

Achieving Efficient Channel Utilization and Weighted Fairness for Data Communications in IEEE 802.11 WLAN under the DCF

Daji Qiao and Kang G. Shin
Real-Time Computing Laboratory
The University of Michigan
Email: {dqiao,kgshin}@eecs.umich.edu

Abstract— Fair allocation of bandwidth and maximization of channel utilization are two important issues when designing a contention-based wireless medium access control (MAC) protocol. However, achieving both design goals at the same time is very difficult, and has not yet been addressed elsewhere. In this paper, we study this challenging problem, particularly for data communications in IEEE 802.11 wireless local-area networks (WLANs).

We propose a priority-based fair medium access control (P-MAC) protocol by modifying the Distributed Coordination Function (DCF) of the IEEE 802.11 MAC. The key idea is that the contention window size for each wireless station is properly selected to (1) reflect the relative weights among data traffic flows, so as to achieve the weighted fairness; and (2) reflect the number of stations contending for the wireless medium, so as to maximize the aggregate throughput. In P-MAC, our approximations to the optimal contention window sizes, which are based on a theoretical analysis, are evaluated numerically and shown to work well under different network configurations and traffic scenarios. Moreover, simulation results show that, with few changes to the original DCF, P-MAC performs significantly better in terms of both fairness and throughput.

I. INTRODUCTION

A. Motivation and Problem Statement

There are two basic types of wireless medium access control (MAC) protocols: *polling-based* and *contention-based*. In a polling-based MAC protocol, a coordinator station is responsible for all the frame transmissions on the shared wireless medium. A wireless station that wants to transmit must wait until it is polled by the coordinator station. The Point Coordination Function (PCF) of the IEEE 802.11 MAC [1] is an example of the polling-based approach. It uses a poll-and-response protocol to eliminate contentions among wireless stations. In contrast, a coordinator station is not required to perform the centralized medium access control in a contention-based MAC protocol. Any wireless station that wishes to transmit does so if the wireless medium is sensed free. The wireless stations are, in fact, contending for the shared medium, and thus, collisions are inevitable. The Distributed Coordination Function (DCF) of the IEEE 802.11 MAC is an example of the contention-based approach.

Fair allocation of bandwidth and maximization of channel utilization have been identified as two important goals when de-

signing a contention-based wireless MAC protocol. Unfortunately, these two design goals create inherent conflicts between them. For example, in an IEEE 802.11 WLAN under the DCF, maximum channel utilization may be achieved if there is only one station transmitting continuously with zero backoff, while all the other stations are starved. Clearly, this is unfair. In general, it is very difficult to maximize the channel utilization subject to the constraint of achieving fairness among traffic flows. In this paper, we study this challenging problem, particularly for data communications in an IEEE 802.11 WLAN, as a starting point.

The ideal weighted fairness is defined as follows. Assume that there are $n (> 1)$ different traffic classes each characterized by a positive weight. Let ϕ_i denote the weight associated with class- i traffic ($1 \leq i \leq n$), and without loss of generality, let's assume that $1 = \phi_1 > \phi_2 > \dots > \phi_n > 0$. Further, assume that each wireless station carries only one traffic flow.¹ Let f_i denote the set of stations carrying class- i traffic, and let $w_i^s(t_b, t_e)$ be the amount of class- i traffic transmitted by station $s \in f_i$ during the time interval $[t_b, t_e]$. To be fair to all the traffic flows, it requires, regardless of where and how small the interval $[t_b, t_e]$ is,

$$\forall i, j \in \{1, \dots, n\}, \forall s \in f_i, \forall s' \in f_j, \frac{w_i^s(t_b, t_e)}{\phi_i} = \frac{w_j^{s'}(t_b, t_e)}{\phi_j}. \quad (1)$$

However, the ideal weighted fairness cannot be accurately achieved in practice, since data transmitted on a real network is packetized. Instead, we define a new fairness objective function for data communications in a WLAN.²

As specified in the IEEE 802.11 standard [1], each data packet generated by the higher layer is fragmented further into smaller MAC frames for transmission. Therefore, it is reasonable to assume that *each data traffic flow has the same MAC frame size*. Let SU_s be the probability that a MAC frame transmission is successful and is from station s . If Eq. (2) holds, all the traffic flows within a WLAN would share the wireless medium fairly in a probabilistic sense, and we claim that the weighted fairness

¹If a wireless station carries multiple traffic flows, the fairness among those traffic flows can be guaranteed by some in-station packet scheduling algorithms, e.g., weighted round-robin.

²In this work, we focus on the fair allocation of bandwidth among multiple data traffic flows, but not their delay/jitter performances.

The work reported in this paper was supported in part by AFOSR under Grant No. F49620-00-1-0327.

intended for data communications in a WLAN is achieved.

$$\forall i, j \in \{1, \dots, n\}, \forall s \in f_i, \forall s' \in f_j, \frac{SU_s}{\phi_i} = \frac{SU_{s'}}{\phi_j}. \quad (2)$$

The objective is to design a contention-based MAC protocol for data communications in a WLAN, which (1) achieves weighted fairness among data traffic flows while maximizing the aggregate throughput; and (2) maintains the compatibility or close resemblance to the DCF of the IEEE 802.11 MAC, so as to facilitate its deployment.

B. Related Work

Much research has been done to develop scheduling algorithms for wireless networks to achieve weighted fairness among traffic flows that share the wireless medium [2][3][4][5]. These algorithms are centralized by design, and therefore, can only be embedded into the polling-based MAC protocols. There has also been some work done on the contention-based MAC protocols. The protocols presented in [6] and [7] attempt to provide equal shares of bandwidth to different stations, and the traffic weights are implicitly assumed to be the same. In [8], although priorities have been taken into consideration when controlling the medium access, this protocol does not perform fair allocation of bandwidth. There are three recent papers [9][10][11] that address the weighted fairness issues in multi-hop wireless networks. The authors of [9] presented a fully-distributed algorithm for scheduling frame transmissions such that different traffic flows are allocated bandwidth in proportion to their weights. The authors of [10] presented a novel mechanism of translating any pre-specified fairness model into a corresponding backoff-based contention resolution algorithm. However, the focus of both papers is to ensure fairness through appropriate MAC layer designs, and neither of them attempts to maximize the channel utilization. [11] is the only paper that attempts to deal with both fairness and utilization maximization simultaneously. It focuses on maximizing the aggregate channel reuse in a multi-hop wireless network subject to a minimum fairness guarantee, which is quite different from the issues we address in this work.

C. Organization

The rest of this paper is organized as follows. Section II introduces the DCF of the IEEE 802.11 MAC and details its backoff behavior, then states the assumptions to be used. The proposed priority-based fair medium access control (P-MAC) protocol is presented in Section III. Section IV presents and discusses the simulation results, and Section V concludes the paper.

II. SYSTEM OVERVIEW

A. DCF of IEEE 802.11 MAC

The DCF [1], as the basic access mechanism of the IEEE 802.11 MAC, achieves automatic medium sharing between compatible stations through the use of Carrier-Sense Multiple

Access with Collision Avoidance (CSMA/CA). Before a station starts transmission, it senses the wireless medium to determine if it is idle. If the medium is idle, the transmission may proceed, else the station will wait until the end of the in-progress transmission. The CSMA/CA mechanism requires a minimum specified gap/space between contiguous frame transmissions. A station will ensure that the medium has been idle for the specified inter-frame interval before attempting to transmit.

The *distributed inter-frame space* (DIFS) is used by stations operating under the DCF to transmit data and management frames. A station using the DCF has to follow two medium access rules: (1) the station will be allowed to transmit only if its carrier-sense mechanism determines that the medium has been idle for at least DIFS time; and (2) in order to reduce the collision probability among multiple stations accessing the medium, the station will select a random backoff interval after deferral, or prior to attempting to transmit another frame after a successful transmission.

One important characteristic of the IEEE 802.11 MAC is that an acknowledgment (ACK) frame will be sent by the receiver upon successful reception of a data frame. Only after receiving an ACK frame correctly, the transmitter assumes successful delivery of the corresponding data frame. The *short inter-frame space* (SIFS), which is smaller than DIFS, is the time interval between reception of a data frame and transmission of its ACK frame. Using this small gap between transmissions within the frame exchange sequence prevents other stations — which are required to wait for the medium to be idle for a longer gap (e.g., at least DIFS time) — from attempting to use the medium, thus giving priority to completion of the in-progress frame exchange sequence. Moreover, the DCF defines an optional mechanism, which requires that the transmitter and receiver exchange short Request-To-Send (RTS) and Clear-To-Send (CTS) control frames prior to the actual data frame transmission.

B. Backoff Behavior of IEEE 802.11 DCF

The DCF adopts a *slotted binary exponential backoff* mechanism to select the random backoff interval (in number of *time slots*). This random number is drawn from a uniform distribution over the interval $[0, CW-1]$, where CW is the contention window size and its initial value is $aCWmin$. In the case of an unsuccessful transmission, CW is doubled. Once CW reaches $aCWmax$, it will remain at this value until it is reset to $aCWmin$. In the case of a successful transmission, the CW value is reset to $aCWmin$ before the random backoff interval is selected. Each station decrements its backoff counter as long as the wireless medium is sensed to be idle for at least DIFS time. If the counter has not reached zero and the medium becomes busy again, the station freezes its counter. When the counter finally reaches zero, the station starts its transmission.

Fig. 1 illustrates such an operation of decrementing the backoff counter. After the successful transmission and acknowledgment of frame A1, station A waits for DIFS time and selects a backoff interval equal to 6, before attempting to transmit the

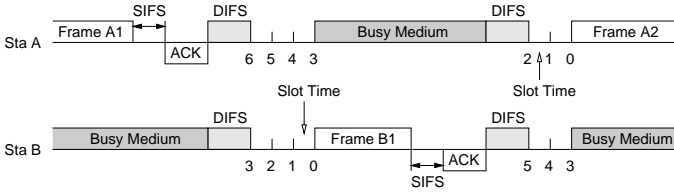


Fig. 1. An example of data frame transmissions and backoff decrements

next frame A2. Assume that station B selects a smaller backoff interval equal to 3 after it has sensed the medium to be idle for DIFS time. Since the backoff counter of station B reaches zero before that of station A, frame B1 is transmitted. As a result of the medium sensed busy, the backoff counter of station A is frozen at 3, and decrements again after the medium is sensed idle for DIFS time.

C. Assumptions

In this paper, the fairness and throughput performances of an IEEE 802.11 WLAN are analyzed under the assumption of ideal channel conditions (i.e., no transmission errors and no hidden terminals). Besides, although the analysis and conclusions presented in this paper do not depend on the technology adopted at the physical layer (PHY), the PHY does determine some network parameters, such as SIFS and DIFS. In the simulation, we assume that each wireless station operates at the IEEE 802.11a PHY mode-8, and the related network parameters are summarized in Table I. For more details of the IEEE 802.11a PHY, refer to [12] and [13].

TABLE I

IEEE 802.11a PHY MODE-8 PARAMETERS AND ADDITIONAL NETWORK PARAMETERS USED IN THE SIMULATION

Parameters	Value	Comments
$r_{Transmit}$	54 Mbps	data transmission rate
BpS	27	bytes per OFDM symbol
$t_{SlotTime}$	9 μs	an idle slot time
$t_{SIFSTime}$	16 μs	SIFS time
$t_{DIFSTime}$	34 μs	DIFS = SIFS + 2 \times Slot
$t_{PropDelay}$	1 μs	propagation delay
$a_{RTSLength}$	20 octets	RTS frame length
$a_{CTSLength}$	14 octets	CTS frame length
$a_{Payload}$	2304 octets	data frame payload length
$a_{ACKLength}$	14 octets	ACK frame length
$a_{MACOverhead}$	28 octets	MAC layer overhead
$t_{PHYOverhead}$	20 μs	PHY layer overhead
t_{Symbol}	4 μs	OFDM symbol interval

III. PRIORITY-BASED FAIR MEDIUM ACCESS CONTROL

Our proposed MAC protocol is motivated by [14], where a simple analytical model is presented to compute the saturation

throughput of the IEEE 802.11 DCF. From this paper, we have an interesting and important observation: *if the number of contending stations within an IEEE 802.11 WLAN is known, then by setting the probability that a station transmits in a randomly-chosen time slot to an optimal value — which is a function of the number of contending stations, the aggregate throughput can be maximized.* In other words, by simply changing the backoff rule, the maximum aggregate throughput won't be improved. The critical factor of improving the maximum aggregate throughput is not how to design a new backoff rule, but how to adjust the parameters of a backoff scheme so that each station can be tuned to run at its *optimal* point. For example, in the binary exponential backoff scheme proposed in the IEEE 802.11 DCF, two parameters (aCW_{min} and aCW_{max}) need to be adjusted, while in the uniform backoff scheme proposed in [14], only one parameter (CW_{opt}) needs to be adjusted. However, the author of [14] did not consider the fairness issue. For simplicity, we assume a uniform backoff scheme in our protocol. A similar analysis can also be applied to the binary exponential backoff scheme, although it is much more complicated.

Clearly, to achieve the desired fairness among traffic flows, the stations carrying higher-weight traffic flows should be assigned higher priorities for their accesses to the shared wireless medium. The basic idea of our protocol is that the prioritized access to the wireless medium is controlled through different backoff intervals. The shorter backoff time a station waits, the higher priority it will get. For this reason, our proposed MAC protocol is called *priority-based fair medium access control* (P-MAC), in which the contention window size for each station is properly selected to (1) reflect the relative weights among traffic flows, so as to achieve the weighted fairness; and (2) reflect the number of stations contending for the wireless medium, so as to maximize the aggregate throughput.

A. Fairness and Throughput Analysis

Consider the scenario when there are n greedy contending stations — stations that always have packets to transmit — in the network. Assume that one of the contending stations, $u \in f_i$, uses a contention window CW_u to access the wireless medium,³ and initially, its backoff value $b_u(t)$ is uniformly selected from the range $[0, CW_u - 1]$. As illustrated in Fig. 1, $b_u(t)$ is decremented at the end of each time slot, which could be either an idle period of length $t_{SlotTime}$, or a busy period due to a collision, or a busy period due to a successful frame transmission. Note that t is a discrete time point corresponding to the end of a time slot. Then, as indicated in [15], the stochastic process $b_u(t)$ can be modeled by the following discrete-time Markov chain:

$$P\{b_u(t+1) = k\} =$$

³Since we assume a uniform backoff in our study, station u will use this contention window CW_u to select the backoff intervals for all of its frame transmission attempts.

$$\left\{ \begin{array}{ll} P\{b_u(t) = k + 1\} + \frac{P\{b_u(t)=0\}}{CW_u} & \text{for } k = 0, \dots, CW_u - 2, \\ \frac{P\{b_u(t)=0\}}{CW_u} & \text{for } k = CW_u - 1, \end{array} \right. \quad (3)$$

where the term $P\{b_u(t) = k + 1\}$ corresponds to decrementing the backoff value at the end of each time slot. The term $\frac{P\{b_u(t)=0\}}{CW_u}$ accounts for the fact that, after a frame transmission attempt, the new backoff value is uniformly selected from the range $[0, CW_u - 1]$, regardless whether the frame transmission was successful or not. The steady state probabilities of this Markov chain are:

$$\lim_{t \rightarrow \infty} P\{b_u(t) = k\} = \frac{2 \cdot (CW_u - k)}{CW_u \cdot (CW_u + 1)}. \quad (4)$$

For more details of this Markov chain, see [15].⁴ Recall that, when the backoff counter finally reaches zero, the station starts its transmission. Therefore, the probability that station u transmits in a randomly-chosen time slot is

$$p_u = \lim_{t \rightarrow \infty} P\{b_u(t) = 0\} = \frac{2}{CW_u + 1}. \quad (5)$$

The probability that at least one station attempts to transmit in a slot, or equivalently, the probability that a slot is not idle, is given by

$$P_{tr} = 1 - \prod_{\forall v} (1 - p_v), \quad (6)$$

and the probability that a transmission is successful and is from station u can be calculated as

$$SU_u = p_u \cdot \prod_{v \neq u} (1 - p_v). \quad (7)$$

To achieve the desired fairness among the stations carrying the same traffic class, say, class i , we must have

$$\begin{aligned} \forall u, v \in f_i, \quad & \frac{SU_u}{\phi_i} = \frac{SU_v}{\phi_i} \\ \iff \forall u, v \in f_i, \quad & p_u(1 - p_v) = p_v(1 - p_u) \\ \iff \forall u, v \in f_i, \quad & p_u = p_v \\ \iff \forall u, v \in f_i, \quad & CW_u = CW_v. \end{aligned} \quad (8)$$

The interpretation of Eq. (9) is trivial: in order to be fair to the traffic flows of the same priority, the source stations should use the same contention window size. Now, we use p_i and SU_i to denote the probability that a station carrying class- i traffic transmits in a time slot, and the probability that a transmission is successful and is from a station carrying class- i traffic, respectively. We can rewrite Eqs. (6) and (7) as

$$P_{tr} = 1 - \prod_{i=1}^n (1 - p_i)^{|f_i|}, \quad (10)$$

and

$$SU_i = p_i \cdot (1 - p_i)^{|f_i|-1} \cdot \prod_{j \neq i} (1 - p_j)^{|f_j|}. \quad (11)$$

Similarly, to achieve the desired fairness among the stations carrying different traffic classes, we must have

$$\begin{aligned} \forall i, j \in \{1, \dots, n\}, \quad & \frac{SU_i}{\phi_i} = \frac{SU_j}{\phi_j} \\ \iff \forall i, j \in \{1, \dots, n\}, \quad & \frac{p_i(1 - p_j)}{\phi_i} = \frac{p_j(1 - p_i)}{\phi_j} \\ \iff \forall j \in \{2, \dots, n\}, \quad & p_j = \frac{p_1}{\frac{1-p_1}{\phi_j} + p_1} \end{aligned} \quad (12)$$

$$\iff \forall j \in \{2, \dots, n\}, \quad CW_j = \frac{CW_1 - 1}{\phi_j} + 1. \quad (13)$$

As expected, if a wireless station carries a low-priority traffic flow, it should use a larger contention window to access the wireless medium, thus favoring the higher-weight traffic flows. Based on the above analysis, we draw the following conclusion: *if the contention window sizes of all the stations are carefully selected to satisfy both Eq. (9) and Eq. (13), the desired weighted fairness among traffic flows can be achieved.*

Since the probability that a transmission is successful is given by

$$SU = \sum_{i=1}^n |f_i| \cdot p_i \cdot (1 - p_i)^{|f_i|-1} \cdot \prod_{j \neq i} (1 - p_j)^{|f_j|}, \quad (14)$$

and $(1 - P_{tr})$ is the probability of an idle slot, the aggregate throughput can be calculated as

$$\mathcal{T} = \frac{SU \cdot aPayload}{SU \cdot \ell_{succ} + (P_{tr} - SU) \cdot \ell_{coll} + (1 - P_{tr}) \cdot tSlotTime}, \quad (15)$$

where $aPayload$ is the data frame payload length, ℓ_{succ} is the length of a successful frame transmission, ℓ_{coll} is the collision length, and $tSlotTime$ is the length of an idle time slot. Eq. (15) can be rewritten as

$$\mathcal{T} = \frac{aPayload}{\ell_{succ} - \ell_{coll} + \frac{P_{tr} \cdot \ell_{coll} + (1 - P_{tr}) \cdot tSlotTime}{SU}}. \quad (16)$$

Since $aPayload$, ℓ_{succ} , ℓ_{coll} , and $tSlotTime$ are constant for all the stations, maximization of the aggregate throughput is equivalent to maximization of

$$\mathcal{T}' = \frac{SU}{P_{tr} \cdot \ell_{coll} + (1 - P_{tr}) \cdot tSlotTime}. \quad (17)$$

Note that the collision lengths — refer to [13] for calculation details — in the cases with and without RTS/CTS support are given by

$$\begin{aligned} \ell_{coll(basic)} &= tPHYOverhead \\ &+ \left[\frac{aMACOverhead + aPayload + 2.75}{BpS} \right] \cdot tSymbol \\ &+ tPropDelay + tDIFSTime, \end{aligned} \quad (18)$$

⁴The following fairness analysis is new, although it is based on the discrete-time Markov chain model described in [15].

and

$$\begin{aligned} \ell_{coll(rts/cts)} &= tPHYOverhead \\ &+ \left[\frac{aRTSLength + 2.75}{BpS} \right] \cdot tSymbol \\ &+ tPropDelay + tDIFSTime, \end{aligned} \quad (19)$$

respectively. It is clear from Eqs. (17), (10) and (5) that \mathcal{T}' depends on CW_j ($j = 1, \dots, n$) via p_j . Since for each $j \in \{2, \dots, n\}$, p_j is a function of p_1 , we can first solve the following equation to get the optimal value of p_1 to maximize the aggregate throughput:

$$\frac{d\mathcal{T}'}{dp_1} = 0, \quad (20)$$

and then apply Eqs. (12) and (5) to obtain the optimal values of p_j 's and the corresponding contention window sizes. However, it is difficult to solve Eq. (20) directly to get a closed-form expression for the optimal p_1 when $n \geq 3$. Therefore, we will instead give an approximate solution.

Let's first look at the simplest case where there are only two stations in the wireless network and each station carries a traffic flow with a different weight. Eq. (17) is simplified to

$$\mathcal{T}'_{(simple)} = \frac{p_1 \cdot (1 - p_2) + p_2 \cdot (1 - p_1)}{\ell_{coll} [1 - (1 - p_1)(1 - p_2)] + tSlotTime(1 - p_1)(1 - p_2)}, \quad (21)$$

and the optimal p_1 can be derived as

$$p_{1(simple)}^* = \frac{1}{1 + \sqrt{\frac{\ell_{coll}}{tSlotTime} \cdot \phi_2}}. \quad (22)$$

In particular, if $\phi_2 = 0$, then we have $p_1^* = 1$ and $CW_1^* = 1$, which means that the only active station in the network transmits continuously without any backoff, thus maximizing the bandwidth utilization. Now, consider the general case. Notice that, under the assumption of " $CW_1 \gg |f_j|\phi_j$ for each $j \in \{1, 2, \dots, n\}$ ", $|f_j|p_j = \frac{2|f_j|}{CW_{j+1}} \approx \frac{2|f_j|\phi_j}{CW_1} \ll 1$, we can make the following approximations:

$$\frac{\sum_{j=2}^n |f_j|p_j \cdot (1 - |f_1|p_1)}{\sum_{j=2}^n \frac{|f_j|}{|f_1|}\phi_j} \approx \frac{|f_1|p_1 \cdot (1 - \sum_{j=2}^n |f_j|p_j)}{\phi_1}, \quad (23)$$

and Eq. (24) (on next page). Comparing Eq. (24) with Eq. (21), we can get an approximate optimal value of p_1 :

$$p_1^* \approx \frac{1}{|f_1| + \sqrt{\frac{\ell_{coll}}{tSlotTime} \cdot |f_1| \cdot (\sum_{j=2}^n |f_j|\phi_j)}}. \quad (25)$$

As a result, the optimal contention window size for the stations carrying class-1 traffic can be approximated as:

$$CW_{1(I)}^* \approx 2|f_1| + 2\sqrt{\frac{\ell_{coll}}{tSlotTime} \cdot |f_1| \cdot \left(\sum_{j=2}^n |f_j|\phi_j\right)} - 1. \quad (26)$$

We call Eq. (26) the *approximation function I*. Recall that this approximate solution is obtained under the assumption of " $CW_1 \gg |f_j|\phi_j$ ". Suppose that we define " \gg " as "ten times greater than", then for $CW_{1(I)}^*$ to be a good approximation, we require

$$CW_{1(I)}^* \gg |f_1|, \quad (27)$$

or equivalently,

$$\sum_{j=2}^n |f_j|\phi_j \geq \frac{16 tSlotTime}{\ell_{coll}} \cdot |f_1|. \quad (28)$$

In other words, when $|f_j|$ or ϕ_j is small, the approximation function I is expected to *not* perform well. Besides, by carefully examining the above approximation process, we can see that $CW_{1(I)}^*$ is actually obtained by replacing the contending stations with two virtual stations with probabilities $|f_1|p_1$ and $\sum_{j=2}^n |f_j|p_j$, respectively, to attempt a frame transmission in a randomly-chosen time slot. As a result, this approximation process always tends to select smaller contention window sizes than the actual optimal values. Therefore, any other approximate value that is less than $CW_{1(I)}^*$ is even farther away from the actual optimal value, and hence not acceptable.

On the other hand, when $|f_j| = 0$ for each $j \in \{2, \dots, n\}$, a close approximation to the optimal value of CW_1 is given in [15] as:

$$CW_{1(II)}^* \approx |f_1| \cdot \sqrt{\frac{2\ell_{coll}}{tSlotTime}}, \quad (29)$$

and we call it the *approximation function II*. Therefore, when the traffic on the wireless medium is dominated by class-1 traffic, we may use this approximation instead, but we require

$$CW_{1(II)}^* > CW_{1(I)}^*, \quad (30)$$

or equivalently,

$$\sum_{j=2}^n |f_j|\phi_j < \left[\frac{1}{2} + \frac{tSlotTime}{\ell_{coll}} - \sqrt{\frac{2 tSlotTime}{\ell_{coll}}} \right] \cdot |f_1|. \quad (31)$$

Finally, based on the above analysis, we combine the two approximation functions and propose the following approximation to the optimal value of CW_1 , which depends on the traffic scenario on the wireless medium:

$$CW_1^* \approx \begin{cases} 2|f_1| + 2\sqrt{\frac{\ell_{coll}}{tSlotTime} \cdot |f_1| \cdot (\sum_{j=2}^n |f_j|\phi_j)} - 1, & \text{if } \sum_{j=2}^n |f_j|\phi_j \geq \frac{|f_1|}{r}, \\ |f_1| \cdot \sqrt{\frac{2\ell_{coll}}{tSlotTime}}, & \text{otherwise,} \end{cases} \quad (32)$$

where

$$r = \max \left(\frac{\ell_{coll}}{16 tSlotTime}, \frac{1}{\frac{1}{2} + \frac{tSlotTime}{\ell_{coll}} - \sqrt{\frac{2 tSlotTime}{\ell_{coll}}}} \right).$$

$$\mathcal{T}' \approx \frac{|f_1|p_1 \cdot \left(1 - \sum_{j=2}^n |f_j|p_j\right) + \sum_{j=2}^n |f_j|p_j \cdot (1 - |f_1|p_1)}{\ell_{coll} \cdot \left[1 - (1 - |f_1|p_1) \left(1 - \sum_{j=2}^n |f_j|p_j\right)\right] + tSlotTime \cdot (1 - |f_1|p_1) \left(1 - \sum_{j=2}^n |f_j|p_j\right)}. \quad (24)$$

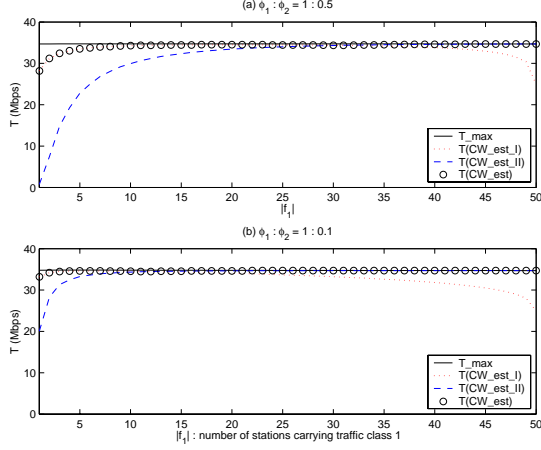


Fig. 2. Throughput performance under the 2-traffic-class scenario

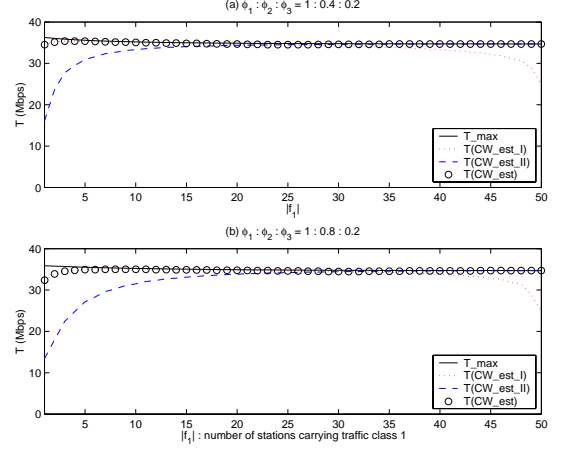


Fig. 3. Throughput performance under the 3-traffic-class scenario

Our approximation scheme is evaluated numerically. Fig. 2 shows the throughput performance when there are only two traffic classes on the wireless medium without RTS/CTS support. Assume that there are a total of 50 stations in an IEEE 802.11 WLAN, and each station carries a traffic flow belonging to either of the two traffic classes. Let the values along the X-axis represent $|f_1|$, the number of stations carrying traffic class-1. The solid lines in the figure represent the maximum achievable throughputs. As expected, the approximation function I (dotted lines in the figure) is not well-behaved when $|f_1|$ is large (hence $|f_2|$ is small), because the “ $CW_1 \gg |f_1|$ ” assumption does not hold in this case. Furthermore, by comparing the dotted lines in the two sub-figures, we can see that the throughput performance of the approximation function I is even worse with a smaller ϕ_2 value. Fortunately, due to the capability of adaptively switching from the approximation function I to the approximation function II (dashed lines in the figure), our final approximation (circles in the figure) works well in this range. However, when $|f_1|$ is small (e.g., ≤ 5 in this figure), noticeable gaps can be observed between our approximated maximum aggregate throughputs and the actual values. The reason for this is that, when $|f_1|$ is small, the “ $CW_1 \gg |f_2|\phi_2$ ” assumption does not hold. In general, if $|f_1|$ is small, then since CW_1 is not significantly larger than $\sum_{j=2}^n |f_j|\phi_j$ (based on Eq. (32)), the “ $CW_1 \gg |f_j|\phi_j$ ” assumption must not hold for some $j \in \{2, \dots, n\}$. Therefore, similar gaps can be observed under all the traffic scenarios (see Fig. 3, for example).

Fig. 3 illustrates the throughput performance of our approximation scheme when there are three traffic classes on the wireless medium without RTS/CTS support. Assume that there are an equal number of stations carrying class-2 and class-3 traffic, respectively. Notice that all the observations from Fig. 2 also

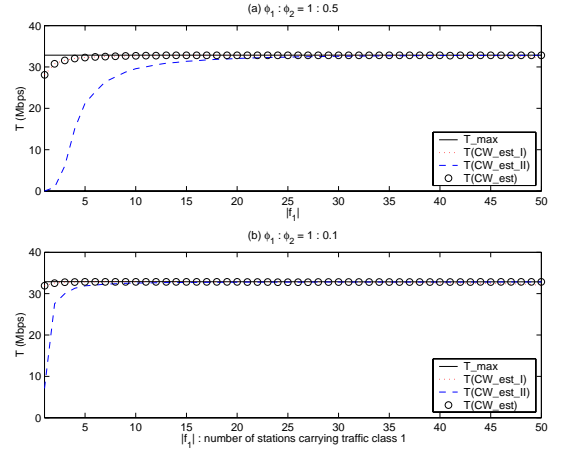


Fig. 4. Throughput performance under the 2-traffic-class scenario with RTS/CTS support

hold under this new traffic scenario. Figs. 4 and 5 show the results for both traffic scenarios, respectively, when the RTS/CTS mechanism is adopted. We have two observations. First, since the introduction of the RTS/CTS mechanism significantly reduces the collision length, i.e., $\ell_{coll}(rts/cts) \ll \ell_{coll}(basic)$, the deviation of our approximate contention window sizes won't result in as big gaps as those without RTS/CTS support. We can see that even the approximation function I itself performs well for most network configurations.⁵ Second, the maximum achievable throughputs (solid lines in the figures) are smaller than those without RTS/CTS support, which is due to the added

⁵In Figs. 4 and 5, most parts of the dotted lines (approximation function I) and the circles (final approximation function) are overlapped with each other and are close to the solid lines (maximum achievable throughputs).

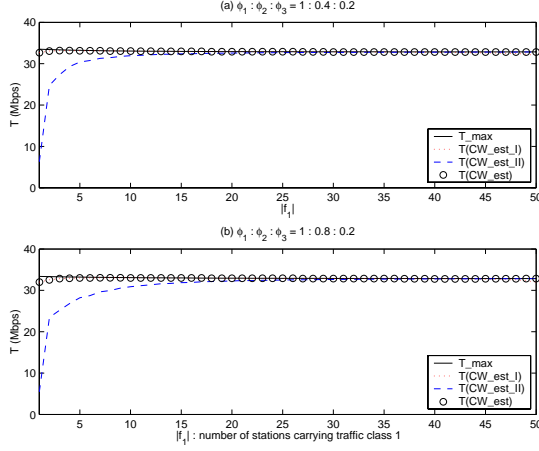


Fig. 5. Throughput performance under the 3-traffic-class scenario with RTS/CTS support

overhead of the RTS/CTS frame exchanges. Actually, the effectiveness of the RTS/CTS mechanism depends on the length of the data frame payload. The longer the payload length, the more effective the mechanism will be. Note, however, that the tradeoff between advantages and disadvantages of the RTS/CTS mechanism is not the focus of this work.

B. The Proposed P-MAC Protocol

Our proposed priority-based fair medium access control (P-MAC) protocol is similar to the IEEE 802.11 DCF, except that the slotted binary exponential backoff mechanism is replaced by a new scheme to determine the optimal backoff values. In P-MAC, the contention window size for each wireless station is properly selected to (1) reflect the relative weights among data traffic flows, so as to achieve the weighted fairness; and (2) reflect the number of stations contending for the wireless medium, so as to maximize the aggregate throughput. The details of P-MAC are presented as follows.

P-MAC requires that each wireless station keeps sensing the channel and monitoring the activities on the wireless medium when it is not transmitting. Therefore, each station knows whether at each time slot the wireless medium is busy or idle, whether a busy period corresponds to a collision or not, and which traffic class a successfully-transmitted frame belongs to. Let avg_idle and avg_wait_i ($i = 1, 2, \dots, n$) denote the average number of consecutive idle slots on the wireless medium and the average number of time slots between two consecutive successful class- i frame transmissions, respectively, and they can be calculated as

$$avg_idle = \frac{1}{P_{tr}} - 1, \quad (33)$$

and

$$avg_wait_i = \frac{1}{P_{tr(i)}} - 1. \quad (34)$$

Here, P_{tr} is the probability that at least one station attempts to transmit in a slot, and it is given by Eq. (10). $P_{tr(i)}$ is the

probability of a successful frame transmission from any station that carries class- i traffic, and it is given by

$$P_{tr(i)} = |f_i| \cdot p_i \cdot (1 - p_i)^{|f_i|-1} \cdot \prod_{j \neq i} (1 - p_j)^{|f_j|}, \quad (35)$$

Notice that we have the following relation:

$$\frac{1 - P_{tr}}{P_{tr(i)}} = \frac{1 - p_i}{|f_i| p_i}, \quad (36)$$

and hence,

$$|f_i| = \frac{(CW_i - 1) \cdot (avg_idle + 1)}{2 \cdot avg_idle \cdot (avg_wait_i + 1)}. \quad (37)$$

Based on the measurements of avg_idle and avg_wait_i by monitoring the medium activities, each station can estimate the values of $|f_i|$'s using Eq. (37) and then determine the optimal contention window sizes using Eqs. (32) and (13).

Fig. 6 (on next page) shows the pseudo-coded algorithm executed by each station to adjust its contention window size. The number of traffic classes (n) and the associated weight for each class (ϕ_i) are assumed to be available *a priori* to each station. Each station maintains a set of random variables, "IDLE" and "WAIT(i)" for each $i \in \{1, 2, \dots, n\}$. The contention window size for the stations carrying class-1 traffic, $cw(1)$, is initialized to cw_start , a design parameter. Let an idle-busy-cycle be the time interval between the ends of two adjacent busy periods on the wireless medium. The observation window size w_{obs} , another design parameter, represents the number of idle-busy-cycles within which the measurements of avg_idle and avg_wait_i are taken, and the count w_{count} for monitored idle-busy-cycles is reset to 0. As shown in the pseudo-code, "IDLE" is updated after each idle-busy-cycle, while "WAIT(i)" is updated only if the busy period corresponds to a successful class- i frame transmission. At the end of each observation window, the values of $|f_i|$'s are estimated, and the contention window sizes are adjusted according to these estimates. Finally, based on the traffic class a station is carrying, it can determine which contention window size to use for its next frame transmission attempt. Notice that t_{curr} and $t_{prev}(j)$ are the discrete time points measured in time slots, and α and β are both smoothing factors.

P-MAC is intended to achieve efficient channel utilization and weighted fairness for data communications in a WLAN. It is particularly effective when the network is saturated and multiple data traffic flows (e.g., long-lived FTP sessions) with different weights are contending for the shared wireless medium. However, P-MAC may not be fair in terms of delay/jitter, since the delay/jitter requirements of traffic flows are not taken into consideration in the protocol design.

IV. PERFORMANCE EVALUATION AND DISCUSSION

In this section, we evaluate the effectiveness of the proposed P-MAC scheme by simulation. The network parameters used in

```

01.  $cw(1) := cw\_start; w\_count := 0;$ 
02. for ( $i = 2 : n$ )
03.    $cw(i) := \lceil cw(1) - 1 \rceil / \phi_i + 1;$ 
04. while (the wireless network is alive) {
05.   if (a new idle-busy-cycle has been monitored on the wireless medium) then {
06.      $w\_count := w\_count + 1;$ 
07.      $IDLE := \alpha \times IDLE + (1 - \alpha) \times new\_idle;$ 
08.     if (the busy period is due to a successful frame transmission) then {
09.        $j :=$  the traffic class of the delivered frame;
10.        $WAIT(j) := \alpha \times WAIT(j) + (1 - \alpha) \times [t_{curr} - t_{prev}(j) - 1];$ 
11.        $t_{prev}(j) := t_{curr};$ 
12.     }
13.     if ( $w\_count = 0 \pmod{w_{obs}}$ ) then {
14.       for ( $i = 1 : n$ )
15.          $|f_i|_{est} = \beta \times |f_i|_{est} + (1 - \beta) \times \frac{(cw(i)-1) \times (IDLE+1)}{2 \times IDLE \times (WAIT(i)+1)};$ 
16.         ( calculate  $cw(1)$  using Eq. (32) based on the new  $|f_i|_{est}$  values )
17.         for ( $i = 2 : n$ )
18.            $cw(i) := \lceil cw(1) - 1 \rceil / \phi_i + 1;$ 
19.           ( reset  $IDLE$ ,  $WAIT(i)$ , and  $t_{prev}(i)$  )
20.         }
21.       }
22.     }

```

Fig. 6. The algorithm executed by each station to adjust the contention window size

the simulation are listed in Table I, and the RTS/CTS mechanism is not used. P-MAC is compared against the DCF of the IEEE 802.11 MAC, which uses the slotted binary exponential backoff mechanism with $aCWmin = 16$ and $aCWmax = 1024$, as specified in [12], in terms of fairness and throughput. In the simulation runs of P-MAC, the observation window w_{obs} is set to 50, while the smoothing factors α and β are chosen to be 0.9 and 0.8, respectively. The duration of each simulation run is 6 minutes unless specified otherwise.

First, P-MAC is compared with the IEEE 802.11 DCF under the single-class traffic scenario. Let m denote the number of active stations in the network, and each station carries a class-1 data traffic flow. Each flow is backlogged throughout the simulation duration. The throughputs for all the flows in the case of $m = 5, 15$, and 50 are shown (on next page) in Figs. 7(a), (b), and (c), respectively. Clearly, P-MAC shows much better fairness than the IEEE 802.11 DCF, regardless of the network configuration.

In order to evaluate the fairness of a MAC scheme quantitatively, we introduce a new measure called the *fairness index* as follows. Let \mathcal{T}_f denote the throughput of traffic flow f , and ϕ_f be the associated weight. The fairness index, \mathcal{F} , is then defined as:

$$\mathcal{F} = \frac{\mu(\mathcal{T}_f / \phi_f)}{\mu(\mathcal{T}_f / \phi_f) + \sigma(\mathcal{T}_f / \phi_f)}, \quad (38)$$

where μ and σ are, respectively, the mean and the standard deviation of \mathcal{T}_f / ϕ_f over all the active data traffic flows. When the perfect fairness is achieved, the ratio \mathcal{T}_f / ϕ_f is the same for all flows, and the fairness index is equal to 1. In general, the fair-

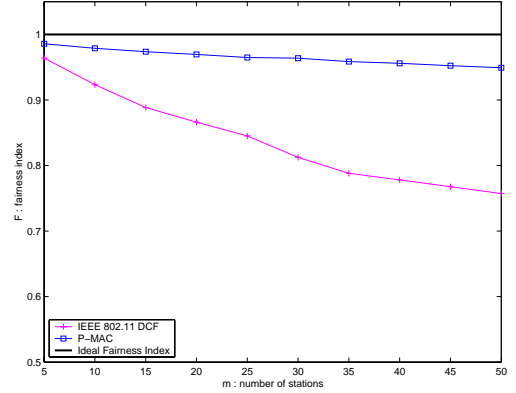


Fig. 8. Comparison of P-MAC and the IEEE 802.11 DCF using fairness index

ness index is a real value between 0 and 1, and the closer to 1 the fairness index, the fairer. Now, the fairness and throughput of P-MAC and the IEEE 802.11 DCF are compared for different network sizes, and the results are plotted in Figs. 8 and 9, respectively. The ideal fairness index and the maximum achievable throughput are shown as the thick solid lines in the figures. We can see that P-MAC achieves excellent fairness, while the fairness of the IEEE 802.11 DCF is generally much poorer. This is consistent with the previous observations in Fig. 7. On the other hand, the IEEE 802.11 DCF achieves lower aggregate throughput, and the throughput performance is even worse when the network size increases. This is due to the increased frequency of frame collisions as a result of more contending stations in the

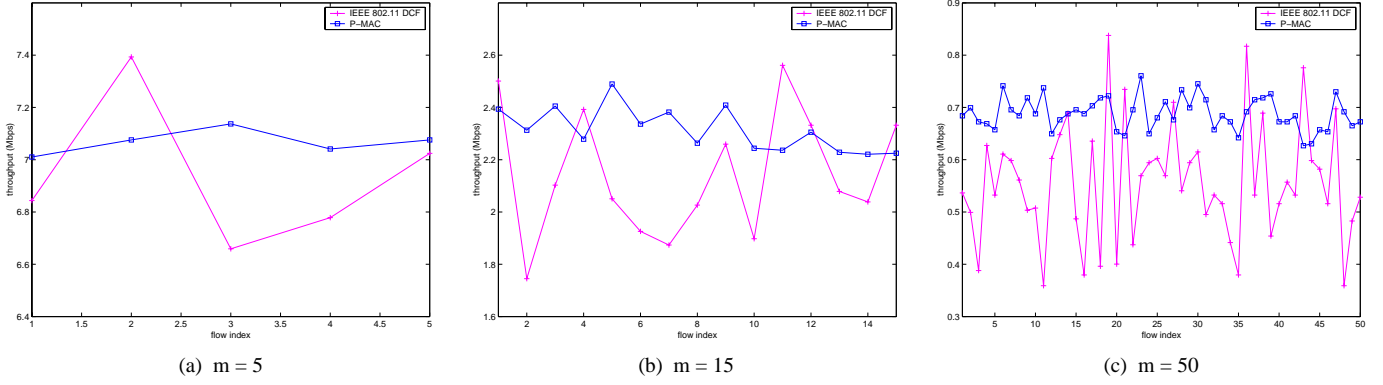


Fig. 7. Comparison of P-MAC and the IEEE 802.11 DCF

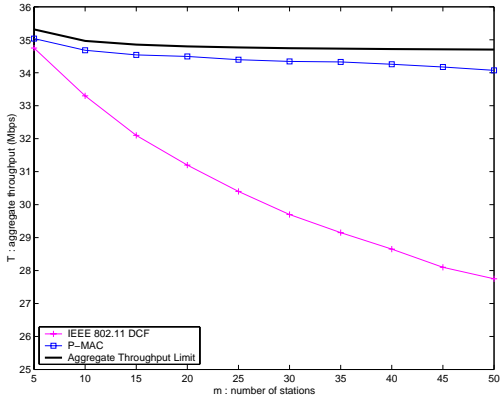


Fig. 9. Comparison of P-MAC and the IEEE 802.11 DCF using aggregate throughput

network. In contrast, since P-MAC always tends to select the contention window size carefully to keep a very low frame collision probability, it can still maintain a high aggregate throughput close to the theoretical limit, even when the number of the contending stations is large.

Second, the performance of P-MAC is evaluated under multiple-traffic-class scenarios. There are a total of 50 active stations in the network, and each station carries a data traffic flow.⁶ Again, each flow is backlogged throughout the simulation duration, and the results are plotted in Figs. 10 and 11. We can see that P-MAC shows very good fairness ($\mathcal{F} > 0.9$) and high aggregate throughput for both 2-traffic-class and 3-traffic-class scenarios. Recall that the selection of the contention window size in P-MAC is based on an approximation scheme that does not work well when $|f_i|$ is small. Therefore, we can observe relatively larger gaps between the throughput simulation results of P-MAC and the theoretical limits of the aggregate throughput, which are consistent with the analysis results shown in Figs. 2 and 3.

Third, in order to verify that P-MAC is well-behaved in presence of traffic fluctuations, we simulate a network with a vari-

⁶For the 3-traffic-class scenario, the numbers of stations carrying traffic class-2 and class-3 are set to be equal.

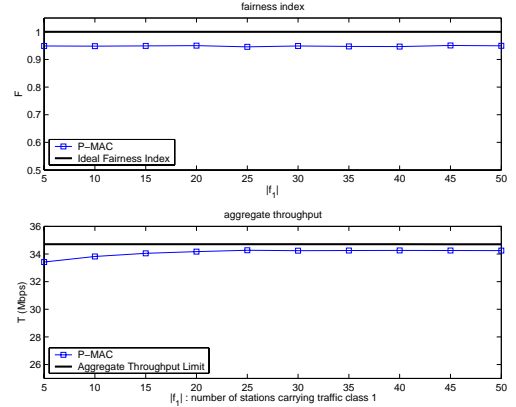


Fig. 10. P-MAC performance with 2 traffic classes: $\phi_1/\phi_2=1/0.5$

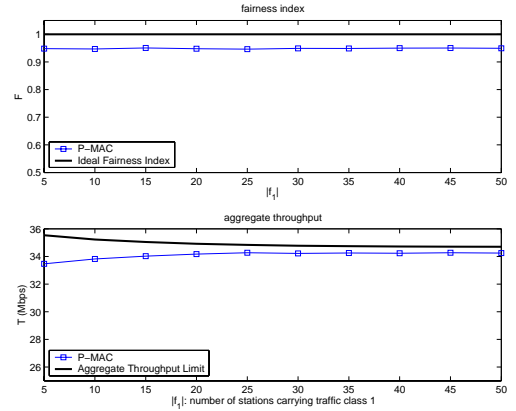


Fig. 11. P-MAC performance with 3 traffic classes: $\phi_1/\phi_2/\phi_3=1/0.8/0.2$

able number of active stations. The variation patterns of the active stations carrying different traffic classes are represented by the thick stair case curves in Fig. 12. For example, in the simulation period between 4 and 5 minutes, there are 25 class-1, 15 class-2, and 5 class-3 data traffic flows contending for the shared wireless medium. As shown in figure, the estimates of $|f_1|$, $|f_2|$, and $|f_3|$ are recorded by the solid curve, the dashed curve, and the dotted curve, respectively. Clearly, these estimates follow

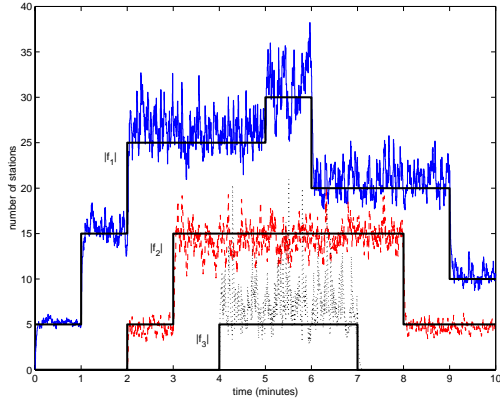


Fig. 12. An example of the on-line estimation in P-MAC ($\phi_1/\phi_2/\phi_3=1/0.8/0.2$)

the variation of the actual numbers of active stations. Observe that the estimate of $|f_3|$ is the most inaccurate among the three. By referring to the pseudo-code of P-MAC (shown in Fig. 6), this is because a small variation of WAIT(3) results in a large variation of $|f_3|_{est}$ due to the large value of $cw(3)$.

V. CONCLUSION AND FUTURE WORK

In this paper, we propose a priority-based fair medium access control (P-MAC) protocol, by modifying the IEEE 802.11 DCF, to maximize the wireless channel utilization subject to the weighted fairness among multiple data traffic flows. The basic idea of P-MAC is that the contention window size for each wireless station is properly selected to (1) reflect the relative weights among traffic flows, so as to achieve the weighted fairness; and (2) reflect the number of stations contending for the wireless medium, so as to maximize the aggregate throughput. Simulation results have shown that P-MAC outperforms the IEEE 802.11 DCF significantly in terms of both fairness and throughput.

Note that our fairness/throughput analysis presented in this paper is based on the assumption of no hidden stations, which is not always true in practice. In the presence of hidden stations, the wireless network becomes much more complicated and is very difficult to analyze. As part of future work, we plan to extend our P-MAC protocol to deal with hidden stations by including certain heuristic policies. Besides, although P-MAC achieves excellent fairness among data traffic flows in terms of bandwidth allocation, it may not be fair in terms of delay/jitter. This is because the delay/jitter requirements of traffic flows are not taken into consideration in the protocol design. We would like to address how to support real-time services in a DCF WLAN in the future.

REFERENCES

- [1] IEEE 802.11, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, Standard, IEEE, Aug. 1999.
- [2] T. S. Eugene Ng, Ion Stoica, and Hui Zhang, "Packet Fair Queueing Algorithms for Wireless Networks with Location-Dependent Errors," in *Proc. IEEE INFOCOM'98*, San Francisco, CA, Mar. 1998, pp. 1103–1111.
- [3] Songwu Lu, Thyagarajan Nandagopal, and Vaduvur Bharghavan, "A Wireless Fair Service Algorithm for Packet Cellular Networks," in *Proc. ACM MobiCom'98*, Dallas, TX, Oct. 1998, pp. 10–20.
- [4] Parameswaran Ramanathan and Prathima Agrawal, "Adapting Packet Fair Queueing Algorithms to Wireless Networks," in *Proc. ACM MobiCom'98*, Dallas, TX, Oct. 1998, pp. 1–9.
- [5] Thyagarajan Nandagopal, Songwu Lu, and Vaduvur Bharghavan, "A Unified Architecture for the Design and Evaluation of Wireless Fair Queueing Algorithms," in *Proc. ACM MobiCom'99*, Seattle, WA, Aug. 1999, pp. 132–142.
- [6] Vaduvur Bharghavan, Alan Demers, Scott Shenker, and Lixia Zhang, "MACAW: A Media Access Protocol for Wireless LAN's," in *Proc. ACM SIGCOMM'94*, London, England UK, Aug. 1994, pp. 212–225.
- [7] Timucin Ozugur, Mahmoud Naghshineh, Parviz Kermani, C. Michael Olsen, Babak Rezvani, and John A. Copeland, "Balanced Media Access Methods for Wireless Networks," in *Proc. ACM MobiCom'98*, Dallas, TX, Oct. 1998, pp. 21–32.
- [8] Dr-Jiunn Deng and Ruay-Shiung Chang, "A Nonpreemptive Priority-Based Access Control Scheme for Broadband Ad Hoc Wireless ATM Local Area Networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 9, pp. 1731–1739, Sep. 2000.
- [9] Nitin H. Vaidya, Paramvir Bahl, and Seema Gupta, "Distributed Fair Scheduling in a Wireless LAN," in *Proc. ACM MobiCom'00*, Boston, MA, Aug. 2000, pp. 167–178.
- [10] Thyagarajan Nandagopal, Tae-Eun Kim, Xia Gao, and Vaduvur Bharghavan, "Achieving MAC Layer Fairness in Wireless Packet Networks," in *Proc. ACM MobiCom'00*, Boston, MA, Aug. 2000, pp. 87–98.
- [11] Haiyun Luo, Songwu Lu, and Vaduvur Bharghavan, "A New Model for Packet Scheduling in Multihop Wireless Networks," in *Proc. ACM MobiCom'00*, Boston, MA, Aug. 2000, pp. 76–86.
- [12] IEEE 802.11a, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High-speed Physical Layer in the 5 GHz Band*, Supplement to IEEE 802.11 Standard, Sep. 1999.
- [13] Daji Qiao and Sunghyun Choi, "Goodput Enhancement of IEEE 802.11a Wireless LAN via Link Adaptation," in *Proc. IEEE ICC'01*, Helsinki, Finland, Jun. 2001.
- [14] Giuseppe Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
- [15] Giuseppe Bianchi, Luigi Fratta, and Matteo Oliveri, "Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs," in *Proc. IEEE PIMRC*, Taipei, Taiwan, Oct. 1996, pp. 407–411.