Downlink and Uplink Resource Allocation in IEEE 802.11 Wireless LANs

Sung Won Kim, Byung-Seo Kim, *Student Member, IEEE*, and Yuguang Fang, *Senior Member, IEEE*

***Abstract*—Wireless local area networks (WLANs) based on the IEEE 802.11 standard are becoming increasingly popular and widely deployed. It is likely that WLAN will become an important complementary technology for future cellular systems and will typically be used to provide hotspot coverage. In this paper, the complementary use of WLANs in conjunction with mobile cellular networks is studied. We identify the fairness problem between uplink and downlink traffic flows in the IEEE 802.11 distributed coordination function and then propose an easy solution that can be implemented at the access point (AP) in the MAC layer without modification of the standard for stations (STAs). This solution aims at providing a controllable resource-allocation method between uplink and downlink traffic flows and adapting the parameters according to the dynamic traffic load changes. The proposed solution also enhances the system utilization by reducing the probability of frame collision.**

***Index Terms*—Fairness, hotspot, IEEE 802.11 distributed coordination function (DCF), wireless local area network (WLAN).**

I. Introduction

EVEN WITH the emergence of the developed wireless cellular networks such as second-generation personal-communications service (PCS) systems, there is a need for another public wireless access solution to meet the ever-increasing demand for data-intensive applications and ubiquitous communications. To achieve these goals, future third-/fourth-generation (3G/4G) mobile networks will evolve to provide higher data rates and new radio access technologies. One widely anticipated form of evolution is the complementary use of so-called hotspots [1]–[3].

Wireless local area networks (WLANs) based on the IEEE 802.11 standard [4] are becoming increasingly prevalent for offices, public places, and homes. The focus is now turning to deploying these networks over hotspots such as airports, hotels, cafes, and other areas in which people can have untethered public accesses to the Internet. Low-cost and high-speed WLANs can be integrated within the cellular coverage to provide hotspot coverage for high-speed data services. WLAN offers an interesting possibility for cellular operators to offer additional capacity and higher bandwidth for end users without sacrificing the capacity of cellular users, since WLANs operate on unlicensed frequency bands.

The medium-access control (MAC) protocol in the IEEE 802.11 standard consists of two coordination functions: mandatory distributed coordination function (DCF) and optional point coordination function (PCF). In the DCF, a set of wireless stations (STAs) communicate with each other using a contention-based channel-access method, namely carrier sense multiple access with collision avoidance (CSMA/CA). CSMA/CA is known for its inherent fairness between STAs and robustness. It is quite effective in supporting symmetric traffic loads in *ad hoc* networks where the traffic loads between STAs are similar. However, this form of random-access protocol is not recommended for asymmetric traffic loads in hotspots where the traffic loads are converged into an access point (AP) [5]. There are many wireless multimedia applications in which the utilization of radio resource is strongly biased toward the downlink (from AP to STAs) against the uplink (from STAs to AP) or the direct link (from STAs to STAs). For example, let us consider Internet access or mobile computing. In general, the information database and computing power for multimedia services tend to locate at the network side rather than at the STAs.

By using the optional PCF of IEEE802.11 that is a polling-based service, it is possible to overcome the problems of asymmetric traffic loads. To do so, the PCF requires a centralized scheduling algorithm. Moreover, STAs need to inform their traffic load status to AP for the optimal scheduling algorithm. The scheduler and status reporting induce the modifications of the standard MAC protocol for both STAs and AP.

We focus on the DCF mode in IEEE 802.11 that is a contention-based channel-access method and is widely deployed in the world. We identify situations in which the traffic flows for the downlink is completely blocked due to the unfairness of the CSMA/CA MAC protocol in distributed environments. We then propose an easy solution that overcomes such problems. The proposed MAC layer algorithm can be implemented at an AP without modification at STAs. We show that our approach ensures an adequate level of fairness between the uplink and the downlink, and improves the system utilization.

The remainder of this paper is organized as follows. Section II presents related works and Section III describes the proposed MAC protocol. Section IV analyzes the performance of the proposed method. In Section V, we investigate the enhancement of the proposed method on the utilization and the fairness with some numerical results. Finally, this paper is concluded in Section VI.

Manuscript received May 17, 2004; revised July 28, 2004 and September 17, 2004. This work was supported in part by the U.S. Office of Naval Research under Grant N000140210464 (ONR Young Investigator Award) and in part by the Postdoctoral Fellowship Program, Korea Science and Engineering Foundation (KOSEF). The review of this paper was coordinated by Dr. Y.-B. Lin.

The authors are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: ksw@ieee.org; jsnbs@ufl.edu; fang@ece.ufl.edu).

Digital Object Identifier 10.1109/TVT.2004.838887

II. Related Works

A. DCF of IEEE 802.11

The DCF achieves automatic medium sharing between compatible STAs through the use of CSMA/CA. Before a STA starts transmission, it senses the wireless medium to determine if it is idle. If the medium appears to be idle, the transmission may proceed. Otherwise, the STA will wait until the end of the ongoing transmission.

A STA will ensure that the medium has been idle for the specified time interval before attempting to transmit. STAs deliver data frames with arbitrary lengths after sensing that the medium is idle for at least a DCF interframe space (DIFS), the required time interval in the standard. If two or more STAs find that the channel is idle at the same time, a collision occurs. In order to reduce the probability of such collisions, a STA has to perform a backoff procedure before starting a transmission. The duration of this backoff is determined by the contention window (CW) size, which is initially set to $CW_{\min}$. The CW value is used to randomly choose the number of slot times (σ) in the range of $[0, CW-1]$, which is used for backoff duration. In the case of an unsuccessful transmission, the CW value is updated to $CW \times 2$ while it does not exceed $CW_{\max}$. This will guarantee that, in case of a collision, the probability of another collision at the time of next transmission attempt is further decreased.

A transmitter and receiver pair exchanges short request-to-send (RTS) and clear-to-send (CTS) control frames prior to the actual data transmission to avoid the collision of data frames. An acknowledgment (ACK) frame will be sent by the receiver upon successful reception of a data frame. It is only after receiving an ACK frame correctly that the transmitter assumes successful delivery of the corresponding data frame. Short interframe space (SIFS), which is smaller than DIFS, is a time interval between RTS, CTS, data frame, and ACK frame. Using this small gap between transmissions within the frame-exchange sequence prevents other STAs from attempting to use the medium. As a consequence, it gives priority to completion of the ongoing frame exchange sequence.

B. EDCA of IEEE 802.11e

The IEEE 802.11 Working Group initiated a Study Group (SG11e) with the charter to enhance the 802.11 MAC [6], which would provide support for applications with quality-of-service (QoS) requirements, while maintaining backward compatibility with the 802.11 standard.

An enhanced version of DCF, called enhanced distributed channel access (EDCA), differentiates the channel access priority according to the QoS requirements. The IEEE 802.11e draft defines eight traffic categories (TCs) for priority-based traffic and each QoS-capable STA (QSTA) marks their frames to indicate a specific premium service requirement. The QoS-capable AP (QAP) can dynamically adjust the contention window parameters as well as the transmit opportunity (TxOP) limit for each traffic category.

EDCA provides relative QoS differentiation among traffic classes, but does not provide any QoS guarantee. EDCA is relatively simple, but the performance provided by this priority-based scheme is obviously less predictable than a reservation-based method. Due to the fact that a QSTA does not have to go through admission control to get bandwidth, it is hardly possible to reduce the amount of traffic that one STA might send. The QAP can only adjust the contention window and TxOP duration for each TC.

C. Fairness Issue in IEEE 802.11 WLANs

Fair-scheduling algorithms attempt to allocate network resource fairly to each traffic flow in proportion to a given flow weight. An initial study on fair scheduling among STAs is presented by Lu *et al.* [7]. The authors propose a fair queueing-scheduling algorithm performed at the AP, which relies on a TDMA-based MAC algorithm. Other fair-queueing solutions based on the WLAN MAC model appear in [8]–[11]. In these works, the weights are used to differentiate between traffic classes and apportion the bandwidth among them. The weight for high utilization class traffic is usually larger than that for low-utilization class traffic. The authors in [12]–[14] propose to scale the contention window, vary the interframe spacings, and change the backoff period according to the priority level of the traffic flow. Because all these studies are focused on the fairness or priority among STAs in a WLAN, the unfair sharing of bandwidth between the uplink and the downlink still remains.

In [15], the authors observe a significant unfairness between the uplink and the downlink flows when the DCF or EDCA is employed in a WLAN with AP. Since the DCF protocol allows equal access to the media for all hosts, the AP and the STAs have equal utilization to the medium. Thus, when the downlink has much more offered traffic load than the uplink, the downlink becomes a bottleneck of the system capacity and much more APs should be deployed to accommodate such STAs. EDCA allows QoS differentiation, which is an important improvement over legacy DCF. Nevertheless, it presents the same utilization asymmetry, giving an advantage to the uplink transmission as the number of STAs increases.

The TCP fairness issues between the uplink and downlink in WLANs has been studied in [16]. The authors are interested in a solution that results in uplink and downlink TCP flows having an equal share of the wireless bandwidth (utilization ratio of one). They identify the cause of unfairness in the buffer size availability at the AP and propose a solution based on TCP receiver window manipulation. Because this solution operates on the TCP layer, it is not effective when there exist traffic flows other than TCP.

III. Proposed MAC Protocol

A. Channel Access

We propose an improved MAC protocol based on the IEEE 802.11 WLAN standard. Each STA can communicate with the AP (uplink or downlink) or with other STAs (direct link). Since we focus on the resource allocation between uplink and downlink, we do not consider the direct link utilization in this paper, although it is noted that the utilization sharing between uplink and direct link is proportional to the ratio of the number of active STAs for the uplink and that for the direct link.

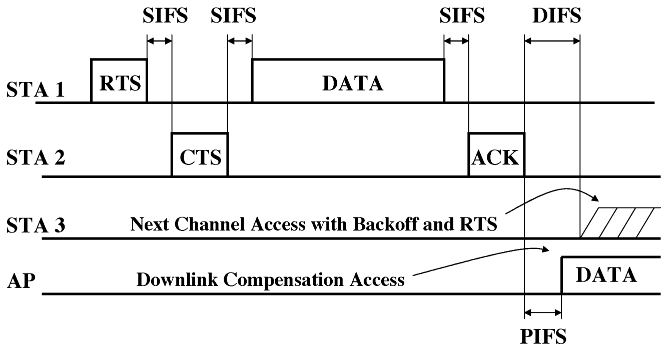


Fig. 1. DCF access mechanism and downlink compensation access mechanism.

STAs and AP use the DCF mechanism with RTS/CTS handshaking, as shown in Fig. 1, where the next channel access should wait for DIFS and backoff window time after a previous ACK frame. A two-way handshaking technique without RTS/CTS handshaking, called the basic access mechanism, is not considered in this paper, although our proposed method can be easily extended to the basic access mechanism.

All the time durations used for successful data transmission on the uplink (downlink) are called the *uplink (downlink) utilization*. The system utilization is the sum of the downlink and uplink utilization. We define the utilization ratio γ such that

$$\gamma = \frac{\text{downlink utilization}}{\text{uplink utilization}}. \tag{1}$$

To control the fair sharing of bandwidth between the uplink and downlink, we define a required utilization ratio as ψ. We say the system is in "fair state" if $\gamma = \psi$. If $\gamma < \psi$, there should be some compromise between the uplink and downlink. In that case, the AP can transmit data frames using point interframe space (PIFS) following the previous ACK frame until it becomes $\gamma \geq \psi$, as shown in Fig. 1. We call this mechanism downlink compensation access (DCA). During the DCA, the handshake mechanism of RTS and CTS is not necessary and the AP can transmit multiple data frames while $\gamma < \psi$. Note that the AP accesses the wireless channel without collision during the DCA because it transmits a data frame using PIFS, which is shorter than DIFS. Also note that AP accesses the wireless channel with the DCF when $\gamma \geq \psi$ or there is no ACK frame transmitted on the wireless channel. In this way, the required utilization ratio is maintained while the utilization increases due to the reduced probability of collision.

To keep up with the dynamic changes of traffic-load conditions, ψ and γ should be updated adaptively. We propose an update method for ψ and γ in the next section.

B. Parameter Update

To allocate the network resource according to the offered load, the value of ψ should be proportional to the offered load. Let ρ_u and ρ_d denote the time average of the accumulated offered load on the uplink and downlink, respectively. Under real situations, ρ_u and ρ_d can be a long-term average value, measured for a predefined duration. For example, the duration can be a daytime, working hour, or busy hour, according to the network-design criteria [17]. Based on these measurements, the AP may change the value of ψ (i.e., $\psi = \rho_d/\rho_u$) or network operator may send the value of ψ to the AP.

The AP can accurately measure ρ_d, since the AP transmits the downlink traffic. The AP estimates ρ_u according to the feedback informations from STAs that are transmitted by the control frames. To reduce the overhead caused by these control frames, we propose a simple update method for ψ.

We define a discrete time scale t_s that increments by one whenever there is a successful data-frame transmission. The AP has an internal memory that records the source MAC addresses of uplink data frames and the destination MAC addresses of downlink data frames for a time window of T_w. Let $\Psi_u(t_s)$ and $\Psi_d(t_s)$ denote the number of the source MAC addresses of the uplink traffic and the number of the destination MAC addresses of the downlink traffic between $(t_s - T_w)$ and t_s, respectively. For example, if an uplink (downlink) data frame is transmitted with a new source (destination) MAC address that is not registered in the AP, $\Psi_u(t_s)$ ($\Psi_d(t_s)$) increases by one and the MAC address is recorded in the AP. When the recorded MAC address becomes older than T_w, it is removed from the AP and $\Psi_d(t_s)$ or $\Psi_u(t_s)$ decreases by one. Thus, $\Psi_u(t_s)$ and $\Psi_d(t_s)$ represent the active number of STAs in an uplink and downlink for a period T_w, respectively. We update the required utilization ratio such that

$$\psi(t_s) = \frac{\Psi_d(t_s)}{\Psi_u(t_s)}. \tag{2}$$

Although this estimation does not exactly reflect the offered load, it is easy to be implemented in the AP and does not require the feedback information from STAs. Moreover, the proposed method does not require the modification of standard for STAs, which makes it compatible with the deployed STAs.

We propose a system parameter $\omega(t_s)$ that is a decision criterion for the DCA at time t_s. The initial value of $\omega(t_s)$ is set to zero, i.e., $\omega(0) = 0$. Let $\Gamma_d(t_s)$ and $\Gamma_u(t_s)$ be the lengths of data frames that are normalized to the maximum value of a data frame length and are successfully transmitted by the downlink and uplink at time t_s, respectively. Note that one of $\Gamma_d(t_s)$ and $\Gamma_u(t_s)$ is zero at t_s. The value of $\omega(t_s)$ is updated at every successful frame transmission such that

$$\omega(t_s) = \omega(t_s - 1) + \Gamma_d(t_s) - \psi(t_s)\Gamma_u(t_s). \tag{3}$$

Note that the case of $\omega(t_s) = 0$ is a fair state and the case of $\omega(t_s) < 0$ is the state that requires the DCA. Thus, we propose that the AP adopts $\omega(t_s)$ to decide the access method. If $\omega(t_s) < 0$ and there is an ACK frame transmitted on the channel, the AP uses the DCA whenever it has pending frames. Otherwise, the AP uses the DCF.

IV. Performance Analysis

A. Utilization of IEEE 802.11 DCF

To compare the performance of the proposed method with that of the IEEE 802.11 DCF, we adopt the analytical model used in [18]–[20]. The channel conditions are assumed to be ideal, i.e., no hidden terminals, no channel error, and no capture

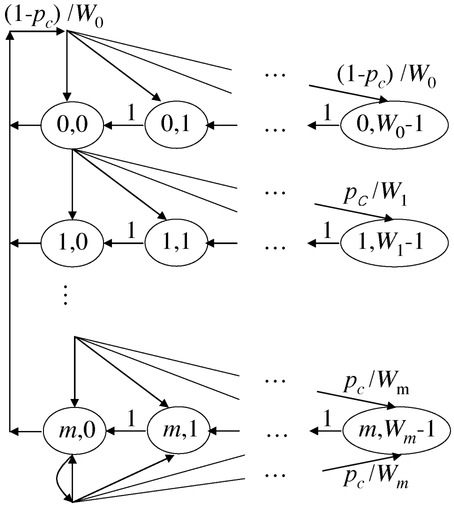


Fig. 2. State-transition diagram for the backoff window size.

effect. The transmission queues of each STA and AP are assumed to be always nonempty. The destinations of frames generated from n STAs are AP while the destinations of frames transmitted from AP are uniformly distributed among n STAs. Let us adopt the notation $W_i = 2\times W_{i-1}$ where $i \in \{0,\ldots,m\}$ is called the backoff stage and m is the maximum backoff stage such that $CW_{\max} = 2^m CW_{\min}$.

A discrete and integer time scale is adopted: t and $t+1$ correspond to the beginnings of two consecutive decrements of backoff time counter. We call the time interval between t and $t+1$ "counter time." Note that the decrement of the backoff time counter is stopped when the channel is sensed busy; thus, the time interval between the beginnings of two consecutive counter time instants may be much longer than the constant slot time size σ, as it may include a frame transmission. Let $a(t)$ be the stochastic process representing the backoff stage $\{0,\ldots,m\}$ of a STA at time t. Let $b(t)$ be the stochastic process representing the backoff time counter for a given STA. The $b(t)$ of each STA decrements at the beginning of each counter time.

If the probability of frame collision is approximated by a constant value p_c regardless of the backoff stage, it is possible to model the two-dimensional (2-D) process $\{a(t), b(t)\}$ with the discrete-time Markov chain depicted in Fig. 2. This assumption is reasonable in the steady state, since p_c mainly depends on the overall network traffic.

Let $P_{(i,j)(k,l)}$ denote the state-transition probability of $P\{(k,l) \rightarrow (i,j)\}$. Referring to Fig. 2, $P_{(i,j)(k,l)}$'s for all possible transitions $(k,l) \rightarrow (i,j)$ are expressed by (4), shown at the bottom of the page. The first case in (4) corresponds to the decrement of backoff time counter at the beginning of each counter time. The second case accounts for a successful frame transmission followed by a new initial backoff time counter randomly chosen in the range of $[0, W_0-1]$. The other cases model an unsuccessful frame transmission. In the third case, when an unsuccessful frame transmission occurs at a backoff stage of $i-1$, the backoff stage increases by one and a new initial backoff time counter is chosen in the range of $[1, W_i-1]$. Finally, the fourth case models the maximum backoff stage m, where it is not increased even after an unsuccessful frame transmission [see (4) at the bottom of the page].

Let $\pi_{i,j}$ denote the stationary probability of a state (i,j). In the steady state, we obtain the following relations:

$$p_c \times \pi_{i-1,0} = \begin{cases} \pi_{i,0} & 0 < i < m \\ (1-p_c)\times \pi_{m,0} & i = m. \end{cases} \tag{5}$$

Because of the chain regularities, for each $0 \le j \le W_i - 1$, $\pi_{i,j}$ can be expressed as

$$\pi_{i,j} = \frac{W_i - j}{W_i} \times \begin{cases} (1-p_c)\times \sum_{k=0}^{m} \pi_{k,0}, & i = 0 \\ p_c \times \pi_{i-1,0}, & 0 < i < m \\ p_c \times (\pi_{m-1,0} + \pi_{m,0}), & i = m. \end{cases} \tag{6}$$

Note that (5) can be obtained from (6) by setting $j = 0$. It is also noted that

$$\sum_{i=0}^{m} \sum_{j=0}^{W_i-1} \pi_{i,j} = 1. \tag{7}$$

From (4)–(7), we can get $\pi_{i,j}$ for any feasible (i,j) by using numerical methods. For details, the readers are referred to [18].

Let us define the probability τ that a STA transmits a data frame in a counter time. As any transmission occurs when the backoff time counter is equal to zero, we have

$$\tau = \sum_{i=0}^{m} \pi_{i,0} \tag{8}$$

and

$$p_c = 1 - (1-\tau)^n. \tag{9}$$

Let P_t be the probability that there is at least one transmission in a counter time and P_s be the probability of successful transmission. Since n STAs and AP contend on the channel with probability τ, we have

$$P_t = 1 - (1-\tau)^{n+1}, \tag{10}$$

$$P_s = \frac{(n+1)\tau(1-\tau)^n}{P_t} = \frac{(n+1)\tau(1-\tau)^n}{1-(1-\tau)^{n+1}}. \tag{11}$$

Let U be the normalized system utilization defined as the fraction of time that the channel is used to successfully transmit payload bits. We express U as the payload information transmitted

$$P_{(i,j)(k,l)} = \begin{cases} 1 & \text{if } 0 \le i \le m, & 0 \le j \le W_i - 2, & k = i, & \text{and } l = j+1 \\ \frac{(1-p_c)}{W_0} & \text{if } i = 0, & 0 \le j \le W_0 - 1, & 0 \le k \le m, & \text{and } l = 0 \\ \frac{p_c}{W_i} & \text{if } 1 \le i \le m, & 0 \le j \le W_i - 1, & k = i-1, & \text{and } l = 0 \\ \frac{p_c}{W_m} & \text{if } i = m, & 0 \le j \le W_m - 1, & k = m, & \text{and } l = 0. \end{cases} \tag{4}$$

in a counter time normalized to the length of a counter time, which is given by

$$U = \frac{P_s P_t T_p}{(1-P_t)\sigma + P_t P_s T_s + P_t(1-P_s)T_c}. \qquad (12)$$

Here, T_s is the average time that the channel is sensed busy because of a successful transmission, T_c is the average time that the channel is sensed busy because of a collision, σ is the duration of an empty slot time, and T_p is the average transmission time of the frame payload. The numerator in (12) represents the average time of payload information transmitted in a counter time and the denominator represents the average length of a counter time. In the RTS/CTS access mechanism, we have

$$\begin{aligned} T_s &= T_{\text{RTS}} + T_{\text{CTS}} + T_H + T_p + T_{\text{ACK}} \\ &\quad + 3T_{\text{SIFS}} + T_{\text{DIFS}} && (13) \\ T_c &= T_{\text{RTS}} + T_{\text{DIFS}} && (14) \end{aligned}$$

where T_{RTS}, T_{CTS}, T_H, T_{ACK}, T_{SIFS}, and T_{DIFS} are the time durations of RTS, CTS, frame header, ACK, SIFS, and DIFS, respectively.

As AP and STAs have the same τ, the downlink utilization U_d, the uplink utilization U_u, and the utilization ratio γ are

$$\begin{aligned} U_d &= U\frac{1}{n+1} && (15) \\ U_u &= U\frac{n}{n+1} && (16) \\ \gamma &= \frac{U_d}{U_u} = \frac{1}{n}. && (17) \end{aligned}$$

It is expected that the larger the number of STAs is, the higher the utilization is devoted to the uplink.

B. Utilization of the Proposed Method

At first, we analyze the performance of the proposed MAC in the case of $\psi = 1$ and then we use the approximate model for other cases.

If we assume that the lengths of data frames are set to the maximum value, $\omega(t_s)$ becomes equal to the number of data frames transmitted by downlink subtracted by the number of data frames transmitted by uplink. Thus, we can model the state transition diagram for $\omega(t_s)$, as shown in Fig. 3(a), where states are changed at every successful data-frame transmission. When a data frame is successfully transmitted from the AP, $\omega(t_s)$ increases by one, while $\omega(t_s)$ decreases by one when one of STAs transmits a data frame successfully. When $\omega(t_s)$ becomes -1, the DCA is activated after the previous ACK frame and $\omega(t_s)$ becomes zero with probability one.

The stationary probability of the state i, π_i, can be found by the following flow balance:

$$\begin{aligned} \pi_{-1} &= \frac{n}{n+1}\pi_0 && (18) \\ \pi_0 &= \pi_{-1} + \frac{n}{n+1}\pi_1 && (19) \\ \pi_i &= \frac{1}{n+1}\pi_{i-1} + \frac{n}{n+1}\pi_{i+1}, \quad \text{for } i \ge 1 && (20) \end{aligned}$$

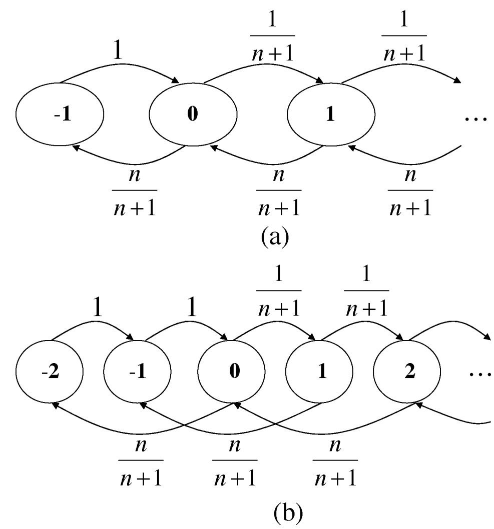


Fig. 3. State transition diagram for $\omega(t_s)$ at (a) $\psi = 1$ and (b) $\psi = 2$.

and

$$\sum_{i=-1}^{\infty} \pi_i = 1. \qquad (21)$$

It is easily shown that

$$\begin{aligned} \pi_{-1} &= \frac{n-1}{2n} && (22) \\ \pi_i &= \frac{n+1}{n^{i+1}}\pi_{-1} \\ &= \frac{n^2-1}{2n^{i+2}}, \quad \text{for } i \ge 0. && (23) \end{aligned}$$

Let P_d be the probability that a data frame is transmitted by the DCA under the condition that the previous transmission was successful. Because the DCA is activated at $\omega(t_s) < 0$, we obtain

$$P_d = \frac{\pi_{-1}}{P_s}. \qquad (24)$$

Then, the system utilization of the proposed MAC is given by

$$U = \frac{[P_s P_t T_p]}{[(1-P_t)\sigma + P_t P_s(1-P_d)T_s + P_t P_s P_d T_d + P_t(1-P_s)T_c]}. \qquad (25)$$

In (25), T_d is the time duration for the DCA and is given by

$$T_d = T_H + T_p + T_{\text{ACK}} + T_{\text{SIFS}} + T_{\text{PIFS}} \qquad (26)$$

where T_{PIFS} is the time durations of PIFS.

Compared with (12), the average counter time in (25) is reduced by $P_t P_s P_d(T_{\text{RTS}} + T_{\text{CTS}} + 2T_{\text{SIFS}} + T_{\text{DIFS}} - T_{\text{PIFS}})$, which comes from the overhead reduction of the DCA.

As shown in Fig. 3(b), it becomes more complicated when $\psi \neq 1$. Thus, we use the approximate model to accommodate the general ψ. When n is large enough, most of the DCF accesses are occupied by STAs and we can neglect the DCF access of AP. Then, the state transitions in Fig. 3 are simplified to those in Fig. 4. Note that there are $\psi+1$ states in Fig. 4. The state with zero corresponds to the DCF and those with negative values correspond to the DCA. We generalize these approximate models

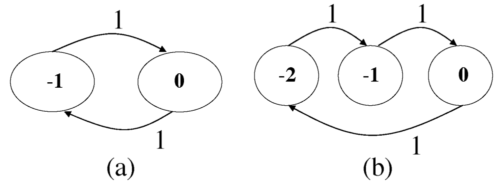


Fig. 4. State-transition diagram for $\omega(t_s)$ at (a) $\psi = 1$ and (b) $\psi = 2$ with an approximate model.

such that P_d and the steady-state probabilities of $\omega(t_s) < 0$ can be approximated by

$$\pi_{\omega(t_s)<0} = \sum_{i<0} \pi_i = \frac{\psi}{1+\psi} \tag{27}$$

$$P_d = \frac{\pi_{\omega(t_s)<0}}{P_s} = \frac{1}{P_s}\frac{\psi}{1+\psi}, \quad \text{for } \psi > 0. \tag{28}$$

The successful data-frame transmission of the downlink is generated at state $\omega(t_s) < 0$ and that of the uplink is generated at state $\omega(t_s) \geq 0$. Thus, the utilizations of uplink and downlink are given by

$$U_d = U\pi_{\omega(t_s)<0} = U\frac{\psi}{1+\psi} \tag{29}$$

$$U_u = U\pi_{\omega(t_s)\geq 0} = U\frac{1}{1+\psi} \tag{30}$$

$$\gamma = \frac{U_d}{U_u} = \psi. \tag{31}$$

The utilization between uplink and downlink are allocated according to the value of ψ.

V. Numerical Results

The values of parameters used to obtain numerical results for both the analytical model and the simulation runs are summarized in Table I. The values of these parameters are based on the IEEE 802.11b direct-sequence spread-spectrum (DSSS) standard [21]. The channel conditions and the traffic models of the simulation runs are the same as those of the analytical model described in Section IV.

The system utilizations of analytic and simulation result are plotted in Fig. 5 at $\psi = 1$. For simplicity, we denote the proposed method as FAIR and the 802.11 DCF as DCF in this section. The FAIR shows enhanced utilization compared with the DCF. This comes from the DCA, which reduces the overhead for data-frame transmission. The difference between the utilizations of DCF and FAIR becomes larger as the number of STAs increases. This is because the collision probability of DCF increases as the number of STAs increases. It is shown that the analytic result is accurate in predicting the system utilization.

Fig. 6 shows the utilizations of downlink and uplink at $\psi = 1$. As the number of STAs increases, most utilization for DCF is used for the uplink traffic, which results in severe unfairness between the uplink and the downlink. For a large number of STAs, AP cannot transmit enough traffic in DCF. However, the uplink and downlink share equal utilization in FAIR because ψ is set to one. This equal sharing is kept constant while the number of STAs increases.

Fig. 7 shows the utilization ratios γ for DCF and FAIR. γ in this figure is the time average value as define in (1). γ of FAIR is maintained at one because the DCA is activated whenever $\omega(t_s)$ in (3) is less than zero. Thus, the dynamic update method of $\omega(t_s)$ operates well to achieve the required utilization ratio ψ.

TABLE I
PARAMETER VALUES

Parameter	Value
$CWmin$	32
$CWmax$	1024
SIFS time	10 μs
PIFS time	30 μs
DIFS time	50 μs
slot time	20 μs
MAC header	272 bits
PHY header	48 bits
Preamble	144 μs
ACK time	304 μs
RTS time	352 μs
CTS time	304 μs
frame payload	8192 bits
channel bit rate	1 Mbps
T_w	30 sec

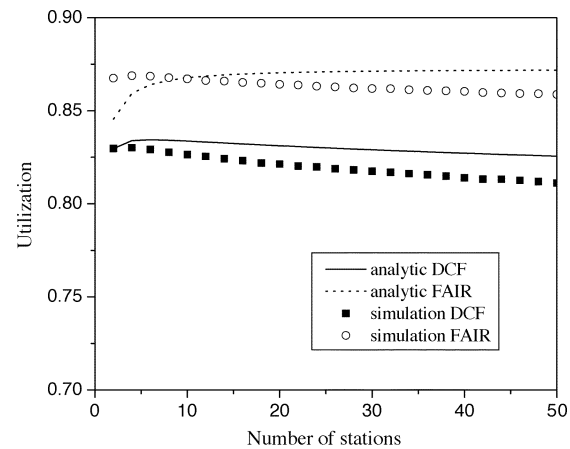


Fig. 5. System utilization U of analysis and simulation at $\psi = 1$.

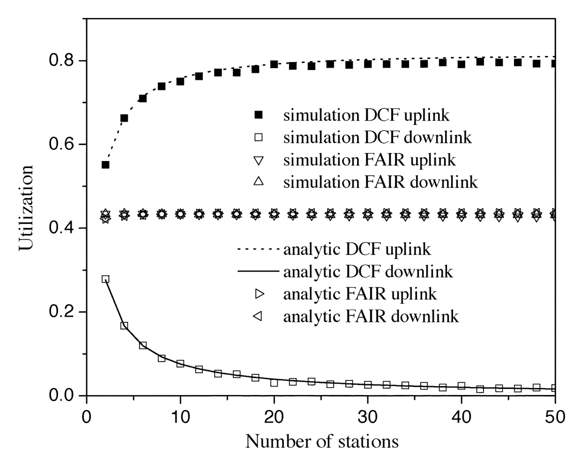


Fig. 6. Comparison between uplink and downlink utilization at $\psi = 1$.

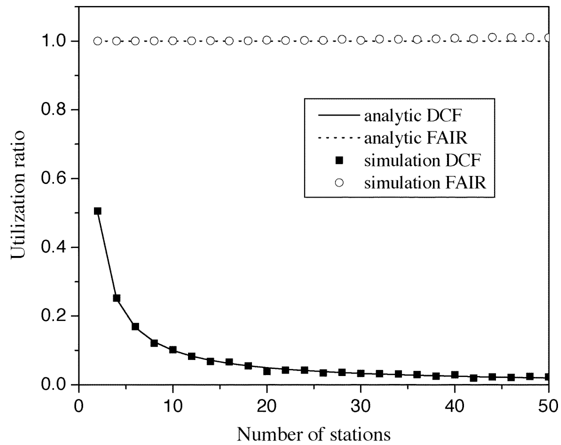


Fig. 7. Utilization ratio γ of analysis and simulation at $\psi = 1$.

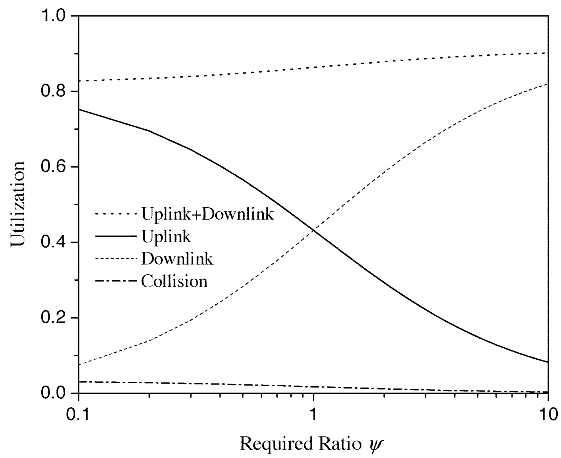


Fig. 8. Utilization sharing of FAIR between uplink and downlink at $n = 25$.

γ of DCF largely depends on the number of STAs as expected in (17).

Fig. 8 shows the simulation results of the FAIR according to the changes of ψ. The analytic results are not shown for the clarity, although it is well matched with the simulation results. As the value of ψ increases, the system utilization that is the sum of downlink and uplink utilization also increases. This is because when the value of ψ increases, the amount of decrement of $\omega(t_s)$ caused by a frame transmission of uplink gets larger, as explained in (3). Thus, as the value of ψ increases, more frames are transmitted during the DCA, which increases the utilization and decreases the collision. This means that the FAIR shows better utilization than the DCF as the traffic load of the downlink becomes heavier than that of the uplink, i.e., $\psi \gg 1$. The case of $\psi \gg 1$ can be widely observed in multimedia communications. The decrease of the utilization wasted by the collision also has the same reason. It is also shown that the FAIR distributes the utilization between the uplink and the downlink according to the value of ψ.

Fig. 9 shows the simulation results of dynamic parameter update for ψ shown in (2), where n_u and n_d represent the average number of active traffic flows that have different MAC addresses in uplink and downlink, respectively. Note that the ideal utilization ratio for $n_d = 40$, $n_d = 20$, and $n_d = 10$ should be 2, 1, and 0.5, respectively. The utilization ratios for the smaller values

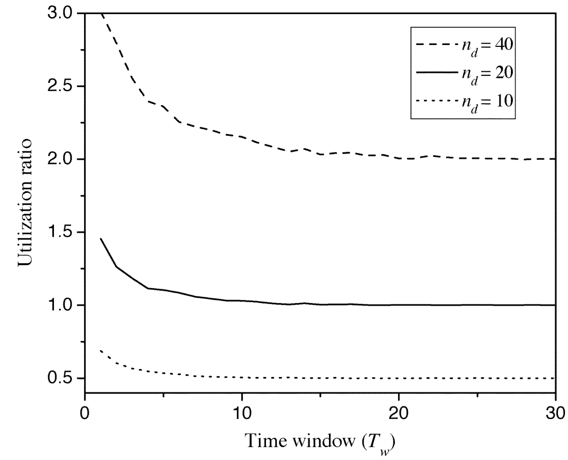


Fig. 9. Utilization ratio γ versus time window size T_w at $n_u = 20$.

of T_w are larger than the ideal values. This is because the lack of information about the number of active flows and the DCA mechanism give higher priority on the downlink utilization. It is also shown that T_w should increase as the downlink traffic load increases to achieve the ideal utilization ratio.

VI. Conclusion

We have proposed an easy implementation method to control the utilization ratio of uplink and downlink traffic and to enhance the system utilization of the IEEE 802.11 DCF. The proposed method can be implemented without the modification of the IEEE 802.11 standard for STAs that are widely deployed. The utilization sharing between the uplink and downlink can be controlled by the network operator or by the offered traffic load.

The efficiency of the proposed system under the saturation load condition has been demonstrated by Markov analysis and computer simulation. The results show that the proposed system enhances the utilization of successful data frame transmission even when the number of STAs increases. The proposed system distributes the utilization between the uplink and downlink according to the required value. This, in turn, drastically reduces the blocking probability of multimedia data frames in the proposed systems compared with that in the IEEE 802.11 DCF, where most of bandwidth is occupied by the uplink. Thus, the proposed system can be a good candidate for hotspots in next-generation cellular mobile communications systems, which aim for multimedia services.

References

[1] H. Honkasalo, K. Pehkonen, M. T. Niemi, and A. T. Leino, "WCDMA and WLAN for 3G and beyond," *IEEE Wireless Commun.*, no. 1, pp. 14–18, Apr. 2002.

[2] A. Doufexi, E. Tameh, A. Nix, and S. Armour, "Hotspot wireless LAN to enhance the performance of 3G and beyond cellular networks," *IEEE Commun. Mag.*, no. 7, pp. 58–65, Jul. 2003.

[3] S. Kishore, L. J. Greenstein, H. V. Poor, and S. C. Schwartz, "Uplink user capacity in a CDMA macrocell with a hotspot microcell: exact and approximate analyzes," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 364–374, Mar. 2003.

[4] Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, Aug. 1999.

[5] D. G. Jeong and W. S. Jeon, "CDMA/TDD system for wireless multimedia services with traffic unbalance between uplink and downlink," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 5, pp. 939–946, May 1999.

[6] Wireless LAN medium access control (MAC) and physical layer (PHY) specification: Medium access control (MAC) quality of service (QoS) enhancement, Feb. 2004.

[7] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Trans. Networking*, vol. 7, no. 4, pp. 473–489, Aug. 1999.

[8] A. Banchs and X. Perez, "Providing throughput guarantees in IEEE 802.11 wireless LAN," in *Proc. IEEE WCNC'02*, Orlando, FL, Mar. 2002, pp. 130–138.

[9] ——, "Distributed weighted fair queuing in 802.11 wireless LAN," in *Proc. IEEE ICC'02*, Apr. 2002, pp. 3121–3127.

[10] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed fair scheduling in a wireless LAN," in *Proc. IEEE/ACM MOBICOM'00*, Boston, MA, Aug. 2000, pp. 167–178.

[11] W. Pattara-Aticom, P. Krishnamurthy, and S. Banerjee, "Distributed mechanisms for quality of service in wireless LANs," *IEEE Trans. Wireless Commun.*, pp. 26–34, Jun. 2003.

[12] 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, 1999.

[13] M. Barry, A. T. Campbell, and A. Veres, "Distributed control algorithms for service differentiation in wireless packet networks," in *Proc. IEEE INFOCOM'01*, Anchorage, AK, Apr. 2001, pp. 582–590.

[14] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11," in *Proc. IEEE INFOCOM'01*, Anchorage, AK, Apr. 2001, pp. 209–218.

[15] A. Grilo and M. Nunes, "Performance evaluation of IEEE 802.11e," in *Proc. IEEE PIMRC'02*, Lisboa, Portugal, Sep. 2002, pp. 511–517.

[16] S. Pilosof, R. Ramjee, D. Raz, Y. Shavitt, and P. Sinha, "Understanding TCP fairness over wireless LAN," in *Proc. IEEE INFOCOM'03*, San Francisco, CA, Mar. 2003, pp. 863–872.

[17] D. G. Jeong and W. S. Jeon, "Time slot allocation in CDMA/TDD systems for mobile multimedia services," *IEEE Commun. Lett.*, vol. 4, no. 2, pp. 59–61, Feb. 2000.

[18] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

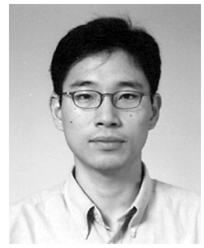
[19] V. Vitsas, "Throughput analysis of linear backoff scheme in wireless LANs," *Electron. Lett.*, vol. 39, no. 1, pp. 99–100, 2003.

[20] J. Hui and M. Devetsikiotis, "Designing improved MAC packet schedulers for 802.11e WLAN," in *Proc. IEEE GLOBECOM'03*, San Francisco, CA, Dec. 2003, pp. 184–189.

[21] Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Higher-speed physical layer extension in the 2.4 GHz Band, Sep. 1999.

Sung Won Kim received the B.S. and M.S. degrees from the Department of Control and Instrumentation Engineering, Seoul National University, Seoul, Korea, in 1990 and 1992, respectively, and the Ph.D. degree from the School of Electrical Engineering and Computer Sciences, Seoul National University, in 2002.

From January 1992 to August 2001, he was a Researcher with the Research and Development Center, LG Electronics, Seoul, Korea. From August 2001 to August 2003, he was a Researcher with the Research and Development Center, AL Tech, Seoul, Korea. He has been a Postdoctoral Researcher in the Department of Electrical and Computer Engineering, University of Florida, Gainesville, since August 2003. His research interests include resource management for wireless networks, performance evaluation for mobile networks, and medium-access control (MAC) protocol.



Byung-Seo Kim (S'04) received the B.S. degree in electrical engineering from In-Ha University, InChon, Korea, in 1998 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, Gainesville, in 2001 and 2004, respectively.

Between 1997 and 1999, he was with Motorola Korea Ltd., PaJu, Korea, as a Computer Integrated Manufacturing (CIM) Engineer. His research interests include the designing and developing of efficient link-adaptable medium-access control (MAC) protocol, cross-layer architecture, and a resource-allocation algorithm for wireless networks.

Yuguang Fang (S'96–M'97–SM'99) received the Ph.D. degree in systems and control engineering from Case Western Reserve University, Cleveland, OH, in 1994 and the Ph.D degree in electrical engineering from Boston University, Boston, MA, in 1997.

From September 1989 to December 1993, he was a Teaching/Research Assistant in the Department of Systems, Control and Industrial Engineering, Case Western Reserve University, where he was a Research Associate from January 1994 to May 1994. He was a Postdoctoral Researcher in the Department of Electrical and Computer Engineering, Boston University, from June 1994 to August 1995. From September 1995 to May 1997, he was a Research Assistant in the Department of Electrical and Computer Engineering, Boston University. From June 1997 to July 1998, he was a Visiting Assistant Professor in the Department of Electrical Engineering, University of Texas, Dallas. From July 1998 to May 2000, he was an Assistant Professor in the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark. In May 2000, he joined the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where he received an early promotion with tenure in August 2003 and has been an Associate Professor since. He has published over 100 papers in refereed professional journals and conferences. His research interests span many areas including wireless networks, mobile computing, mobile communications, automatic control, and neural networks.

Dr. Fang received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He is a Member of the Association for Computing Machinery (ACM) and is an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Editor for IEEE TRANSACTIONS ON MOBILE COMPUTING, a Technical Editor for IEEE WIRELESS COMMUNICATIONS MAGAZINE, an Editor for *ACM Wireless Networks*, and an Area Editor for *ACM Mobile Computing and Communications Review*. He was an Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS from May 1999 to December 2001, an Editor for the *International Journal on Wireless Communications and Mobile Computing* from April 2000 to January 2004 and the Feature Editor for the Scanning the Literature section in IEEE PERSONAL COMMUNICATIONS (now IEEE WIRELESS COMMUNICATIONS) from April 2000 to April 2003. He also has been actively involved with many professional conferences. He is the Program Co-Chair for the Global Internet and Next Generation Networks Symposium at IEEE GLOBECOM'04 and was the Program Vice Chair for the 2000 IEEE Wireless Communications and Networking Conference (WCNC'00), where he received the IEEE Appreciation Award for service to this conference. He has served on many Technical Program Committees, such as IEEE INFOCOM (1998, 2000, 2003, 2004, and 2005), IEEE ICC (2004), IEEE GLOBECOM (2002, 2003, and 2004), IEEE WCNC (1999, 2000, 2002, and 2004), and ACM MobiCom (2001). He served as the Committee Co-Chair for Student Travel Award for 2002 ACM MobiCom and was the Vice-Chair for the IEEE Gainesville Section in 2002 and 2003, and was the Chair in 2004.