# Improving the Aggregate Throughput of Access Points in IEEE 802.11 Wireless LANs

X. James Dong, Mustafa Ergen, Pravin Varaiya, and Anuj Puri
Department of Electrical Engineering & Computer Science
University of California at Berkeley
Berkeley, CA 94720

## Abstract

*In IEEE 802.11 wireless LANs, the DCF access method and the PCF access method operate alternatively within a superframe to service the time-varying traffic demands. Due to different medium access mechanisms deployed by DCF and PCF, they work well for some particular types of traffic scenarios respectively, while their performance may degrade under some other types of traffic demands. None of them can keep high performance under all types of traffic demands. In this paper, we propose an adaptive control algorithm to keep the aggregate throughput of access points close to the best capacity of both DCF mode and PCF mode upon various traffic requests, by dynamically adjusting the duration of DCF mode and PCF mode within a superframe based on recent traffic scenarios in wireless LANs. We modified the MAC module of IEEE 802.11 wireless LANs in OPNET to support our algorithm. Extensive simulations have been done using OPNET. The simulation results show that high performance of access points is achieved during the dynamic traffic demands. The new control algorithm only requires simple extensions to the MAC layer of access points and is compatible with the MAC layer defined in IEEE 802.11 standard.*

## 1 Introduction

Compared to its wired peers, wireless LANs have advantages like low cost, fast setup, flexible configuration, user mobility support, etc. Recently, the introduction of IEEE 802.11 standard provides basis for interoperability between different products [8]. In addition, the 11 Mbps of IEEE 802.11b and 54 Mbps of IEEE 802.11a offer high speed connections like Ethernet [9, 10]. As the cost of these wireless equipments continues to drop, wireless LANs are being massively deployed in public and residential buildings such as classrooms, airports, apartments, etc. More and more users are accepting wireless LANs as the dominant front-end networking facility to receive online services.

There are two ways to organize stations of wireless LANs: infrastructure mode and ad-hoc mode. In ad-hoc mode, each station plays the same role and communicates with other stations within its transmission range directly. The ad-hoc mode of wireless LANs is configured usually for temporary deployment. In infrastructure mode, stations consist of access points and mobile stations. Each mobile station communicates with the access point rather than other mobile stations. The access point may connect to the wired networks as a bridge or router for mobile stations to access the backbone networks. This paper considers the IEEE 802.11 wireless LANs with an access point in infrastructure mode.

Wireless LANs need to handle fully dynamic traffic loads, which are not evenly distributed over time and location. For example, access points in classrooms expect more intensive traffic demands during lecture hours than at night time; Access points in public libraries are supposed to carry more traffic than those in warehouses. Two access methods are defined in IEEE 802.11 wireless LANs to coordinate the wireless medium access: Distributed Coordination Function (DCF) and Point Coordination Function (PCF) [8]. DCF is a mandatory component in all IEEE 802.11 compatible products, while PCF is an optional component. They demonstrate different performance under different traffic condition. As we will discuss in late sections, DCF works well under low traffic demands, while PCF works satisfactorily under high traffic load. In DCF mode, at some hot spots, as the number of mobile stations associated with the access point becomes large and the traffic demands of each mobile station increase, all users will suffer from the fierce competition for wireless medium and frequent collisions. The performance of wireless LANs will degrade dramatically.

Noting the difference between performance of DCF and PCF in different traffic scenarios and the fact that DCF and PCF coexist in the same access point, we propose an adaptive control algorithm to dynamically adjust the ratio of DCF duration and PCF duration upon different traffic demands. Thus the capacity loss and performance degradation of wireless LANs can be alleviated to relieve the hot spot congestion under heavy traffic demands from a large number of mobile stations.

The rest of this paper is organized as follows. Section 2 introduces the DCF access method and the PCF access method briefly. Then the saturation throughput analysis of DCF mode and PCF mode are summarized. Section 4 proposes an adaptive control algorithm to tune the performance of wireless LANs under different traffic scenarios. In section 5, we evaluate the performance of wireless LANs with the new control algorithm via OPNET simulations. The simulation results show that high performance of access points is achieved during highly dynamic traffic demands.

## 2 Introduction to IEEE 802.11

IEEE 802.11 standard defines a network which is composed of Basic Service Sets (BSSs) that are interconnected with a Distribution System (DS). Mobile stations in a BSS gain access to the wired network through an access point. Each BSS has a BSS ID that distinguishes it from other BSSs. A mobile station accesses the wireless medium via associating with the access point which already authenticated it. A mobile station can be associated with only one access point but it can be authenticated by more than one access point. The DS supports mobility and the integration of BSSs in a manner that is transparent to mobile stations.

Known as Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), DCF is the basic access method in the IEEE 802.11 MAC protocol. In addition to the DCF, the standard also incorporates an alternative access method known as PCF. PCF is a centralized scheme which uses a point coordinator (PC, usually it is the access point) to determine which mobile station is allowed to transmit.

### 2.1 DCF

The DCF access method in IEEE 802.11 MAC protocol is based on CSMA/CA scheme [8], which has two important modules: carrier sense and binary exponential backoff. Stations waiting in the idle state sense the medium before any transmission attempt. There are two different carrier sense mechanisms: Virtual Carrier Sense (VCS) and Physical Carrier Sense (PCS). VCS is determined by the Network Allocation Vector (NAV), which is set according to the time specified in the Duration/ID field of packets. This gives stations information about how long the channel is occupied due to an ongoing transmission. On the other hand, PCS is a notification mechanism from physical layer to MAC layer saying that there is no signal detected at the moment. By combining VCS and PCS, MAC implements collision avoidance mechanism of CSMA/CA. Before initializing any transmission, the station checks VCS and then senses the medium for a DIFS time by PCS.

If a station finds the medium busy, it waits until the carrier sense mechanism notifies the station that the medium is idle. Next, the station goes to backoff state from PCS & VCS state and selects a backoff interval uniformly out of a contention window. Contention window doubles after every unsuccessful transmission attempt, consequently the station averagely waits longer in backoff. If the station senses a transmission while it counts down in backoff state, it suspends the transmission and goes back to carrier sense state and waits until the medium becomes idle and then starts counting down from where it stopped. When the backoff counter reaches zero, the station transmits the packet, and waits for ACK frame or CTS frame to make sure that the transmission is successful. In case of unsuccessful transmission, it doubles its contention window and increments its retry counter. When the retry count reaches the maximum value, the station discards that data frame.

To avoid hidden terminal problem and decrease the collision possibility, the RTS/CTS mechanism is introduced. A station reserves the medium by sending a Request-To-Send (RTS) frame. By setting their VCSs appropriately, all stations receiving the RTS frame stop transmitting except the station to whom the RTS is destined. Destined station sends a Clear-To-Send (CTS) frame to acknowledge the transmitting station that it is ready to receive. RTS and CTS frames contains the Duration/ID field in which the other stations learn how long the medium will be busy and set their NAVs accordingly. Stations wait SIFS time between packets. The RTS/CTS mechanism is disabled when the data size is below the RTS threshold, as specified in the standard [8].

### 2.2 PCF

The PCF access method is built using the DCF through the use of an access priority mechanism that provides contention free access to the channel. The point coordinator starts a CF period by broadcasting a beacon frame. The point coordinator seizes the control of the idle channel by transmitting after it senses that the channel has been idle for a time interval of Priority Inter frame Space (PIFS) that is chosen to be smaller than DIFS but larger than SIFS. The transmission of a CF-ACK frame is an acknowledgement by the PC and CF-End frame marks the end of the contention free period by PC. In PCF mode, the point coordinator gives each associated mobile station a turn to transmit in a predetermined order.

The PC sends data to mobile stations in CF-Poll frames which also has the polling function. A mobile station transmits frames after the reception of a CF-Poll frame with the

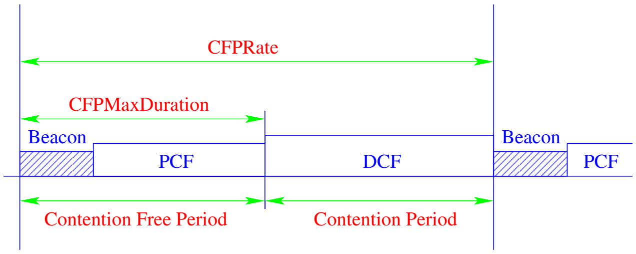

**Figure 1. DCF, PCF, and superframe**

poll bit enabled. The need for separate acknowledgements is avoided by piggybacking the acknowledgement on subsequent frame (by the setting of the appropriate bit).

Frames from a mobile station are transmitted after the channel has been idle for a time interval SIFS as compared to CF-Poll frames that are transmitted once the channel has been idle for PIFS. Thus the PC will transmit the next CF-Poll frame in case there is no frame from a mobile station in time interval PIFS after the transmission of the previous CF-Poll frame. To minimize collisions during the contention free period, each station sets its NAV equal to the maximum allowable length of the CFP. However, a station resets its NAV if a CF-ACK frame is seen by it before its NAV has expired.

## 2.3 Superframe

If both DCF mode and PCF mode are enabled at an access point, transmission slots are divided into Contention Free Period (CFP, during which PCF is active) and Contention Period (CP, during which DCF is active), as shown in Figure 1. CFP and CP appear alternately. Each CFP and the CP immediately after it form a superframe. Superframes are separated by periodical beacon frames. Each beacon frame announces the start of a superframe. The CFP duration is constrained by CFPMaxDuration, a MIB parameter, while the CFP frequency (or beacon frequency) is determined by CFPRate, another MIB parameter. These MIB parameters are writable (controllable to users). So flexibility is provided by IEEE 802.11 standard to optimize the performance of wireless LANs.

## 3 The Saturation Throughput of IEEE 802.11 Wireless LANs

The saturation throughput of the DCF access method has been extensively studied in recent literatures [2, 3, 5, 6, 7, 12, 14]. This paper adopts the analysis presented in [2, 3]. As to the saturation throughput of PCF access method, we'll give a simple treatment in this section. The results in this section are applicable to a wireless LAN that consists of an access point and a number of mobile stations, where the outcoming frame buffer of each active mobile station is always not empty and the access point is passive to receive data frames and exchange control frames with mobile stations.

| | |
|---|---|
| $T$ | number of total mobile stations in a wireless LAN |
| $n$ | number of active stations in a wireless LAN |
| $W$ | minimum window size |
| $m$ | maximum window size is $2^m W$ |
| $p$ | the conditional probability that each frame collides constantly and independently |
| $\tau$ | the probability that a station transmits in a randomly chosen slot time |
| $b(t)$ | the random process of the backoff time counter for a given station |
| $s(t)$ | the random process of the backoff stage for a given station at time $t$ |
| $b_{i,k}$ | $\lim_{t \to \infty} P\{s(t) = i, b(t) = k\}, i \in (0, m), k \in (0, W_i - 1)$ |
| $\delta$ | the propagation delay |
| SIFS | the SIFS time duration |
| DIFS | the DIFS time duration |
| PIFS | the PIFS time duration |
| $\sigma$ | the propagation delay |
| PHY | transmission time of PHY preamble and header |
| MAC | transmission time of MAC header |
| RTS | transmission time of a RTS frame |
| CTS | transmission time of a CTS frame |
| ACK | transmission time of a ACK frame |
| H | transmission time of PHY and MAC header of a frame |
| $P_{tr}$ | the probability with at least one transmission in each slot |
| $P_s$ | the probability with a successful transmission in each slot |
| $T_s$ | the average time that the wireless channel is sensed busy because of a successful transmission |
| $T_c$ | the average time that the wireless channel is sensed busy because of collisions |
| $L$ | transmission time of MAC payload |
| $S$ | the normalized channel utilization rate of successful payload transmission |

**Table 1. Main notations**

To facilitate following discussion, we define some notations in Table 1. They are almost the same as the glossary used in [3]. What we want to clarify here is that the transmission time of the MAC overhead and control frames include the time for the PLCP (Physical Layer Convergence Protocol) Preamble and the PLCP Header as defined in the IEEE 802.11 standard [8]. The basic service rate, at which the control and management frames are transmitted, is assumed to be 1 Mbps. In addition, $T$ means the total number of mobile stations that are associated with the access point. If outcoming frame buffer of the mobile station is not empty within a duration, we call this mobile station an active station for the duration. If no traffic from a mobile station within a duration, we call it an inactive station for the duration. To make the analysis tractable, each station is assumed to be either in active state or inactive state.

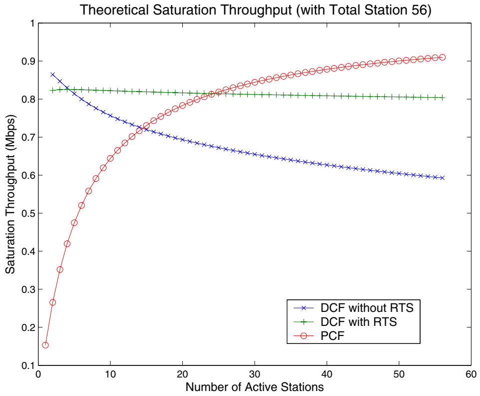

**Figure 2. The theoretical saturation throughput**

### 3.1 DCF Mode

In [2, 3], the author assumed that the channel noise can be ignored, there are no hidden terminal and capture problems, and the stations are in saturation state (their buffers are always not empty during stable state). In stable saturation state, the conditional collision probability $p$ and the transmission probability $\tau$ are assumed to be constant and independent for each station. If $p$ and $\tau$ are known, then the successful transmission probability in each slot can be obtained. Thus the efficient channel utilization rate can be calculated. To get $p$ and $\tau$, a two dimensional Markov chain model based on stationary probability $b_{i,k}$ is deployed to describe the saturation behavior of wireless LANs [2, 3]. Then $p$ and $\tau$ can be obtained by numerically solving following nonlinear system equations:

$$b_{0,0} = \frac{2(1-2p)(1-p)}{(1-2p)(W+1)+pW(1-(2p)^m)} \quad (1)$$

$$\tau = \sum_{i=0}^{m} b_{i,0} = \frac{1-p^{m+1}}{1-p} b_{0,0} \quad (2)$$

$$p = 1-(1-\tau)^{n-1} \quad (3)$$

To calculate $S$, the fraction of time used to transmit payload frames successfully, we need to get $P_{tr}$ and $P_s$ based on $\tau$ by following equations:

$$P_{tr} = 1-(1-\tau)^n \quad (4)$$

$$\begin{aligned} P_s &= \frac{n\tau(1-\tau)^{n-1}}{P_{tr}} \\ &= \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n} \end{aligned} \quad (5)$$

We are interested in the normalized throughput, so the data transmission rate is assumed to be 1 Mbps. Then the saturation throughput $S$ can be obtained by:

$$\begin{aligned} S &= \frac{E[\text{Payload transmission time in a slot time}]}{E[\text{Length of a slot time}]} \\ &= \frac{P_s P_{tr} E[L]}{(1-P_{tr})\delta + P_s P_{tr} T_s + (1-P_s) P_{tr} T_c} \end{aligned} \quad (6)$$

The computation of $T_s$ and $T_c$ depends on the mechanisms used by DCF. If only the basic access mechanism is utilized, we have:

$$T_s^{bas} = \text{H+E[L]+}\sigma\text{+SIFS+ACK+}\sigma\text{+DIFS} \quad (7)$$

$$T_c^{bas} = \text{H+E[L]+}\sigma\text{+SIFS+ACK+DIFS} \quad (8)$$

If the RTS/CTS mechanism is enabled, we have:

$$\begin{aligned} T_s^{rts} =& \text{RTS+}\sigma\text{+SIFS+CTS+}\sigma\text{+SIFS} \\ & \text{H+E[L]+}\sigma\text{+SIFS+ACK+}\sigma\text{+DIFS} \end{aligned} \quad (9)$$

$$T_c^{rts} = \text{RTS+}\sigma\text{+SIFS+CTS+DIFS} \quad (10)$$

As the correction given in [14], here we consider both the ACK timeout and the CTS timeout, which make the analysis more accurate. The numerical results of the saturation throughputs in DCF mode with RTS/CTS disabled or in DCF mode with RTS/CTS enabled are shown in Figure 2.

### 3.2 PCF Mode

In PCF mode, the timing diagram for frame exchange sequences in the stable saturation state is given in Figure 3. The polling order of each mobile station is assumed to be determined by an algorithm with circular round-robin fashion. Define POLL to be the transmission time of a CF-Poll frame and NULL to be the transmission time of a Null frame. Based on the timing diagram, the approximate formulae for the saturation throughput is:

$$S = \frac{nE[L]}{\text{n(2SIFS+H+E[L]+2}\sigma\text{+POLL)+(T-n)(2SIFS+2}\sigma\text{+POLL+NULL)}} \quad (11)$$

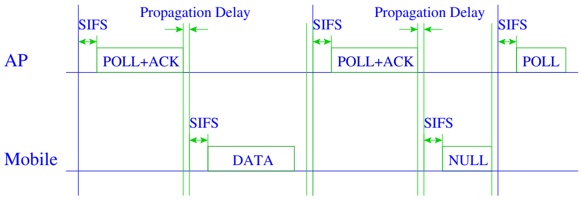

**Figure 3. Timing diagram of PCF**

If the number of total stations, $T$, is 56, the numerical results of saturation throughput in PCF mode are shown in Figure 2. As the ratio of active stations over total stations becomes larger, the efficiency of PCF mode also increases.

# 4 An Adaptive Control Algorithm

## 4.1 The System Dynamics of Wireless LANs

We aim to keep the aggregate throughput of access points in wireless LANs close to the capacity under any saturation traffic demands. Unlike traditional controllable physical systems, wireless LANs may produce different outputs given the same input sequences and environment parameters. It is difficult to quantitatively model the system behavior of a fully dynamic and nonlinear dynamic system like a wireless LAN. Fortunately, in saturation state, the analysis has been well established for both DCF mode and PCF mode, as we discussed in previous section.

In DCF mode, the saturation throughput $F_d$ is a function of backoff window size $W$, maximum backoff stage $m$, retry limit $R$, and the number of active stations $n$. Since $W$, $m$, and $R$ are either defined in the standard or predefined by users, $F_d$ is a function of $n$, denoted as $F_d(n)$. In PCF mode, the saturation throughput $F_p$ is a function of the total number of associated stations $T$, and the number of current active stations $n$. Since $T$ is known to the access point in operation, the saturation throughput in PCF mode can be denoted as $F_p(n)$.

Assuming the ratio of PCF duration within a superframe is $r$, the saturation throughput of the access point is approximately $r * F_p(n) + (1 - r) * F_d(n)$. Actually $r$ can be dynamically modified by changing the writable MIB parameter CFPMaxDuration [8]. For any given $n$, we want to set appropriate CFPMaxDuration to maximize the total saturation throughput, as shown below:

$$\max_r \{r * F_p(n) + (1 - r) * F_d(n)\} \qquad (12)$$

In the optimal case, the actual capacity curve consists of two best parts from DCF mode and PCF mode respectively, as shown in Figure 2. There is an intersection point, at which the number of active station is $n_{cross}$, between the saturation throughput curves of DCF mode and PCF mode. If the number of active station is less than $n_{cross}$, DCF works well to meet the traffic demands from less active stations with less contention for wireless medium. If the number of active station is greater than $n_{cross}$, PCF is better because the polling algorithm of PCF becomes efficient under heavy traffic demands. However, it is not practical to estimate whether the active stations are in saturation state or not or how many stations are in saturation state. Because the control algorithm has to deal with the non-saturation case. We need to find other system state variables that can be easily measured.

## 4.2 The Adaptive Control Framework

The major components of our adaptive control framework consist of two input state variables, a controllable variable, and the control algorithm, as shown in figure 4.

We consider the target system to be a discrete time dynamic system. The measurements are collected and the control actions are exerted at regular intervals. These intervals don't need to be constant. Assuming the time for measurement collections and control actions are $t_1$, $t_2$, ..., $t_i$, ...., the system variable $x(t)$ at each time interval are $x_1$, $x_2$, ..., $x_n$, ..., respectively.

Similar to conventional digital control systems, to find the right sampling interval is not trivial. On one aspect, if the sample period is too long, the wireless LAN can't respond to bursty traffic or instantaneous changes. On the other aspect, if the sample period is too short, the system will be sensitive to measurement variance and more processing workload will be incurred. The unit of sampling interval is assumed to be the number of beacons.

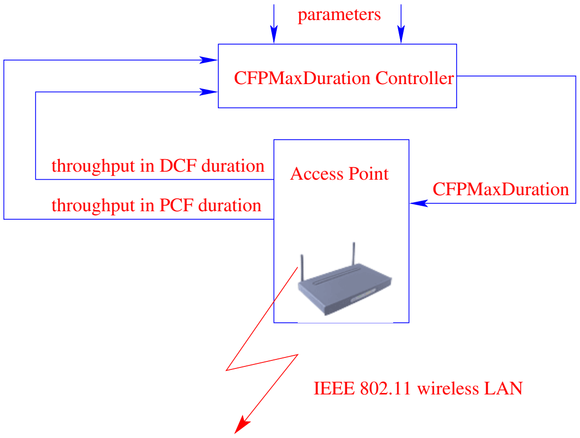

**Figure 4. The feedback control framework**

Two input variables for the control algorithm are used: $f_d(t)$ and $f_p(t)$. $f_d(t)$ is the measured throughput during DCF mode operation of wireless LANs, while $f_p(t)$ is the measured throughput during PCF mode operation of wireless LANs. They are updated every sampling period.

The controllable variable $r$ is the percentage of PCF duration within each superframe. CFPMaxDuration ($r$ times beacon interval) is a MIB parameter in access points that can be updated by SNMP protocol. We take the discrete value of $r$ from a finite set $\{r_1, r_2, ...r_N\}$, in which $\{r_1 \leq r_2 \leq ... \leq r_N\}$. The number of these discrete values is $N$.

The adaptive control algorithm predicts the aggregate throughput of access points based on the system dynamics discussed in previous section. It exerts appropriate control actions by changing $r$ and accepts the feedback of $f_d(t)$ and $f_p(t)$.

### 4.3 The Adaptive Control Algorithm

One of the main purposes of medium access methods is to increase the bandwidth or resource utilization rate and make the resource fairly shared by users. With binary exponential backoff and VCS/PCS mechanisms, etc., the DCF access method is designed to avoid collisions. Given that the traffic load is heavy enough to cause frequent collisions, for example, a large number of active stations with intensive VoIP traffic, DCF will not meet their demands. This is indicated by the throughput degradation. However, under the same intensive traffic demands, PCF can make full use of the medium by eliminating some extra overhead of control frames and backoff duration, which is indicated by the throughput improvement. If the traffic demand is not heavy, for example, occasional web browsing and email exchanging from a large number of stations, both access methods can meet the demand. However, in this case, the DCF access method usually achieves better throughput and delay. Therefore, it is reasonable to accept throughput as the feedback indicator whether current dominant access method fits the traffic pattern.

The pseudo code of the adaptive control algorithm is given below.

```
CFPMaxDurationControl()
measure $f_p(t)$ and $f_d(t)$
if $f_p(t) \leq (1+\beta) f_d(t)$ then
  $i = max\{1, i-1\}$
  $r = r_i$
else
  if $f_d(t) \leq (1+\beta) f_p(t)$ then
    $i = min\{i+1, N\}$
    $r = r_i$
  end if
end if
```

$\beta$ ($0 \leq \beta \leq 1$) is a damping factor, which is introduced to avoid frequent oscillations of $r$ caused by randomness of the measurements.

There are some constraints on the ratio $r$. Firstly, no matter what the value of $r$ is, the PCF duration and the DCF duration should be large enough for the transmission of at least one frame with reasonable length. If no frame gets transmitted successfully in either DCF mode or PCF mode, the measurements of $f_p(t)$ and $f_d(t)$ are not accurate enough to demonstrate the real transmission capability in either mode under given traffic demands. In addition, marginal time slots are needed to transmit beacons and other management frames like association, disassociate, etc. The lower bound and upper bound of $r$ are denoted as $r_{min}$ and $r_{max}$ respectively.

Although the adaptive control of parameter CFPMaxDuration via $r$ is proposed in this paper, we don't mean to say that the DCF access method and the PCF access method is interchangeable from the view point of traffic. On the contrary, it is the traffic pattern that determines which access method performs well and deserves more duration. What the controller does is to identify the access method that recent traffic pattern favors by measuring the throughput $f_d(t)$ in DCF mode and throughput $f_p(t)$ in PCF mode respectively.

## 5 Simulations and Performance Evaluation

This section describes the simulation setup and performance evaluations of different schemes. For a particular traffic scenario, we run simulations in three different modes: DCF mode, PCF mode, and adaptive hybrid mode. In DCF mode or PCF mode, only the DCF or the PCF access method used. In adaptive hybrid mode, the adaptive control algorithm presented in this paper is used to change the ratio $r$ of PCF duration within a superframe. The length of a superframe is fixed.

### 5.1 Simulation Setup

All simulations are made using an event-driven network simulator, OPNET release 8.1.A [15]. The Wireless LAN simulation model is part of standard OPNET model for IT DecisionGuru and OPNET Modeler with radio module. The MAC layer and PHY layer of the wireless LAN model is designed based on the IEEE 802.11 standard [16]. Although the MAC implementation is not complete, it supports most of IEEE 802.11 standard functionalities, such as the backoff, deference, RTS/CTS, NAV, frame exchange sequence, fragmentation, access point functionality, basic service set, lost frame retransmission, and duplicate frame detection, etc. Both DCF mode and PCF mode are supported. The DCF mode was implemented in OPNET by mapping objects defined in the SDL-92 formatted documentation to OPNET objects, to avoid the ambiguity caused by various interpretations of the standard documentation [1]. The PCF module in OPNET is revised to schedule pollings with the circular round-robin fashion. In addition, OPNET supports the 802.11b physical layer model [16]. Both PLCP preamble and PLCP header are implemented. Multirate support is provided.

The network topology of our target simulations is taken from the typical real life classroom wireless networks. In the classroom with area of $30m \times 20m$, 56 mobile stations and an access point form an independent BSS. Each station in this wireless LAN can hear about each other, and has no mobility at all. Therefore, the hidden terminal problem and the capture effects can be ignored. Each active station carries 1 Mbps CBR traffic demand that is large enough to saturate the wireless LAN. Each station is activated one by one every 100 seconds.

| | |
|---|---|
| W | 32 |
| m | 5 |
| PHY | 192 bits |
| MAC | 224 bits |
| RTS | 160+192 bits |
| CTS | 112+192 bits |
| ACK | 112+192 bits |
| Slot Time | 20 $\mu s$ |
| SIFS | 10 $\mu s$ |
| DIFS | 50 $\mu s$ |
| ACK Timeout | 314 $\mu s$ |
| CTS Timeout | 314 $\mu s$ |
| short retry limit | 255 |
| long retry limit | 255 |

**Table 2. Simulation parameters for IEEE 802.11 MAC protocol**

| | |
|---|---|
| MAC Payload | 8000 bits |
| channel bit rate | 1 Mbps |
| $r_{min}$ | 10% |
| $r_{max}$ | 90% |
| $r_{init}$ | 50% |
| the damping factor $\beta$ | 5% |
| beacon interval | 1 second |
| the sampling interval | 2 beacons |

**Table 3. Simulation parameters for the control algorithm**

The configuration of the access point and mobile stations is mostly based on the default parameter values proposed for the Direct Sequence Spread Spectrum (DSSS) system in IEEE 802.11b standard, like the length of RTS, CTS, ACK, MAC header, the slot time, the SIFS, the DIFS, the PLCP preamble, etc, as given in Table 2. Some other parameters, which are related to applications or defined for the dynamic control algorithm, are summarized in Table 3. The length of data frames is constant. There is no fragmentation. The radio link is configured as a lossless wireless channel. Note that the MAC header we have used is 224 bits instead of 272 bits. Because 272 bits are for the MAC header of a generic IEEE 802.11 data frame in which the ToDS and FromDS bits are set to be 1 (ToDS bit is set to 1 if the frame goes to the Distribution System and FromDS bit is set to 1 if the frame comes from the Distribution System). Although the address fields of a generic data frame include up to four addresses (source, destination, transmitting access point, and receiver access point), only three of four addresses are used if all traffic flows within the BSS. In our simulations, since all data frames are transmitted among stations within the BSS, we use 224 bits as the length of MAC header, which may slightly increase bandwidth efficiency when compared to the results obtained from [2, 3, 14].

## 5.2 Measurement and Explanation

For each simulation using DCF mode, we do it with RTS/CTS disabled and with RTS/CTS enabled. Since the values of some parameters in simulations are not the same as those in analysis, some simulation results only match the analysis results approximately.

Figure 5 and Figure 7 show the evolution of the ratio $r$ as the number of active stations increases. The initial value of $r$ is 0.5. When the number of active stations is relatively small, the throughput in PCF mode is less than the throughput in DCF mode because of the wasted poll slots. Therefore, the ratio $r$ decreases from 0.5 to $r_{min}$, and remains $r_{min}$ until the number of active stations reaches a threshold $n_{cross}$. At $n_{cross}$, there is an intersection point between the throughput curves of DCF mode and PCF mode. If the number of active stations exceeds $n_{cross}$, the throughput of PCF mode is better than the throughput of DCF mode in saturation state. Accordingly, $r$ changes gradually from $r_{min}$ to $r_{max}$. Due to the adoption of damping factor, the ratio $r$ shows stable behavior and few oscillations.

In Figure 6 and Figure 8, if there is only one active station, the saturation throughput is quite close to 0.9 Mbps with RTS/CTS disabled or 0.8 Mbps with RTS/CTS enabled, e.g., the bandwidth efficiency is close to 0.9 or 0.8 respectively. The study in [4] verifies this. However, some recent studies reported that the bandwidth utilization rate of IEEE 802.11b without RTS/CTS can be only up to 60% under single user transmission [13]. The main reason behind the difference is that the 11 Mbps bandwidth is only for PHY payload of data frames. Actually the transmission time of the PHY header (PLCP header and PLCP preamble) and all IFS time durations remain the same as those in IEEE 802.11 for the same DSSS physical layer. Assuming the ideal channel condition for single user transmission, the collisions can be ignored. Then the average backoff counter is 15.5 and the average backoff time is 310 $\mu s$. Here we present a simply way to calculate the maximum theoretical throughput under single user transmission for the case that the transmission rate is 1 Mbps and the case that the transmission rate is 11 Mbps. The RTS/CTS mechanisms is assumed to be disabled. If the transmission rate is 1 Mbps and the basic rate for control frames is 1 Mbps, the maximum throughput is:

$$S = \frac{\text{E[L]}}{\text{H+E[L]+}\sigma\text{+SIFS+ACK+}\sigma\text{+DIFS+310}} \tag{13}$$

If the transmission rate is 11 Mbps and the basic rate for control frames is also 11 Mbps, the maximum throughput

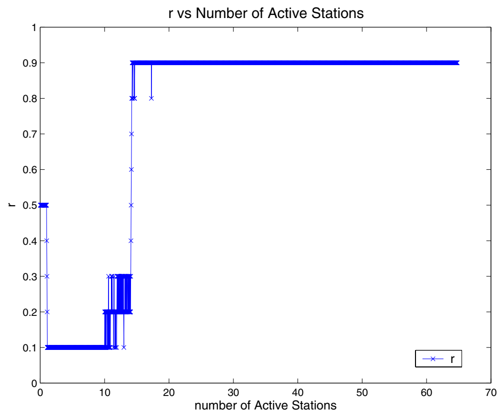

Figure 5. Ratio r vs number of active stations (DCF without RTS/CTS)

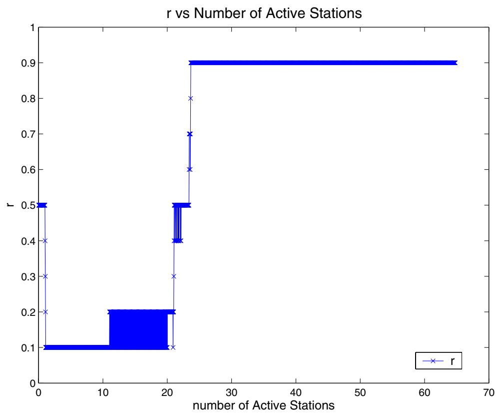

Figure 7. Ratio r vs number of active stations (DCF with RTS/CTS)

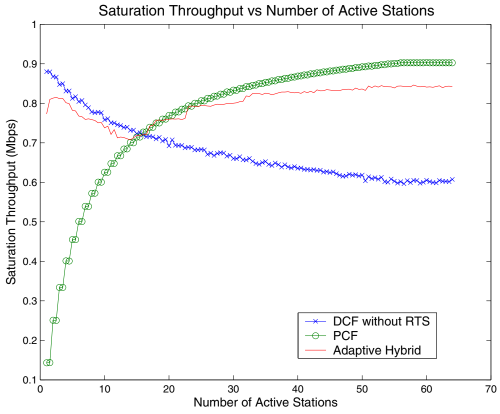

Figure 6. Saturation throughput vs number of active stations (DCF without RTS/CTS)

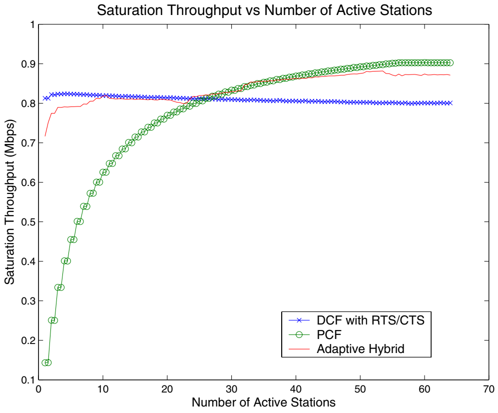

Figure 8. Saturation throughput vs number of active stations (DCF with RTS/CTS)

is:

$$S = \frac{\frac{\text{E[L]}}{11}}{\text{2PHY+SIFS} + \frac{\text{MAC+E[L]+ACK-PHY}}{11} + 2\sigma + \text{DIFS+310}}$$

Let the MAC payload be 1000 bytes, the bandwidth efficiency at 1 Mbps and at 11 Mbps are 87.99% (matches our simulation result) and 48.042% respectively. Let the MAC payload be 1470 bits, then the bandwidth efficiency at 1 Mbps and at 11 Mbps are 91.5% and 57.6% respectively. The bandwidth efficiency at 11 Mbps matches the results reported in recent studies [13]. In addition, the implementation of short radio does improve the utilization rate of IEEE 802.11b by shortening the transmission time of PLCP header. But it is optional in the standard.

Another interesting phenomenon is that the maximum throughput obtained with RTS/CTS enabled is not strictly decreasing as the number of active stations increases, as shown in Figure 8. The reason is that there are more backoff slots between two neighboring transmissions (may not from the same station) in the channel if the number of active stations is extremely small and the stations have to wait for each other during these backoff slots, which is also explained in [12].

Figure 6 and Figure 8 compare the throughputs of all cases. Before the threshold $n_{cross}$, the throughput in pure DCF mode is the best. After the threshold $n_{cross}$, the throughput in pure PCF mode is the best. The throughput in adaptive hybrid mode is always very close to the best one. Before the threshold $n_{cross}$, it is far more better than the throughput in pure PCF mode. After the threshold $n_{cross}$, it is far more better than the throughput in pure DCF mode.

From the simulation results we find that the introduction of the adaptive control algorithm proposed in this paper makes the access point responsive to current traffic condition changes. It exhibits the good performance of DCF mode under low traffic profile and the satisfactory efficiency of PCF mode under intensive traffic demands from large number of active stations.

## 6 Conclusion and Future Work

Due to the scarce radio resource and limited available bandwidth provided by wireless LANs, the channel utilization rate is an important metric of the efficiency of MAC protocols. For most current MAC protocols, their efficiencies also depend on current traffic demands. None of these MAC protocols works perfectly under any types of traffic demands. Specifically in IEEE 802.11 wireless LANs, the DCF access method is a good choice if the traffic demands are not intensive, while the PCF access method is preferred in case that the number of transmitting nodes is large and most traffic demands are video/audio flows or large file transferring. This paper proposes an adaptive control algorithm to tune the performance of IEEE 802.11 wireless LANs, so that traffic demands with different characteristics will be handled mostly by the MAC protocol that fits them.

Our algorithm aims to control the percentage of CFP duration within a superframe by measuring the throughput in PCF mode and the throughput in DCF mode periodically. The percentage of CFP duration slides among a series of evenly distributed points. To avoid the oscillation of the percentage, the damping factor is introduced. In addition, enough duration is reserved for both DCF mode and PCF mode to transmit management frames and keep the throughput measurements valid. We modified the MAC model of IEEE 802.11 wireless LANs in OPNET to support our algorithm. Extensive simulations have been done using OPNET. The simulation results show that the control algorithm can keep the throughput close to the capacity of the wireless LAN under various traffic demands with different characteristics.

The new control algorithm only requires simple extensions to the MAC layer of access points and is compatible with the MAC layer defined in IEEE 802.11 standard. The mobile hosts don't need to change anything.

## References

[1] R. Baldwin, N. Davis IV, S. Midkiff, Implementation of an IEEE 802.11 Wireless LAN Model using OPNET, 1998 OPNETWORK, Washington D.C., June 1998.

[2] G. Bianchi, IEEE 802.11-Saturation throughput analysis, IEEE Communincations Letters, vol. 2, Dec. 1998, pp.318-320.

[3] G. Binachi, Performance Analysis of the IEEE 802.11 Distributed Coordination Function, IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, March 2000.

[4] B. Bing, Measured Performance of the IEEE 802.11 Wireless LAN, in Proc. 26th Conference on Local Computer Networks (LCN 1999), Lowell, Massachusetts, USA, 17-20 October, 1999, pp34-42.

[5] F. Cali, M. Conti, and E. Gregori, Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit, IEEE/ACM Trans. on Networking, vol.8, No.6, Dec. 2000, pp.785-799.

[6] H. S. Chayya and S. Gupta, Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol, Wireless Networks, Vol 3, 1997, pp.217-234.

[7] T. S. Ho and K. C. Chen, Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs, in Proc. IEEE PIMRC, Taipei, Taiwan, Oct. 1996, pp. 392396.

[8] IEEE 802.11 Standard Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999.

[9] Supplement to IEEE 802.11 Standard Part II: Higher-speed Physical Layer Extension in the 5 Ghz Band, 1999.

[10] Supplement to IEEE 802.11 Standard Part II: Higher-Speed Physical Layer Extension in the 2.4GHz Band, 1999.

[11] IEEE 802.11 Working Group, http://grouper.ieee.org/groups/802/11/.

[12] Tay, Y., and K. Chua. 2001. A Capacity Analysis for the IEEE 802.11 Mac protocol. Wireless Networks, 7(2): 159-71.

[13] A. Vasan and A. Udaya Shankar, An Empirical Characterization of Instantaneous Throughput in 802.11b WLANs, citeseer.nj.nec.com/vasan02empirical.html.

[14] Haitao Wu, Yong Peng, Keping Long, and Shiduan Cheng and Jian Ma, Performance of Reliable Transport Protocol over IEEE 802.11 Wireless LAN: Analysis And Enhancement, INFOCOM 2002.

[15] IT DecisionGuru and OPNET Modeler (http:// www.opnet.com)

[16] OPNET Technologies, Inc. Wireless LAN Model Descrption. November 15, 2000.