A Performance Modeling of WiMedia UWB MAC

Chunxiao Ma

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science (Electrical Engineering) at

Concordia University

Montreal, Quebec, Canada

April 2010

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Chunxiao Ma

Entitled: A Performance Modeling of WiMedia UWB MAC

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair

_____ Examiner

_____ Examiner

_____ Supervisor

Approved by _____
Chair of Department or Graduate Program Director

_____
Dean of Faculty

Date _____

ABSTRACT

A Performance Modeling of WiMedia UWB MAC

Chunxiao Ma

The WiMedia Ultra-Wide Band (UWB) specifications include Multi-band Orthogonal Frequency Division Multiplexing (OFDM) physical layer and distributed Medium Access Control (MAC) layer. Physical layer offers a range of coded data rates from 53.3 Mb/s to 480Mb/s. WiMedia MAC consists of Beacon period (BP), Distributed Reservation Protocol (DRP) and Prioritized Contention Access (PCA). DRP allows access to the medium through reservation, while PCA provides access to the bandwidth not utilized by DRP through contention. In this thesis, we present an integrated performance analysis of DRP and PCA schemes which includes the multi-rate transmission at the physical layer. We assume that DRP traffic consists of several classes of calls and we use a call blocking model to determine the performance of DRP scheme in terms of call blocking probabilities and bandwidth utilization of the system. Then we assume PCA traffic with multi-priority levels. We determine the throughput and mean total delay for PCA traffic in the presence of DRP traffic under the non-saturation condition. The numerical results show that under heavy traffic throughput of low priority traffic drops and its delay increases sharply.

ACKNOWLEDGEMENTS

First and foremost I offer my sincerest gratitude to my supervisor, Dr. Mustafa K. Mehmet Ali, who provided me invaluable assistance, support and guidance throughout my research and thesis. Without his knowledge and advice, this thesis would not have been completed.

I would like to thank all my friends and colleagues, who gave me wise advice and helped me to get through the difficult times. I am especially grateful to Guangfu Shi, who supported me in many different ways and stand beside me all the time. I am also thankful to Zuhui Ma, Tong Li and Yi Li for helping me a lot in my learning and life in Montreal.

Lastly, and most important, I would like to thank my parents, Yulan Xu and Yaquan Ma for their support and gently love. They raised me, taught me and gave me the wonderful life I wanted. I cannot express my gratitude for them in words. Without them, I would not be here.

Dedicated to my dear parents ...

Table of Contents

List of Figures

List of Tables

List of Symbols

T_{MAS}	Duration of a superframe slot
T	Duration of a superframe measured in microseconds
N_{SF}	The number of channels in a superframe
N_o	The number of channels reserved for PCA traffic in a superframe
N_{SD}	The maximum number of channels available for DRP traffic in a superframe
N_{DRP}	The number of channels currently reserved for DRP traffic in a superframe
N_{PCA}	The number of channels available for PCA traffic in a superframe
α_k	Mean arrival rate of class k calls measured in calls per second
α	Total arrival rate of calls measured in calls per second
μ_k	Service rate of class k calls measured in calls per second
$\bar{X}_k$	Mean duration of class k calls in the system
ρ_k	Traffic load offered by class k calls
b_k	The number of channels per superframe reserved for a class k call, $b_k \leq N_{SD}$

n_k	The number of class k calls in the system
Γ_k	The size of information payload of class k calls in number of bits
$\bar{\Gamma}_k$	The average of Γ_k
Δ	The number of distinct physical transmission rates
R_k	Physical transmission rate of class k calls
F_k	The number of bits per superframe that may be transferred for a class k call
$\bar{N}_k$	Average number of superframes needed to transmit a class k call
n	The network state of the DRP calls, $n = (n_1, \ldots, n_k, \ldots, n_K)$
n_k^+	The network state that the k'th component differ from the state n by plus a unit, $n_k^+ = (n_1, \ldots, n_{k-1}, n_k + 1, n_{k+1}, \ldots, n_K)$
n_k^-	The network state that the k'th component differ from the state n by minus a unit, $n_k^- = (n_1, \ldots, n_{k-1}, n_k - 1, n_{k+1}, \ldots, n_K)$
Ω	The set of permissible network states for DRP calls, $\Omega = \{n \mid 0 \leq n \cdot b \leq N_{SD}\}$
Ω'	The set of permissible network states for unblocking DRP calls, $\Omega' = \{n \mid 0 \leq n \cdot b \leq N_{SD} - b_k\}$
δ_k^+	A variable that indicates whether n_k^+ is a permissible state or not

$\delta_{\bar{k}}$	A variable that indicates whether $n_{\bar{k}}$ is a permissible state or not
$P(n)$	The probability that there are n calls in the system
$G(\Omega)$	Normalization constant
$G(\Omega')$	Normalization constant
Q_k	The blocking probability of a new arriving class k call
$q(c)$	The probability that the number of channels reserved is equal to c
$E(n_k \mid c)$	The expectation value of the number of class k calls given a total of c channels in a superframe is reserved
$\hat{q}(c)$	A variable defined as $\hat{q}(c) = q(j)/q(0)$, where $\hat{q}(0) = 1$
$\overline{N}_{DRP}$	The average number of channels reserved for DRP traffic
η	The number of PCA stations
N	The number of access categories in PCA traffic
i	The priority level of an access category, where $i = 0 \ldots N-1$
AC_i	The access category assigned with priority i
λ	The arrival rate of packets for each PCA station measured in packets per second
λ_i	The arrival rate of packets for AC_i per station measured in packets per second

λ_T	The total arrival rate of packets to all PCA stations measured in packets per second
$CW_{i,min}$	The initial minimum backoff window size for AC_i
$CW_{i,max}$	The maximum backoff window size for AC_i
$b(i,t)$	The random process representing the value of the backoff counter of AC_i at time t
$s(i,t)$	The random process representing the backoff stage j ($j = 0,1, \dots, L_{i,retry}$) for AC_i at time t
$v(i,t)$	The random process representing whether there is at least a packet waiting in the queue of AC_i at the time of packet service completion
l	The current value of the backoff counter, $l = 0,1, \dots, W_{i,j}$
j	The current backoff stage, $j = 0,1, \dots, L_{i,retry}$
e	A binary variable that denotes the status of the AC_i queue as empty or not empty
$L_{i,retry}$	The retry limit for AC_i
$W_{i,j}$	The value of backoff window size for AC_i in the j'th backoff stage
ρ_i	The probability that an AC_i queue is busy at the completion of packet service time

ρ_i^*	The probability that an AC_i queue is not empty at the moment when a packet transmission is successful or it is discarded
p_i	The probability that a transmitted AC_i packet collides during a PCA slot
q_i	The probability that at least one packet arrival during a generic slot
q_i^*	The probability that at least one packet arrival during the time interval between the two slots that the backoff counter may be decremented
$b_{i,j,l,e}$	The steady-state distributions of the Markov chain
τ_i	The probability that a station transmits a packet from AC_i queue in a generic slot
p_{busy}	The probability that the channel is busy in a PCA period
p_i^{loss}	The probability that an AC_i packet will be discarded
$\tilde{g}(\zeta)$	The PDF of the duration of a generic slot
$\bar{\tilde{g}}$	The mean duration of a generic slot
$P_i(\varsigma, \tilde{t})$	The probability of ς packets arrivals at the AC_i queue during a time duration of $\tilde{t}$ seconds
$p_{s,i}$	The probability that an AC_i packet is transmitted successfully from any station during a generic slot
p_s	The probability that a generic slot carries a successful packet transmission

S_i	The throughput of PCA for AC_i
T_e	The real-time duration of an empty PCA slot
L	The length of the payload
$T_{E(L)}$	The time to transmit a payload with average length $E(L)$ in real time
$T_{s(i)}$	The average time of a successful PCA transmission for AC_i in real time
T_c	The average time of a collision in PCA period in real time
T_H	The time to transmit the header
T_{SIFS}	The duration SIFS period
T_{ACK}	The time to transmit an acknowledgement
T_{RTS}	The time to transmit a Request-To-Send (RTS) frame
T_{CTS}	The time to transmit a Clear-To-Send (CTS) frame
$T_{AIFS(i)}$	The Arbitration Inter-Frame Space (AIFS) time for AC_i
β_k	The probability that the transmission rate in a station is R_k
$\bar{N}_{PCA}$	The average number of channels available for PCA traffic
T'_e	The duration of empty PCA slot measured in virtual time slots
$T'_{s(i)}$	The mean duration of packet successful transmissions from AC_i queue measured in virtual time slots

T_c'	The mean duration of collisions measured in virtual time slots
a_i	The number of generic slots that it takes to decrement the counter by one
$A_i(z)$	The PGF of a_i
x_i	The duration of the time in number of virtual slots that a counter remains frozen during a generic slot
$X_i(z)$	The PGF of x_i
y_i	The amount of time in number of virtual slots that it takes to decrement counter by one
$Y_i(z)$	The PGF of y_i
v_i	The total amount of time it takes to decrement the counter by one
$\bar{v}_i$	The mean value of v_i
$V_i(z)$	The PGF of v_i
$r_{i,j}$	The chosen value of the backoff counter in the j'th stage for AC_i
$R_{i,j}(z)$	The PGF of $r_{i,j}$
$g_{i,j}$	The number of virtual slots that AC_i spends in the j'th backoff stage
$G_{i,j}(z)$	The PGF of $g_{i,j}$
$B_{i,j,s}(z)$	A substitute for $\prod_{k=s}^{j} G_{i,k}(z)$

p_i^{EM}	The probability of no arrivals during the post-backoff procedure
$M_i(z)$	The PGF of packet service time of AC_i
$\overline{M}_i$	Then mean packer service time of AC_i
$\bar{Q}_i$	The mean queuing delay for AC_i queue
$\overline{M}_i{}'$	The mean packet service time excluding the post-backoff procedure
$M_i{}'(z)$	The PGF of $\overline{M}_i{}'$
$\overline{D}_i$	The mean total delay of a packet to be transmitted

List of Acronyms

AC	Access Category
ACK	Acknowledgment
AIFS	Arbitration Inter-Frame Space
BM	Burst Mode
BP	Beacon Period
BPST	Beacon Period Start Time
BS	Beacon slot
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
DCM	Dual Carrier Modulation
DRP	Distributed Reservation Protocol
EIRP	Equivalent Isotropically Radiated Power
FCC	Federal Communications Commission
FCFS	First Come First Served
GBE	Global Balance Equation
IE	Information Element
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
IFS	Inter-Frame Space
ISM	Industrial, Scientific and Medical
LBE	Local Balance Equation
LLC	Logical Link Control

MAC	Medium Access Control
MAS	Medium Access Slot
MIB	Management Information Base
MIFS	Minimum Interframe Spacing
MPDU	MAC Protocol Data Unit
MSDU	MAC Service Data Unit
NAV	Network Allocation Vector
OFDM	Orthogonal Frequency Division Multiplexing
PAN	Personal Area Network
PCA	Prioritized Contention Access
PDU	Protocol Data Unit
PER	Packet Error Rate
PHY	Physical (layer)
PHY SAP	PHY Service Access Point
PLCP	Physical Layer Convergence Protocol
PMD	Physical Medium Dependent
PN	Pseudo-Noise
PPDU	PLCP Protocol Data Unit
PSDU	PHY Service Data Unit
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RX	Receive or Receiver
RTS	Request To Send
SAP	Service Access Point

SIFS	Short Interframe Spacing
TX	Transmit or Transmitter
TXOP	Transmission Opportunity
UWB	Ultra Wideband
ZPS	Zero-Padded Suffix

CHAPTER 1

INTRODUCTION

1.1 Introduction

During the past few decades, the electronic devices have been extensively adopted in the home applications and office facility. With the development of the electronic industry, providing high data rate network access with the Quality of Service (QoS) became a challenge. The traditional solution has been the Ethernet, where devices are connected with the network by cables. However, as more portable terminals such as notebook computers, personal digital assistants (PDAs) and mobile phones became part of the daily life; wireless networks on account of flexibility are preferred instead of the wired networks. The short-range wireless technology enable devices move around freely in a local coverage area without piles of cables and keep providing connectivity to the network. It also has the characteristics of low expense and ease of installation. Therefore, the short-range wireless technologies have a huge potential on personal and business networks and it has received an enormous research attention in recent years. The subject of this thesis is the performance of short-range wireless technologies.

Generally, the short-range wireless networks can be classified into two kinds: wireless local area networks (WLANs) and the wireless personal area networks (WPANs). The WLAN permits indoor communication within the range of 46 meters and outdoor communication within 92 meters, the main standards of which are developed by

Institute of Electrical and Electronics Engineers (IEEE) 802.11 working group. On the other hand, WPAN, also known as indoor networks, has a range of 10 meters, whose standards are developed by IEEE 802.15 working group. Several types of WPAN technologies have achieved success in the market, such as Zigbee based on IEEE 802.15.4 for low rate WPANs [1], Bluetooth based on IEEE 802.15.1 standard for high rate WPANs [2] and Ultra-Wide Band (UWB) based on the ECMA-368 and ECMA-369 standards proposed by WiMedia Alliance [3][4]. UWB is a new generation of WPAN technology and is a good choice both for low and high rate WPANs, due to its low transmission power and low interference features.

Ultra-wideband transmission is a short-range wireless communication technology that transmits short pulses over a large bandwidth. To avoid interference to existing radio services, different standards and restrictions was established by the Federal Communications Commission (FCC) for three types of UWB devices: the imaging systems, vehicular radar systems, and communications and measurement systems. In 2002, FCC authorized the unlicensed use of UWB system in the frequency range from 3.1 to 10.6 GHz. The UWB was defined by FCC as the wireless transmission system whose instantaneous width is greater than 20% of the center frequency or fractional bandwidth is greater than 500MHz [5]. The FCC Equivalent Isotropically Radiated Power (EIRP) emission of UWB devices is lower than the -41.3 dBm/MHz level. The spectrum mask of the UWB signal is shown in Figure 1.1. There are two proposals for the implementation of the UWB system under the restrictions of FCC. One is the pulse-based approach, which transmits pulses during extremely short durations over a rich frequency spectrum. The other one is multiband-orthogonal frequency-division multiplexing (MB-

OFDM)-based approach, which transmits data simultaneously using multiple-carriers and OFDM fashion [6]. After several years' of discussion, the former approach was finally added into IEEE 802.15.4a standard finally; MB-OFDM UWB, supported by the WiMedia Alliance, is published as ECMA-368 and 369 standards [3][4]. In March 2007, WiMedia UWB Common Radio Platform was approved for release as an ISO/IEC International Standard, which was the first UWB international standard.

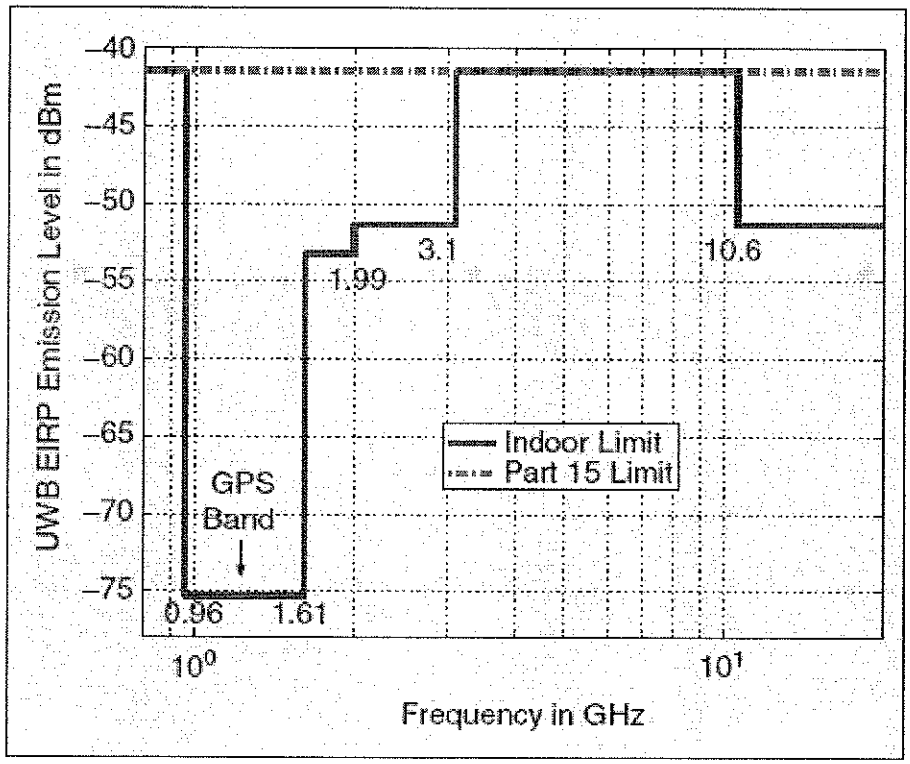


Figure 1.1 FCC spectral mask for indoor commercial system [5]

1.2 Short-range Wireless Technologies

As the development of the short-range wireless technologies continues, WPAN has been gaining much attention due to its high throughput and low power consumption features. While the WLAN works in the indoor range of about 46 meters, WPAN typically works for the range under 10 meters. In this section, we will briefly introduce the main technologies for the short-range wireless networks.

1.2.1 WiFi

WiFi is a typical WLAN technology that is based on the IEEE802.11 standard. WiFi devices have been widely used in personal applications, such as PC, mobile phones, video game consoles and PDAs, enabling them to connect with the Internet without unsightly cables. WiFi technology also works well for the business, such as the devices in corporations and campuses due to its ease of deployment and mobility [7].

Several specifications of 802.11 standards have been proposed by IEEE for WLAN technology, each of which is identified by a lower case letter following 802.11 such as 802.11a. To date, there are four primary specifications available, defined by a set of characteristics that relate to unlicensed frequency band and performance, which are briefly described below,

- **802.11a**

  802.11a works in the 5GHz frequency band, and provides up to 54Mbps PHY transmission rate. It uses an orthogonal frequency division multiplexing (OFDM) modulation scheme.

- **802.11b**

  802.11b provides up to 11 Mbps transmission rate in the 2.4 GHz band. 802.11b uses direct sequence spread spectrum (DSSS) scheme.

- **802.11g**

  802.11g also works in the 2.4 GHz band, but uses OFDM modulation scheme as in 802.11a and it supports up to 54 Mbps PHY transmission rate.

- **802.11e**

  802.11e is an enhancement to 802.11a and 802.11b which adds Quality of Service (QoS) support. It uses enhanced distributed channel access (EDCA) mechanism to control channel access.

Next, we briefly explain some of the concepts mentioned in the above.

DSSS used in 802.11b is a modulation technique that disperses the information signal over the wider bandwidth in the frequency band of the device. At the transmitter side, each data bit is modulated by a pseudo-noise (PN) binary code sequence called chips. These chips have much higher bit rate than data signals. The structure of each chip is known by both transmitter and receiver. Thus, the receiver can reconstruct the information signal by multiplexing the same PN sequence with the received signal. The

inherent redundancy of the chip enables the receiver to check and correct the error that occurs during the transmission and thereby reduce the effects of noise.

Different from DSSS, OFDM modulation scheme distributes data over a large number of sub-carriers, each of which owns a unique frequency which distinguishes them from each other. These sub-carriers are designed to be orthogonal to each other, so that the interference between the sub-channels is very little. In addition, since the high speed data is divided into the sub-channels with lower rates, OFDM has resistance to Inter Symbol Interference (ISI) and fading caused by multipath propagation.

EDCA is a distributed access mechanism defined by 802.11e, which classifies the traffic into multiple priority classes and provides QoS support to multimedia applications. These priority classes are differentiated by the Transmission Opportunity (TXOP), the minimum and maximum contention window size and the Arbitration Interframe Space (AIFS). The traffic of higher priority classes have to wait less before they are transmitted so that they have higher opportunity to access the channel.

Although WiFi has become the most prominent WLAN technology, it still has many limitations. Today, most of WLAN products support 802.11b and 802.11g standards, both of which operates in 2.4 GHz Industrial, Scientific and Medical (ISM) frequency band. However, many applications use this frequency band, such as microwave ovens and the Bluetooth technology, as a result WiFi devices experience interference and oversubscription issues which cause latency and jitter. Furthermore, though the maximum raw data rate of 802.11 standards is 54 Mbps, the data rate that can be achieved in application layer is less than half of the raw data rate due to the high overhead in PHY

and MAC layers. The power requirement and energy consumption are also the issues need to be faced. In addition, the 802.11 standard is mainly based on CSMA/CA mechanism, so devices need to contend for the channel access, which may cause long waiting times before the data can be sent. Thus, WiFi is not a good choice for the real-time applications such as high quality video stream. A new generation of 802.11 named 802.11n was just released in October 2009. Since it has yet not been widely used, we do not discuss it in this thesis.

1.2.2 Bluetooth

Bluetooth is a protocol designed for WPAN based on 802.15.1 standard, which is used within a short transmission range and has low power consumption. This technology allows wireless devices such as mobile phones, PC input and output devices and game consoles to interconnect with each other over a range of up to 10 meters. Bluetooth has a lower data rate than WiFi, so that it has lower power requirement. However, Bluetooth encounters the same issue as WiFi since it also uses 2.4 GHz ISM band; it is easily affected by the interference from other devices operating in the same frequency band. Besides, the low data rate and short communication range are also the main drawbacks of the Bluetooth.

1.2.3 Zigbee

Zigbee is a specification for WPAN based on IEEE 802.15.4 standard, which defines low-power, low-cost networks for supporting communications within approximately 10 meters of distance. Zigbee technology also operates at 2.4 GHz ISM band and it has 16 channels with 5MHz bandwidth per channel. The radio uses DSSS modulation technique supporting 250 kbps raw data rate. The features of Zigbee are low cost and low power consumption, with the latter leading to long battery life. The main limitation of Zigbee is its low data rate. As a result, this technology is mainly used in wireless control and monitoring applications, such as home awareness, hospital care and access control in a commercial building.

1.2.4 UWB

Compared to other three short-range wireless technologies mentioned above, UWB is the best choice for high speed communications in WPAN due to its high PHY transmission rate and many other unique features. UWB uses an extremely wide frequency band (more than 7 GHz) to transmit data, which achieves higher transmission rates than other short range technologies. A number of sub-bands are defined in WiMedia UWB each with 500 MHz bandwidth; a device can transmit the data by interleaving the signals across these sub-bands and attain the low power consumption. In addition, UWB

works on the new legalized 3.1 GHz to 10.6 GHz spectrum, so it can get rid of the interference of the narrow-band devices nearby.

UWB standard specifies Physical layer (PHY) and Medium Access Control layer (MAC) protocols. WiMedia PHY uses multi-band orthogonal frequency division modulation (MB-OFDM) modulation method, so it maintains all the advantages of OFDM such as resiliency to inter-frame interference and multipath effects. It provides PHY data rates from 53.3 Mbps to 480 Mbps, which enable communications among multiple devices using different transmission rates. In UWB MAC, two access protocols are defined: distributed reservation protocol (DRP) and prioritized contention access (PCA) protocol. The former protocol is to provide Quality of Service to devices through channel reservation for exclusive transmission, while PCA is used for isochronous and asynchronous data transfer without reservation. Thus, UWB is more flexible than other technologies and it especially provides QoS required in real-time communications.

The objective of this thesis is to study the performance of WiMedia UWB MAC protocol. We present an integrated performance analysis of DRP and PCA protocol, which allows multiple physical transmission rates. We assume that several classes of calls reserve the channel by using DRP protocol and multiple priority levels of packets access the channel following PCA protocol. We present the call blocking probabilities for DRP traffic, the throughput and mean packet delay for PCA traffic under the non-saturation condition. Furthermore, we study effects of varying the call arrival rates and packet arrival rates and investigate its impact on the performance of PCA traffic.

CHAPTER 2

WIMEDIA UWB

In this chapter, we first introduce the UWB standard defined by WiMedia Alliance. Then, PHY and MAC layers of this standard are described in detail. We briefly review the previous work relevant to UWB MAC. Finally, the major contributions of the thesis are explained and the thesis organization is provided at the end of the chapter.

2.1 WiMedia UWB Standard Overview

The standard of WiMedia UWB specifies a distributed MAC sublayer and a PHY layer for UWB communications. WiMedia PHY utilizes an extremely wide frequency band and can support multiple data rates up to 480 Mbps. MB-OFDM scheme is adopted by UWB PHY to transmit information, which provides high spectral flexibility and resiliency to RF interference and multi-path effects [8]. In addition to the key features of PHY, the WiMedia MAC layer also has some remarkable characteristics. There are two kinds of MAC protocols defined for different services. One is distributed reservation protocol (DRP), which can be used to reserve bandwidth for periodic data transmission. The other one is prioritized contention access (PCA) protocol, which is

used to transmit bursts of data without reservations. Thus, the WiMedia MAC can support different types of services by using these two protocols simultaneously.

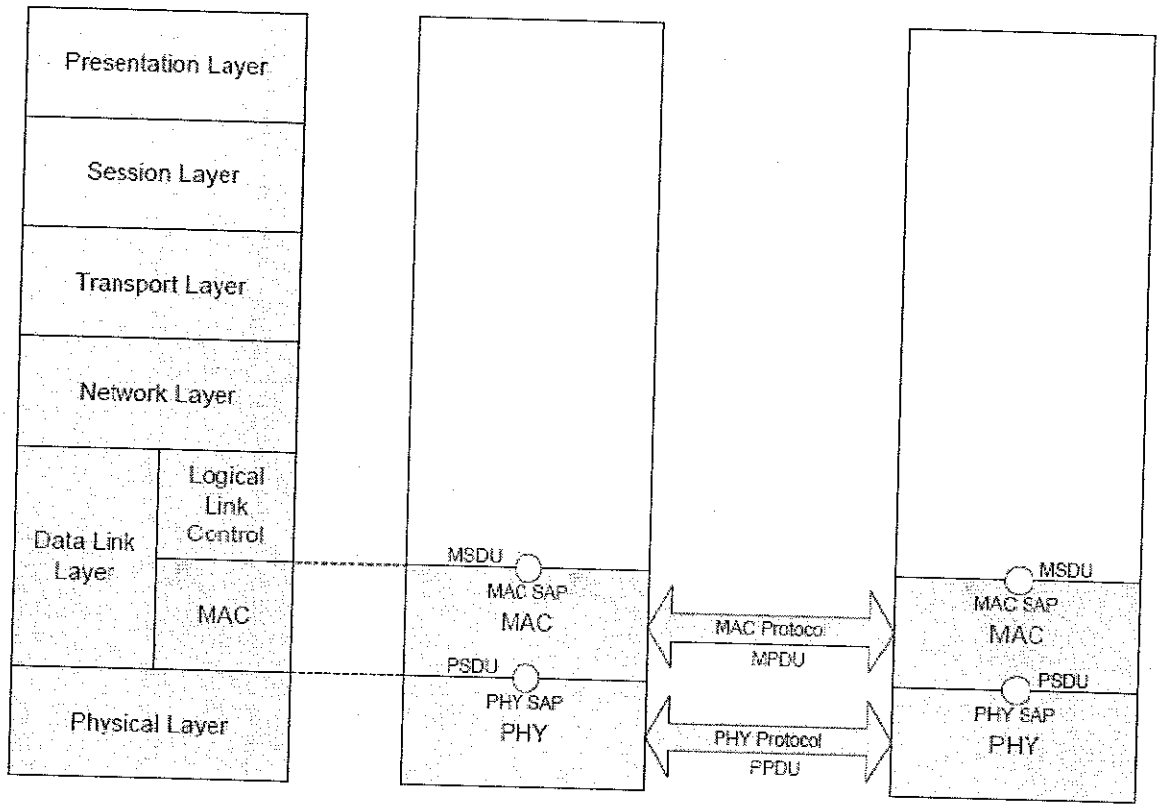


Figure 2.1 Architectural reference model [3]

As illustrated in Figure 2.1, PHY and data link layers (DLL) are the lowest layers of the OSI reference model and MAC layer is the sublayer of the data link layer. PHY layer provides service to MAC sublayer through the PHY service access point (PHY SAP) which will be described in the following section. On the other hand, the MAC layer provides service to the MAC client in the higher layer by means of the MAC service access point (MAC SAP). The MAC protocol applies between MAC sublayer peers at the source and destination nodes.

2.2 WiMedia PHY

The WiMedia UWB PHY provides communications at multiple data rates in the unlicensed spectrum from 3.1 to 10.6 GHz ISM radio band. The spectrum is divided into 14 different bands, each of which has the bandwidth of 528 MHz. These 14 bands are grouped into six band groups. Band groups 1 to 4 are over the span of bands 1 to 12, with three bands each; band group 5 consists bands 13 and 14; band group 6 consists bands 9, 10 and 11. Figure 2.2 shows the allocation of these six band groups and the center frequency of each band.

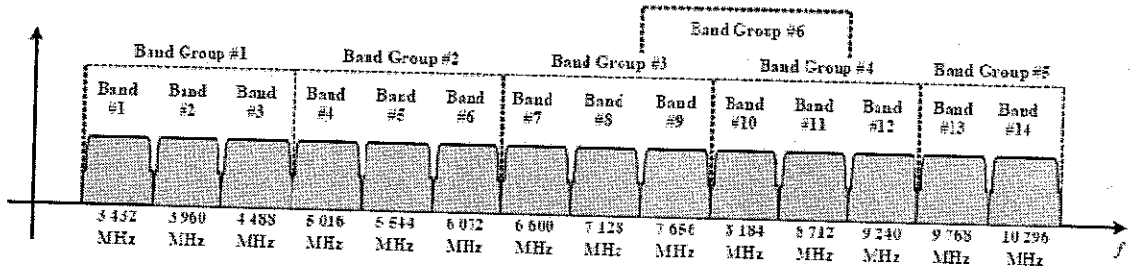


Figure 2.2 Diagram of the band group allocation [3]

The WiMedia UWB defines two sub-layers within the PHY layer. One is physical medium dependent (PMD) sublayer, which provides medium dependent methods for stations to transmit and receive data. The other one is physical layer convergence protocol (PLCP) sublayer. The function of PLCP sublayer is to convert the

PLCP service data units (PSDU) into the PLCP protocol data units (PPDU). Then, PMD sublayer transmits and receives PPDU by using MB-OFDM scheme.

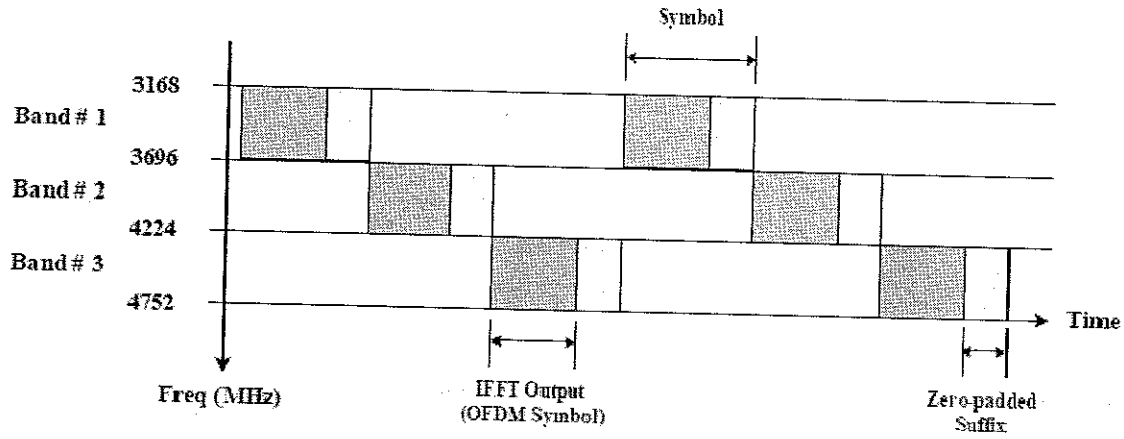


Figure 2.3 Example realization of a transmitted signal using three bands [3]

The OFDM symbol is created from a waveform by using an Inverse Fast Fourier Transform (IFFT) function. All OFDM symbols have the same length and are transmitted at the raw data rate of 640 Mbps. Data is modulated onto the carriers of the OFDM symbol and the frequency is changed between consequent symbols. Moreover, a zero-padded suffix (ZPS) is appended to each OFDM symbol, which enables the avoidance of the multi-path effects and provides a guard interval between the transmissions of different symbols. The OFDM symbols can hop across three bands in a band group following a predefined order. Figure 2.3 shows an example for the band group 1 which illustrates how the mechanism works. In the figure, the first symbol is transmitted on a centre frequency of 3 432 MHz, the second symbol changes its centre frequency to 3 960 MHz,

the third symbol changes its centre frequency to 4 488 MHz, then the fourth symbol changes its centre frequency back to 3 432 MHz, and so on.

By applying different coding techniques, such as quadrature phase shift keying (QPSK) and Dual Carrier Modulation (DCM), WiMedia PHY layer can provide transmissions at data rates of 53.3, 80, 106.7, 160, 200, 320, 400 and 480Mb/s. Since the signal to noise ratio (SNR) is affected by the distance between devices and the local electromagnetic environment, devices can choose different data rates for information transmission to satisfy their SNR requirements. The coverage ranges of data rates are shown in Table 2.1. Long range communication will cause low transmission rate. The biggest data rate can be achieved within 4 meters. Therefore, the wireless networks using Multi-rate devices are more flexible.

Data rate (Mbps)	Calculated coverage range (m)
53.3	< 18
80	< 14
106.7	< 13
160	< 10
200	< 9
320	< 6
400	< 5
480	< 4

Table 2.1 The coverage ranges of different data rate [9]

PHY layer provides both single frame and burst mode transmissions. In single frame transmission, the frame timing is fully controlled by MAC sublayer. On the contrary, in

the burst mode transmission only the first frame timing is controlled by MAC sublayer. The timing for each of the remaining frames from the same device is provided by PHY layer in the burst. PHY supplies information to the MAC sublayer regarding the frame transmission mode, data rate, length of the frame payload, frame preamble and other PHY-related parameters [5].

2.3 WiMedia MAC

The MAC layer is a sub layer of the Data Link Layer, which provides the interface between logical link control (LLC) layer and PHY layer. Due to the characteristics of UWB PHY layer, WiMedia MAC layer provides specific services for the WiMedia PHY layer, such as synchronization, power management, information transmission, etc. The WiMedia MAC has a fully distributed architecture; hence there is no device in the system that operates as a central controller. The MAC layer manages access of users to the physical medium and supports both reservation and contention based access to the medium.

2.3.1 Superframe Structure

In WiMedia MAC, the transmission time is organized into superframes with the duration of 65535 μs. The superframe is composed of 256 Medium Access Slots (MASs),

each of which lasts 256 μs. The superframe structure is presented in Figure 2.4. Due to different usages, a superframe is divided into two parts: Beacon Period (BP) and Data Transmission Period (DTP). BP is used to reserve MASs, while the remaining MASs are used by DTP to transmit user information. During the DTP, devices can access the channel by using either contention-free Distributed Reservation Protocol (DRP) and/or Prioritized Contention Access (PCA) mechanism.

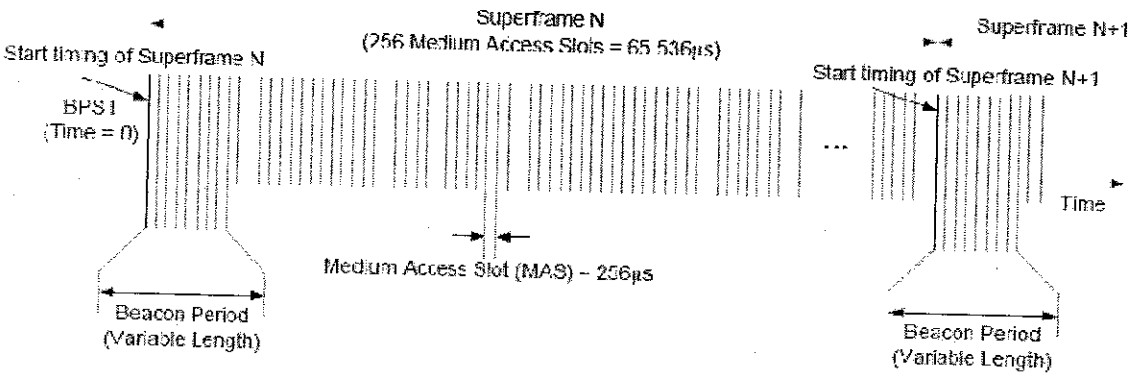


Figure 2.4 WiMedia MAC superframe structure [3]

2.3.2 Beacon Period

As stated earlier, there is no central controller in WiMedia MAC, so each device needs to broadcast its presence and capability by using beacon frames. Devices broadcast and receive beacons during beacon period (BP) to reserve the communication slots. BP contains variable number of beacon slots (BS), which are located at the front end of each superframe. The length of BP is based on the number of devices in the system, where the first two beacon slots are always reserved as signaling slots. As Figure 2.5 illustrates,

beacon slots are numbered starting from zero and cannot exceed the maximum index of the beacon slot defined as mMaxBPLength-1. The duration of a beacon slot is 85 μs, so each MAS contains three beacon slots. The interval between the end of the last beacon slot and the start of the next MAS will not be used.

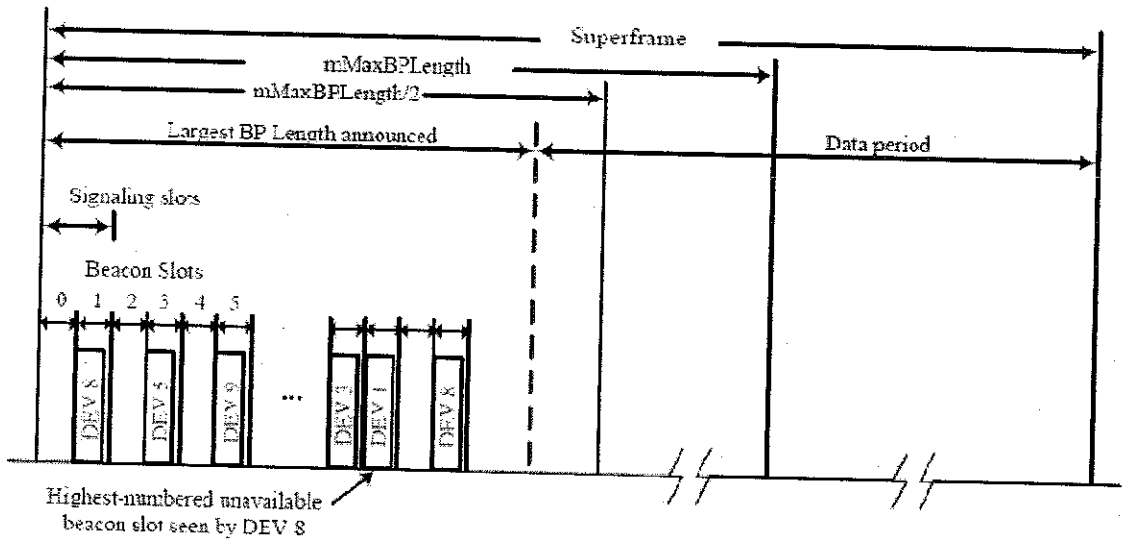


Figure 2.5 An example of BP structure

Before transmitting any frames, a device will scan for beacon period for at least one superframe. If there is no beacon period existing in the network, it will create a new BP and send a beacon in the first beacon slot after the signaling slots. Otherwise, it will choose a BS in the existing BP to transmit its beacon frames in order to exchange the timing and control message with other devices for synchronization and coordinating channel access.

In WiMedia standard, logical groups are formed around each device to facilitate contention free transmission, which are indicated by beacon groups. Any device with the beacon that belongs to a device's beacon group can be seen as the neighbor of this device. A device protects its BP and its neighbor's BP for exclusive use of the beacon protocol. Otherwise, the beacon which does not belong to the beacon group of a device, called alien beacon, may be protected by this device by announcing a reservation in its own beacon frames.

2.3.3 Distributed Reservation Protocol

Distributed Reservation Protocol enables a device to reserve one or more MASs to transmit isochronous traffic during the DTP, while the unreserved MASs in a superframe can be used to transmit frames asynchronously according to the Prioritized Contention Access (PCA) protocol. DRP guarantees the transmission by multiple devices, each of which has the exclusive access to the channel. All devices using DRP for transmissions shall announce the utilization of MASs in the BP of each superframe by sending DRP availability information element (IE) in the BP. The device that wants to initiate a DRP connection will negotiate the reservation with all neighbours to find out the available MASs, which are not necessary to be consecutive.

There are five reservation types specified by DRP IE, which are briefly described below.

- **Alien BP reservation**

MASs can be reserved to protect alien BPs. No frames other than beacons can be transmitted during an alien BP reservation.

- **Hard reservation**

  A device can reserve MASs for exclusive use of the reservation owner and target by using hard reservation. No device other than reservation owner can initiate the frame transactions during the reserved MASs.

- **Soft reservation**

  By using a soft reservation, devices can access the medium with the PCA protocol. The reservation owner can initiate a frame transaction at the beginning of each reservation block with the highest priority. The neighbours of the reservation owner shall access the medium following PCA protocol while the devices other than the reservation owner and its neighbours shall not access the medium.

- **Private reservation**

  Like hard reservation, private reservation also provides the exclusive access to the medium for the reservation owner and targets. But the access protocol and frame exchange sequences are out of scope of the UWB standard [5];

- **PCA reservation**

  In PCA reservation, any device may access the medium using PCA protocol.

2.3.4 Prioritized Contention Access

PCA is a contention-based random access scheme, which uses prioritized CSMA/CA mechanism for medium access. Since the WiMedia MAC is fully distributed, each device needs to sense whether the channel is idle before transmitting. Backoff algorithm is used by PCA protocol to avoid collisions, which is applied based on the number of devices and the traffic priority. For the bursty traffic, PCA is an efficient way to enable sharing of bandwidth by multiple devices.

WiMedia PCA protocol is similar to the EDCA mechanism defined in the IEEE 802.11e draft standard, since both of them provide differential channel access by using different traffic priorities. In PCA, four Access Categories (ACs) have been defined to handle different types of traffic, such as background traffic, best effort traffic, voice transmission and video transmission. Different access parameters are assigned to different ACs, which enable higher priority ACs have more opportunity to access the medium.

2.3.4.1 Frame Transaction

The PCA protocol provides differentiate, distributed contention access to the wireless medium for frames waiting in a device for transmission. Since there is no central controller in the system, the collision occurs when more than one device tries to access the medium during the same slot. Thus, CSMA/CA mechanism is used by PCA to control the access in order to reduce the collision probability. When collision happens, the device

will perform a backoff procedure, the packet will be retransmitted after a random backoff time.

A frame transaction contains a single frame, an Acknowledgement frame (ACK) and an optional Request To Send / Clear To Send (RTS/CTS) frame. There are two modes used in the frame transaction: the basic access mode and the RTS/CTS mode. In basic access mode, each device senses the channel first. If the channel is sensed idle, the device waits for arbitration inter-frame space (AIFS) period and then starts the transmission. Otherwise, the device waits for AIFS period after the channel becomes idle, and then initiates the backoff procedure. In RTS/CTS access mode, before transmitting a packet, the device sends a RTS frame to the receiver to acquire the channel for transmission. Once the receiver receives the RTS frame, it will send a CTS frame back to the transmitter which indicates that the channel is available for their transmission. Then packets can be exchanged between these two devices without the interruption of neighbors. In this mode, the collisions may occur during the transmissions of RTS and CTS frames. If a transmitter does not receive the expected CTS frame, it goes to the backoff procedure.

2.3.4.2 Inter-frame Space

WiMedia PCA defines three types of Inter-frame Spaces (IFSs) depending on the PHY layer, which are the minimum inter-frame space (MIFS), the short inter-frame space (SIFS), and AIFS. The interval between the start of a frame with non-zero length payload and the end of the previous frame cannot be less than MIFS; while the interval

between the frame without payload and the end of the previous frame cannot be less than SIFS.

- **MIFS**

  The Minimum Inter-Frame Space is the duration between frames in a burst, which is related to the burst mode used in the PHY layer.

- **SIFS**

  The Short Inter-Frame Space is the interval used to separate all frames during the frame transmission.

- **AIFS**

  The Arbitration Inter-Frame Space is the interval that a device has to wait for after the channel becomes idle. The device can obtain Transmission opportunity (TXOP) or begin the backoff procedure after AIFS specified by differential AC. The value of AIFS is calculated as follows,

$$AIFS[AC] = pSIFS + mAIFSN[AC] \times pSlotTime$$

  where $AIFS[AC]$, $pSIFS$, $mAIFSN[AC]$ and $pSlotTime$ are management information base (MIB) attributes.

The relationship between AIFS and other IFSs is presented in Figure 2.6, where AC_VO and AC_BE indicate the access categories for voice application and for background traffic respectively. Voice has the highest priority and background traffic has the lowest priority. The higher priority device is assigned the shorter AIFS, thus it will gain faster access to the channel since other devices still waiting for their AIFS to expire.

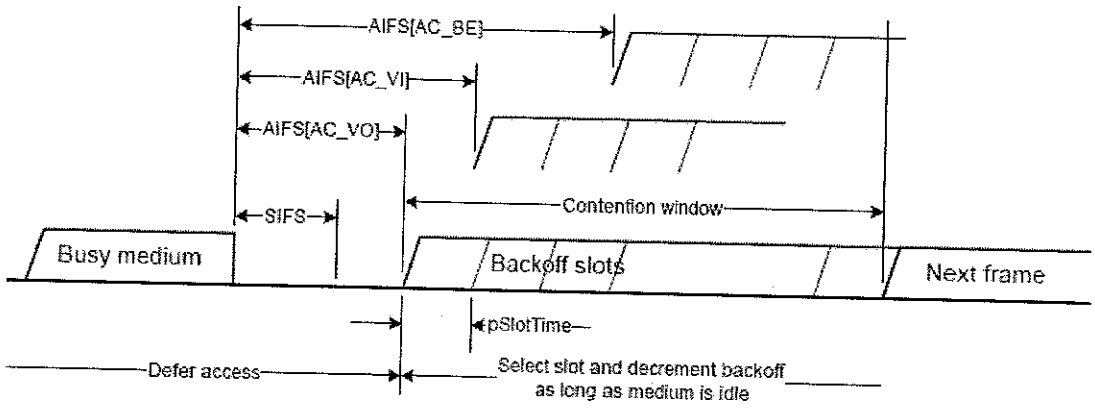


Figure 2.6 IFS relationship for PCA [3]

2.3.4.3 Backoff Procedure

In PCA protocol, the device sets a backoff counter for each AC. The value of the backoff counter is sampled as a uniformly distributed integer from the interval [0, CW], where CW indicates the size of the contention window. The frame transmission is accommodated within PCA slots. All PCA slots have a length of $pSlotTime$, which indicates the time that any station needs to detect the busy condition of the channel, which depends on the PHY layer. A station is allowed to transmit data at the beginning of each PCA slot.

For an AC, the backoff counter shall initiate the CW from the minimum contention window when the backoff counter is invoked, which is specified by $mCWmin[AC]$. When a device senses that the channel has been idle for an $AIFS[AC]$, from then on, it decrements the backoff counter for that AC by one during each slot that the channel remains idle. When the channel is sensed busy, the backoff counter shall be

frozen by the device, until the channel becomes idle again for more than an AIFS. When the value of counter reaches zero, the device can start the transmission. If there is another device that starts to send a packet in the same time slot, the transmission will end up in a collision. The device will choose a new backoff time from a new contention window, and invoke the backoff procedure again. The new size of the backoff window is set to be the smaller of $2 \times CW[AC] + 1$ and $mCWmax[AC]$, where $CW[AC]$ and $mCWmax[AC]$ indicate the last CW value for this AC and the maximum size of contention window for this AC respectively. The values of $mCWmin[AC]$ and $mCWmax[AC]$ are specified for the PHY layer which are reported in the WiMedia standard [3]. The above procedure is repeated until either the packet is transmitted successfully or reaches the retry limit. If the retry limit is reached and the transmission still ends up in a collision, then the packet is discarded.

2.4 Literature Survey

Next, we discuss the previous work on the performance modeling of WiMedia. First, we review the existing work on DRP. Several modelings of DRP have been attempted through extensions of the analyses of centralized TDMA protocol [10][11][12]. In [10], stations were assumed to have infinite buffers and packets are transmitted during the slots reserved for them by using a conflict-free distributed protocol. It has been suggested that uniform slot reservation has better performance than contiguous slot reservation protocol. In [11], circuit-switched TDMA (CS-TDMA) is used for multiple-

access communications to share the channel. Each station is allocated fixed number of slots during each frame. The generating function of the queue size and the waiting time distribution has been derived using a discrete-time G/G/N queuing system. In [12], the effects of arbitrary traffic patterns on the performance of a generalized TDMA protocol have been studied through queuing theory. It is shown that the queue size can be optimized by judiciously choosing slot assignment among heterogeneous stations. In the centralized system, the assignment of slots in the superframe is desired to be either contiguous or uniformly in order to reduce the delay variation. However, such a reservation pattern is not applicable in distributed systems, because the availability of time slots in distributed systems is random. In [13], the DRP mode has been modeled as a multi-server queue with Poisson packet arrival process and constant packet transmission times. The number of servers is determined by the number of reserved MASs in a superframe and reserved MASs may not be contiguous. The main objective of the paper is to study the effect of the uneven reservation pattern of MASs on the queue size. We think that DRP has been introduced to provide users with deterministic transmission capacity when they have a significant amount of information to transfer. Therefore, the assumption that queue size will reach to steady-state may not be applicable. The modeling of DRP as a queue has also been studied in [14] using matrix geometric technique. The system has been modeled as a MAP/PH/1 queue with the number of reserved MASs chosen as the number of phases of PH distribution [14]. As in the preceding work, assuming constant packet transmission time they determine the steady-state queue size and packet waiting time distributions.

In parallel to the study of DRP, many attempts have been made on performance modeling of the PCA protocol. In the literature, there are many works on the analysis of the performance of CSMA/CA mechanism based protocols. One widely used analytical model is a bi-dimensional Markov chain model proposed by Bianchi for the Distributed Coordination Function (DCF) scheme in IEEE 802.11 [15][16]. This model assumes that all stations in the system are saturated, which means that each station has at least one packet waiting to be sent all the time. Also, the probability of collision is assumed to have a constant value, which is independent of the number of collisions experienced until that time. The result of the work is accurate as long as the contention window size and the number of stations are high. Moreover, Bianchi assumed that the backoff counter is decremented in each slot no matter whether the channel is sensed idle or not. But in fact, the backoff counter will be frozen when the channel is sensed busy and will be reactivated when it becomes idle again. Subsequently, some models have been proposed which extend the Bianchi's work. In [17], Ziouva and Antonakopoulos extended the analysis to determine both the throughput and the saturated delay of IEEE 802.11. The results are more accurate because counter freezing has been accounted for. In [18], Wu et al. has extended the work in [16] to determine the saturated throughput of 802.11 with an upper limit on the number of retransmissions of a frame. In [19], Xiao presented the throughput upper limit and the delay lower limit for IEEE 802.11. In [20], a model was proposed by Malone et al. to analyze the throughput under non-saturated and heterogeneous condition. The works in [16][17][18][19] do not model multiple transmission rates or traffic with multiple classes of priority for QoS support.

To support QoS in MAC layer, EDCA mechanism was introduced to the IEEE 802.11e standard which is a priority based scheme. In [21], another paper from Xiao, a model is proposed to derive the performance for different EDCA services under the saturation condition. In the model, different classes of traffic employ different initial and maximum window sizes and retry limits. The model fails to take into account inter-frame space. In [22], Tsang et al. presented an analytical model that investigated the effects of the window size and AIFS on the throughput and mean access delay. PCA contention-based scheme is very similar to the EDCA in IEEE 802.11e; they both use CSMA/CA mechanism and provide QoS service. As a result, many of the proposed models for PCA are extensions of the models of EDCA. In [23], a model was proposed for Distributed Prioritized MAC Protocols for both saturation and non-saturation cases. It discussed the analysis for a network with two access classes and calculated the throughput and average packet service time per station. The minimum contention window size was fixed for all access categories; the results are only relevant to the impact of AIFS.

The works mentioned above deal only with DRP or PCA MAC, and there are only few works that attempt to analyse the two protocols simultaneously. In [24], the performance of PCA scheme has been analyzed in the presence of DRP traffic. It has been assumed that DRP traffic occupies constant number of MASs in a superframe. Under the assumption that each MAS is equally likely to serve to DRP traffic, they derive the distribution of the contiguous number of MASs available for PCA traffic. They assume that there are N priority classes of PCA traffic at each station. Then, they model the behaviour of the PCA scheme using a three dimensional Markov chain. The analysis takes into account that a PCA period may not have a sufficient duration for the

transmission of a packet because of contention. They determine the saturation throughput for each priority class of traffic with the number of PCA stations as a parameter. The main drawbacks of this work are that the number of MASs allocated for DRP traffic is constant and also the PCA traffic has been assumed to be under saturation. In [25], the effect of DRP on PCA frame service time was presented. Four DRP clusters are assumed in a superframe with different percentages: 6.25%, 12.5%, 25% and 50%. In neither of these works, the analyses allow stations to have different transmission rates.

2.5 Main Contributions of the Thesis

Next, we describe the main contributions of the thesis,

- As may be seen from the literature, most of previous works studied the DRP and PCA protocols separately under the saturation condition. However, from the characteristics of UWB MAC, the DRP provides QoS support for devices while PCA enables immediate access to the channel, both protocols are important for analyzing the system performance. Further, the bandwidth available for PCA usage is dependent on the bandwidth reserved by DRP. Therefore, it is more appropriate to analyze them jointly.
- Most of the previous work on DRP assumes that the DRP traffic is queued which will not be suitable to real-time applications. We present a blocking model of DRP scheme that is more appropriate for stream type applications.
- We present an analysis of PCA traffic under non-saturation condition.

- Our model is designed for multi-rate wireless network.

2.6 Thesis Organization

The remainder of this thesis is organized as follows,

Chapter 2 presents an overview of the WiMedia UWB and describes both physical (PHY) and Medium Access Control (MAC) layers of this standard in detail. It also gives a literature review of the previous modeling relevant to UWB MAC and provides the contribution and the organization of this thesis.

Chapter 3 presents a performance modeling of distributed reservation protocol (DRP). We assume that DRP traffic consists of several classes of calls and each class of calls arrives according to a Poisson process. Classes are determined through their transmission rates, bandwidth requirement and call durations. We determine the joint distribution of number of calls from each class in the system and then determine the call blocking probabilities for each class and the bandwidth utilization of the system.

Chapter 4 presents a performance modeling of the PCA traffic in the presence of DRP traffic under the non-saturation condition. We assume that there are four access categories in the system and model each access category in a station as an M/G/1 queuing. Then, we present a discrete-time Markov chain model of packet service in each queue and determine the distribution of packet service time. Finally, we determine the system throughput and the mean of total packet delay in an access queue.

Chapter 5 provides a brief summary and conclusions of the thesis along with some suggestions for future research.

CHAPTER 3

A PERFORMANCE MODELING OF DRP

In this Chapter, we will give a performance modeling of DRP protocol. We assume that DRP traffic consists of multiple classes of calls. Each class of calls arrives according to a Poisson distribution and they differ in their bandwidth requirements and transmission rates. We determine the joint distribution of number of calls from each class in the system, then the call blocking probabilities and the bandwidth utilization.

3.1 Background

As described in Chapter 1, DRP traffic has priority over PCA traffic in accessing the channel. The channels not reserved by DRP can be accessed by the station using PCA protocol. We assume that a minimum number of channels are reserved for PCA traffic to prevent overflowing of PCA queues. In the analysis, we neglect the BP duration without loss of generality and assume that all the channels in a superframe are available to transmit data.

Let us make the following definitions,

T_{MAS} : duration of a superframe slot, $T_{MAS} = 256\mu s$

T : duration of a superframe measured in microseconds

N_{SF} : number of channels in a superframe, $N_{SF} = 256$

N_o : number of channels reserved for PCA traffic in a superframe

N_{SD} : maximum number of channels available for DRP traffic in a superframe

N_{DRP} : number of channels currently reserved for DRP traffic in a superframe

N_{PCA} : number of channels available for PCA traffic in a superframe

Then we have the following relation for the duration of a superframe,

$$T = T_{MAS} N_{SF} \tag{3.1}$$

The maximum number of channels available for DRP traffic and the number of channels currently available for PCA traffic can be expressed as follows respectively,

$$N_{SD} = N_{SF} - N_o \tag{3.2}$$

$$N_{PCA} = N_{SF} - N_{DRP} \tag{3.3}$$

Let P_{PCA} denote the probability that currently an MAS belongs to a PCA period, then we have,

$$P_{PCA} = \frac{N_{PCA}}{N_{SF}} = \frac{N_{SF} - N_{DRP}}{N_{SF}} \tag{3.4}$$

3.2 Model Assumptions

We assume that there are K DRP classes of calls that access the network and we let C_k denote class k, where $k = 1, \ldots, K$. Classes are determined through their transmission rates, bandwidth requirements and call durations. We assume that the arrival of each class of calls is according to a Poisson process. We note that a new arriving call will be lost if there are no sufficient number of unreserved channels available to meet the bandwidth requirement of the call.

Let us define,

α_k : mean arrival rate of class k calls measured in calls per second

α : total arrival rate of calls measured in calls per second

μ_k : service rate of class k calls measured in calls per second

$\bar{X}_k$: mean duration of class k calls in the system

ρ_k : traffic load offered by class k calls

b_k : number of channels per superframe reserved for a class k call, $b_k \leq N_{SD}$

n_k : number of class k calls in the system

Γ_k : size of information payload of class k calls in number of bits and $\bar{\Gamma}_k$ is its average

Δ : number of distinct physical transmission rates

R_k : physical transmission rate of class k calls

F_k : number of bits per superframe that may be transferred for a class k call

$\overline{N}_k$: average number of superframes needed to transmit a class k call

Since class k calls arrive according to a Poisson process with the rate of α_k calls per second, we have the following relation for the total arrival rate of calls,

$$\alpha = \sum_{k=1}^{K} \alpha_k \tag{3.5}$$

Then, the probability that a call belongs to class k denoted by θ_k can be given by,

$$\theta_k = \frac{\alpha_k}{\alpha} \tag{3.6}$$

We assume that the information payload size of a class k call is exponentially distributed. Next we explain calculation of the mean duration of class k calls.

As we defined above, a class k call requires b_k channels per superframe to transmit its information with the channel transmission rate of R_k and slot duration of T_{MAS}. So, the number of bits per superframe transferred for a class k call is equal to b_k times the number of bits transferred per channel, which is given by,

$$F_k = b_k * T_{MAS} * R_k \tag{3.7}$$

Then, by dividing the average size of information payload of a class k call with the number of bits per call per superframe, we obtain the average number of superframes needed to transmit a class k call,

$$\bar{N}_k = \frac{\bar{I}_k}{F_k} \tag{3.8}$$

Since each superframe last T seconds, the average duration of the class k calls will be $\bar{N}_k T$, given by,

$$\bar{X}_k = \bar{N}_k * T \tag{3.9}$$

Since information payload size is exponentially distributed, we assume that service time of a class k call is also exponentially distributed with parameter μ_k, which is given by,

$$\mu_k = \frac{1}{\bar{X}_k} \tag{3.10}$$

Then, the traffic load offered by class k calls can be calculated by,

$$\rho_k = \frac{\alpha_k}{\mu_k} \tag{3.11}$$

3.3 The Steady-state Probability Distribution of the Number of DRP Calls in the System

Next, we will use the available results in the literature to determine the joint distribution of the number of calls from each class in the system at the steady-state. In circuit-switch networks, bandwidth is assigned permanently for the call duration and the information is transferred in order. In DRP scheme, since the channels need to be reserved before a call is accepted, the system can be seen as a circuit-switched network and the probability of call blocking is the relevant criterion, specifically. Therefore, we will use existing results on call blocking for circuit-switched networks to determine the performance of DRP scheme [31][32][33]. In this section, first, we will give the state description of the network. We assume that the bandwidth is shared completely by all the calls.

Since n_k denotes the number of class k calls in the system, then the network state of the DRP calls may be denoted by the vector n,

$$n = (n_1, \dots, n_k, \dots, n_K)$$

We let the n_k^+ and n_k^- denote the network states, in which the k'th component differ from the state n by plus or minus unity respectively,

$$n_k^+ = (n_1, \dots, n_{k-1}, n_k + 1, n_{k+1}, \dots, n_K)$$

$$n_k^- = (n_1, \dots, n_{k-1}, n_k - 1, n_{k+1}, \dots, n_K)$$

Then, the number of channels in use by DRP traffic is given by,

$$N_{DRP} = n \cdot b = \sum_{k=1}^{K} n_k b_k \tag{3.12}$$

Letting Ω denote the set of permissible network states for DRP calls, then, it is given by,

$$\Omega = \{n | 0 \leq n \cdot b \leq N_{SD}\} \tag{3.13}$$

Let us define,

$$\delta_k^+ = \begin{cases} 1, & n_k^+ \in \Omega \\ 0, & \text{otherwise} \end{cases} \quad , \quad \delta_k^- = \begin{cases} 1, & n_k^- \in \Omega \\ 0, & \text{otherwise} \end{cases}$$

Let $P(n)$ denote the probability that there are n calls in the system. We may model the state of the network as a multi-dimensional birth death process. The arrival of class k calls with the rate of α_k can be seen as a birth process while the departure of class k calls following call completion can be seen as a death process. Then the global balance equation (GBE) for state n may be written as,

$$P(n) \sum_{k=1}^{K} (\alpha_k \delta_k^+ + n_k \mu_k \delta_k^-) = \sum_{k=1}^{K} \alpha_k \delta_k^- P(n_k^-) + \sum_{k=1}^{K} (n_k + 1) \mu_k \delta_k^+ P(n_k^+) \tag{3.14}$$

The right-hand side of (3.14) is the total probability flow into state n, while the left-hand side is the total probability flow out of state n.

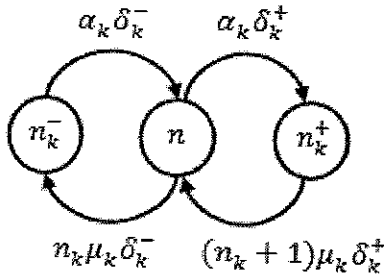


Figure 3.1 The state-transition diagram of the system for state n

The local balance equation (LBE) for state n is given by,

$$n_k\mu_k\delta_k^- P(n) = \alpha_k P(n_k^-) \tag{3.15}$$

The above equation balances the flow from state n to state n_k^- with the flow from state n_k^- into state n. The state transition flow diagram is presented in Figure 3.1, where the flows entering and leaving the states are depicted. Then we obtain the joint probability distribution of the number of calls in the system by a simple iteration on (3.15), given by,

$$P(n) = \frac{1}{G(\Omega)} \prod_{k=1}^{K} \frac{\rho_k^{n_k}}{n_k!} \tag{3.16}$$

where, $G(\Omega)$ is the normalization constant, given by,

$$G(\Omega) = \sum_{n\in\Omega} \left(\prod_{k=1}^{K} \frac{\rho_k^{n_k}}{n_k!} \right) \tag{3.17}$$

The traffic load offered by class k calls ρ_k , is given by (3.11).

3.4 Blocking Probability

Next, we will determine the blocking probability seen by class k calls. A class k call will be blocked if there is no sufficient number of idle channels to meet the bandwidth requirement of this call. Let us define blocking probability of a new arriving class k call as Q_k, then, is given by

$$Q_k = \sum_{\{n \in \Omega | n_k^+ \notin \Omega\}} P(n) \tag{3.18}$$

In the above, we are summing up the probabilities of the states that a call will be blocked, then, the above equation can be rewritten as,

$$Q_k = \frac{G(n \in \Omega | n_k^+ \notin \Omega)}{G(\Omega)} \tag{3.19}$$

Let us define

$$\Omega' = \{n | 0 \leq n \cdot b \leq N_{SD} - b_k\} \tag{3.20}$$

and

$$G(\Omega') = \sum_{n \in \Omega'} \left(\prod_{k=1}^{K} \frac{\rho_k^{n_k}}{n_k!} \right) \tag{3.21}$$

Then, the blocking probability Q_k may be written as,

$$Q_k = 1 - \frac{G(\Omega')}{G(\Omega)} \tag{3.22}$$

Let us define $q(c)$ as the probability that the number of channels reserved by the DRP traffic N_{DRP}:

$$q(c) = \Pr(N_{DRP} = c \text{ MASs})$$

From (3.16), the distribution of c is given by

$$q(c) = \sum_{\{n:n\cdot b=c\}} \prod_{k=1}^{K} \frac{\rho_k^{n_k}}{n_k!} \frac{1}{G(\Omega)} \tag{3.23}$$

From (3.15), the LBE can be written as

$$E(n_k|c)\mu_k q(c) = \alpha_k q(c - b_k) \tag{3.24}$$

where $E(n_k|c)$ denotes the expectation value of the number of class k calls given that the total number of channels reserved for DRP traffic is $N_{DRP} = c$.

Since $\rho_k = \alpha_k/\mu_k$, we have,

$$E(n_k|c)q(c) = \rho_k q(c - b_k) \tag{3.25}$$

By applying the iteration on (3.24), we can determine the probability distribution of number of channels busy with DRP traffic in a superframe as follows,

$$\sum_{k=1}^{K} \rho_k b_k q(c - b_k) = \sum_{k=1}^{K} b_k E(n_k|c)q(c)$$

$$= q(c)E(\sum_{k=1}^{K}(b_k n_k|c))$$

$$= cq(c)\ ,\ c = 0,1,\dots,N_{SD} \tag{3.26}$$

, where $q(x) = 0$ for $x < 0$ and

$$\sum_{c=0}^{N_{SD}} q(c) = 1 \tag{3.27}$$

Let us define

$$q(c) = q(0)\hat{q}(c),\ \ c = 1,\dots,N_{SD} \tag{3.26}$$

$$\hat{q}(0) = 1 \tag{3.27}$$

Substituting the above equations into (3.24), we have,

$$\hat{q}(c) = \frac{1}{c}\sum_{k=1}^{K}\rho_k b_k \hat{q}(c-b_k), \qquad c = 0,1,\dots,N_{SD} \tag{3.28}$$

From the normalizing condition, $\sum_{c=0}^{N_{SD}} q(c) = 1$, is required, we have,

$$q(0) = \left\{\sum_{c=0}^{N_{SD}} \hat{q}(c)\right\}^{-1} \tag{3.29}$$

Then, the blocking probability in (3.18) can be rewritten as,

$$Q_k = \Pr\ (N_{SD} - N_{DRP} < b_k) = \sum_{k=0}^{b_k-1} q(N_{SD}-k)\,, \qquad k = 1,..,K \tag{3.30}$$

3.5 Channel Utilization

Next, we will determine the channel utilization of the DRP traffic, which is the average number of busy channels in a superframe denoted by $\bar{N}_{DRP}$. In the section 3.4, we calculated $q(c) = \Pr(N_{DRP} = c)$ for all $c = 0,1, \ldots, N_{SD}$. By using the result in (3.26), the bandwidth utilization is given by,

$$\bar{N}_{DRP} = \sum_{c=0}^{N_{SD}} [c * q(c)] \tag{3.31}$$

3.6 Numerical Results

In this section, we present some numerical results regarding the analysis of this chapter. We assume that the calls have constant length payloads, $\bar{\Gamma}_k = \Gamma$. For simplicity, we also assume that the PHY and MAC layer overheads are included in Γ. The values of WiMedia MAC sublayer parameters used in the numerical examples are tabulated in Table 3.1 without loss of any generality, we assume that $N_o = 0$.

Parameter	Value
N_{SF}	256

T_{MAS}	256 μs
DRP payload length (Γ)	1 Mb

Table 3.1 Parameters used in DRP scheme

3.6.1 Call Blocking Probability

First, we present numerical results for the blocking probability with K = 8, 10 classes of calls with class parameters given in Tables 3.2 and 3.3 respectively. As may be seen, the transmission rate of each class corresponds to one of the physical rates defined in the standard and the bandwidth requirement of each call class is increasing with the decreasing transmission rate.

3.6.1.1 *Example 1*

For each value of k, we consider three cases, which differ from each other through channel requirements and call duration. Mean duration of a call has been determined from (3.11).

C_k	R_k	λ_k	Case I		Case II		Case III	
			b_k	$\bar{X}_k$	b_k	$\bar{X}_k$	b_k	$\bar{X}_k$
1	53.3	0.1α	8	0.6004	16	0.3002	32	0.1501
2	80	0.1α	7	0.4571	14	0.2286	28	0.1143
3	106.7	0.1α	6	0.3999	12	0.1999	24	0.1000
4	160	0.1α	5	0.3200	10	0.1600	20	0.0800
5	200	0.1α	4	0.3200	8	0.1600	16	0.0800
6	320	0.1α	3	0.2667	6	0.1333	12	0.0667
7	400	0.2α	2	0.3200	4	0.1600	8	0.0800
8	480	0.2α	1	0.5333	2	0.2667	4	0.1333

Table 3.2 Parameter values of classes of calls with $K = 8$

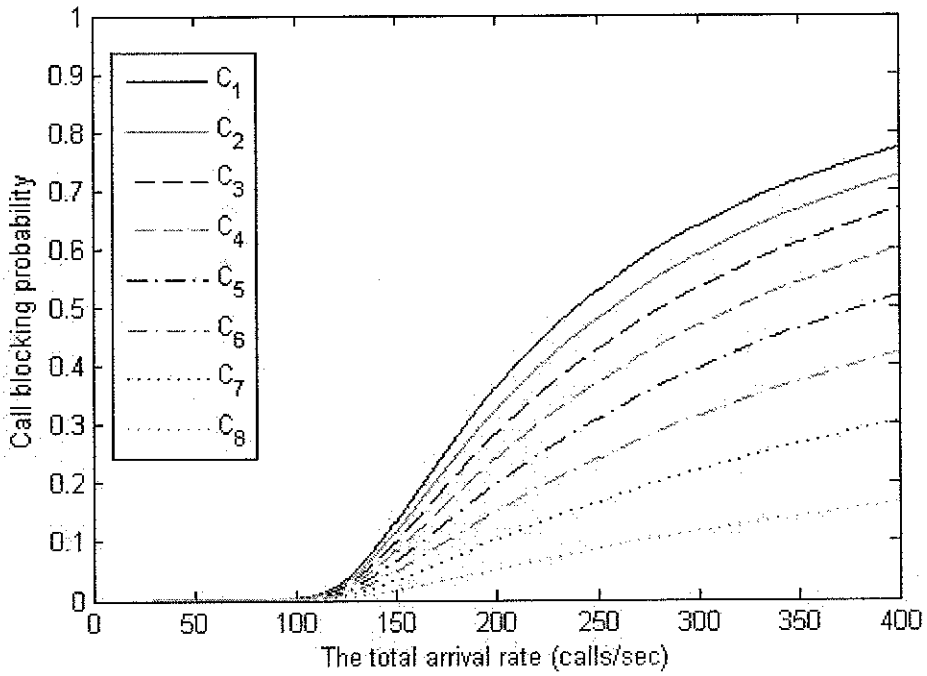


Figure 3.2 Blocking probability versus total arrival rate for Case I with $K = 8$

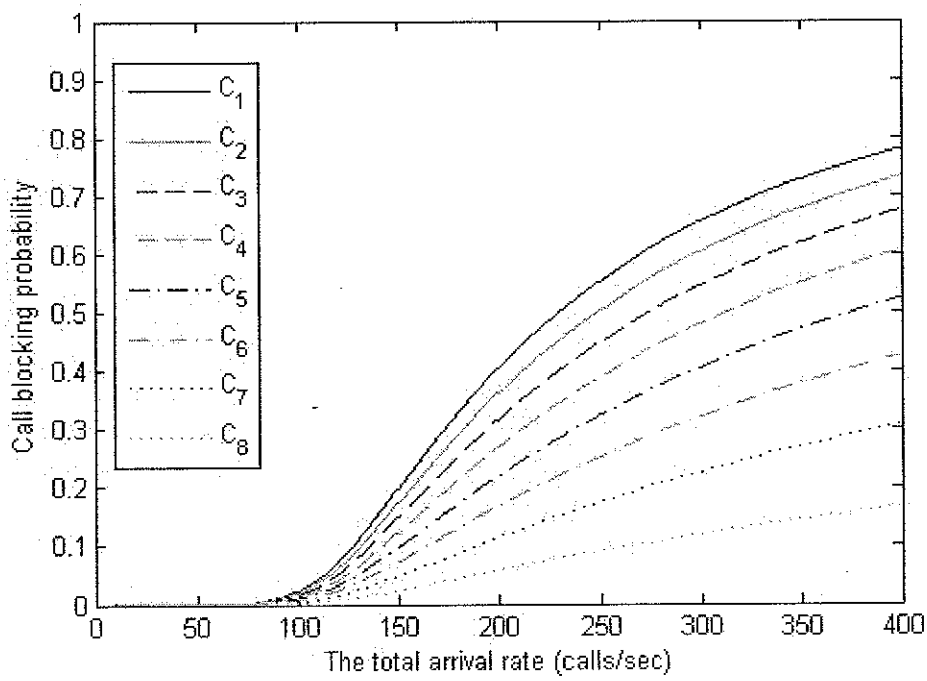


Figure 3.3 Blocking probability versus total arrival rate for Case II with $K = 8$

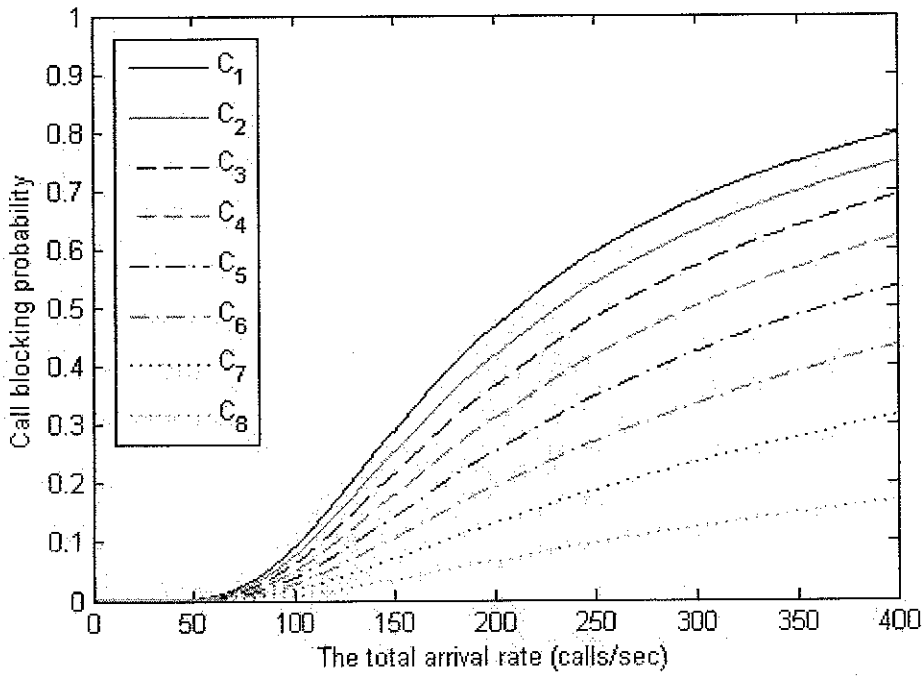


Figure 3.4 Blocking probability versus total arrival rate for Case III with $K = 8$

In Figures 3.2, 3.3, 3.4, we present call blocking probabilities for different DRP classes as a function of the total call arrival rate. It may be seen that the blocking probabilities increase with the call arrival rate and the classes with higher bandwidth demands experience higher blocking probabilities.

3.6.1.2 *Example 2*

In this example, the total call arrival rate is assumed to be constant, which is set to 100 calls/sec with $K = 8$ classes of calls. We vary the payload size per call from 0 to 10 Mbits. The call blocking probabilities are determined for the parameter values of the classes of Case I given in Table 3.2.

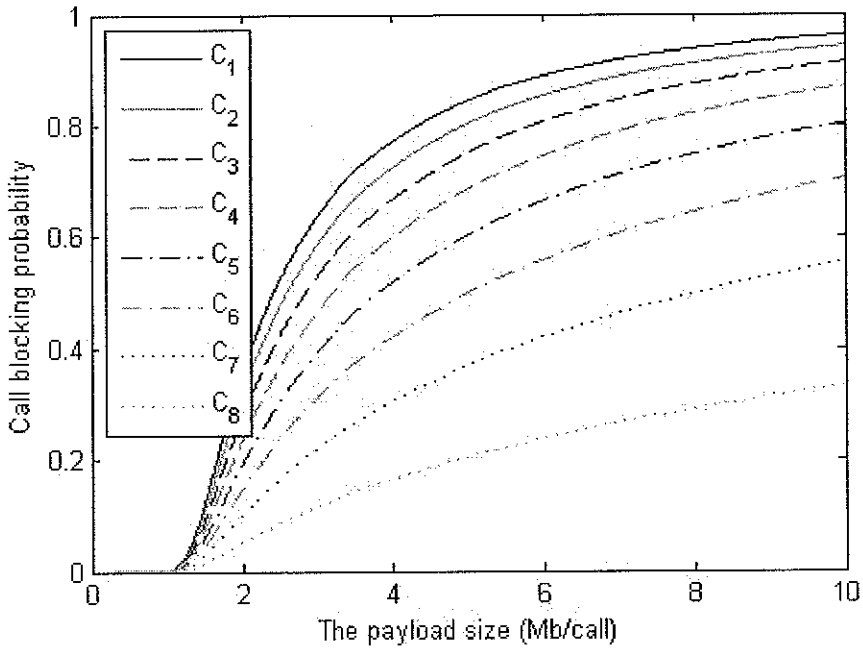


Figure 3.5 Blocking probability versus the payload size per call for Case I with $K = 8$

Figure 3.5 presents the blocking probabilities as a function of the information payload size per call, Γ, for Case I. As the size of payload increases, the blocking probabilities also increase.

3.6.1.3 *Example 3*

In this example, we illustrate that each class of calls with different data rates can be further divided into sub-classes. We assume that there are $K = 10$ classes in the system. The parameter values are shown in Table 3.3 and the result is shown in Figure 3.6. As may be seen, classes 7 and 8 have identical data rates but different channel requirements. Classes 9 and 10 also exhibit this property.

C_k	R_k	λ_k	b_k	X_k
1	53.3	0.1α	10	0.4803
2	80	0.1α	9	0.3556
3	106.7	0.1α	8	0.2999
4	160	0.1α	7	0.2286
5	200	0.1α	6	0.2133
6	320	0.1α	5	0.1600
7	400	0.1α	4	0.1600
8	400	0.1α	3	0.2133
9	480	0.1α	2	0.2667
10	480	0.1α	1	0.5333

Table 3.3 Parameter values of 10 classes of calls

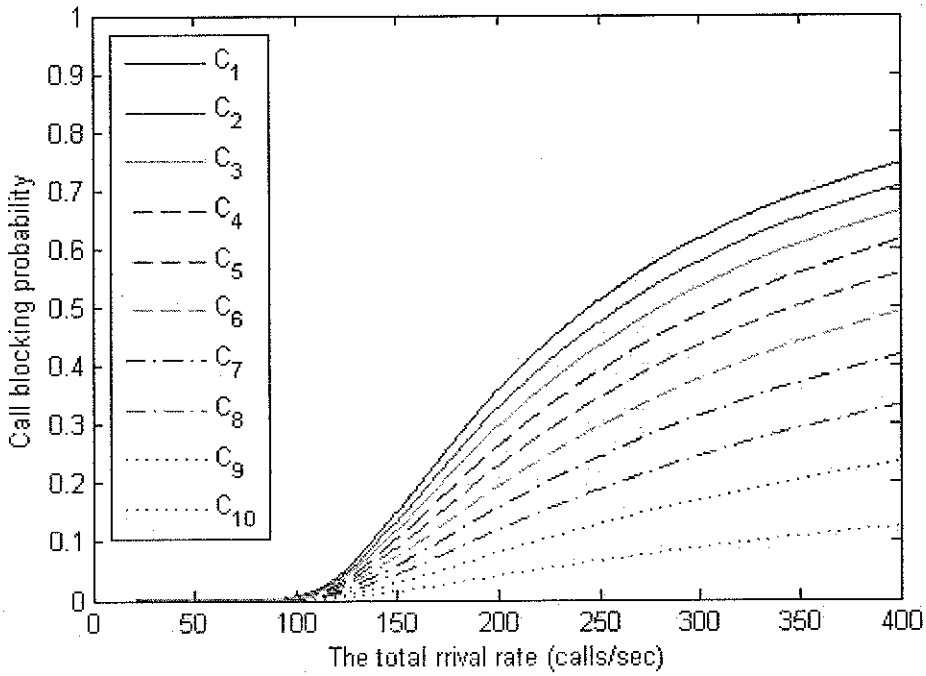


Figure 3.6 Blocking probability versus total arrival rate for 10 classes of calls

From Figure 3.6, two classes of calls with the same data rate but different channel requirements experience different blocking probabilities. The call blocking probabilities increase with bandwidth demand.

3.6.2 Average number of busy channels

In this section, we determine the average number of busy channels in a superframe, $\bar{N}_{DRP}$, from (3.26) and (3.31). We present the result for K = 8 classes with class parameter values of Case I given in Table 3.2.

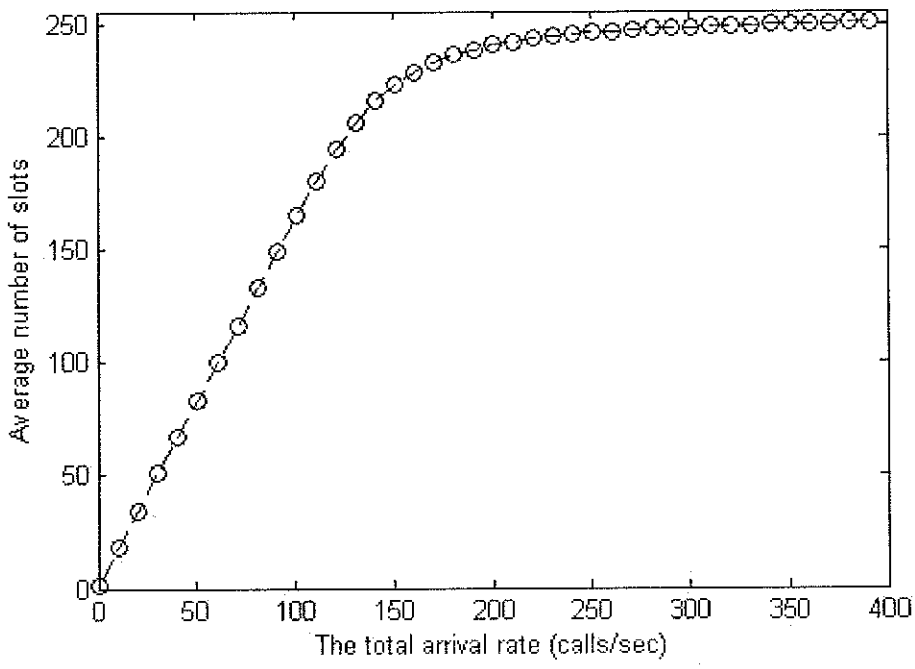


Figure 3.7 Average number of busy slots in a superframe as a function of total call arrival rates, α

Figure 3.7 shows the average number of channels occupied by DRP traffic as a function of the total call arrival rate of the calls to the system. We observe that the average number of busy channels increases with the total call arrival rate, reaching almost to the number of channels in a superframe, $N_{SF} = 256$.

CHAPTER 4

A PREFORMANCE MODELING OF PCA

In the previous chapter, we presented a performance modeling of DRP scheme by assuming multiple physical transmission rates. Since the number of slots reserved by DRP during a superframe is not constant, the number of remaining channels available for PCA will vary, therefore the performance of PCA traffic depends on the state of DRP mode. In this chapter, we present a performance modeling of the PCA scheme in the presence of DRP traffic. We model each access category in a station as an M/G/1 queue, and use a Markov chain model to derive the packet service time distribution. Then, we determine the mean of total packet delay in an access queue. Each access queue serves to the packets on a First Come First Served (FCFS) basis, thus new arriving packets to a non-empty queue has to wait their service turn.

4.1 Model Assumptions

We assume that a minimum number of channels are reserved for PCA traffic to prevent overflowing of PCA queues, which is denoted by N_o as mentioned in Chapter 3. PCA traffic will use N_o channels reserved for its exclusive use as well as any channels currently not in use by DRP traffic. The number of channels available for PCA traffic in

any superframe is given by N_{PCA} in (3.2). Let $\overline{N}_{PCA}$ denote the average number of channels available for PCA traffic in any superframe. Since we have determined the average number of busy MASs for DRP traffic in (3.31), we can easily obtain $\overline{N}_{PCA}$.

The performance of the system will be carried out through the extension of the analysis of 802.11e under non-saturation condition to WiMedia PCA scheme. While in 802.11e there is a single channel serving the traffic, as we have seen in WiMedia, there may be multiple channel transmission capacity available to serve PCA traffic. In order to make use of the results for 802.11e, we will assume that PCA traffic will be served by a single equivalent channel which is N_{PCA} times faster, instead of N_{PCA} multiple channels in parallel. This approximation will be good under heavy traffic when all the channels available for PCA system are busy. On the other hand, under light traffic it will give a performance better than expected, since in practice, when the number of busy queues is less than the number of available channels only a subset of available channels will be used. Clearly, from the application's point of view the behaviour of the system under heavy traffic is more important. In the following, we will draw from the IEEE 802.11e EDCA analysis in [26][27][28][29][30].

We assume that there are η PCA stations in the system with homogeneous traffic arrivals across stations with λ packets/sec/station. It is assumed that the arrival traffic to each PCA station consists of N access categories, which will be denoted by AC_i, $i = 0 \ldots N-1$. Access category i, AC_i, will have the priority level i. We note that $i = 0$ has the lowest priority level. The arrival of traffic to each AC_i in a PCA station is assumed to follow a Poisson process with the rate of λ_i packets per second. PCA stations

maintain separate queues for each AC_i, as shown in Figure 4.1. We will assume that all ACs have equal arrival rates.

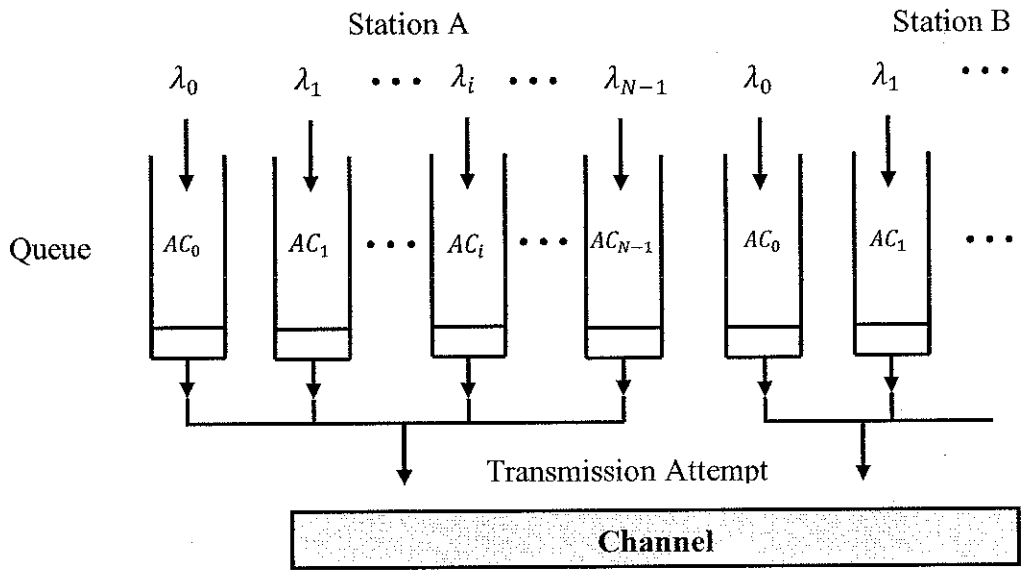


Figure 4.1 Model overview

Let λ_T denote the total arrival rate of packets to all stations, then, λ_i, λ and λ_T are related to each other through the following expressions,

$$\lambda = \frac{\lambda_T}{\eta} \tag{4.1}$$

$$\lambda_i = \frac{\lambda}{N} \tag{4.2}$$

During PCA periods, the channel will have idle and busy periods. The time axis will be slotted, which will be referred as PCA slots. A PCA slot will have time duration

of T_e sec and it will be referred to as being empty if it is not part of a successful packet transmission or a collision.

We will define a generic slot as an empty PCA slot, a successful packet transmission time or a collision interval. Let t denote the time in number of generic slots. The system assigns different initial backoff window sizes to each access category (AC) for handling AC priorities. We let $CW_{i,min}$ and $CW_{i,max}$ denote the initial minimum and maximum backoff window sizes for AC_i. Then, we let $s(i,t)$ denote the random process representing the backoff stage j ($j = 0,1,\dots,L_{i,retry}$) for AC_i at time t, where $L_{i,retry}$ is the retry limit. The service time of a packet will begin with the value of the backoff stage set to zero and the value will increase by one each time a collision happens during the transmission. When the number of trials of a packet reaches its maximum value and the transmission still ends up in a collision, the packet will be discarded. After a successful transmission or discarding of a packet, the backoff stage is reset and the above procedure is repeated with the next packet.

Let $b(i,t)$ be the random process representing the value of the backoff counter for AC_i at time t. The initial counter value for backoff stage j is uniformly sampled from the values $\{0,1,\dots,W_{i,j}\}$, where $W_{i,j}$ is set to the smaller of $CW_{i,max}$ and $2W_{i,j-1}+1$ for $j \geq 1$ and $W_{i,0} = CW_{i,min}$. Then, the value of the backoff window size for AC_i may be expressed as,

$$W_{i,j} = \begin{cases} (CW_{i,min}+1)2^j - 1 & , \ 0 \leq j \leq m_i - 1 \\ CW_{i,max} & , \ m_i \leq j \leq L_{i,retry} \end{cases} \tag{4.3}$$

where,

$$m_i = \log_2 \frac{CW_{i,max} + 1}{CW_{i,min} + 1}$$

At the beginning of each empty PCA slot, the counter value of AC_i will be decremented by one. If the channel is sensed busy, the backoff counter will be frozen until the channel is sensed idle again. A frozen period contains either a successful packet transmission or a collision. There may be repeated frozen periods between two consecutive empty PCA slots. Once the backoff counter value reaches zero, the station attempts to transmit the packet.

The packet in service will be either successfully transmitted or discarded, following that the AC_i will go through post-backoff procedure, irrespective of the status of its queue. The timing structure of packet service is presented in Figure 4.2. Post-backoff procedure is to ensure fairness among queues; otherwise currently transmitting queue may dominate the channel. The post-backoff procedure is implemented by choosing a uniformly distributed value from $\{0, 1,\ldots, W_{i,0}\}$ and counting down to zero. We will assume that the service time of a packet is completed at the end of the following post-backoff procedure. The service time of the next packet begins at the end of the current backoff procedure if the AC_i queue is nonempty at the conclusion of this post-backoff procedure or as soon as a packet arrives at an empty queue. A packet will be transmitted immediately if its service time begins following completion of post-backoff procedure. On the other hand, if service time begins when a packet arrives at an empty queue, if the channel is idle, packet transmission will begin, otherwise a backoff process will be initiated.

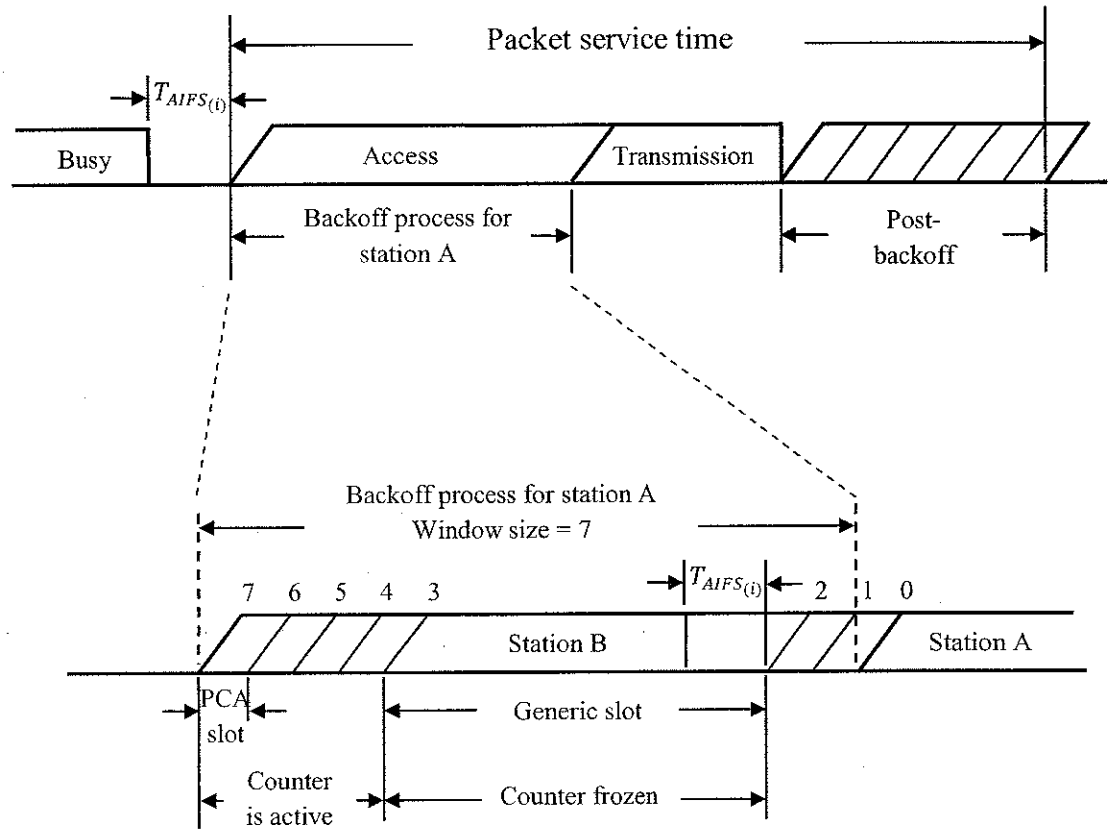


Figure 4.2 The timing structure of packet service from station A

Next we summarize steps of the protocol implemented by an AC_i,

- **Step 1.** Service time of the next packet begins when the post-backoff procedure of the current packet is completed if the AC_i queue is nonempty at that time, otherwise as soon as the next packet arrives to the empty queue.
- **Step 2.** If the service time of a packet begins following completion of post-backoff procedure, this packet will be transmitted immediately. On the other hand, if service time begins when a packet arrives at an empty queue, the station senses

the channel. If the channel is idle, packet transmission will begin, otherwise a backoff process will be initiated.

- **Step 3.** The backoff counter is decremented during empty PCA slots. When backoff counter value reaches to zero, the packet is transmitted.
- **Step 4.** If packet transmission is successful, the station goes into post-backoff procedure. At the end of the post-backoff procedure, packet service is completed.
- **Step 5.** If packet transmission ends in a collision and packet try limit has not been reached, backoff stage is incremented by one and a new initial backoff value is chosen from values $\{0, 1,\ldots, W_{i,j}\}(j = 1, \ldots, L_{i,retry})$. Then, go to step 3. If packet try limit has been reached, the packet is discarded and post-backoff procedure is entered.

4.2 Markov Chain Model of Packet Service

In this section, we present a discrete-time Markov chain model of packet service in an AC_i queue. In a fully saturated model, there is at least one packet ready to be transmitted in the queue at the time that a packet is transmitted successfully or is discarded. On the other hand, in the non-saturated traffic model, there is not necessarily always a packet waiting in the queue for transmission at that moment. Therefore, the bi-dimensional random process $\{s(i,t), b(i,t)\}$ described in the previous section is extended to a three dimensional discrete-time Markov chain $\{s(i,t), b(i,t), v(i,t)\}$, where the third

dimension $v(i,t)$ denotes whether there is a packet waiting in the queue of AC_i at the time of packet service completion. We let $v(i,t)$ denote a Bernoulli random variable which assumes the values of one and zero with probabilities ρ_i^* and $1-\rho_i^*$ respectively, where ρ_i^* is the probability that AC_i queue is not empty at the moment when a packet transmission is successful or it is discarded. We define ρ_i as the probability that an AC_i queue is busy at the completion of packet service time. Then, let us define the following two probabilities.

$$q_i = \text{Prob (at least one packet arrival during a generic slot)} \tag{4.4}$$

$$q_i^* = \text{Prob}\begin{pmatrix}\text{at least one packet arrival during the time interval between}\\ \text{the two slots that the backoff counter may be decremented}\end{pmatrix} \tag{4.5}$$

A transmitted packet encounters a collision when at least one other PCA station transmits during a PCA slot. We will let p_i denote the probability that a transmitted packet collides during a PCA slot and we assume that p_i is independent of the state $s(i,t)$of the station. Since a PCA station senses the channel busy when one or more PCA stations transmits during a PCA slot, p_i also equal to the probability that a PCA station in the backoff stage for AC_i senses the channel busy during a PCA slot. If a packet experiences a collision, the value of the backoff stage is increased by one, and a new backoff counter value is uniformly chosen. When the contention window size reaches its maximum value at the backoff stage m_i, the contention window size remains at its maximum value for $m_i \leq j \leq L_{i,retry}$, as shown in (4.3). When the number of trials reaches to $L_{i,retry}$ and the packet still ends up in a collision, the packet is discarded.

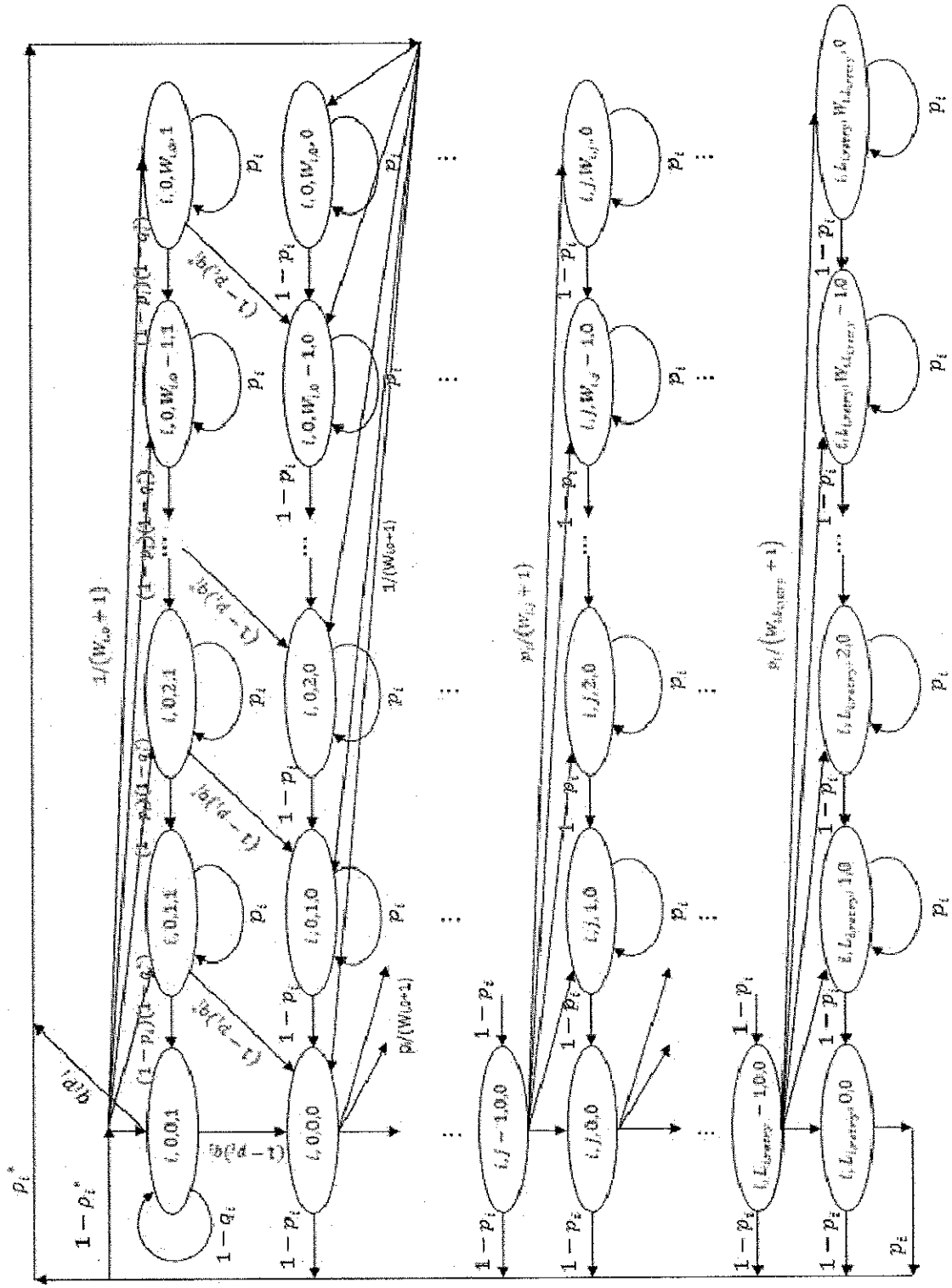


Figure 4.3 State transition diagram for AC_i

Figure 4.3 illustrates the state transition diagram of the Markov chain for AC_i. For mathematical simplification, we use a short notation $\{i, j, l, e\}$ to describe the state vector $\{s(i,t), b(i,t), v(i,t)\}$ for AC_i at time t. In $\{i, j, l, e\}$, i stands for access category (AC), j stands for the backoff stage ($j = 0,1, \ldots, L_{i,retry}$), l stands for the current value of the backoff counter ($l = 0,1, \ldots, W_{i,j}$) and e denotes the status of the AC_i queue as empty or not empty. Thus, in $\{i, j, l, e\}$, we have the following correspondence, $s(i,t) \rightarrow j$, $b(i,t) \rightarrow l$, $v(i,t) \rightarrow e$. We note that when the transmission of a packet terminates the post-backoff procedure begins in state $\{i, 0, l, 0\}$ if the queue is nonempty otherwise in state $\{i, 0, l, 1\}$. If the backoff procedure begins with an empty queue, then, if a new packet arrives in a state $\{i, 0, l, 1\}$, the system makes the transition to state $\{i, 0, l, -1\}$.

4.2.1 The Transition Probability Matrix of the Markov Chain

In this section, we determine the state transition probability matrix of the Markov chain shown in Figure 4.3. Let us define the following notation,

$$p\{i, j_1, l_1, e_1 | i, j_0, l_0, e_0\} = p\begin{Bmatrix} s(i,t+1) = j_1, b(i,t+1) = l_1, e(i,t+1) = e_1 | \\ s(i,t) = j_0, b(i,t) = l_0, v(i,t) = e_0 \end{Bmatrix} \quad (4.6)$$

Since the initial backoff counter value is uniformly sampled from $\{0,1, \ldots, W_{i,j}\}$ at the j'th stage, the probability distribution of the initial value is given by $1/(W_{i,j} + 1)$.

Then, we can express the non-zero transition probabilities of the Markov chain as follows,

1) The self-transition probability in a state that no packet arrives in the following generic slot when the queue of AC_i is empty is given by,

$$p\{i,0,0,1|i,0,0,1\} = 1 - q_i \quad , \quad j = 0 \quad , \quad l = 0 \tag{4.7}$$

2) The transition probability to a state following the arrival of a new packet to an empty AC_i queue after the end of the post-backoff procedure is given by,

$$p\{i,0,0,0|i,0,0,1\} = q_i(1-p_i) + \frac{q_i p_i}{W_{i,0}+1} \quad , \quad j = 0 \quad , \quad l = 0 \tag{4.8}$$

$$p\{i,0,l,0|i,0,0,1\} = \frac{q_i p_i}{W_{i,0}+1} \quad , \quad j = 0 \quad , \quad 1 \le l \le W_{i,0} \tag{4.9}$$

3) The self-transition probability in a state that the backoff counter is frozen due to the channel being sensed busy is given by,

$$p\{i,0,l,1|i,0,l,1\} = p_i \quad , \quad j = 0 \quad , \quad 1 \le l \le W_{i,0} \tag{4.10}$$

$$p\{i,j,l,0|i,j,l,0\} = p_i \quad , \quad 0 \le j \le L_{i,retry} \quad , \quad 1 \le l \le W_{i,j} \tag{4.11}$$

4) The transition probabilities when the counter will be decremented by one as the channel is sensed idle in a PCA slot are given by,

$$p\{i,0,l-1,1|i,0,l,1\} = (1-p_i)(1-q_l^*) \quad , \quad j = 0 \quad , \quad 1 \le l \le W_{i,0} \tag{4.12}$$

$$p\{i,0,l-1,0|i,0,l,1\} = (1-p_i)q_i^* \quad , \quad j=0 \quad , \quad 1 \le l \le W_{i,0} \quad (4.13)$$

$$p\{i,j,l-1,0|i,j,l,0\} = 1-p_i \quad , \quad 0 \le j \le L_{i,retry} \quad , \quad 1 \le l \le W_{i,j} \quad (4.14)$$

5) The transition probabilities that there is no packet waiting in the queue when a packet is transmitted successfully or it is dropped are given by,

$$p\{i,0,l,1|i,j,0,0\} = \frac{(1-\rho_i{}^*)(1-p_i)}{W_{i,0}+1} \quad , \quad 0 \le j \le L_{i,retry}-1 \quad , \quad 1 \le l \le W_{i,0} \quad (4.15)$$

$$p\{i,0,l,1|i,L_{i,retry},0,0\} = \frac{(1-\rho_i{}^*)}{W_{i,0}+1} \quad , \quad j = L_{i,retry} \quad , \quad 1 \le l \le W_{i,0} \quad (4.16)$$

6) On the other hand, if there is at least one packet waiting in the queue when a packet service is completed, the probability of transition to each state in backoff stage zero from backoff stage j is given by,

$$p\{i,0,l,0|i,j,0,0\} = \frac{\rho_i{}^*(1-p_i)}{W_{i,0}+1} \quad , \quad 0 \le j \le L_{i,retry}-1 \quad , \quad 1 \le l \le W_{i,0} \quad (4.17)$$

$$p\{i,0,l,0|i,L_{i,retry},0,0\} = \frac{\rho_i{}^*}{W_{i,0}+1} \quad , \quad j = L_{i,retry} \quad , \quad 1 \le l \le W_{i,0} \quad (4.18)$$

7) The value of the backoff stage is incremented by one, when a collision happens.

$$p\{i,j+1,l,0|i,j,0,0\} = \frac{p_i}{W_{i,j+1}+1} \quad , \quad 0 \le j \le L_{i,retry} \quad , \quad 1 \le l \le W_{i,j+1} \quad (4.19)$$

4.2.2 The Steady-State Probability Distribution of the Markov Chain

From the state transition probabilities we have determined in section 4.2.1, we will determine the steady-state probability distribution of the above Markov chain in this section. Let $b_{i,j,l,e}$ denote the steady-state distribution of the Markov chain, which is defined as,

$$b_{i,j,l,e} = \lim_{t\to\infty} p\{s(i,t) = j, b(i,t) = l, v(i,t) = e\} \quad ,$$

$$0 \le j \le L_{i,retry} \quad , \quad 0 \le l \le W_{i,j} \qquad (4.20)$$

Since the unsuccessful transmission probability is given by p_i, the probability of station being in the j'th backoff stage is given by,

$$b_{i,j,0,0} = {p_i}^j b_{i,0,0,0} \qquad (4.21)$$

When the backoff counter reaches zero for each backoff stage,

$$b_{i,j,0,0}$$

$$= \begin{cases} (1-p_i)b_{i,j,1,0} + q_i(1-p_i)b_{i,0,0,1} + q_i^*(1-p_i)b_{i,0,1,1} \\ \qquad + \dfrac{{\rho_i}^*}{W_{i,j}+1} b_{i,L_{i,retry},0,0} + \dfrac{{\rho_i}^*(1-p_i)}{W_{i,j}+1} \sum_{k=1}^{L_{i,retry}-1} b_{i,k,0,0} \quad , \quad j = 0 \\ \\ (1-p_i)b_{i,j,1,0} + \dfrac{1}{W_{i,j}+1} p_i b_{i,j-1,0,0} \quad , \quad 1 \le j \le L_{i,retry} \end{cases}$$

(4.22)

For the backoff counter value $1 \leq l \leq W_{i,j} - 1$,

$$b_{i,j,l,0}$$

$$= \begin{cases} (1-p_i)b_{i,j,l+1,0} + p_i b_{i,j,l,0} + q_i^*(1-p_i)b_{i,0,l+1,1} \\ \qquad + \dfrac{\rho_i{}^*}{W_{i,j}+1} b_{i,L_{i,retry},0,0} + \dfrac{\rho_i{}^*(1-p_i)}{W_{i,j}+1} \sum_{k=1}^{L_{i,retry}-1} b_{i,k,0,0} & , \quad j = 0 \\ (1-p_i)b_{i,j,l+1,0} + p_i b_{i,j,l,0} + \dfrac{1}{W_{i,j}+1} p_i b_{i,j-1,0,0} & , \quad 1 \leq j \leq L_{i,retry} \end{cases}$$

(4.23)

When the backoff counter value is $l = W_{i,j}$,

$$b_{i,j,l,0}$$

$$= \begin{cases} p_i b_{i,j,l,0} + q_i^*(1-p_i)b_{i,0,l+1,1} \\ \qquad + \dfrac{\rho_i{}^*}{W_{i,j}+1} b_{i,L_{i,retry},0,0} + \dfrac{\rho_i{}^*(1-p_i)}{W_{i,j}+1} \sum_{k=1}^{L_{i,retry}-1} b_{i,k,0,0} & , \quad j = 0 \\ p_i b_{i,j,l,0} + \dfrac{1}{W_{i,j}+1} p_i b_{i,j-1,0,0} & , \quad 1 \leq j \leq L_{i,retry} \end{cases}$$

(4.24)

From (4.21), we have

$$\sum_{j=1}^{L_{i,retry}-1} b_{i,j,0,0} = \sum_{j=1}^{L_{i,retry}-1} p_i{}^j b_{i,0,0,0} = \frac{b_{i,0,0,0} - b_{i,L_{i,retry},0,0}}{1-p_i} \tag{4.25}$$

Combining the equations (4.22) to (4.25), we have,

$$b_{i,j,l,0} = \frac{W_{i,j}+1-l}{W_{i,j}+1}\frac{1}{1-p_i}b_{i,j,0,0} \quad , \quad 1 \le j \le L_{i,retry} \quad , \quad 1 \le l \le W_{i,j} \quad (4.26)$$

$$b_{i,0,l,1} = \frac{(1-\rho_i{}^*)b_{i,0,0,0}}{(W_{i,0}+1)(1-p_i)}\frac{1-(1-q_i^*)^{W_{i,0}+1-l}}{q_i^*} \quad , \quad j=0 \quad , \quad 1 \le l \le W_{i,0} \quad (4.27)$$

$$b_{i,0,0,1} = \frac{(1-\rho_i{}^*)b_{i,0,0,0}}{(W_{i,0}+1)q_i}\frac{1-(1-q_i^*)^{W_{i,0}+1}}{q_i^*} \quad , \quad j=0 \quad , \quad l=0 \quad (4.28)$$

and

$$b_{i,0,l,0} = \frac{W_{i,j}+1-l}{W_{i,0}+1}\frac{1}{1-p_i}(b_{i,0,0,0}+q_ip_ib_{i,0,l,1}) - b_{i,0,l,1} \quad ,$$

$$j=0 \quad , \quad 1 \le l \le W_{i,0} \quad (4.29)$$

From the normalization condition, the sum of the probabilities of all the states in the Markov chain should be equal to one,

$$\sum_{j=0}^{L_{i,retry}}\sum_{l=0}^{W_{i,j}}\sum_{e=0}^{1} b_{i,j,l,e} = 1 \quad (4.30)$$

Thus, we can determine the probability of the initial state by substituting (4.26) to (4.29) into (4.30), which is given by,

$$\frac{1}{b_{i,0,0,0}} = \sum_{j=0}^{L_{i,retry}}\left[1+\frac{1}{1-p_i}\sum_{l=0}^{W_{i,j}}\frac{W_{i,j}+1-l}{W_{i,j}+1}\right]{p_i}^j$$

$$+\frac{1-\rho_i{}^*}{q_i}\frac{1-(1-q_i^*)^{W_{i,0}+1}}{(W_{i,0}+1)q_i^*}\left(1+\frac{W_{i,0}q_ip_i}{2(1-p_i)}\right) \quad (4.31)$$

4.2.3 Determination of the Parameters of the Markov Chain Model

4.2.3.1 Determination of the packet transmission probability

Let τ_i denote the probability that a station transmits a packet from AC_i queue in a generic slot. Once the value of backoff counter reaches zero at any backoff stage, the transmission begins, so we have

$$\tau_i = \sum_{j=0}^{L_{i,retry}} b_{i,j,0,0} = b_{i,0,0,0} \frac{1 - {p_i}^{L_{i,retry}+1}}{1 - p_i} \tag{4.32}$$

4.2.3.2 Determination of p_i

The channel is sensed busy in PCA period when at least one station transmits during a PCA slot. The probability that the channel is busy in a PCA period, denoted by p_{busy}, is given by

$$p_{busy} = 1 - \prod_{h=0}^{N-1} (1 - \tau_h)^{\eta} \tag{4.33}$$

Access categories will be different channel busy probabilities because of their different contention window sizes. An external observer will see the channel idle, if all the 4η queues are empty or in backoff process. An access category i, AC_i, will see the channel idle if the remaining $4\eta - 1$ queues are empty or in backoff process. Since ACs have

different contention window sizes, the probability that AC_i will see a busy channel, p_i, will be different, which is given by,

$$p_i = 1 - \prod_{h=0, h\neq i}^{N-1} (1-\tau_h)^{\eta} = 1 - \frac{1-p_{busy}}{1-\tau_i} \tag{4.34}$$

Let P_i^{loss} denote the probability that an AC_i packet will be discarded, which is given by,

$$P_i^{loss} = {p_i}^{L_{i,retry}+1} \tag{4.35}$$

4.2.3.3 Calculation of q_i and q_i^*

In Figure 4.3, q_i denote the probability that at least one packet arrives at the AC_i queue during a generic slot while the system is in the state $\{i, 0,0,1\}$, which corresponds to an empty queue at the beginning of the slot. The probability distribution function (PDF) of the duration of a generic slot denoted by $\tilde{g}(\zeta)$ for AC_i is given by,

$$\tilde{g}(\zeta) = p_s\delta\left(\zeta - T_{s(i)}\right) + \left(1-p_{busy}\right)\delta(\zeta - T_e) + \left(p_{busy} - p_s\right)\delta(\zeta - T_c) \tag{4.36}$$

Where $\delta(\,)$ is the impulse function and T_e, $T_{s(i)}$ and T_c denote the real-time duration of a PCA slot, the average time of a successful AC_i packet transmission, and the average time of a collision in PCA period, respectively. p_s is the probability that the transmission of a packet is successful, the calculation of which is provided later on in section 4.3.

Let $\bar{\tilde{g}}$ denote the expected duration of a generic slot for AC_i, then,

$$\bar{\tilde{g}} = p_s T_{s(i)} + (1 - p_{busy})T_e + (p_{busy} - p_s)T_c \tag{4.37}$$

Under the assumption that the arrival of the packets for AC_i is according to a Poisson distribution with parameter λ_i, the probability of ς packets arrivals at the AC_i queue during a time duration of $\tilde{t}$ seconds is given by,

$$P_i(\varsigma, \tilde{t}) = \frac{e^{-\lambda_i \tilde{t}}(\lambda_i \tilde{t})^{\varsigma}}{\varsigma!} \quad , \quad \varsigma = 0,1,2 \ldots \tag{4.38}$$

Then, $P_i(0, \tilde{t})$, the probability that no packet arrives at the AC_i queue during $\tilde{t}$ seconds, is given by,

$$P_i(0, \tilde{t}\,) = e^{-\lambda_i \tilde{t}} \tag{4.39}$$

Thus, the probability that at least one packet arrives at the queue of AC_i during a generic slot is given by,

$$q_i = 1 - \int_0^{\infty} P_i(0, \zeta)\tilde{g}(\zeta)\, d\zeta = 1 - \int_0^{\infty} e^{-\lambda_i \zeta}\tilde{g}(\zeta)\, d\zeta$$

$$q_i = 1 - \left(p_s e^{-\lambda_i T_{s(i)}} + (1 - p_{busy})e^{-\lambda_i T_e} + (p_{busy} - p_s)e^{-\lambda_i T_c}\right) \tag{4.40}$$

As we defined in (4.5), q_i^* denotes the probability that at least one packet arrives during the time that it takes to decrement the counter by one, which is given below,

$$q_i^* = 1 - V_i(z)|_{z=e^{-\lambda_i \sigma}} \tag{4.41}$$

where $V_i(z)$ is the PGF of the time to decrement the counter by one in number of virtual slots of duration σ determined in (4.64). Hence,

$$q_i^* = 1 - \frac{e^{-\lambda_i T_e}(1 - p_i)}{1 - p_i(\frac{p_s}{p_{busy}} e^{-\lambda_i T_{s(i)}} + \frac{p_{busy} - p_s}{p_{busy}} e^{-\lambda_i T_c})} \quad (4.42)$$

4.2.3.4 Calculation of ρ_i and ρ_i^*

Let $\bar{M}_i$ denote the mean packet service time for an AC_i queue. In a single server queue with Poisson arrivals, the utilization of an AC_i queue, ρ_i, is given by $\lambda_i \bar{M}_i$, where λ_i is the arrival rate. In this case, the packet service time is equal to the duration from the time a packet reaches to head of its queue to the end of post-backoff procedure following the packet's successful transmission or discarding. For stability, ρ_i cannot greater than one, so we have,

$$\rho_i = \min(1, \lambda_i \bar{M}_i) \quad (4.43)$$

where, $\bar{M}_i$ is determined in Section 4.5.

ρ_i^* denotes the probability that the queue is nonempty at the time when a packet is transmitted successfully or discarded. The relationship of ρ_i^* and ρ_i is shown in Figure 4.4, which is given by,

$$1 - \rho_i = P_i^B (1 - \rho_i^*) \quad (4.44)$$

where P_i^B is the probability that there is no packet arrival during the post-backoff procedure, which is given by,

$$P_i^B = \frac{1-(1-q_i{}^*)^{(W_{i,0}+1)}}{(W_{i,0}+1)q_i{}^*} \tag{4.45}$$

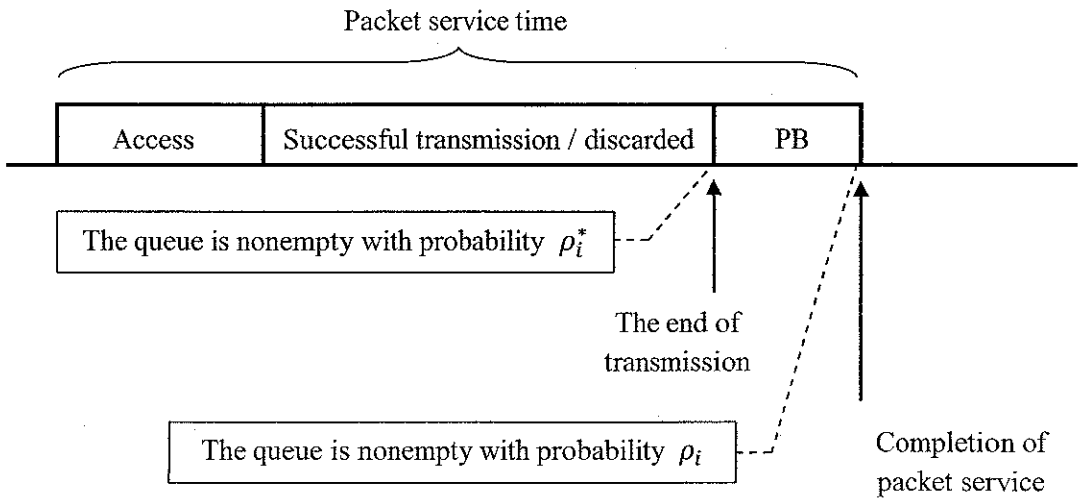


Figure 4.4 The relationship between $\rho_i{}^*$ and ρ_i

(PB denotes the post-backoff procedure)

4.3 Throughput for PCA Traffic

Let $p_{s,i}$ denote the probability that a packet is transmitted successfully from any AC_i queue in a generic slot. Since there are η stations and each station contains queues for all ACs, we have,

$$p_{s,i} = \eta\tau_i(1 - p_i) \tag{4.46}$$

Then, the probability that a generic slot carries a successful packet transmission is given by,

$$p_s = \sum_{i=0}^{N-1} p_{s,i} \tag{4.47}$$

Let S_i denote the throughput of PCA for AC_i. The expression of S_i from [22] is given by,

$$S_i = \frac{E(\text{duration of a generic slot carrying successful } AC_i \text{ transmission})}{E(\text{duration of a generic slot})}$$

$$S_i = \frac{p_{s,i}T_{E(L)}}{\bar{\bar{g}}} = \frac{p_{s,i}T_{E(L)}}{(1 - p_{busy})T_e + p_sT_{s(i)} + (p_{busy} - p_s)T_c} \tag{4.48}$$

where, T_e, L, $T_{E(L)}$, $T_{s(i)}$ and T_c denote the real-time duration of an empty PCA slot, the length of the payload, the time to transmit payload with average length $E(L)$, the average time of a successful PCA transmission, and the average time of a collision in PCA period, respectively. In the above, we note that the denominator is given by (4.37).

Let T_H, T_{SIFS}, T_{ACK}, T_{RTS}, T_{CTS}, and $T_{AIFS(i)}$ denote the time to transmit the header (including PLCP preamble and PLCP header consisting of MAC header, physical layer header, header check sequence (HCS),Reed-Solomon parity bits and tail bits), the SIFS time, the time to transmit an acknowledgement, the time to transmit a Request-To-Send (RTS) frame, the time to transmit a Clear-To-Send (CTS) frame and an Arbitration Inter-Frame Space (AIFS) time for AC_i.

For the basic access mode, $T_{s(i)}$ and T_c are respectively given by

$$T_{S(i)} = T_H + T_{E(L)} + T_{SIFS} + T_{ACK} + T_{AIFS(i)} \quad , i = 0,1,\dots,N-1 \tag{4.49}$$

$$T_C = T_H + T_{E(L)} + T_{SIFS} + T_{ACK} + \max\left(T_{AIFS(i)}\right) \quad , i = 0,1,\dots,N-1 \tag{4.50}$$

For the RTS/CTS access mode, $T_{S(i)}$ and T_c are respectively given by

$$T_{S(i)} = T_{RTS} + T_{CTS} + T_H + T_{E(L)} + 3T_{SIFS} + T_{ACK} + T_{AIFS(i)}$$

$$, i = 0,1,\dots,N-1 \tag{4.51}$$

$$T_C = T_{RTS} + T_{SIFS} + T_{CTS} + \max\left(T_{AIFS(i)}\right)$$

$$, i = 0,1,\dots,N-1 \tag{4.52}$$

Let us assume that there are Δ transmission rates available to a station, R_k, $k = 1 \dots \Delta$ and let β_k denote the probability that the transmission rate in a station is R_k. Assuming that each station is equally likely to have any transmission rate, then, we have $\beta_k = 1/\Delta$. Since the arrival rate of the packets to all stations is the same, the probability that a packet will be transmitted at the rate R_k is given by β_k. Given that there are N_{PCA} channels available for PCA traffic, the transmission rate of the single equivalent channel will be R_k (N_{PCA} / N_{SF}). Let $\bar{N}_{PCA}$ denote the average number of channels available for PCA traffic, we have,

$$\bar{N}_{PCA} = N_{SF} - \bar{N}_{DRP} \tag{4.53}$$

where $\bar{q}$ is the average number of busy channels for DRP in a superframe obtained in (3.31). Thus the expected transmission time of a packet is given by,

$$T_{E(L)} = \frac{1}{\Delta}\sum_{k=1}^{\Delta}\frac{E(L)}{R_k \frac{\bar{N}_{PCA}}{N_{SF}}} \tag{4.54}$$

Then, we can obtain the throughput for AC_i traffic by substituting (4.54) into (4.48).

4.4 Mean packet Delay

Next, we are ready to determine the mean total packet delay in terms of mean packet service time and the mean queuing delay. First, we will derive the probability generating function (PGF) of the packet service time and the packet queuing delay. Then, we will determine the mean packet service time and mean queuing delay. Finally, we will determine the mean total packet delay. In the following, we will apply a discrete time analysis. As a result, we will assume that time is divided into virtual slots, which has duration of σ seconds. The PCA slot, packet successful transmission time and collision duration will be assumed to be integer multiples of virtual time slots given below,

$$T_e{}' = \frac{T_e}{\sigma} \tag{4.55}$$

$$T_{s(i)}{}' = \frac{T_{s(i)}}{\sigma} \tag{4.56}$$

$$T_c{}' = \frac{T_c}{\sigma} \tag{4.57}$$

4.4.1 The Probability Generating Function of the Packet Service Time in Number of Virtual Slots

First, let us determine the probability generating function (PGF) of the packet service time. The packet service time has been defined as the duration from the time a packet arrives at the head of its queue to the end of following post-backoff procedure, as shown in Figure 4.4.

When the transmission of a packet collides with another packet during the transmission, the backoff stage will be incremented by one and a new backoff counter value will be chosen for retransmission. Each time the channel is sensed idle during a PCA slot, the counter value will be decremented by one. On the contrary, if the channel is sensed busy, the backoff counter will be frozen. Next, we will determine the various components of packet service time.

- **Duration of the time to decrement the counter by one**

Let a_i denote the number of generic slots that it takes to decrement the counter by one, which follows a geometric distribution. Since the probabilities that the channel is sensed busy and idle during a time slot are given by p_i and $1 - p_i$, respectively, we have,

$$\Pr(a_i = k \text{ generic slots}) = (1 - p_i)p_i{}^k \quad , \quad k = 0, 1, 2, \ldots \tag{4.58}$$

Then, the PGF of a_i is given by,

$$A_i(z) = \frac{1 - p_i}{1 - p_i z} \tag{4.59}$$

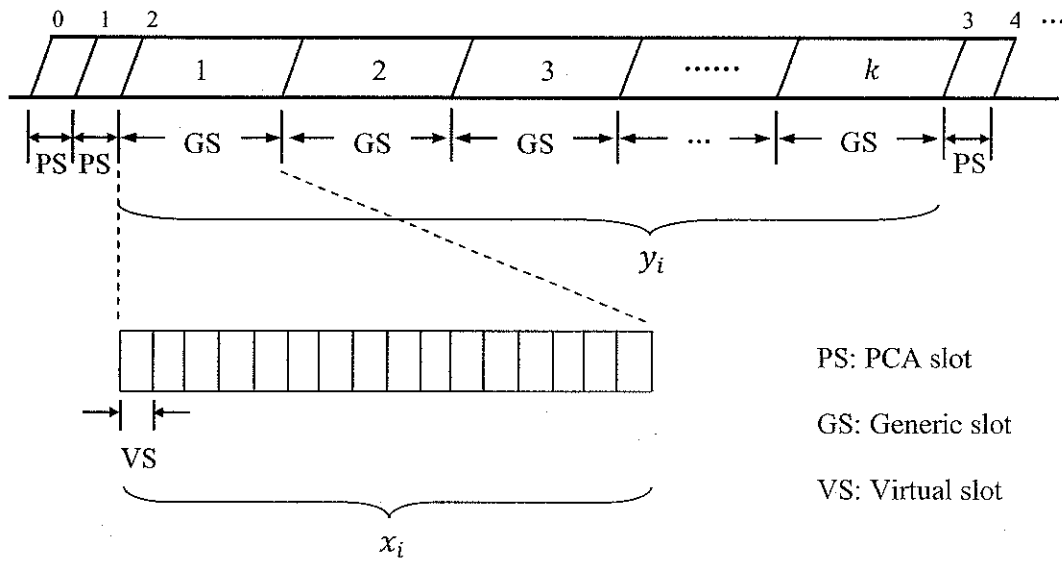


Figure 4.5 The relationship between generic slot and virtual slot

We note that during the k generic slots that the counter is frozen and each slot carries either a successful transmission or a collision. Let x_i denote the duration of the time in number of virtual slots that a counter remains frozen during a generic slot, as shown in Figure 4.5. Then, the expected value of x_i is given by,

$$E(x_i) = \frac{p_s}{p_{busy}} T_{s(i)}' + \frac{p_{busy} - p_s}{p_{busy}} T_c' \tag{4.60}$$

Let $X_i(z)$ denote the PGF of x_i . Then we have,

$$X_i(z) = E(z^{x_i}) = \frac{p_s}{p_{busy}} z^{T_{s(i)}'} + \frac{p_{busy} - p_s}{p_{busy}} z^{T_c'} \tag{4.61}$$

Let y_i denote the amount of time in number of virtual slots that it takes to decrement counter by one. We have,

$$y_i = \sum_{k=0}^{a_i} x_{i,k} \tag{4.62}$$

All $x_{i,k}$ are independent identically distributed with PGF $X_i(z)$, so the PGF of y_i is given by,

$$Y_i(z) = A_i(z)|_{z=X_i(z)} = \frac{1-p_i}{1-p_i X_i(z)} \tag{4.63}$$

Following the unfreezing of the counter, the station will spend an empty PCA slot before it begins to decrement the counter. Then, the PGF of the total amount of time it takes to decrement the counter by one, is given by,

$$V_i(z) = z^{T_e{}'} Y_i(z) \tag{4.64}$$

- **Duration of the time that an AC spends in a backoff stage**

Let $r_{i,j}$ denote the chosen value of the backoff counter in the j'th stage for AC_i, then $r_{i,j}$ has a uniform probability distribution given by,

$$\Pr(r_{i,j}) = \frac{1}{W_{i,j}+1} \quad , \ 0 \le r_{i,j} \le W_{i,j} \tag{4.65}$$

The PGF of $r_{i,j}$ denoted by $R_{i,j}(z)$ is given by,

$$R_{i,j}(z) = \sum_{r_{i,j}=0}^{W_{i,j}} z^{r_{i,j}} \Pr(r_{i,j}) = \frac{1}{W_{i,j}+1} \sum_{r_{i,j}=0}^{W_{i,j}} z^{r_{i,j}} = \frac{1}{W_{i,j}+1} \frac{1-z^{W_{i,j}+1}}{1-z} \quad (4.66)$$

Let us define $g_{i,j}$ as the number of virtual slots that AC_i spends in the j'th backoff stage, which is given by,

$$g_{i,j} = \sum_{k=0}^{r_{i,j}} v_{i,k} \quad (4.67)$$

All $v_{i,k}$ are independent identically distributed with PGF $V_i(z)$. Thus the total duration of the count time in the j'th backoff stage denoted by $G_{i,j}(z)$ can be expressed as,

$$G_{i,j}(z) = R_{i,j}(z)\big|_{z=V_i(z)} = \frac{1}{W_{i,j}+1} \frac{1-V_i(z)^{W_{i,j}+1}}{1-V_i(z)} \quad (4.68)$$

For simplicity we will introduce $B_{i,j,s}(z)$ as:

$$B_{i,j,s}(z) = \prod_{k=s}^{j} G_{i,k}(z) \quad (4.69)$$

Next, we determine the PGF of the packet service time. The service time of the next packet begins if AC_i queue is nonempty at the end of the current post-backoff procedure and the packet will be transmitted immediately. On the other hand, if the queue is empty, the service time of the packet begins as soon as the packet arrives. If the channel is sensed idle, the packet will be transmitted immediately, otherwise, a backoff procedure will be initiated by uniformly choosing a value from $\{0, 1, \ldots, W_{i,0}\}$. From

(4.38), the probability that no packet arrives at the AC_i queue during a virtual slot is given by,

$$P_i(0,\sigma) = e^{-\lambda_i \sigma}$$

Let P_i^{EM} denote the probability of no arrivals during the post-backoff procedure. Since the PGF of the duration of the post-backoff duration, $G_{i,0}(z)$, has been determined in (4.67), we have,

$$P_i^{EM} = G_{i,0}(z)\Big|_{z=e^{-\lambda_i \sigma}}$$

For a packet which has been transmitted successfully, the packet service time contains the backoff delay of all stages that packet went through, the collision delay T_c for each stage it passed through and a transmission delay $T_{s(i)}$ for the last stage. For a packet which is discarded, the packet service time will contain the delay of all the stages from 0 to $L_{i,retry}$. Let $M_i(z)$ denote the PGF of packet service time, then, it is given by,

$$M_i(z) = (1-\rho_i^*)P_i^{EM} p_i G_{i,0}(z) + (1-p_i)\sum_{j=0}^{L_{i,retry}} {p_i}^j z^{{T_{s(i)}}' + j{T_c}'} B_{i,j,0}(z)$$

$$+ {p_i}^{L_{i,retry}+1} z^{(L_{i,retry}+1){T_c}'} B_{i,L_{i,retry},0}(z) \qquad (4.70)$$

In the above, the first term corresponds to the additional backoff stage of a packet arriving at an empty queue and busy channel. The second and third terms correspond to the service time of a successfully transmitted and discarded packet, respectively.

4.4.2 Mean Packet Service Time

Let us denote the mean packet service time of AC_i as $\bar{M}_i$. Taking the first order derivative of (4.70) and substituting $z = 1$, we have,

$$\bar{M}_i = (M_i(z))'|_{z=1}$$

$$= \frac{W_{i,0}}{2}(1-\rho_i^*)P_i^B p_i \bar{v}_i + (1-{p_i}^{L_{i,retry}})\left(T_{s(i)}{}' + T_c{}'\frac{p_i}{1-p_i}\right) + \bar{v}_i \sum_{j=0}^{L_{i,retry}} {p_i}^j \frac{W_{i,j}}{2} \tag{4.71}$$

where $\bar{v}_i$ is the mean value of the time to decrement the counter by one, which can be found directly from the transform in (4.64).

$$\bar{v}_i = \left(V_i(z)\right)'\Big|_{z=1} = T_e{}' + \left(\frac{p_s}{p_{busy}}T_{s(i)}{}' + \frac{p_{busy}-p_s}{p_{busy}}T_c{}'\right)\frac{p_i}{1-p_i} \tag{4.72}$$

4.4.3 Mean Queuing Delay of a Packet

In the previous section, we have determined the mean packet service time, $\bar{M}_i$ given in (4.71). Assuming that the packet service times are independent of each other, we can determine the queuing delay by applying M/G/1 model in each queue. Then, let $\bar{Q}_i$

denote the average queuing delay for AC_i queue. From [32], the average queuing delay in a M/G/1 queue is given by,

$$\bar{Q}_i = \frac{\lambda_i \bar{M}_i^2}{2(1-\rho_i)} \tag{4.73}$$

where $\bar{M}_i^2$ is the second moment of the packet service time.

The second moment of the packet service time is determined by successive differentiation and manipulation of $M_i(z)$ (4.70), which is given by,

$$\bar{M}_i^2 = \bar{M}_i + (M_i(z))''|_{z=1}$$

$$= \frac{W_{i,0}}{2}(1-\rho_i^*)P_i^B p_i \bar{v}_i^2 + (1-p_i)\sum_{j=0}^{L_{i,retry}} \left(T_{s(i)}' + jT_c'\right)^2 {p_i}^j + 2(1-p_i)$$

$$+ \sum_{j=0}^{L_{i,retry}} \left(T_{s(i)}' + jT_c'\right){p_i}^j \bar{B}_{i,j,0} + (1-p_i)\sum_{j=0}^{L_{i,retry}} {p_i}^j \bar{B}_{i,j,0}^2$$

$$+ \left(\left(L_{i,retry}+1\right)T_c'\right)^2 {p_i}^{L_{i,retry}+1} + 2\left(L_{i,retry}+1\right)T_c' {p_i}^{L_{i,retry}+1}\bar{B}_{i,j,0}$$

$$+ {p_i}^{L_{i,retry}+1}\bar{B}_{i,j,0}^2 \tag{4.74}$$

where

$$\bar{B}_{i,j,0} = \left(B_{i,j,0}(z)\right)'\Big|_{z=1} = \bar{v}_i \sum_{s=0}^{j} \frac{W_{i,s}}{2} \tag{4.75}$$

and

$$\bar{B}_{i,j,0}^2 = \left(B_{i,j,0}(z)\right)''\Big|_{z=1} + \left(B_{i,j,0}(z)\right)'\Big|_{z=1}$$

$$= (\bar{v}_i)^2\left[\left(\sum_{s=0}^{j}\frac{W_{i,s}}{2}\right)^2 - \sum_{s=0}^{j}\left(\frac{W_{i,s}}{2}\right)^2 + \sum_{s=0}^{j}\frac{W_{i,s}(W_{i,s}-1)}{3}\right] + \bar{v}_i^2\sum_{s=0}^{j}\frac{W_{i,s}}{2} \qquad (4.76)$$

The first moment of countdown delay of a state, $\bar{v}_i$, is given by (4.72) and its second moment, $\bar{v}_i^2$, is given by,

$$\bar{v}_i^2 = \left(V_i(z)\right)''\Big|_{z=1} + \left(V_i(z)\right)'\Big|_{z=1}$$

$$= {T_e}^2 + \left[\frac{p_s}{p_{busy}}{T_{s(i)}}'\left(2{T_e}' + {T_{s(i)}}'\right) + \frac{p_{busy}-p_s}{p_{busy}}{T_c}'\left(2{T_e}' + {T_{s(i)}}'\right)\right]\frac{p_i}{1-p_i}$$

$$+2\left[\frac{p_s}{p_{busy}}{T_{s(i)}}' + \frac{p_{busy}-p_s}{p_{busy}}{T_c}'\right]^2\left(\frac{p_i}{1-p_i}\right)^2 \qquad (4.77)$$

Then, we can obtain the second moment of the packet service time, $\bar{M}_i^2$, by substituting (4.75) to (4.77) into (4.74), which is given by,

$$\bar{M}_i^2 = \frac{W_{i,0}}{2}(1-\rho_i^*)P_i^B p_i \bar{v}_i^2 + (1-{p_i}^{L_{i,retry}+1})\left({T_{s(i)}}'^2 + {T_c}'^2\frac{p_i}{1-p_i}\right)$$

$$+2{T_c}'\left(1-\left(L_{i,retry}+1\right){p_i}^{L_{i,retry}} + L_{i,retry}{p_i}^{L_{i,retry}+1}\right)\left({T_{s(i)}}' + {T_c}'\frac{p_i}{1-p_i}\right)\frac{p_i}{1-p_i}$$

$$+\bar{v}_i\left[\left({T_{s(i)}}' + {T_c}'\frac{p_i}{1-p_i}\right)(\omega 1_i - {p_i}^{L_{i,retry}+1}\omega 2_i) + {T_c}'\omega 3_i\right]$$

$$+(\bar{v}_i)^2\left(\frac{\omega 4_i}{3} + \frac{\omega 5_i - \omega 3_i}{2}\right) + \bar{v}_i^2\frac{\omega 1_i}{2} \qquad (4.78)$$

where $\omega 1_i \ldots \omega 5_i$ are defined as,

$$\omega 1_i = \sum_{j=0}^{L_{i,retry}} p_i{}^j \, W_{i,j} \tag{4.79}$$

$$\omega 2_i = \sum_{j=0}^{L_{i,retry}} W_{i,j} \tag{4.80}$$

$$\omega 3_i = \sum_{j=1}^{L_{i,retry}} j p_i{}^j \, W_{i,j} \tag{4.81}$$

$$\omega 4_i = \sum_{j=0}^{L_{i,retry}} p_i{}^j \, W_{i,j}\left(W_{i,j} - 1\right) \tag{4.82}$$

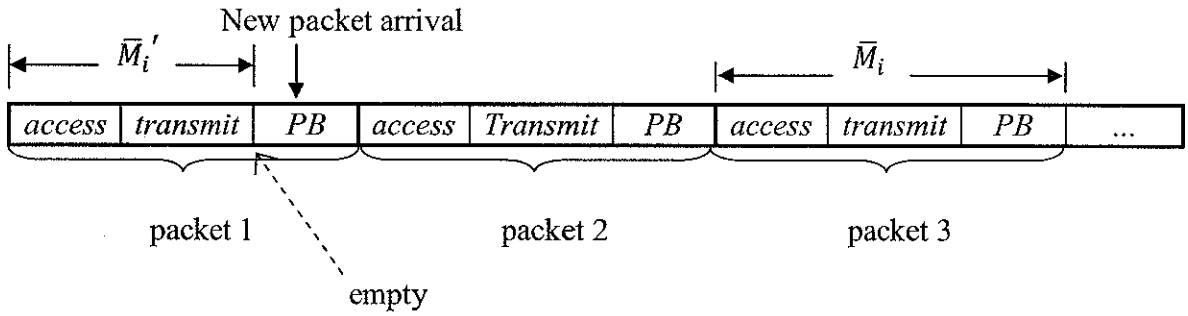
$$\omega 5_i = \sum_{j=1}^{L_{i,retry}} p_i{}^j \, W_{i,j} \sum_{s=0}^{j-1} W_{i,s} \tag{4.83}$$

Finally, we can determine the mean queuing delay by substituting (4.78) into (4.73).

4.4.4 Mean Total Delay of a Packet

The mean total delay of a packet is given by the sum of its mean packet service time and mean queuing delay. Since the delay of a packet is completed following its successful transmission or discarding, we should exclude the post-backoff procedure from the packet delay. Let $\bar{M}_i{}'$ denote the mean packet service time excluding the post-backoff procedure as shown in Figure 4.6.

Case I:



Case II:

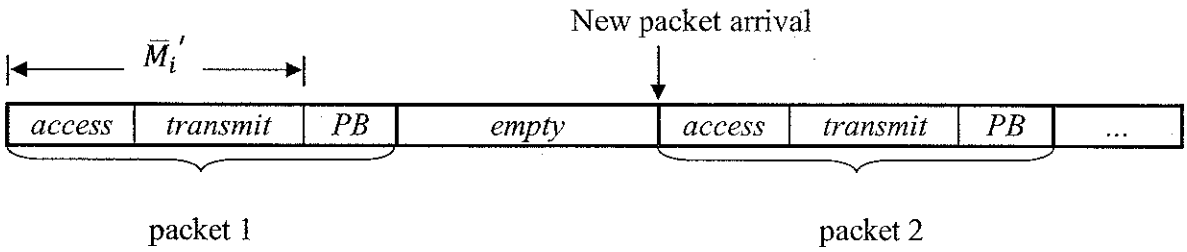


Figure 4.6 Packet service time excluding the post-backoff procedure

(PB denotes the post-backoff procedure)

Let $M_i'(z)$ denote the PGF of the packet service time excluding the post-backoff procedure, then,

$$M_i'(z) = M_i(z) - G_{i,0}(z) \tag{4.84}$$

Then, taking the first order derivatives of (4.84) and substituting $z = 1$, the mean packet service time without post-backoff duration is given by,

$$\bar{M}_i{}' = \left(M_i{}'(z)\right)'\Big|_{z=1}$$

$$= \bar{M}_i - \bar{v}_i \frac{W_{i,0}}{2} \tag{4.85}$$

Let $\bar{D}_i$ denote the mean total delay of a packet to be transmitted. Then, it is given by,

$$\bar{D}_i = \bar{M}_i{}' + \bar{Q}_i \tag{4.86}$$

4.5 Numerical Results

In this section, we present some numerical results regarding the analysis in this chapter. We assume that there are five PCA stations in the system and each station has N =4 ACs. The values of WiMedia MAC sublayer parameters and the values of the PHY dependent parameters used by the MAC sublayer used in the PCA scheme are tabulated in Tables 4.1 and 4.2, respectively. For simplicity, we chose virtual slot duration to be equal to the PCA slot. The duration for a packet transmission and collision, $T_{s(i)}$ and T_c, are assumed to be much greater than the duration of a PCA slot.

	AC_0	AC_1	AC_2	AC_3
$L_{i,retry}$	7	7	7	7
$CW_{i,min}$	15	15	7	3
$CW_{i,max}$	1023	1023	511	255

Table 4.1 MAC sublayer parameters in PCA scheme

Parameter	Value
T_{MAS}	256 μs
T_{sync}	9.375 μs
T_{hdr}	3.75 μs
T_{RTS}	14.45 μs
T_{CTS}	14.45 μs
T_{ACK}	15.5μs
T_{SIFS}	10 μs
$T_{AIFS(i)}$	46 μs
T_e	9 μs
σ	9 μs
DRP packet length (Γ)	1 Mb
PCA packet length (L)	1 Mb
DRP class (K)	8
PCA class (N)	4
N_{SF}	256

Table 4.2 PHY-dependent MAC sublayer parameters in PCA scheme

From (4.54), we determined the mean transmission time of a packet for different call arrival rates and different value of the minimum number of channels reserved for PCA traffic. The results are tabulated in Tables 4.3.

α	N_0	$\bar{q}$	$\overline{N}_{PCA}$	$T_{E(L)}$
40 calls/sec	4	48	208	9200 μs
100 calls/sec	4	119	137	13900 μs
100 calls/sec	50	97	159	12000 μs

Table 4.3 Mean transmission time of a packet

4.5.1 Results of ρ_i, p_i and P_i^{loss}

- **Utilization of an AC_i queue**

First, we present the numerical results of utilization of an AC_i queue for PCA traffic under the assumption that the total call arrival rate of DRP traffic has a constant value, α =100 calls/sec and $N_o = 4$. In Table 4.3, we have determined that the $\overline{N}_{PCA}$ in this case is equal to 137.

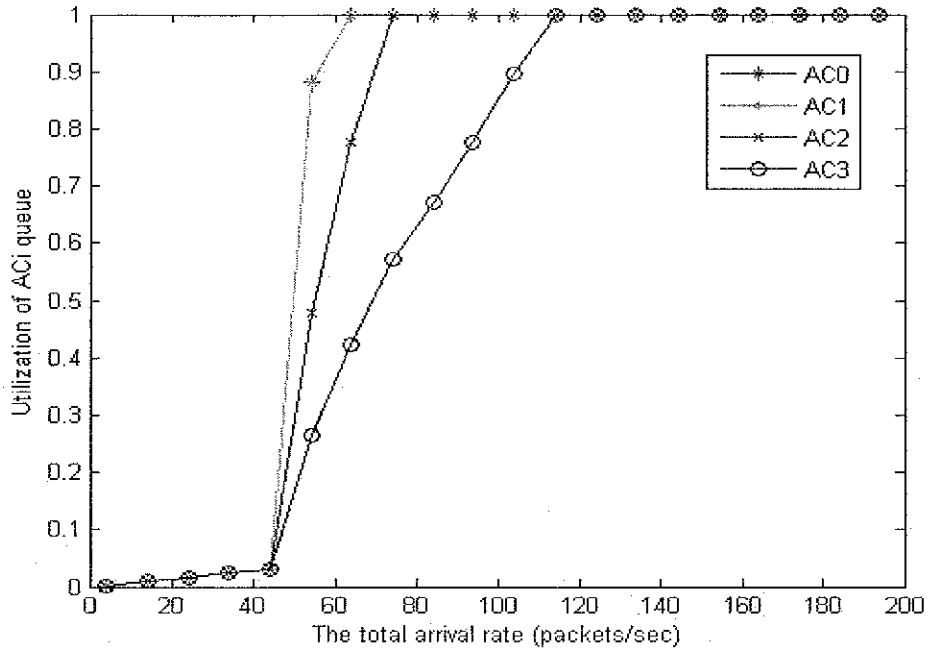


Figure 4.7 ρ_i versus the total arrival rate

From Figure 4.7, it may be seen that the utilization for all ACs increases with the total packet arrival rate and the utilization of the lower priority levels is greater than that of the higher priority levels.

- **The Probability that a transmitted packet encounters a collision**

Next, we present the probability that a transmitted packet encounters a collision, p_i , for each AC in the same PCA system.

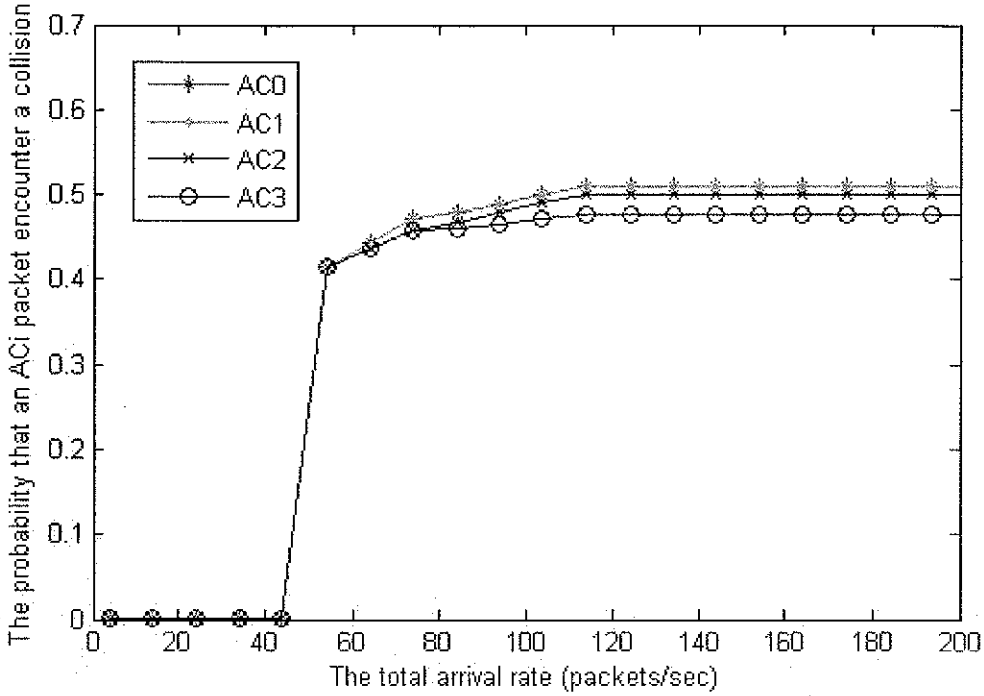


Figure 4.8 p_i versus the total arrival rate

From Figure 4.8, we observe that p_i is very small when the total packet arrival rate is low. On the other hand, when the total packet arrival rate is over about 60 packets/sec, p_i reaches approximately 50% soon and tends to be constant when queue is saturated.

- **The Probability that a Packet from an AC_i Queue will be Discarded**

Then, we plot the result of probability that a packet from an AC_i queue will be discarded, P_i^{loss}, for each AC in the same system.

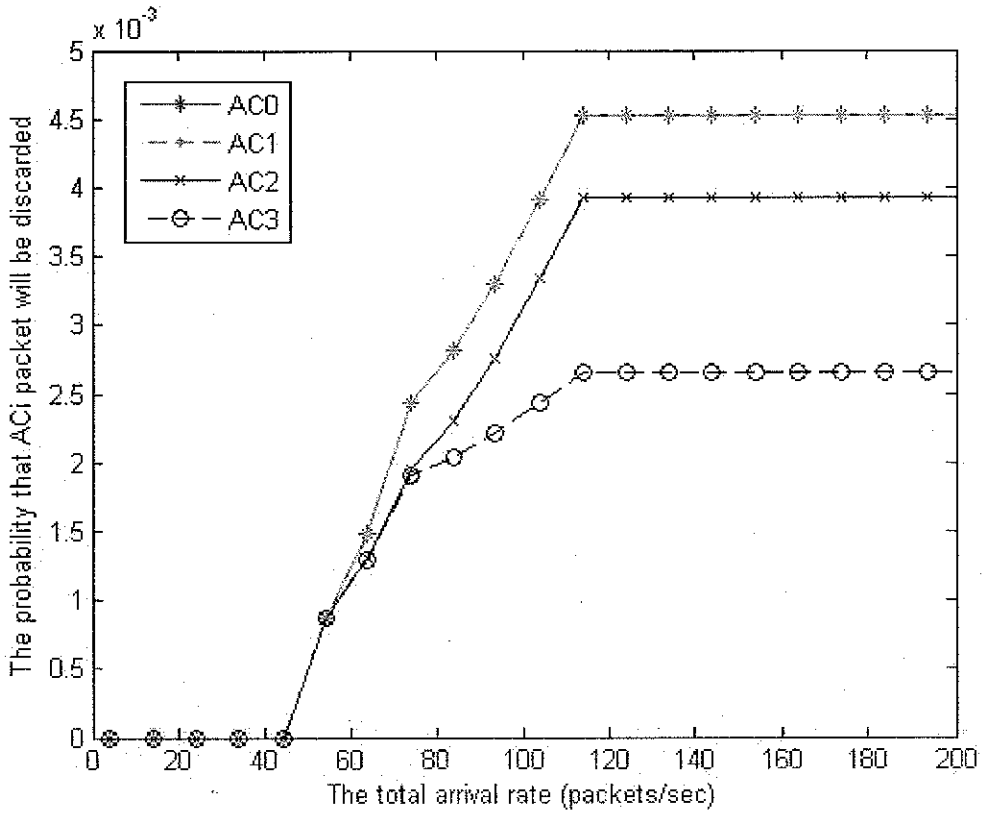


Figure 4.9 P_i^{loss} versus the total arrival rate

From Figure 4.9, it may be seen that the probabilities that an AC_i packet will be discarded are almost equal to zero when the traffic is light. They increase with the rise of the total packet arrival rate. The packets of the lower priority levels have the higher packet discarding probabilities.

4.5.2 Throughput for PCA Traffic

Next, we present the results of the throughput of each access category of PCA as a function of the total packet arrival rate for the same PCA system.

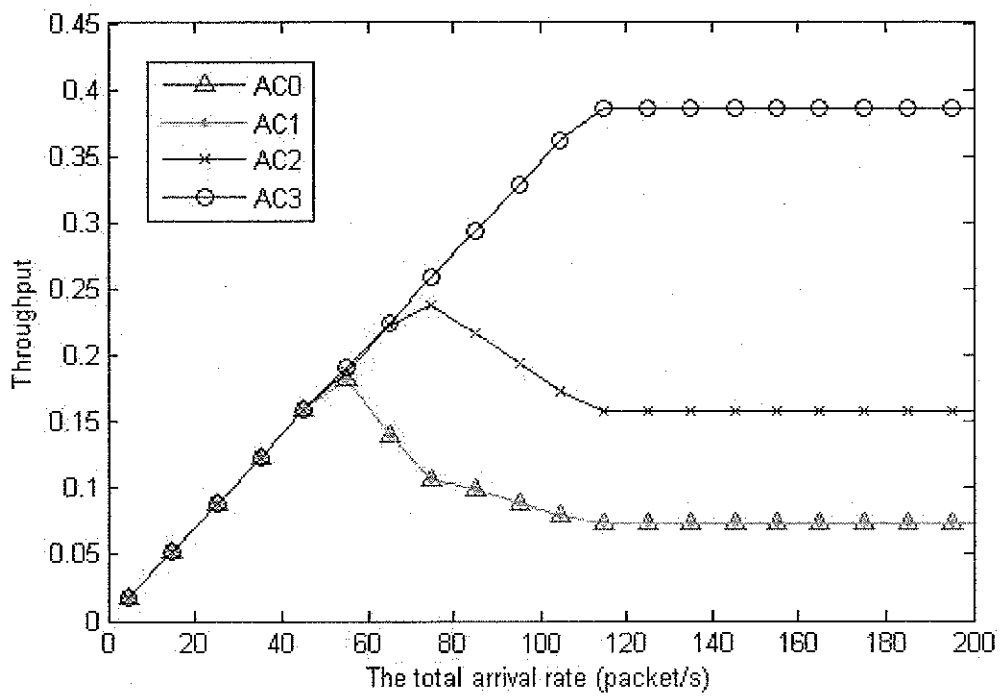


Figure 4.10 Throughput with different arrival rates

From Figure 4.10, we observe that the throughput for all ACs increase linearly with the increasing total packet arrival rate until the system reaches to saturation. We note that the saturation occurs when the total arrival rate is about 60 packets/sec and the throughput per AC_i is approximately 20%. Since all ACs have equal throughputs at this

point, the total throughput of PCA traffic is 80% and the remainder of the channel is taken by collisions and counter counting. From there on, the throughput of the lower priority levels decreases with the total packet arrival rate. We note that the curves of AC_0 and AC_1 are the same, because from Table 4.1, $CW_{i,min}$ and $CW_{i,max}$ for the first two categories are equal.

4.5.3 Mean Service Time of a Packet

Next, we calculate the mean packet service time as a function of the total packet arrival rate by applying different calls arrival rates for DRP traffic and different N_0 and compare the results severally.

- **Comparison of the Mean Packet Service Time for Different Values of N_o**

  We compare the results for two cases:

  Case I. $N_o = 4$

  Case II. $N_o = 50$

In Table 4.3, we have presented the number of MASs available for PCA for these two cases, which are $\overline{N}_{PCA} = 137$ for Case I and $\overline{N}_{PCA} = 159$ for Case II, respectively.

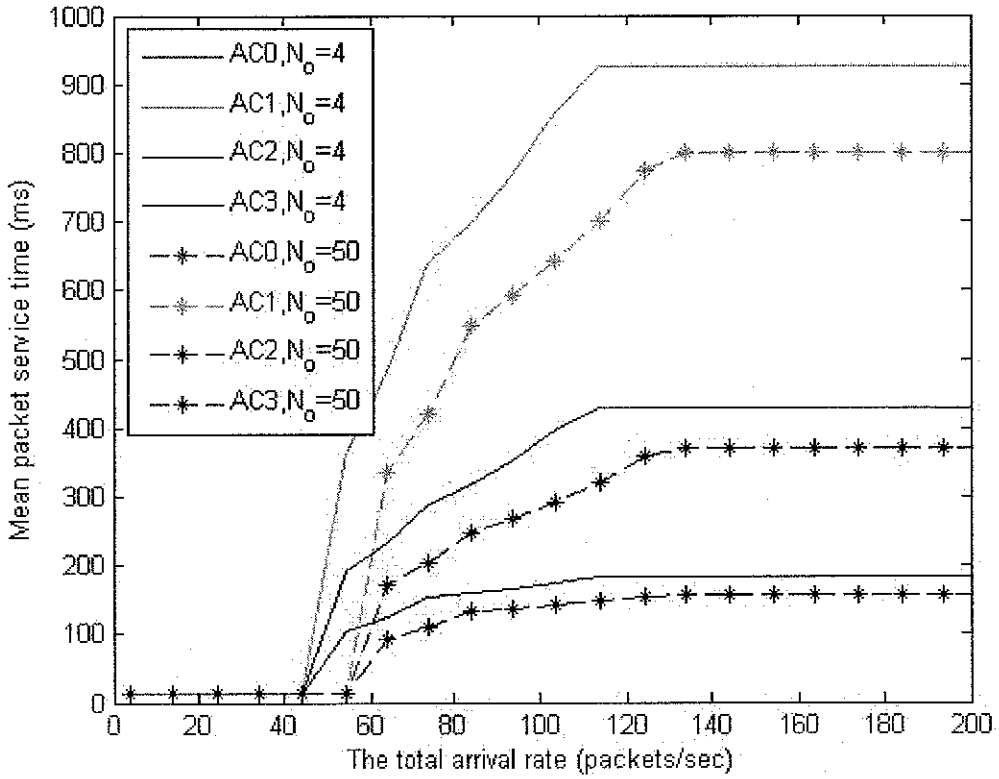


Figure 4.11 Mean packet service time comparison with different value of N_o

Figure 4.11 shows the results of the two cases. As may be seen, for both cases, all ACs have similar packet service time when the total arrival rate is small. It is because that each AC is assumed to have the same arrival rate and when packet arrival rate is low, the queues are not saturated. The packet service times increase with the increase of total arrival rate and the packet service time of the lower priority levels is greater than that of the higher priority levels. The reason is that the probability of packet collision increases as the traffic load gets heavier and the packets of higher priority levels can access to the channel in shorter times.

For Case I, the mean packet service time increases significantly when the total arrival rate reaches 45 packets/second and it tends to be constant when the AC_i queues get saturated. As expected, for Case II, the mean packet service time is always smaller than that of Case I.

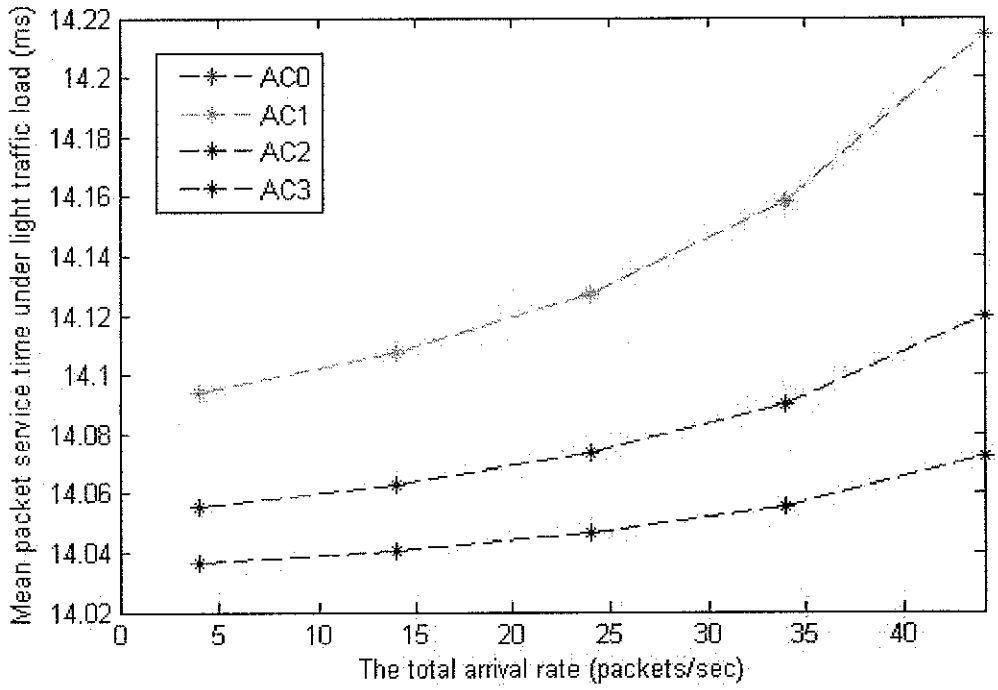


Figure 4.12 The mean packet service under the light traffic load

Taking Case I as an example, we plot the results of mean packet service time under light traffic load. From Figure 4.12, it may be clearer that the mean packet service time of the lower priority levels is higher than that of the higher priority levels.

- **Comparison of the Mean Packet Service Time for Different Call Arrival Rates for DRP Traffic**

From Chapter 3, we have known that the number of MASs available for PCA traffic varies, depending on the number of MASs reserved for DRP transmission. In section 3.6.2, we determined the average number of MASs occupied by DRP traffic as a function of the total call arrival rate of the calls to the system, α. Next, we present the comparison of mean packet service times for different values of α with N_o= 4. Here we compare two cases:

Case III. α = 100 calls/second

Case IV. α = 40 calls/second

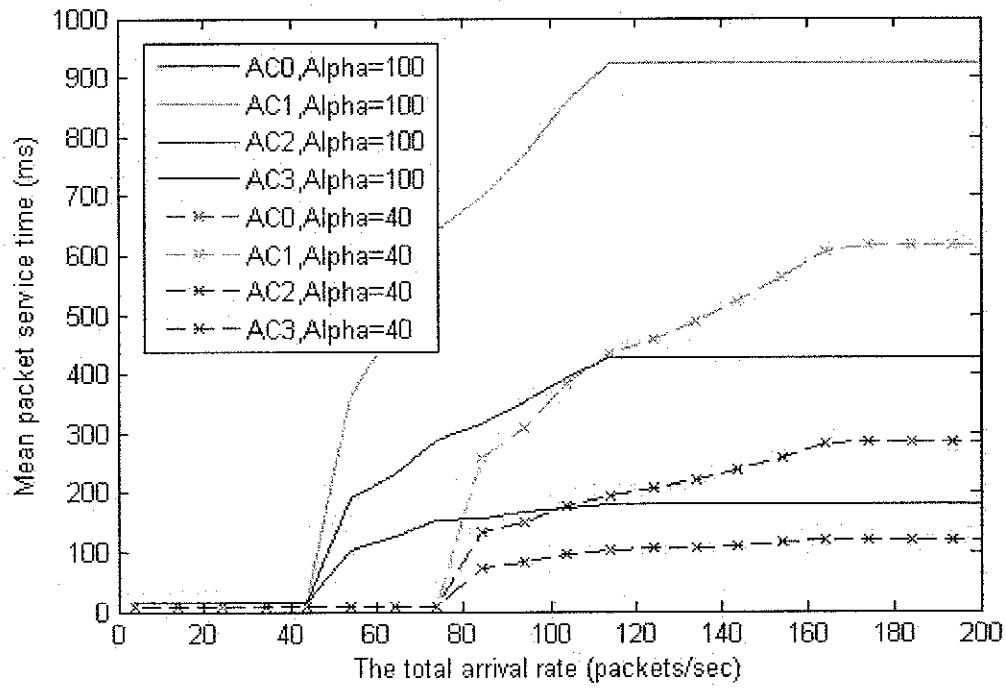


Figure 4.13 Mean packet service time comparison with different α

In Figure 4.13, we observe that the mean packet service time for Case III is greater than that of Case IV for all ACs. The reason is that for α = 100 and 40 calls/second the average number of busy channels in a superframe $\overline{N}_{DRP}$ is equal to 119 and 48, respectively. $\overline{N}_{DRP}$ is an increasing function of α, which results in the decrease of $\overline{N}_{PCA}$ and causes the increase of the packet service time ultimately. From Table 4.3, we can get $\overline{N}_{PCA}$ for the two cases as 137 and 208, respectively.

4.5.4 Mean Queuing Delay of a Packet

Next, we present the mean queuing delay for each AC packet as a function of the total packet arrival rate for $N_o = 4$ and $\alpha = 100$.

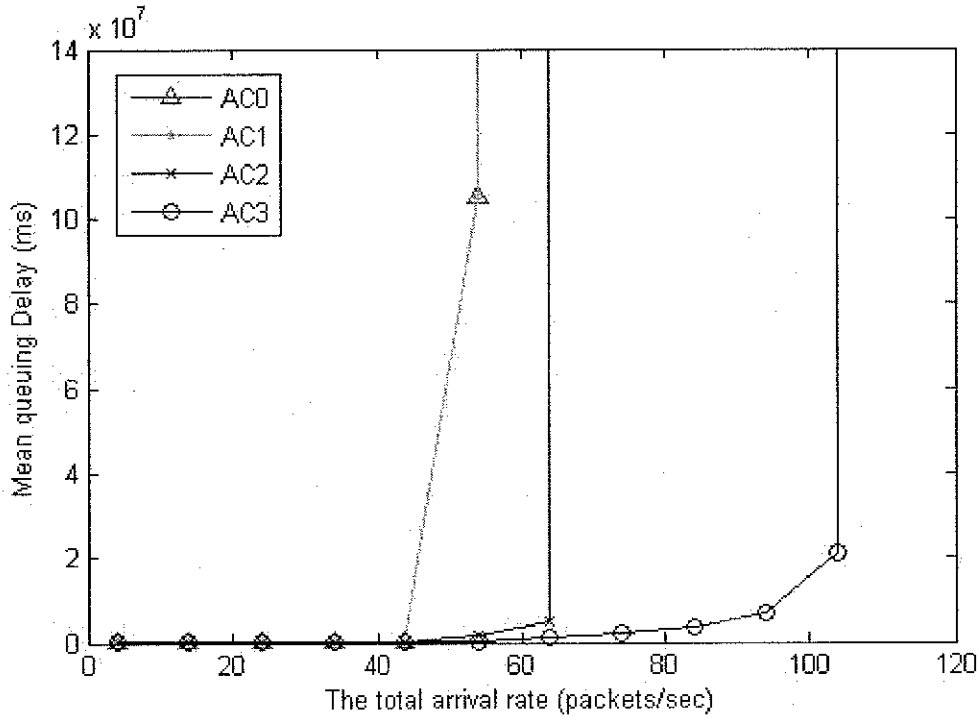


Figure 4.14 Mean queuing delay versus total arrival rate

Figure 4.14 presents the mean queuing delay results for the four ACs. It may be seen that the queuing delay of a packet increases with the total arrival rate and it goes to infinity when the queue gets saturated. The delay for lower priority traffic increases to infinity faster than the higher priority traffic. The reason is that as the traffic load increases the lower priority traffic is not able to access the channel anymore.

4.5.5 Mean Total Delay of a Packet

From (4.86), we can obtain mean total delay of a packet to be transmitted by summing the mean packet service time without post-backoff duration and mean queuing delay. We assume that $N_o = 4$.

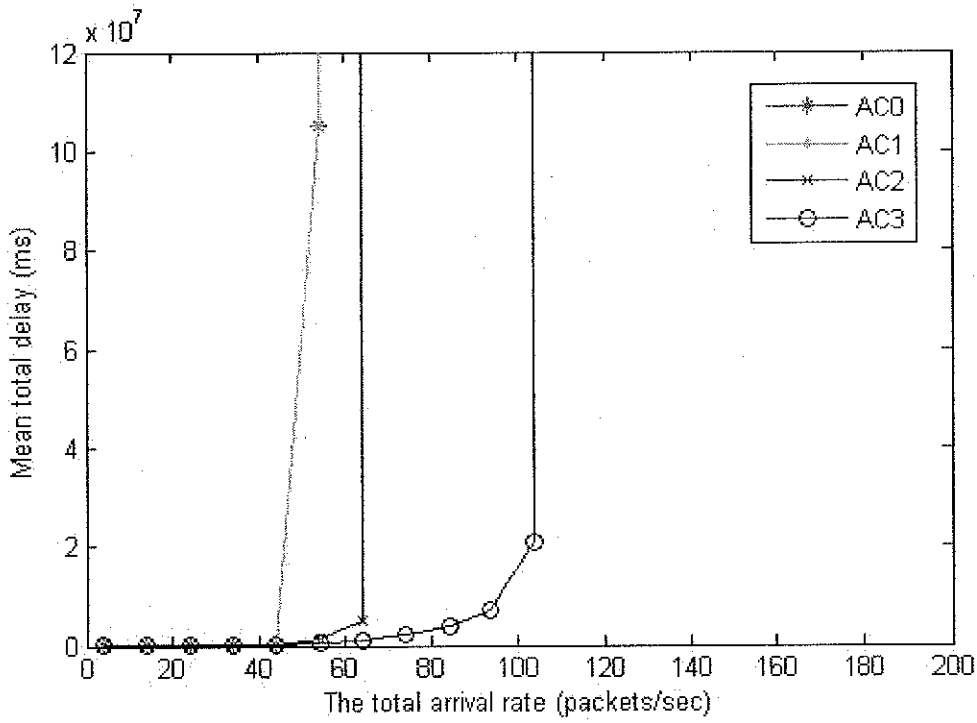


Figure 4.15 Mean total delay versus total arrival rate

The mean total delays of all four ACs versus the total arrival rate are presented in Figure 4.15. When the total arrival rate is low, the queues are not saturated, so the packets are transmitted within a short time. In this condition, the mean total delay mainly

corresponds to the mean packet service time. When the queue tends to be saturated, the queuing delay rises significantly and is much greater than the mean packet service time. Thus, when the total arrival rate gets too high, the mean total packet delay is dominated by the mean queuing delay.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

In this thesis, we present a joint analytical model that determines the performance of the WiMedia DRP and PCA schemes. Most of the previous work have been concerned only with the modeling of one of these two schemes and little attention has been paid to their simultaneous modeling. Further, previous analysis fails to take into account the availability of multiple transmission rates at the physical layer, which is one of the important characteristics of UWB. In our work, we model DRP as stream traffic and capture its impact on the performance of PCA traffic. We assume that DRP traffic consists of multiple classes of calls, which differ in their bandwidth requirements and transmission rates. The arrival of each class of calls is assumed to be according to a Poisson process. We use a blocking model for DRP traffic and determine the call blocking probabilities of each class and bandwidth utilization of the system. We believe that this is the first work that uses blocking model for DRP traffic.

Then we present a performance modeling of PCA in the presence of DRP traffic. We discuss the analysis of four access categories in the system and model each of them in a station as an M/G/1 queue. We derive the steady-state probability distribution of packet service time using a three-dimensional Markov chain model. Then we determine the throughput and mean packet total delay for all PCA access categories. The numerical results show that low priority PCA traffic experiences low throughput and high delay

under heavy traffic and the performance of the system for the higher priority levels is always better than that for the lower priority levels. We propose that a minimum number of channels to be reserved for PCA traffic to prevent their starvation. From the numerical results, we observe that the average number of channels being reserved for DRP increases is incremented with the call arrival rates. We propose reservation of a number of channels for exclusive by the PCA traffic to improve their performance.

In this thesis, we assumed identical $AIFS$ value for all ACs with service differentiation being determined only by minimum and maximum contention window sizes. In future work, different $AIFS$s can be assigned to different ACs to further differentiate the service.

The arrivals of the calls and packets are assumed to be according to Poisson processes. However, the information traffic is dynamic in real networks, especially in UWB networks. Hence, our model may be extended to general arrival processes, so that the traffic can be modeled more accurately.

CHAPTER 6

REFERENCES

[1] IEEE Std 802.15.4-2006 IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. - Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs), 2006.

[2] IEEE Std 802.15.1-2005 IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. - Part 15.1: Wireless medium access control (MAC) and physical layer (PHY) specifications for wireless personal area networks (WPANs), 2005.

[3] Standard ECMA-368: High Rate Ultra Wideband PHY and MAC Standard, 3st Edition, Dec. 2008.

[4] Standard ECMA-369: MAC-PHY Interface for ECMA-368, 3st Edition, Dec. 2008.

[5] K. Siwiak, P. Withington and S. Phelan, "Ultra-wide Band Radio: the Emergence of an Important New Technology," *IEEE Vehicular Technology Conference*, 2001, pp. 1169 – 1172.

[6] Federal Communications Commission (FCC), "Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems," First Report and Order, ET Docket 98-153, FCC 02-48; Adopted: Feb. 2002; Released: Apr. 2002.

[7] Website of the WiFi Alliance, "Discover and Learn." Internet: http://www.wi-fi.org/discover_and_learn.php [Nov. 13, 2009].

[8] MultiBand OFDM Alliance Special Interest Group (MBOA-SIG) White Paper, "Ultrawideband: High-speed, short-range technology with far-reaching effects." Internet: http://www.alereon.com/technology/white-papers, Sept. 1, 2004 [Nov. 13, 2009].

[9] Ruediger Kays, "Efficient Wireless Distribution of Electronic Media in Ultra-High Rate Home Networks Based on Radio." Internet:

http://www.ict-omega.eu/fileadmin/documents/presentations/training-seminar-5-Feb-09/Kays_Tutorial_Radio_2009-02-05.pdf [Nov. 13, 2009].

[10] M. Hofri and Z. Rosberg, "Packet delay under the golden ratio weighted TDM policy in a multiple-access channel," *IEEE Transactions on Information Theory*, vol. 33, no. 3, pp. 341–349, May. 1987.

[11] I. Rubin and Z. Zhang, "Message delay and queue-size analysis for circuitswitched TDMA systems," *IEEE Transactions on Communications*, vol. 39, no. 6, pp. 905–914, Jun. 1991.

[12] M. K. Khan and H. Peyravi, "Delay and jitter analysis of generalized demand-assignment multiple access (DAMA) protocols with general traffic," *Annual Hawaii International Conference on System Science*, 2005, pp. 304a–304a.

[13] H. Wu, Y. Xia, and Q. Zhang, "Delay analysis of DRP in MBOA UWB MAC," *IEEE International Conference on Communications*, 2006, pp. 229–233.

[14] N. Arianpoo, Y. Lin, V. W. S. Wong, and A. S. Alfa, "Analysis of distributed reservation protocol for UWB-based WPANs with ECMA-368 MAC," *IEEE Wireless Communications and Networking Conference*, 2008, pp. 1553-1558.

[15] G. Bianchi, "IEEE 802.11—saturated throughput analysis," *IEEE Communications Letters*, vol. 12, no. 2, pp. 318–320, Dec. 1998.

[16] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no 3, pp. 535–547, Mar. 2000.

[17] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: Throughput and delay analysis," *Computer Communications*, vol. 25, no. 3, pp. 313–321, Feb. 2002.

[18] Y. Wu, K. Long, S. Cheng, and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LANs: Analysis and enhancement," *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, 2002, pp. 599–607.

[19] Y. Xiao and J. Rosdahl, "Throughput and delay limits of IEEE 802.11," *IEEE Communication Letters*, vol. 6, no. 8, pp. 355–357, Aug. 2002.

[20] D. Malone, K. Duffy and D.J. Leith, "Modeling the 802.11 distributed coordination function in non-saturated heterogeneous condition", *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 159-172, Feb. 2007.

[21] Y. Xiao, "Performance Analysis of IEEE 802.11e EDCF under saturation condition," *IEEE International Conference on Communications*, 2004, pp. 170-174.

[22] Z. Kong, D. Tsang, B. Bensaou and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 10, pp. 2095–2106, Dec. 2004.

[23] X. Ling, K.-H. Liu, Y. Cheng, X. Shen, and J. W. Mark, "A novel performance model for distributed prioritized MAC protocols," *IEEE Globel Telecommunications Conference*, 2007, pp. 4692-4696.

[24] D.T.C. Wong, F.P.S. Chin, M.R. Shajan and Y.H. Chew, "Performance Analysis of Saturated Throughput of PCA in the Presence of Hard DRPs in WiMedia MAC," *IEEE Wireless Communications and Networking Conference*, 2007, pp. 423-429.

[25] R. Ruby and J. Pan, "Performance Analysis of WiMedia UWB MAC," *IEEE International Conference on Distributed Computing Systems Workshops*, 2009, pp. 504-510.

[26] P.E. Engelstad and O.N. Østerbø, "Non-Saturation and Saturation Analysis of IEEE 802.11e EDCA with Starvation Prediction," *ACM International Symposium on Modeling, Analysis & Simulation of Wireless and Mobile Systems*, 2005, pp. 224-233.

[27] P.E. Engelstad and O.N. Østerbø, "Analysis of the Total Delay of IEEE 802.11e EDCA and 802.11 DCF," *IEEE International Conference on Communication*, 2006, pp. 552-559.

[28] P.E. Engelstad and O.N. Østerbø, "The Delay Distribution of IEEE 802.11e EDCA and 802.11 DCF," *IEEE International Performance Computing and Communications Conference*, 2006, pp. 96-102.

[29] P.E. Engelstad and O.N. Østerbø, "Queueing Delay Analysis of 802.11e EDCA," *The Third Annual Conference on Wireless On demand Network Systems and Services*, 2006, pp. 123-133.

[30] P.E. Engelstad and O.N. Østerbø, "Delay and Throughput Analysis of IEEE 802.11e EDCA with Starvation Prediction," *The 30th Annual IEEE Conference on Local Computer Networks*, 2005, pp. 647-655.

[31] J.S. Kaufman, "Blocking in a Shared Resource Environment," *IEEE Transactions on Communications*, vol. 29, no. 10, pp. 1474-1481, Oct. 1981.

[32] J.F. Hayes and T.V.J.G. Babu, *Modeling and Analysis of Telecommunications Networks*. Hoboken, New Jersey: John Wiley & Sons, 2004.

[33] M.M. Ali, "Call-burst blocking and call admission control in a broadband network with bursty sources", *Performance Evaluation*, vol. 38, pp. 1-19, 1999.