component in providing networking infrastructure for data delivery. This wireless data revolution has been made possible by the introduction of new networking technologies and paradigms such as [15, 77]

- wireless PANs (personal area networks, e.g., Bluetooth [89] or ZigBee [58])
- wireless LANs (local area networks, e.g., Wi-Fi [82])
- wireless MANs (metropolitan area networks, e.g., WiMAX [53])
- wireless WANs (wide area networks, e.g., MBWA [54])

In particular, IEEE 802.11-based wireless LANs, sometimes referred to as wireless Ethernet or wireless fidelity (Wi-Fi), have become very popular and deployed widely in many areas such as airport lounges, hotels, campuses, rail-way stations, and even in private homes. The prevalence of wireless LANs is now a standard feature for laptops, personal digital assistants (PDAs), video game consoles, and mobile phones. Reasons for this explosive popularity and rapid evolving are its simplicity, convenience, mobility, and most of all, high-speed access to the Internet. Meanwhile, a wide variety of applications and services, ranging from best-effort to real-time, are running over wireless LANs, and now wireless LANs have become a part of our everyday lives.

Generally, a wireless LAN consists of two main components; wireless enabled devices, i.e., wireless terminals (WTs), and an access point (AP). The AP forms a bridge between wired and wireless networks, and as shown in Fig. 1.1, wireless terminals connect to the Internet via the AP. This type of networks is referred to as infrastructure networks. There is another type of networks called ad hoc networks, where there is no APs and wireless terminals communicate directly with each other on a peer-to-peer mode. This dissertation focuses on widely spread infrastructure networks.

Although wireless LANs are simple and easy to use, they have certain drawbacks as a result of changes in today's Internet. The IEEE 802.11 protocol is designed to achieve per-station fairness (or station-based fairness), so that all stations (note that in

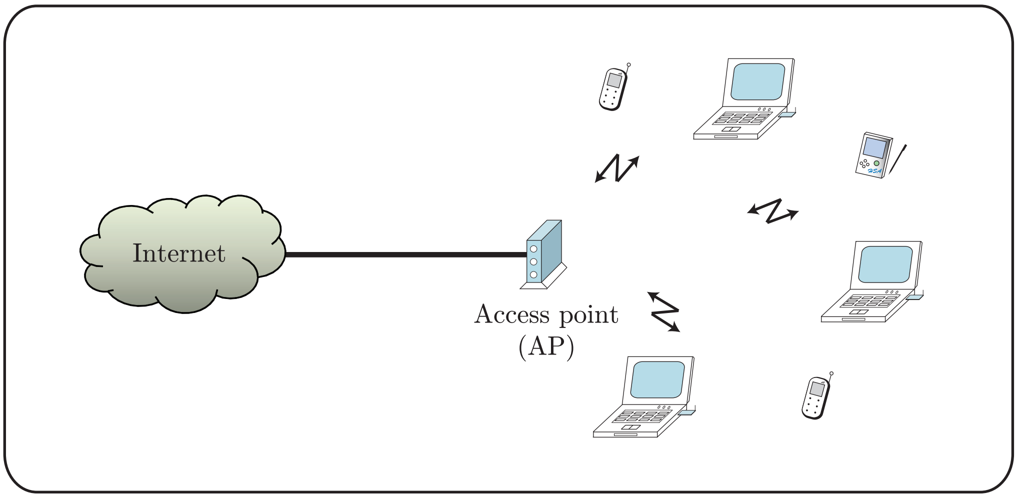


Figure 1.1: Wireless local area network (LAN).

this dissertation, we use term *station* to refer to any wireless device i.e., AP and wireless terminal accessing the wireless channel) accessing the wireless channel share the wireless bandwidth fairly. This works well when all stations in a wireless LAN are identical to each other. However, achieving per-station fairness is not always reasonable, especially when some of stations in the wireless LAN behave in different ways. This dissertation discusses such three key fairness issues occurred in IEEE 802.11 wireless LANs;

(i) per-flow fairness issue in single-rate wireless LANs,

(ii) per-flow fairness issue of best-effort flows in QoS-oriented wireless LANs, and

(iii) per-flow fairness issue under the proportional fairness in multi-rate wireless LANs.

We discuss these issues in detail in Chapters 2, 3, and 4, respectively. First of all, let us see how the IEEE 802.11 protocol works.

1.1 IEEE 802.11 Protocol

The IEEE 802.11 protocol [82] focuses on medium access control (MAC) and physical layers. The original standard supports three types of implementations at the physical layer; frequency hopping spread spectrum (FHSS), direct sequence spread spectrum (DSSS), and infrared (IR). FHSS and DSSS use radio frequencies on the 2.4 GHz industrial, scientific and medical (ISM) band, while IR uses infrared light. FHSS and IR implementations support 1 Mbps data transmission rate with an optional 2 Mbps extension, while DSSS supports both. Later, higher physical layer extensions such as IEEE 802.11a [83], IEEE 802.11b [84], and IEEE 802.11g [86] were standardized. The IEEE 802.11 MAC protocol defines:

- medium access schemes
- authentication, association, and re-association services
- encryption and decryption procedures
- power management techniques
- multiple data transmission rates

The IEEE 802.11 MAC protocol specifies two medium access schemes, DCF (distributed coordination function) and PCF (point coordination function). The latter is an optional mechanism operated on DCF with a point coordinator. This dissertation focuses on DCF which follows a CSMA/CA (carrier sense multiple access with collision avoidance) mechanism to reduce and resolve collisions of data frames.

1.1.1 CSMA/CA Mechanism

CSMA/CA is a listen before talk (LBT) mechanism and it is illustrated in Fig. 1.2. When a station has a data frame to transmit, it first senses the wireless channel for a *DIFS* (DCF interframe space) period given by

$$DIFS = SIFS + 2 \cdot SlotTime,$$

where *SIFS* and *SlotTime* denote the short interframe space and the duration of a time slot, respectively. When the station finds the wireless channel idle for a *DIFS* period, the random backoff procedure starts.

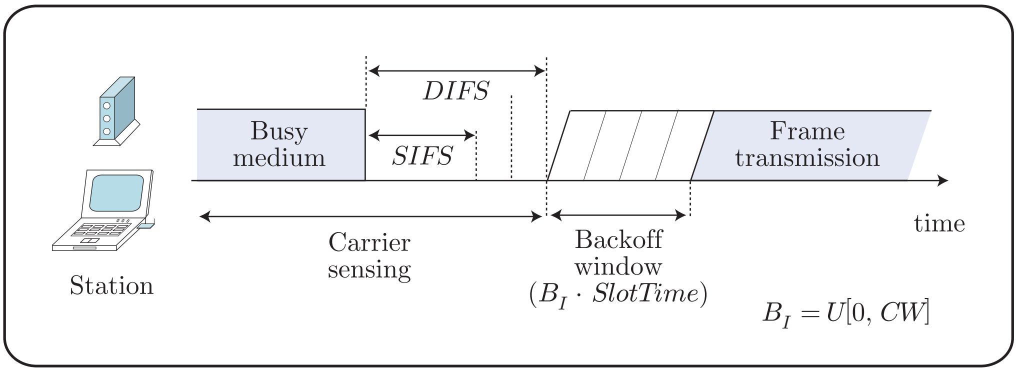


Figure 1.2: CSMA/CA mechanism.

The station generates a random backoff interval (also referred to as backoff window) as an additional deferral time before starting its frame transmission, and the backoff counter decreases while the wireless channel is idle. Note here that the backoff counter is frozen when the wireless channel is sensed busy, and it is reactivated when the wireless channel is sensed idle again for a *DIFS* period. When the backoff counter reaches zero, the station acquires the transmission right and starts the frame transmission. The IEEE 802.11 DCF defines two frame transmission schemes, basic scheme and RTS/CTS scheme (request-to-send/clear-to-send). We discuss these schemes in detail in Sections 1.1.2 and 1.1.3, respectively.

The random backoff interval is given by $B_I \cdot SlotTime$, where B_I is the initial value of the backoff counter and is a random integer determined by each station individually. The random integer B_I follows a uniform distribution on $[0, CW]$, where CW is referred to as contention window. Parameter CW is initially set to be its minimum value $CW_{\min}$ and for the rth ($r = 1, 2, \ldots, r_{\max}$) retransmission attempt, it is given by

$$CW = \min\left(2^r(CW_{\min} + 1) - 1, CW_{\max}\right), \tag{1.1}$$

where $r_{\max}$ denotes the maximum number of retransmission attempts of a data frame, and $CW_{\max}$ denotes the maximum value of CW. Figure 1.3 illustrates this process. If the retransmission of a data frame fails $r_{\max}$ times successively, the data frame is dropped and never transmitted again. When a data frame succeeds in transmission or it is dropped, CW is reset to be $CW_{\min}$.

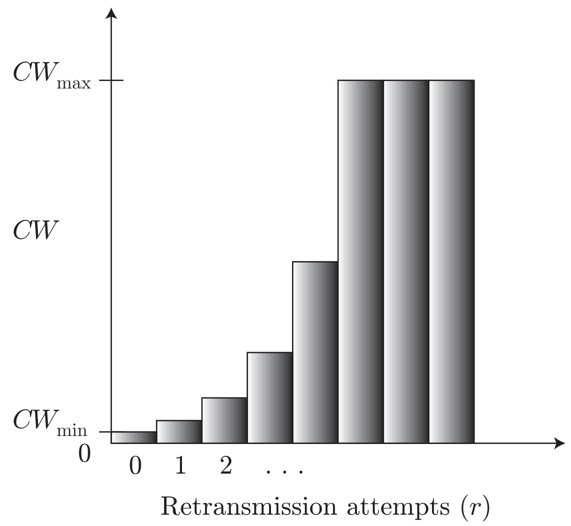


Figure 1.3: Random backoff procedure.

1.1.2 Basic Scheme

The IEEE 802.11 DCF basic scheme is a two-way handshaking mechanism and it is used for short data frames which are not longer than *RTSThreshold* defined in the IEEE 802.11 MAC protocol. As depicted in Fig. 1.4, in the basic scheme, the sender station (station A) starts to transmit the data frame immediately after it acquires the transmission right. The IEEE 802.11 MAC protocol utilizes a positive acknowledgments (ACKs) to ensure reliable transmission, and the receiver station (station B) sends an ACK frame a *SIFS* period after the correct receipt of the data frame. Note that *SIFS* is shorter than *DIFS*. If the sender station fails to receive the ACK frame, it tries to retransmit the data frame. Note that the maximum number of retransmission attempts $r_{\max}$ in the basic scheme is set to be *ShortRetryLimit* of 7.

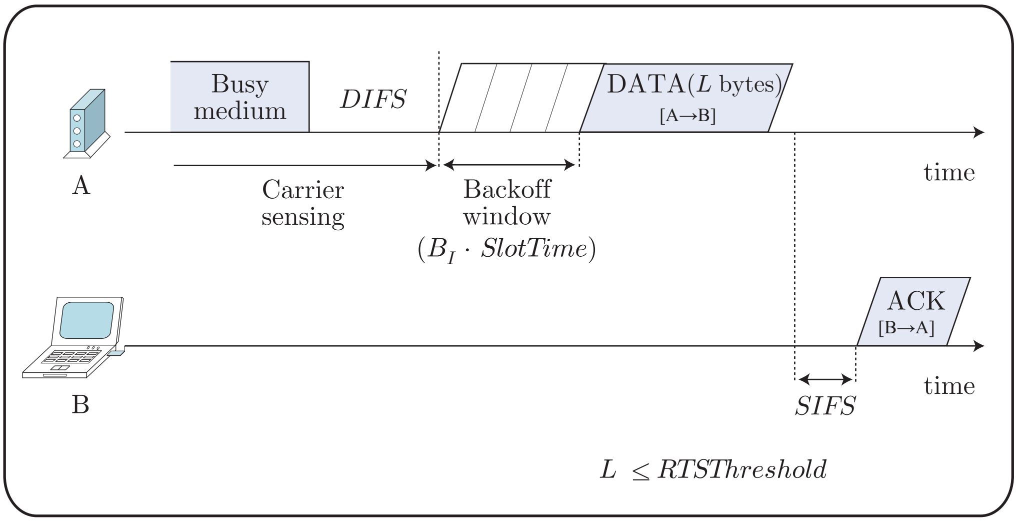


Figure 1.4: Basic scheme.

1.1.3 RTS/CTS Scheme

In addition to the basic scheme explained above, the IEEE 802.11 MAC protocol defines a four-way handshaking mechanism which is known as RTS/CTS scheme. In IEEE 802.11-based wireless LANs, long data frames which are longer than *RTSThreshold* are transmitted using RTS/CTS scheme. As illustrated in Fig. 1.5, in the RTS/CTS scheme, sender and receiver stations (stations A and B, respectively) exchange RTS and CTS control frames prior to the data frame transmission, in order to avoid collisions of long data frames and to resolve the hidden-terminal problem. Note that the interframe intervals between RTS, CTS, data, and ACK frames are all equal to *SIFS*. In the RTS/CTS scheme, the maximum number of retransmission attempts $r_{\max}$ is set to be *LongRetryLimit* of 4.

Table 1.1 summarizes the default MAC parameter values defined in the IEEE 802.11a and IEEE 802.11b physical layer extensions. Note that *SlotTime* and *SIFS* values in Table 1.1 are in μsec unit.

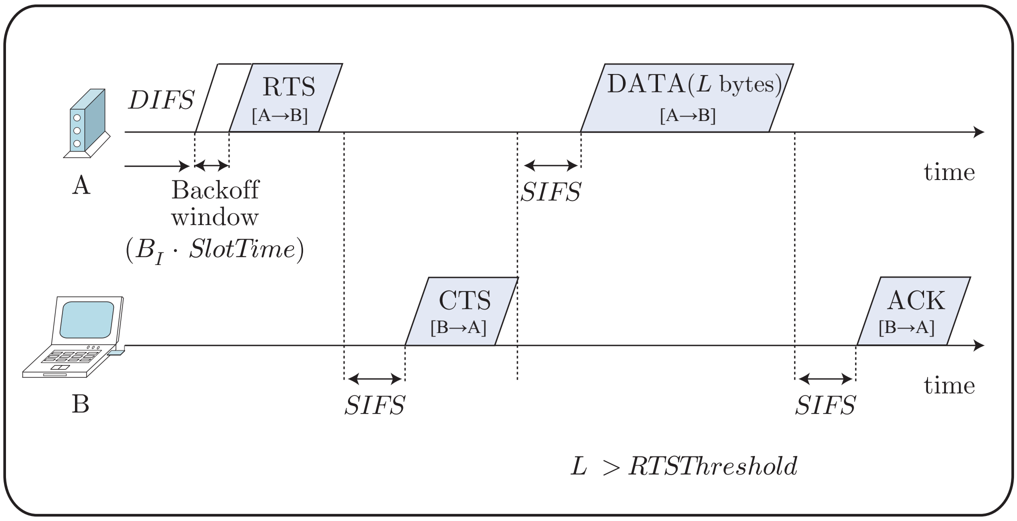


Figure 1.5: RTS/CTS scheme.

Table 1.1: Default MAC parameter values for different physical layer extensions.

Physical layer extension	*SlotTime*	*SIFS*	$CW_{\min}$	$CW_{\max}$
IEEE 802.11a	16	9	15	1023
IEEE 802.11b	10	20	31	1023

1.2 Features and Challenges of Wireless LANs

This section briefly overviews some key features and challenges of IEEE 802.11 wireless LANs. These include

- security
- roaming and handover
- quality of services (QoS)
- multi-rate transmission
- system performance

1.2.1 Security

Wireless LAN security problems have been widely publicized and have been a key barrier to take up. In a wireless LAN, the data frame is broadcast from the sender station in the hope that the receiver station is within its transmission range. The drawback to this mechanism is that any other station within this range also receives the data frame, and without a security mechanism of some sort, it can process the data. The IEEE 802.11 standard defines authentication mechanisms, including open system and shared key, and data encryption technique called wired equivalent privacy (WEP) to provide station authentication and data privacy, respectively. When the IEEE 802.11 draft standard was introduced in 1997, WEP was intended to provide confidentiality comparable to that of traditional wired networks [72].

Beginning in 2001, however, the vulnerability of WEP has been identified [7, 20]. Since WEP cannot provide strong link-layer level security, a revised version of WEP, known as WEP2 was proposed. After it became clear that overall WEP algorithm was deficient and would require even more fixes, both WEP and WEP2 were superseded by Wi-Fi protected access (WPA) which is a subset of the IEEE 802.11i [87] standard. The IEEE 802.11i enhancement defines advanced encryption standard (AES) and temporal key integrity protocol (TKIP), and AES and TKIP are considered to be so promising as a cure for wireless LANs security problems [10, 30, 60].

1.2.2 Roaming and Handover

There are three different handover scenarios in IEEE 802.11 wireless LANs [65].

(i) movement within the basic service set (BSS),

(ii) movement from one AP to another within the same extended service set (ESS), and

(iii) movement from a BSS in one ESS to a BSS in a different ESS.

Most of solutions provide handover for the first two scenarios; these can be done using link-layer. The last scenario requires the involvement of layer 3.

The way an application operates directly correlates to its resilience during the roaming process. Connection-oriented applications, such as TCP-based applications, are more tolerant to packet loss incurred during the roaming process because TCP is a reliable and connection oriented protocol. Data loss during the roaming process and handover delay, however, might cause a noticeable impact to UDP-based connectionless applications, such as voice over IP (VoIP) and video. In order to solve such issues and to provide secure access, several protocols, which enable fast roaming, have been proposed so far [13]. These include control and provisioning of wireless access points (CAPWAP) [27], handover keying (HOKEY) [14], and IEEE 802.11r [88].

1.2.3 QoS

The IEEE 802.11 DCF access scheme described in Section 1.1 provides best-effort traffic services. Real-time traffic services such as IP telephony (i.e., VoIP) and video conferencing, however, demand various QoS requirements, such as bandwidth guarantees, low delay, low jitter, and low packet loss rate. So far, numerous research efforts have been conducted to support QoS in IEEE 802.11-based wireless LANs.

1.2.3.1 QoS Schemes for IEEE 802.11 Wireless LANs

Most existing QoS mechanisms for IEEE 802.11 wireless LANs can be classified into three main categories [94].

(i) service differentiation schemes,

(ii) admission control and bandwidth reservation schemes, and

(iii) link adaptation schemes.

Service differentiation schemes include IEEE 802.11e EDCA [85], persistent factor DCF (P-DCF) [25], distributed weighted fair queue (DWFQ) [3, 4], distributed fair scheduling (DFS) [4, 80], and distributed deficit round robin (DDRR) [1]. Due to inefficiency of IEEE 802.11 protocol, service differentiation does not perform well under high traffic load conditions [51]. In such situations, admission control and bandwidth reservation become necessary in order to guarantee QoS of existing traffic flows. Admission control and bandwidth reservation schemes include virtual MAC (VMAC) [5], probe packet scheme [75], ARME [4], and AACA [52]. On the other hand, link adaptation mechanisms, which selects proper data transmission rate according to the wireless channel condition, include received signal strength (RSS) [61], packet error rate (PER) prediction [49], and link adaptation with success/fail (S/F) thresholds [11].

In the next subsection, we present overview of widely adopted IEEE 802.11e EDCA amendment which was standardized in 2005.

1.2.3.2 IEEE 802.11e EDCA Protocol

The IEEE 802.11e protocol defines a new coordination function called hybrid coordination function (HCF), which is composed of a contention-based channel access part and a centrally controlled channel access part. The contention-based channel access is referred to as enhanced distributed channel access (EDCA) and the centrally controlled channel access is referred to as HCF controlled channel access (HCCA).

In the IEEE 802.11e EDCA, traffic is classified into four access categories AC_i $(i = 0, 1, 2, 3)$ according to QoS requirements, and each of which follows a CSMA/CA mechanism explained in Section 1.1.1. These access categories are differentiated by means of different MAC parameter values. The carrier sensing time $AIFS[i]$ of AC_i is given by

$$AIFS[i] = SIFS + AIFSN[i] \cdot SlotTime,$$

where $AIFSN[i]$ denotes an integer greater than one. Minimum and maximum contention windows of AC_i are given by $CW_{\min}[i]$ and $CW_{\max}[i]$, respectively. The IEEE 802.11e EDCA is illustrated in Fig. 1.6.

With the default parameter value set, AC_i $(i = 0, 1, 2)$ is given more chances to access to the wireless channel than AC_j $(i + 1 \leq j \leq 3)$ (see Table 1.2). Thus, as shown in

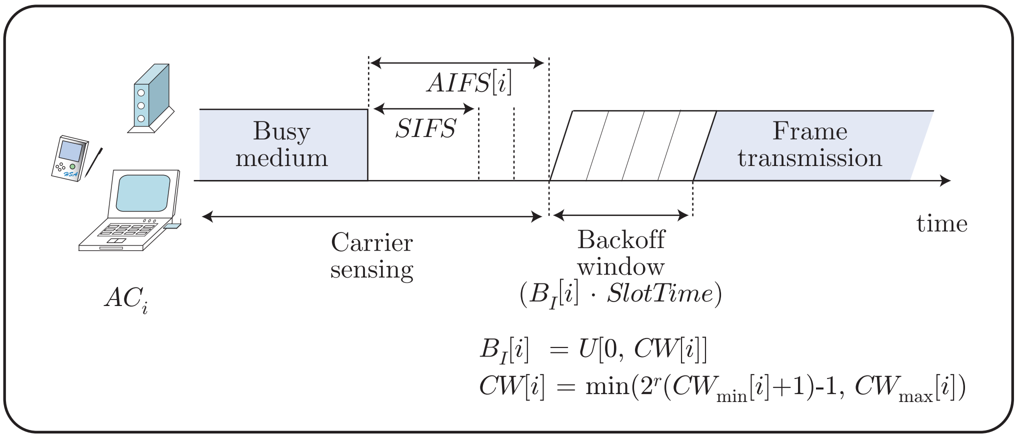


Figure 1.6: IEEE 802.11e EDCA access scheme.

Fig. 1.7, delay sensitive voice and video flows are mapped into AC_0 and AC_1, respectively, while best-effort (BE) flows are mapped into AC_2 and background (BK) flows are mapped into AC_3.

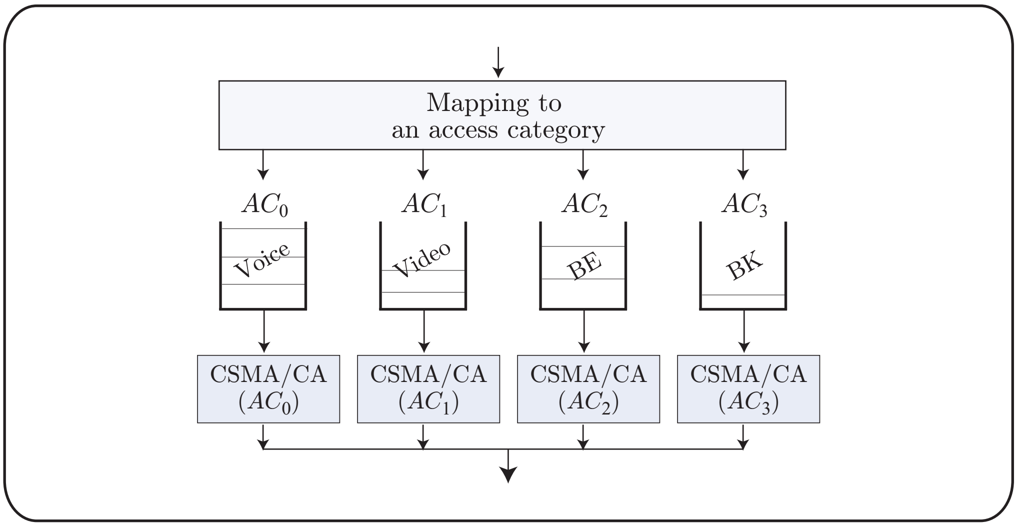


Figure 1.7: IEEE 802.11e EDCA access categories.

If backoff counters of more than one access category in a station expire at the same time slot, so-called an internal collision occurs. The internal collision is resolved within the station; the frame of the access category with the highest priority, involved in the internal collision, is transmitted and others behave as if they underwent an external frame collision on the wireless channel. As in the IEEE 802.11 DCF, once an AC_i acquires the transmission right, it starts to transmit the data frame using either basic scheme or RTS/CTS scheme, depending upon the length of the data frame.

Furthermore, as shown in Fig. 1.8, the IEEE 802.11e EDCA allows the AC_i to send multiple data frames within the duration of transmission opportunity $TXOP[i]$. Note that $TXOP[i]$ is given in msec unit, and if $TXOP[i] = 0$, only one data frame transmission is allowed. The default parameter values for the IEEE 802.11e EDCA protocol are summarized in Table 1.2.

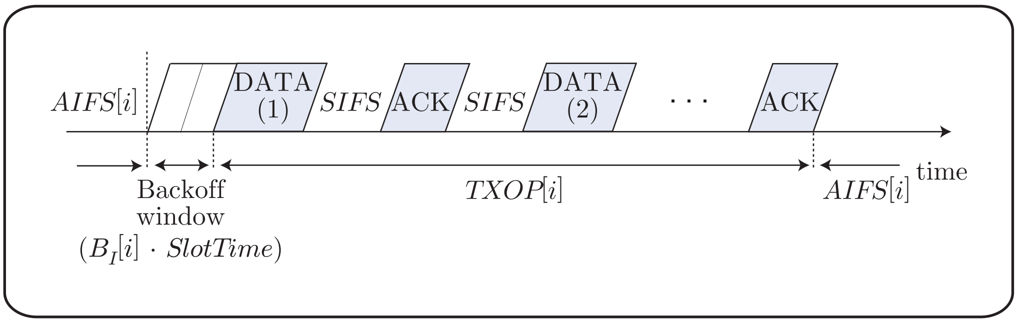


Figure 1.8: IEEE 802.11e EDCA transmission opportunity.

Table 1.2: IEEE 802.11e EDCA default parameter values.

Parameter	IEEE 802.11a				IEEE 802.11b			
	Access category, AC_i (Traffic type)				Access category, AC_i (Traffic type)			
	$i = 0$ (Voice)	$i = 1$ (Video)	$i = 2$ (BE)	$i = 3$ (BK)	$i = 0$ (Voice)	$i = 1$ (Video)	$i = 2$ (BE)	$i = 3$ (BK)
$CW_{\min}[i]$	3	7	15	15	7	15	31	31
$CW_{\max}[i]$	7	15	1023	1023	15	31	1023	1023
$AIFSN[i]$	2	2	3	7	2	2	3	7
$TXOP[i]$	1.504	3.008	0	0	3.264	6.016	0	0

1.2.4 Multi-Rate Transmission

In wireless systems, the radio propagation environment varies over time and space due to factors such as signal attenuations and fading, motion of objects, and interference which lead to variations in the received signal-to-noise ratio (SNR). A high level modulation can be used when the channel SNR is sufficiently high such that the received signal can be properly decoded.

Thus the IEEE 802.11 physical layer extensions were designed to support multiple data transmission rates by employing different modulations and channel coding schemes. Each station in a wireless LAN individually selects an appropriate data transmission rate *DataRate* according to its channel condition, and transmits data frames using the selected *DataRate*. Stations in a good channel condition typically employ a higher data transmission rate, while stations in a poor channel condition do a lower data transmission rate so as to prevent frequent frame losses due to transmission errors on the wireless channel. Note, however, that control frames such as RTS, CTS, and ACK frames are transmitted at the basic transmission rate *BasicRate* (e.g., 1 or 2 Mbps in IEEE 802.11b wireless LANs and 6, 12, or 24 Mbps in IEEE 802.11a wireless LANs) so as to be perceived by all stations in the wireless LAN. Supported *DataRate* values for different physical layer extensions are summarized in Table 1.3.

Table 1.3: Supported data rates for different physical layer extensions.

Physical layer extension	Supported *DataRate*s
IEEE 802.11a	6, 9, 12, 18, 24, 36, 48, and 54 Mbps
IEEE 802.11b	1, 2, 5.5, and 11 Mbps

So far, many rate adaptation algorithms have been proposed to switch the data transmission rate. These can be grouped into two categories [90]; (i) statistic-based such as ARF [41], LA-2 [69], and AARF/AMRR [48], and (ii) signal measurement-based such as RBAR [33], OAR [73], and RSS measurement [56].

The multi-rate mechanism with a dynamic rate switching is a way to improve the performance of individual stations. In IEEE 802.11-based wireless LANs, however, when at least one station uses a low data transmission rate, the overall system performance degrades badly [31,74]. It is because unlike time division multiple access (TDMA) systems, in wireless LANs, the access time of stations to the shared medium is not fixed, and thus stations with low data rates use much longer time to finish their data transmissions. As a result most of network resources are occupied by the stations with low data transmission rates. We discuss this issue in detail in Chapter 4.

1.2.5 System Performance

Over last few years, applications running over wireless LANs, traffic types passing through it, and users' demands and requirements have been changed dramatically. At the time the IEEE 802.11 task group finalized the IEEE 802.11 protocol, most of traffic was generated by client/server applications, and traffic was flowing from a server to clients. Thus the wireless bandwidth was dominated by downlink flows, i.e., flows from an AP to wireless terminals. With the appearance of so-called P2P (peer-to-peer) applications such as file sharing, however, both uplink and downlink flows now compete for the bandwidth, because they generate bi-directional traffic. Furthermore multimedia traffic such as VoIP and video are also now running over in wireless LANs, and as we discussed in Section 1.2.3, they need to be satisfied some QoS requirements. As we see later in this dissertation, in such cases the performance of the wireless LAN degrades badly because the legacy standard has not designed to support such situations.

These system performance issues can be classified into (i) fairness issues, (ii) efficiency issues, and (iii) load balancing issues. Usually, these kinds of performance issues do not appear until the wireless channel capacity is overloaded, i.e., saturated situation. However, with greedy applications with elastic traffic, such as file transfer with TCP, saturated situations frequently arise in wireless LANs. Note here that load balancing issues [23, 34, 93] occur where there exist more than one AP within their transmission ranges, i.e., multi-cell environment. On the other hand, fairness issues and efficiency issues arises even in single-cell wireless LANs due to the unexpected usages described above. As we noted in the beginning of this chapter, this dissertation focuses on such key fairness and efficiency issues occurred in IEEE 802.11-based wireless LANs. In the next section we present an brief overview of some fairness criteria which we use in subsequent chapters.

1.3 Fairness Concepts

The notion of fairness arises in the context of packet-switched networks carrying elastic traffic between node pairs, i.e., the traffic streams that can exhaust, perhaps within certain bounds, any bandwidth that is assigned to them [63]. The most important example of such traffic is the best-effort traffic carried under the TCP; another example is the available bit rate traffic in asynchronous transfer mode (ATM).

Consider a network with elastic traffic flows. The elasticity means that each traffic flow can consume any assigned aggregated bandwidth. A general problem with this type of networks is how to assign bandwidth to traffic flows so that the capacities of links are not exceeded and that the actual aggregated bandwidth volumes assigned to each flow are distributed in a fair way. One way to do this is to apply the well-known max-min fairness principle [6]; an alternative to max-min fairness is the concept of utility fairness. In next subsections we discuss these fairness criteria in detail.

1.3.1 Max-Min Fairness

The basic idea behind the max-min fairness is to first allocate equal network resources (i.e., bandwidth) to all contending flows. If a flow cannot utilize the given bandwidth, because of constraints elsewhere, then the residual bandwidth is distributed among others.

Figure 1.9 clarifies this notion. We assume a network with four elastic traffic flows. One flow, i.e., flow 3, goes through all shared links AB and BC, and each other flow does only one shared link. Capacity of each link is assumed to be C Mbps. In this network, it is plausible to limit flows 1, 2, and 3 to a rate of $C/3$ Mbps each, since this gives each of these flows as much rate as the others. It would be rather pointless, however, to restrict flow 4 to rate of $C/3$ Mbps. Flow 4 might better be limited to $2C/3$ Mbps, since any lower limit than $2C/3$ Mbps would waste some of the capacity of the rightmost BC link without benefiting flows 1, 2, or 3, and any higher limit than $2C/3$ Mbps would be unfair because it would further restrict flow 3.

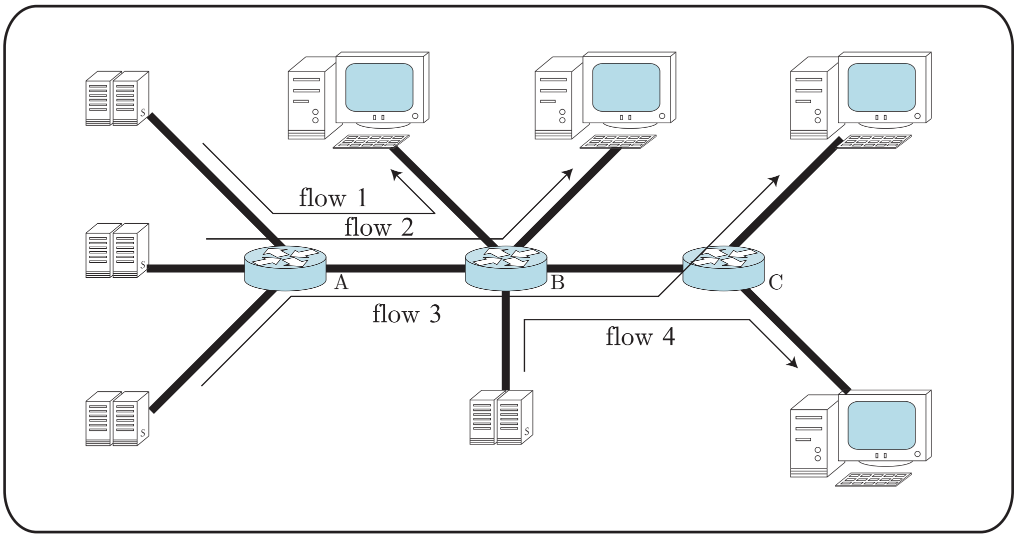


Figure 1.9: A network scenario with four elastic traffic flows.

The example in Fig. 1.9 leads to the idea of maximizing the network use allocated to the flows with minimum allocation, thus giving rise to the term *max-min*. After these most poorly treated flows are given the greatest possible allocation, there might be considerable latitude left for choosing allocations for other flows. It is then reasonable to maximize the allocation for the most poorly treated of these other flows, and so forth, until all allocations are specified. An alternative way to express this intuition is to maximize the allocation of each flow i subject to the constraint that an incremental increase in i's allocation does not cause a decrease in allocations of some other flows that is not greater than i's.

We assume a direct graph network with a set of traffic flows. Each flow i ($i = 1, 2, \cdots, N$) has an associated fixed path in the network. We denote by r_i the allocated rate for flow i (i.e., the throughput of flow i). The total allocated rates for all flows on link a of the network is then given by

$$F_a = \sum_{\text{all } i \text{ on } a} r_i$$

Letting C_a be the capacity of link a, we have the following constraints on the vector $\boldsymbol{r} = \{r_1, r_2, \cdots, r_N\}$ of allocated rates:

$$r_i \geq 0, \text{ for all } i, \text{and}$$
$$F_a \leq C_a, \text{ for all } a.$$

A vector $\boldsymbol{r}$ satisfying these constraints is said to be feasible. Furthermore, a vector $\boldsymbol{r}$ is said to be max-min fair if it is feasible and for each i, r_i cannot be increased while maintaining feasibility without decreasing r_j for some flow j for which $r_j \leq r_i$. It is known that there exists only one such solution when the resources of links and paths of all flows are both finite [63].

Given a feasible rate vector $\boldsymbol{r}$, we say that link a is a bottleneck link with respect to $\boldsymbol{r}$ for a flow i crossing a if $F_a = C_a$ and $r_i \geq r_j$ for all flows j crossing the link a. For example, bottleneck links of flow 1, 2, 3, and 4 in Fig. 1.9 are links AB, AB, AB, and BC, respectively. We now conclude with a preposition that a feasible vector $\boldsymbol{r}$ is said to be max-min fair if and only if each flow has a bottleneck link respect to $\boldsymbol{r}$. The proof of this preposition and some algorithms for computing max-min fair rate vectors can be found in [6].

We observe that the total system throughput in the scenario depicted in Fig. 1.9 under the max-min fairness criterion is only $5C/3$ Mbps. Since the achievable maximum system throughput is $2C$ Mbps, it is clear that, max-min fairness degrades the system throughput. The reason is that the same rate is assigned to every flow passing through the same bottleneck link (i.e., flows 1, 2, and 3) whatever the number of links on its path may be. In other words, when flows are not identical to each other, achieving max-min fairness degrades the system throughput.

Note that we can easily maximize the total system throughput by restricting rates of flows 1, 2, 3, and 4 to $C/2$, $C/2$, 0, and C Mbps, respectively. This is, however, highly unfair because some of flows fully occupy the link and achieve maximum of C Mbps throughput while some achieves nothing. Hence, a natural question arises whether there is some compromise solution between max-min fairness and throughput maximization that has better total system throughput than max-min fairness, yet is not as unfair as pure throughput maximization. The answer is yes, and such fair allocation principle is called utility fairness.

1.3.2 Utility Fairness

Utility fairness is often used as an alternative, a less egalitarian approach to max-min fairness [70]. It corresponds to the *utility* metric $\sum_i U(r_i)$, where r_i is the rate of flow i and U is a concave function called utility function. The concept of utility is a convenient way to represent user preferences, and a utility function U can be interpreted as a user satisfaction [57].

The properties of utility fairness depends on the choice of utility function U. The most often used class of utility functions is of form

$$U_\epsilon(r) = \begin{cases} \log r, & \text{if } \epsilon = 1, \\ (1-\epsilon)^{-1} r^{1-\epsilon}, & \text{otherwise,} \end{cases}$$

proposed in [59]. Note that when $\epsilon = 0$, the utility function $U_\epsilon(r)$ maximizes the throughput, while it achieves max-min fairness when $\epsilon = \infty$. Among all, the most widely used utility function is $U(r) = \log r$, i.e., $\epsilon = 1$. In such a case utility fairness is called *proportional fairness* [43].

1.3.3 Proportional Fairness

The idea behind the proportional fairness is to maximize the overall performance while giving at least some amount of rate to each flow [43]. The proportional fairness principle uses the revenue objective which consists in maximizing the sum of natural logarithms of the rates assigned to flows. The use of logarithmic function, instead of other functions, makes it impossible to assign zero rate to any flow (because $\log 0 = -\infty$), and at the same time, makes it not profitable to assign too much rate to any individual flow (because the derivative of $\log r$, i.e., $1/r$, rapidly decreases with the increase of r).

A vector $\boldsymbol{r} = \{r_1, r_2, \cdots, r_N\}$ of allocated rates is said to be proportionally fair if it is feasible and it maximizes the function

$$f = \sum_{\text{all } i} \log r_i.$$

In other words, for any other feasible vector $\boldsymbol{r^*} = \{{r_1}^*, {r_2}^*, \cdots, {r_N}^*\}$,

$$\sum_{\text{all } i} \frac{{r_i}^* - r_i}{r_i} \leq 0.$$

Thus with the proportional fairness, a worse treated flow may see its utility decreased if this allows a large enough increase to an already better treated flow. Again, in the case of finitely many links and paths, the vector $\boldsymbol{r}$ of proportionally fair rate shares is unique.

Let us return to the example in Fig. 1.9. We can determine the solution for the proportional fairness criterion for the scenario shown in Fig. 1.9 by maximizing the function $f(x) = \log{(1-x)/2} + \log{(1-x)/2} + \log x + \log{(1-x)}$. Thus we obtain $3C/8$, $3C/8$, $C/4$, and $3C/4$ Mbps as the rates of flows 1, 2, 3, and 4, respectively.

From a user's point of view, the proportional fairness solution is less fair than the max-min fairness; the throughput of the long hop flow, i.e., flow 3, is smaller than that of short flows. Because of favoring short flows, however, the proportional fairness allocation is more efficient in terms of total system throughput; in this case the total system throughput becomes $7C/4$ Mbps. In other words, proportional fairness solution does better than the max-min fairness solution in terms of total system throughput, at the expense of fairness given to end users.

A comparison of throughput of each flow and total system throughput under different fairness criteria for the example in Fig. 1.9 are given in Table 1.4. A well-known fairness index FI [37] ($0 < FI \leq 1$) and an efficiency index EI [70] ($0 \leq EI \leq 1$) are also compared in Table 1.4. Note that,

$$FI = \frac{\left(\sum_{i=1}^{N} r_i\right)^2}{N \sum_{i=1}^{N} {r_i}^2}, \qquad EI = \frac{\sum_{i=1}^{N} r_i}{\max \sum_{i=1}^{N} r_i},$$

where r_i and N denote the throughput of flow i and the total number of flows exist in the network, respectively.

Table 1.4: Throughput (Mbps), FI, and EI under different fairness criteria.

	Max-min fairness ($\epsilon = \infty$)	Throughput maximization ($\epsilon = 0$)	Proportional fairness ($\epsilon = 1$)
flow-1	$C/3$	$C/2$	$3C/8$
flow-2	$C/3$	$C/2$	$3C/8$
flow-3	$C/3$	0	$C/4$
flow-4	$2C/3$	C	$3C/4$
total	$5C/3$	$2C$	$7C/4$
FI	25/28	2/3	49/58
EI	5/6	1	7/8

1.4 Fairness in Wireless LANs of Wired-Cum-Wireless Networks

In this section we discuss fairness among flows in a wired-cum-wireless network where the wireless portion is the IEEE 802.11 wireless LAN. As noted in Section 1.1, the IEEE 802.11 MAC protocol is designed to achieve per-station fairness, and in the link-layer level, stations are identified by their MAC addresses. Thus when we focus on the link-layer, a flow can be defined as a sequence of data frames with the same sender and receiver MAC addresses. With this definition, we can assume that each wireless terminal holds only one flow at most in link-layer level, because wireless terminals transmit data frames only to AP, in infrastructure wireless LANs.

Up until 10 to 15 years ago, ISDN (integrated digital service network) was the most common way to connect to the Internet from private homes. ISDN typically provided a maximum of 64 kbps data rate in both uplink and downlink. With the emerge of broadband technologies, such as cable modem, xDSL (digital subscriber line), and FTTH (fiber to the home), the link capacity of wired networks has been increased rapidly. It is now common to see wired connections with bandwidth of more than 100 Mbps.

On the other hand, channel capacity of wireless LANs also has been increased in recent years. At the time the IEEE 802.11 protocol was standardized, the maximum data transmission rate was 2 Mbps. In the development of IEEE 802.11b protocol, the maximum data transmission rate was increased to 11 Mbps and then to 54 Mbps with the introduction of IEEE 802.11a and IEEE 802.11g protocols. Achievable effective throughput in a wireless LAN, however, highly depends upon the channel condition of the wireless channel, and it is not higher than that of typical wired LANs. Thus wireless LANs are very likely to become the bottleneck link along the end-to-end path of a traffic flow in wired-cum-wireless LANs, and it is worthwhile to consider fairness in such scenarios.

Thus we now discuss the fairness in a wireless LAN of a wired-cum-wireless network assuming that the wireless LAN is the bottleneck link of all flows going through it. Obviously, we can discuss the fairness in such situations in the same manner as we did in wired networks. There is, however, one difference; wired links are generally configured as full-duplexed, while wireless LANs are half-duplexed. Thus when we discuss fairness in a wireless LAN, we should take into the account of both uplink and downlink flows that share the wireless channel, because in a half-duplexed medium, only one direction flow can be utilized at a time. Since the performance of single- and multi-rate wireless LANs are different, in Section 1.4.1, we first discuss fairness in single-rate wireless LANs, where all flows are identical to each other. In Section 1.4.2, we then discuss fairness in multi-rate wireless LANs, where flows are not identical to each other.

1.4.1 Fairness in Single-Rate Wireless LANs

We consider a wired-cum-wireless network with a single-rate wireless LAN as shown in Fig. 1.10. The wireless LAN consists of one AP and three wireless terminals WT_1, WT_2, and WT_3. WT_1 transmits data frames towards the AP, i.e., WT_1 holds an uplink flow, while others receive data frames from the AP, i.e., downlink flows. We assume that bandwidths of wired links are large enough and the wireless LAN is the bottleneck link of all three flows. The channel capacity of the wireless LAN is assumed to be C Mbps.

As we discussed in Section 1.1, IEEE 802.11-based wireless LANs are designed to achieve per-station fairness in link-layer, i.e., layer 2. Thus all frame transmitting stations achieve the same access chance, i.e., transmission right, on average, and share the wireless bandwidth fairly among these stations. Since there are two stations, WT_1 and AP, transmitting frames in the scenario shown in Fig. 1.10, each of them achieves throughput of $C/2$ Mbps, i.e., the throughput of flow 1 is equal to $C/2$ Mbps. However, the AP handles two flows, flow 2 and flow 3, and thus their throughput is limited to $C/4$ Mbps.

It is obvious that the throughput of downlink flows degrades and the extent of the

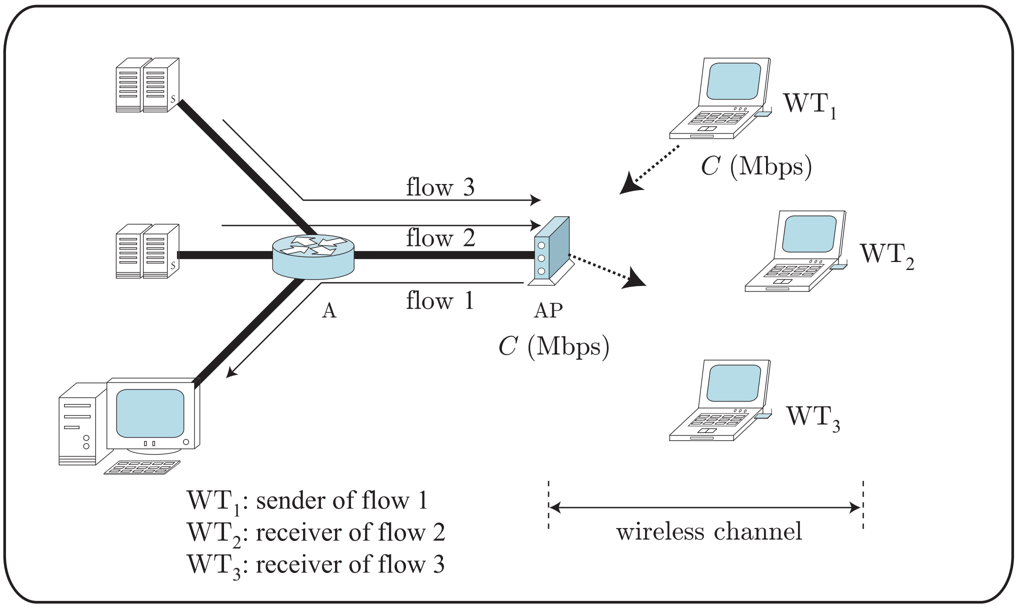


Figure 1.10: Wired-cum-wireless LAN with identical flows (a single-rate wireless LAN).

unfairness between uplink and downlink flows increases with the number of downlink flows that the AP handles. Generally, when the AP handles N_D downlink flows, the throughput of an uplink flow is N_D times as large as that of a downlink flow, under the per-station fairness property. Thus whenever we consider fairness among flows in a wireless LAN where there exist both uplink and downlink flows, achieving per-station fairness in link-layer level is not a suitable or a acceptable solution. Note, however, that if all wireless terminals in a single-rate wireless LAN are receivers, or all are senders, per-station fairness is reasonable, because it will result in max-min fairness.

Since wireless LANs are half-duplexed, the wireless channel can be illustrated as shown in Fig. 1.11, when both uplink and downlink flows share the wireless channel. If this wireless channel is the bottleneck link for all three flows, we can achieve max-min fairness by limiting all flows to rate of $C/3$ Mbps. Here, a general question arises: How to control the rate of each flow? Since the transmission right of each flow is determined at the link-layer under the CSMA/CA mechanism, we can control the rate of each flow by controlling the access chances of each station. For the scenario in Fig. 1.10, we can achieve such max-min fairness by giving twice as transmission rights to the AP as the contending wireless terminal WT_1.

1.4.2 Fairness in Multi-Rate Wireless LANs

We now think about the fairness in a multi-rate wireless LAN illustrated in Fig. 1.12, where we assume that different flows with different data transmission rates are contending

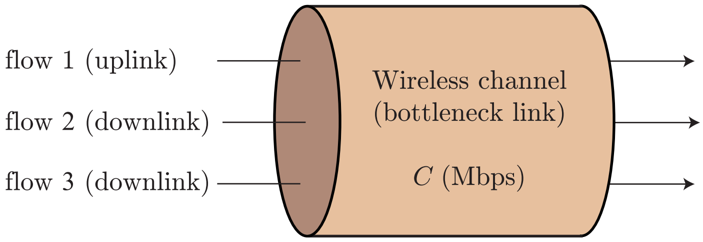


Figure 1.11: Half-duplexed wireless channel.

for the wireless channel which is the bottleneck link for all flows. In this scenario, we assume that two wireless terminals, WT_1 and WT_2 transmits data frames towards the AP. We assume that WT_1 is placed closer to the AP and thus it transmits data frames at transmission rate of C Mbps, while WT_2 is placed far away from the AP and it transmits data frames at transmission rate of $C/2$ Mbps. Thus in transmitting a data frame (for the simplicity, frame lengths of data frames are assumed to be 1 Mbit), WT_1 occupies the wireless channel for $1/C$ sec while WT_2 does for $2/C$ sec. With the max-min fairness criterion each wireless terminal transmits one data frame, on average, in $1/C + 2/C$ sec time period, and hence the throughput of each flow becomes $C/3$ Mbps. Thus the total system throughput is limited to $2C/3$ Mbps.

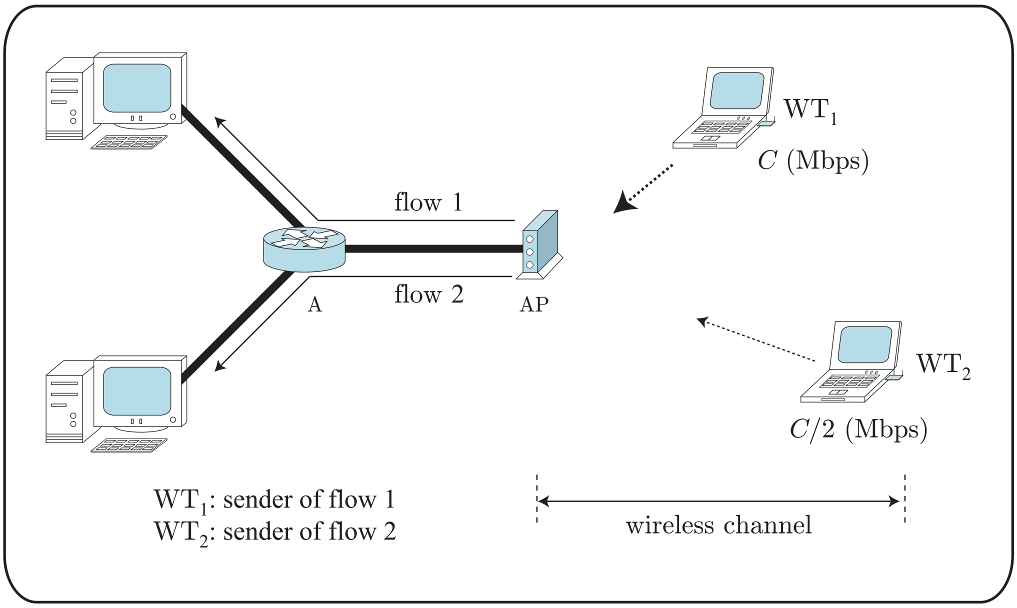


Figure 1.12: Wired-cum-wireless LAN with different flows (a multi-rate wireless LAN).

Behavior of the above wireless LAN with flows of high and low data transmission rates is similar to a wired network with short and long hop flows. Throughput fairness among these flows can be achieved under the max-min fairness criterion, but it degrades the total system throughput. On the other hand, the total system throughput can be easily maximized by shutting off the flow with low data transmission rate. As we discussed in Section 1.3.3, a compromise is required between the total system throughput (i.e., network resource usage) and fairness among flows.

We then consider achieving proportional fairness in a multi-rate wireless LAN of a wired-cum-wireless network. Note that the wireless channel is the bottleneck link for all flows and thus it is saturated. We assume that WT_1 in Fig. 1.12 utilizes the wireless channel for t sec in a unit time. Throughput of WT_1 is then given by Ct Mbps. On the other hand, throughput of WT_2 is given by $C(1-t)/2$ Mbps, because WT_2 transmits data frames at transmission rate of $C/2$ Mbps for a time period of $1-t$ sec in a unit time. Thus we can obtain the proportional fairness by maximizing the function $f(t) = \log Ct + \log C(1-t)/2$. The solution can be determined as $t = 1/2$, and this means that, when flows are proportionally fair, they share the bottleneck link fairly in terms of the channel utilization time. In other words, in wireless LANs, the proportional fairness is equivalent to the air-time fairness, i.e., fairness in channel occupancy time [38].

As shown in [38], we can obtain above observation even for a general case. We assume a saturated wireless LAN with N frame transmitting stations. Let T_i be the total amount of air-time used by station i ($i = 1, 2, \cdots, N$) measured over a very long period. The fraction of air-time t_i used by station i is then given by $t_i = T_i / \sum_{i=1}^{N} T_i$. Let R_i be the data transmission rate of station i. Then the throughput Θ_i of station i can be calculated as $\Theta_i = R_i T_i / \sum_{i=1}^{N} T_i$.

As explained in Section 1.3.3, the proportional fairness is achieved by maximizing $\sum_{i=1}^{N} \log \Theta_i$, and it is equivalent to maximizing $\prod_{i=1}^{N} \Theta_i$. Note here that

$$\begin{aligned}
\prod_{i=1}^{N} \Theta_i &= \prod_{i=1}^{N} \frac{R_i T_i}{\displaystyle\sum_{i=1}^{N} T_i}, \\
&= \prod_{i=1}^{N} R_i \prod_{i=1}^{N} \frac{T_i}{\displaystyle\sum_{i=1}^{N} T_i}.
\end{aligned}$$

Since R_is are constants for the optimization problem, maximizing $\prod_{i=1}^{N} \Theta_i$ is equivalent to maximizing $\prod_{i=1}^{N} \{T_i / \sum_{i=1}^{N} T_i\}$. As $\sum_{i=1}^{N} \{T_i / \sum_{i=1}^{N} T_i\} = 1$, $\prod_{i=1}^{N} \Theta_i$ is maximized when

$$T_1 = T_2 = \cdots = T_N.$$

Furthermore, because $t_1 = t_2 = \cdots = 1/N$, the resulting throughput Θ_i of station i under the proportional fairness can be determined as

$$\Theta_i = \frac{R_i}{N}.$$

We now summarize the properties of the proportional fairness in wireless LANs as follows:

(i) the wireless channel is occupied equally by all competing stations, and

(ii) given a fixed number of stations, the throughput of each station depends only on its data transmission rate and it does not vary according to the data transmission rates of others.

This means that, under the proportional fairness, the throughput of a station is equal to the throughput that the station would achieve in a single-rate wireless LAN in which all competing stations are running at its data transmission rate. Thus we believe that achieving proportional fairness is an adequate compromise in multi-rate wireless LANs [70].

Returning to our example in Fig. 1.12, if our objective is to achieve proportional fairness, we should give transmission rights to WT_1 twice as much as to WT_2 so that average channel utilization time of both wireless terminals to be the same. In this case, the throughput of WT_1 and WT_2 becomes $2/(1/C + 1/C + 2/C) = C/2$ Mbps and $1/(1/C+1/C+2/C) = C/4$ Mbps, respectively, and the total system throughput increases to $3C/4$ Mbps. We see that, with the proportional fairness, each flow in the scenario in Fig. 1.12 has to satisfy with a throughput of 1/2 of their data transmission rate.

In Table 1.5, we summarize the throughput of each flow, total system throughput, *FI*, and *EI* under different three fairness criteria for examples in Figs. 1.10 and 1.12.

Table 1.5: Throughput (Mbps), *FI*, and *EI* under three fairness criteria.

	Scenario in Fig. 1.10		Scenario in Fig. 1.12	
	Per-station fairness	Max-min fairness	Max-min fairness	Proportional fairness
flow-1	$C/2$	$C/3$	$C/3$	$C/2$
flow-2	$C/4$	$C/3$	$C/3$	$C/4$
flow-3	$C/4$	$C/3$	-	-
total	C	C	$2C/3$	$3C/4$
FI	8/9	1	1	9/10
EI	1	1	2/3	3/4

1.5 Conclusion

This chapter first discussed the research background and then provided an introduction to the IEEE 802.11 protocol. We then discussed some features and challenges of the IEEE 802.11-based wireless LANs. In this chapter, we also presented an overview of the IEEE 802.11e EDCA protocol, which was standardized to support QoS in wireless LANs.

Furthermore we presented an overview of some fairness concepts such as max-min fairness and proportional fairness, which we use in subsequent chapters. At the end of this chapter we discussed the fairness in a wireless LAN of a wired-cum-wireless network, in which the wireless LAN is the bottleneck link along the path of all flows. We showed that the max-min fairness criterion works well in single-rate wireless LANs, where all flows are identical to each other. On the other hand, the proportional fairness is reasonable for a multi-rate wireless LANs, where some of flows are different to others.

Chapter 2

Per-Flow Fairness in Single-Rate Wireless LANs

THIS chapter discusses the per-flow fairness in single-rate wireless LANs where the wireless channel is assumed to be the bottleneck link for all flows passing through it. As we discussed briefly in Chapter 1, in typical client/server applications, data traffic is asymmetric, i.e., it flows from a server to a client. In wireless LANs, this feature implies that downlink flows, i.e., flows from the AP to wireless terminals, dominate the bandwidth of the wireless channel. On the other hand, in P2P applications, data traffic is bi-directional and therefore uplink and downlink flows compete for the bandwidth. Recent studies show that when both uplink and downlink flows exist in IEEE 802.11-based wireless LANs, uplink flows attain significantly greater throughput than the competing downlink flows [22, 26, 46, 50, 62].

As we see later in detail, this problem occurs due to the per-station fairness property of the legacy IEEE 802.11 MAC protocol. To ameliorate fairness among flows in a wireless LAN with both uplink and downlink flows, in this chapter, we consider a simple yet highly effective scheme by modifying the random backoff mechanism in the IEEE 802.11 MAC protocol. In our scheme, APs dynamically control their minimum contention window size $CW_{\min}$, a parameter of the random backoff mechanism, in order to adjust the ratio R of the total packet rate of downlink flows to the packet rate of an uplink flow. Our scheme has the following features. First of all, no modification is required at wireless terminals. Furthermore, the optimal $CW_{\min}$ at APs is given by an explicit function of R, regardless of the number of uplink flows. Thus APs can easily compute the optimal $CW_{\min}$ to achieve fairness between uplink and downlink flows. Through simulation experiments with UDP and TCP flows, we show that our scheme can achieve fairness.

The rest of this chapter is organized as follows. Section 2.1 overviews the fairness issue between uplink and downlink flows. Section 2.2 presents related work. Then Section 2.3 describes our dynamic contention window control scheme. In Section 2.4, our scheme is evaluated with simulation experiments. Finally, the conclusion is given in Section 2.5.

2.1 Problem Overview

The IEEE 802.11 MAC protocol is designed to achieve per-station fairness, so that all stations accessing the wireless channel share the bandwidth fairly. Therefore APs access the wireless channel with the same authority as wireless terminals in the IEEE 802.11-based wireless LANs, even though APs aggregate several downlink flows. It is clear that this feature leads to serious unfairness between uplink and downlink flows. For example, assume that an AP aggregates N_{D} downlink flows and it shares the wireless channel with N_{U} wireless terminals. In the saturated situation, the available bandwidth B_W is shared equally among stations. Therefore the throughput of a station, i.e., the throughput of an uplink flow is equal to $B_W/(N_{\mathrm{U}}+1)$, while the throughput of a downlink flow is limited to $B_W/(N_{\mathrm{D}}(N_{\mathrm{U}}+1))$. Thus, the throughput of an uplink flow is N_{D} times as large as that of a downlink flow.

Unfairness between uplink and downlink flows also emerges even when a single downlink TCP flow competes with uplink TCP flows for the wireless channel. Since the destination nodes of uplink flows send back TCP ACK segments, the AP handles both TCP data segments of the downlink flow and TCP ACK segments of uplink flows, and some of them may get lost at the AP due to buffer overflow. Generally, a data segment loss reduces the TCP transmission rate. On the other hand, the loss of TCP ACK segments does not strongly affect the TCP transmission rate because TCP employs the cumulative acknowledgment [78]. Therefore buffer overflow at the AP has a greater impact on the downlink TCP flow than uplink TCP flows.

2.2 Related Work

So far, several papers have addressed the fairness issue between uplink and downlink flows. Pilosof et al. first revealed this issue and proposed adjusting the advertised window size in the header of TCP ACK segments [62]. This scheme, however, is applicable only to TCP flows because it uses the underlying flow control mechanism of TCP.

Gopalakrishnan et al. proposed a packet aggregation/fragmentation scheme to alleviate unfairness in IEEE 802.11b-based wireless LANs [26]. In their scheme, the AP first aggregates multiple packets destined for multiple wireless terminals into a large MSDU. The MAC layer then divides the MSDU into smaller fragments. Once the first fragment acquires the transmission right, the MAC protocol allows the AP to transmit the subsequent fragments without performing the random backoff procedure. Note that this scheme works well only for small packets because the maximum MSDU length is specified to be 2304 bytes in the IEEE 802.11 standard [82].

Leith et al. proposed a solution of unfairness between uplink and downlink TCP flows in IEEE 802.11e-based wireless LANs [50]. In their scheme, downlink TCP data segments and TCP ACK segments of uplink flows are classified into different access categories, and downlink TCP data segments are transmitted with a longer TXOP than other flows. The length of TXOP for downlink TCP data segments is dynamically controlled based on the number of downlink flows.

The common feature in [26, 50, 62] is that in acquiring the transmission right, all stations follow the CSMA/CA mechanism with the same parameters. There exists another way to alleviate unfairness between uplink and downlink flows. Fukuda et al. and Kim et al. independently proposed a static scheme to provide an ample opportunity for APs to acquire the transmission right [22, 46]. In the default IEEE 802.11 MAC protocol, all stations perform carrier sensing for a *DIFS* period before transmitting data frames. In [22, 46], APs perform carrier sensing only for a *PIFS* period, where $\mathit{PIFS} = \mathit{SIFS} + \mathit{SlotTime}$. Since $\mathit{PIFS} < \mathit{DIFS}$, their scheme gives higher priority to APs, comparing to wireless terminals, in accessing the wireless channel. Note however that this scheme lacks flexibility in response to the number of flows because *PIFS* is a fixed parameter in the IEEE 802.11 MAC protocol.

2.3 Dynamic Contention Window Control

We aim at achieving fairness between uplink and downlink flows by adjusting the ratio R of the total packet rate of downlink flows to the packet rate of an uplink flow, where the packet rate is defined as the mean number of successful transmissions in the unit time. To do so, we dynamically control the minimum contention window size $CW_{\min}$ at APs, where all wireless terminals are assumed to follow the default IEEE 802.11 MAC protocol. In a saturated situation, the packet rate ratio R between uplink and downlink flows is considered as a function of the numbers of uplink and downlink flows, as well as the minimum contention window sizes $CW_{\min}$ at the AP and wireless terminals. In Section 2.3.1, we first provide a simplified analysis of the packet rate ratio R, whose result suggests that the packet rate ratio R does not strongly depend on the number N_{U} of uplink flows. In Section 2.3.2, we conduct a mean field approximation analysis for saturated wireless LANs. We then derive an explicit formula for the optimal $CW_{\min}$ at the AP in Section 2.3.3, assuming the number of uplink flows is one, i.e., $N_{\mathrm{U}} = 1$, and we consider using it as a quasi-optimal $CW_{\min}$ at the AP for all $N_{\mathrm{U}} \geq 1$.

2.3.1 Fair-Share between Uplink and Downlink Flows

Consider a wireless LAN with N_{D} downlink flows and N_{U} uplink flows. A flow is defined in terms of source and destination MAC addresses (i.e., in this wireless LAN, N_{U} wireless terminals and one AP share the bandwidth). The wireless LAN model is depicted in Fig. 2.1. We assume that all flows are identical to each other. We also assume that wireless channel is error free, i.e., frame transmissions do not fail due to transmission errors. We observe the system only in contention periods, during which stations with positive backoff counters can decrease those values. For simplicity, we assume that stations are synchronized and time is divided into slots.

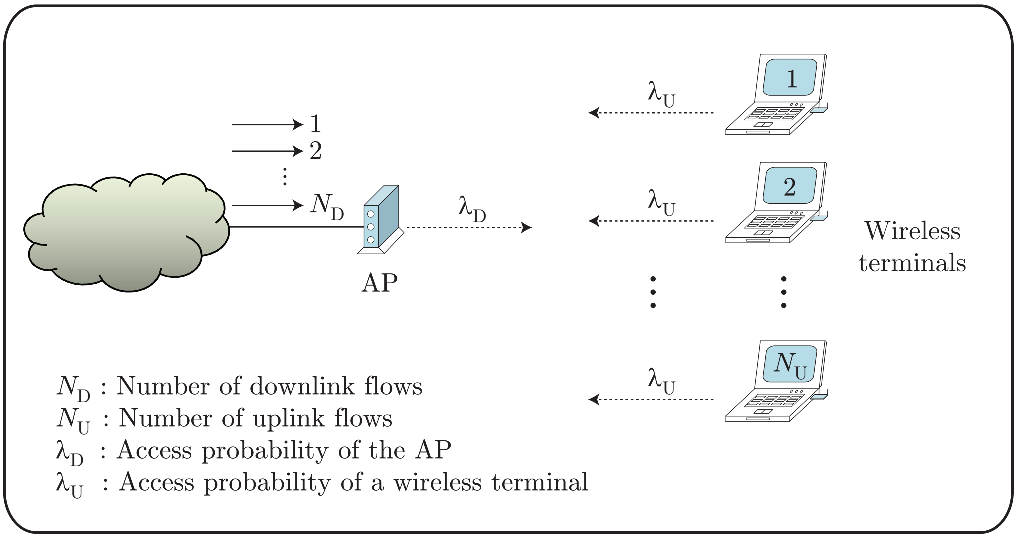


Figure 2.1: Wireless LAN model.

Let λ_{D} (resp. λ_{U}) denote the probability that the AP (resp. a wireless terminal) tries to transmit a frame at a randomly chosen time slot in contention periods. Also, let P_{S_D} (resp. P_{S_U}) denote the probability of a successful frame transmission of the AP (resp. a wireless terminal). We then have

$$\begin{aligned} P_{S\_\text{D}} &= \lambda_{\text{D}}(1-\lambda_{\text{U}})^{N_{\text{U}}}, \\ P_{S\_\text{U}} &= \lambda_{\text{U}}(1-\lambda_{\text{D}})(1-\lambda_{\text{U}})^{N_{\text{U}}-1}. \end{aligned}$$

Note here that after a successful data frame transmission, the station randomly chooses a backoff interval B_I from $[0, CW_{\min}]$, and when $B_I = 0$, the next transmission succeeds with probability one, while backoff counters of all other stations remain frozen. Since,

$$\Pr(B_I = 0) = \frac{1}{(CW_{\min}+1)},$$

the conditional probability $P_S(k)$ ($k = 1, 2, \ldots$) of k frames transmitted successively given that a frame succeeds in transmission is obtained to be

$$P_S(k) = \left(1 - \frac{1}{CW_{\min}+1}\right)\left(\frac{1}{CW_{\min}+1}\right)^{k-1}.$$

Therefore, once a station acquires the transmission right, it can successively transmit $1 + 1/CW_{\min}$ frames on average, without competing against other stations.

Let $CW_{\min\_D}$ and $CW_{\min\_U}$ denote the minimum contention window sizes at the AP and wireless terminals, respectively. Note that the total packet rate of downlink flows is proportional to $P_{S\_D}(1 + 1/CW_{\min\_D})$, while the packet rate of an uplink flow is proportional to $P_{S\_U}(1 + 1/CW_{\min\_U})$. Thus the ratio R of the total packet rate of downlink flows to the packet rate of an uplink flow is given by

$$R = A \cdot \frac{\lambda_D(1 - \lambda_U)}{\lambda_U(1 - \lambda_D)}, \tag{2.1}$$

where

$$A = \frac{1 + 1/CW_{\min\_D}}{1 + 1/CW_{\min\_U}}. \tag{2.2}$$

Note here that the packet rate R does not depend explicitly on the number N_U of uplink flows. Note also that the access probability λ_D of the AP is related closely to the minimum contention window size $CW_{\min\_D}$ at the AP. Thus the above discussion suggests that the optimal $CW_{\min\_D}$ at the AP to attain the target packet rate ratio R^* can be determined independent of the number N_U of uplink flows.

2.3.2 Mean Field Approximation Analysis

In this subsection, we conduct a mean field approximation analysis to obtain the optimal $CW_{\min\_D}$ at the AP in a system with N_D downlink flows and N_U uplink flows. We first consider the dynamics of the AP. Time instants immediately after frame transmissions are chosen as imbedded Markov points and we construct the imbedded Markov chain $\{X_n;\ n = 1, 2, \ldots\}$, where X_n denotes the number of retransmissions of the outstanding frame at the nth imbedded Markov point. We conduct the mean field approximation, i.e., every frame transmission of the AP fails with probability

$$\alpha_D = 1 - (1 - \lambda_U)^{N_U}, \tag{2.3}$$

where λ_U denotes the average access probability of a wireless terminal. Thus for all i ($i = 0, 1, \ldots$), the state transition from $X_n = i$ to $X_{n+1} = i + 1$ happens with probability α_D and the state transition from $X_n = i$ to $X_{n+1} = 0$ happens with probability $1 - \alpha_D$. Figure 2.2 shows the state transition diagram for the AP.

Let $\pi_D(i)$ ($i = 0, 1, \ldots$) denote the steady state probability of the imbedded Markov chain of the AP being in state i. It then follows that

$$\pi_D(i) = (1 - \alpha_D){\alpha_D}^i.$$

Next we derive the steady state probability of the AP at a randomly chosen slot in contention periods. The first transmission of a frame starts at a slot subsequent to

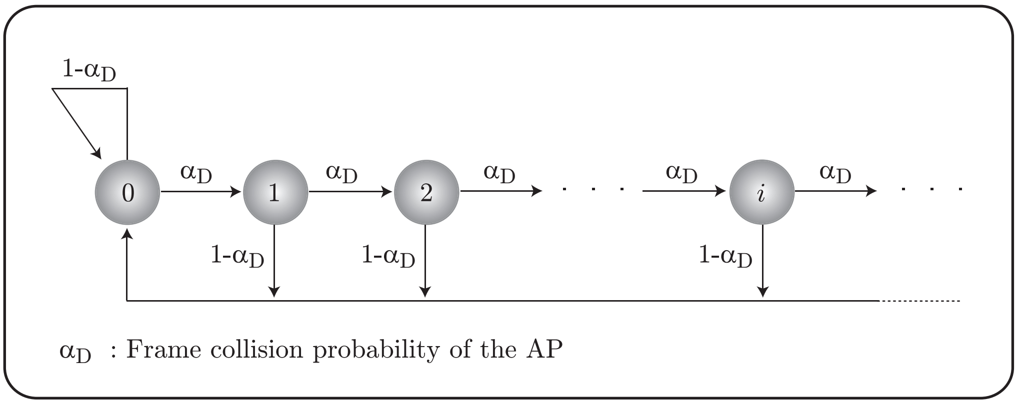


Figure 2.2: State transition diagram for the AP.

the last slot in the backoff period whose length is chosen randomly from $[0, CW_{\mathrm{min\_D}}]$. Therefore the mean number $CW_{\mathrm{D}}(0)$ of slots required for the first transmission is given by $CW_{\mathrm{min\_D}}/2+1$. For the ith ($i=1,2,\ldots$) retransmission, the length of the backoff period follows a uniform distribution on $[0, 2^i(CW_{\mathrm{min\_D}}+1)-1]$ (see (1.1)). Thus the mean number $CW_{\mathrm{D}}(i)$ ($i=1,2,\ldots$) of slots required for the ith retransmission is given by $2^{i-1}(CW_{\mathrm{min\_D}}+1)+1/2$. As a result, the stationary probability $\pi^*_{\mathrm{D}}(i)$ ($i=0,1,\ldots$) that the AP is within the ith contention window is given by

$$\pi^*_{\mathrm{D}}(i) = \frac{CW_{\mathrm{D}}(i)\pi_{\mathrm{D}}(i)}{\sum_{j=0}^{\infty} CW_{\mathrm{D}}(j)\pi_{\mathrm{D}}(j)}, \tag{2.4}$$

where $\alpha_{\mathrm{D}} < 1/2$ is necessary for the convergence of the infinite sum in the denominator on the right hand side of (2.4).

On the other hand, the average access probability $p_{\mathrm{D}}(i)$ ($i=0,1,\ldots$) of the AP for the ith retransmission is given by

$$p_{\mathrm{D}}(i) = \frac{1}{CW_{\mathrm{D}}(i)}.$$

Note that the above equation also holds for the transmission of new frames with $i=0$. Therefore the average access probability λ_{D} of the AP is given by

$$\begin{aligned}\lambda_{\mathrm{D}} &= \sum_{i=0}^{\infty} p_{\mathrm{D}}(i)\pi^*_{\mathrm{D}}(i) \\ &= \frac{2(1-2\alpha_{\mathrm{D}})}{(CW_{\mathrm{min\_D}}+1)(1-\alpha_{\mathrm{D}})+1-2\alpha_{\mathrm{D}}}.\end{aligned} \tag{2.5}$$

Similarly, the average access probability λ_{U} of a wireless terminal can be found to be

$$\lambda_{\mathrm{U}} = \frac{2(1-2\alpha_{\mathrm{U}})}{(CW_{\mathrm{min\_U}}+1)(1-\alpha_{\mathrm{U}})+1-2\alpha_{\mathrm{U}}}, \tag{2.6}$$

where α_{U} denotes the frame collision probability of a wireless terminal. Note that,

$$\alpha_{\mathrm{U}} = 1 - (1 - \lambda_{\mathrm{D}})(1 - \lambda_{\mathrm{U}})^{N_{\mathrm{U}}-1}. \tag{2.7}$$

Given N_{D}, N_{U}, $CW_{\mathrm{min\_U}}$, and $CW_{\mathrm{min\_D}}$, (2.5) and (2.6) are considered as a system of nonlinear equations for the access probabilities λ_{D} and λ_{U}. To solve it numerically, we may simply compute (2.7), (2.6), (2.3), and (2.5) in this order repeatedly until λ_{D} and λ_{U} converge, where the initial values of α_{D} and α_{U} are set to be small enough. Recall that the resulting packet rate ratio R is given by (2.1). Since the optimal $CW_{\mathrm{min\_D}}$ is not greater than $CW_{\mathrm{min\_U}}$, we can easily find the optimal $CW_{\mathrm{min\_D}}$ achieving the target packet rate ratio R^* by computing the packet rate ratios R's for all $CW_{\mathrm{min\_D}} \leq CW_{\mathrm{min\_U}}$.

2.3.3 Quasi-Optimal $CW^*_{\mathrm{min\_D}}$ to Achieve Fairness

Through numerical experiments for $N_{\mathrm{D}} \in [1, 15]$, $N_{\mathrm{U}} \in [1, 10]$, and $CW_{\mathrm{min\_U}} = 31$[1], we found that the optimal $CW_{\mathrm{min\_D}}$ is a nondecreasing function of N_{U} for any fixed $N_{\mathrm{D}} \in [1, 15]$. Specifically, for $N_{\mathrm{D}} = 1, \ldots, 4, 7, \ldots, 9$, and $11, \ldots, 14$, the optimal $CW_{\mathrm{min\_D}}$ remains constant for all $N_{\mathrm{U}} \in [1, 10]$. On the other hand, for $N_{\mathrm{D}} = 5$, 6, 10, and 15, the optimal $CW_{\mathrm{min\_D}}$ increases only by one with N_{U}. Thus the optimal $CW_{\mathrm{min\_D}}$ is rather insensitive to N_{U}, which supports our claim in Section 2.3.1. Note that this observation strongly suggests that for any N_{U}, we can use the optimal $CW_{\mathrm{min\_D}}$ for $N_{\mathrm{U}} = 1$ as a quasi-optimal $CW_{\mathrm{min\_D}}$. Thus in this subsection, we derive an explicit expression of the optimal $CW^*_{\mathrm{min\_D}}$ for $N_{\mathrm{U}} = 1$, achieving the target packet rate ratio $R^* \geq 1$, which will be served as a quasi-optimal $CW_{\mathrm{min\_D}}$ for all $N_{\mathrm{U}} \geq 1$.

When $N_{\mathrm{U}} = 1$, it follows from (2.3) and (2.7) that $\alpha_{\mathrm{D}} = \lambda_{\mathrm{U}}$ and $\alpha_{\mathrm{U}} = \lambda_{\mathrm{D}}$. Note also that from (2.1), the access probability λ_{D} of the AP is given in terms of the packet rate ratio R:

$$\lambda_{\mathrm{D}} = \frac{R\lambda_{\mathrm{U}}}{R\lambda_{\mathrm{U}} + A(1 - \lambda_{\mathrm{U}})}. \tag{2.8}$$

Substituting (2.8) into (2.6) and rearranging terms yield

$$[(CW_{\mathrm{min\_U}} + 2)A + R]\lambda_{\mathrm{U}}^2 - [(CW_{\mathrm{min\_U}} + 2)A + 2(R + A)]\lambda_{\mathrm{U}} + 2A = 0. \tag{2.9}$$

Recall that λ_{U} $(= \alpha_{\mathrm{D}})$ should satisfy $0 < \lambda_{\mathrm{U}} < 1/2$. It is easy to see that (2.9) has a unique solution $\lambda_{\mathrm{U}} = \lambda^*_{\mathrm{U}}$ in $(0, 1/2)$ when $CW_{\mathrm{min\_U}} \geq 2$.

On the other hand, it follows from (2.5) and (2.8) that

$$\frac{R\lambda_{\mathrm{U}}}{R\lambda_{\mathrm{U}} + A(1 - \lambda_{\mathrm{U}})} = \frac{2(1 - 2\lambda_{\mathrm{U}})}{(CW_{\mathrm{min\_D}} + 1)(1 - \lambda_{\mathrm{U}}) + 1 - 2\lambda_{\mathrm{U}}}. \tag{2.10}$$

[1]The default value of CW_{min} is equal to 31 in IEEE 802.11b wireless LANs.

Substituting the solution λ_{U}^* of (2.9) into (2.10) and manipulating the resulting equation with $[(CW_{\mathrm{min\_U}}+2)A+R]\lambda_{\mathrm{U}}^*(1-\lambda_{\mathrm{U}}^*) = 2A-(2A+R)\lambda_{\mathrm{U}}^*$, we obtain

$$CW_{\mathrm{min\_D}} = 2 + \frac{A}{R}(CW_{\mathrm{min\_U}} - 2). \tag{2.11}$$

We now derive the optimal $CW_{\mathrm{min\_D}}^*$ at the AP, attaining the target packet rate ratio $R^* \geq 1$. Letting $R = R^*$, substituting (2.2) into (2.11), and rearranging terms yield

$$CW_{\mathrm{min\_D}}^2 - 2\left(1+\frac{B}{R^*}\right)CW_{\mathrm{min\_D}} - \frac{2B}{R^*} = 0, \tag{2.12}$$

where B is given by

$$B = \frac{CW_{\mathrm{min\_U}}(CW_{\mathrm{min\_U}}-2)}{2(CW_{\mathrm{min\_U}}+1)}. \tag{2.13}$$

We define $f(x)$ as $f(x) = x^2 - 2(1+B/R^*)x - 2B/R^*$ and suppose $CW_{\mathrm{min\_U}} \geq 3$. We then have $f(0) = -CW_{\mathrm{min\_U}}(CW_{\mathrm{min\_U}}-2)/((CW_{\mathrm{min\_U}}+1)R^*) < 0$ and $f(CW_{\mathrm{min\_U}}) = (1-1/R^*)CW_{\mathrm{min\_U}}(CW_{\mathrm{min\_U}}-2) \geq 0$. Thus, when $CW_{\mathrm{min\_U}} \geq 3$, (2.12) has a unique solution in $(0, CW_{\mathrm{min\_U}}]$ and the other solution is negative.

Let $CW_{\mathrm{min\_D}}^*$ denote the positive solution of (2.12) which is rounded to the nearest integer. For $CW_{\mathrm{min\_U}} \geq 3$, $CW_{\mathrm{min\_D}}^*$, the optimal $CW_{\mathrm{min\_D}}$ of the AP, which attains the target packet rate ratio $R^* \geq 1$ can be expressed as

$$CW_{\mathrm{min\_D}}^* = \left\lfloor \frac{3}{2} + \frac{B}{R^*} + \sqrt{\left(1+\frac{B}{R^*}\right)^2 + \frac{2B}{R^*}} \right\rfloor, \tag{2.14}$$

where $\lfloor x \rfloor$ stands for the maximum integer that is not greater than x.

Since the optimal $CW_{\mathrm{min\_D}}$ is less sensitive to the number N_{U}, we consider using $CW_{\mathrm{min\_D}}^*$ in (2.14) as a quasi-optimal $CW_{\mathrm{min\_D}}$ at the AP for all $N_{\mathrm{U}} \geq 1$. Note that when $R^* = 1$, (2.14) yields $CW_{\mathrm{min\_D}}^* = CW_{\mathrm{min\_U}}$, which is consistent with the per-station fairness property of the IEEE 802.11 MAC protocol.

2.4 Simulation Experiments and Results

In this section, we evaluate the performance of our scheme through simulation experiments with an original simulator written in C++.

2.4.1 Simulation Model

Figure 2.3 shows the network topology used in the simulation experiments. We assume that N_{D} downlink flows and N_{U} uplink flows exist in the wireless LAN. We employ an IEEE 802.11b-based wireless LAN, whose parameters are shown in Table 1.1. Note that *RTSThreshold* is set to be 3,000 bytes. We assume that all stations transmit data frames at *DataRate*=11 Mbps and control frames at *BasicRate*=1 Mbps. The buffer size of each station is set to be 100 packets. All wired links have bandwidth of 100 Mbps and propagation delay of 25 msec. Simulation time for each scenario is set to be 2,000 seconds.

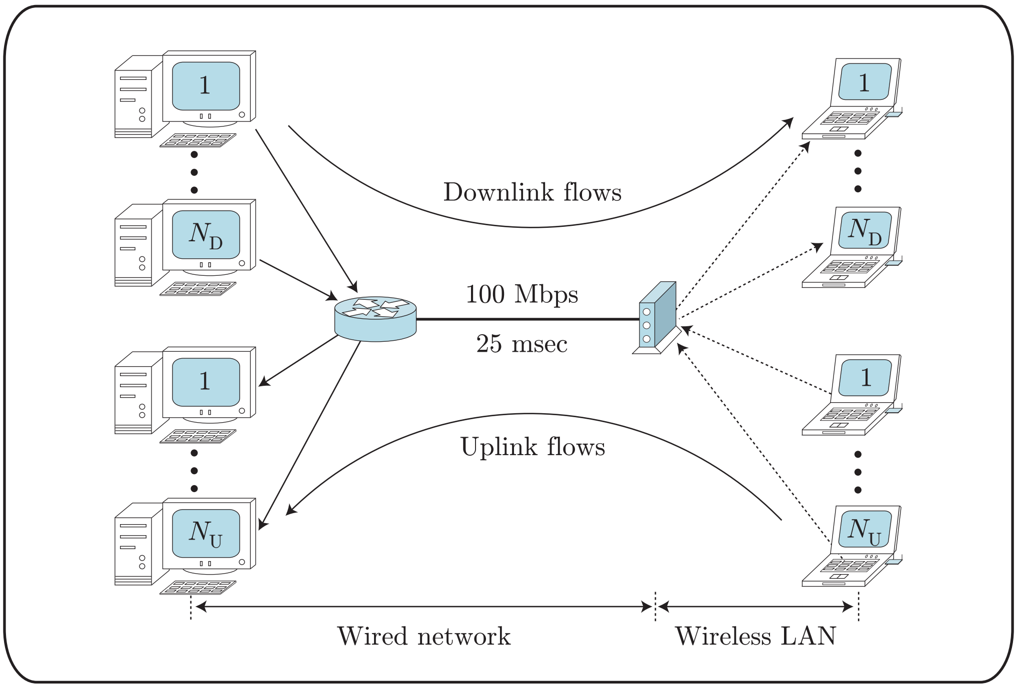


Figure 2.3: Network topology for simulation experiments.

2.4.2 Performance Measures

The performance of our scheme is evaluated in terms of the two fairness indices $FI_{m,n}$ and Γ, and total system throughput.

We first introduce fairness index proposed by Jain et al. [37] to quantify per-flow fairness. Let $\theta_{\mathrm{D},i}$ ($i = 1, 2, \ldots, N_{\mathrm{D}}$) denote the throughput of the ith downlink flow and $\theta_{\mathrm{U},j}$ ($j = 1, 2, \ldots, N_{\mathrm{U}}$) denote the throughput of the jth uplink flow. We then define fairness index $FI_{m,n}$ as

$$FI_{m,n} = \frac{\left(\displaystyle\sum_{i=1}^{m} \theta_{\mathrm{D},i} + \sum_{j=1}^{n} \theta_{\mathrm{U},j}\right)^2}{(m+n)\left(\displaystyle\sum_{i=1}^{m} {\theta_{\mathrm{D},i}}^2 + \sum_{j=1}^{n} {\theta_{\mathrm{U},j}}^2\right)}.$$

Note that $0 < FI_{N_{\mathrm{D}},N_{\mathrm{U}}} \leq 1$, and $FI_{N_{\mathrm{D}},N_{\mathrm{U}}} = 1$ (i.e., $FI_{m,n}$ with $m = N_{\mathrm{D}}$ and $n = N_{\mathrm{U}}$) when all $N_{\mathrm{D}} + N_{\mathrm{U}}$ flows enjoy complete fair-share of the bandwidth. On the other hand,

$FI_{N_{\mathrm{D}},0}$ quantifies per-flow fairness within N_{D} downlink flows and $FI_{0,N_{\mathrm{U}}}$ quantifies per-flow fairness within N_{U} uplink flows.

Next, we define another fairness index Γ to evaluate fairness *between* uplink and downlink flows as

$$\Gamma = \frac{\max(\overline{\theta}_{\mathrm{D}}, \overline{\theta}_{\mathrm{U}})}{\min(\overline{\theta}_{\mathrm{D}}, \overline{\theta}_{\mathrm{U}})},$$

where $\overline{\theta}_{\mathrm{D}}$ (resp. $\overline{\theta}_{\mathrm{U}}$) denotes the average throughput of downlink (resp. uplink) flows:

$$\overline{\theta}_{\mathrm{D}} = \frac{1}{N_{\mathrm{D}}} \sum_{i=1}^{N_{\mathrm{D}}} \theta_{\mathrm{D},i}, \quad \overline{\theta}_{\mathrm{U}} = \frac{1}{N_{\mathrm{U}}} \sum_{j=1}^{N_{\mathrm{U}}} \theta_{\mathrm{U},j}.$$

Note that $\Gamma \geq 1$, and $\Gamma = 1$ when uplink and downlink flows enjoy fair-share of the bandwidth.

2.4.3 Validation of the Analytical Result of $CW^*_{\mathrm{min\_D}}$

In Section 2.3, we claimed that the optimal $CW_{\mathrm{min\_D}}$ to achieve fairness is

(i) fairly insensitive to the number N_{U} of uplink flows and

(ii) it is given by $CW^*_{\mathrm{min\_D}}$ in (2.14).

We validate these claims by simulation experiments for UDP flows. The lengths of all segments are assumed to be 1,000 bytes and they are generated by CBR (constant bit rate) traffic source. The offered load of each flow is always set to be 10 Mbps, i.e., a saturated situation. As a result, $R^* = N_{\mathrm{D}}$, i.e., the AP handling N_{D} downlink flows should attain the packet rate N_{D} times as large as the packet rate of an uplink flow, in order to achieve per-flow fairness.

Let $\widetilde{CW}^*_{\mathrm{min\_D}}$ denote the optimal $CW_{\mathrm{min\_D}}$ that gives the minimum value of index Γ, which is obtained by simulation experiments. In order to evaluate the influence of the number N_{U} of uplink flows on $\widetilde{CW}^*_{\mathrm{min\_D}}$, we conducted several simulation experiments for various scenarios by changing the numbers N_{D} and N_{U}. With these simulations, we found that $\widetilde{CW}^*_{\mathrm{min\_D}}$ varies along with the number N_{D} of downlink flows, whereas $\widetilde{CW}^*_{\mathrm{min\_D}}$ did not vary with the number N_{U} of uplink flows. These results clearly support our claim (i). We will discuss this again in Section 2.4.6.

Next we investigate our claim (ii). Table 2.1 shows $CW^*_{\mathrm{min\_D}}$ in (2.14) with $R^* = N_{\mathrm{D}}$ and $\widetilde{CW}^*_{\mathrm{min\_D}}$, where $N_{\mathrm{U}} = 1$. We observe that $CW^*_{\mathrm{min\_D}}$ is identical with $\widetilde{CW}^*_{\mathrm{min\_D}}$, except for $N_{\mathrm{D}} = 10$, 14, and 25. Note that when $N_{\mathrm{D}} = 10$, $(FI_{10,1}, \Gamma) = (0.99053, 1.19983)$ and $(0.99682, 1.17590)$ for $CW_{\mathrm{min\_D}} = 5$ and 6, respectively. Similarly, when $N_{\mathrm{D}} = 14$, $(FI_{14,1}, \Gamma) = (0.99431, 1.24769)$ and $(0.99561, 1.22037)$ for $CW_{\mathrm{min\_D}} = 4$ and 5, respectively, and when $N_{\mathrm{D}} = 25$, $(FI_{25,1}, \Gamma) = 0.99107, 1.43227)$ and $(0.99368, 1.40631)$ for

$CW_{\text{min\_D}} = 3$ and 4, respectively; the difference in any case is very small. We then conclude that our formula (2.14) provides a quasi-optimal value of the minimum contention window size, even when it is not the optimal one.

Table 2.1: $CW^*_{\text{min\_D}}$, $\widetilde{CW}^*_{\text{min\_D}}$, and estimated achievable throughput ratio R_{est}.

R^*, N_{D}	$CW^*_{\text{min\_D}}$	$\widetilde{CW}^*_{\text{min\_D}}$	R_{est}
1	31	31	1.00
2	17	17	1.98
3	12	12	3.04
4	10	10	3.86
5	8	8	5.27
6, 7	7	7	6.42
8, 9	6	6	8.19
10	5	6	11.24
11, ..., 13	5	5	11.24
14	4	5	17.56
15, ..., 24	4	4	17.56
25	3	4	37.46
26, ..., 78	3	3	37.46

2.4.4 Fairness among UDP Flows

We now evaluate our scheme for UDP flows whose characteristics are identical with those in Section 2.4.3. Therefore we set $R^* = N_{\text{D}}$ in our scheme. We conduct simulation experiments by varying numbers N_{D} and N_{U}.

We first discuss the performance of our scheme when there exist small number of flows. We found that $FI_{N_{\text{D}},0}$ and $FI_{0,N_{\text{U}}}$ of both the default system (i.e., $CW_{\text{min\_D}} = 31$) and our scheme with $CW^*_{\text{min\_D}}$ in (2.14) are almost equal to one regardless of N_{D} and N_{U}. Therefore, instead of showing results of $FI_{N_{\text{D}},0}$ and $FI_{0,N_{\text{U}}}$, in Figs. 2.4 and 2.5, we compare $FI_{N_{\text{D}},N_{\text{U}}}$ and Γ, respectively.

As shown in Fig. 2.4, $FI_{N_{\text{D}},N_{\text{U}}}$ of the default system decreases with the increase of N_{D}. In contrast, $FI_{N_{\text{D}},N_{\text{U}}}$ of our scheme is very close to one regardless of N_{D} and N_{U}.

With the results shown in Fig. 2.5, we observe per-station fairness in the default system, i.e., Γ in the default system is very close to N_{D}. On the other hand, in our scheme Γ is always equal to one. Thus our scheme in (2.14) achieves fairness for any number of flows.

Figure 2.6 compares the total system throughput of both systems. We observe that, our scheme keeps high bandwidth utilization as in the default system. Furthermore, Table 2.2 compares the minimum and maximum throughput of all $N_{\text{D}} + N_{\text{U}}$ UDP flows.

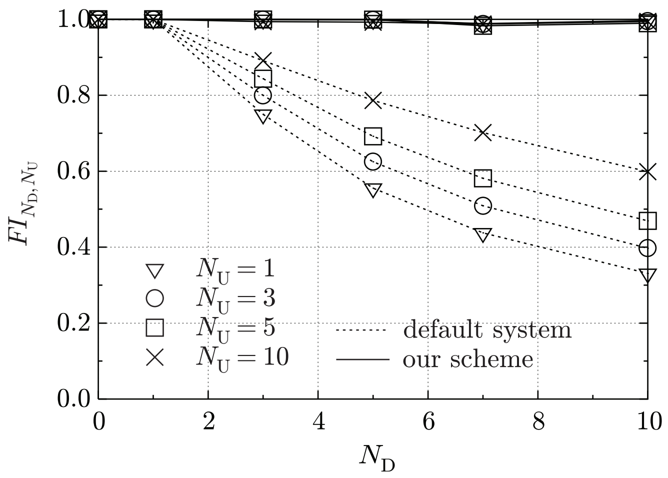


Figure 2.4: Per-flow fairness index $FI_{N_\mathrm{D},N_\mathrm{U}}$ of UDP flows.

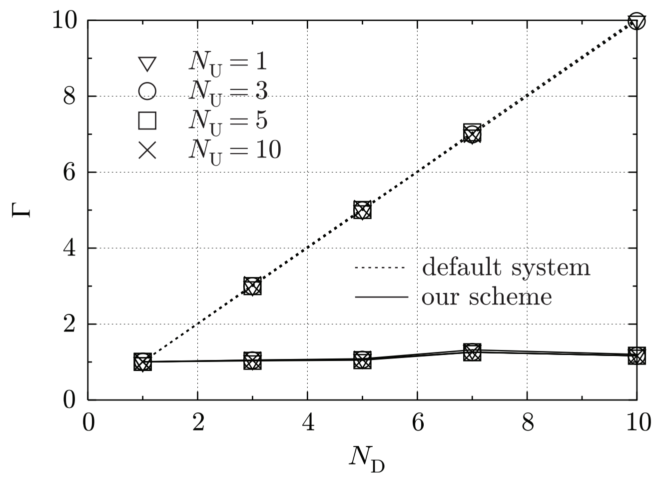


Figure 2.5: Uplink/downlink fairness index Γ of UDP flows.

Thus our scheme not only achieves fairness but also keeps high bandwidth utilization for UDP traffic when the number of flows is small.

We now examine the performance of our scheme when there are many flows. We observe in Table 2.3 that the total throughput of both systems become smaller with the increase of the number N_U of uplink flows. The total throughput of the default system does not vary with the number N_D of downlink flows, because the AP is always saturated. Recall that in our scheme, the AP uses a very small value of $CW_{\min}$ for large N_D (see Table 2.1). Therefore one might expect throughput degradation due to frequent frame collisions. As shown in Table 2.3, however, the total throughput of our scheme is almost the same as that of the default system.

To examine why the total throughput does not degrade even for large N_D, we have a

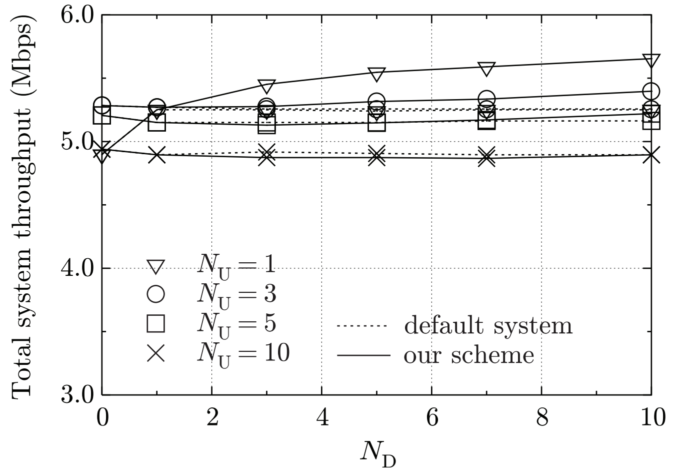


Figure 2.6: Total system throughput of UDP flows.

Table 2.2: Minimum and maximum throughput of UDP flows (in Mbps).

R^*, N_D	N_U	Default system		Our scheme	
		Minimum	Maximum	Minimum	Maximum
1	1	2.62	2.62	2.62	2.62
	5	0.85	0.86	0.85	0.86
	10	0.44	0.45	0.44	0.45
5	1	0.52	2.62	0.91	0.98
	5	0.17	0.86	0.50	0.53
	10	0.09	0.45	0.31	0.33
10	1	0.26	2.62	0.50	0.60
	5	0.08	0.86	0.33	0.38
	10	0.04	0.45	0.23	0.26

look at frame collision probabilities. Table 2.4 shows the overall frame collision probability P_c, as well as the frame collision probabilities $P_c^{(AP)}$ and $P_c^{(WT)}$ of the AP and a wireless terminal. We observe that P_c's of both systems increase with N_U. However, when the number of downlink flows is large (say $N_D \geq 15$), P_c of our scheme is smaller than that of the default system. We explain this interesting phenomenon, assuming $N_U = 1$. Recall that when a frame collision occurs, both the AP and the wireless terminal double their current CW. Since $CW^*_{\text{min\_D}} \ll CW_{\text{min\_U}} = 31$, the AP can access the wireless channel several times before the wireless terminal does, once a frame collision occurs. Therefore the number of frame transmission attempts of the AP is larger than that of the wireless terminal. Note that when $N_U = 1$, the AP and the wireless terminal suffer from the same number of frame collisions because frame collisions happen only when those two stations transmit their frames at the same time. As a result, we have $P_c^{(AP)} < P_c^{(WT)}$, and for

Table 2.3: Total system throughput of UDP flows when N_D and N_U are large.

R^*, N_D	N_U	Total system throughput (Mbps)	
		Default system	Our scheme
15	15	4.95	5.03
	30	4.66	4.76
	50	4.42	4.44
30	15	4.95	5.05
	30	4.67	4.77
	50	4.41	4.46
50	15	4.95	5.06
	30	4.66	4.77
	50	4.41	4.46

large N_D, the overall frame collision probability of our scheme is smaller than that of the default system. These observations imply that our scheme is scalable to systems with many flows, in terms of the total throughput.

Table 2.4: Frame collision probability of UDP flows.

N_U	$P_\mathrm{c}, P_\mathrm{c}^{(\mathrm{AP})}, P_\mathrm{c}^{(\mathrm{WT})}$	Default system	Our scheme		
		$N_\mathrm{D} \geq 1$	$R^* = N_\mathrm{D}$		
			15	30	50
1	P_c	0.06	0.03	0.02	0.03
	$P_\mathrm{c}^{(\mathrm{AP})}$	0.06	0.02	0.01	0.02
	$P_\mathrm{c}^{(\mathrm{WT})}$	0.06	0.42	0.54	0.54
15	P_c	0.35	0.31	0.24	0.25
	$P_\mathrm{c}^{(\mathrm{AP})}$	0.35	0.19	0.14	0.14
	$P_\mathrm{c}^{(\mathrm{WT})}$	0.35	0.45	0.50	0.51
30	P_c	0.45	0.43	0.38	0.38
	$P_\mathrm{c}^{(\mathrm{AP})}$	0.45	0.28	0.23	0.23
	$P_\mathrm{c}^{(\mathrm{WT})}$	0.45	0.50	0.53	0.53
50	P_c	0.53	0.52	0.49	0.49
	$P_\mathrm{c}^{(\mathrm{AP})}$	0.53	0.36	0.31	0.31
	$P_\mathrm{c}^{(\mathrm{WT})}$	0.53	0.56	0.57	0.57

Next we examine fairness between uplink and downlink flows in our scheme for large N_D and N_U. Recall that our scheme selects the optimal value of $CW_{\mathrm{min\_D}}$ from integers in [3, 31]. Thus the degree of freedom is limited and therefore our scheme cannot always attain complete fair-share of the bandwidth, especially when there are many flows. Note here that achievable fairness in the current scenario can be estimated in the following way.

When N_{D} is given, the optimal $CW^*_{\mathrm{min\_D}}$ is determined by (2.14) with $R^* = N_{\mathrm{D}}$, and the resulting throughput ratio R_{est} is estimated with (2.11) and (2.2), as shown in Table 2.1. The index Γ is then estimated to be $\max(R_{\mathrm{est}}/N_{\mathrm{D}}, N_{\mathrm{D}}/R_{\mathrm{est}})$. Table 2.5 compares Γ_{est} and Γ_{sim}, where Γ_{est} denotes the Γ value estimated according to this procedure and Γ_{sim} denotes the Γ value obtained by simulation experiments. We observe that Γ_{est} is fairly close to Γ_{sim}. Together with Table 2.3, we conclude that for large N_{D} and N_{U}, our scheme greatly ameliorates fairness between uplink and downlink flows without degradation of the system throughput, even if it cannot achieve complete fair-share of the bandwidth.

Table 2.5: Comparison of Γ_{est} and Γ_{sim}.

R^*, N_{D}	Γ_{est}		Γ_{sim}	
	Default system	Our scheme	Default system	Our scheme
5	5.00	1.05	4.99	1.04
15	15.00	1.17	14.98	1.13
30	30.00	1.25	30.01	1.28
50	50.00	1.33	49.97	1.27

2.4.5 Fairness among TCP Flows

In this subsection, we evaluate our scheme for TCP New Reno flows [19] with data segments of 1,000 bytes. To do so, we consider a wireless LAN with N_{D} downlink TCP flows and N_{U} uplink TCP flows. In this case, the influence of TCP ACK segments should be taken into account, because the AP handles N_{U} TCP ACK segment flows, as well as N_{D} TCP data segment flows. When delayed-ACK is not implemented, the mean number of TCP ACK segments arriving in a unit time is equal to the mean number of TCP data segments generated in a unit time. This implies that the packet rate of a TCP ACK segment flow should be equal to that of the corresponding data segment flow. Thus the optimal $CW^*_{\mathrm{min\_D}}$ is determined with $R^* = N_{\mathrm{D}} + N_{\mathrm{U}}$. Note that if n out of N_{U} uplink flows take delayed-ACK option, $CW^*_{\mathrm{min\_D}}$ would be determined with $R^* = N_{\mathrm{D}} + (N_{\mathrm{U}} - n) + n/2$ because with delayed-ACK option, TCP ACK segments are generated every other TCP data segment. For the simplicity, we assume that none of TCP flows uses the delayed-ACK option.

Figures 2.7 and 2.8 compare per-flow fairness index $FI_{N_{\mathrm{D}},N_{\mathrm{U}}}$ and uplink/downlink fairness index Γ of TCP flows. We observe that uplink and downlink flows in the default system share the bandwidth equally when $N_{\mathrm{D}} + N_{\mathrm{U}} \leq 5$. When $N_{\mathrm{D}} + N_{\mathrm{U}} \geq 6$, however, the throughput of uplink flows is greater than that of downlink flows and the extent of unfairness increases with $N_{\mathrm{D}} + N_{\mathrm{U}}$. On the other hand, our scheme achieves fair-share of the bandwidth, regardless of the number of flows.

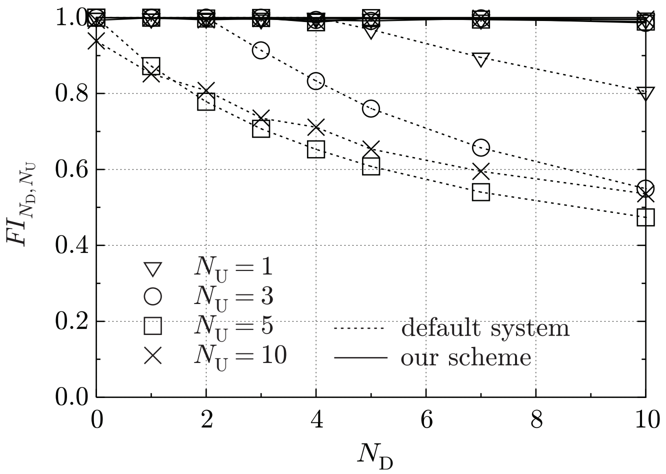


Figure 2.7: Per-flow fairness index $FI_{N_{\mathrm{D}},N_{\mathrm{U}}}$ of TCP flows.

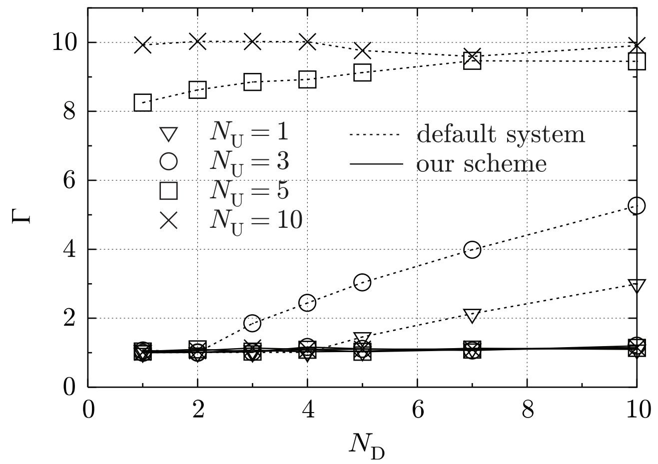


Figure 2.8: Uplink/downlink fairness index Γ of TCP flows.

The above phenomenon in the default system can be explained as follows. Table 2.6 shows the range of the packet loss probability at the AP due to buffer overflow in the default system. We observe that the packet loss does not occur at the AP when $N_{\mathrm{D}}+N_{\mathrm{U}} \leq 5$. In this simulation experiment, the buffer size of the AP is set to be 100 packets and advertised window of TCP flows is set to be 20 packets. Since delayed-ACK is not implemented in this simulation experiment, no packet loss occurs if $(N_{\mathrm{D}}+N_{\mathrm{U}})\times 20 \leq 100$. In such a situation, TCP connections reach the steady state and the self-clocking effect emerges [36]. Thus TCP ACK segments regulate the transmission rate of data segments at wireless terminals, and therefore the bandwidth is shared equally among all flows. When

$N_{\mathrm{D}}+N_{\mathrm{U}} \geq 6$, however, the packet loss is observed and the packet loss probability $P_{\mathrm{loss}}^{(\mathrm{AP})}$ at the AP increases with $N_{\mathrm{D}}+N_{\mathrm{U}}$. As a result, the throughput of downlink flows decreases, as explained in Section 2.1. Thus, in the default system, per-flow fairness among TCP flows can be achieved when the buffer size of the AP is large enough to prevent packet loss. Note that large buffer leads to large queueing delay, and it becomes difficult to guarantee QoS for real-time traffic. Furthermore, the increase of queueing delay may result in throughput degradation of TCP flows because the throughput performance of TCP depends on RTT (round trip time). Thus, in the default system, the performance of respective flows would be sacrificed to achieve fairness.

Table 2.6: Packet loss probability of TCP flows at the AP, due to buffer overflow, in the default system.

$N_{\mathrm{D}}+N_{\mathrm{U}}$	$P_{\mathrm{loss}}^{(\mathrm{AP})}$
5 or less	0.0
6	$1.5 \times 10^{-3} \sim 1.3 \times 10^{-2}$
7	$2.9 \times 10^{-3} \sim 5.8 \times 10^{-2}$
8	$4.1 \times 10^{-3} \sim 1.5 \times 10^{-1}$
9	$5.2 \times 10^{-3} \sim 1.1 \times 10^{-1}$
10	$6.2 \times 10^{-3} \sim 1.4 \times 10^{-1}$

Figure 2.9 compares the total system throughput of TCP flows. We observe that there is little difference between the total throughput of the default system and that of our scheme. Furthermore, Table 2.7 compares the minimum and maximum throughput of all $N_{\mathrm{D}}+N_{\mathrm{U}}$ TCP flows. With these results, we conclude that our scheme ameliorates fairness of TCP flows without any serious side effects.

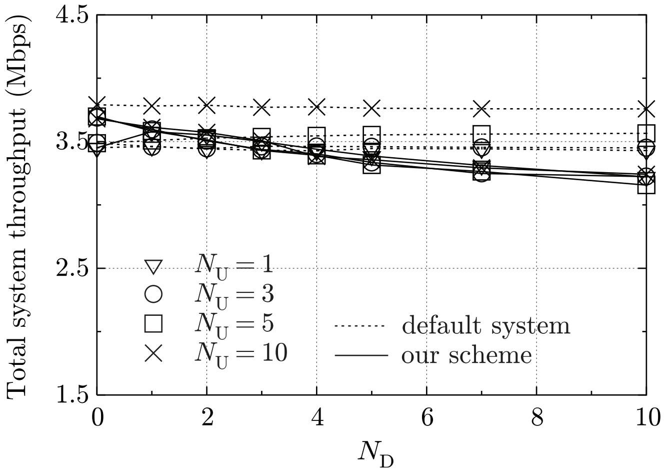


Figure 2.9: Total system throughput of TCP flows.

Table 2.7: Minimum and maximum throughput of TCP flows (in Mbps).

R^*, N_D	N_U	Default system		Our scheme	
		Minimum	Maximum	Minimum	Maximum
1	1	1.68	1.77	1.69	1.76
	5	0.08	0.69	0.59	0.61
	10	0.03	0.38	0.25	0.29
5	1	0.49	0.79	0.53	0.58
	5	0.06	0.66	0.32	0.35
	10	0.02	0.39	0.20	0.23
10	1	0.22	0.81	0.26	0.31
	5	0.05	0.62	0.19	0.22
	10	0.02	0.37	0.13	0.16

2.4.6 Dynamic Behavior of Our Scheme

Finally we discuss the dynamic behavior of our scheme for UDP and TCP flows. To do so, we change the numbers of uplink and downlink flows during a single simulation experiment, as shown in Table 2.8. Note that in our scheme, $CW^*_{\mathrm{min\_D}}$ is dynamically determined with $R^* = N_\mathrm{D}$ for UDP flows and $R^* = N_\mathrm{D} + N_\mathrm{U}$ for TCP flows.

Table 2.8: Simulation scenario for dynamic behaviors.

interval (sec)	[0,300)	[300,600)	[600,900)
$(N_\mathrm{D}, N_\mathrm{U})$	$(3, 3)$	$(3, 4)$	$(3, 5)$
interval (sec)	[900,1200)	[1200,1500)	[1500,1800)
$(N_\mathrm{D}, N_\mathrm{U})$	$(4, 5)$	$(5, 5)$	$(5, 4)$
interval (sec)	[1800,2100)	[2100,2400)	[2400,2700]
$(N_\mathrm{D}, N_\mathrm{U})$	$(5, 3)$	$(4, 3)$	$(3, 3)$

Figure 2.10 plots the average throughput of uplink UDP flows and downlink UDP flows of the default system, our scheme, and the PIFS scheme proposed in [22, 46], where the average throughput is calculated every 60 seconds. In the default system, the number of downlink flows does not affect the throughput of uplink flows. Therefore they remain almost constant during time intervals [600,1500) and [1800,2700]. The behavior of the PIFS system is similar to the default system, even though it slightly improves fairness compared to the default system, because *PIFS* is a constant parameter. On the other hand, our dynamic contention window control scheme works very well and significantly ameliorates fairness between uplink and downlink UDP flows in any situation. Since the number N_D of downlink flows does not change during time intervals [0,900) and [1200,2100), $CW_{\mathrm{min\_D}}$ of the AP in our scheme is fixed during these time intervals. Nonetheless, Figure 2.10 shows that our scheme achieves fair-share of the bandwidth, and

this reconfirms our claim (i) discussed in Section 2.4.3.

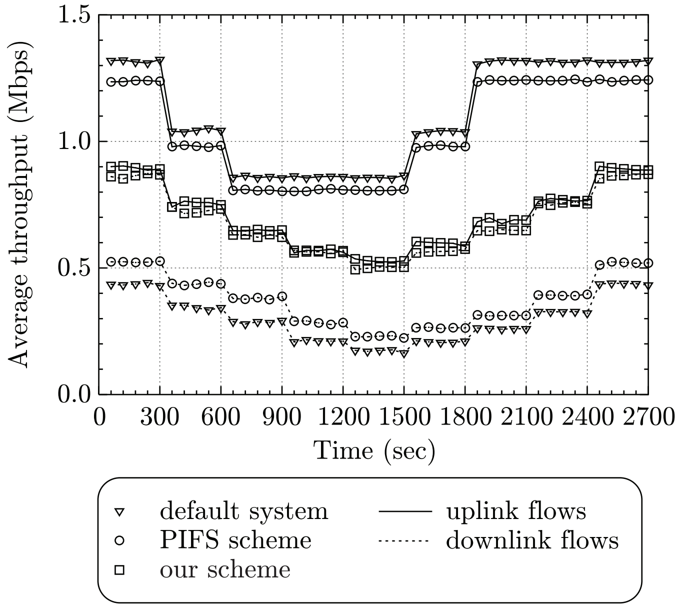


Figure 2.10: Dynamic behaviors of UDP flows.

Figure 2.11 plots the average throughput of uplink TCP flows and downlink TCP flows of the default system, our scheme, and the TXOP scheme proposed in [50], where the average throughput is calculated every 60 seconds. We observe that our dynamic contention window control scheme works very well; it achieves fairness between uplink and downlink TCP flows and keeps the system stable in any situation, as in the case of UDP flows. On the other hand, we observe that the throughput in the TXOP scheme fluctuates. This phenomenon can be explained as follows. Recall that in the TXOP scheme, the *TXOP* duration of the AP varies according to the number of downlink flows, while providing the AP and wireless terminals with equal opportunities to acquire the transmission right of TCP data segments. This feature leads to intermittent burst transmissions of TCP data segments from the AP, and therefore the TXOP scheme cannot achieve short-term fairness between uplink and downlink flows.

2.5 Conclusion

In this chapter, we discussed the fairness issue between uplink and downlink flows in single-rate IEEE 802.11 wireless LANs. To achieve fairness, we presented a dynamic contention window control scheme, where the minimum contention window $CW_{\min}$ of an AP is controlled according to the target packet rate ratio R^* between uplink and downlink flows. We conducted various simulation experiments with UDP and TCP flows,

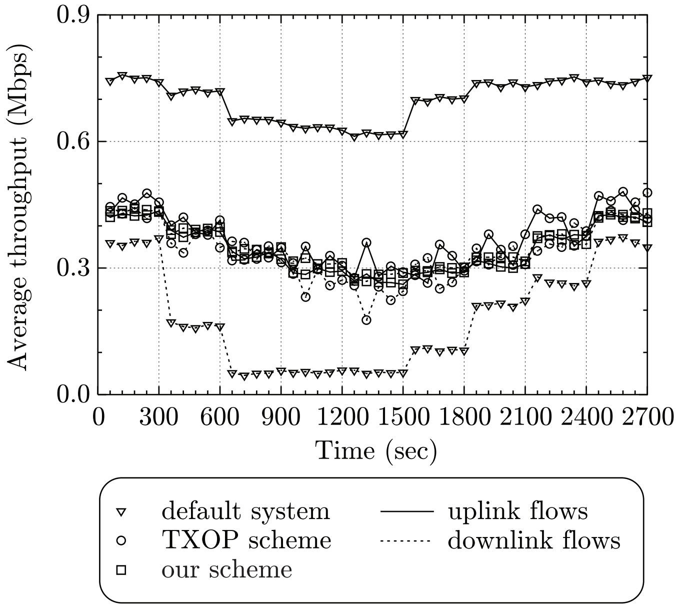


Figure 2.11: Dynamic behaviors of TCP flows.

and demonstrated that our scheme can significantly ameliorate fairness between uplink and downlink flows.

In determining the target packet rate ratio R^*, our scheme may require APs to know several information about downlink flows, including the number N_{D} of downlink flows. To do so, it is necessary to implement an additional function on APs to collect information of layers 3 and 4, e.g., IP addresses, port numbers, and transport protocol.

Chapter 3

Per-Flow Fairness in QoS-Oriented Wireless LANs

So far, many research efforts have been devoted to modeling and evaluating the performance of the IEEE 802.11e EDCA [9, 21, 28, 42, 47, 55, 66, 76, 91]. Most of those, however, aimed to improve QoS of real-time flows, and only a few studies have examined the performance of best-effort flows [8,50]. In this chapter, we discuss a fairness issue between uplink and downlink best-effort flows in the QoS-oriented IEEE 802.11e EDCA-based wireless LANs.

The IEEE 802.11e EDCA assigns different parameter values to different access categories in order to differentiate flows with different QoS requirements. As a result, a bundle of flows transmitted from an access category in an AP is treated in the same way as an individual flow transmitted from the same access category in wireless terminals, because the same set of parameter values is used in all stations. Thus, as shown in Section 3.1, unfairness between uplink and downlink best-effort flows emerges at the best-effort access category (i.e., AC_2) in the IEEE 802.11e EDCA-based wireless LANs, which leads to a serious throughput degradation of downlink best-effort flows, compared with the contending uplink best-effort flows. Giving more access chances to downlink best-effort flows will resolve this throughput unfairness. If we did it in a naive way, however, the performance of real-time traffic at AC_0 and AC_1 would be degraded and QoS requirements of these real-time traffic would not be guaranteed.

In this chapter, we consider a dynamic contention window control scheme to ameliorate fairness between uplink and downlink best-effort flows, while guaranteeing QoS requirements for real-time flows. In our scheme, the minimum contention window size $CW_{\min}$ of the best-effort access category AC_2 at APs is first determined based on the number of TCP flows, in such a way that this unfairness is resolved. $CW_{\min}$s and the maximum contention window sizes $CW_{\max}$s for real-time traffic at APs are then determined so as to guarantee QoS requirements for these traffic. Note that our scheme does not require any modifications at wireless terminals.

The rest of this chapter is organized as follows. Section 3.1 overviews the fairness issue between uplink and downlink flows occurred in IEEE 802.11e EDCA wireless LANs, and Section 3.2 presents related work. Section 3.3 describes our dynamic contention window control scheme. In Section 3.4, our scheme is evaluated with simulation experiments. Finally, the conclusion is given in Section 3.5.

3.1 Problem Overview

The IEEE 802.11e EDCA differentiates traffic flows in terms of their QoS requirements, and therefore flows transmitted from an AP and wireless terminals in the same access category are treated equally. Thus unfairness described in Chapter 2 arises at AC_2, because AC_2 handles best-effort traffic that typically uses TCP. Note that unless AC_0 and AC_1 are saturated, their QoS is guaranteed and fairness issue does not arise at AC_0 and AC_1. In this chapter, we present a dynamic contention window control scheme to alleviate unfairness between uplink and downlink best-effort flows in the IEEE 802.11e EDCA-based wireless LANs.

3.2 Related Work

So far, a number of works have been conducted to evaluate and enhance the performance of the IEEE 802.11e wireless LANs. Some of them have proposed window control schemes too, which control $CW_{\min}$s or/and $CW_{\max}$s [21, 42, 66, 76, 91]. However, most of them aim at improving the performance of real-time flows and they do not consider enhancing the performance of best-effort flows.

Casetti et al. proposed a static solution of unfairness between uplink and downlink best-effort flows in the IEEE 802.11e-based wireless LANs [8], where $AIFS[i]$ and $CW_{\min}[i]$ of downlink flows were set to be smaller values. This scheme, however, lacks flexibility in response to the number of flows because it uses fixed parameter values.

Leith et al. proposed adjusting $TXOP[2]$, in order to achieve fairness [50]. In their scheme, downlink TCP data segments and TCP ACK segments of uplink flows are classified into different access categories, and downlink TCP data segments are transmitted with a longer TXOP than other flows. The length of TXOP for downlink TCP data segments is dynamically controlled based on the number of downlink flows, whereas $TXOP[0]$ and $TXOP[1]$ remain in default values. Thus the performance of real-time flows degrades with the increase of the number of downlink TCP data segment flows.

3.3 Dynamic Contention Window Control

We aim at achieving fairness between uplink and downlink best-effort flows, while guaranteeing QoS requirements for real-time flows, where all wireless terminals are assumed to follow the default IEEE 802.11e EDCA protocol. In other words, we provide sufficient bandwidth with real-time traffic and then provide the remaining bandwidth equally with uplink and downlink best-effort flows. To do so, our scheme dynamically adjusts $CW_{\min}[i]$s and $CW_{\max}[i]$s at the AP. The idea behind our scheme is as follows. Note first that the default IEEE 802.11e EDCA guarantees QoS of real-time traffic by setting (see Table 1.2)

$$\begin{aligned} CW_{\max}[0] &= CW_{\min}[1], \\ CW_{\max}[1] &= CW_{\min}[2], \text{ and} \\ CW_{\max}[i] &= 2CW_{\min}[i] + 1 \ (i = 0, 1). \end{aligned}$$

We follow this principle, too, which is described in Table 3.1, where $CW_{\min\_D}[i]$ and $CW_{\max\_D}[i]$ $(i = 0, 1, 2, 3)$ denote $CW_{\min}[i]$ and $CW_{\max}[i]$ of AC_i at the AP. In this way, QoS of real-time traffic is expected to be guaranteed.

Table 3.1: $CW_{\min}[i]$ and $CW_{\max}[i]$ $(i = 0, 1)$ in our scheme.

$CW_{\min\_D}[i], CW_{\max\_D}[i]$	AC_0 of the AP	AC_1 of the AP
$CW_{\min\_D}[i]$	$\frac{CW_{\min\_D}[2]+1}{4} - 1$	$\frac{CW_{\min\_D}[2]+1}{2} - 1$
$CW_{\max\_D}[i]$	$\frac{CW_{\min\_D}[2]+1}{2} - 1$	$CW_{\min\_D}[2]$

On the other hand, to achieve fairness between uplink and downlink best-effort flows, we dynamically adjust $CW_{\min}[2]$ at the AP, based on the number of best-effort flows that the AP handles. Note that $CW_{\min}[i]$ and $CW_{\max}[i]$ $(i = 0, 1)$ at the AP are also determined uniquely once $CW_{\min\_D}[2]$ at the AP is determined (see Table 3.1). In Chapter 2, we presented a way of adjusting $CW_{\min}$ at the AP for the legacy IEEE 802.11 wireless LANs. Our scheme is based on the number of downlink flows, where for an arbitrarily fixed R, the total packet rate of downlink flows is R times as large as the packet rate of an individual uplink flow. We apply this scheme to achieve fairness in uplink and downlink best-effort flows in the IEEE 802.11e wireless LANs.

We now summarize our scheme. Let $CW_{\min\_U}[i]$ and $CW_{\max\_U}[i]$ $(i = 0, 1, 2, 3)$ denote $CW_{\min}[i]$ and $CW_{\max}[i]$ of AC_i at wireless terminals. Let M denote the number of best-effort flows that the AP handles at AC_2. The target packet rate R is then given by

$$R = \max(1, M),$$

where we assume that the mean frame lengths of best-effort flows are identical. Thus from (2.14), we set $CW_{\text{min\_D}}[2]$ (≥ 3) to be

$$CW_{\text{min\_D}}[2] = \left\lfloor \frac{3}{2} + \frac{B}{R} + \sqrt{\left(1 + \frac{B}{R}\right)^2 + \frac{2B}{R}} \right\rfloor, \tag{3.1}$$

where $\lfloor x \rfloor$ stands for the maximum integer that is not greater than x and

$$B = \frac{CW_{\text{min\_U}}[2](CW_{\text{min\_U}}[2] - 2)}{2(CW_{\text{min\_U}}[2] + 1)}.$$

As stated in Section 2.4.5, If best-effort traffic consists of TCP flows, the AP handles TCP ACK segment flows in AC_2, as well as TCP data segment flows. In such a case, M is given by the total number of those flows. For example, when there are $N_{\text{BE\_D}}$ downlink TCP flows and $N_{\text{BE\_U}}$ uplink TCP flows, the AP handles $N_{\text{BE\_D}}$ downlink data segment flows and $N_{\text{BE\_U}}$ downlink ACK segment flows, so that $M = N_{\text{BE\_D}} + N_{\text{BE\_U}}$. Our scheme then sets $CW_{\text{min\_D}}[i]$ and $CW_{\text{max\_D}}[i]$ ($i = 0, 1$) at the AP as shown in Table 3.1, while other parameters are fixed to their default values in Table 1.2. Note that when $M = 0$ or 1, $CW_{\text{min\_D}}[2] = 31 = CW_{\text{min\_U}}[2]$, and therefore all parameter values are equal to the default ones.

3.4 Simulation Experiments and Results

In this section, we evaluate the performance of our scheme through simulation experiments.

3.4.1 Simulation Model

Figure 3.1 shows the network topology used in the simulation experiments, which consists of N_{VO} VoIP calls (VoIP sessions), $N_{\text{VI\_D}}$ downlink video flows, $N_{\text{VI\_U}}$ uplink video flows, $N_{\text{BE\_D}}$ downlink TCP flows, and $N_{\text{BE\_U}}$ uplink TCP flows. Since each VoIP session generates both uplink and downlink flows, there are N_{VO} uplink VoIP flows and N_{VO} downlink VoIP flows. We assume that each voice traffic is encoded into a VoIP flow with ITU-T G.711 [68], which has the average source bit rate of 64 kbps and the packet size of 160 bytes. We also assume that each video traffic flow is encoded with ITU-T H.263 [81], which has the average bit rate of 512 kbps and the packet size of 512 bytes. UDP is used as layer-4 protocol for both voice and video flows. We employ an IEEE 802.11b-based wireless LAN, whose parameters are shown in Table 1.1. Note that *RTSThreshold*, *DataRate*, and *BasicRate* is set to be 800 bytes, 11 Mbps, and 1 Mbps, respectively. Note also that every wireless terminal in our scheme follows the default IEEE 802.11e EDCA whose parameters are given in Table 1.2. The buffer size for each access category is set to be 100 packets. All wired links have bandwidth of 100 Mbps and propagation delay of 25 msec. Simulation time for each scenario is set to be 2,000 seconds.

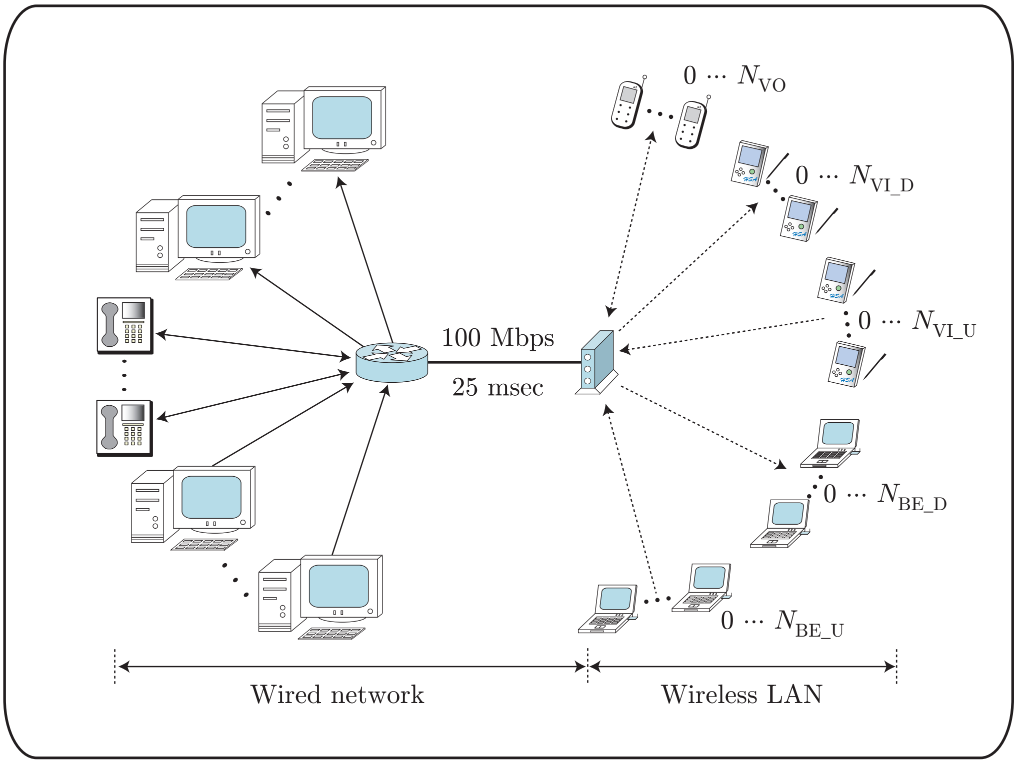


Figure 3.1: Network topology for simulation experiments.

3.4.2 Performance Measures

The performance of real-time traffic is evaluated in terms of the 99th percentile delay, the 99th percentile jitter, and the average packet loss ratio. The maximum allowable values of these QoS metrics are shown in Table 3.2 [17, 35]. We regard these as strict QoS constraints. Thus, in all simulation experiments, real-time packets whose end-to-end delay or/and jitter is not less than their constraints are discarded at the decoder and they count as lost packets. Therefore, if the packet loss ratio of a real-time flow is smaller than its constraint, QoS requirement of the real-time flow is satisfied.

On the other hand, the performance of best-effort traffic is evaluated in terms of two fairness indices, $FI_{m,n}$ and Γ defined in Section 2.4.2, and the total best-effort throughput.

3.4.3 Capacity for Accommodating Real-Time Traffic

This subsection evaluates the capacity for accommodating real-time traffic in our scheme. When the number M of downlink best-effort flows at AC_2 in the AP is very large,

Table 3.2: Maximum allowable values of QoS metrics [17, 35].

Traffic type	Delay (msec)	Jitter (msec)	Packet loss ratio
VoIP	150	10	3%
Video	150	20	1%

our scheme sets $CW_{\mathrm{min\_D}}[0]$ and $CW_{\mathrm{min\_D}}[1]$ at the AP to be 0 and 1, respectively. We call this the *extremal situation* because the AP tries to transmit frames most aggressively.

We first consider the case where only VoIP traffic flows exist in the wireless LAN. We perform simulation experiments by varying the number N_{VO} of VoIP sessions from 1 to 15. Figures 3.2 and 3.3 plot the 99th percentile end-to-end delay and the 99th percentile jitter of uplink and downlink VoIP flows, respectively, in the default system and in our scheme under the extremal situation. As we see in Fig. 3.2, when the number N_{VO} of VoIP sessions is small (say, $N_{\mathrm{VO}} = 4$), the 99th percentile delay of downlink flows is almost the same as that of the uplink flows, both in the default system and in our scheme under the extremal situation. In the default system, however, delay of the downlink VoIP flows increases rapidly with N_{VO} compared with that of the uplink VoIP flows. This is caused by the queueing delay at the AP. On the other hand, when our scheme runs under the extremal situation, $CW_{\mathrm{min\_D}}[0] = 0$ at the AP, and under such a situation the AP accesses to the wireless channel very aggressively. Thus uplink flows have to wait a long time, on average, in accessing the wireless channel and this results in a comparatively higher delay for the uplink VoIP flows, while the delay of the downlink VoIP flows remains almost constant.

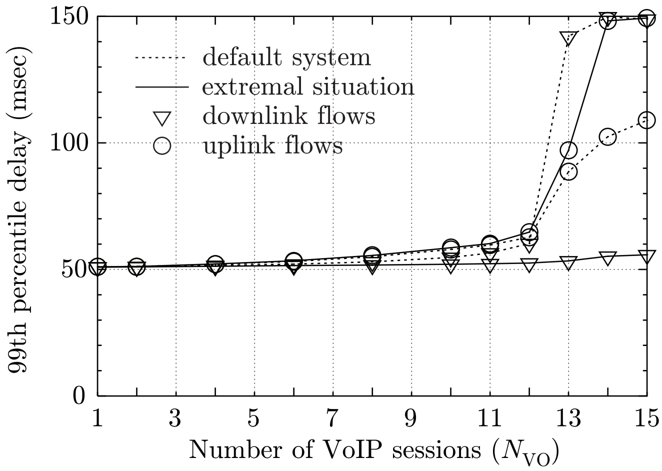


Figure 3.2: 99th percentile end-to-end delay of VoIP flows.

Comparing to delay performance in Fig. 3.2, as we observe in Fig. 3.3, the 99th percentile jitter of uplink VoIP flows both in the default system and in our scheme under

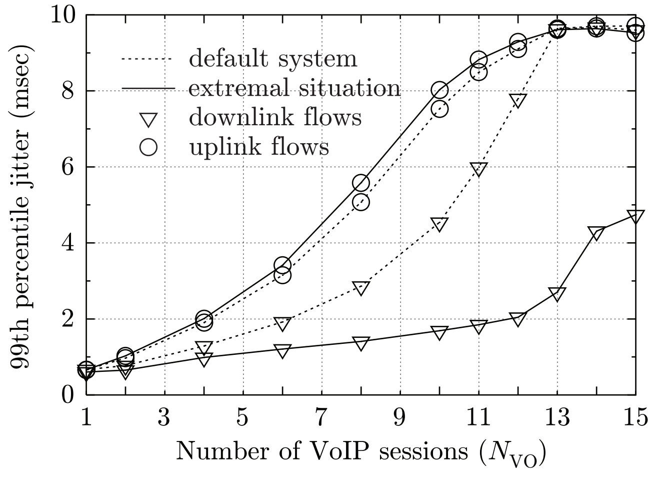


Figure 3.3: 99th percentile jitter of VoIP flows.

the extremal situation is higher than that of the downlink flows in most cases. Also, the jitter increases with N_{VO} in both systems.

Figure 3.4 shows the average packet loss ratio in the default system and in our scheme under the extremal situation. Note that the packet losses in Fig. 3.4 include the total packet losses due to QoS constraints at the application layer (i.e., at the decoder), due to the retransmission limitation at the MAC layer, and due to buffer overflows at stations. As illustrated in the figure, when $N_{\mathrm{VO}} < 6$, no packet losses occur. Packet loss ratio, however, increases with N_{VO} for $N_{\mathrm{VO}} > 6$. This figure shows that our scheme can accommodate 11 QoS guaranteed VoIP sessions at most, with the QoS requirements given in Table 3.2, which is the same capacity as in the default system.

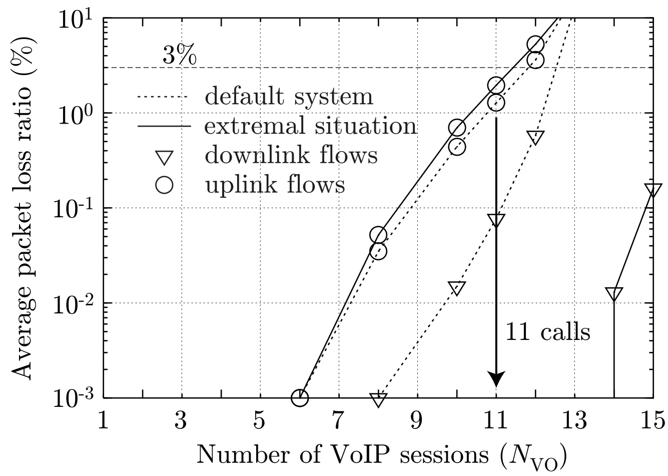


Figure 3.4: Average packet loss ratio of VoIP flows.

Now we have a look at packet losses due to QoS constraints. As stated in Section 3.4.2, VoIP packets whose end-to-end delay is not less than 150 msec or/and whose jitter is not less than 10 msec are discarded at the decoder. Table 3.3 shows the average packet loss ratios, observed at the decoder, of the uplink and downlink VoIP flows, in the default system and in our scheme under the extremal situation, where $P_{\text{loss}}^{(\text{Del})}$, $P_{\text{loss}}^{(\text{Jit})}$, and $P_{\text{loss}}^{(\text{Del\&Jit})}$ denote the average ratios of dropped packets only due to delay constraint, only due to jitter constraint, and due to both delay and jitter constraints, respectively (see Fig. 3.5). We observe that, when the number of VoIP flows is small (i.e., $N_{\text{VO}} \leq 12$), packet loss, at the decoder, occurs only due to jitter constraint, in both systems. When the number of VoIP flows is large (i.e., $N_{\text{VO}} \geq 13$), however, queueing delay of VoIP packets becomes large and thus packet loss occurs due to delay constraint, too. These results agree with the results in Figs. 3.2 and 3.3.

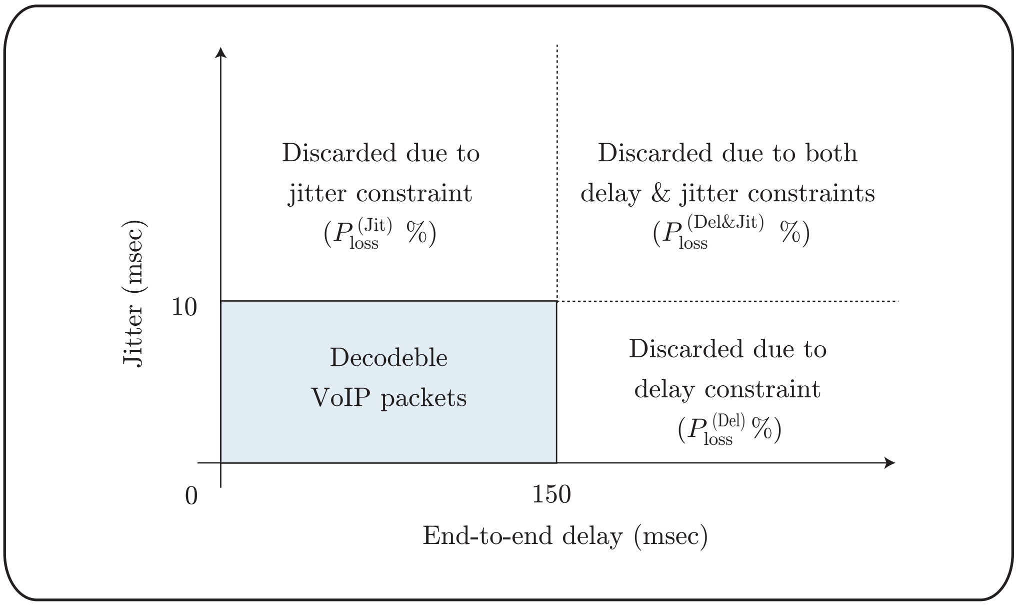


Figure 3.5: VoIP packet discard regions.

Next we have a look at packet losses due to the retransmission limitation at the MAC layer. With the simulation results, we found that, packet loss ratio due to the retransmission limitation was very small in both systems. For example, when $N_{\text{VO}} = 11$, it was only 2.7×10^{-4}% in the default system and 9.0×10^{-5}% in our scheme even it runs under the extremal situation. Therefore, rather than showing such packet loss ratios, we show frame collision probabilities. Figure 3.6 compares the overall frame collision probability P_c, as well as the average frame collision probabilities $P_c^{(\text{AP})}$ and $P_c^{(\text{WT})}$ at the AP and a wireless terminal. We observe that when $N_{\text{VO}} = 1$, frame collision probabilities are almost equal to zero and they increase with N_{VO} in both systems. We also observe that, when N_{VO} is large, frame collision probability at the AP in our scheme is smaller than that of in the default system and this leads to a smaller overall frame collision

Table 3.3: Average packet loss ratios of VoIP flows, observed at the decoder, due to QoS constraints ($P_{\text{loss}}^{(\text{Del})}$ %, $P_{\text{loss}}^{(\text{Jit})}$ %, $P_{\text{loss}}^{(\text{Del\&Jit})}$ %).

N_{VO}	Uplink flows		Downlink flows	
	Default system	Extremal situation	Default system	Extremal situation
8	(0.00, 0.03, 0.00)	(0.00, 0.05, 0.00)	(0.00, 0.00, 0.00)	(0.00, 0.00, 0.00)
10	(0.00, 0.44, 0.00)	(0.00, 0.70, 0.00)	(0.00, 0.05, 0.00)	(0.00, 0.00, 0.00)
11	(0.00, 1.28, 0.00)	(0.00, 1.96, 0.00)	(0.00, 0.08, 0.00)	(0.00, 0.00, 0.00)
12	(0.00, 3.61, 0.00)	(0.00, 5.28, 0.00)	(0.00, 0.58, 0.00)	(0.00, 0.00, 0.00)
13	(0.00, 17.18, 0.00)	(0.05, 16.17, 0.04)	(2.84, 9.98, 0.20)	(0.00, 0.00, 0.00)
15	(0.04, 26.15, 0.03)	(91.31, 0.54, 0.16)	(69.83, 0.00, 0.00)	(0.00, 0.16, 0.00)

probability in our scheme than that in the default system. We observed this interesting phenomenon in the legacy IEEE 802.11 DCF-based wireless LANs, and it can be explained as follows in the same manner as in Section 2.4.4.

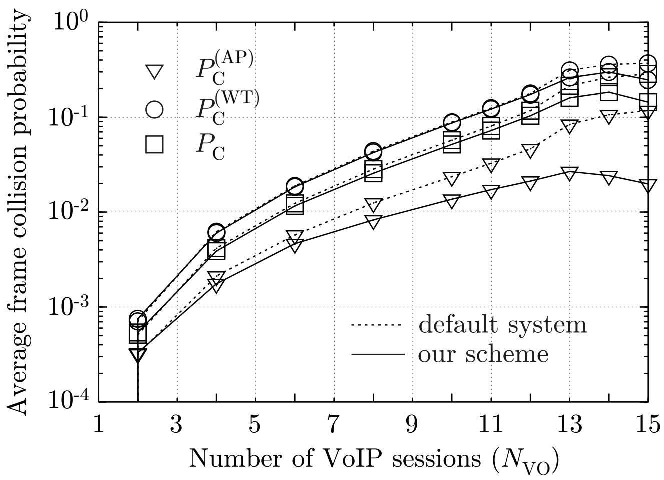


Figure 3.6: Average frame collision probability of VoIP flows.

Recall that when a frame transmitted by the AP collided with a frame transmitted by a wireless terminal, both the AP and the wireless terminal double their current contention windows. Since, in our scheme, $CW_{\text{min_D}}[i]$s at the AP is very smaller than to $CW_{\text{min_U}}[i]$s at the wireless terminal (in the extremal situation, $CW_{\text{min\_D}}[0] = 0$ and $CW_{\text{min\_U}}[0] = 7$), the AP can access to the wireless channel several times before the wireless terminal does. As a result we have $P_c^{(\text{AP})} < P_c^{(\text{WT})}$, and for large N_{VO}, the overall frame collision probability P_c of our scheme becomes smaller than that of the default system.

Next we examine the capacity for accommodating video flows. For QoS guaranteed N_{VO} (i.e., $N_{\text{VO}} \leq 11$), we consider two systems only with uplink video flows and only with downlink video flows, each of which coexists with N_{VO} VoIP sessions. The maximum number of QoS guaranteed video flows in each case are summarized in Table 3.4. We observe that our scheme can handle at least the same number of video flows as the default system can, even when it runs under the extremal situation. We then conclude that even in the most severe condition, our scheme can accommodate at least the same number of QoS guaranteed real-time flows as the default system can.

Table 3.4: Maximum number of QoS guaranteed video flows.

Number of VoIP flows (N_{VO})		0	1	3	5	7	9	10	11
Number of uplink video flows ($N_{\mathrm{VI\_U}}$)	Default system	7	7	5	4	3	1	1	0
	Extremal situation	7	7	5	4	3	1	1	0
Number of downlink video flows ($N_{\mathrm{VI\_D}}$)	Default system	8	7	5	4	3	1	1	0
	Extremal situation	9	7	5	4	3	1	1	0

3.4.4 Fairness in Best-Effort Traffic

We now examine the fairness issue in best-effort traffic. We assume that all best-effort flows follow TCP New Reno [19] with data segments of 1,000 bytes. To evaluate the performance of best-effort flows coexisting with real-time flows, we first have to find a proper combination for the numbers of VoIP and video flows. Otherwise, best-effort flows may not get chances to transmit their frames under the situation that the wireless bandwidth is fully occupied by real-time flows. Thus, with the results in Table 3.4, we first fix N_{VO}, $N_{\mathrm{VI\_D}}$, and $N_{\mathrm{VI\_U}}$ so that there is sufficient bandwidth for best-effort. In such a way, we fix $N_{\mathrm{VO}} = 5$ and $N_{\mathrm{VI\_D}} = N_{\mathrm{VI\_U}} = 1$, and we vary $N_{\mathrm{BE\_D}}$ from 1 to 10 and $N_{\mathrm{BE\_U}}$ from 1 to 5.

Recall that in our scheme, $CW_{\mathrm{min\_D}}[2]$ in (3.1) is determined with $M = N_{\mathrm{BE\_D}} + N_{\mathrm{BE\_U}}$, and $CW_{\mathrm{min\_D}}[i]$ and $CW_{\mathrm{max\_D}}[i]$ $(i = 0, 1)$ are set as shown in Table 3.1. Before discussing fairness in best-effort traffic, we shall confirm that QoS of real-time traffic is guaranteed. Table 3.5 shows the 99th percentile end-to-end delay, the 99th percentile jitter, and the average packet loss ratio of VoIP and video flows, where $N_{\mathrm{BE\_D}} = 10$ and $N_{\mathrm{BE\_U}} = 5$. We observe that QoS of real-time flows are guaranteed both in the default system and in our scheme, even though downlink real-time flows in our scheme receive a slightly preferential treatment, compared with the default system. The performance of uplink real-time flows, however, worsens slightly compared with the default system. The reason for this performance degradation is that our scheme gives higher priority to downlink flows by assigning smaller $CW_{\mathrm{min}}[i]$s at the AP, and thus uplink flows suffer some difficulties in accessing the wireless channel, compared with the default system.

We now discuss fairness in best-effort flows. Figures 3.7 and 3.8 illustrate fairness

Table 3.5: 99th percentile delay, 99th percentile jitter, and average packet loss ratio of real-time flows ($N_{\mathrm{VO}} = 5$, $N_{\mathrm{VI\_D}} = N_{\mathrm{VI\_U}} = 1$, $N_{\mathrm{BE\_D}} = 10$, $N_{\mathrm{BE\_U}} = 5$).

VoIP flows	Uplink		Downlink	
	Default system	Our scheme	Default system	Our scheme
Delay (msec)	58.10	65.48	56.88	53.99
Jitter (msec)	6.14	7.81	5.12	3.66
Packet loss ratio (%)	0.12	0.68	0.04	0.00

Video flows	Uplink		Downlink	
	Default system	Our scheme	Default system	Our scheme
Delay (msec)	64.39	83.80	63.77	61.31
Jitter (msec)	11.48	13.17	11.03	9.51
Packet loss ratio (%)	0.11	0.62	0.07	0.01

indices $FI_{\mathrm{BE\_D,BE\_U}}$ and Γ, respectively. In the default system, unfairness between uplink and downlink best-effort flows is observed when the total number $N_{\mathrm{BE\_D}} + N_{\mathrm{BE\_U}}$ of best-effort flows is large. On the other hand, in our scheme, both $FI_{\mathrm{BE\_D,BE\_U}}$ and Γ remain near one, regardless of the numbers of best-effort flows. Thus, our scheme greatly improves both per-flow fairness and fairness between uplink and downlink best-effort flows.

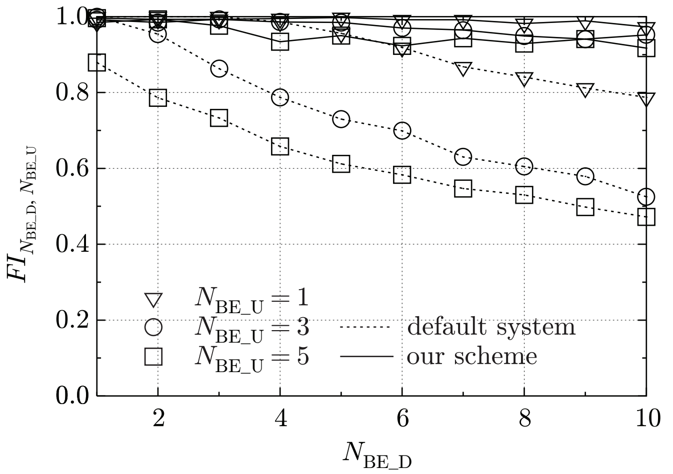


Figure 3.7: Per-flow fairness index $FI_{\mathrm{BE\_D,BE\_U}}$ of TCP flows ($N_{\mathrm{VO}}=5$, $N_{\mathrm{VI\_D}}=N_{\mathrm{VI\_U}}=1$).

In order to look at unfairness of the default system closely, we consider the packet loss probability due to buffer overflow of TCP flows at each station. We found that packet loss due to buffer overflow never occurred at wireless terminals. Thus, in Table 3.6, we show the probabilities of packet loss, i.e., data segment loss and ACK segment loss,

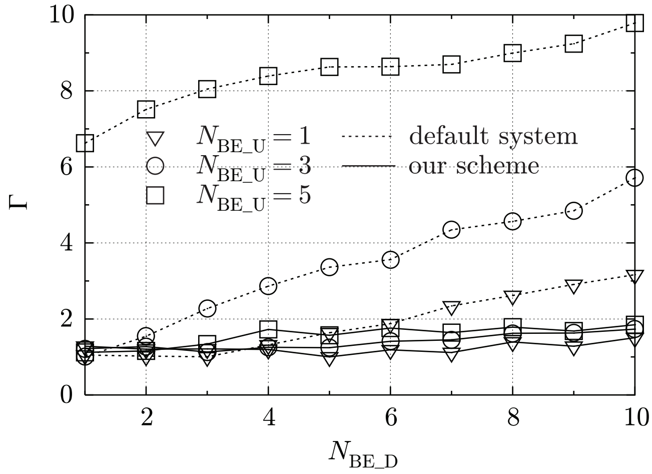


Figure 3.8: Uplink/downlink fairness index Γ of TCP flows ($N_{\mathrm{VO}}=5$, $N_{\mathrm{VI\_D}}=N_{\mathrm{VI\_U}}=1$).

at AC_2 in the AP, where the result in our scheme is also shown for reference. When

Table 3.6: Packet loss probability of TCP flows due to buffer overflow at stations ($N_{\mathrm{VO}} = 5$, $N_{\mathrm{VI\_D}} = N_{\mathrm{VI\_U}} = 1$).

$N_{\mathrm{BE\_D}}$	$N_{\mathrm{BE\_U}}$	Data segment loss probability		ACK segment loss probability	
		Default system	Our scheme	Default system	Our scheme
1	1	0.00	0.00	0.00	0.00
	5	0.09	0.00	0.09	0.00
5	1	0.01	0.00	0.01	0.00
	5	0.12	0.02	0.09	0.01
10	1	0.03	0.02	0.02	0.01
	5	0.19	0.03	0.16	0.02

$N_{\mathrm{BE\_D}} = N_{\mathrm{BE\_U}} = 1$, the packet loss does not occur. In our simulation experiment, the buffer size of AC_2 is set to be 100 packets and the advertised window of TCP flows is set to be 20 packets, and thus, no data segment loss occurs if $N_{\mathrm{BE\_D}} + N_{\mathrm{BE\_U}}$ is small. As explained in Section 2.4.5, in such a situation, TCP connections reach the steady state and the self-clocking effect emerges [36]. Thus TCP ACK segments regulate the transmission rate of data segments, and therefore the bandwidth is shared almost equally among TCP flows. When $N_{\mathrm{BE\_D}}+N_{\mathrm{BE\_U}}$ is large, however, both the data segment loss and the ACK segment loss are observed in the default system. As a result, the throughput of downlink TCP flows in the default system decreases. In other words, the default system can achieve per-flow fairness among best-effort flows only when the buffer size is large enough to prevent packet losses. Note that a large buffer leads to large queueing delay and it may result in throughput degradation of TCP flows because the throughput performance of TCP depends on RTT.

Finally, Fig. 3.9 compares the total throughput of best-effort flows. We observe that the total best-effort throughput in our scheme is greater than that in the default system. Thus our scheme not only achieves fairness but also keeps high bandwidth utilization for best-effort traffic. Readers may wonder why the performance of our scheme does not degrade even when there are many best-effort flows. Recall that our scheme changes $CW_{\min}$s and $CW_{\max}$s only at the AP and it does not make any modification at wireless terminals. Thus, as discussed in Section 3.4.3, the frame collision probability in our scheme becomes smaller than that of the default system and this results in a good throughput performance. We observed this phenomenon in legacy IEEE 802.11 DCF wireless LANs (see Section 2.4.4). With the results in Fig. 3.9, we conclude that our scheme works well without any serious side effects.

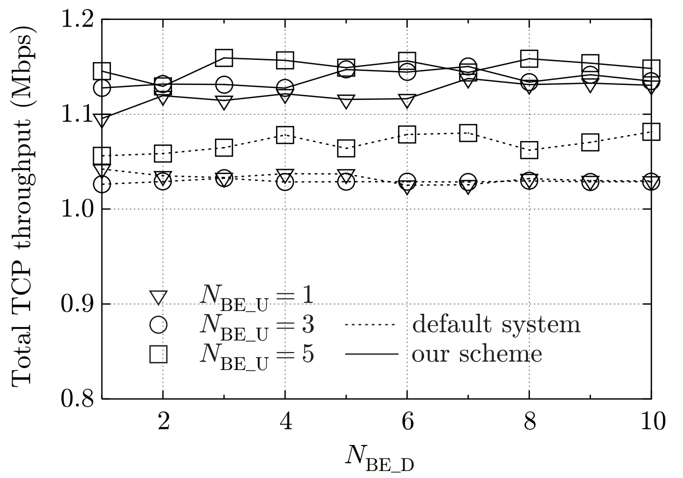


Figure 3.9: Total system throughput of TCP flows ($N_{\mathrm{VO}} = 5$, $N_{\mathrm{VI\_D}} = N_{\mathrm{VI\_U}} = 1$).

3.5 Conclusion

This chapter discussed a fairness issue between uplink and downlink best-effort flows in QoS-oriented IEEE 802.11e EDCA-based wireless LANs. To achieve max-min fairness among best-effort flows, we presented a dynamic contention window control scheme, which adjusts the minimum and maximum contention window sizes at the AP. In our scheme, the AP first determines the optimal $CW_{\min}$ for the best-effort access category according to the number of best-effort flows it handles. It then adjusts the $CW_{\min}$s and $CW_{\max}$s of higher priority access categories so as to guarantee QoS requirements for real-time traffic. In such a way our scheme gives sufficient bandwidth for real-time flows and then try to share the remaining bandwidth among all best-effort flows. Note that our scheme does not require any modification at wireless terminals. Simulation results showed that our scheme achieved fairness between uplink and downlink best-effort flows, as well as per-flow fairness among best-effort flows, while guaranteeing QoS requirements for real-time traffic.

Chapter 4

Per-Flow Fairness in Multi-Rate Wireless LANs

THIS chapter considers *performance anomaly* in multi-rate IEEE 802.11 wireless LANs, where stations with the lowest data transmission rate regulate the throughput of all other stations and the total system throughput degrades badly.

As we stated in Chapter 1, the IEEE 802.11 physical layer extensions support multiple data transmission rates by employing different modulations and channel coding schemes. For example, the IEEE 802.11b [84] physical layer extension provides four data transmission rates: 1, 2, 5.5, and 11 Mbps. Each station individually selects an appropriate data transmission rate according to its channel condition. Stations in a good channel condition typically employ a higher data transmission rate, while stations in a poor channel condition do a lower data transmission rate so as to prevent frequent frame losses due to transmission errors on the wireless channel [44, 67].

Generally, stations with a high data transmission rate (called high-rate stations hereafter) can achieve higher throughput than stations with a low data transmission rate (called low-rate stations hereafter). When stations with different data transmission rates coexist in a multi-rate wireless LAN, however, stations with the lowest data transmission rate regulate the throughput of all other stations and it is forced to be the same as the throughput of stations with the lowest data transmission rate. As a result, the total system throughput degrades badly in multi-rate wireless LANs [31, 74]. This phenomenon is well-known as performance anomaly. Note that the performance anomaly is caused by the following two facts:

(i) the default MAC protocol equally gives all stations the channel access right, i.e., the per-station fairness property, and

(ii) once acquiring the channel, low-rate stations occupy the wireless channel for a longer time than high-rate stations do, because in IEEE 802.11 wireless LANs, channel access time is not fixed.

The total system throughput can be improved readily by restraining low-rate stations from accessing the wireless channel. It is clear, however, that this would lead to serious unfairness in individual throughput between high- and low-rate stations. Thus, improving the total system throughput and achieving throughput fairness among individual stations are two conflicting goals in multi-rate wireless LANs, and we have to compromise one way or another. As we showed in Section 1.4.2, the proportional fairness [43] which is equivalent to air-time fairness, i.e., fairness in channel occupancy time, is a reasonable compromise between these two goals.

So far, numerous research efforts have been devoted to alleviate the performance anomaly [2, 9, 12, 16, 18, 24, 32, 39, 40, 45, 64, 71, 73, 74, 79, 92]. Most of them, however, focus on air-time fairness only in *station-level*, i.e., they aim to make the channel occupancy times of respective stations equal. Since APs handle multiple downlink flows, however, air-time fairness in station-level leads to serious performance degradation of downlink flows [62] (see Chapters 2 and 3 for details).

Aiming at air-time fairness in *flow-level*, this chapter considers a dynamic contention window control mechanism that works well for both uplink and downlink flows. In our scheme, flows are classified into several classes according to their data transmission rates, and at APs, downlink flows in respective classes are stored in separate buffers, as in the IEEE 802.11e [85] protocol. Further our scheme assigns different minimum contention window sizes $CW_{\min}$s to those classes according to their data transmission rates and target packet rates. Through simulation experiments, we show the effectiveness of our scheme even when there are many downlink flows at the AP, i.e., our scheme alleviates the unfairness in multi-rate wireless LANs.

The rest of this chapter is organized as follows. Section 4.1 overviews the performance anomaly and Section 4.2 reviews related work. In Section 4.3, we describe our contention window control mechanism and in Section 4.4, our scheme is evaluated through simulation experiments with UDP and TCP flows. Finally, the conclusion is given in Section 4.5.

4.1 Problem Overview

To illustrate the performance anomaly [31], we consider an IEEE 802.11b-based wireless LAN with two wireless terminals, WT_0 and WT_1, and one AP. WT_0 and WT_1 transmit data frames towards the AP at data transmission rates of $R^{[0]}$ Mbps and $R^{[1]}$ Mbps, respectively. Simulation scenario is depicted in Fig. 4.1. The length of MSDUs is set to be 1,000 bytes. We conduct a simulation experiment for 300 seconds by varying $R^{[0]}$ and $R^{[1]}$ as shown in Table 4.1.

Figure 4.2 shows the average throughput of WT_0 and WT_1, calculated every 10 seconds. We observe that

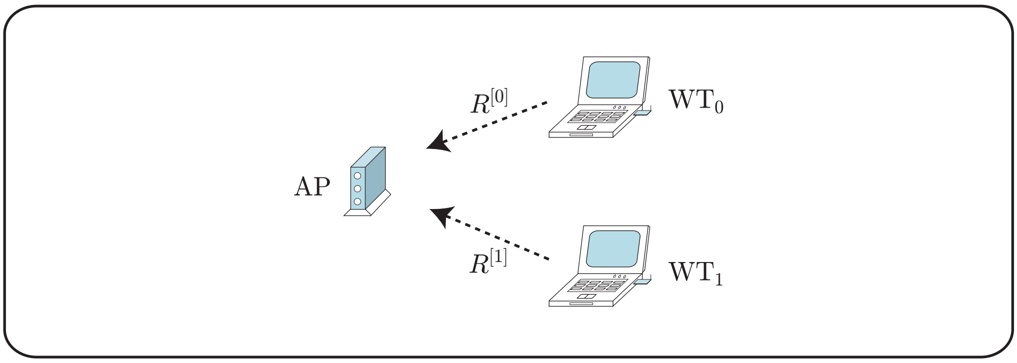


Figure 4.1: Simple scenario of a multi-rate wireless LAN.

Table 4.1: Simulation scenario for performance anomaly.

Time period / Data rate	[0,100)	[100,200)	[200,300)
$R^{[0]}$ (Mbps)	11	1	1
$R^{[1]}$ (Mbps)	11	11	1

(i) the average throughput of those two wireless terminals are always identical and

(ii) in the interval [100,200), the average throughput of each wireless terminal is about 0.73 Mbps, which is less than 30% of the average throughput of 2.63 Mbps in the interval [0,100).

Note that phenomenon (i) is a key feature in the IEEE 802.11 DCF protocol: all stations can acquire the transmission right equally in the saturated situation because CSMA/CA is not influenced by data transmission rates of stations. As a result, all stations share the bandwidth fairly when the average lengths of payloads are identical.

Next we consider phenomenon (ii). Since the average data transmission rates in the intervals [0,100) and [100,200) are equal to 11 Mbps and 6 Mbps, respectively, readers who are not familiar with the performance anomaly might expect that the throughput in the interval [100,200) would be about half of that in the interval [0,100). In reality, however, more than 70% of throughput degradation occurs. To understand this phenomenon, we approximately analyze the average throughput.

Figures 4.3(a) and 4.3(b) show frame formats of a data frame and an ACK frame, respectively, defined in the IEEE 802.11b protocol. Let $T_i^{(\mathrm{DATA})}$ μsec denote the time needed to transmit a data frame with MSDU payload of L_i bytes at transmission rate of R_i Mbps. Also, let $T^{(\mathrm{ACK})}$ μsec denote the time needed to transmit an ACK frame. We

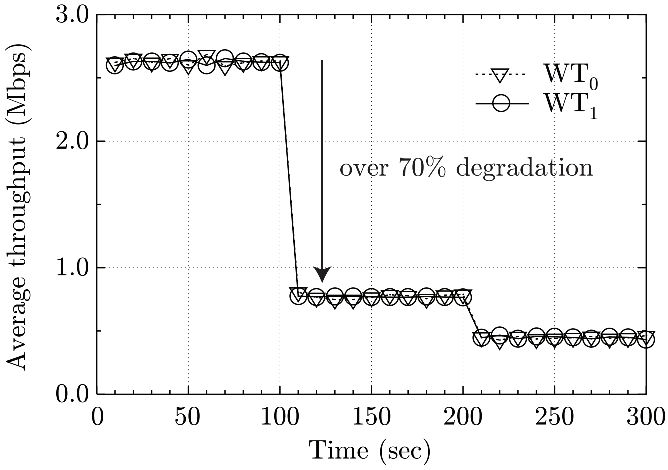


Figure 4.2: Performance anomaly in a multi-rate wireless LAN.

then have

$$T_i^{(\mathrm{DATA})} = 192 + \frac{8(28 + L_i)}{R_i} \quad (\mu\mathrm{sec}),$$
$$T^{(\mathrm{ACK})} = 192 + \frac{8 \times 14}{BasicRate} = 304 \quad (\mu\mathrm{sec}).$$

Together with Fig. 4.4, the channel occupancy time (called air-time) Z_i in transmitting an MSDU of L_i bytes at transmission rate of R_i Mbps is given by

$$\begin{aligned} Z_i &= T_i^{(\mathrm{DATA})} + SIFS + T^{(\mathrm{ACK})} + DIFS \\ &= 556 + \frac{8(28 + L_i)}{R_i} \quad (\mu\mathrm{sec}). \end{aligned} \tag{4.1}$$

In particular, when two wireless terminals, WT_i and WT_j with data transmission rates of R_i Mbps and R_j Mbps, respectively, coexist, the average throughput $\theta^{[\mathrm{WT}_k]}_{(R_i,R_j)}$ Mbps of WT_k $(k = i, j)$ is given by

$$\theta^{[\mathrm{WT}_k]}_{(R_i,R_j)} \approx \frac{1}{2} \cdot \frac{8L_k}{\dfrac{CW_{\min} \times SlotTime}{4} + \dfrac{1}{2}(Z_i + Z_j)},$$

where we assume that no collisions happen and the mean length of contention periods is given by $CW_{\min} \times SlotTime/4$.

The result indicates that when two wireless terminals with R_i and R_j coexist and the lengths of their MSDUs are identical, the average throughput $\Theta_{(R_i,R_j)}$ of each wireless terminal is given in a form:

$$\Theta_{(R_i,R_j)} = \frac{a}{b + \dfrac{1}{R_i} + \dfrac{1}{R_j}},$$

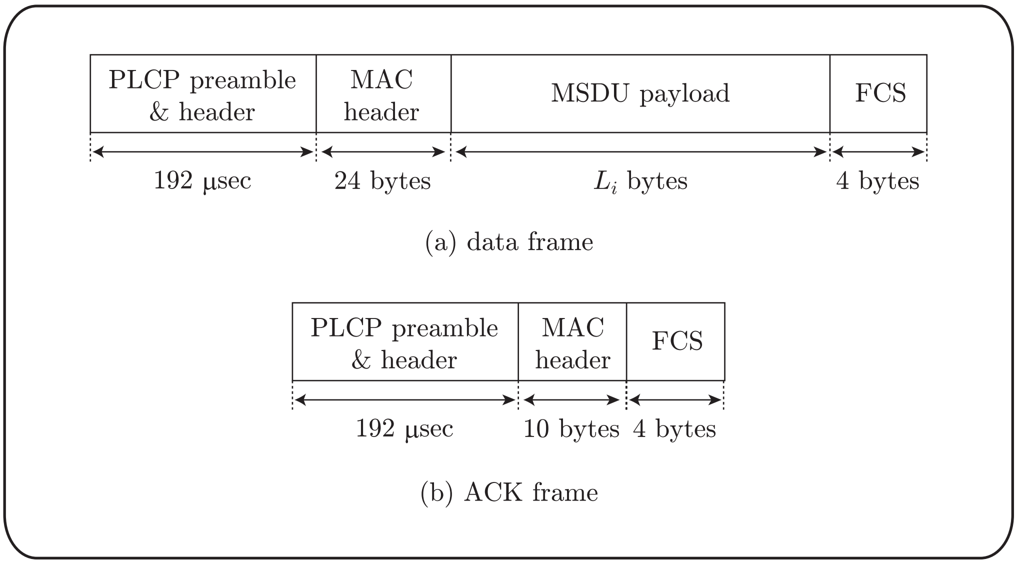


Figure 4.3: Frame formats defined in the IEEE 802.11b protocol.

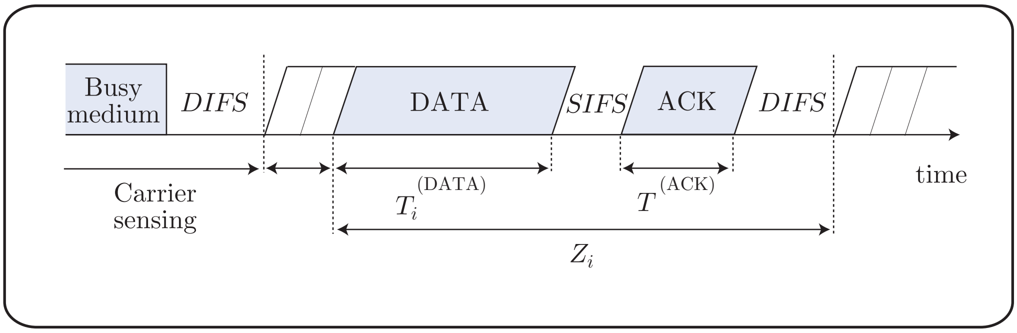


Figure 4.4: Channel occupancy time under the basic scheme.

where a and b are positive constants, and in the scenario of the above simulation, they are approximately given by 0.973 and 0.173, respectively. Thus we have $\Theta_{(11,11)} \approx 2.74$ Mbps and $\Theta_{(1,11)} \approx 0.77$ Mbps, which is consistent with Fig. 4.2. In summary, when wireless terminals with different data transmission rates coexist and the lengths of their MSDUs are identical, the wireless terminals have equal opportunities to access the channel and the throughput is given in terms of the harmonic mean of their data transmission rates, which lead to the performance anomaly.

4.2 Related Work

As solutions of the performance anomaly, several methods have been proposed so far for achieving air-time fairness in multi-rate wireless LANs. They are classified into two types;

(i) adjusting channel access probabilities of stations by varying parameter values defined in CSMA/CA mechanism [2, 12, 32, 39, 45, 74, 92] and

(ii) performing frame aggregation/fragmentation or allowing multiple data frame transmission to high-rate stations [18, 64, 71, 73, 92].

Note that all of above schemes aim to air-time fairness in station-level, i.e., they try to make the channel occupancy times of respective stations equal. Note that an AP handles multiple downlink flows and the bandwidth that the AP acquires is shared with those downlink flows. Therefore air-time fairness in station-level will lead to a serious throughput degradation of downlink flows [62].

Tan et al. propose a scheme called TBR (Time-based regulator) which can be used for both uplink and downlink flows [79]. In the TBR scheme, each wireless terminal implements a TBR agent and communicates with the main TBR agent employed at the AP. In order to achieve air-time fairness in flow-level, packets in each flow stored in an individual buffer and the main TBR agent reschedules the transmission order of frames by employing a leaky bucket algorithm and informs other TBR agents.

In [16], Dunn et al. propose shortening frame size based on the number of uplink and downlink flows and their data transmission rates. They aim to achieve air-time fairness separately among uplink flows and among downlink flows. Note that the scheme proposed in [16] ignores protocol overhead, so that air-time fairness cannot be achieved in a strict sense. Further it assigns short frame size to low-rate stations, so that the total system throughput degrades due to the overhead. Another possible way to achieve air-time fairness is to assign long frame size to high-rate stations. In this kind of schemes, however, the throughput of each flows fluctuates so much and therefore it cannot achieve fairness in a short period of time (see Fig. 2.11).

4.3 Dynamic Contention Window Control

To alleviate the performance anomaly, we consider a dynamic contention window control mechanism that works for both uplink and downlink flows. Our scheme aims to achieve air-time fairness in flow-level by controlling $CW_{\min}$s, where a flow is defined as a sequence of frames with the same sender and receiver station MAC addresses.

4.3.1 Flow Classification

We refer to flows whose data transmission rate is equal to R_i Mbps ($i = 1, 2, \dots, i_{\max}$) as flows in class i, where $R_1 > R_2 > \cdots > R_{i_{\max}}$. We also refer to wireless terminals transmitting class i flows as wireless terminals in class i (see Fig. 4.5(a)). Unlike wireless terminals, APs handle multiple downlink flows destined to wireless terminals in different classes. Thus flows in different classes can exist at APs. Therefore, at the AP, we provide $i_{\max}$ buffers, each of which stores data frames in a specific class (see Fig. 4.5(b)).

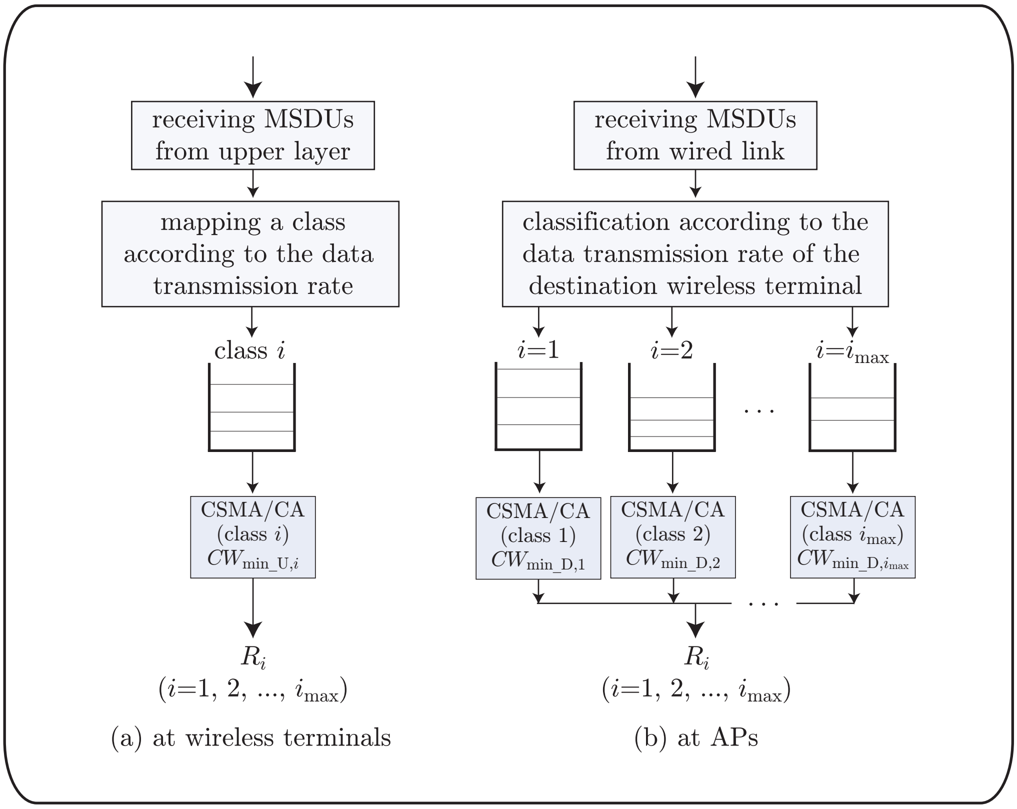


Figure 4.5: Behavior of our scheme.

We then differentiate among flows in different classes by setting different $CW_{\min}$s to different classes in such a way that all flows share the wireless channel equally in terms of the channel occupancy time. Note that each class at the AP performs CSMA/CA individually using the assigned $CW_{\min}$. If backoff counters of more than one class in the AP expire at the same time slot, among them the class with the highest data transmission rate is allowed to transmit its data frame, while other classes behave as if a collision occurred. Note that this mechanism is very similar to the IEEE 802.11e [85] protocol, which provides differentiated QoS in wireless LANs.

4.3.2 Mean Field Approximation Analysis

A simple mean field approximation analysis for single-rate IEEE 802.11 wireless LANs is presented in Section 2.3.2, by assuming $CW_{\max} = \infty$ and $r_{\max} = \infty$. In this subsection we extend it to multi-rate IEEE 802.11 wireless LANs.

Let $N_{\text{U},i}$ and $N_{\text{D},i}$ $(i = 1, 2, \ldots, i_{\max})$ denote the number of uplink flows (i.e., the number of wireless terminals) in class i and the number of class i downlink flows that the AP handles, respectively. We denote $CW_{\min}$ for class i uplink flows by $CW_{\min\_\mathrm{U},i}$ and $CW_{\min}$ for class i downlink flows by $CW_{\min\_\mathrm{D},i}$. We define $\lambda_{\mathrm{U},i}$ ($i = 1, 2, \ldots, i_{\max}$) as the steady-state probability that a wireless terminal in class i accesses the channel in a randomly chosen slot of contention periods. Also $\lambda_{\mathrm{D},i}$ ($i = 1, 2, \ldots, i_{\max}$) is defined as the steady-state probability that the AP accesses the channel for the transmission of a class i flow in a randomly chosen slot of contention periods. Following exactly the same approach as in Section 2.3.2, we obtain for $X = \mathrm{U}, \mathrm{D}$

$$\lambda_{X,i} = \frac{2(1 - 2\alpha_{X,i})}{(CW_{\min\_X,i} + 1)(1 - \alpha_{X,i}) + 1 - 2\alpha_{X,i}}, \tag{4.2}$$

where $\alpha_{\mathrm{U},i}$ and $\alpha_{\mathrm{D},i}$ denote the transmission failure probabilities of a wireless terminal in class i and a class i flow transmitted by the AP, respectively. Note that

$$\alpha_{\mathrm{U},i} = 1 - \frac{1}{1 - \lambda_{\mathrm{U},i}} \prod_{k=1}^{i_{\max}} (1 - \lambda_{\mathrm{U},k})^{N_{\mathrm{U},k}} (1 - \lambda_{\mathrm{D},k}), \tag{4.3}$$

$$\alpha_{\mathrm{D},i} = 1 - \prod_{k=1}^{i_{\max}} (1 - \lambda_{\mathrm{U},k})^{N_{\mathrm{U},k}} \prod_{k=1}^{i-1} (1 - \lambda_{\mathrm{D},k}). \tag{4.4}$$

Since a station with minimum contention window of $CW_{\min}$ can successively transmit $1 + 1/CW_{\min}$ frames on average (see Section 2.3.1), the average number $\Theta_{\mathrm{U},i}$ (resp. $\Theta_{\mathrm{D},i}$) of frames in class i flows transmitted by a wireless terminal (resp. the AP) in a unit time is given by

$$\Theta_{X,i} = \left(1 + \frac{1}{CW_{\min\_X,i}}\right) \lambda_{X,i}(1 - \alpha_{X,i}), \tag{4.5}$$

and the corresponding channel occupancy time is given by

$$Z_{X,i} = \Theta_{X,i} Z_i, \tag{4.6}$$

for $X = \mathrm{U}, \mathrm{D}$, where Z_i is given in (4.1).

4.3.3 Quasi-Optimal $CW_{\min}$ to Achieve Air-Time Fairness

To achieve air-time fairness in flow-level, the mean channel occupancy time $Z_{\mathrm{U},i}$ of class i uplink flows should be identical regardless of classes, and the total channel occupancy time $Z_{\mathrm{D},i}$ of class i downlink flows should be $N_{\mathrm{D},i}$ times as large as the mean channel occupancy time $Z_{\mathrm{U},i}$ of class i uplink flows. In this chapter, we fix $CW_{\min\_\mathrm{U},1}$ to the default $CW_{\min}$ value (i.e., 31 in the IEEE 802.11b wireless LANs) and consider adjusting $CW_{\min\_\mathrm{U},i}$ ($i = 2, 3, \ldots, i_{\max}$) and $CW_{\min\_\mathrm{D},i}$ ($i = 1, 2, \ldots, i_{\max}$) in such a way that the following equations hold.

$$\frac{Z_{\mathrm{U},i}}{Z_{\mathrm{U},1}} = 1, \quad i = 2, 3, \ldots, i_{\max}, \tag{4.7}$$

$$\frac{Z_{\mathrm{D},i}}{Z_{\mathrm{U},i}} = N_{\mathrm{D},i}, \quad i = 1, 2, \ldots, i_{\max}. \tag{4.8}$$

When $N_{\mathrm{U},i}$, $N_{\mathrm{D},i}$ ($i = 1, 2, \ldots, i_{\max}$), and $CW_{\mathrm{min\_U},1}$ are given, other CW_{min}s can be obtained numerically using (4.7) and (4.8), and (4.2), (4.3), (4.4), (4.5), and (4.6).

4.4 Simulation Experiments and Results

In this section, we evaluate the performance of our scheme through simulation experiments.

4.4.1 Simulation Model

We employ an IEEE 802.11b-based wireless LAN with the default MAC parameters shown in Table 1.1. For the simplicity, we assume that there exist only flows in class 1 and class 4. Figure 4.6 shows the network topology used in the simulation experiments. Note that $R_1 = 11$ Mbps and $R_4 = 1$ Mbps. Note also that *BasicRate* is set to be 1 Mbps and *RTSThreshold* is set to be 3,000 bytes. Each wired link has bandwidth of 100 Mbps and propagation delay of 25 msec. The size of each buffer is set to be 100 packets. Simulation time for each scenario is set to be 2,000 seconds.

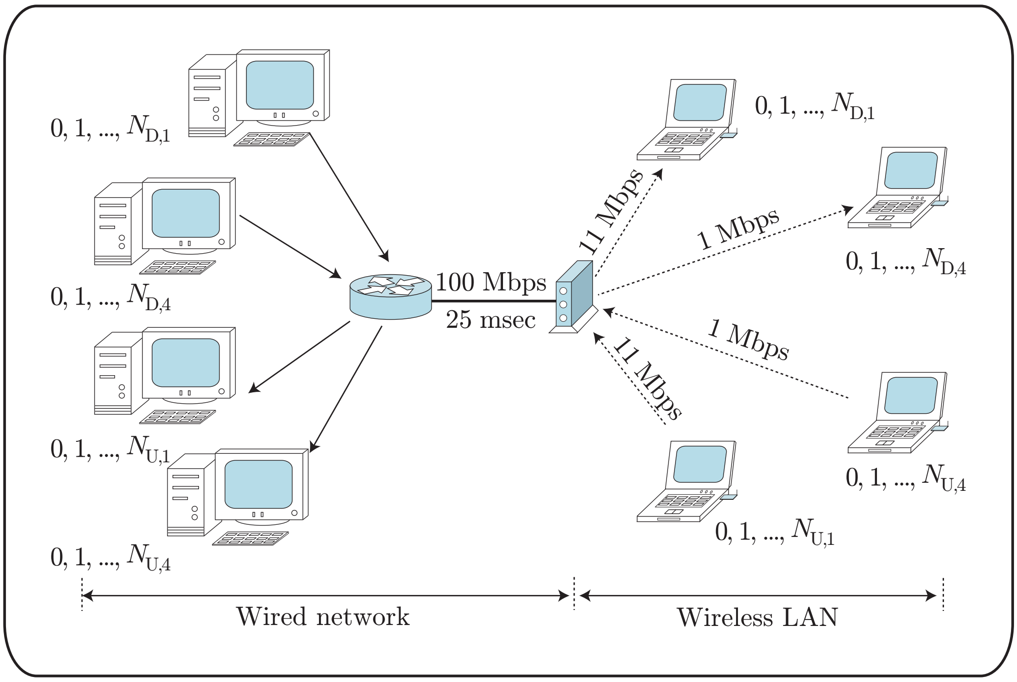


Figure 4.6: Network topology for simulation experiments.

4.4.2 Performance Measures

The performance of our scheme is evaluated in terms of the total system throughput and fairness index FI proposed by Jain et al. [37]. Let $Z_{\mathrm{U},i}^{[k]}$ (resp. $Z_{\mathrm{D},i}^{[k]}$) denote the channel occupancy time of the kth uplink (resp. downlink) flow in class i. FI is then given by

$$FI = \frac{\left(\sum_{X=\mathrm{U,D}} \sum_{i=1,4} \sum_{k=1}^{N_{X,i}} Z_{X,i}^{[k]} \right)^2}{N \sum_{X=\mathrm{U,D}} \sum_{i=1,4} \sum_{k=1}^{N_{X,i}} \left(Z_{X,i}^{[k]} \right)^2},$$

where $N = N_{\mathrm{U},1} + N_{\mathrm{U},4} + N_{\mathrm{D},1} + N_{\mathrm{D},4}$. Note that $0 < FI \leq 1$ and $FI = 1$ when channel occupancy times of all flows are identical.

4.4.3 Performance Evaluation for UDP Flows

In this subsection, we evaluate our scheme for UDP flows. The length of all segments are assumed to be 1,000 bytes and they are generated by a CBR traffic source at 10 Mbps, i.e., saturated situation.

Recall that in the default IEEE 802.11 system, $CW_{\mathrm{min\_}X,i} = 31$ ($X = \mathrm{U,D}$ and $i = 1,4$), while in our scheme, $CW_{\mathrm{min\_U},1} = 31$ and other CW_{min}s are determined numerically as explained in Section 4.3.3. We confirm that quasi-optimal CW_{min}s to achieve air-time fairness are less sensitive to numbers of uplink flows [2]. For example, $(CW_{\mathrm{min\_U},1}, CW_{\mathrm{min\_U},4})$ = (31,192) and (31,190) for $(N_{\mathrm{U},1}, N_{\mathrm{U},4})$ = (1,1) and (10,10), respectively. Thus in our scheme, we always set $CW_{\mathrm{min\_U},4}$ to be 192 and determine $CW_{\mathrm{min\_D},1}$ and $CW_{\mathrm{min\_D},4}$ dynamically by assuming $N_{\mathrm{U},1} = N_{\mathrm{U},4} = 1$. In this way, for example, we obtain $(CW_{\mathrm{min\_D},1}, CW_{\mathrm{min\_D},4})$ = (32,192), (13,66), (9,41), and (6,22) for $(N_{\mathrm{D},1}, N_{\mathrm{D},4})$ = (1,1), (3,3), (5,5), and (10,10), respectively.

Figure 4.7 illustrates the simulation results of air-time fairness index FI of UDP flows. As we see in Fig. 4.7, FI in the default system is always less than one. In our scheme, however, FI is almost equal to one for any number of flows. Thus our scheme achieves air-time fairness in flow-level in any case.

Figure 4.8 compares the total system throughput of UDP flows. With the results in Fig. 4.8, we observe that the total system throughput in the default system is very small. In our scheme, however, the total system throughput is greatly improved by 210% to 240%. Thus our scheme is valid and works well for UDP traffic flows.

4.4.4 Performance Evaluation for TCP flows

We next evaluate our scheme for TCP flows, assuming that all TCP flows follow TCP new Reno [19] with segment size of 1,000 bytes. Note that delayed-ACK is not

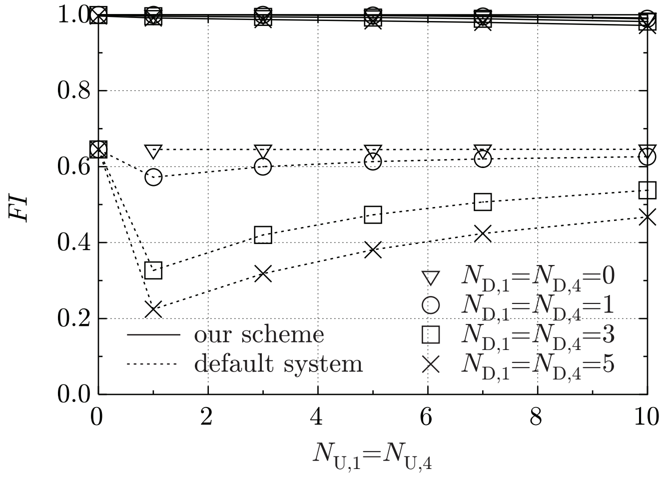


Figure 4.7: Air-time fairness index FI of UDP flows.

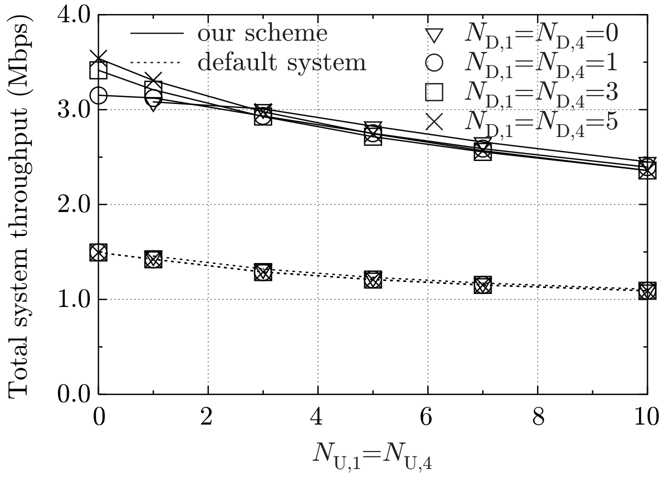


Figure 4.8: Total system throughput of UDP flows.

implemented and advertised window is set to be 20 packets.

Unlike UDP, TCP flows consist of TCP ACK segment and data segment flows. For example, when there exist $N_{\mathrm{D},i}$ TCP downlink flows and $N_{\mathrm{U},i}$ TCP uplink flows, the AP handles $N_{\mathrm{U},i}$ TCP ACK segment flows as well as $N_{\mathrm{D},i}$ TCP data segment flows. As we discussed in Section 2.4.5, in such a case, the target packet rate of class i downlink flows in (4.8) is given by the total number of those flows. Simulation results of air-time fairness index FI of TCP flows is shown in Fig. 4.9. We observe that our scheme achieves higher FI comparing to that of the default system.

Finally, the total system throughput of TCP flows for the default IEEE 802.11 MAC protocol and our scheme is depicted in Fig. 4.10. We observe that our scheme achieves higher system throughput. Thus our scheme works well for TCP flows, too.

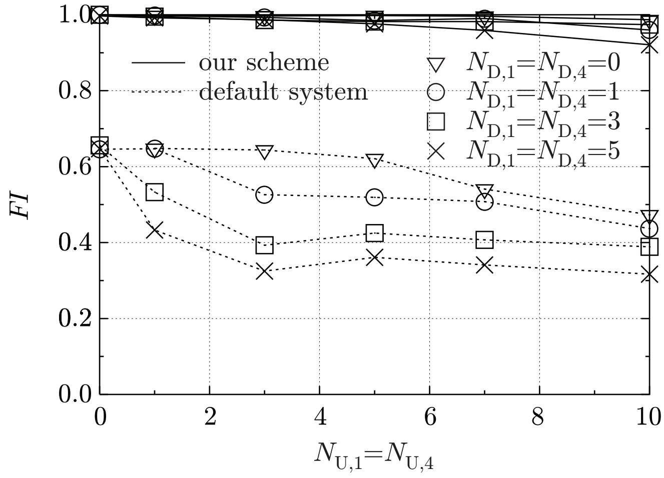


Figure 4.9: Air-time fairness index *FI* of TCP flows.

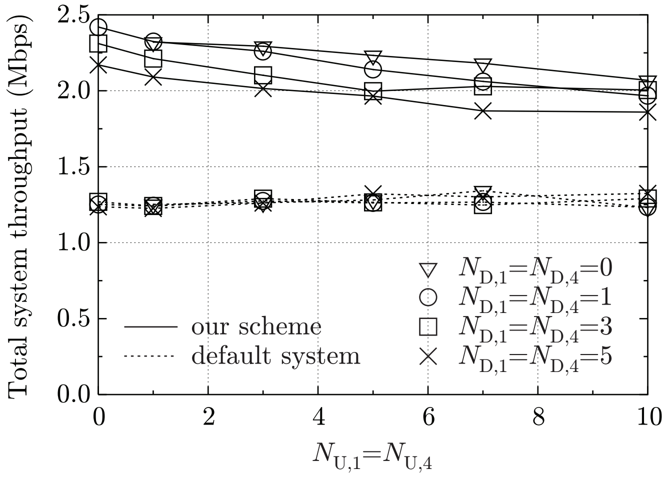


Figure 4.10: Total system throughput of TCP flows.

4.5 Conclusion

This chapter discussed the performance anomaly in multi-rate IEEE 802.11 wireless LANs, where stations with the lowest data transmission rate regulate the throughput of all other stations and the total system throughput degrades badly. In this chapter, we presented a dynamic window control of uplink and downlink flows in order to achieve air-time fairness in flow-level. In our scheme, flows are classified into several classes according to their data transmission rates, and at APs, downlink flows in respective classes are stored in separate buffers. Further our scheme assigns different $CW_{\min}$s to those classes according to their data transmission rates and the target packet rates. Through simulation experiments, we showed the excellent performance of our scheme.

Chapter 5

Conclusions

THROUGHOUT this dissertation, we discussed per-flow fairness in IEEE 802.11-based wireless LANs. We considered fairness in both single- and multi-rate wireless LANs, and presented simple yet effective window control mechanisms to improve the fairness among flows. We now summarize all of the results and observations which we obtained through the study. At the end of this chapter, we also summarize some future works and implementation issues of our schemes.

In Chapter 1, we first discussed the background and motivations behind this research. We then provided an introduction to the IEEE 802.11 protocol, especially we focused on the CSMA/CA mechanism and explained the per-station fairness property in the IEEE 802.11 MAC protocol. We then introduced some key features and challenges of IEEE 802.11-based wireless LANs. In Chapter 1, we also presented an overview of the IEEE 802.11e EDCA protocol, which was standardized to support QoS in IEEE 802.11-based wireless LANs.

Furthermore, Chapter 1 presented a brief outline of some fairness concepts such as max-min fairness and proportional fairness. At the end of Chapter 1, we discussed the fairness in wireless LANs of wired-cum-wireless networks, in which wireless LANs are the bottleneck link along the path of all flows passing through them. We showed that the max-min fairness criterion works well and achieves per-flow fairness in single-rate wireless LANs, while the proportional fairness is reasonable for multi-rate wireless LANs. Moreover, it is known that the proportional fairness is equivalent to the air-time fairness, i.e., the fairness in channel occupancy time, in wireless LANs.

With the basic concepts and ideas summarized in Chapter 1, we discussed three key fairness issues occurred in IEEE 802.11 wireless LANs in Chapters 2, 3, and 4. We first considered a per-flow fairness issue in single-rate wireless LANs in Chapter 2. Next in Chapter 3, we focused on a per-flow fairness issue in QoS-oriented wireless LANs, and at last in Chapter 4, we addressed a per-flow fairness issue in multi-rate wireless LANs.

In Chapter 2, we showed that the per-station fairness property of the CSMA/CA mechanism degrades the throughput of downlink flows badly comparing to that of uplink flows in legacy IEEE 802.11 wireless LANs, when APs handle multiple downlink flows. To achieve per-flow fairness, we presented a dynamic contention window control scheme, where the minimum contention window $CW_{\min}$ of an AP is controlled according to the target packet rate ratio R^* between uplink and downlink flows. Note that our scheme does not require any modification at wireless terminals. We conducted various simulation experiments with UDP and TCP flows, and demonstrated that our scheme can significantly ameliorate max-min fairness.

In Chapter 3, we discussed a per-flow fairness issue occurred among best-effort flows in single-rate IEEE 802.11e EDCA-based wireless LANs. To achieve max-min fairness among best-effort flows, we considered a dynamic contention window control scheme, which adjusts the minimum and maximum contention window sizes at the AP. In our scheme, the AP first determines the optimal $CW_{\min}$ for the best-effort access category according to the number of downlink best-effort flows it handles. It then adjusts the $CW_{\min}$s and $CW_{\max}$s of higher priority access categories so as to guarantee QoS requirements for real-time traffic. In this way, our scheme gives sufficient bandwidth for real-time flows and then try to share the remaining bandwidth among all best-effort flows. Note that our contention window control scheme does not require any modification at wireless terminals. Simulation results showed that our scheme achieved max-min fairness among best-effort flows, while guaranteeing QoS requirements for real-time traffic.

Next in Chapter 4, we focused on the well-known performance anomaly in multi-rate IEEE 802.11 wireless LANs, where stations with the lowest data transmission rate regulate the throughput of all other stations and the total system throughput degrades badly. Since flows are not identical to each other in such a multi-rate wireless LAN, we aimed at achieving proportional fairness in order to strike a balance between the throughput fairness among stations and the total system throughput. As the proportional fairness in a wireless LAN is equivalent to the air-time fairness, in Chapter 4, we presented a dynamic window control of uplink and downlink flows that achieves air-time fairness in flow-level. In our scheme, flows are classified into several classes according to their data transmission rates, and at APs, downlink flows in respective classes are stored in separate buffers. Further our scheme assigns different $CW_{\min}$s to those classes according to their data transmission rates and the target packet rates. Through simulation experiments with UDP and TCP, we showed the excellent performance of our dynamic contention window control scheme.

As summarized above, we addressed three fairness issues and for each we presented a dynamic contention window scheme to alleviate unfairness. All three schemes we introduced are based on controlling packet rates at stations in link-layer level. Therefore, we believe that even when flows have different characteristics, such as different packet lengths, our schemes can be applicable if the target packet rate ratio is known. In determining the target packet rate ratio R^* (when all flows are identical, R^* is given by the number of downlink flows at the AP), however, our schemes may require APs to know several information about downlink flows, including the number of active downlink flows. To do so, it is necessary to implement an additional function on APs to collect information of layers 3 and 4, such as IP addresses, port numbers, and transport protocol.

Throughout this dissertation, we have assumed a single-cell wireless LAN environment. In a multi-cell environment, however, the behavior and the performance of each station varies depending upon the locations (positions) of all stations as well as the access scheme, i.e., basic scheme or RTS/CTS scheme, they use to transmit data frames [29]. Thus in such cases, further discussions and more evaluations are required. We leave such issues for future works.

References

[1] W. P. Aukom, S. Banerjee, and P. Krishnamurthy, "Starvation prevention and quality of service in wireless LANs," in *Proceedings of Wireless Personal Multimedia Communications (WPMC'02)*, Hawaii, October 27–30, 2002, pp. 1078–1082.

[2] A. V. Babu and L. Jacob, "Fairness analysis of IEEE 802.11 multirate wireless LANs," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 3073–3088, September 2007.

[3] A. Banchs and X. Perez, "Distributed weighted fair queuing in 802.11 wireless LAN," in *Proceedings of IEEE International Conference on Communications (ICC'02)*, New York, USA, April 28–May 2, 2002, pp. 3121–3127.

[4] A. Banchs and X. Perez, "Providing throughput guarantees in IEEE 802.11 wireless LAN," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'02)*, Orlando, USA, March 17–21, 2002, pp. 130–138.

[5] M. Barry, A. T. Campbell, and A. Veres, "Distributed control algorithms for service differentiation in wireless packet networks," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'01)*, Alaska, USA, April 22–26, 2001, pp. 582–590.

[6] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed., New Jersey, USA: Prentice-Hall International, 1992.

[7] H. Boland and H. Mousavi, "Security issues of the IEEE 802.11b wireless LAN," in *Proceedings of Canadian Conference on Electrical and Computer Engineering (CCECE'04)*, Ontario, Canada, May 2–5, 2004, pp. 333–336.

[8] C. Casetti and C. F. Chiasserini, "Improving fairness and throughput for voice traffic in 802.11e EDCA," in *Proceedings of IEEE Personal, Indoor, and Mobile Radio Communications (PIMRC'04)*, Barcelona, Spain, September 5–8, 2004, pp. 525–530.

[9] C. Casetti, C. F. Chiasserini, L. Merello, and G. Olmo, "Supporting multimedia traffic in 802.11e WLANs," in *Proceedings of IEEE Vehicular Technology Conference (VTC'05-Spring)*, Stockholm, Sweden, May 30–June 1, 2005, pp. 2340–2344.

[10] J.-C. Chen, M.-C. Jiang, and Y.-W. Liu, "Wireless LAN security and IEEE 802.11i," *IEEE Wireless Communications Magazine*, vol. 12, no. 1, pp. 27–36, February 2005.

[11] P. Chevillat, J. Jelitto, A. N. Barreto, and H. L. Truong, "A dynamic link adaptation algorithm for IEEE 802.11a wireless LANs," in *Proceedings of IEEE International Conference on Communications (ICC'03)*, Alaska, USA, May 11–15, 2003, pp. 1141–1145.

[12] C. T. Chou, K. G. Shin, and N. S. Shankar, "Contention-based airtime usage control in multirate IEEE 802.11 wireless LANs," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1179–1192, December 2006.

[13] T. Clancy, "Secure handover in enterprise WLANs: CAPWAP, HOKEY, and IEEE 802.11r," *IEEE Wireless Communications Magazine*, vol. 15, no. 5, pp. 80–85, October 2008.

[14] T. Clancy, M. Nakhjiri, V. Narayanan, and L. Dondeti, "Handover key management and re-authentication problem statement," RFC 5169, March 2008.

[15] C. M. Cordeiro and D. P. Agrawal, *Ad Hoc & Sensor Networks*, Massachusetts, USA: World Scientific, 2006.

[16] J. Dunn, M. Neufeld, A. Sheth, D. Grunwald, and J. Bennett, "A practical cross-layer mechanism for fairness in 802.11 networks," *Mobile Networks and Applications*, vol. 11, no. 1, pp. 37–45, February 2006.

[17] *End-User Multimedia QoS Categories*, ITU-T Std. G.1010, 2001.

[18] M. Ergen and P. Varaiya, "Formulation of distribution coordination function of IEEE 802.11 for asynchronous networks: Mixed data rate and packet size," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 436–447, January 2008.

[19] S. Floyd and T. Henderson, "The new Reno modification to TCP's fast recovery algorithm," RFC 2582, April 1999.

[20] S. Fluhrer, I. Mantin, and A. Shamir, "Weaknesses in the key scheduling algorithm of RC4," *Selected Areas in Cryptography*, vol. 2259, pp. 1–24, August 2001.

[21] J. Freitag, N. L. S. da Fonseca, and J. F. de Rezende, "Tuning of 802.11e network parameters," *IEEE Communications Letters*, vol. 10, no. 8, pp. 611–613, August 2006.

[22] Y. Fukuda, A. Fujiwara, M. Tsuru, and Y. Oie, "Analysis of access point selection strategy in wireless LAN," in *Proceedings of IEEE Vehicular Technology Conference (VTC'05-Fall)*, Texas, USA, September 25–28, 2005, pp. 2532–2536.

[23] Y. Fukuda and Y. Oie, "Decentralized access point selection architecture for wireless LANs—deployability and robustness," in *Proceedings of IEEE Vehicular Technology Conference (VTC'04-Fall)*, Los Angeles, USA, September 26–29, 2004, pp. 1103–1107.

[24] E. Garcia, D. Viamonte, R. Vidal, and J. Paradells, "Achievable bandwidth estimation for stations in multi-rate IEEE 802.11 WLAN cells," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'07)*, Helsinki, Finland, June 18–21, 2007.

[25] Y. Ge and J. Hou, "An analytical model for service differentiation in IEEE 802.11," in *Proceedings of IEEE International Conference on Communications (ICC'03)*, Alaska, USA, May 11–15, 2003, pp. 1157–1162.

[26] P. Gopalakrishnan, D. Famolari, and T. Kodama, "Improving WLAN voice capacity through dynamic priority access," in *Proceedings of IEEE Global Communications conference (GLOBECOM'04)*, Texas, USA, November 29–December 3, 2004, pp. 3245–3249.

[27] B. O. Hara, P. Calhoun, and J. Kempf, "Configuration and provisioning for wireless access points problem statement," RFC 3990, February 2005.

[28] D. He and C. Q. Shen, "Simulation study of IEEE 802.11e EDCF," in *Proceedings of IEEE Vehicular Technology Conference (VTC'03-Spring)*, Jeju, Korea, April 22–25, 2003, pp. 685–689.

[29] K. Heck, "Wireless LAN performance in overlapping cells," in *Proceedings of IEEE Vehicular Technology Conference (VTC'03-Fall)*, Orlando, USA, October 6–9, 2003, pp. 2895–2900.

[30] P. S. Henry and H. Luo, "WiFi: what's next?" *IEEE Communications Magazine*, vol. 40, no. 12, pp. 66–72, December 2002.

[31] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'03)*, San Francisco, USA, March 30–April 3, 2003, pp. 836–843.

[32] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: An optimal access method for high throughput and fairness in rate diverse wireless LANs," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 4, pp. 121–132, October 2005.

[33] G. Holland, N. Vaidya, and P. Bahl, “A rate-adaptive MAC protocol for multi-hop wireless networks,” in *Proceedings of ACM/IEEE Mobile Computing and Networking (MOBICOM'01)*, Rome, Italy, July 16–21, 2001, pp. 236–251.

[34] V. Holopainen and R. Kantola, “End to end principle in access point selection,” in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'07)*, Helsinki, Finland, June 18–21, 2007.

[35] M. Ivanovici and R. Beuran, “User-perceived quality assessment for multimedia applications,” in *Proceedings of Optimization of Electrical and Electronic Equipments (OPTIM'06)*, Brasov, Romania, May 18–19, 2006, pp. 55–60.

[36] V. Jacobson, “Congestion avoidance and control,” in *Proceedings of ACM Symposium on Communications Architectures and Protocols (SIGCOMM'88)*, Stanford, USA, August 16–18, 1988, pp. 314–329.

[37] R. Jain, D. M. Chiu, and W. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared systems,” *Digital Equipment Corporation Technical Report*, DEC-TR-301, September 1984.

[38] L. B. Jiang and S. C. Liew, “Proportional fairness in wireless LANs and ad hoc networks,” in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'05)*, New Orleans, USA, March 13–17, 2005, pp. 1551–1556.

[39] T. Joshi, A. Mukherjee, Y. Yoo, and D. P. Agrawal, “Airtime fairness for IEEE 802.11 multirate networks,” *IEEE Transactions on Mobile Computing*, vol. 7, no. 4, pp. 513–527, April 2008.

[40] T. Joshi, A. Mukherjee, and D. P. Agrawal, “Analytical modeling of the link delay characteristics for IEEE 802.11 DCF multi-rate WLANs,” in *Proceedings of Canadian Conference on Electrical and Computer Engineering (CCECE'06)*, Ottawa, Canada, May 7–10, 2006, pp. 2164–2167.

[41] A. Kamerman and L. Monteban, “WaveLAN-II: A high-performance wireless LAN for the unlicensed band,” *Bell Labs Technical Journal*, vol. 2, no. 3, pp. 118–133, August 2002.

[42] K. Kawamura, T. Hiraguri, and M. Ogasawara, “Technique for dynamically updating EDCA access parameters for WLANs,” *NTT Technical Review*, vol. 5, no. 11, November 2007.

[43] F. Kelly, “Charging and rate control for elastic traffic,” *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, January/February 1997.

[44] B.-S. Kim, Y. Fang, T. F. Wong, and Y. Kwon, "Throughput enhancement through dynamic fragmentation in wireless LANs," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1415–1425, July 2005.

[45] H. Kim, S. Yun, I. Kang, and S. Bahk, "Resolving 802.11 performance anomalies through QoS differentiation," *IEEE Communications Letters*, vol. 9, no. 7, pp. 655–657, July 2005.

[46] S. W. Kim, B.-S. Kim, and Y. Fang, "Downlink and uplink resource allocation in IEEE 802.11 wireless LANs," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 1, pp. 320–327, January 2005.

[47] D. Koscielnik, "Simulation study of IEEE 802.11e wireless LAN—enhancements for real time applications," in *Proceedings of IEEE International Symposium on Industrial Electronics (ISIE'06)*, Montreal, Canada, July 9–13, 2006, pp. 1875–1880.

[48] M. Lacage, M. H. Manshaei, and T. Turletti, "IEEE 802.11 rate adaptation: A practical approach," in *Proceedings of ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM'04)*, Venice, Italy, October 4–6, 2004, pp. 126–134.

[49] M. Lampe, H. Rohling, and J. Eichinger, "PER-prediction for link adaptation in OFDM systems," in *International OFDM Workshop'02*, Hamburg, Germany, September 10–11, 2002.

[50] D. J. Leith, P. Clifford, D. Malone, and A. Ng, "TCP fairness in 802.11e WLANs," *IEEE Communications Letters*, vol. 9, no. 11, pp. 964–966, November 2005.

[51] A. Lindgren, A. Almquist, and O. Schelen, "Evaluation of quality of service schemes for IEEE 802.11 wireless LANs," in *Proceedings of IEEE Local Computer Networks (LCN'01)*, Florida, USA, November 14–16, 2001, pp. 348–351.

[52] K. Liu, T. Wong, J. Li, L. Bu, and J. Han, "A reservation-based multiple access protocol with collision avoidance for wireless multihop ad hoc networks," in *Proceedings of IEEE International Conference on Communications (ICC'03)*, Alaska, USA, May 11–15, 2003, pp. 1119–1123.

[53] *Local and Metropolitan Area Networks: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std. 802.16, 2004.

[54] *Local and Metropolitan Area Networks: Air Interface for Mobile Broadband Wireless Access Systems Supporting Vehicular Mobility - Physical and Media Access Control Layer Specification*, IEEE Std. 802.20, 2008.

[55] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS support in wireless LANs," *IEEE Wireless Communications Magazine*, vol. 10, no. 6, pp. 40–50, December 2003.

[56] M. H. Manshaei, T. Turletti, and M. Krunz, "A media-oriented transmission mode selection in 802.11 wireless LANs," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'04)*, Atlanta, USA, March 21–25, 2004, pp. 1228–1233.

[57] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*, New York, USA: Oxford University Press, 1995.

[58] *Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE Std. 802.15.4, 2006.

[59] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, October 2000.

[60] J. S. Park and D. Dicoi, "WLAN security: Current and future," *IEEE Internet Computing*, vol. 7, no. 5, pp. 60–65, September/October 2003.

[61] J. P. Pavon and S. Choi, "Link adaptation strategy for IEEE 802.11 WLAN via received signal strength measurement," in *Proceedings of IEEE International Conference on Communications (ICC'03)*, Alaska, USA, May 11–15, 2003, pp. 1108–1113.

[62] S. Pilosof, R. Ramjee, D. Raz, Y. Shavitt, and P. Sinha, "Understanding TCP fairness over wireless LAN," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'03)*, San Francisco, USA, March 30–April 3, 2003, pp. 863–872.

[63] M. Pioro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*, San Francisco, USA: Morgan Kaufmann, 2004.

[64] D. Pong and T. Moors, "Fairness and capacity trade-off in IEEE 802.11 WLANs," in *Proceedings of IEEE Local Computer Networks (LCN'04)*, Florida, USA, November 15–17, 2004, pp. 310–317.

[65] A. R. Prasad and N. R. Prasad, *802.11 WLANs and IP networking: Security, QoS, and mobility*, Massachusetts, USA: Artech House, 2005.

[66] R. Pries, S. Menth, D. Staehle, M. Menth, and P. Tran-Gia, "Dynamic contention window adaptation (DCWA) in IEEE 802.11e wireless local area networks," in *Proceedings of International Conference on Communications and Electronics (ICCE'08)*, HoiAn, Vietnam, June 4–6, 2008, pp. 92–97.

[67] J. G. Proakis and M. Salehi, *Digital Communications*, New York, USA: McGraw-Hill, 2008.

[68] *Pulse Code Modulation (PCM) of Voice Frequencies*, ITU-T Std. G.711, 1988.

[69] D. Qiao, S. Choi, and K. G. Shin, "Goodput analysis and link adaptation for IEEE 802.11a wireless LANs," *IEEE Transactions on Mobile Computing*, vol. 1, no. 4, pp. 278–292, October/December 2002.

[70] B. Radunovi and J. Le Boudec, "Rate performance objectives of multihop wireless networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 334–349, October/December 2004.

[71] T. Razafindralambo, I. Lassous, L. Iannone, and S. Fdida, "Dynamic packet aggregation to solve performance anomaly in 802.11 wireless networks," in *Proceedings of ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM'06)*, Malaga, Spain, October 2–6, 2006, pp. 247–254.

[72] P. Roshan and J. Leary, *802.11 Wireless LAN Fundamentals*, Indiana, USA: Cisco Press, 2004.

[73] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "OAR: An opportunistic auto-rate media access protocol for ad hoc networks," *Wireless Networks*, vol. 11, nos. 1–2, pp. 39–53, January 2005.

[74] K. Saitoh, Y. Inoue, A. Iiuka, and M. Morikura, "An effective data transfer method by integrating priority control into multirate mechanisms for IEEE 802.11 wireless LANs," in *Proceedings of IEEE Vehicular Technology Conference (VTC'02-Spring)*, Alabama, USA, May 6–9, 2002, pp. 55–59.

[75] S. H. Shah, K. Chen, and K. Nahrstedt, "Dynamic bandwidth management for single-hop ad hoc wireless networks," in *Proceedings of IEEE International Conference of Pervasive Computing and Communications (PerCom'03)*, Texas, USA, March 23–26, 2003, pp. 195–203.

[76] V. A. Siris and G. Stamatakis, "Optimal CWmin selection for achieving proportional fairness in multi-rate 802.11e WLANs: Test-bed implementation and evaluation," in *Proceedings of Wireless Network Testbeds, Experimental Evaluation & Characterization (WiNTECH'06)*, Los Angeles, USA, September 29, 2006, pp. 41–48.

[77] C. Smith and D. Collins, *3G Wireless Networks*, 2nd ed., New York, USA: McGraw-Hill, 2007.

[78] W. R. Stevens, *TCP/IP Illustrated*, vol. 1, Indiana, USA: Addison-Wesley, 1999.

[79] G. Tan and J. Guttag, "Time-based fairness improves performance in multi-rate WLANs," in *Proceedings of USENIX Annual Technical Conference (USENIX'04)*, June 27–July 2, Boston, USA, 2004, pp. 269–282.

[80] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed fair scheduling in a wireless LAN," in *Proceedings of ACM/IEEE Mobile Computing and Networking (MOBICOM'00)*, Boston, USA, August 6–11, 2000, pp. 167–178.

[81] *Video Coding for Low Bit Rate Communication*, ITU-T Std. H.263, 2005.

[82] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. 802.11, 1999.

[83] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Std. 802.11a, 1999.

[84] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, IEEE Std. 802.11b, 1999.

[85] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements*, IEEE Std. 802.11e, 2005.

[86] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band*, IEEE Std. 802.11g, 2003.

[87] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Medium Access Control (MAC) Security Enhancements*, IEEE Std. 802.11i, 2004.

[88] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: Fast Basic Service Set (BSS)*, IEEE Std. 802.11r, 2008.

[89] *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs)*, IEEE Std. 802.15.1, 2005.

[90] Q. Xia and M. Hamdi, "Smart sender: A practical rate adaptation algorithm for multirate IEEE 802.11 WLANs," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1764–1775, May 2008.

[91] Y. Xiao, "QoS guarantee and provisioning at the contention-based wireless MAC layer in the IEEE 802.11e wireless LANs," *IEEE Wireless Communications Magazine*, vol. 13, no. 1, pp. 14–21, February 2006.

[92] D.-Y. Yang, T.-J. Lee, K. Jang, J.-B. Chang, and S. Choi, "Performance enhancement of multirate IEEE 802.11 WLANs with geographically scattered stations," *IEEE Transactions on Mobile Computing*, vol. 5, no. 7, pp. 906–919, July 2006.

[93] L.-H. Yen, T.-T. Yeh, and K.-H. Chi, "Load balancing in IEEE 802.11 networks," *IEEE Internet Computing*, vol. 13, no. 1, pp. 56–64, January/February 2009.

[94] H. Zhu, M. Li, I. Chlamtac, and B. Prabhakaran, "A survey of quality of service in IEEE 802.11 networks," *IEEE Wireless Communications Magazine*, vol. 11, no. 4, pp. 6–14, August 2004.

Publications

A. Journal Papers

1. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control mechanism to achieve fairness between uplink and downlink flows in IEEE 802.11 wireless LANs," *IEEE Transactions on Wireless Communications*, vol. 7, no. 9, pp. 3517–3525, September 2008.

2. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control scheme in IEEE 802.11e EDCA-based wireless LANs," *IEICE Transactions on Communications*, vol. E93-B, no. 1, pp. 56–64, January 2010.

B. Refereed Conference Papers

1. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control to achieve fairness between uplink and downlink flows in IEEE 802.11 WLANs," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'07)*, Hong Kong, March 11–15, 2007, pp. 2109–2114.

2. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control scheme in IEEE 802.11e wireless LANs," presented at *IEEE Vehicular Technology Conference (VTC'09-Spring)*, Barcelona, Spain, April 26–29, 2009.

3. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control of uplink and downlink flows in multi-rate IEEE 802.11 wireless LANs," accepted for presentation at *International Conference on Information Networking (ICOIN'10)*, Busan, Korea, January 27–29, 2010.

C. Non-Refereed Technical Papers

1. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Window control for fairness issue between uplink and downlink flows in IEEE 802.11 wireless LANs," *IE-*

ICE Technical Report, vol. 105, no. 628, IN2005-171, pp. 91–96, March 2006 [in Japanese].

2. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic window control mechanism to achieve per-flow fairness in IEEE 802.11 WLANs," *IEICE Technical Report*, vol. 106, no. 578, IN2006-195, pp. 89–94, March 2007 [in Japanese].

3. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control mechanism in IEEE 802.11e wireless LANs," *IEICE Technical Report*, vol. 107, no. 524, NS2007-189, pp. 319–324, March 2008 [in Japanese].

4. B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Contention window control in multi-rate IEEE 802.11 wireless LANs," *IEICE Technical Report*, vol. 108, no. 286, NS2008-100, pp. 69–74, November 2008 [in Japanese].