國 立 交 通 大 學

電信工程學系碩士班

碩 士 論 文

企業/家庭 無線區域網路閘道之乏晰允諾控制

Fuzzy CAC for WLAN Gateway in SOHO/Home Networks

研 究 生：吳俊憲

指導教授：張仲儒 博士

中華民國九十三年十月

企業/家庭 無線區域網路閘道之乏晰允諾控制

Fuzzy CAC for WLAN Gateway in SOHO/Home Networks

研 究 生：吳俊憲　　　　　　　　　　Student: Jiun-Shian Wu

指導教授：張仲儒　博士　　　　　　　Advisor: Dr. Chung-Ju Chang

國立交通大學
電信工程學系碩士班
碩士論文

A Thesis
Submitted to Institute of Communication Engineering
College of Electrical Engineering and Computer Science
National Chiao Tung University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In
Electrical Engineering
October 2004
Hsinchu, Taiwan, Republic of China

中華民國九十三年十月

企業/家庭 無線區域網路閘道之乏晰允諾控制

研究生：吳俊憲　　　　　　　　　　指導教授：張仲儒 博士

國立交通大學電信工程學系碩士班

中文摘要

近年來企業/家庭無線區域網路的普及發展使得多種無線傳輸服務應用將成為主要的趨勢，針對各種不同應用的服務品質，如何有效保證其服務品質並同時增進頻寬使用效益將是系統設計的一大考驗。允諾控制為此提供了一個方法有效保證服務品質同時亦能最大化頻寬使用效益。在此，允諾控制考慮新進服務的特性，要求的服務品質和系統的剩餘資源以及系統量測品質做出決策。

針對各種不同問題，乏晰理論提供了一個有效簡單的邏輯設計。在本篇論文中我們提出了一個乏晰允諾控制法則，針對企業/家庭 無線區域網路下的多種應用提供一個保障服務品質的設計。並說明其設計的理念，確保此設計的正確性。我們提出一個比例性的時間優先頻率排成提供不同服務品質之間的使用者有著比例分配的服務品質，並同時藉由乏晰允諾控制提供不同使用者達到其要求服務品質。

Fuzzy CAC for WLAN in SOHO/Home Networks

Student: Jiun-Shien Wu   Advisor: Dr. Chung-Ju Chang

Institute of Communication Engineering
National Chiao Tung University

Abstract

In recent years, the popular development of SOHO/Home WLANs make the various service transmissions a major tendency in the future. For various quality of services (QoS), a real challenge of system design issues is how to guarantee these QoS and increase the utilization of system at the same time. Call admission control provides a effective method for QoS guarantee and also maximizing utilization. Here, call admission control consider the traffic characteristics and required QoS of a new call, the remaining resource of system and quality measured by system.

Fuzzy theorem provides a simple logic design for various problems. In this thesis, we provide fuzzy rules for call admission control in SOHO/Home networks, and the design could guarantee the various QoS in WLAN. We explain the reasons and principles for the design procedure. We also provide a weighted priority time period for the different services to make their quality as fair ratio relationship and we achieve the various QoS requirements through the fuzzy call admission control.

誌　謝

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Wireless local area networks (WLAN) are being developed to provide high bandwidth to users in a limited geographical area, and WLAN will then be attached to backbone networks like the Internet, or interface to wireless wide area networks (WAN) for range extension. In recent years, the IEEE 802.11 WLAN has become a prevailing broadband wireless technology [1], and the IEEE 802.11 WLAN can be considered as a wireless version of Ethernet, which supports best-effort service. However, not only data transmissions are concerned but also multimedia traffic requiring constant, variable, and available bit rate services will be employed on WLAN in this future.

At present, the IEEE 802.11 standard MAC protocol supports two kinds of access methods: DCF (Distributed Coordination Function) and PCF (Point Coordination Function). The performance of the DCF has already been studied in many literature researchers [2-5]. Regarding the PCF, we can find performance evaluation taking into account voice or video transmission with the PCF [6-9]. In [6], the combined performance of data transmission with the DCF and voice transmission with PCF is evaluated. The authors studied the combined performance of data transmission with DCF and video transmission with the PCF in [7]. Performance evaluation considering video transmission with the PCF could also be found in [8]. In [9], the performance of priority-based multimedia transmission with PCF was evaluated.

Also, for quality of service (QoS) in application services, IEEE 802.11 Task Group E currently defines enhancements to 802.11 MAC, called 802.11e [10], which introduces a hybrid coordination function (HCF). The HCF consists of a contention-based channel access called enhanced distributed channel access (EDCA) mechanism for contention-based transfer and a controlled channel access, referred to as HCF Controlled Channel Access (HCCA) mechanism, for contention free transfer.

This new standard provides the means of prioritizing the radio channel access by different stations and media access control (MAC) data streams. Stations, which operate under the 802.11e and provide the QoS facility, are called QoS stations (QSTAs), and a QoS

station, which works as the centralized controller for all other QSTAs within the same basic service set, is called Hybrid Coordinator (HC). The HC will typically reside within an 802.11e QoS access point (QAP) connecting its QSTAs with infrastructure.

A basic unit of allocation the right to transmit onto the wireless medium (WM) is the transmission opportunity (TXOP). A TXOP is defined as an interval of time when a particular QSTA has the right to initiate transmissions, and defined by a starting time and a maximum duration. The TXOP may be obtained by QSTA winning an instance of EDCF during the contention period (CP) called EDCA TXOP , or by QSTA receiving a QoS(+) CF-Poll during the CP or contention free period (CFP) called HCCA TXOP or polled TXOP. The duration of an EDCA TXOP is limited by TXOP Limit distributed by the HC in beacon frames and includes the time required to transmit frames sent as an immediate response to TXOP holder. An essential characteristic of 802.11e is that HC operates during both the CP and CFP which gives a HC the highest priority in medium access. And a controlled access phase (CAP) is a time period when the HC maintains control of the medium, after sensing the channel to be idle for a PCF Interframe Space (PIFS) duration and gaining access to the medium. A CAP may span multiple consecutive transmission opportunities (TXOPs) and may contain several polled TXOPs.

The EDCA is based on Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA), and supports four access categories (ACs). The four ACs are best effort, background, video, and voice, and each AC is a common set of EDCA parameters that are information of AIFS, CW_{min}, CW_{max}, and TXOP Limit used by a QSTA for contention of the channel for the transmission of MAC service data units (MSDUs) with certain priorities, where AIFS is the arbitration interframe space (AIFS), and CW_{min}/CW_{max} is the minimum/maximum of the contention window (CW) of the AC. Different EDCA parameters of the four ACs are broadcasted in the beacon by the HC, and the minimum AIFS for a QAP is PCF interframe space (PIFS) smaller than the DCF interfrane space (DIFS), the minimum AIFS for a non-QAP QSTA. The CW_{min}/CW_{max} has the range of 0 to 32767 from $2^i - 1$ where i is the integer from 0 to 15. There are four CWs of associated AC, as referred to CW[AC], which shall be initialized to the value of the parameter CW_{min}[AC]. And for each AC, the corresponding backoff timer is set to an integer value chosen randomly with a uniform distribution in the range [0, CW[AC]]. The CW[AC] may be updated to the value (CW[AC]+1)*2-1 if CW[AC] is less than CW_{max}[AC], or remain unchanged for the remainder of retries if the CW[AC] is equal to the CW_{max}[AC].

In this thesis, we adopt the HCF which a CAP is generated periodically in the beginning of a superframe after sensing the channel to be idle for PIFS duration, and our system is based on 802.11e draft 6.0 with 802.11a for physical layer transmission.

In chapter 2, we proposed a modified method with inter frame space (IFS) for data transmission avoiding much idle time for data contending access. The system we consider is with full queue traffic model, voice service access would be transmitted in the CAP and CP and video service is transmitted in the CAP. In section 2.1, system model is described including system architecture and system operation on 802.11. In the section 2.2, we describe traffic sources which are voice, video and data traffics in the system. In section 2.3, we discuss the simulation results under the figures about throughput, drop rate, inter-arrival delay of the related service. Finally, we have concluding remarks in section 2.4.

The chapter 3 is outlined as follows. In section 3.1 some respects of call admission control are introduced under various traffics in the wireless system. In section 3.2, the system model is described including the system architecture, the related call request operation, system operation in 802.11 and traffic sources are described including video, voice, and data traffic. In section 3.3, the fuzzy call admission control is described including the goal, the fuzzy call admission controller and fuzzy rules. In the section 3.4, we describe the Gaussian approximation used for the equivalent CAP estimator. Finally, concluding remarks are given in section 3.5.

At last, the conclusion is given by chapter 4.

Chapter 2

Simulation Study of IEEE 802.11e HCF on Hybrid Services

2.1   System Model
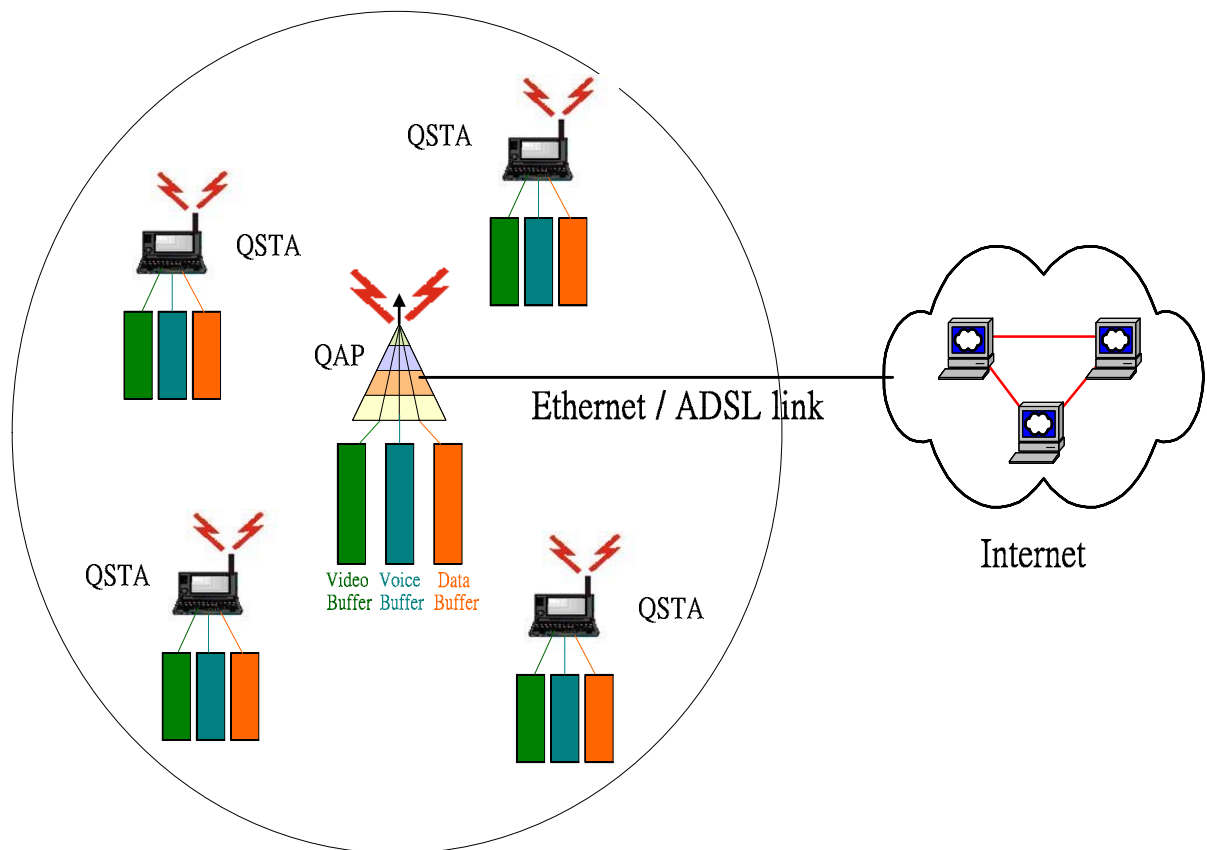
Figure 2.1: Indoor SOHO WLAN Infrastructure Environment

2.1.1   System Architecture

For an indoor SOHO WLAN environment shown in figure 2.1, there is a QoS access point (QAP) having only two interfaces, an Ethernet interface and an 802.11e interface. The Ethernet/ADSL link supports the data traffic and the real time service traffic sent from or to

the Internet, and the 802.11e interface is used to transmission of the packets in the wireless environment. In this thesis, the end host is called the quality of service station (QSTA) for which we intend to provide the quality of service (QoS) facility. In the infrastructure environment of Small Office, Home Office (SOHO), we assume that all video, voice, and data packets are sent through the QAP even though in some cases, when both QSTAs are within in the coverage area of an QAP, voice and data packets could not be sent directly between the QSTAs.
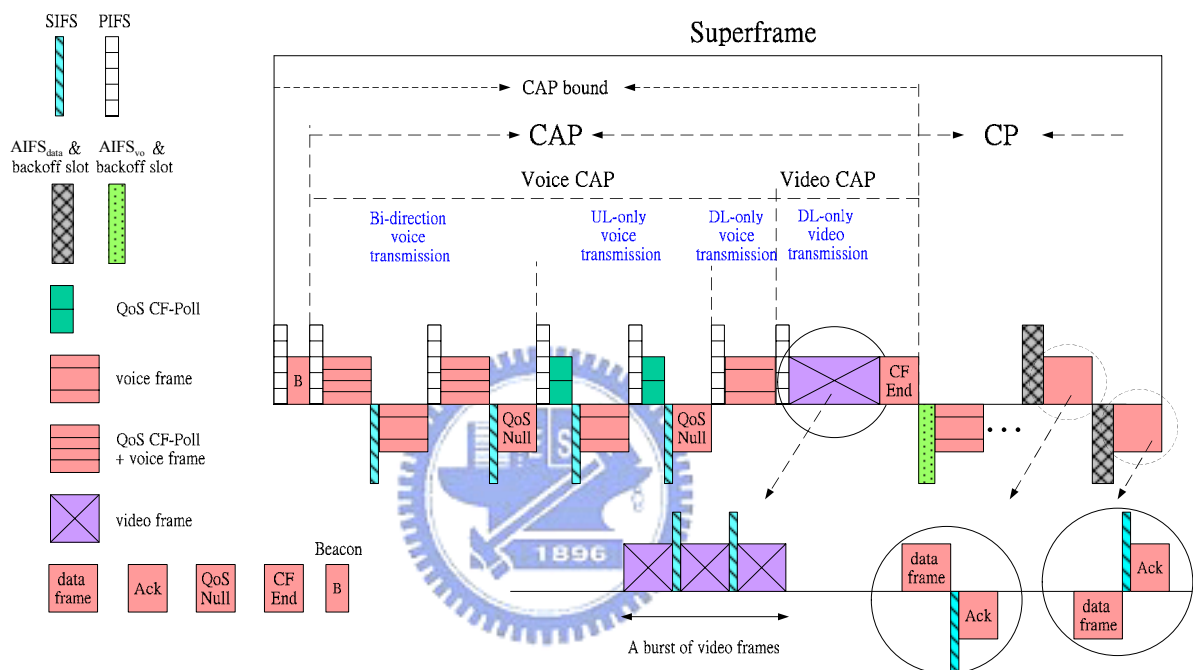
Figure 2.2: An example of a superframe timing

2.1.2   System Operation

In a superframe, it can be divided into two periods: controlled access phase (CAP) and contention period (CP) as shown in Fig. 2.2. In the CAP, the QAP polls the QSTAs to get the uplink voice frames or transmits the downlink voice frames and the downlink video packets directly without acknowledge based on HCF controlled channel access. In the CP, the non-real time data frames and intended contending voice frames of QSTAs would use different priorities to contend the channel based on EDCF.

The CAP within a superframe can be divided into two transmission phase: Voice CAP, and Video CAP. In addition, the voice service within a Voice CAP also can be sequentially served into three transmission period: bi-direction voice transmission, uplink-only (UL-only)

voice transmission, and downlink-only (DL-only) voice transmission. In a superframe, there is at most one voice frame for uplink transmission of a QSTA, and also the QAP has at most one downlink voice frame for each QSTA with the conversation. At the beginning of the Voice CAP, the QAP decides the polling sequence based on the polling list. The first voice service of polling sequence is bi-direction voice transmission. If the QAP has a downlink voice frame queued for transmission to the QSTA in the polling list at the QAP, the downlink voice frame and QoS(+)CF-Poll can be combined and transmitted as a single frame by the QAP after the channel is sensed to be idle for a PIFS interval. After receiving the frame, the destination of the QSTA can transmit an uplink voice frame to the QAP after a SIFS interval. This frame exchange procedure is called bi-direction voice transmission. After bi-direction voice transmission, for the QSTA in polling list but the QAP does not have a voice frame queued for it, the QAP sends a sole QoS(+)CF-Poll frame to the QSTA after the channel is sensed to be idle for a PIFS interval, and waits to receive an uplink voice frame from the QSTA after a SIFS interval. This frame exchange procedure is called uplink-only (UL-only) voice transmission. If the QAP receives a QoS Null frame from the QSTA during bi-direction voice transmission or UL-only voice transmission, the QAP regards the QSTA entering into silence period and removes the QSTA from the polling list in the QAP. After UL-only voice transmission, the QAP consecutively sends the remaining downlink voice frames with no acknowledge after waiting a PIFS channel idle interval during downlink-only (DL-only) voice transmission.

In the Video CAP, the QAP sends the multiple video packets belong to the QSTA in a burst and Round Robin manner after the channel is sensed to be idle for a PIFS interval. During the frame interval, the QAP receives several video packets of the frame interval belong to the QSTA from Ethernet/ADSL link and the QAP may transmit these video packets during the following three CAPs in a burst without acknowledge which are separated by SIFS interval in the video CAP. These video packets in the QAP are lasting availably for the following three CAPs or until they are transmitted completely. While these video packets are not completely transmitted by QAP during the following three CAPs, the QAP would not transmit these packets which are dropped by the QAP, and the QAP will transmit the at most packets of next frame interval. During the following two CAPs, if there are enough residual time in the CAP bound for a video packet transmission, the QAP sends the packet belong to the QSTA which is at service turn and also sends the remaining packets at next video CAP in first after the channel is sensed to be idle for a PIFS interval. The video packets belong to a

QSTA may be transmitted in two consecutive video CAPs. Each video burst transmission on the wireless medium is corresponding to several video packets arriving at the QAP during a frame interval intended to a QSTA. After the QAP transmits a burst of video packets to a QSTA, the QAP continues to transmit another video burst of the next QSTA in a cyclic manner (that is so called Round Robin) after a PIFS channel idle interval. In addition, after all bursts of the frame interval are transmitted, the QAP consecutively transmits the bursts of the next frame interval if they are all arrived at QAP. In addition, until all bursts of video packets arriving at QAP before the superframe are transmitted completely by the QAP, and QAP sends a CF-end frame as termination of the video CAP. Fig. 2.2 shows an example of a superframe timing.

In the CP, the voice frames for uplink but not polled during the voice CAP in the same superframe will contend with data frames. The data frames including QAP downlink data frames and QSTA uplink data frames. The QAP should assign different AIFS, $\mathrm{CW_{min}}$, and $\mathrm{CW_{max}}$ for the contending uplink voice service and data service. After the QAP receives an uplink contending successful voice frame from the QSTA, the QAP re-adds the QSTA in the polling list. When the counter of a data frame contending for transmission counts down to zero and there is enough time for a complete data frame exchange in CP, the QSTA would transmit the data frame. Also, it happens to a data frame having zero-counter in the beginning of the CP because there is not enough access time for the frame exchange in the last CP. In the beginning of the CP, if a data frame is first transmitted, the voice service would delay for at least a data frame transmission period in spite of the collision happened or not. In order to guarantee the voice delay as small as possible, the AIFS for voice service is set to equal to PIFS, denote as $\mathrm{AIFS_{vo}} = \mathrm{PIFS}$. And the $\mathrm{CW_{min}}$ of a voice frame, denote as $\mathrm{CW_{min\_vo}}$, and $\mathrm{CW_{max}}$ of a voice packet, as $\mathrm{CW_{max\_vo}}$ are given by

$$\mathrm{CW_{min\_vo}} = \mathrm{CW_{max\_vo}} = \begin{cases} \lceil E[N_C] \rceil & \text{if } N_{NP} \leq 1 \\ \max\left\{ 2, \ \lceil E[N_C] \rceil \right\} & \text{if } N_{NP} \geq 2 \end{cases} \qquad (2.1)$$

where N_C is the number of voice packets contending in a superframe, N_{NP} is the number of QSTA with voice conversation not in the polling list, and $E[N_C]$ is the average contending voice packet number of the superframe which may be zero if N_{NP} is equal to zero. Also the $\mathrm{CW_{min\_vo}}$ and $\mathrm{CW_{max\_vo}}$ is the smallest integer greater or equal to the mean value of N_C. Here, we guarantee that the contending uplink voice service has a higher priority

than non-real-time data service. That is the contending voice frames are always served before the data frames. In a CP, when a data frame of a QSTA is served in the CP, it stands for contending voice frames has been completely served and hereafter only the data frames contend with each other in the CP. In order to avoid the uplink contending voice frame to be terminated by non-real-time data frames, the AIFS of non-real-time data service, denote as AIFS$_{data}$ is given by

$$AIFS_{data} = \begin{cases} PIFS + SlotTime * (CW_{max\_vo} + 1) & \text{for the first frame transmitted of a QSTA in a CP} \\ DIFS & \text{others} \end{cases} \quad (2.2)$$

where the "+1" is for contending uplink voice frames avoiding collision with a data frame having zero-counter. The minimum and maximum contention window of a data frame, CW_{min_data} and CW_{max_data}, could be variable set by the retransmit times of the data frame. And CW_{min_data} and CW_{max_data} of the first frame of a QSTA in a superframe are the same and not changed even during the retransmissions. If a data frame is retransmitted five times and it is still not successful, the data frame is dropped by the QSTA.
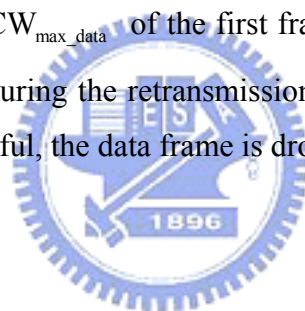
2.2 Traffic Source Models

Now we describe our traffic source models. The services we want to support are voice, video and data traffic.

Voice traffic: As shown in Fig. 2.3, the conversational dynamics of a communication pair is modeled as a four-state model, and the random process of conversational pairs in the coverage of a QAP is assumed to be independent and identically distributed. The four-state is described as follows that uplink and downlink are both talking, called Double Talk; Uplink and downlink are both silent, called Mutual Silence; Uplink is talking and downlink is silent, defined as Talkspurt; Uplink is silent and downlink is talking, defined as Pause. In the model, the sojourn time of each state in which a conversation pair state is assumed to be exponentially distributed. The state transition rates, λ_{ij}, can be obtained by fitting the mean duration of talkspurt, pause, double talk, and mutual silence, given in [11]. Within each

talking spurt interval for each member of a conversational pair, 8 kb/s pulse code modulated (PCM) digital voice is assumed and the mean rate of a communication pair given parameters in Fig. 2.3 is 6.8 kb/s. During a talking spurt interval, a voice frame is coded every superframe interval and the coder generates some voice frames which are all the same size depending on the coding rate and superframe interval. We consider the superframe interval as the maximum acceptable value for one-way voice transfer delay over the WLAN delay. When a voice frame is not transmitted successfully in this superframe, it is dropped by the QAP or QSTA

The state transition parameters
(unit: 1/sec)
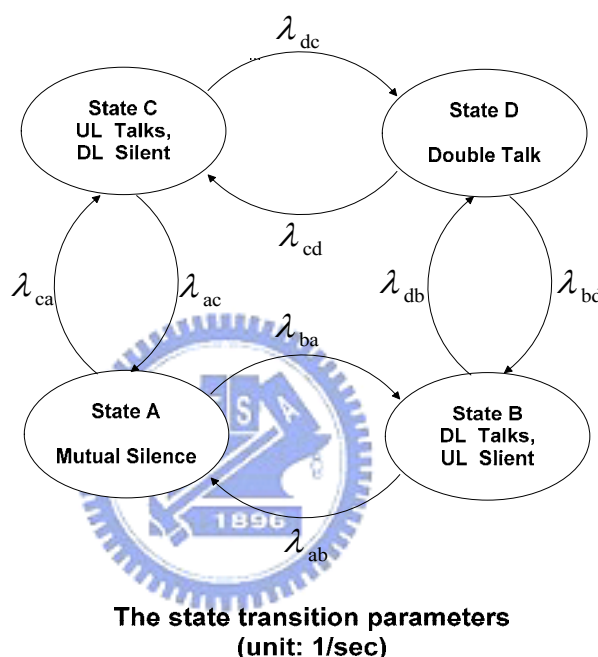
λ_{ab}	λ_{ba}	λ_{bd}	λ_{db}
λ_{ac}	λ_{ca}	λ_{cd}	λ_{dc}
1.63	0.846	0.279	2.24

Figure 2.3: The four-state Markov model of a conversation pair

Video traffic: In the article, the frame generation process for a video coder was assumed to have two motion states: the low motion state for the rate of interframe coding, and the high motion state for rate of intraframe coding. The rate of intraframe coding was further divided into two parts: the first part had the same rate as the interframe coding and the second part, called difference coding, was the difference rate between intraframe coding and interframe coding [12]. The interframe coding and the difference coding were both modeled as discrete-state Markov-modulated poisson process (MMPP) with basic rates A_r and A_a

[13]. The state-transition diagram is shown in Fig. 2.3(a) and 2.3(b). The video source will alternate between interframe and intraframe, depending on the video source activity factor. As shown in Fig. 2.4(c), there is a transition rate c in the interframe state and a transition rate d in the intraframe state. Let $\lambda_a(t), \lambda_r(t), \lambda_a'(t)$ denote the frame generation rates for intraframe coding, interframe coding, and difference coding at the time t, respectively, from the video coder. Clearly, $\lambda_a(t) = \lambda_r(t) + \lambda_a'(t)$ The process of $\lambda_r(t)$ is an $(M_r + 1)$-state MMPP process. The state-transition diagram for $\lambda_r(t)$ uses the label $m_r A_r$ to indicate the frame generation rate of interframe coding of a state and use the labels $(M_r - m_r)\gamma$ and $m_r\omega$ to denote the transition rate from state $m_r A_r$ to state $(m_r + 1)A_r$, and from state $m_r A_r$ to state $(m_r - 1)A_r$, respectively. Similarly, the process for $\lambda_a'(t)$ is an $(M_a + 1)$-state birth-death Markov process. The state-transition diagram for $\lambda_a'(t)$ uses the label $m_a A_a$ to indicate the additional frame generation rate of a state due to intraframe coding and uses the labels $(M_a - m_a)\phi$ and $m_a\varphi$ to denote the transition rate from state $m_a A_a$ to state $(m_a + 1)A_a$ and from state $m_a A_a$ to state $(m_a - 1)A_a$, respectively. One should be note that the long-term correlation behavior of a video source results from the process $\lambda_a(t)$. The values of γ, ω, M_r, A_r, ϕ, φ, M_a, A_a, c and d can be obtained from the traffic variables of R_p, R_m and T_p [13,14]. It was assumed that $R_p = 4.95$ Mbps, $R_m = 1.65$ Mbps, $T_p = 0.5$ second, which give $M_r = M_a = 20$, $A_r = 2.01 \times 10^5$ bit/sec, $A_a = 4.71 \times 10^4$ bit/sec, $\gamma = 1.33$, $\omega = 2$, $\phi = \varphi = 1$, $c = 2$, and $d = 18.01$ (1/ sec).
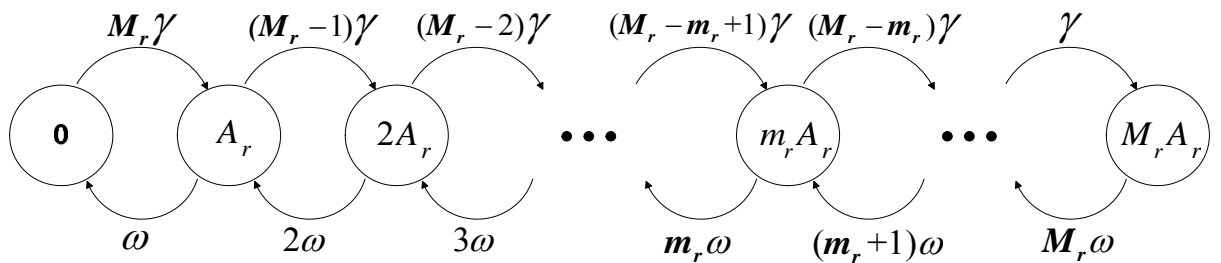
Figure 2.4(a): State transition diagram for interframe coding $\lambda_r(t)$
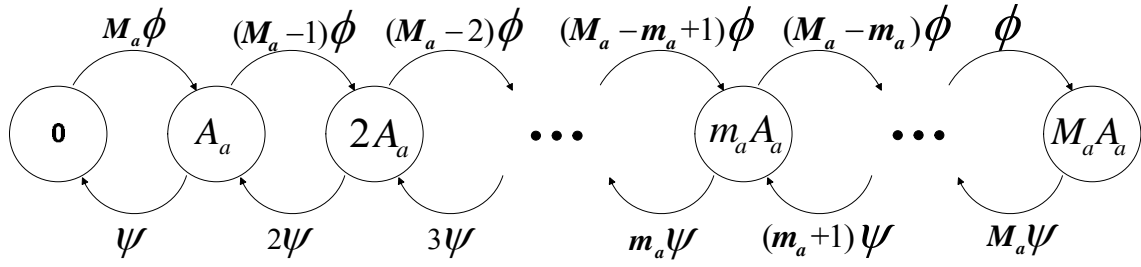
Figure 2.4(b): State transition diagram for difference coding $\lambda'_a(t)$
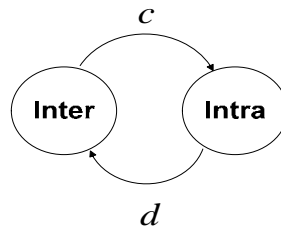
Figure 2.4(c): Interframe and intraframe alternate model for a video source

Data traffic: We have an assumption for data traffic that there is always a data frame of a QSTA ready to be sent. When a data frame is transmitted or dropped by a QSTA/QAP, there is always another data frame in QSTA/QAP intended for contenting the wireless medium at instant.

2.3   Simulation Results

Parameters of the simulation environment are set in the Table 1.1, and the values of PHY-related parameters were referred to IEEE 802.11a [16].

Environment parameter	Symbol	Value
Superframe period	T_S	20 ms
CAP bound	CAP_{bound}	15 ms
Transmission rate	R_T	12, 24 Mbps
Voice coding rate	Rc	1000 bit/sec
Slot time	σ	9 µs

SIFS time	SIFS	16 μs
PIFS time = SIFS time + Slot time	PIFS	25 μs
PIFS time = SIFS time + 2*Slot time	DIFS	34 μs
Minimum contention window size for data service	CW_{min}	32
PLCP preamble duration	tPLCPPreamble	16 μs
PLCP SIGNAL field duration	tPLCP_SIG	4 μs
OFDM symbol interval	tSymble	4 μs
MAC payload	P_{MAC}	1500 Bytes
MAC header (QoS data type)	H_{MAC}	36 Bytes

Table 2.1: System parameters

Figure 2.5 shows the average CAP period versus the voice pair number with six scenarios, and figure 2.6 shows the average voice CAP period versus the voice pair number with four scenarios. In figure 2.5, for the two scenarios of transmission rate 12 Mbps & video user 5 and 7, the average CAP period is saturated and almost the same whether the voice pair increases or not. The saturation phenomenon under the two scenarios results in that transmission rate 12 Mbps & CAP_{bound} (15 ms) could not support more than five video users with a low drop rate. At the scenario of transmission rate 12 Mbps & video user 5, when the voice pair number is 10, the average CAP period is lightly smaller than the saturation parallel line, and video frames have been already dropped in some degree. The gap between the average maximum CAP period and CAP_{bound} is about half of a video packet transmission interval with payload is 1500 bytes. The factor of this gap is resulted from the 802.11e transmission rule, and the value of this gap is about 0.5 ms for the transmission rate 12 Mbps and the payload of a video packet size, 1500 bytes. For the other four scenarios, the increasing amount of the average CAP period for transmission rate 12/24 Mbps with voice pair number increasing is equal to the increasing of average voice CAP period for transmission rate 12/24 Mbps with voice pair number increasing in Figure 2.6. The increasing amount of average CAP period while increasing voice pair number could be approximated as follows:

$$\Delta = \frac{1}{T_s} \sum_{i=1}^{4} \frac{R_{voice}}{P_{voice\text{-}payload}} p(i) \cdot t(i) \tag{2.3}$$

where $P_{voice\text{-}payload}$ is the payload of the voice packet (20 byte), R_{voice} is the mean rate of a voice pair (6.8 kbps) in figure 2.3, $p(i)$ is corresponding probability at the state in figure 2.3 and $t(i)$ is the time interval for a voice pair transmission at the state. The Δ=0.12 and 0.9 (ms per pair) for transmission rate 12 and 24 Mbps are the same as figure 2.6. In figure 2.6, the average voice CAP period is independent of the video users because the transmission priority of voice transmission is higher than the video packet transmission.

Figure 2.7 shows the average contention period (CP) versus the voice pair number, and the sum of figure 2.6 and figure 2.7 is the superframe period (T_s= 20 ms). For the two scenarios of transmission rate 12 Mbps & video user 5 and 7, the average contention period achieves the minimum when the average CAP period achieves the maximum value and video packets has been dropped obviously. For the other four scenarios, while the voice pair number increases, the average contention period reduces because the voice transmission is prior to data transmission, and the reducing amount of average contention period for transmission rate 12/24 Mbps with voice pair number increasing is equal to the increasing amount of average voice CAP period for transmission rate 12/24 Mbps with voice pair number increasing in Figure 2.6.

Figure 2.8 shows the average used superframe period versus the voice pair number. The average used superframe period is all the same just depending on the transmission rate. While the gap interval between the average used superframe period and superframe period (T_s) is about the half of a data frame transmission interval depending on the transmission rate.

Figure 2.9 shows the average real time service period versus the voice pair number. The period includes the CAP period and voice contending period. Figure 2.10 shows the average voice contending period versus the voice pair number. The average voice contending period increase with the voice pair number increasing. The figure 2.9 can be added form the figure 2.1 and figure 2.10, and because the average voice contending period is much smaller than the average CAP period, the figure 2.1 and figure 2.9 are similar.

Figure 2.11 shows the average video frame drop rate versus the video users. For the transmission rate 12 Mbps, video frames begin to be dropped when there are four video users,

and dramatically increases with video user increasing. While the voice pair number increases, the average video frame drop rate also increases because the video CAP period is reducing. When the video user is five, the average CAP period reaches the maximum and hereafter the increasing of voice pair number would make the average video drop rate increase with a constant value. We can see the same phenomenon for video user six and seven. The phenomenon results from the same increasing voice transmission would make the same reduced period of video transmission. For the transmission rate 24 Mbps, the average frame drop rate is about zero for video user 3 to 7 and voice pair number 10 to 30. It also can be seen from the figure 2.5 that the average CAP period is much smaller than the CAP_{bound} for transmission rate 24 Mbps & voice pair number 10 to 30.

Figure 2.12 shows the average inter-arrival period of video frame versus the video user. When the average inter-arrival period of video frame is video frame interval (video frame interval = 2*superfame = 40 ms), the corresponding average video frame drop rate in figure 2.8 is almost zero. For transmission rate 24 Mbps & voice pair number 10, 20, and 30, the average inter-arrival period of video frame is 40 ms. and their average video drop rates are almost zero which also means that the average CAP period is much smaller than the CAP_{bound}, 15 ms. For transmission rate 12 Mbps, the average inter-arrival period of video frame is larger than 40 ms which means there are video frames dropped, and the average inter-arrival period increases dramatically with increasing of video users. The average inter-arrival period of video frame and the corresponding average video frame drop rate could be transformed from each other as follows:

$$1 - \frac{2 \cdot T_s}{t_{inter}} = p_{drop} \tag{2.4}$$

where T_s is superframe period, t_{inter} is the average inter-arrival period of video frame, and p_{drop} is the average video frame drop rate. we take a example to verify that for the scenario of transmission rate 12 Mbps & voice pair 20, when the video user is six, the corresponding average inter-arrival period is 60 ms (3*superframe) in figure 2.12 and the corresponding average video frame drop rate is 1-(40/60)=0.33 in figure 2.11.

Figure 2.13 shows the data service throughput versus the voice pair number. For the scenario of transmission rate 12 Mbps & video user 5, the data service throughput is independent of the voice pair number, because the average CAP period achieves the maximum close to CAP_{bound}. And the data service throughput is the same whether the voice

pair number increases or not. We can see that the minimum contention period is 5 ms. which means data bandwidth is allocated to 12*(5/20) = 3 Mbps. We see that the efficiency of data contention is

$$\mu = \frac{2.13}{3} = 0.71 \tag{2.5}$$

For other scenarios in figure 2.13, the data service throughput reduces with the increasing with voice pair number for the reduced contention period. And, for transmission rate 12 Mbps & video user 3, when the voice pair number increases 10, the average contention period would reduce 1.2 ms. and total data service throughput would reduce by $12 \cdot (\frac{1.2}{20}) \cdot \mu = 0.51$ Mbps, we can prove it by the values in figure 2.13. The data service throughput can be get form the average CP in figure 2.10 as follows:

$$throughput_{data} = R_T \cdot \frac{t_{CP}}{T_S} \cdot \mu \tag{2.6}$$

where t_{CP} is the average CP period in the figure 2.7, R_T is the transmission rate, μ is data contention efficiency, and T_S is the period of superframe.

Figure 2.14 shows the video service throughput versus the voice pair number. For the scenario of transmission rate 12 Mbps & video user 5, the average CAP period achieves the maximum closed to CAP_{bound}. While the voice pair number increases, the video CAP period reduces and the video throughput also reduces. For the other scenarios, even the number of voice pair increases the video throughput not changes because the CAP period does not achieve the maximum, and the increasing of voice pair number would not reduce the period of video transmission. In addition, we see that for the scenario of transmission rate 12, and 24 & video user 5, the video throughput is different resulted from drop rate. The ratio of difference to the total video throughput is as the same as the video frame drop rate in figure 2.8.

2.4 Concluding Remarks

In this chapter, we modified the inter frame space (IFS) of data frames after the contending transmission of voice frames, and we find the efficiency under the full loading of data service in contention period.

For video downlink service, we study the related frame drop rate corresponding to the relative inter-arrival delay under video user and voice pair number. And we find the fair drop

rate of each video user under Round Robin scheduling. The dynamic of voice frames transmission in CAP and CP avoid the redundant voice polling time, and the throughput of data service or supporting users under delay requirement could increase apparently. For the property of large variance in video traffic, the video frame drop rate increases apparently while increasing the video user. And our simulation discussion can provide a design reference about data throughput, voice pair number, and video users in the real environment system.
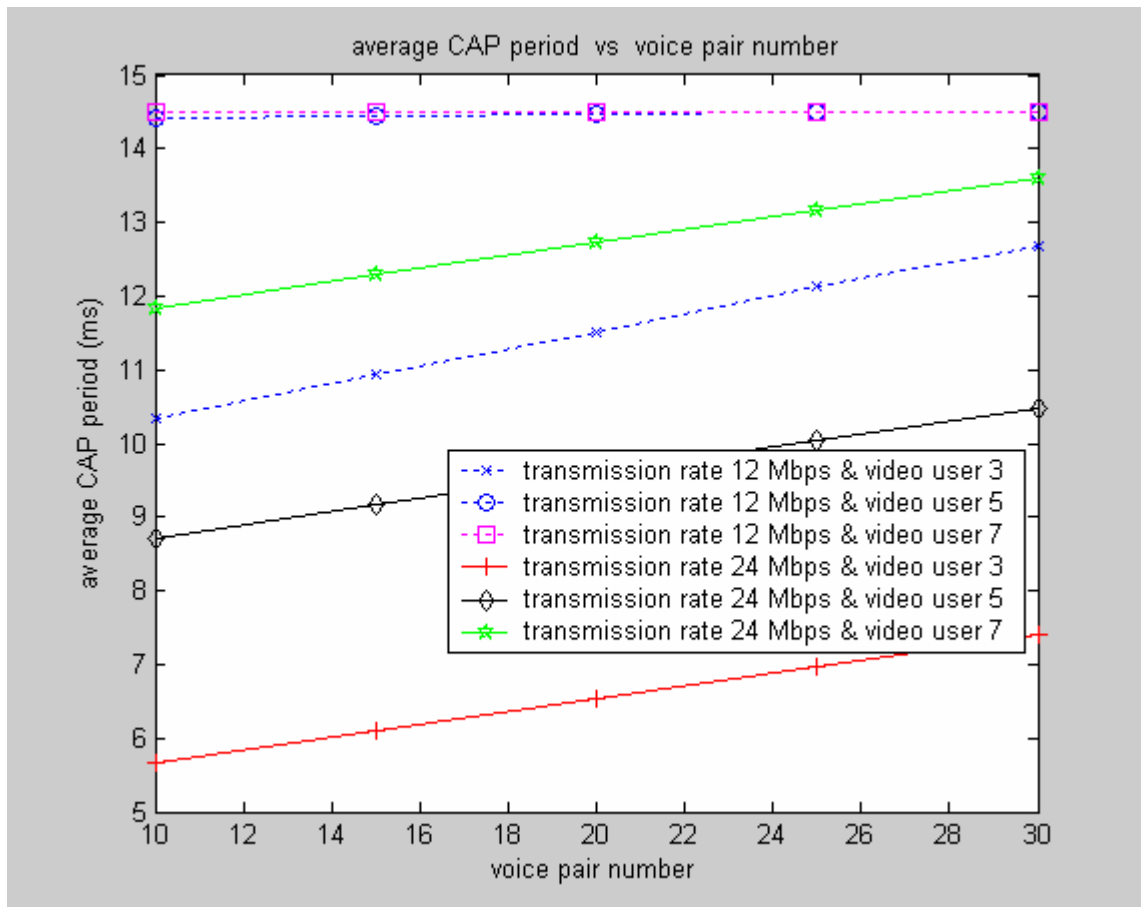
Figure 2.5: Average CAP period versus number of voice pairs
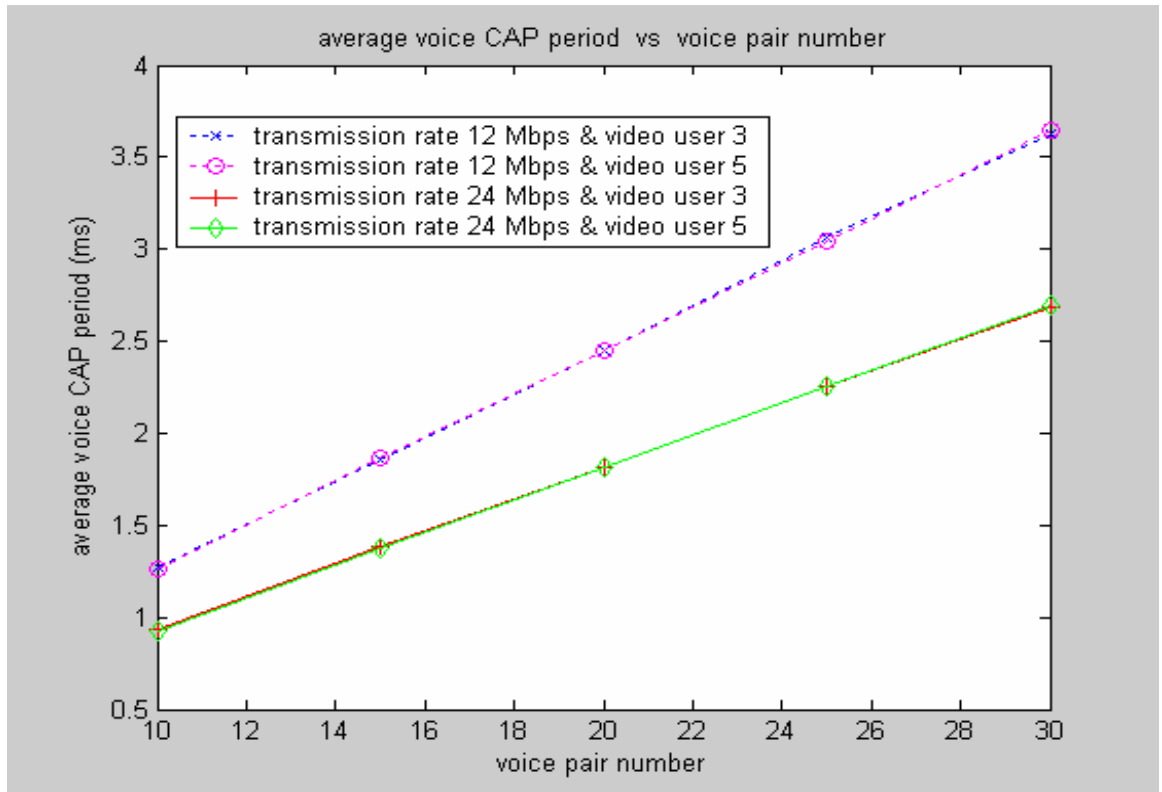
Figure 2.6: Average voice CAP period versus number of voice pairs
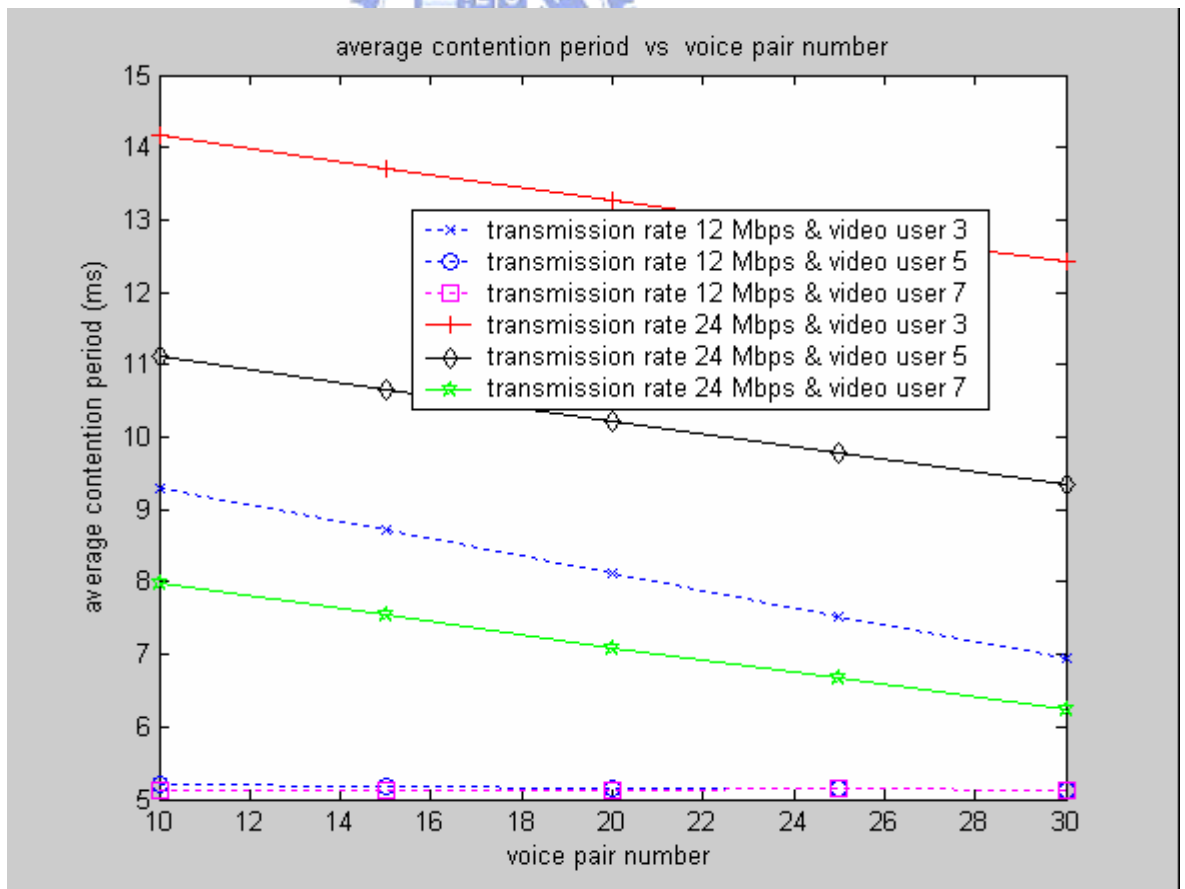
Figure 2.7: Average CP versus number of voice pairs
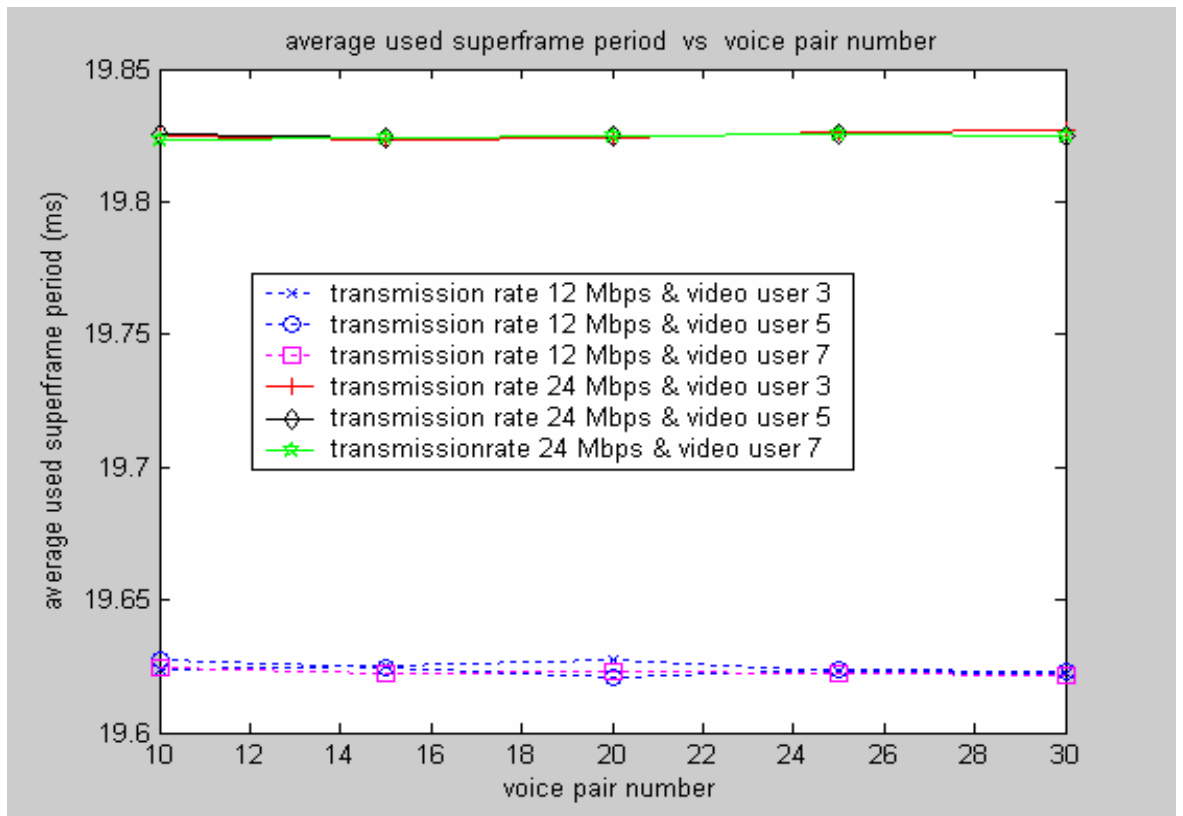
Figure 2.8: Average used superframe period versus voice pair number
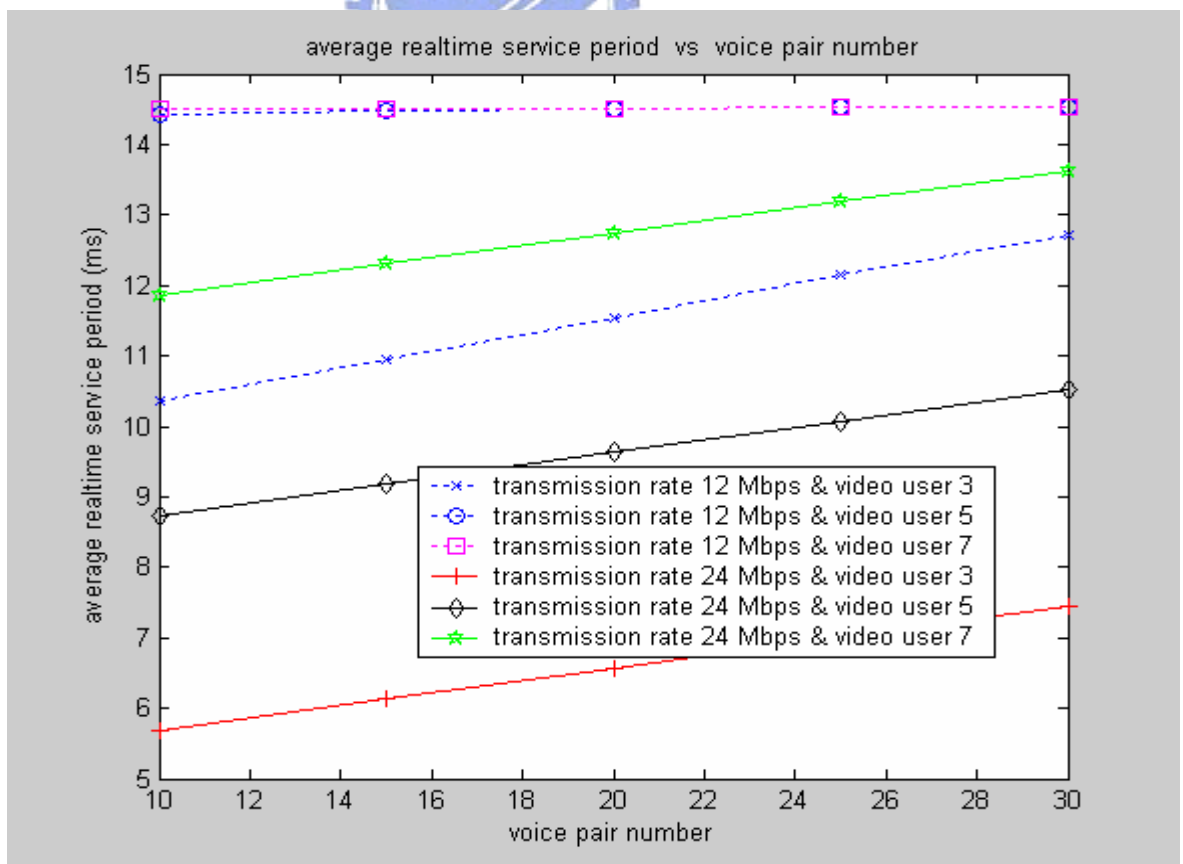
Figure 2.9: Average real-time service period versus voice pair number

18

Figure 2.10: Average voice contending period versus number of voice pairs

Figure 2.11: Average frame drop rate versus number of video users

Figure 2.12: Average inter-arrival period of video frame versus number of video users
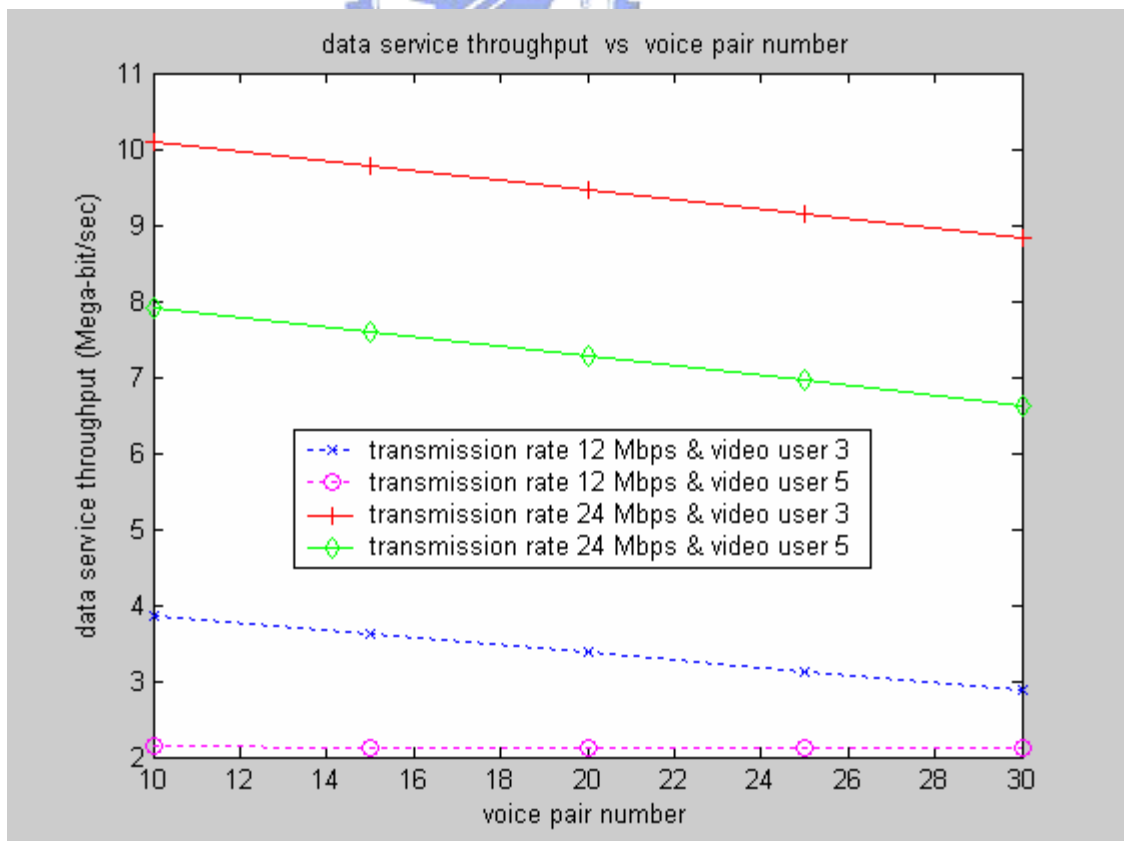
Figure 2.13: Average data service throughput versus number of voice pairs

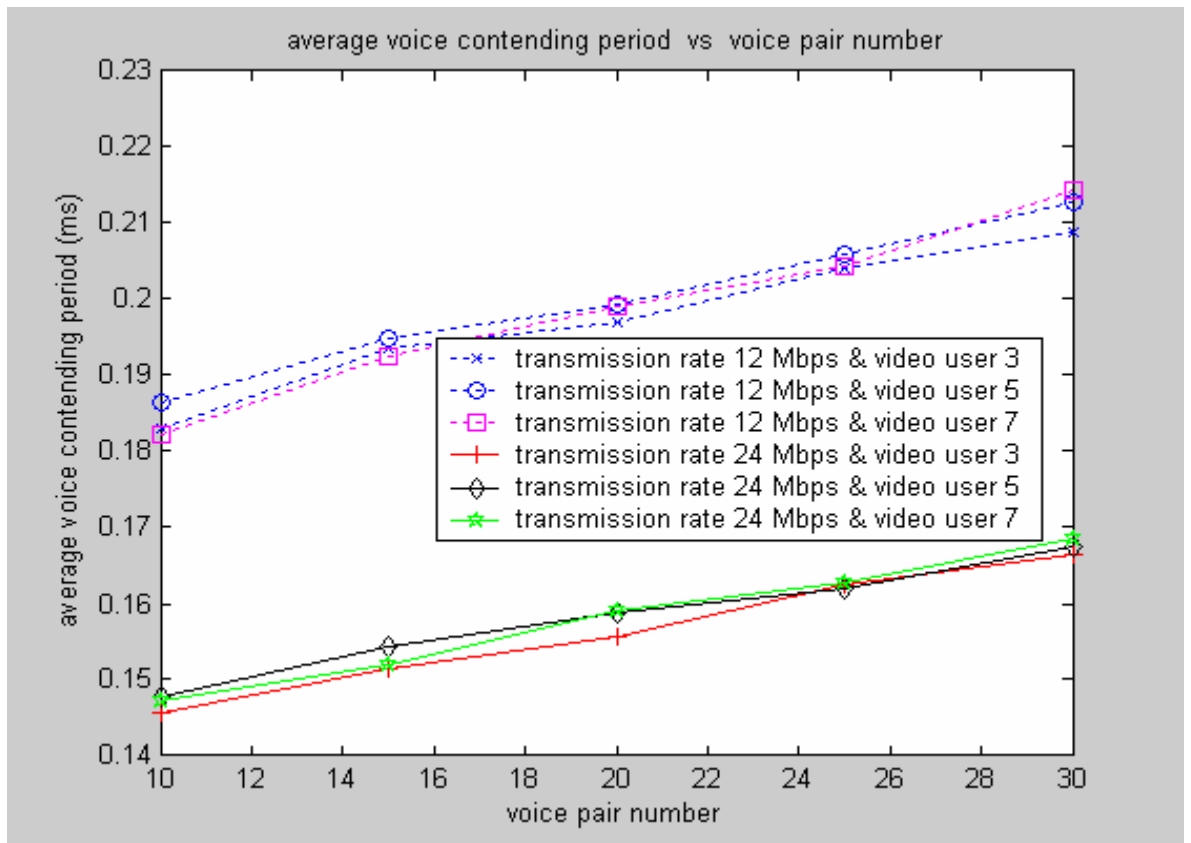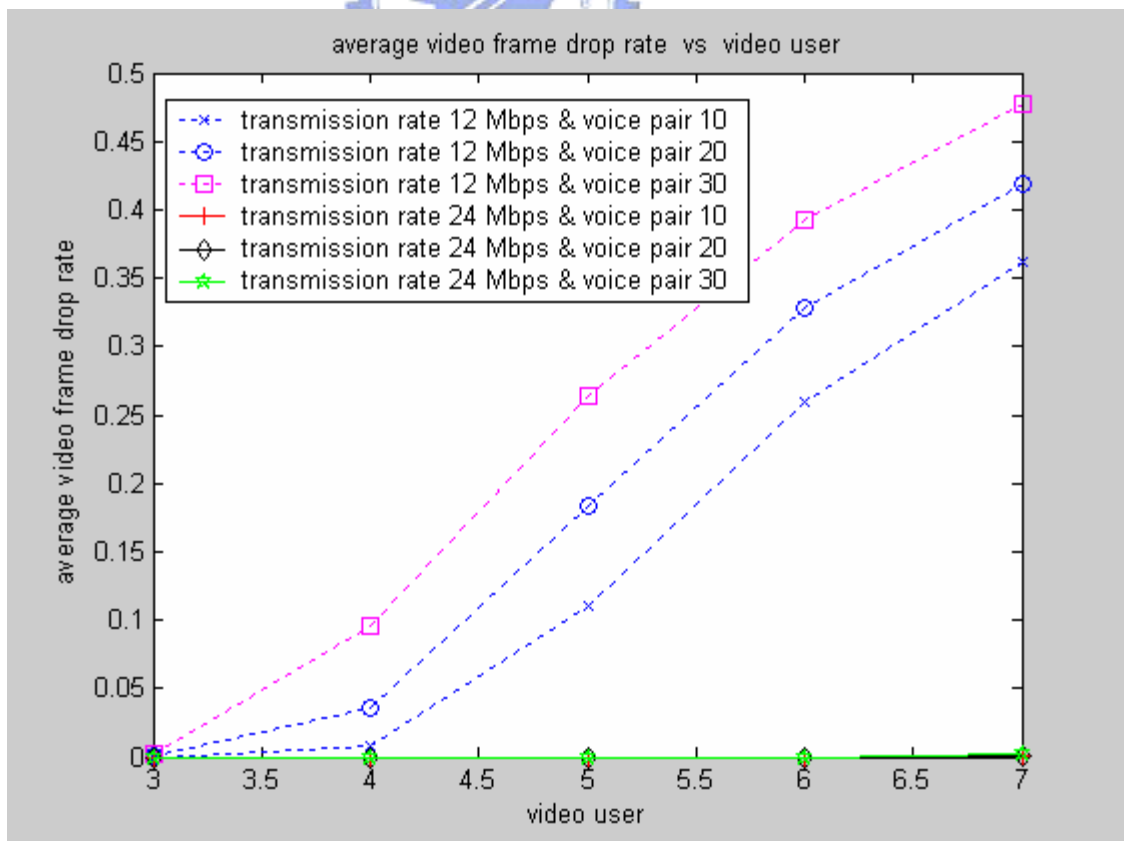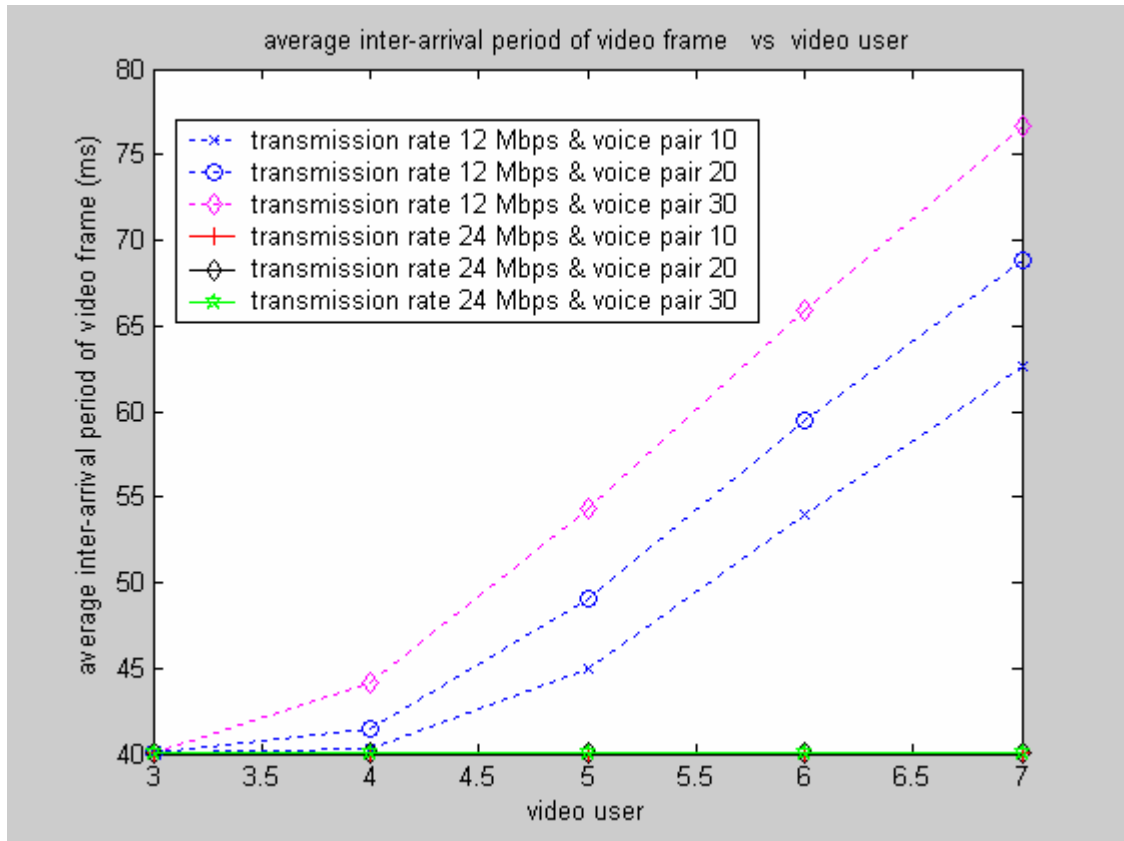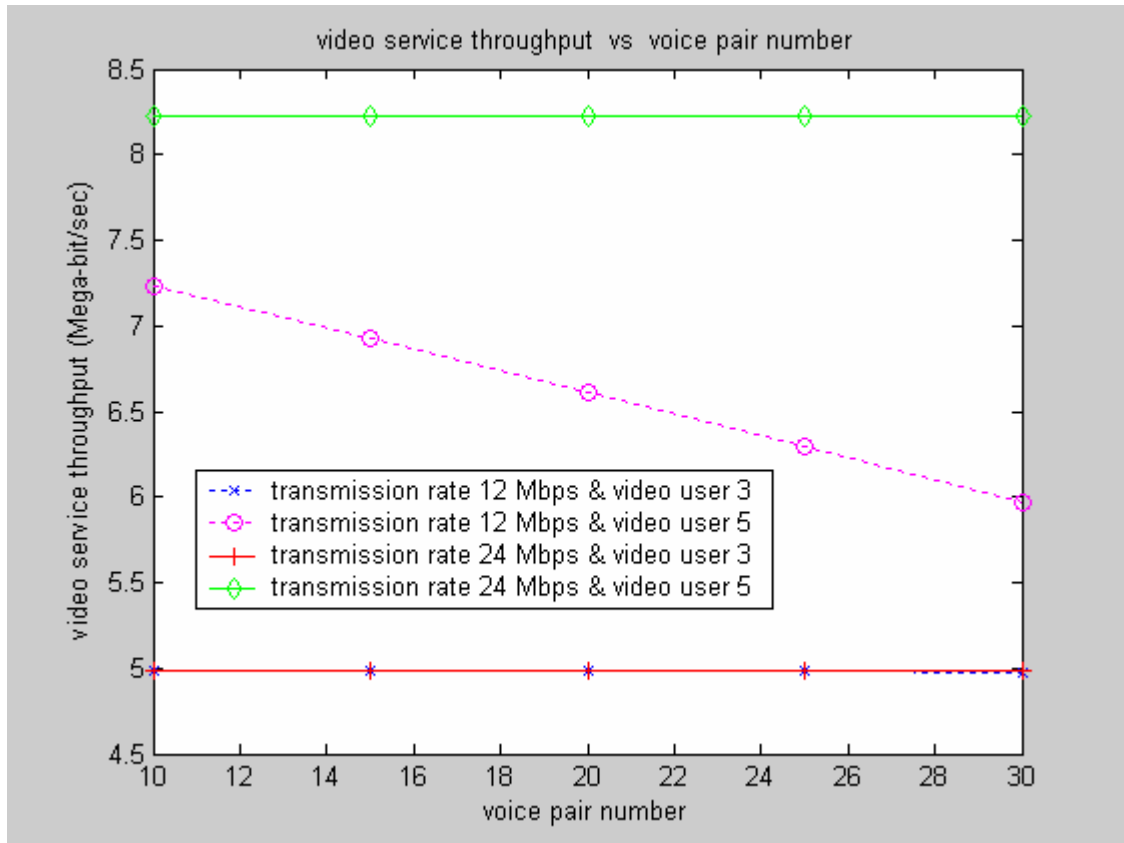Figure 2.14: Video service throughput versus number of voice pairs

Chapter 3

Fuzzy CAC for WLAN Gateway in

SOHO/Home Networks

3.1   Introduction

Wireless Local Area Networks (WLANs) are becoming popular both in enterprise and the residential environment. The main attractions of WLANs include cost effectiveness, ease of installation, flexibility and high bandwidth capacity. With 802.11e, to new enhancements of the 802.11 MAC protocol is designed to provide quality of service (QoS) for time-sensitive applications, such as voice and video which are typical applications in Small-Office/Home-Office (SOHO) and home networks. A challenge issue in future wireless networks are how to guarantee quality QoS requirements over the lifetime of the connections, and at the same time to make efficient use of the network resources. The achievement of guaranteeing QoS requirements depends on the resource in the system being able to load the required bandwidth of various QoS requirements, and a new connection accepted for efficient utilization sometimes worsens the QoS of other service connections. However, connection admission control (CAC) is one method to estimate the system resource whether supporting the required bandwidth while a new service connection is requested. In particular, the admission control is necessary when there are various types of traffic with various QoS requirements and when the system operates in the vicinity of its full capacity. Ganz and Wongthavarawat [17] propose a scheme for fast association suitable for handoff of multimedia applications.

In this paper, we propose a fuzzy CAC algorithm for 802.11e in SOHO/home networks. Our discussion on the 11e will be based on the version of draft. The most likely changes in future versions will be discussed as well.

3.2   System Model

3.2.1 System Architecture

For an indoor SOHO/Home WLAN environment shown as Figure 3.1, there is a QoS

access point (QAP) having only two interfaces, an Ethernet/ADSL interface and an wireless interface. The Ethernet/ADSL link supports the data traffic and the real time service traffic sent from or to the Internet, and the wireless interface is used to transmission of the packets in the wireless 802.11. In this paper, there are several quality of service stations (QSTAs) for which we intend to provide the quality of service (QoS) facility. In the infrastructure environment of Small Office, Home Office (SOHO), we assume that all voice, video and data packets are sent through the QAP even though in the case, when both QSTAs are within in the coverage area of an QAP, voice, video and data packets could not be sent directly between the QSTAs. There is a fuzzy controller within the QAP to control the new service call whether be accepted or rejected. Here the services we consider are duplex video conferencing, downlink video on demand, duplex voice conversation and duplex data services.
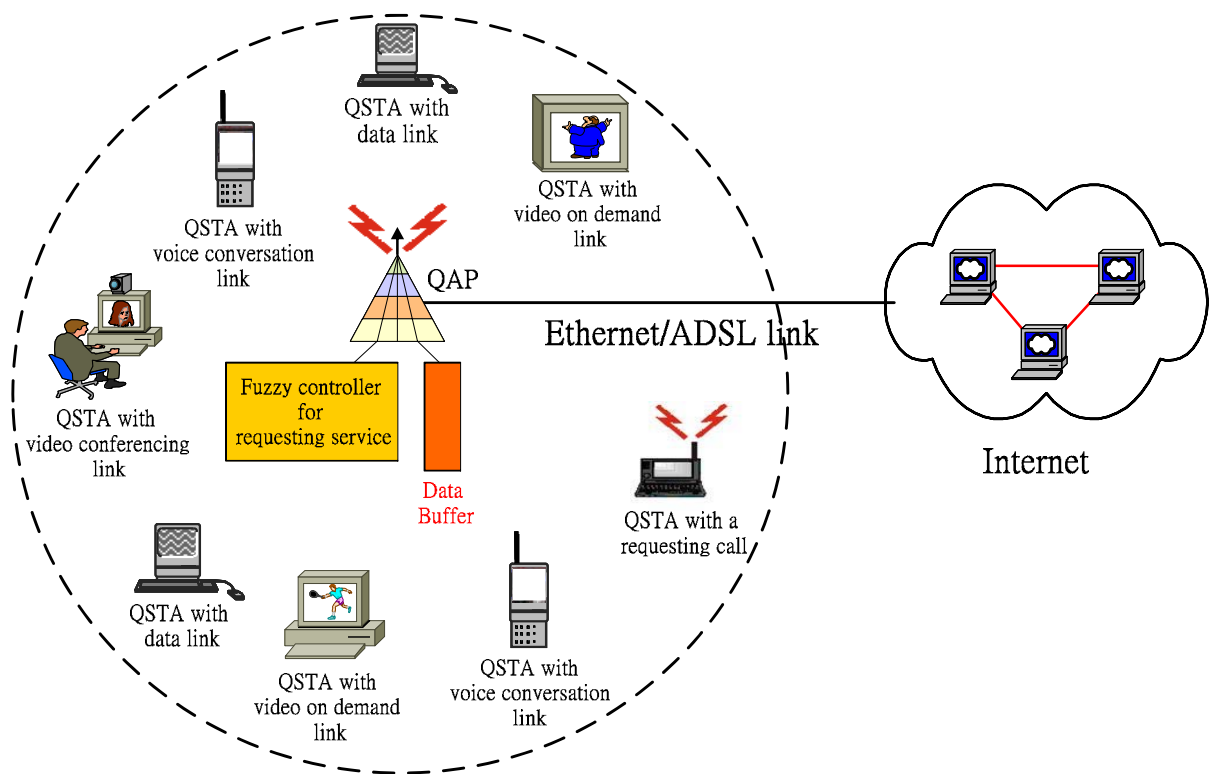
Figure 3.1: An infrastructure of SOHO/Home networks

3.2.2 Related Work on WLAN Operations

In 802.11e draft standard, admission control is a part of QoS, especially for shared wireless media, and because HC is the central controller for polling-based controlled channel

access, the admission of polling-base service is also straightforwardly controlled by HC. The QSTAs need to setup the traffic stream before it can start using the controlled channel access. When a new traffic stream is intended to be served, the QSTA needs to submit an ADDTS request QoS Action frame to the HC and informs the HC of specific QoS requirements for polling transmission with a detailed Traffic Specification (TSPEC) of the traffic stream in the request frame, containing parameters such as mean rate, peak rate, maximum burst size, delay bound, etc. On the receipt of an ADDTS request QoS Action frame, the HC can then make a decision as to whether accept or deny contained in the ADDTS response QoS Action frame which is replied to the request of the QSTA by the HC. At the end of the traffic stream, the QSTA could send a DELTS QoS Action frame to the HC and inform HC as the end. Here we assume that a QSTA can only have a service each time and there are no internal frame collisions between data frames and ADDTS request QoS Action frame at the QSTA.
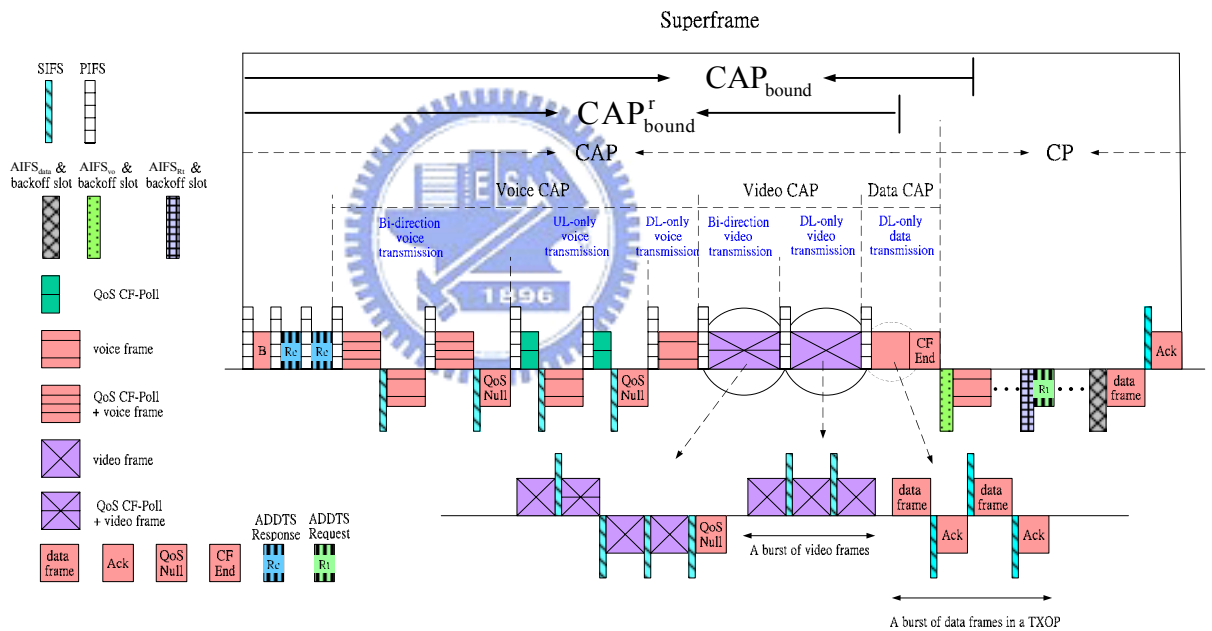
Figure 3.2: An example of a timing arrangement during a superframe

Figure 3.2 shows the timing arrangement of wireless medium in a superframe. In a superframe, it can be divided into two periods: controlled access phase (CAP) and contention period (CP). There is a maximum transmission period for CAP, and we denote it as CAP_{bound}. There is another maximum real-time transmission period for real-time CAP, and we denote it as CAP_{bound}^{r}. In the CAP, after the response management frame transmissions, the QAP polls the QSTAs to get the uplink voice frames and video packets and the QAP could transmit the

downlink voice frames, video packets directly without acknowledge and data frames with acknowledge based on HCF controlled channel access. In the CP, the non-real time data frames, intended contending voice frames of QSTAs and new service request management frames would use different priorities to contend the channel based on EDCF.

The CAP starts with a beacon generated by HC after the channel is sensed to be idle for PIFS interval. After the beacon transmission, the HC sends the ADDTS response QoS Action frame after PIFS idle interval to the QSTA which requested the new service during previous superframe. If the QSTA which sent the request frame in a contention-based way during previous superframe does not receive the according response frame during the following CAP, the request frame is seen as a collision of transmission and the QSTA continues to request the service by contending in this superframe. After the response transmission, the rest of a CAP is sequentially divided into three transmission phases: Voice CAP, Video CAP, and Data CAP. In addition, the voice service within Voice CAP also can be sequentially served into three transmission periods: bi-direction voice transmission, uplink-only (UL-only) voice transmission, and downlink-only (DL-only) voice transmission. After accepting a voice request service, the QSTA with the voice communication service is added in the voice polling list at QAP.

In Voice CAP, each time when HC starts transmission of a QSTA, HC should wait for PIFS idle interval. After accepting a new voice call, the QAP would add the QSTA in polling list, and at the beginning of the Voice CAP, the QAP decides the polling sequence based on the polling list. The first voice service of polling sequence is bi-direction voice transmission. If the QAP has a downlink voice frame queued for transmission to the QSTA in the polling list at the QAP, the downlink voice frame and QoS(+)CF-Poll can be combined and transmitted as a single frame by the QAP after the channel is sensed to be idle for a PIFS interval. After receiving the frame, the destination of the QSTA can transmit an uplink voice frame to the QAP after a SIFS interval. This frame exchange procedure is called bi-direction voice transmission. After bi-direction voice transmission, for the QSTA in polling list but the QAP does not have a voice frame queued for it, the QAP sends a sole QoS(+)CF-Poll frame to the QSTA after the channel is sensed to be idle for a PIFS interval, and waits to receive an uplink voice frame from the QSTA after a SIFS interval. This frame exchange procedure is called uplink-only (UL-only) voice transmission. If the QAP receives a QoS Null frame from the QSTA during bi-direction voice transmission or UL-only voice transmission, the QAP regards the QSTA entering into silence period and removes the QSTA from the polling list in

the QAP. After UL-only voice transmission, the QAP consecutively sends the remaining downlink voice frames with no acknowledge after waiting a PIFS channel idle interval during downlink-only (DL-only) voice transmission. The transmission of Voice CAP would end until all the voice packets for QSTAs are finished and all the QSTA in the voice polling list are polled, and the QAP stars the Video CAP for video packet transmissions.

The Video CAP also can be divided into two transmission periods: Bi-direction video transmission, and downlink-only (DL-only) video transmission. Bi-direction video transmission is used for active duplex service such as video conferencing, and DL-only video transmission is used for simplex service such as video on-demand. After accepting the duplex video service call, the QSTA with video conferencing is added in the video polling list at QAP. We assume that the interval of two superframes is equal to a video frame interval and the video packets arriving/generated during two superframes of a video frame interval are lasting availably at QAP/QSTA for the following three superframes or until they are transmitted completely. If these video packets delay exceeding the following next third superframe, these packets are dropped by QAP or QSTAs. For the fairness quality of the video services, the QAP schedules the transmission orders with a cyclic manner (that is so called Round Robin) for individual video services such as video on-demand and video conferencing services.

In Bi-direction video transmission, during two superframes of a video frame interval, the QAP receives several video packets from Ethernet/ADSL link belong to the QSTA in the video polling list, and the QAP may transmit these video packets during the following three Video CAPs after the channel is sensed to be idle for a PIFS interval in a burst separated by SIFS interval without acknowledge. And the last video packet of the burst can be combined with QoS(+)CF-Poll as a single packet to poll the receiving QSTA by the QAP. When the downlink packets of a video frame interval could not be transmitted totally within the limits of CAP_{bound}^{r}, the QAP also transmits the remaining packets of the video frame interval at first after PIFS idle interval in the next Video CAP, and also polls the receiving QSTA. After the QSTA receives the QoS CF-Poll + video packet, the QSTA can uplink a burst of video packets generated during two superframes of a video frame interval, and the transmissions of video packets are separate by SIFS interval without acknowledge. During the transmission of uplink video packets, the QSTA should add the transmission of a QoS Null frame at last after SIFS interval to indicate the end for finish or the unfinished transmission of the burst to the QAP due to insufficient period within the limits of CAP_{bound}. If the QAP receives the QoS

Null frame with the reason of unfinished transmission, the QAP will send the QoS CF-Poll to poll again the QSTA at first in the Video CAP of the next superframe and the polled QSTA could only transmit the remaining packets of the video frame interval. Each time the transmission volume of video packets to QSTA/QAP are video packets generated/arriving during a video frame interval or the remaining packets of a video frame interval. If the remaining packets of a video frame interval are not totally transmitted at the next third superframe, these packets would be dropped and the packets of the next video frame interval belong to the same QSTA would be transmitted instead. After the duplex transmission for a QSTA, the QAP continues to transmit another video burst of the next QSTA in a cyclic manner after a PIFS channel idle interval (that is so called Round Robin). The transmission of Bi-direction video transmission would end until all the previous available video packets queued for QSTAs in the video polling list at QAP are transmitted and the HC starts the DL-only video transmission.

In DL-only video transmission, during two superframes of a video frame interval, the QAP receives several video packets from Ethernet/ADSL link belong to the QSTA which is accepted for downlink video service, and the QAP may transmit these video packets during the following three Video CAPs after the channel is sensed to be idle for a PIFS interval in a burst without acknowledge which there is a SIFS idle interval between the two successive video packet transmissions. When the downlink packets of a video frame interval could not be transmitted totally within the limits of CAP_{bound}^r, the QAP could only transmit the remaining packets of the video frame interval at first after PIFS idle interval in the DL-only transmission of the next Video CAP. Each time the transmission volume of video packets to QSTA are video packets arriving at the QAP during a video frame interval or the remaining packets of a video frame interval. If the remaining packets of a video frame interval are not totally transmitted at the next third superframe, these packets would be dropped and the packets of the next video frame interval belong to the same QSTA would be transmitted instead. After the QAP transmits a burst of video packets to a QSTA, the QAP continues to transmit another video burst of the next QSTA in a cyclic manner after a PIFS channel idle interval (that is so called Round Robin). The transmission of a video CAP would end until all the previous available video packets at QAP are transmitted or in the limits of the CAP_{bound}^r, and the HC starts the Data CAP for data downlink transmission.

In Data CAP, HC transmits the downlink data frames in first in first out (FIFO) with

acknowledgement after PIFS idle interval in the limit of the $\text{CAP}_{\text{bound}}$. If the consecutive frames are queued for the same QSTA, the HC sends these data frames in a burst which the transmission of data frames are separated by SIFS interval. Each time when a QSTA receives a data frame, it can response an Ack frame after SIFS interval. At the last of CAP, the QAP sends a CF-end as termination of the CAP and indicating the start of the CP.

In the CP, the voice frames for uplink but not polled during the voice CAP in the same superframe will contend with uplink data frames of QSTAs and ADDTS request QoS Action frames from QSTAs asking new services. The QAP should assign different AIFS, CW_{min}, and CW_{max} for the contending uplink voice service and data service. After the QAP receives an uplink contending successful voice frame from the QSTA, the QAP re-adds the QSTA in the voice polling list. In CP, each uplink data frame of QSTA requires to contend the channel, that is when the QSTA contends the channel, the QAP sends only a data frame. When the counter of a data frame contending for transmission counts down to zero and there is enough time for a complete data frame exchange in CP, the QSTA would transmit the data frame. Also, it happens to a data frame having zero-counter in the beginning of the CP because there is not enough access time for the frame exchange in the last CP. In the beginning of the CP, if a data frame is first transmitted, the voice service would delay for at least a data frame transmission period in spite of the collision happened or not. In order to guarantee the voice delay as small as possible, the AIFS for voice service is set to equal to PIFS, denoted as $\text{AIFS}_{\text{vo}} = \text{PIFS}$. And the CW_{min} of a voice frame, denote as $\text{CW}_{\text{min\_vo}}$, and CW_{max} of a voice packet, as $\text{CW}_{\text{max\_vo}}$ are given by

$$\text{CW}_{\text{min\_vo}} = \text{CW}_{\text{max\_vo}} = \begin{cases} \lceil E[N_C] \rceil & \text{if } N_{NP} \leq 1 \\ \max\left\{ 2, \lceil E[N_C] \rceil \right\} & \text{if } N_{NP} \geq 2 \end{cases} \tag{3.1}$$

where N_C is the number of voice packets contending in a superframe, N_{NP}^{voice} is the number of QSTA with voice conversation not in the polling list, and $E[N_C]$ is the average contending voice packet number of the superframe which may be zero if N_{NP}^{voice} is equal to zero. Also the $\text{CW}_{\text{min\_vo}} / \text{CW}_{\text{max\_vo}}$ is the smallest integer greater or equal to the mean value of N_C. Here, we guarantee that the contending uplink voice service has a higher priority than non-real-time data service. That is the contending voice frames are always served before the

data frames. In a CP, when a data frame of a QSTA is served in the CP, it stands for contending voice frames has been completely served and hereafter only the data frames contend with each other in the CP. In order to avoid the uplink contending voice frame to be terminated by non-real-time data frames, the AIFS of non-real-time data service, denote as $AIFS_{data}$ is given by

$$AIFS_{data} = \begin{cases} PIFS + SlotTime * (CW_{max\_vo} + 1) & \text{for the first frame transmitted of a QSTA in a CP} \\ DIFS & \text{others} \end{cases} \quad (3.2)$$

where the "+1" is for contending uplink voice frames avoiding collision with a data frame having zero-counter. Each initial transmission of a data frame, the CW of the data frame is equal to CW_{min_data} if the CW reaches the maximum due to exponential increase from collisions, it would not change until the data frame is successfully transmitted. For the ADDTS request QoS Action frames, the AIFS for the request frames is set to equal to DIFS, denoted as $AIFS_{Rt} = DIFS$, and the CW_{min} and CW_{max} of the request frame are as the same as those of a data frame.
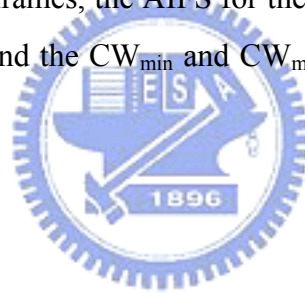
3.2.3 Scheduling for Weighted QoS in Video Services

There are three phase of real-time transmission during the CAP period, denoted as Voice CAP, Video DL-only and Video Bi-direction transmission. we take Round Robin scheduling for each phase for fairness quality of users in the corresponding service. For the differential drop rates of video services, the weighting of priority scheduling times is defined as follows:

$$\frac{t_{vi}^1}{t_{vi}^2} = \frac{n_{vi}^1 \cdot D_{bound}^{vi-1}}{n_{vi}^2 \cdot D_{bound}^{vi-2}} \quad (3.3)$$

where n_{vi}^1 is the number of video one-way transmission in the system. And n_{vi}^2 is the number of video duplex transmission in the system. t_{vi}^1 and t_{vi}^2 are the weights of the priority scheduling during a mean time of video service. D_{bound}^{vi-1} and D_{bound}^{vi-2} are the

delay bound for simplex and duplex type services, respectively. Through the weighted Round Robin scheduling, the video service could achieve the maximum efficiency while under the acceptable delay bounds. And each user in the corresponding type is also fairly sharing the resources. Also if the drop bounds are the same, and weighted times are just depending on the user number of each type.

3.2.4 Traffic Source

The traffic sources we consider in our system include video on-demand, video conferencing, voice conversation, and asymmetric data transmissions. Each traffic source includes the call arrival process, sojourn time distribution, and traffic model. And call arrival process and sojourn time distribution of each type traffic source are Poisson process and exponential distribution, respectively. The mean time of sojourn time for video on-demand, video conferencing, voice conversation and asymmetric data transmissions are 8, 3, 3, 10 (minutes), respectively. The rates of Poisson process for video on-demand, video conferencing, voice conversation, and asymmetric data call are λ, λ, λ and 1/30 (1/sec), respectively. And each traffic model is described as follows:

Voice traffic: As shown in Figure 3.3, the conversational dynamics of a communication pair is modeled as a four-state model, and the random process of conversational pairs are assumed to be independent and identically distributed. The four-state is described as follows that uplink and downlink are both talking, called Double Talk; Uplink and downlink are both silent, called Mutual Silence; Uplink is talking and downlink is silent, defined as Talkspurt; Uplink is silent and downlink is talking, defined as Pause. In the model, the sojourn time of each state in which a conversation pair state is assumed to be exponentially distributed. The state transition rates, λ_{ij}, can be obtained by fitting the mean duration of talkspurt, pause, double talk, and mutual silence, given in [11]. Within each talking spurt interval for each member of a conversational pair, 8 kb/s pulse code modulated (PCM) digital voice is assumed and the mean rate of a communication pair given parameters in Figure 3.3 is 6.8 kb/s. During a talking spurt interval, a voice frame is coded every superframe interval and the coder generates some voice frames which are all the same size depending on the coding rate (8 kb/s), and superframe interval. We consider the superframe interval as the maximum acceptable value for one-way voice transfer delay over the WLAN delay. When a voice frame is not transmitted successfully in this superframe, it is dropped by the QAP or QSTA.
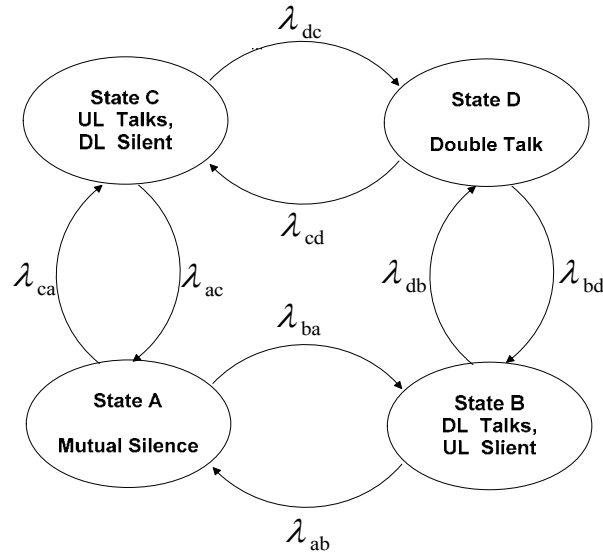
The state transition parameters (unit: 1/sec)

λ_{ab}	λ_{ba}	λ_{bd}	λ_{db}
λ_{ac}	λ_{ca}	λ_{cd}	λ_{dc}
1.63	0.846	0.279	2.24

Figure 3.3: The four-state Markov model of a conversation pair

Video traffic: Here we describe the one way simplex video traffic model, and we model the duplex video transmission with two independent simplex video traffic streams. In the article, the byte generation process for a video coder is assumed to have two motion states: one is the low motion state for interframe coding rate, and the other is the high motion state for intraframe coding rate [12]. The rate of intraframe coding is further divided into two parts: the first part has the same rate as the interframe coding; the second part, called difference coding, is the difference rate between intraframe coding and interframe coding. The interframe coding and the difference coding are both modeled as discrete-state Markov-modulated poisson process (MMPP) with basic rates A_r and A_a [13]. The state-transition diagram is shown in Fig. 3.4(a) and 3.4(b). The video source will alternate between interframe and intraframe, depending on the video source activity factor. As shown in Fig. 3.4(c), there is a transition rate c in the interframe state and a transition rate d in the intraframe state. Let $\lambda_a(t), \lambda_r(t)$, and $\lambda'_a(t)$ denote the byte generation rates for intraframe coding, interframe coding, and difference coding at the time t respectively from the video coder. Clearly, $\lambda_a(t) = \lambda_r(t) + \lambda'_a(t)$ The process of $\lambda_r(t)$ is an

31

(M_r+1)-state MMPP process. The state-transition diagram for $\lambda_r(t)$ uses the label $m_r A_r$ to indicate the byte generation rate of interframe coding of a state and use the labels $(M_r - m_r)\gamma$ and $m_r \omega$ to denote the transition rate from state $m_r A_r$ to state $(m_r + 1)A_r$, and from state $m_r A_r$ to state $(m_r - 1)A_r$, respectively. Similarly, the process for $\lambda_a'(t)$ is an $(M_a + 1)$-state MMPP process. The state-transition diagram for $\lambda_a'(t)$ uses the label $m_a A_a$ to indicate the additional frame generation rate of a state due to intraframe coding and uses the labels $(M_a - m_a)\phi$ and $m_a \varphi$ to denote the transition rate from state $m_a A_a$ to state $(m_a + 1)A_a$ and from state $m_a A_a$ to state $(m_a - 1)A_a$, respectively. One should be note that the long-term correlation behavior of a video source results from the process $\lambda_a(t)$.

The values of γ, ω, M_r, A_r, ϕ, φ, M_a, A_a, c and d can be obtained from the traffic variables of R_p, R_m and T_p [14, 15]. The parameters for video on-demand service are those R_p = 1.98 Mbps, R_m = 0.8 Mbps, and T_p = 0.5 second, which give M_r = M_a = 20, A_r = 6.07×10^4 bps, A_a = 1.43×10^4 bps, γ =1.33, ω = 2, ϕ = ψ = 1, c = 2, and d = 18.01 (1/sec). And the one way video traffic parameters of video conferencing service are those R_p = 318 kbps, R_m = 128 kbps, and T_p = 0.5 second, which give M_r = M_a = 20, A_r = 1.55×10^4 bps, A_a = 0.36×10^4 bps, γ =1.33, ω = 2, ϕ = ψ = 1, c = 2, and d = 18.01 (1/sec), and we model the duplex video transmission with two independent simplex video traffic streams.
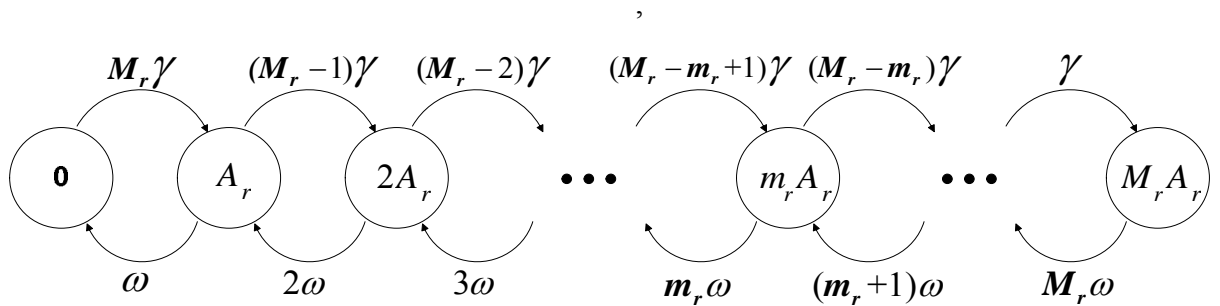

Figure 3.4(a): State transition diagram for interframe coding $\lambda_r(t)$
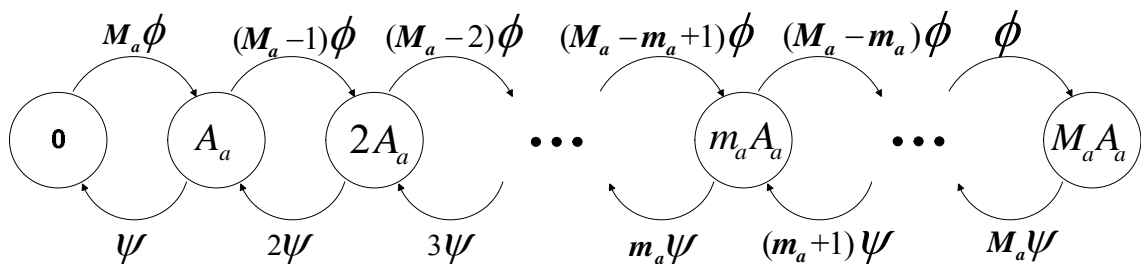
Figure 3.4(b): State transition diagram for difference coding $\lambda'_a(t)$
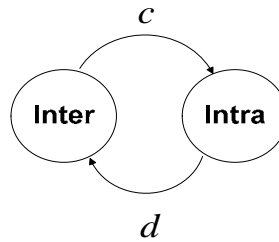

Figure 3.4(c): Interframe and intraframe alternate model for a video source

Data traffic: The data traffic model includes both uplink and downlink transmissions. An uplink/downlink ratio of 1/12 has been chosen for the data traffic in our environment in order to emulate the asymmetrical behavior of Web-browsing-like services, and the downlink data traffic with a mean rate, 12 kbps. And the data traffic is described as follows: The data traffic is self-similar and long-range dependence (LRD) with uniformly distributed packet size ranged from 64 to 1518 bytes. To generate self-similar, we used the method described in [18], where the resulting traffic is an aggregation of multiple streams. The structure of the synthetic self-similar traffic generator is shown in Figure 3.5 Each source is performed by ON/OFF Parato-distributed model. The design of the number of sources, K, in a generator is based on experiment result discussed in [19]. It shows that the bursts of the traffic (Hurst parameter) does not change with K if the total load is fixed. Each ON/OFF model allows the ON and OFF periods to have infinite variance (high variability or Noah Effect), and the superposition of many such sources produces aggregate traffic that exhibits long-range dependence (also called the Joseph Effect) [18].
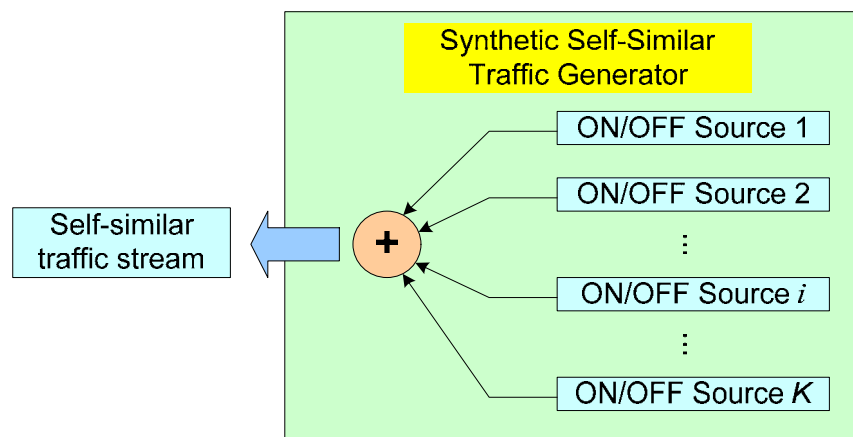
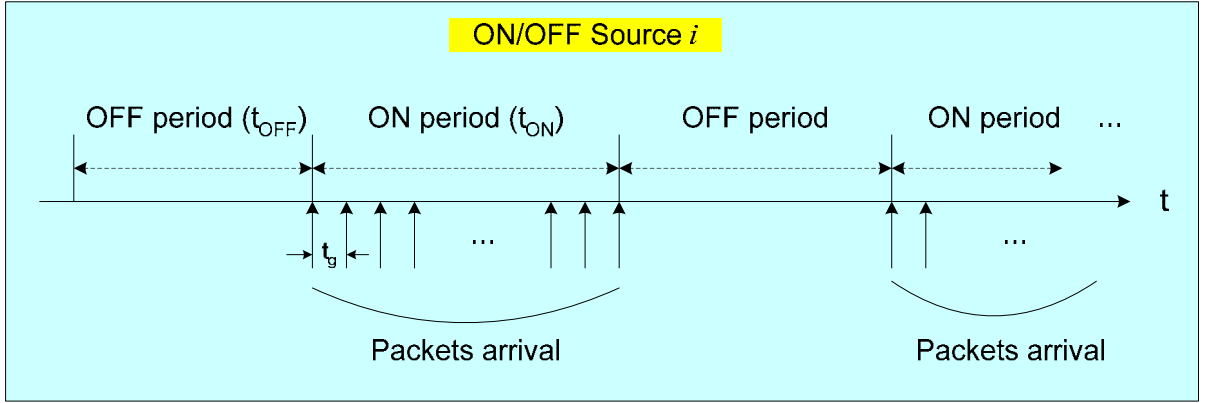Figure 3.5: Synthetic self-similar traffic generator

Figure 3.6: ON/OFF Parato-distributed source i

Now we discuss the detail of each ON-OFF source. Figure 3.6 shows the model of source i, and the parameters of this source are described as follows:

The number of packets generated by data source i, denoted as $N_{P(i)}$, during ON period follows Pareto distribution with a minimum of 1 and maximum of 2^{16}-1. Pareto distribution can be defined as follows:

$$\begin{cases} f_X(x) = \dfrac{\alpha k^{\alpha}}{x^{\alpha+1}}, x \geq k, \\ \mu = \dfrac{k\alpha}{\alpha-1}, \alpha > 1, \\ \sigma^2 = \dfrac{k^2\alpha}{(\alpha-2)\cdot(\alpha-1)^2}, \alpha > 2, \end{cases} \qquad (3.4)$$

where α is a shape parameter, and k is a location parameter. We set the shape parameter $\alpha = 1.4$. The choice of α was prompted by measurements on actual Ethernet traffic [20]. They reported the measured Hurst parameter of 0.8 for moderate network load. The relationship between the Hurst parameter and the shape parameter α is $H = (3-\alpha)/2$ [18]. Thus, $\alpha = 1.4$ should result in $H = 0.8$.

During ON period, the packet assumed to immediately follow the previous packet with minimum inter-packet gap t_g. We choose t_g, which equals to the transmission period of the standard preamble (8 bytes) of Ethernet packet.

Every source has a constant packet size from uniform distribution between 64 and 1518 (in bytes). We denote the packet size generated by source i is $P_{S(i)}$. Then, the duration of

ON period (t_{ON}) of source i can be described as

$$t_{ON} = (P_{S(i)} + t_g) \times N_{P(i)} \times \text{bytetime} , \tag{3.5}$$

where bytetime depends on the transmission rate of Ethernet.

OFF periods (intervals between the packet trains) also follow the Parato distribution with the shape parameter $\alpha = 1.2$. We used heavier tail for the distribution of the OFF periods represent a stable state in a network, i.e., a network can be in OFF state (no packet transmission) for an unlimitedly long time, while the durations of the ON periods are ultimately limited by network resources and (necessarily finite) file sizes. By aggregating streams from K independent sources, the realistic self-similar traffic is generated. We use the traffic generator to emulate best-effort data traffic source in our simulation.

3.3 Fuzzy Call Admission Control

3.3.1 The Goal

We consider the call admission control for real-time services at the WLAN gateway in SOHO/Home networks. The objectives of the fuzzy call admission control (CAC) algorithm are to maximize the utilization of network resources and to minimize the block rate of new calls if the utilization reaches the maximum, while guaranteeing the QoS of existing calls. The QoS requirements are: drop rates of voice and video service, and loss rate of data service in the QAP for downlink which are denoted as D_{bound}^{vo}, D_{bound}^{vi}, and L_{bound}, respectively. The QAP should avoid the much data loss resulted from overflow from the wired link and guarantee the drop rate of real-time services due to exceeding the delay requirement which are the period of two superframes.

3.3.2 Fuzzy Call Admission Controller

In general, call admission control (CAC) can be defined as the procedure of deciding whether or not to accept a new call. One of the fundamental aspects of CAC is to evaluate the impact of a new connection on the current traffic load. This usually involves determining the resources (bandwidth) needed for a new call with its specific QoS and its impact on the existing calls. The block diagram of fuzzy call admission control is shown as the figure 3.7.

Here we use the equivalent CAP estimator to evaluate the estimated CAP, denoted as CAP_e of a new real-time service call according to mean rate, peak rate of the new call traffic characteristics and its corresponding QoS requirement which is $D_{\text{bound}}^{\text{vo}}$ or $D_{\text{bound}}^{\text{vi}}$ depending on the type of the new call. The estimation procedure of CAP_e from the CAP estimator is attached in the appendix. And, we take the average CAP of real-time service, denoted as CAP_{ave}^r, which is the average system loading of real-time service during the time period from the last accepted/rejected call, and is defined as follows:

$$CAP_{ave}^r = \frac{1}{w} \sum_{i=1}^{w} CAP_r(i) \tag{3.6}$$

where w is the period interval in unit of superframe from the last accepted or left call, and $CAP_r(i)$ is the i_{th} real-time CAP of the w CAPs. There are three input linguistic variables of the fuzzy call admission controller which are Data loss index, CAP index, and Real-time drop index denoted as D_{loss}, CAP, R_{drop}, respectively. We describe these variables as follows:

(1) Data loss index: (D_{loss})

D_{loss} describes the information on data packets for downlink which are dropped by the QAP due to overflow, and is measured by the QAP. D_{loss} is the ratio of total downlink data loss volume due to overflow to the total data arrival volume at the QAP.

(2) CAP index: (CAP)

CAP describes the information on the capability of the resource loading the new call for real-time services, and is defined as follows:

$$CAP = 1 - \frac{CAP_{ave}^r - \text{CAP}_e}{\text{CAP}_{\text{bound}}^{\text{r}}} \tag{3.7}$$

where $\text{CAP}_{\text{bound}}^{\text{r}}$ is the bound period of real-time CAP, CAP_e is the estimated CAP period for the real-time new call, and CAP_{ave}^r is the average time period for real-time transmission from the last arrival/left call.

(3) Real-time drop index: (R_{drop})

R_{drop} describes the information on drop rate of real-time services during the period from the last accepted or leaved call, and we define it as follows:

$$R_{drop} = R_{drop-vo} + R_{drop-vi}$$

$$= \sum_{i=1}^{n_{vo}} \frac{1}{w} \sum_{j=1}^{w} 2 \cdot \frac{d_{vo}(i,j)}{a_{vo}(i,j)} + \sum_{i=1}^{n_{vi}^1} \frac{1}{w} \sum_{j=1}^{w} \frac{d_{vi}^1(i,j)}{a_{vi}^1(i,j)} + \sum_{i=1}^{n_{vi}^2} \frac{1}{w} \sum_{j=1}^{w} 2 \cdot \frac{d_{vi}^2(i,j)}{a_{vi}^2(i,j)} \qquad (3.8)$$

where w is the period interval in unit of superframe from the last accepted or left call. And n_{vo}, n_{vi}^1 and n_{vi}^2 are the number of voice pairs, video on-demand user and video conferencing pairs still existed in the system, respectively. $R_{drop-vo}$ and $R_{drop-vi}$ are the average drop rate of the voice frames and video packets during the W superframes, respectively. d_{vo}/a_{vo} is the voice service drop/arrival volume of the i_{th} in n_{vo} users during the j_{th} of the w superframes. d_{vi}^1/a_{vi}^1 is the video downlink service drop/arrival volume of the i_{th} in n_{vi}^1 users during the j_{th} of the w superframes. d_{vi}^2/a_{vi}^2 is the video conferencing service drop/arrival volume of the i_{th} in n_{vi}^2 users during the j_{th} of the w superframes.
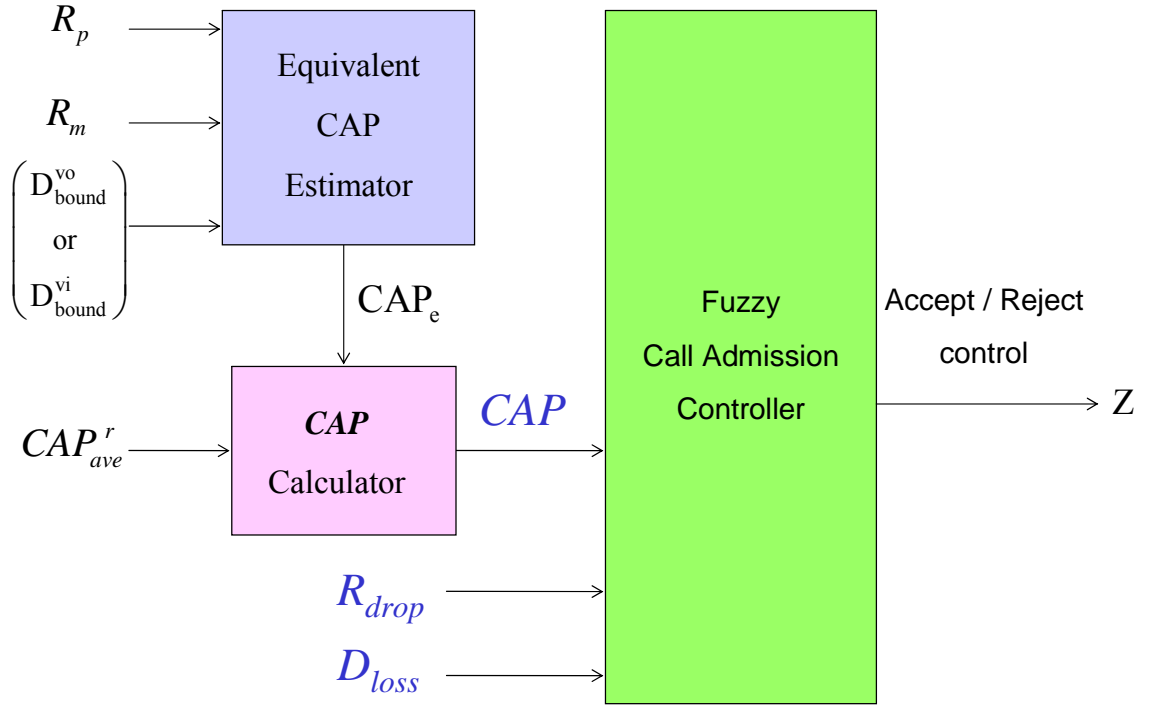
Fig. 3.7: Block diagram of the fuzzy call admission control

Each time when a new call request is coming, the CAC controller makes a decision as "Accept" or "Reject" according to the fuzzy rules. The term set for the Data loss index is defined as $T(D_{loss})=\{$ Small (S), Medium (M), Large (L) $\}$. The term set for the CAP is $T(CAP)=\{$Positive (P), Negative Small (NS), Negative Large (NL) $\}$. The term set for the Real-time drop index is $T(R_{drop})=\{$Small (S), Medium (M), Large(L) $\}$. In order to provide a soft admission decision, not only "Accept" and "Reject" but also "Weak Accept" and "Weak Reject" are employed to describe the accept/reject decision. Thus the term set of the output linguistic variable is defined as $T(z)=\{$ Accept (A), Weak Accept (WA), Weak Reject (WR), Reject (R) $\}$. The membership functions for $T(D_{loss})$ is shown in Fig. 3.8 and L_{bound} is the required data loss rate at the QAP. The membership function for $T(CAP)$ is shown in figure 3.9, and the membership function for $T(R_{drop})$ is as shown in figure 3.10, and D_{bound} is the system drop bound for real-time service and is defined as follows:

$$D_{bound} = \begin{cases} 2 \cdot n_{vo} \cdot D_{bound}^{vo} & \text{if } n_{vi}^1 \text{ and } n_{vi}^2 = 0 \\ (n_{vi}^1 + 2 \cdot n_{vi}^2) \cdot D_{bound}^{vi} & \text{if } n_{vi}^1 \text{ or } n_{vi}^2 \neq 0 \end{cases} \tag{3.9}$$

where we assume the QoS requirement of uplink is the same as the downlink. The membership function for $T(z)$ is shown in Fig. 3.11.

38

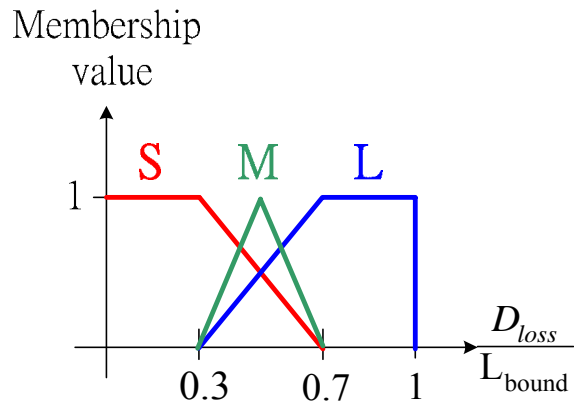Figure 3.8: The membership function of the term set for D_{loss}
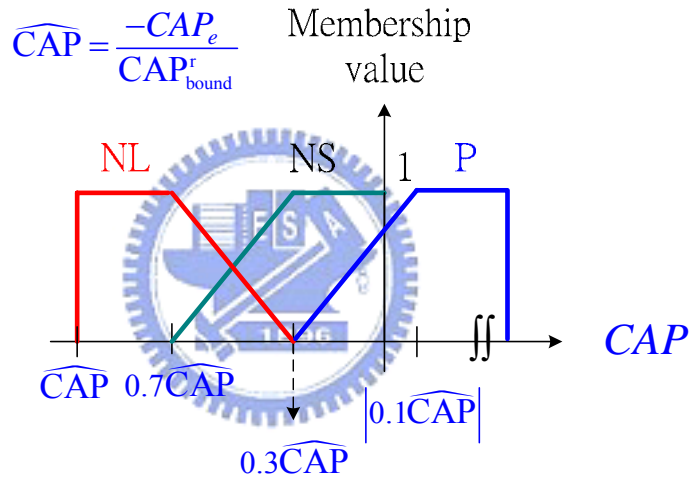

$$\widehat{CAP} = \frac{-CAP_e}{CAP_{bound}^r}$$

Figure 3.9: The membership function of the term set for *CAP*
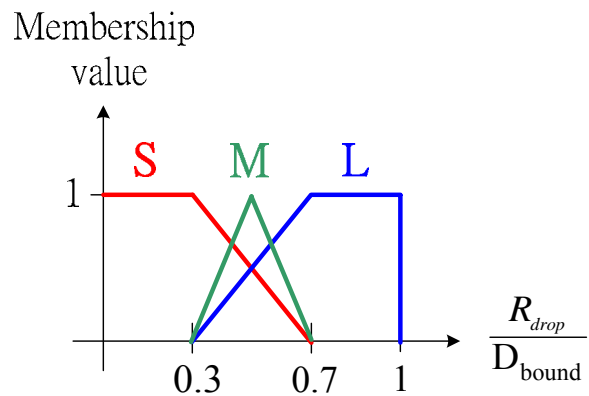

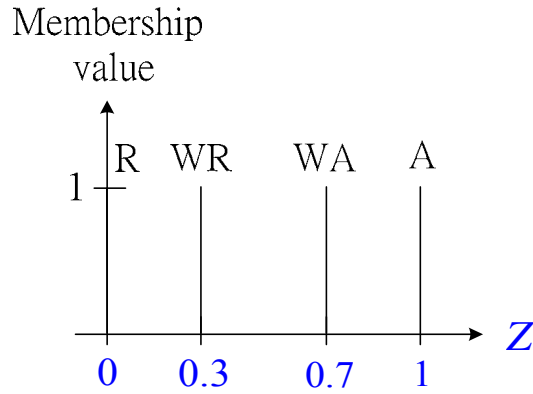Figure 3.10: The membership function of the term set for R_{drop}

Fig. 3.11: The membership function of the term set for Z

The fuzzy admission control rules are designed in Table 3.1.

In the rule 1 that D_{loss} is small, R_{drop} is small and *CAP* is positive, for increasing utilization of the goal, the output is "Accept". In the rule 7 that D_{loss} is small and R_{drop} is large, means the system could not guarantee the QoS of real-times if accepting the new call, so the output is "Reject". Also during rule 1 to rule 7, the output may be "WA"/"WR" in a sequence way, while the QoS get worse/better in comparison the QoS of "A"/"R". In the rule 10, 11 and 14, the R_{drop} is medium or large, and the output trends toward the "R". In the rule 12, 13 and 14, the D_{loss} is large, so the output is "R". Also we take a gradual change during rule 8 to 14, the output could be determined.

The fuzzy call admission controller uses the max-min inference method for the inference engine because it is designed for real-time operation [21], and also uses the center of area for defuzzification method [22]. A new call request can be accepted if the output of the fuzzy call admission controller, z, is greater than an acceptable threshold Z_a, which we set the value is 0.5, and be rejected if z is smaller than Z_a.

Rule	D_{loss}	R_{drop}	CAP	Z
1	S	S	P	A
2	S	S	NS	WA
3	S	S	NL	WR
4	S	M	P	WA
5	S	M	NS	WR
6	S	M	NL	R
7	S	L	P, NS, NL	R

8	M	S	P	A
9	M	S	NS, NL	WA
10	M	M	P, NS	WR
11	M	M	NL	R
12	L	S	P	WR
13	L	S	NS, NL	R
14	L	M, L	P, NS, NL	R

Table 3.1: The rule structure for the fuzzy admission controller

3.4 A Gaussian Approximation for Equivalent CAP Estimator

For the fairness of real-time transmission, we take Round Robin scheduling method. When the video packets arrives during the two successive superframes can not be transmitted during the following three superframes, these video are dropped by the QAP. We assume the CAP_{bound}^r could support the N users at a certain drop probability, and then we could formulate as follows:

$$p(\sum_{i=1}^{N} CAP_i(n) + \sum_{i=1}^{N} CAP_i(n\text{-}1) > 3\ CAP_{bound}^r) \le p_{drop} \qquad (3.10)$$

where p_{drop} is the bound we want to guarantee which is the drop bound of the real-time call. we assume the CAP_i during the two successive superframes are independent and identically distributions. Then (3.10) could be written to (3.11)

$$p(\sum_{i=1}^{2N} CAP_i(n) > 3\ CAP_{bound}^r) \le p_{drop} \qquad (3.11)$$

we take a Gaussian approximation method for the summation of $CAP_i(n)$ as (3.12) and the mean (μ) and variance (σ^2) of the new real-time CAP_i could be known by the QAP from the peak rate and mean rate of the new call. That is

set $\Omega = \sum\limits_{i=1}^{2N} CAP_i(n)$ is Gaussian distribution with $\begin{cases} \hat{\mu} = 2N\mu \\ \hat{\sigma} = \sqrt{2N\sigma^2} \end{cases}$

We use the approximation table look method, so the N and CAP_e could be derived.

$$p\left(\frac{(\Omega - \hat{\mu})}{\hat{\sigma}} > \frac{3\ CAP_{bound}^r - \hat{\mu}}{\hat{\sigma}} \right) \le p_{drop} \qquad (3.12)$$

$$\frac{3\ CAP_{bound} - \hat{\mu}}{\hat{\sigma}} \ge \beta \quad \to N \to CAP_e \qquad (3.13)$$

3.5 Concluding Remarks

In this chapter, for real-time service requesting calls, we proposed a fuzzy CAC algorithm at QAP for WLAN in SOHO/Home networks. The input linguistic variables we consider are used for checking the system load and whether satisfying the required QoS of the existed traffic in the system. We take a Gaussian approximation to the CAP estimator which is discussed in the 3.4. The QoS requirements are drop rates of real-time services, and loss rate of data downlink due to overflow at QAP. We have a procedure to design a fuzzy CAC algorithm for the wireless 802.11e MAC protocol and we consider the application real-time services including voice, video on-demand, and video conferencing.

Chapter 4
Conclusion

In this thesis, we first study the simulation phenomenon for 802.11e with hybrid services in chapter2. Three types of services are considered, ie. real-time voice, real-time video, and data services. For real-time service, the delay-considered drop is investigated and we find the data throughput is dynamical with the real-time service users. The difference priority in contention period guarantees the real-time service efficiently accessing the medium. We simulate the video traffic and to understand video throughput within the limit of the delay bound influenced by the voice pair number. We also modify the IFS of data which could enhance the medium efficiency under the hybrid contention traffics.

In chapter 3, we design a call admission control scheme in the SOHO/Home networks for WLAN gateway with fuzzy theorem. And we use a weighted Round Robin scheduling to achieve the different type video user with a ratio weighted delay-based drop quality. Through the weighted Round Robin scheduling, the video service could achieve the maximum efficiency while under the acceptable delay bounds. And each user in the corresponding type is also fairly sharing the resources. The scheduling can help our CAC guarantee the maximum throughput and avoid much impact on the QoS of the existed users if the new call were accepted. In the CAC, we provide a approximation for estimate the new call bandwidth under the delay-based drop requirement. We design a fuzzy algorithm to apply in the WLAN protocol with SOHO/Home networks which is more popular in the future.

Bibliography

[1] IEEE Standard for Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extention in the 2.4 GHz Band, IEEE Std 802.11 1999, Sep. 1999.

[2] H. S. Chhaya and S. Gupta, "Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol," *Wireless networks*, vol. 3, pp. 217-234, Aug. 1997.

[3] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Select Areas Commun.*, vol. 18, no. 3, pp. 535-546, March 2000.

[4] A. Kamerman and G. Aben, "Throughput performance of wireless LANs operating at 2.4 and 5 GHz," *Proc. PIMRC'2000,* pp. 190-194, Sep. 2000.

[5] D. J. Deng and R.S. Chang, "A priority scheme for IEEE 802.11 DCF access method , " *IEICE Trans. Commun.*, vol. E82-B, no. 1,pp. 96-102, Jan. 1999.

[6] B.P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, " *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 116-126, Sep. 1997

[7] T. Suzuki and S. Tasaka, "Performance evaluation of video transmission with the PCF of the IEEE 802.11 standard MAC protocol," *IEICE Trans. Commun.*, vol. E83-B, no. 9, pp. 2068-2076, Sep. 200

[8] S. K. Sharma, K. Kim, C. Oh, and A. Ahmad, "Performance evaluation of IEEE 802.11 MAC protocol for multimedia services," *Proc. APCC/ICC's98*, pp. 181-185, 1998

[9] T. Suzuki and S. Tasaka, "Performance evaluation of priority-based multimedia transmission with the PCF in an IEEE 802.11 standard wireless LAN," *Personal, Indoor and Mobile Radio Communications*, vol. 2, pp.G-70 -G-77, Oct. 2001.

[10] IEEE Standard for Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE Std 802.11e/D6.0, November 2003.

[11] P. T. Brady, "A statistical analysis of on-off patterns in 16 conversations," *Bell Syst.Tech.J.*, vol. 47, pp. 73-91, Jan. 1968.

[12] D. L. Gall, "MPEG: A video compression standard for multimedia applications," *Commun. of the ACM*, vol. 34, no. 4, pp. 46-58, Apr. 1991.

[13] N. L. S. Fonseca and J. A. Silvester, "Modeling the output process of an ATM multiplexer with Markov modulated arrivals," *IEEE ICC '94* , pp. 721-725, 1994.

[14] J. J. Bae, and T. Suda, "Survey of traffic control schemes and protocols in ATM networks," *Proc*. of the *IEEE*, pp. 170-189, Feb. 1991.

[15] J. W. Roberts, "Variable-bit-rate traffic control in B-ISDN," *IEEE Commun. May.,* pp. 50-56, Sep. 1991.

[16] D. Qiao, S. Choi and K.G. Shin, "Goodput analysis and link adaptation for IEEE 802.11a wireless LANs," *IEEE transactions on mobile computing*, vol. 1, no. 4, pp. 278-292, Oct. 2002.

[17] A. Ganz and K. Wongthavarawat, "IEEE 802.11 wireless LAN association procedure for multimedia applications," *IEEE/MILCOM,* vol 2, pp. 1287-1291, Nov. 1999)

[18] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," *Proc. ACM SIGCOMM 1995*, pp. 100-113, Aug. 1995

[19] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: A Dynamic Protocol for an Ethernet PON (EPON)," *IEEE Commun. Mag*., vol. 40, no. 2, pp. 74-80, Feb.2002.

[20] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1-15, Feb. 1994

[21] H.-J. Zimmermann, "Fuzzy set theory and its applications," 2nd revised edition, Kluwer Academic Publishers, pp. 11-17, 1991.

[22] R. G. Cheng, C. J. Chang, and L. F. Lin, "A QoS-provisioning neural fuzzy connection admission controller for multimedia high-speed networks," *IEEE/ACM Trans. Networking*, vol. 7, no. 1, pp. 111-121, Feb. 1999.

Vita

姓名:吳俊憲

學歷:

      2002~2004   國立交通大學電信工程研究所

      1997~2002   國立交通大學土木工程系

      1994~1997   台北市立建國高級中學

E-mail: jswu.cm91g@nctu.edu.tw