Hindawi Publishing Corporation
Mobile Information Systems
Volume 2015, Article ID 383081, 16 pages
http://dx.doi.org/10.1155/2015/383081

Research Article

Hybrid Control of Contention Window and Frame Aggregation for Performance Enhancement in Multirate WLANs

Muhammad Adnan and Eun-Chan Park

Department of Information and Communication Engineering, Dongguk University, Seoul 100-715, Republic of Korea

Correspondence should be addressed to Eun-Chan Park; ecpark@dongguk.edu

Received 6 December 2013; Accepted 12 August 2014

Academic Editor: David Taniar

The IEEE 802.11 standard has been evolved to support multiple transmission rates in wireless local area networks (WLANs) to cope with diverse channel conditions and to increase throughput. However, when stations with different transmission rates coexist, the basic channel access mechanism of WLAN, distributed coordination function (DCF), not only fails to assure airtime fairness among competing stations but also decreases overall network throughput, because DCF was designed to provide fair opportunity of channel access, regardless of transmission rate. As an effective solution to this problem, we propose a hybrid control mechanism that integrates contention window control and frame aggregation. The former adjusts the size of contention window and differentiates the channel access opportunity depending on the transmission rates of stations. The latter controls the number of packets in the aggregated frame to tightly assure per-station airtime fairness with the reduced channel access overheads. Moreover, we derive an analytical model to evaluate the performance of the proposed mechanism in terms of throughput and fairness. Along with the analysis results, the extensive simulation results confirm that the proposed mechanism significantly increases the overall throughput by about three times compared to the conventional DCF, while assuring airtime fairness strictly.

1. Introduction

The development of technologies for mobile devices and wireless communications has resulted in the explosive increase in the demand for wireless Internet access. The wireless local area network (WLAN) based on IEEE 802.11 [1], also referred to as Wi-Fi, becomes the most prevailing technology to provide wireless connectivity for mobile devices. To cope with the increasing demand for higher throughput in WLANs, the IEEE 802.11n standard was introduced [2]. The main feature of IEEE 802.11n is the increase of transmission rate at the physical (PHY) layer up to 600 Mb/s. It is attained by employing several advanced PHY layer technologies such as multiple-input-multiple-output antennas, orthogonal frequency division multiplexing, adaptive channel coding, and channel bonding. Another key feature of IEEE 802.11n is the fast link adaptation mechanism with which each station (STA) adjusts its transmission rate depending on the channel condition. Therefore, STAs with different transmission rates coexist and they compete for the shared channel. However, the coexistence among heterogeneous STAs hinders the efficient channel sharing and prevents STAs in good channel conditions from achieving higher throughput than STAs in poor channel conditions. This problem results from the basic channel access mechanism of WLAN, distributed coordination function (DCF). In [3], an analytic model of DCF was derived using Markov chain and it was shown that DCF assures throughput fairness among STAs if all the STAs have the same packet size and transmission rate. Later, it was also shown that this property of DCF still holds even in the multirate environment [4–6]. As long as STAs have the same packet size, they attain the comparable throughput regardless of their transmission rates; that is, a STA with a higher transmission rate gets throughput similar to that with a lower transmission rate. Consequently, DCF remarkably decreases the overall achievable throughput in multirate WLANs. The main reason is that DCF was originally designed to provide all competing STAs with fair channel access opportunity; and thus a low-rate STA occupies the channel for a longer period of time to transmit a packet as compared to a high-rate STA.

In literature, this problem is termed as *performance anomaly* in multirate WLANs [4]. Many solutions have been proposed to resolve this problem. We can classify them into two categories: (i) control of channel access probability and (ii) control of channel occupation time. The first approach of controlling channel access probability was proposed in [7–12]. The key idea of this approach is that the contention window (CW) size is adjusted inversely proportional to the transmission rate. As a result, high-rate STAs have the smaller CW and they get more chances of channel access, whereas low-rate STAs have the larger CW and they get less chances of channel access. This approach assures that all the STAs occupy the channel for a comparable time in the long term of time, which is referred to as airtime fairness or temporal fairness. However, this approach essentially cannot be free from the problem originating from the trade-off between channel access delay and collision probability in setting the value of CW; (i) when network is sparse with low-rate STAs, the channel access delay becomes unnecessarily long; that is, the channel is underutilized; (ii) when network is dense with high-rate STAs, the collision probability becomes inevitably high. In order to provide temporal fairness in multirate WLANs, another approach proposed in [13–15] controls the channel occupation time once a STA gets the chance of channel access. It makes use of transmission opportunity (TXOP) or frame aggregation (FA). The TXOP and FA were originally proposed to improve quality of service (QoS) and efficiency of channel usage, respectively; they can also be used for enhancing fairness. By setting the same size of TXOP limit for all the STAs, the high-rate STA in the good channel condition is allowed to transmit a large number of frames while the low-rate STA in the bad channel condition can only transmit a small number of frames. In a similar way, by means of FA, the number of frames aggregated can be controlled in proportion to the transmission rate so that each STA occupies the airtime comparably, regardless of its transmission rate. Although these mechanisms are effective for airtime fairness, they have several drawbacks; (i) as the value of TXOP limit or the number of frames aggregated increases, the probability of transmission failure may increase due to variation of channel quality. Moreover, if collision occurs, the whole frames aggregated should be retransmitted, which is exacerbated when the number of STAs increases; (ii) conversely as the value of TXOP limit or the number of frames aggregated decreases, the channel access overheads (e.g., backoff time, PHY header, and acknowledgement (ACK) frame transmission time) increase, which reduce the efficiency of channel usage; (iii) both mechanisms of TXOP and FA cannot provide airtime fairness in a fine degree because they can only control the airtime in the time unit of frame transmission.

Besides these two approaches, another approach was recently proposed in [5, 16], where two mechanisms are combined to mitigate the performance anomaly problem. In [5, 16], the sophisticated model was derived to analyze the performance of multirate WLANs and several mechanisms were proposed; one of them is to differentiate both CW and frame size depending on the transmission rate of STA. It was asserted that, to relieve the problem, the size of CW should be set inversely proportional to the transmission rate while the frame size for the low-rate STA should be reduced compared to that for the high-rate STA. The performance was evaluated in diverse aspects by setting different values of CW and frame size for each STA. These studies proved the potentiality of the combined approach; however, the work in [5] did not present any specific mechanism to control the values of CW and frame size, and the work in [16] resorted to the relative differentiation of CW and frame size without proposing the proper reference values. Moreover, they did not consider the limit on the maximum size of IEEE 802.11 frame and the frame transmission failure due to collision. A similar mechanism proposed in [17] jointly controls CW and TXOP limit to improve the aggregate throughput and fairness of multicell WLANs. In this mechanism, the size of CW is controlled based on the packet error rate due to collision and the value of TXOP limit is adjusted such that the throughput of the worst link can be higher than a predefined threshold value. The objective of this work is to mitigate the effect of hidden terminal on the throughput and fairness in multicell WLANs, which is irrespective of performance anomaly in multirate WLANs.

In this paper, we propose a hybrid approach to deal with the problem of performance anomaly. The proposed mechanism combines both approaches of CW control and FA. It adjusts the size of CW depending on the number of STAs and transmission rate, in order to provide differentiated channel access probability without resulting in the unnecessary high collision probability or long channel access delay. At the same time, the number of packets in the aggregated frame is controlled to further assure that the channel occupation time of each STA becomes comparable, regardless of transmission rate and/or packet size. Compared to the existing mechanisms, the proposed mechanism holds the following key features, which are the contributions of this paper.

(i) The proposed hybrid mechanism incorporates the control of channel access probability with the control of channel occupation time, aiming to improve airtime fairness and channel efficiency in multirate WLANs. It supports *intergroup differentiation* and *intragroup differentiation* in terms of CW and FA.

(ii) The value of CW is not only controlled based on the transmission rate of STA but also scaled based on the number of STAs. In this way, the proposed mechanism can avoid underutilization of channel and high probability of collision, while mitigating the unfairness among STAs.

(iii) The number of packets aggregated is controlled to tightly assure equal amount of channel occupation time for each STA, regardless of transmission rate and packet size. The control mechanism of FA also reduces the channel access overheads, contributing to increase of throughput.

(iv) By deriving an accurate analytical model, it is proven that the proposed mechanism assures airtime fairness and that per-STA throughput becomes proportional to transmission rate.

The preliminary version of this paper was presented in [18]. In this paper, we significantly extend our previous work by elaborating the control mechanism, deriving the analysis model, and performing extensive simulations.

The rest of the paper is organized as follows. Section 2 presents the background and motivation of this study. The proposed mechanism is described in Section 3 and its analytical model is derived in Section 4. In Section 5, the performance of the proposed mechanism is evaluated via extensive simulations under various conditions. The conclusion follows in Section 6.

2. Background

2.1. Distributed Coordination Function (DCF). The DCF mechanism is based on principle of carrier sense multiple access (CSMA). Before transmitting a frame, the STA first checks whether the channel is busy. If the channel is idle during the interval of *distributed interframe space* (DIFS), the STA starts transmission. Otherwise it defers the transmission in order to avoid collision and backs off for a random amount of time before transmission attempt. During the backoff state, the STA listens to the channel and if it is sensed idle, the backoff counter is decremented. If the channel is sensed busy, the backoff counter is frozen until the channel becomes idle again. Once the backoff counter reaches zero, the STA starts transmitting. At the receiver side, if the frame is received successfully, which can be determined by checking the error detection code in the frame header, a positive acknowledgment (ACK) is sent back after a short interval called *short interframe space* (SIFS). Otherwise if the error is detected in the received frame, the receiver does not send any ACK. If the transmitter does not receive an ACK within the expected time, it assumes that collision has occurred and it will retransmit the frame with another new backoff time. The random backoff counter is chosen within the range $[0, W-1]$, where W is contention window. In the IEEE 802.11 standard, it is initially set to $\mathrm{CW}_{\min}$, and its value is doubled when detecting the transmission failure, with a maximum possible value of $\mathrm{CW}_{\max}$. If the transmission succeeds, the value of W is reset to $\mathrm{CW}_{\min}$. This procedure is referred to as *binary exponential backoff* and is helpful to mitigate collision. In this way, the random backoff mechanism of DCF can provide all competing STAs with fair channel access opportunity.

2.2. Contention Window Control. The contention window plays a key role in controlling the probability of channel access and collision; that is, as the value of CW increases, the collision probability decreases at the cost of increase in the delay of channel access. In the standard of IEEE 802.11e [19], the idea of differentiating CW has been introduced in order to improve QoS for different services. The standard defines four access categories (ACs) with different priorities, depending on QoS requirements. For a service classified as the high-priority AC (e.g., real-time service like voice over IP), the value of $\mathrm{CW}_{\min}$ is set to a small value, while it is set to a large value for a service classified as the low-priority AC (e.g., data service). Therefore, the channel can be accessed in a prioritized manner, and the frame with strict QoS requirements can be transmitted first compared to the frame with loose QoS requirements.

The differentiation of CW can also be used to deal with the problem of performance anomaly. Since the problem arises from the equal channel access opportunity among STAs regardless of transmission rate, it can be solved by giving different opportunity in proportion to the transmission rate of STA. For this purpose, the mechanism in [7, 8, 10] differentiates the value of $\mathrm{CW}_{\min}$ as

$$\mathrm{CW}_{\min}(R_d) = \mathrm{CW}_{\min}(R_{\max}) \cdot \left(\frac{R_{\max}}{R_d}\right), \quad (1)$$

where $\mathrm{CW}_{\min}(R_d)$ and $\mathrm{CW}_{\min}(R_{\max})$ are the values of $\mathrm{CW}_{\min}$ for the STA whose transmission rate is R_d and the maximum value of $R_{\max}$, respectively. It can be shown from (1) that, as R_d decreases from $R_{\max}$, $\mathrm{CW}_{\min}(R_d)$ increases and the probability of channel access decreases accordingly; that is, the low-rate STA gets less chance of channel access compared to the high-rate STA. This approach is simple and seems to be effective to mitigate the unfairness in terms of airtime; however, it has several disadvantages. First, the overall channel utilization may decrease. If there are few or no high-rate STAs but many low-rate STAs, the backoff time for low-rate STAs becomes unnecessarily long. Second, the collision may occur frequently and the airtime fairness cannot be attained due to the BEB mechanism. Consider that the channel is shared by few low-rate STAs and many high-rate STAs. In this case, since high-rate STAs have narrow range of backoff counter, frequent collisions occur among them, which cause retransmissions and decrease in the channel efficiency. Moreover, the collisions in high-rate STAs will increase the value of CW according to the BEB mechanism and deprive them of the preferential channel access, eventually resulting in the degradation of airtime fairness.

We can consider another type of CW differentiation as opposite to (1); that is,

$$\mathrm{CW}_{\min}(R_d) = \mathrm{CW}_{\min}(R_{\min}) \cdot \left(\frac{R_{\min}}{R_d}\right), \quad (2)$$

where $\mathrm{CW}_{\min}(R_{\min})$ is the value of $\mathrm{CW}_{\min}$ for the STA whose transmission rate is the minimum value of $R_{\min}$. According to (2), the value of CW decreases from the reference value of $\mathrm{CW}_{\min}(R_{\min})$ as R_d increases from $R_{\min}$. However, this approach also suffers from the drawbacks stated above.

Based on these rationales, we can make the following conclusions regarding the approach of CW differentiation.

(i) In differentiating the value of CW with respect to the transmission rate, its reference value (e.g., $\mathrm{CW}_{\min}(R_{\max})$ in (1) or $\mathrm{CW}_{\min}(R_{\min})$ in (2)) should be carefully determined to harmonize the trade-off between collision probability and channel utilization. Accordingly, the reference value of CW needs to be adjusted depending on the number of STAs sharing the channel.

(ii) The approach of CW differentiation combined with the BEB mechanism is effective only when the collision probability is negligible. Otherwise, it cannot assure airtime fairness strictly.

(iii) In assuring the comparable channel occupation time among STAs, several rate-independent overheads (e.g., PHY header and ACK transmission time) should be considered. In this sense, it is not always desirable for the high-rate STA to have smaller value of CW, because the effective channel occupation time by the high-rate STA is much less than that by the low-rate STA and the frequent channel access by the high-rate STA decreases the efficiency of channel usage.

2.3. TXOP and Frame Aggregation. The techniques of TXOP and frame aggregation can also be used to improve temporal fairness.

The mechanism of TXOP was introduced in IEEE 802.11e [19]. It defines a time, referred to as *TXOP limit*, within which the STA can transmit multiple frames once it gets the chance of channel access. After transmitting a frame followed by the corresponding ACK, the STA is allowed to transmit other backlogged frames after the time interval of SIFS. This mechanism reduces the effective backoff time, because multiple frames are transmitted with a single procedure of backoff. Furthermore, it can also be utilized for the temporal fairness among STAs. If the value of TXOP limit is set to be equal for all the STAs, we can intuitively expect that the channel occupation time per access can be comparable for all the STAs. However, this naive approach is not effective because of the following reasons. The back-to-back frame transmission within the TXOP limit is allowed only if the first frame is successfully transmitted without collision. Otherwise, the following frame transmissions also fail due to collision. The individual ACK for each frame is effective to avoid these consecutive collisions. However, this mechanism has two drawbacks when used to assure airtime fairness. First, it is effective only when the collision probability is properly controlled or minimized. Second, the control granularity is coarse; the value of TXOP limit cannot always be a multiple of frame transmission time for all the STAs that have different values of frame size and transmission rate. Thus, the fairness cannot be tightly attained and some portion of TXOP limit may be wasted.

The IEEE 802.11n standard [2] introduces two types of FA schemes, aggregation of MAC service data unit (A-MSDU) and aggregation of MAC protocol data unit (A-MPDU). The main objective of FA is to increase the channel efficiency by decreasing the MAC-layer overheads including backoff time, header overhead, and ACK transmission. The sender in the A-MSDU scheme constructs and transmits an MPDU aggregated with multiple MSDUs. Compared to TXOP, A-MSDU further removes the overheads because multiple packets (In this paper, the terms of "packet" and "frame" are interchangeably used with the terms of "MSDU" and "MPDU," respectively. Note that an MPDU consists of an MSDU and MAC header.) are transmitted with a single PHY/MAC header and a single ACK is used for the whole MSDUs. In the case of A-MPDU scheme, the sender transmits multiple MPDUs with a common PHY header and separate MAC header for each MPDU and the receiver informs the sender of the successful reception of each MPDU by sending a block ACK after receiving all the MPDUs. As well as improving channel efficiency, both mechanisms can be used to enforce temporal fairness by controlling the number of MSDUs or MPDUs aggregated. However, these mechanisms have several problems. First, A-MSDU has to unnecessarily retransmit the whole MSDUs even though only one MSDU is corrupted due to collision or channel error. A-MPDU is more effective than A-MSDU when the channel is error prone. If the transmission of some MPDUs fails, A-MPDU can selectively retransmit only the corrupted MPDU thanks to the block ACK mechanism. However, A-MPDU is still ineffective in the case of transmission failure due to collision. Second, the size of aggregated frame cannot be increased arbitrarily due to several reasons including limitations imposed by the IEEE 802.11 standard, variation in the channel quality, and short-term fairness. Third, both FA schemes have the problem of coarse control granularity, similar to the TXOP mechanism.

2.4. Motivation. The approach of either CW differentiation or FA is not effective to overcome the problem of performance anomaly due to several drawbacks discussed above. They can be summarized as follows.

(i) The approach of CW differentiation may cause channel underutilization or high collision probability, unless its reference value is properly determined by considering the number of STAs.

(ii) The approach of FA or TXOP can successfully provide airtime fairness only if the collision probability is minimized, which is not controllable without adjusting the value of CW. Moreover, this approach cannot control the airtime in a fine degree because of the coarse control granularity.

(iii) In order to assure airtime fairness and to improve channel efficiency, (i) for the high-rate STA, it is not desirable to aggressively increase the channel access probability by decreasing CW, because the frequent frame transmissions at the high rate decrease the channel efficiency due to the rate-independent overheads; (ii) for the low-rate STA, it is not possible to decrease the channel occupation time less than the transmission time of a single frame without frame fragmentation; that is, there is no choice but to decrease the channel access probability by increasing CW, in order to enforce airtime fairness.

These points motivate us to develop the hybrid approach combining the control of CW and frame aggregation, and they provide the desirable properties of the hybrid approach.

3. Proposed Mechanism

In this section, we propose a simple yet effective hybrid mechanism that controls both channel access probability and channel occupation time, by means of controlling the size

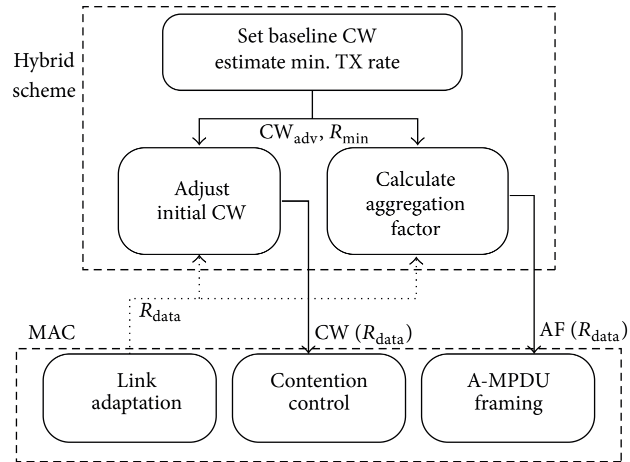


FIGURE 1: Block diagram of the proposed hybrid mechanism.

of CW and the number of packets aggregated, respectively. The objective of our mechanism is twofold, (i) assuring equal channel occupation time for each STA, regardless of transmission rate and/or packet size and (ii) improving channel efficiency by reducing collision probability and channel access overheads. We design the hybrid mechanism based on the motivation discussed in Section 2.4. We divide STAs into two groups, low-rate STAs and high-rate STAs, depending on the transmission rate of each STA, and apply different control strategies to these two groups. We first determine the reference value of CW, which is scaled with respect to the number of STAs. Then, we differentiate the value of CW between high-rate STAs and low-rate STAs, to give high-rate STAs more channel access chances (i.e., *intergroup differentiation*). We further control the value of *aggregation factor* (AF), defined as the number of MPDUs aggregated according to the A-MPDU scheme, to assure airtime fairness strictly and to reduce the channel access overheads. As well as intergroup differentiation, the *intragroup differentiation* is supported in controlling the value of AF; we set the value of AF in proportion to the transmission rate of each STA and differentiate it between low-rate STAs and high-rate STAs. The proposed mechanism works with three basic components as shown in Figure 1. In the subsequent subsections, we discuss the details of each component.

3.1. Estimation of Baseline Contention Window. We consider a single infrastructure WLAN where all the STAs are associated with an access point (AP) and they can sense the transmission of other STAs; that is, there is no hidden STA. The AP calculates and advertises the reference value of CW, denoted as CW_{adv}, which will be used in differentiating CW by each STA. The most important point in setting the value of CW_{adv} is that it should be set to minimize the probability of collision, not only to avoid retransmission but also to make the frame aggregation work successfully for assuring airtime fairness. Moreover, it is proven in [3, 20] that the optimal value of CW that maximizes the channel efficiency in a single-rate WLAN is proportional to the number of STAs, N_{sta}. Based on these rationales, the AP sets CW_{adv} as

$$CW_{adv} = CW_{min} \times N_{sta}. \quad (3)$$

Here, CW_{min} is the minimum value of CW proposed in the IEEE 802.11 standard (e.g., 16 in 802.11a/g). It seems that the value of CW_{adv} is set conservatively compared to the IEEE 802.11 standard, so it probably leads to the underutilization of channel. However, this is not the case. The value of CW_{adv} increases with respect to the value of N_{sta} as shown in (3), but the increase of CW_{adv} does not decrease the channel utilization; that is, the fraction of idle slots due to backoff does not proportionally increase as CW_{adv} increases. The backoff counters of all the STAs are decreased in a single idle slot; and thus the fraction of idle slots can be maintained constant on the whole if CW_{adv} is proportional to N_{sta}, which will be confirmed by simulations in Section 5. To calculate CW_{adv}, the AP keeps track of the value of N_{sta}. It is not difficult for the AP to estimate N_{sta} because all the STAs are associated with the AP and they transmit/receive frames to/from the AP. The calculation of CW_{adv} in (3) is quite simple, without requiring any information that is hard to measure or estimate.

There should be a signalling mechanism with which the AP can inform STAs of the value of CW_{adv} (or N_{sta}), because each STA calculates its own value of CW based on CW_{adv}. We consider that the AP periodically advertises the value of CW_{adv} in the beacon frame. The AP can also inform STAs of the value of N_{sta} by using the request/response mechanism for radio resource measurement defined in the IEEE 802.11k.

3.2. Control of Contention Window and Aggregation Factor. Each STA controls its own CW and AF based on CW_{adv} and the frame transmission rate. The STA i is classified as the low-rate STA if $R_i \leq \gamma \widetilde{R}_{min}$ and as the high-rate STA otherwise, where γ (>1) is a design parameter, R_i is the transmission rate of STA i, and $\widetilde{R}_{min}$ is the minimum transmission rate among active STAs belonging to the same basic service set (BSS). It is not difficult for each STA to learn $\widetilde{R}_{min}$ because the transmission rate is indicated in the PHY header, which is transmitted at the most robust rate (i.e., the lowest rate) irrespective of data frame transmission rate, and every STA within the same BSS can overhear the PHY header and successfully decode it. Note that $\widetilde{R}_{min}$ is not a constant value but it is updated whenever there is a change in the minimum value of frame transmission rate among active STAs.

We define $CW_{min,i}(R_i)$ as the initial value of CW for a STA that has the transmission rate of R_i and calculate it as

$$CW_{min,i}(R_i) = \begin{cases} [\alpha \cdot CW_{adv}], & \text{if } R_i \leq \gamma \widetilde{R}_{min}, \\ \left[\frac{\alpha}{2} \cdot CW_{adv}\right], & \text{otherwise.} \end{cases} \quad (4)$$

Here, α (>0) is a control parameter and $[x]$ represents the round-off value of x. Since the channel access probability is inversely proportional to the value of CW, the control rule of (4) makes high-rate STAs access the channel twice frequently than low-rate STAs and contributes to improving intergroup airtime fairness. As long as STAs belong to the same group, they have the same value of CW; that is, the control of CW supports only intergroup differentiation but no intragroup differentiation, which will be dealt with the control of AF. It is important to note that our mechanism decreases the value of

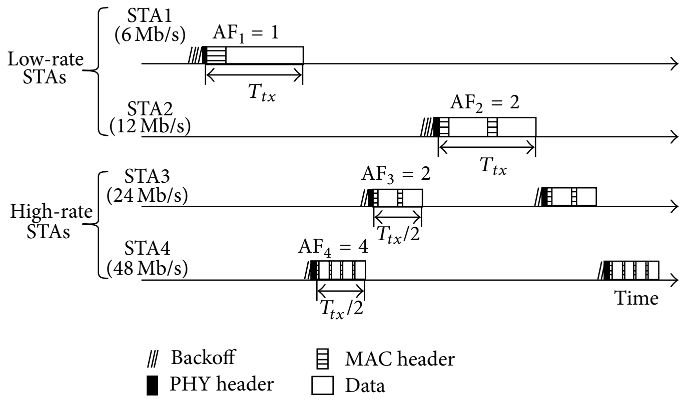


Figure 2: Operational example of channel access and transmission time with the hybrid mechanism.

CW for high-rate STAs less aggressively as compared to the existing approaches in [7, 8], in order to reduce the channel access overheads caused by high-rate STAs.

We employ A-MPDU as the frame aggregation scheme. We denote $\mathrm{AF}_i(R_i)$ as the value of AF for the STA that has the frame transmission rate of R_i. For the sake of simplicity, we make two assumptions, which will be removed in the next subsection. The first assumption is that the size of MPDU frame is the same for all the STAs, and the second one is that any transmission rate is even multiple of the minimum rate; that is, there exists a positive even integer $k(\geq 2)$ such that $R_i = k\widetilde{R}_{\min}$ for any $R_i(>\widetilde{R}_{\min})$. Under these assumptions, we design the control rule for AF as

$$\mathrm{AF}_i(R_i) = \begin{cases} \beta\left(\dfrac{R_i}{\widetilde{R}_{\min}}\right), & \text{if } R_i \leq \gamma\widetilde{R}_{\min}, \\ \dfrac{\beta}{2}\left(\dfrac{R_i}{\widetilde{R}_{\min}}\right), & \text{otherwise.} \end{cases} \quad (5)$$

Here, β is a control parameter and it is a positive integer. Note that, unlike the control of CW, the control of AF supports intragroup differentiation, as well as intergroup differentiation.

Proposition 1. *The combined control of CW and AF in (4) and (5) assures airtime fairness under the assumptions above and the ideal condition that there is neither collision nor channel error.*

Proof. Let us define $T_{tx,i}(R_i)$ as the transmission time of aggregated frame with the aggregation factor of $\mathrm{AF}_i(R_i)$. Here, $T_{tx,i}(R_i)$ includes the time to transmit MAC header and MPDU subheaders but excludes the transmission time of PHY header. Note that the transmission time of PHY header is fixed and it is independent of R_i while the MAC header (and MPDU subheaders) is transmitted at the rate of R_i. Defining L_{mpdu} as the size of an individual MPDU frame in bits, $T_{tx,i}(R_i)$ can be calculated from (5) as

$$T_{tx,i}(R_i) = \frac{L_{\mathrm{mpdu}}}{R_i}\mathrm{AF}_i(R_i) = \begin{cases} \beta\dfrac{L_{\mathrm{mpdu}}}{\widetilde{R}_{\min}}, & \text{for low-rate STAs,} \\ \dfrac{\beta}{2}\dfrac{L_{\mathrm{mpdu}}}{\widetilde{R}_{\min}}, & \text{for high-rate STAs.} \end{cases} \quad (6)$$

It is noteworthy in (6) that $T_{tx,i}(R_i)$ is independent of the transmission rate of each STA and it is the same for all the STAs as long as they belong to the same group. Therefore, the STAs within the same group have the same channel access probability and the same transmission time of aggregated frame from (4) and (6), respectively; that is, the hybrid mechanism assures airtime fairness among STAs belonging to the same group.

We define $N_{a,lo}$ and $N_{a,hi}$ as the amount of channel access during a given time interval of T by a low-rate STA and high-rate STA, respectively. Because of the intergroup differentiation of CW in (4) and the assumption of ideal channel condition, $N_{a,hi} = 2N_{a,lo}$. Then, during the time interval of T, the total time consumed for a high-rate STA to transmit the aggregated frames can be represented from (6) as

$$N_{a,hi}T_{tx,i} = (2N_{a,lo})\frac{\beta}{2}\frac{L_{\mathrm{mpdu}}}{\widetilde{R}_{\min}} = N_{a,lo}T_{tx,i}, \quad (7)$$

which is identical to the total transmission time consumed by a low-rate STA. The result in (7) means that the deficiency in the airtime for the high-rate STA (see (6)) is compensated by more channel access (see (4)) and that the combined control of CW and AF assures the equal airtime for all the STAs. This primitive proof shows the key idea of the proposed mechanism; the rigorous proof will be given with the complicated analysis model in Section 4. □

We give a simple illustrative example where there are four STAs with $R_i = 6, 12, 24, 48$ Mb/s, each of which is denoted as STA1, STA2, STA3, and STA4. If $\gamma = 3$, STA1 and STA2 are classified as low-rate STAs, and the other

INPUT: AF'_i (noninteger value in (8))
OUTPUT: $AF_{int,i}$ (integer value close to AF'_i)
$X \leftarrow$ rand(0, 1); //random value uniformly distributed between 0 and 1
if $X < \lceil AF'_i \rceil - AF'_i$ **then**
$AF_{int,i} \leftarrow \lfloor AF'_i \rfloor$
else
$AF_{int,i} \leftarrow \lceil AF'_i \rceil$
end if

Algorithm 1: Adjustment of aggregation factor with two integer values.

STAs are classified as high-rate STAs. We consider that β is one. Figure 2 shows how these STAs access the channel and transmit the aggregated frame under the proposed hybrid control mechanism. In Figure 2, the transmission of block ACK is intentionally omitted. Since $R_2 = 2R_1$ and $AF_2 = 2AF_1 = 2$ from (5), $T_{tx,2} = 2L_{mpdu}/R_2 = L_{mpdu}/R_1 = T_{tx,1}$. In the case of high-rate STAs, $AF_3 = 2$ and $AF_4 = 4$ from (5) and the corresponding transmission times become $T_{tx,3} = 2L_{mpdu}/R_3 = 2L_{mpdu}/(4R_1) = 0.5T_{tx,1}$ and $T_{tx,4} = 4L_{mpdu}/R_4 = 4L_{mpdu}/(8R_1) = 0.5T_{tx,1}$. The control of CW in (4) lets STA3 and STA4 get channel access twice more than STA1 and STA2. Eventually, the total channel occupation time by all the STAs becomes equal, regardless of transmission rate, and the total number of frames sent becomes proportional to the transmission rate of each STA.

3.3. Elaboration of AF Control Algorithm. We extend the control algorithm of AF by considering the cases where the MPDU size is different for each STA and the transmission rate has any arbitrary value. In this case, the value of AF_i in (5) cannot always become an integer and it causes a rounding error, which cannot be ignored and degrades the performance. To cope with this problem, we elaborate the control algorithm of AF.

By considering the difference in the MPDU size, we modify the value of AF, denoted as AF'_i, as

$$AF'_i(R_i) = AF_i(R_i)\frac{L_{ref}}{L_{mpdu,i}}, \quad (8)$$

where $AF_i(R_i)$ is the same as (5), $L_{mpdu,i}$ is the MPDU size of STA i, and L_{ref} is the reference value of MPDU size for all the STAs. We can see from (8) that the value of AF'_i increases as $L_{mpdu,i}$ decreases to assure fairness in terms of airtime. Next, we devise a simple algorithm to achieve the noninteger value of AF'_i with two integer values of $\lceil AF'_i \rceil$ ($\geq AF'_i$) and $\lfloor AF'_i \rfloor$ ($< AF'_i$), where $\lceil x \rceil$ and $\lfloor x \rfloor$ are the smallest integer that is equal to or greater than x (>0) and the largest integer that is smaller than x, respectively. The following algorithm shows how AF'_i is realized with the integer values of $AF_{int,i}$ ($= \lceil AF'_i \rceil$ or $\lfloor AF'_i \rfloor$). This algorithm is performed each time when the STA constructs the aggregated frame after getting the chance of channel access.

Proposition 2. *The frame aggregation scheme with two integer values in Algorithm 1 results in the effective value of AF identical to AF'_i.*

Proof. As indicated in Algorithm 1, $AF_{int,i}$ becomes either $\lfloor AF'_i \rfloor$ or $\lceil AF'_i \rceil$ with the probability of $w(= \lceil AF'_i \rceil - AF'_i)$ and $(1-w)$, respectively. Therefore, the average value of $AF_{int,i}$, denoted as $\overline{AF}_{int,i}$, can be represented as

$$\begin{aligned}\overline{AF}_{int,i} &= \lfloor AF'_i \rfloor w + \lceil AF'_i \rceil (1-w) \\ &= \left(\lceil AF'_i \rceil - 1\right) w + \lceil AF'_i \rceil (1-w) \\ &= -w + \lceil AF'_i \rceil = AF'_i,\end{aligned} \quad (9)$$

which proves that $\overline{AF}_{int,i} = AF'_i$. □

Proposition 3. *The frame aggregation scheme in Algorithm 1 assures airtime fairness for any frame size and transmission rate of STAs under the same condition as Proposition 1.*

Proof. Similar to (6), we calculate $T_{tx,i}$ as

$$\begin{aligned}T_{tx,i}(R_i) &= \frac{L_{mpdu,i}}{R_i}\left(\lfloor AF'_i \rfloor w + \lceil AF'_i \rceil (1-w)\right) \\ &= \frac{L_{mpdu,i}}{R_i} AF'_i \\ &= \begin{cases} \dfrac{L_{mpdu,i}}{R_i}\left(\beta \dfrac{R_i}{\widetilde{R}_{min}} \dfrac{L_{ref}}{L_{mpdu,i}}\right) = \beta \dfrac{L_{ref}}{\widetilde{R}_{min}}, & \text{for low-rate STAs,} \\ \dfrac{L_{mpdu,i}}{R_i}\left(\dfrac{\beta}{2} \dfrac{R_i}{\widetilde{R}_{min}} \dfrac{L_{ref}}{L_{mpdu,i}}\right) = \dfrac{\beta}{2} \dfrac{L_{ref}}{\widetilde{R}_{min}}, & \text{for high-rate STAs,} \end{cases}\end{aligned} \quad (10)$$

from (5), (8), and (9). The result in (10) along with (4) proves airtime fairness of the frame aggregation scheme in Algorithm 1. □

3.4. Discussion on the Hybrid Mechanism. The proposed hybrid control mechanism has three design parameters, α, β, and γ; α and β are the scaling factors that affect the size of CW and aggregated frame, respectively, and γ is used to distinguish STAs as the low-rate STA or high-rate STA. Here, we discuss how these parameters affect the performance, which can be used as the guideline for setting their values.

We first focus on the effect of α (see (4)). As the value of α increases, the collision probability decreases at the cost of increase in the channel access delay. Thus, α can be used to tune the trade-off between collision probability and channel access delay. We can expect that there exists an optimal value of α that maximizes the overall throughput and that the value of α has little effect on the fairness since the fairness is mainly controlled by the AF control. Using the analysis model derived in the next section, we verified that these intuitions are valid and found that the optimal value of α is about one for several typical conditions. Note that it is hard to obtain the optimal value of α analytically because of nonlinearity of the system model.

Next, we consider the effect of β. The value of β should be set properly so that AF' in (8) should be larger than one. Otherwise if AF' is less than one, $\mathrm{AF}_{\mathrm{int},i}$ becomes zero with a nonzero probability; that is, the STA cannot transmit any frame by acquiring the chance of channel access, and thus the airtime can be wasted. On the other hand, as the value of β increases, the gain of frame aggregation also increases; that is, the header and timing overheads are reduced and the throughput increases accordingly. However, the increase of β causes several issues. First, the large value of β increases the transmission failure due to channel error. The link adaptation mechanism determines a proper transmission rate to cope with the time-varying channel quality. As the transmission time of aggregated frame increases, the transmission becomes more liable to be failed because of fast change in the channel quality like multipath fading. Moreover, if collision occurs, the increase of β accordingly increases the number of frames that should be retransmitted. Second, the increase of β degrades the short-term fairness among STAs and increases the variation in the packet transmission delay since the STA occupies the shared channel for a longer time once it gets the chance of channel access. Third, there exists the limit on the maximum size of aggregated frame in the IEEE 802.11 standard. The maximum size of A-MPDU is specified as 64 KB in IEEE 802.11n and the length of block ACK information field is limited so that at most 64 MPDUs can be aggregated. Taking all of these points into consideration, we consider that the appropriate value of β ranges between one and three.

Lastly, we discuss the effect of γ. The feasible range of γ is between one and $R_{\max}/R_{\min}$ (e.g., 9 in the case of IEEE 802.11a/g). If it is set to these two extreme values, the inter-group differentiation of CW or AF becomes nearly disabled. The value of γ can be used to balance two independent gains that can be achieved by CW control and AF control. In the proposed hybrid mechanism, a high-rate STA can get higher throughput than a low-rate STA due to two reasons; (i) it gets more chances of channel access by the CW control; (ii) it transmits more frames per channel access by the AF control. In order to assure airtime fairness and to improve the total throughput, it is desirable to allow the high-rate STA to transmit more frames per channel access, instead of getting more channel access, because the overhead time per channel access (e.g., backoff time, transmission time of PHY header, and ACK) is independent of R_i and it becomes relatively higher for the high-rate STA compared to the low-rate STA. Because of these reasons, it is better to set γ to a large value so that the hybrid approach sets CW in a conservative manner (i.e., most STAs have the CW scaling factor of α and few high-rate STAs have that of $\alpha/2$) and sets AF in an aggressive manner (i.e., most STAs have the AF scaling factor of β while few STAs have that of $\beta/2$). For the typical conditions of WLANs, we consider that the appropriate value of γ ranges between three and five.

4. Analysis of the Hybrid Mechanism

In this section, we derive the analysis model of the proposed hybrid control mechanism. Using the analysis model, we prove that the proposed mechanism assures airtime fairness and attains the per-STA throughput proportional to the transmission rate, in the presence of collisions and changes of CW due to the BEB mechanism. For the sake of tractability, we make the following reasonable assumptions; (i) there are no hidden STAs and the transmission failure results from either collision or channel error; (ii) every STA always has large number of packets to aggregate and transmit; (iii) the transmission rate of STA is fixed and determined by the link adaptation mechanism. We classify the STAs into M classes depending on its transmission rate, that is, $R_i \in \mathbb{R}$ and $|\mathbb{R}| = M$, and consider that there exist n_i STAs in the ith class $(1 \le i \le M)$.

4.1. Markov Chain Model. To derive the analysis model, we use a two-dimensional discrete-time Markov chain, as similar to the well-known Bianch model [3] and its variations [5, 6, 16]. In the Markov chain, the states are set as $(B_i(t), C_i(t))$, where $B_i(t)$ and $C_i(t)$ are the stochastic processes representing the backoff stage and backoff counter of the class-i STA, respectively. We define $W_{k,i}$ as the size of contention window when $B_i(t) = k$, which can be represented as

$$W_{k,i} = \begin{cases} 2^k W_{0,i}, & 0 \le k \le m_i, \\ 2^{m_i} W_{0,i}, & m_i < k \le L_{r,i}, \end{cases} \tag{11}$$

where $W_{0,i}$ and m_i are the minimum contention window and the maximum backoff stage of the class-i STA, respectively; that is, $W_{0,i} = \mathrm{CW}_{\min,i}$ as given in (4) and $m_i = \log_2(\mathrm{CW}_{\max,i}/\mathrm{CW}_{\min,i})$, and $L_{r,i}$ denotes the maximum number of retransmissions, which is assumed to be larger than m_i. For the backoff stage of $B_i(t) = k$, $C_i(t)$ at slot time t is a uniform random variable in the range of $[0, W_{k,i} - 1]$. We denote (k, l) as the states of Markov chain, where $0 \le k \le L_{r,i}$ and $0 \le l \le W_{k,i} - 1$, and define the state transition probability as

$$\begin{aligned} &P\{k, l \mid k_0, l_0\} \\ &\quad = P\{B_i(t+1) = k,\ C_i(t+1) = l \mid B_i(t) = k_0,\ C_i(t) = l_0\}. \end{aligned} \tag{12}$$

We can consider the following five cases of events and the corresponding transitions of the state variables in the Markov chain.

(C1) The backoff counter is decremented by one when the channel is sensed idle.

(C2) The backoff counter is frozen when the channel is sensed busy.

(C3) If the transmission fails at the backoff stage k ($<L_{r,i}$), the backoff stage increases by one and a new random backoff counter is selected with the increased CW of $W_{k+1,i}$ for retransmission.

(C4) If the transmission succeeds at the backoff stage $k(<L_{r,i})$, the backoff stage is reset and a new backoff counter is selected with the initialized CW of $W_{0,i}$ for the next transmission.

(C5) After the transmission at the backoff stage $k = L_{r,i}$, the backoff stage is reset and a new backoff counter is selected, regardless of transmission failure or success.

The state transition probabilities of these five cases can be calculated as

$$\begin{aligned}
&\text{(C1)}: P\{k, l-1 \mid k, l\} = 1 - P_{b,i}, \\
&\qquad 0 \le k \le L_{r,i}, \quad 0 < l < W_{k,i}, \\
&\text{(C2)}: P\{k, l \mid k, l\} = P_{b,i}, \\
&\qquad 0 \le k \le L_{r,i}, \quad 1 \le l < W_{k,i}, \\
&\text{(C3)}: P\{k+1, l \mid k, 0\} = \frac{P_{c,i}}{W_{k+1,i}}, \\
&\qquad 0 \le k < L_{r,i}, \quad 0 \le l < W_{k,i}, \\
&\text{(C4)}: P\{0, l \mid k, 0\} = \frac{(1 - P_{c,i})}{W_{0,i}}, \\
&\qquad 0 \le k < L_{r,i}, \quad 0 \le l < W_{0,i}, \\
&\text{(C5)}: P\{0, l \mid L_{r,i}, 0\} = \frac{1}{W_{0,i}}, \\
&\qquad 0 \le l < W_{0,i},
\end{aligned} \tag{13}$$

where $P_{b,i}$ and $P_{c,i}$ are the probabilities that the class-i STA senses the channel busy and that the transmission by the class-i STA collides, respectively.

Now, we derive the steady state distribution of this Markov chain, defined as $p_{ss,i}(k,l) = \lim_{t\to\infty} P\{B_i(t) = k,\ C_i(t) = l\}$. From the state transition probabilities in (13) and the balance equations of Markov chain in the steady state, we can get the following relations:

$$\begin{aligned}
&p_{ss,i}(k,0) = (P_{c,i})^k\, p_{ss,i}(0,0), \quad 0 \le k \le L_{r,i}, \\
&p_{ss,i}(k,l) = \frac{W_{k,i} - l}{W_{k,i}} \frac{1}{1 - P_{b,i}} p_{ss,i}(k,0), \\
&\qquad 0 \le k \le L_{r,i}, \quad 1 \le l < W_{k,i}.
\end{aligned} \tag{14}$$

From (14) and the fact that $\sum_{k=0}^{L_{r,i}} \sum_{l=0}^{W_{k,i}-1} p_{ss,i}(k,l) = 1$, $p_{ss,i}(0,0)$ can be found as

$$\begin{aligned}
p_{ss,i}(0,0) &= \left[\sum_{k=0}^{L_{r,i}} (P_{c,i})^k \left(1 + \frac{1}{1 - P_{b,i}} \sum_{l=1}^{W_{k,i}-1} \frac{W_{k,i} - l}{W_{k,i}}\right)\right]^{-1} \\
&= \left[\sum_{k=0}^{L_{r,i}} (P_{c,i})^k \left(1 + \frac{W_{k,i} - 1}{2(1 - P_{b,i})}\right)\right]^{-1}.
\end{aligned} \tag{15}$$

Next, we calculate several probabilities of $P_{a,i}$, $P_{b,i}$, $P_{c,i}$, and $P_{s,i}$. Here, $P_{a,i}$ and $P_{s,i}$ are the probabilities that the class-i STA attempts to access the channel and it succeeds in the transmission, respectively. From (14) and (15), we can calculate the probability of transmission attempt by the class-i STA as

$$P_{a,i} = \sum_{k=0}^{L_{r,i}} p_{ss,i}(k,0) = \frac{1 - (P_{c,i})^{(L_{r,i}+1)}}{1 - P_{c,i}} p_{ss,i}(0,0), \tag{16}$$

because the STA is allowed to transmit when its backoff counter reaches zero. Note that $P_{b,i}$ is the probability that the channel is determined to be busy under the condition that the class-i STA is sensing the channel while $P_{c,i}$ is the collision probability under the condition that it is transmitting. The channel is sensed busy if at least one STA is transmitting except for the sensed STA in the class i; therefore, $P_{b,i}$ becomes

$$P_{b,i} = 1 - (1 - P_{a,i})^{n_i - 1} \prod_{k=1, k\neq i}^{M} (1 - P_{a,k})^{n_k}. \tag{17}$$

The transmission by the class-i STA collides if more than one STA in the class i or class $k(\neq i)$ is transmitting at the same time. Thus, the collision probability $P_{c,i}$ is represented as

$$P_{c,i} = 1 - (1 - P_{a,i})^{n_i - 1} \prod_{k=1, k\neq i}^{M} (1 - P_{a,k})^{n_k}. \tag{18}$$

It is important to note that both $P_{b,i}$ and $P_{c,i}$ have the same formula as shown in (17) and (18); however, they have different conditions; that is, $P_{b,i}$ is the probability under the condition of channel sensing, while $P_{c,i}$ is the probability under the condition of transmission. Since the transmission becomes successful when only one STA is transmitting, the probability of successful transmission $P_{s,i}$ becomes

$$P_{s,i} = n_i P_{a,i} (1 - P_{a,i})^{n_i - 1} \prod_{k=1, k\neq i}^{M} (1 - P_{a,k})^{n_k}. \tag{19}$$

Finally, we can find $P_{a,i}$, $P_{b,i}$, $P_{c,i}$, and $P_{s,i}$ numerically from (14) to (19).

4.2. Collision Modeling under the Hybrid Mechanism. In order to derive the throughput that can be achieved by the class-i STA, we need to further analyze the system by considering *intergroup* and *intragroup* collisions which occur in the proposed hybrid control mechanism.

First, we introduce two groups, G_LO and G_HI, and we consider that the STA belongs to either G_LO or G_HI if its transmission rate is smaller than or equal to $\gamma \tilde{R}_{\min}$ or not, respectively. We define $\mathbb{S}_{\mathbb{LO}}$ and $\mathbb{S}_{\mathbb{HI}}$ as the sets of STAs belonging to G_LO and G_HI and define M_{LO} and M_{HI} as their cardinalities; that is, $M_{\mathrm{LO}} = |\mathbb{S}_{\mathbb{LO}}|$ and $M_{\mathrm{HI}} = |\mathbb{S}_{\mathbb{LO}}|$, respectively. From (4), (11), and (15)–(18), we can see that the transmission attempt probability is equal for all the classes within the same group since they have the same value of $W_{0,i} = \mathrm{CW}_{\min,i}$. We define $P_{a,\mathrm{LO}}$ and $P_{a,\mathrm{HI}}$ as the transmission attempt probabilities of STA belonging to G_LO and G_HI, respectively; that is,

$$P_{a,\mathrm{LO}} \triangleq P_{a,i\in\mathbb{S}_{\mathbb{LO}}} = P_{a,j(\neq i)\in\mathbb{S}_{\mathbb{LO}}},$$
$$P_{a,\mathrm{HI}} \triangleq P_{a,i\in\mathbb{S}_{\mathbb{HI}}} = P_{a,j(\neq i)\in\mathbb{S}_{\mathbb{HI}}}. \tag{20}$$

Next, we introduce two types of collision probabilities, intragroup and intergroup collision probabilities. The former refers to the probability of collision among STAs belonging to the same group, while the latter refers to the probability that more than two STAs in different groups collide. Defining $P_{c,\mathrm{LO}}$ and $P_{c,\mathrm{HI}}$ as the intragroup collision probabilities for G_LO and G_HI, respectively, they are represented as

$$\begin{aligned} P_{c,\mathrm{LO}} &= \Big(1 - \big((1-P_{a,\mathrm{LO}})^{M_{\mathrm{LO}}} \\ &\qquad + M_{\mathrm{LO}} P_{a,\mathrm{LO}} (1-P_{a,\mathrm{LO}})^{(M_{\mathrm{LO}}-1)}\big)\Big) \\ &\quad \times (1-P_{a,\mathrm{HI}})^{M_{\mathrm{HI}}}, \\ P_{c,\mathrm{HI}} &= \Big(1 - \big((1-P_{a,\mathrm{HI}})^{M_{\mathrm{HI}}} \\ &\qquad + M_{\mathrm{HI}} P_{a,\mathrm{HI}} (1-P_{a,\mathrm{HI}})^{(M_{\mathrm{HI}}-1)}\big)\Big) \\ &\quad \times (1-P_{a,\mathrm{LO}})^{M_{\mathrm{LO}}}. \end{aligned} \tag{21}$$

In a similar way, we also obtain the intergroup collision probability, denoted as $P_{c,L-H}$, as

$$P_{c,L-H} = \left(1-(1-P_{a,\mathrm{LO}})^{M_{\mathrm{LO}}}\right) \times \left(1-(1-P_{a,\mathrm{HI}})^{M_{\mathrm{HI}}}\right). \tag{22}$$

4.3. Analysis Results. We finally calculate per-station throughput and aggregate throughput from the models derived in Sections 4.1 and 4.2. We consider three states involved in the frame transmission: backoff (or idle), collision, and success. We define three timing variables, $\overline{T_{\mathrm{bo}}}$, $\overline{T_{\mathrm{col}}}$, and $\overline{T_{\mathrm{suc}}}$, each of which denotes the average time consumed in backoff, collision, and successful transmission, respectively. From (16) to (22), they can be represented as

$$\overline{T_{\mathrm{bo}}} = \left(\prod_{k=1}^{M} (1-P_{a,k})^{n_k}\right) \cdot T_{\mathrm{slot}}, \tag{23}$$

$$\begin{aligned} \overline{T_{\mathrm{col}}} &= P_{c,\mathrm{LO}} \cdot T_{f,i\in\mathbb{S}_{\mathbb{LO}}} + P_{c,\mathrm{HI}} \cdot T_{f,i\in\mathbb{S}_{\mathbb{HI}}} \\ &\quad + P_{c,L-H} \cdot \max\left(T_{f,i\in\mathbb{S}_{\mathbb{LO}}}, T_{f,i\in\mathbb{S}_{\mathbb{HI}}}\right), \end{aligned} \tag{24}$$

$$\overline{T_{\mathrm{suc}}} = \sum_{k=1}^{M} \left(P_{s,k} \cdot T_{f,k}\right). \tag{25}$$

Here, T_{slot} is the duration of a slot and $T_{f,i}$ is the transmission time of aggregated frame by the class-i STA; that is,

$$T_{f,i} = T_{\mathrm{oh}} + \frac{\mathrm{AF}_i' \cdot L_{\mathrm{mpdu},i}}{R_i}, \tag{26}$$

where T_{oh} is the overhead time including DIFS, SIFS, and transmission time of PHY header and block ACK, which is independent of R_i. From (5) and (8), (26) can be rewritten as

$$T_{f,i\in\mathbb{S}_{\mathbb{LO}}} = T_{\mathrm{oh}} + \beta \frac{L_{\mathrm{ref}}}{\tilde{R}_{\min}},$$
$$T_{f,i\in\mathbb{S}_{\mathbb{HI}}} = T_{\mathrm{oh}} + \frac{\beta}{2} \frac{L_{\mathrm{ref}}}{\tilde{R}_{\min}}. \tag{27}$$

Note that the first and second terms in the right side of (24) represent the average time due to intragroup collision, while the third term represents the time due to intergroup collision, which lasts until finishing to transmit the longer aggregated frame; that is, $\max(T_{f,i\in\mathbb{S}_{\mathbb{LO}}}, T_{f,i\in\mathbb{S}_{\mathbb{HI}}}) = T_{f,i\in\mathbb{S}_{\mathbb{LO}}}$ from (27).

4.3.1. Throughput. The per-STA throughput of the class-i STA, th_i, can be represented as

$$\mathrm{th}_i = \frac{p_{s,i} \cdot L_{p,i}}{\overline{T_{\mathrm{bo}}} + \overline{T_{\mathrm{col}}} + \overline{T_{\mathrm{suc}}}}, \tag{28}$$

where $L_{p,i}$ is the size of aggregated frame except for the PHY and MAC headers. We define $L_{\mathrm{pkt},i}$ and L_{mh} as the packet size (in bits) of class-i STA and the size (in bits) of MAC header including MPDU-related subheader; that is, $L_{\mathrm{mpdu},i} = L_{\mathrm{pkt},i} + L_{\mathrm{mh}}$. Assuming $L_{\mathrm{mh}} \ll L_{\mathrm{pkt},i}$, that is, $L_{\mathrm{mpdu},i} \approx L_{\mathrm{pkt},i}$, $L_{p,i}$ can be approximated as

$$L_{p,i} = \mathrm{AF}_i' \cdot L_{\mathrm{pkt},i} \approx \begin{cases} \beta L_{\mathrm{ref}} \dfrac{R_i}{\tilde{R}_{\min}}, & \text{for } i \in \mathbb{S}_{\mathbb{LO}}, \\ \dfrac{\beta}{2} L_{\mathrm{ref}} \dfrac{R_i}{\tilde{R}_{\min}}, & \text{for } i \in \mathbb{S}_{\mathbb{HI}}. \end{cases} \tag{29}$$

In (28), note that $p_{s,i} = P_{s,i}/n_i$ is the per-STA probability of successful transmission, while $P_{s,i}$ is the per-class probability (see (19)). The throughput model in (28) is still effective even if the transmission fails due to channel error. As long as the A-MPDU scheme selectively retransmits the corrupted frame by virtue of block ACK and a proper link adaptation mechanism determines the modulation and coding scheme such that the MPDU frame error rate (FER) is maintained around the target value, the throughput will decrease by a factor of $1 - p_{\mathrm{err}}$, where p_{err} is the target value of FER. On the other hand, the aggregate throughput, TH, can be obtained as

$$\mathrm{TH} = \sum_{i=1}^{M} (n_i \cdot \mathrm{th}_i). \tag{30}$$

4.3.2. Fairness. Now, we focus on the performance of the proposed mechanism in terms of fairness. It is shown from [6, 16] that if $\mathrm{CW}_{\min,i} \gg 1$ and $m_i \approx m_j \gg 1$, then, for $i \neq j$, $P_{c,i} \approx P_{c,j}$ and

$$p_{s,i} W_{0,i} \approx p_{s,j} W_{0,j}. \quad (31)$$

We verified that the result in (31) agrees well with our analysis model given in (15)–(19) (We obtained the value of $p_{s,i}$ numerically and found that the value of $|p_{s,i}W_{0,i} - p_{s,j}W_{0,j}|$ normalized to its average value is insignificant for most cases.).

Proposition 4. *The proposed hybrid mechanism assures airtime fairness among STAs, even in the presence of collisions. Moreover, the fairness is attained irrespective of the control parameters α and β.*

Proof. Consider a unit time interval, $T_u = \overline{T_{\mathrm{bo}}} + \overline{T_{\mathrm{col}}} + \overline{T_{\mathrm{suc}}}$. During T_u, the effective time consumed to transmit the aggregated frame successfully becomes $p_{s,i}T_{f,i}$. First, consider the case where $i, j \in \mathbb{S}_{\mathbb{LO}}$ or $i, j \in \mathbb{S}_{\mathbb{HI}}$. In this case, $W_{0,i}$ is identical to $W_{0,j}$, and it is obvious from (26) and (31) that $p_{s,i}T_{f,i} = p_{s,j}T_{f,j}$. Next, we consider the case where $i \in \mathbb{S}_{\mathbb{LO}}$ and $j \in \mathbb{S}_{\mathbb{HI}}$; then

$$\begin{aligned} \frac{p_{s,i}T_{f,i}}{p_{s,j}T_{f,j}} &\approx \frac{W_{0,j}}{W_{0,i}} \frac{T_{f,i}}{T_{f,j}} \\ &= \frac{(\alpha/2)\,\mathrm{CW}_{\mathrm{adv}}}{\alpha \mathrm{CW}_{\mathrm{adv}}} \frac{\left(T_{\mathrm{oh}} + \beta\left(L_{\mathrm{ref}}/\widetilde{R}_{\min}\right)\right)}{\left(T_{\mathrm{oh}} + (\beta/2)\left(L_{\mathrm{ref}}/\widetilde{R}_{\min}\right)\right)} \approx 1, \end{aligned} \quad (32)$$

from (4) and (27) and the assumption that $T_{\mathrm{oh}} \ll L_{\mathrm{ref}}/\widetilde{R}_{\min}$. Note that (32) holds generally for any values of α and β. Consequently, this result asserts that the airtime consumed for successful transmission of the aggregated frame by each STA is almost the same, regardless of transmission rate, frame size, and the control parameters α and β. □

Proposition 5. *The proposed hybrid mechanism assures that per-station throughput becomes proportional to the transmission rate, regardless of frame size.*

Proof. From (28) and (31),

$$\frac{\mathrm{th}_i}{\mathrm{th}_j} = \frac{p_{s,i}L_{p,i}}{p_{s,j}L_{p,j}} \approx \frac{W_{0,j}}{W_{0,i}} \frac{L_{p,i}}{L_{p,j}}. \quad (33)$$

As long as $i, j (\neq i) \in \mathbb{S}_{\mathbb{LO}}$, $W_{0,i} = W_{0,j}$ from (4) and (33) can be rewritten from (29) as

$$\frac{\mathrm{th}_i}{\mathrm{th}_j} \approx \frac{L_{p,i}}{L_{p,j}} = \frac{R_i}{R_j}. \quad (34)$$

When $i, j (\neq i) \in \mathbb{S}_{\mathbb{HI}}$, this property holds in the same way. If $i \in \mathbb{S}_{\mathbb{LO}}$ and $j \in \mathbb{S}_{\mathbb{HI}}$, (33) becomes from (4) and (29)

$$\frac{\mathrm{th}_i}{\mathrm{th}_j} \approx \frac{(\alpha/2)\,\mathrm{CW}_{\mathrm{adv}}}{\alpha \mathrm{CW}_{\mathrm{adv}}} \frac{\left(\beta L_{\mathrm{ref}}\left(R_i/\widetilde{R}_{\min}\right)\right)}{\left((\beta/2)\,L_{\mathrm{ref}}\left(R_j/\widetilde{R}_{\min}\right)\right)} = \frac{R_i}{R_j}. \quad (35)$$

Table 1: IEEE 802.11 PHY/MAC parameters used in simulations.

Parameters	Value
Packet size	1500 bytes
MAC header with MPDU subframe header	38 bytes
Block ACK frame size	30 bytes
PHY preamble and header time	32 μs
DIFS, SIFS	34, 16 μs
T_{slot}	9 μs
$\mathrm{CW}_{\min}$, $\mathrm{CW}_{\max}$	16, 1024

Therefore, the per-STA throughput becomes proportional to the transmission rate and it is not affected by the frame size. □

5. Simulation

In this section, we validate the analytical model presented in Section 4 by comparing the analysis results with simulation results. Also, we extensively compare the performance of the proposed mechanism with existing approaches in terms of throughput, utilization, and fairness.

5.1. Simulation Setup. We implemented the simulator that models IEEE 802.11 MAC/PHY layers with MATLAB. The system parameters are listed in Table 1. In simulations, we assumed that every STA always has sufficiently large number of packets to send and the transmission rate for each STA is fixed. The values of parameters α, β, and γ in the proposed mechanism were set to 1, 2, and 4, respectively, according to the design guideline discussed in Section 3.4. The simulation time was set to one million slot times.

We consider the following four scenarios for performance evaluation.

(i) *Scenario 1.* There exist four STAs and each STA has different transmission rate but the same packet size; that is, $N_{\mathrm{sta}} = M = 4$, $R_i = 6, 12, 24, 48$, Mb/s, and $L_{\mathrm{pkt},i} = 1500$ bytes for $i = 1, 2, 3, 4$.

(ii) *Scenario 2.* There are two classes and one STA per class; that is, $N_{\mathrm{sta}} = M = 2$, and each STA has different values of transmission rate and/or packet size.

(iii) *Scenario 3.* There exist four classes and multiple STAs per class; $R_1 = 6$ Mb/s and n_1 ranges between 4 and 14, while $R_i = 12, 24, 48$ Mb/s and $n_i = 4$ for $i = 2, 3, 4$.

(iv) *Scenario 4.* In contrast to Scenario 3, $R_4 = 48$ Mb/s and n_4 ranges between 4 and 14, while $R_i = 6, 12, 24$ Mb/s and $n_i = 4$ for $i = 1, 2, 3$.

We evaluate the performance with the following three indices.

(i) *Throughput.* The per-STA throughput is defined as $(L_{\mathrm{pkt},i} \cdot N_{\mathrm{pkt},i})/T_{\mathrm{sim}}$, where $N_{\mathrm{pkt},i}$ is the number of packets successfully transmitted during the simulation time of T_{sim}. The aggregate throughput is the sum of per-STA throughputs.

(ii) *Utilization.* This is defined as the total effective transmission time ($T_{f,i}$) for all STAs divided by the whole

Table 2: Performance validation of the hybrid mechanism as a solution to the problem of performance anomaly (Scenario 1).

Criteria	DCF	Hybrid
Per-STA throughput (Mb/s)		
STA1 ($R_1 = 6$ Mb/s)	2.142	1.267
STA2 ($R_2 = 12$ Mb/s)	2.134	2.531
STA3 ($R_3 = 24$ Mb/s)	2.141	5.047
STA4 ($R_4 = 48$ Mb/s)	2.148	10.713
Aggregate throughput (Mb/s)	8.566	19.558
Utilization	0.801	0.919
Fairness index	0.726	0.997

simulation time. The transmission time excludes the time for backoff and unsuccessful transmission but includes the transmission time for PHY/MAC header and ACK.

(iii) *Fairness Index*. This is the modified version of Jain's fairness index (FI) [21] defined as

$$\text{FI} = \frac{\left(\sum_{i=1}^{M}\sum_{j=1}^{n_i} T_{f(i,j)}\right)^2}{N_{\text{sta}}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\left(T_{f(i,j)}\right)^2}, \tag{36}$$

where $T_{f(i,j)}$ is the total successful transmission time of the aggregated frame by the jth STA in the class i. The value of FI becomes one in the ideal case where all the STAs have the same channel occupation time.

5.2. Performance Validation of the Hybrid Mechanism. The objective of this simulation is to validate the problem of performance anomaly in multirate WLANs and then to confirm that the proposed mechanism resolves this problem effectively. We compare the performances of DCF and the hybrid mechanism in Scenario 1. As shown in Table 2, in the case of DCF, all the STAs achieve the similar throughput, regardless of transmission rate, which degrades aggregate throughput and FI remarkably. However, the hybrid mechanism lets per-STA throughput become almost proportional to the transmission rate by assuring airtime fairness; that is, the value of FI is almost one. In addition, compared to DCF, the proposed mechanism increases the aggregate throughput by more than two times and improves the channel utilization by more than 10%. This performance gain of the hybrid mechanism is derived from the decrease of collision probability and overhead time by controlling the values of CW and AF.

We further investigate the performance of the hybrid mechanism in terms of airtime fairness. For this purpose, we perform simulations under Scenario 2, where two STAs have different transmission rates and packet sizes as indicated in Table 3. Here, we set L_{ref} to 1000 bytes. We observe the throughput ratio between two STAs to evaluate fairness with the hybrid mechanism. Table 3 shows the throughput ratios between two STAs obtained from analysis and simulation, each of which is denoted as γ_{anal} and γ_{sim}, respectively. Table 3 also shows the ideal value of throughput ratio, γ_{ideal}, which is independent of packet size and depends only on the ratio of transmission rates as proved in Proposition 5 (see (34) and (35)). In Table 3, the percentage error between γ_{anal} and γ_{sim} is calculated as

$$e_\gamma(\%) = \frac{|\gamma_{\text{anal}} - \gamma_{\text{sim}}|}{\gamma_{\text{ideal}}} \times 100. \tag{37}$$

From Table 3, we observe the following.

(i) For all the sets of (R_i, R_j) and $(L_{\text{pkt},i}, L_{\text{pkt},j})$, γ_{sim} is almost equal to γ_{anal}; that is, e_γ does not exceed 0.55%.

(ii) The proposed mechanism works well even though the ratio between R_i and R_j is not an integer.

(iii) The throughput ratio depends on the ratio of R_i and R_j but is hardly affected by the packet size.

(iv) The value of FI is almost equal to the ideal value of one in most cases; the airtime fairness is strictly assured, regardless of the differences in the transmission rate and packet size.

Note that there is negligible difference between the values of γ_{ideal}, γ_{anal}, and γ_{sim} when $\gamma_{\text{ideal}} \le 4 (= \gamma)$, but the value of γ_{anal} or γ_{sim} is somewhat deviated from that of γ_{ideal} when $\gamma_{\text{ideal}} > 4$. This difference results from the intergroup collision, which makes the approximation error in (31). As long as the STAs belong to the same group (i.e., the ratio of transmission rates between two STAs is not greater than $\gamma (= 4)$), the proposed mechanism differentiates only the value of AF while keeping the same value of CW. In this case, the value of success probability ($p_{s,i}$) is almost equal for all the STAs. However, when STAs belong to different groups, the intergroup collision happens, which leads to a small difference between $p_{s,i}W_{0,i}$ and $p_{s,j}W_{0,j}$. Even in this case of $\gamma_{\text{ideal}} > 4$, there is negligible difference between the values of γ_{anal} and γ_{sim}. These results in Table 3 confirm the validity of the analysis model and the effectiveness of the proposed mechanism in terms of airtime fairness.

5.3. Validation of Analysis Model. We validate the analysis model of the hybrid mechanism by observing per-STA throughput, aggregate throughput, and utilization. The simulations were performed under Scenario 3 and Scenario 4; the results of Scenario 4 were quite similar to those of Scenario 3, so we only present the results of Scenario 3.

We first investigate per-STA throughput of the proposed mechanism. Let us define $\text{th}_{\text{anal},i}$ as the throughput of STA in class i, which is obtained from the analysis model in Section 4 and is identical for all n_i STAs belonging to class i. Also, we define $\overline{\text{th}}_{\text{sim},i}$ as the average value of per-STA throughputs for class-i STAs obtained from simulations. We found from simulations that the difference of throughput among n_i STAs in class i is insignificant and it does not exceed 0.05 Mb/s for all the cases. Figure 3(a) compares $\text{th}_{\text{anal},i}$ with $\overline{\text{th}}_{\text{sim},i}$, each of which is represented with lines and marks, respectively. We observe from Figure 3(a) that $\overline{\text{th}}_{\text{sim},i}$ is almost equal to $\text{th}_{\text{anal},i}$ and that they decrease as N_{sta} increases. Moreover, the results in Figure 3(a) confirm that STAs achieve their throughputs in proportion to the transmission rate.

Table 3: Per-STA throughput of the hybrid mechanism with different transmission rates and packet sizes (Scenario 2).

(R_i, R_j)	$(L_{\text{pkt},i}, L_{\text{pkt},j})$	Throughput ratio (th_i/th_j) ideal (γ_{ideal})	anal. (γ_{anal})	sim. (γ_{sim})	Percentage error (e_γ)	Fairness index
(48, 6)	(1000, 1000)	8	9.088	9.097	0.111	0.9994
(48, 6)	(1000, 1500)	8	9.088	9.077	0.136	0.9991
(48, 9)	(1500, 500)	5.333	6.059	6.058	0.008	0.9997
(54, 12)	(500, 1000)	4.5	5.112	5.087	0.551	0.9960
(48, 12)	(1500, 500)	4	4.000	4.000	0.008	0.9996
(24, 6)	(1000, 1000)	4	4.000	3.996	0.105	1.0000
(24, 9)	(1500, 1000)	2.667	2.667	2.671	0.146	1.0000
(24, 12)	(500, 1500)	2	2.000	1.992	0.040	0.9996
(12, 6)	(1500, 1000)	2	2.000	2.000	0.000	1.0000
(18, 12)	(1500, 1000)	1.5	1.500	1.500	0.000	1.0000
(9, 6)	(1500, 500)	1.5	1.500	1.500	0.007	0.9995

Next, we focus on the aggregate throughput and utilization. We observe from Figure 3(b) that the simulation results agree well with the analysis results; the differences of aggregate throughput and utilization obtained from analysis and simulation are at most 0.15 Mb/s and 0.005, respectively. The aggregate throughput decreases with respect to the increase of N_{sta}, which results from the increase of n_1 (the number of 6-Mb/s STAs). Unlike the aggregate throughput, the utilization is almost immune to the change of N_{sta}. For the entire range of N_{sta}, the value of utilization obtained from simulations lies between 0.909 and 0.914; that is, the fraction of time consumed due to backoff or collision is less than 10%, regardless of N_{sta}. This result indicates that, by scaling the value of CW based on N_{sta}, the proposed mechanism controls the collision probability and the idle time due to backoff such that they do not increase with respect to the increase of N_{sta} and they are maintained almost constantly.

5.4. Performance Comparison. From now, we compare the performance of the hybrid mechanism with those of existing mechanisms. We consider the following mechanisms.

(i) `DCF`. This is the baseline scheme without any changes to improve throughput and fairness.

(ii) `CW-DIFF`. This mechanism differentiates the value of $\text{CW}_{\min}$ inversely proportional to the transmission rate, as similar to [7, 8].

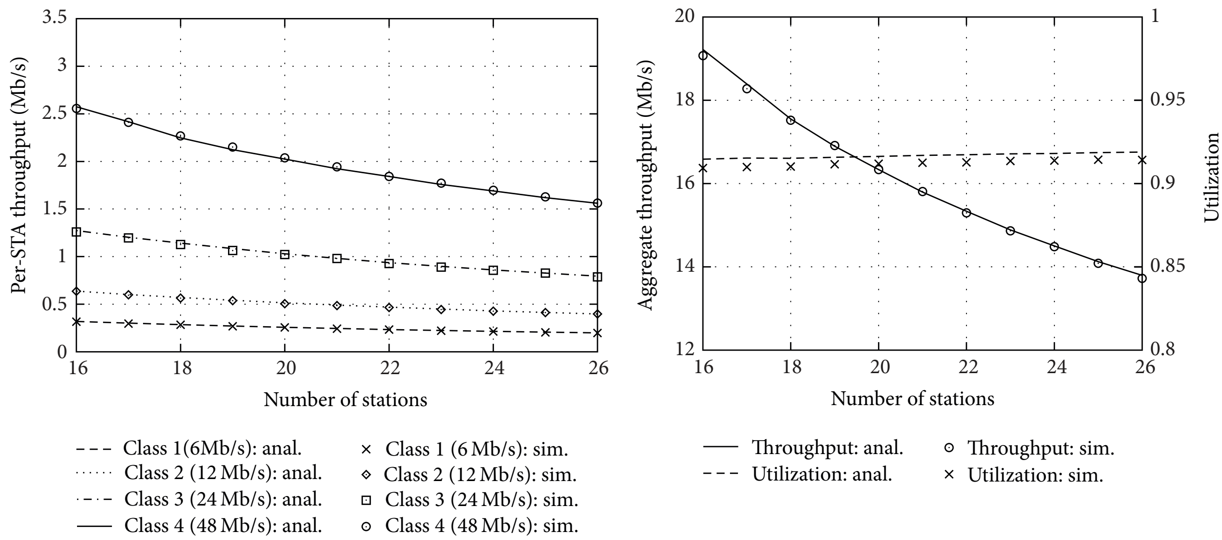
(iii) `TXOP`. In this mechanism, multiple packets are transmitted within the TXOP limit, which is set to be equal for all the STAs.

(iv) `HYBRID`. This is the proposed mechanism that integrates the control of CW with frame aggregation.

Note that `CW-DIFF` or `TXOP` can be considered as a representative mechanism for airtime fairness that controls either the probability of channel access or the time of channel occupation, respectively, and `HYBRID` integrates both approaches. For fair comparison between `HYBRID` and other mechanisms, the vale of $\text{CW}_{\min}(R_{\max})$ (see (1)) in `CW-DIFF` is set to the same value of $\text{CW}_{\min}$ in `HYBRID` and the value of TXOP limit in `TXOP` is set to $\beta L_{\text{ref}}/R_{\min}$, which is comparable to the transmission time of aggregated frame in `HYBRID`.

5.4.1. Aggregate Throughput. Figures 4(a) and 4(b) show the aggregate throughput of several schemes in Scenario 3 and Scenario 4, respectively. In both scenarios, `HYBRID` significantly outperforms the other schemes; the throughput of `HYBRID` is higher than those of `DCF`, `CW-DIFF`, and `TXOP` by up to 3.18, 1.93, and 1.72 times, respectively. The outstanding performance of `HYBRID` stems from its features; firstly, it reduces the collision probability by scaling CW with respect to the number of STAs and secondly it uses the frame aggregation which in turn decreases several overheads in the process of channel access. In Scenario 3, the aggregate throughput decreases for all the schemes as n_1 (the number of 6-Mb/s STAs) increases; however, it increases as n_4 (the number of 48-Mb/s STAs) increases in Scenario 4. In Scenario 3, there is no significant difference between the throughputs of `CW-DIFF` and `TXOP`; however, the throughput of `CW-DIFF` is somewhat lower than that of `TXOP` as n_4 increases in Scenario 4. This result can be explained as follows. `CW-DIFF` fixes the value of CW for high-rate STAs as $\text{CW}_{\min}$, while it increases that of CW for low-rate STAs inversely proportional to the transmission rate. Thus, `CW-DIFF` exacerbates collisions as n_4 increases in Scenario 4, leading to the decrease of throughput. On the other hand, `TXOP` improves the throughput compared to `CW-DIFF` as it reduces the channel access overheads. However, `TXOP` also suffers from frequent collisions as the number of STAs increases. The relative throughput gain of `HYBRID` over the other schemes is almost constant in Scenario 3; the throughput of `HYBRID` is higher than those of `CW-DIFF` and `TXOP` by about 56%~58% and 52%~58%, respectively; however, in Scenario 4, the relative gain of `HYBRID` over `CW-DIFF` and `TXOP` increases from 56% and 52% when $n_4 = 4$ to 93% and 72% when $n_4 = 14$, respectively. The reason is that `HYBRID` increases the throughput not only by decreasing the channel access overheads due to frame aggregation but also by reducing the collision probability due to the control of CW.

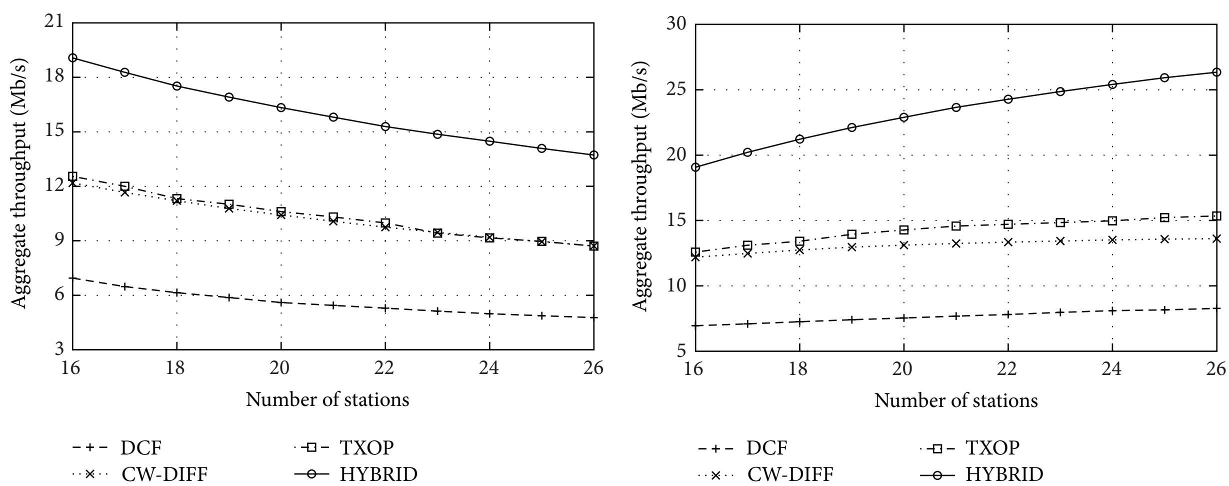
5.4.2. Utilization. We observe the utilization of several comparative schemes from Figure 5. The most noteworthy point is that the utilization of `HYBRID` is almost immune to the change of N_{sta} in both scenarios; however, those of the other schemes decrease as N_{sta} increases. `HYBRID` maintains higher



(a) Per-STA throughput

(b) Aggregate throughput and utilization

Figure 3: Comparison of analysis results with simulation results in the hybrid mechanism (Scenario 3).



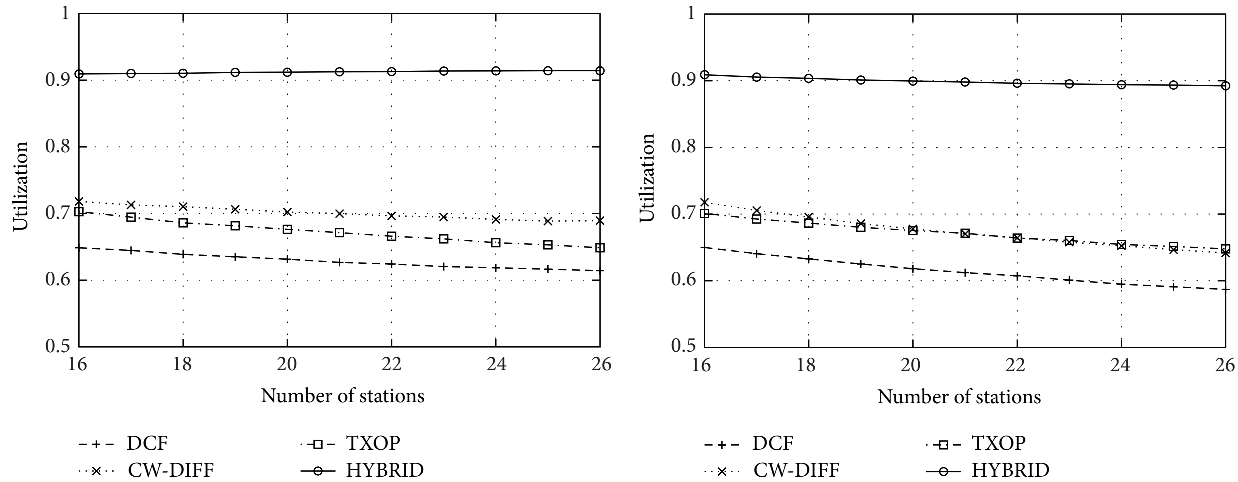
(a) Scenario 3

(b) Scenario 4

Figure 4: Comparison of aggregate throughput for several schemes.

utilization than the other schemes; it is at least 89% for the entire range of N_{sta} in both scenarios, but the utilization of CW–DIFF and TXOP is at most 70% for most cases and that of DCF does not exceed 65%. The decrease of utilization in the schemes except for HYBRID is mainly due to the increase of collision probability with respect to the increase of N_{sta}. In Scenario 3, the utilization of CW–DIFF is somewhat higher than that of TXOP (see Figure 5(a)), which results from the decrease of collision probability due to the increase of CW for low-rate STAs; however its throughput is almost equal to that of TXOP (see Figure 4(a)), implying that the increase of CW in low-rate STAs increases the backoff time and makes the negative effect on the throughput. On the other hand, the utilization of HYBRID is robust to the change of N_{sta}, because HYBRID effectively controls the collision probability and backoff time thanks to the adjustment of CW and frame aggregation.

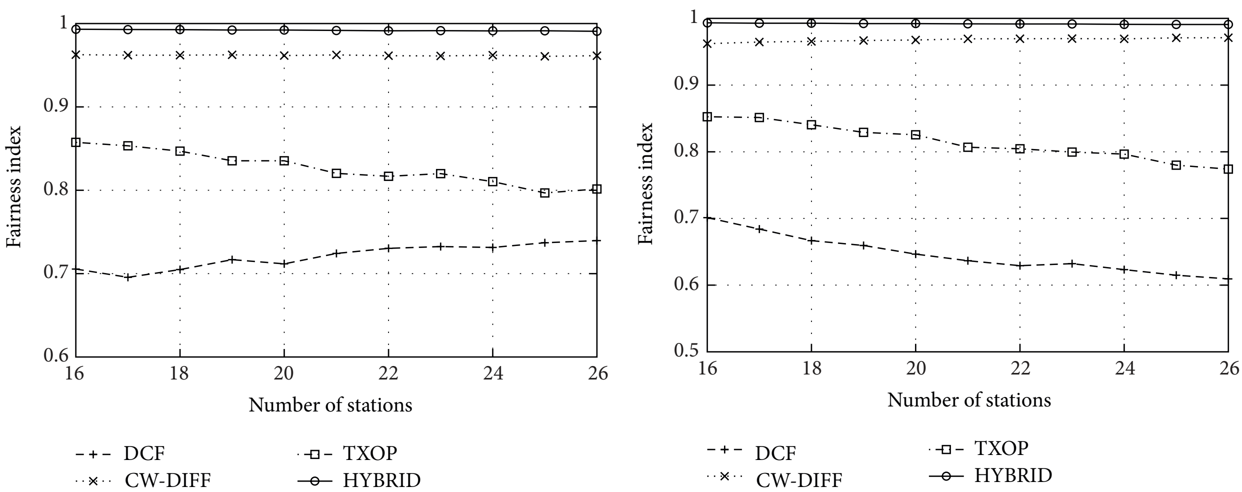
5.4.3. Fairness. Now, we compare the performance of several schemes in terms of fairness index from Figure 6. For the entire range of N_{sta} in both scenarios, HYBRID maintains the value of FI close to the ideal value (i.e., it does not fall below 0.991); moreover it is hardly affected by N_{sta}. By comparing Figure 6 with Figure 4, we observe an interesting result regarding the trade-off between throughput and fairness in CW–DIFF and TXOP; although the throughput of CW–DIFF is mostly lower than or equal to that of TXOP, the FI of CW–DIFF is quite higher than that of TXOP and it is maintained constantly around 0.96~0.97. From these results, we can conclude that, compared to TXOP, CW–DIFF is effective to



(a) Scenario 3

(b) Scenario 4

Figure 5: Comparison of utilization for several schemes.



(a) Scenario 3

(b) Scenario 4

Figure 6: Comparison of airtime fairness for several schemes.

assure airtime fairness at the cost of throughput decrease. The value of FI in `TXOP` decreases with respect to the increase of N_{sta}; it decreases from 0.86 and 0.85 to 0.80 and 0.77 as N_{sta} increases from 16 to 26 in Scenarios 3 and 4, respectively. This is related to the operation of `TXOP` in response to collisions. As N_{sta} increases, the collision occurs frequently and the collision in the first frame makes the STA not fully utilize the remaining time within the TXOP limit, and thus the airtime fairness becomes degraded. Unlike `CW–DIFF` and `TXOP`, the FI of `DCF` shows different aspects between Scenarios 3 and 4; it rather increases as n_1 increases in Scenario 3 but decreases as n_4 increases in Scenario 4. As n_1 increases in Scenario 3, the probability of collision between low-rate STAs may increase and the successful channel occupation time by low-rate STAs decreases accordingly, which contributes to improving airtime fairness. In Scenario 4, the result becomes opposite to Scenario 3; that is, the increased collision between high-rate STAs further deteriorates the airtime fairness.

6. Conclusion

In this paper, we studied the problem of performance degradation in multirate WLANs and proposed a simple yet effective solution to this problem. The proposed mechanism aims to assure airtime fairness and to improve aggregate throughput at the same time. In order to achieve this objective, it integrates the control of CW with the frame aggregation scheme and applies the differentiated control strategy between high-rate STAs and low-rate STAs. The size of CW is controlled from two aspects; it is firstly scaled depending on the number of STAs to effectively decrease the overall collision probability and secondly differentiated depending on the group to which the STA belongs to improve airtime fairness. Moreover, the number of packets aggregated is controlled not only to increase throughput by decreasing several channel access overheads but also to strictly assure airtime fairness by supporting intra/intergroup differentiation. As a result, the proposed hybrid mechanism controls both channel access probability and channel occupation time and it improves fairness and efficiency of channel sharing among STAs. Also, we derived the rigorous analysis model based on Markov chain and proved the desirable properties of the hybrid mechanism. Together with the analysis results, the simulation results confirmed the outstanding performance of the proposed mechanism in terms of throughput, utilization, and fairness.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science, and Technology, under Grant NRF-2013R1A2A2A01009883.

References

[1] IEEE 802.11 Working Group, "Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications," IEEE Std. 802.11-2007, 2007.

[2] IEEE 802.11 Working Group, "Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications. Amendment 5: enhancements for higher throughput," IEEE Std. 802.11n-2009, 2009.

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

[4] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications*, vol. 2, pp. 836–843, IEEE, San Francisco, Calif, USA, April 2003.

[5] D.-Y. Yang, T.-J. Lee, K. Jang, J.-B. Chang, and S. Choi, "Performance enhancement of multirate IEEE 802.11 WLANs with geographically scattered stations," *IEEE Transactions on Mobile Computing*, vol. 5, no. 7, pp. 906–919, 2006.

[6] B. Li, R. Battiti, and Y. Fang, "Achieving optimal performance by using the IEEE 802.11 MAC protocol with service differentiation enhancements," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 3, pp. 1374–1387, 2007.

[7] H. Kim, S. Yun, I. Kang, and S. Bahk, "Resolving 802.11 performance anomalies through QoS differentiation," *IEEE Communications Letters*, vol. 9, no. 7, pp. 655–657, 2005.

[8] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless LANs," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 4, pp. 121–132, 2005.

[9] C.-T. Chou, K. G. Shin, and S. N. Shankar, "Contention-based airtime usage control in multirate IEEE 802.11 wireless LANs," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1179–1192, 2006.

[10] Y. Chetoui and N. Bouabdallah, "Adjustment mechanism for the IEEE 802.11 contention window: an efficient bandwidth sharing scheme," *Computer Communications*, vol. 30, no. 13, pp. 2686–2695, 2007.

[11] A. Banchs, P. Serrano, and H. Oliver, "Proportional fair throughput allocation in multirate IEEE 802.11e wireless LANs," *Wireless Networks*, vol. 13, no. 5, pp. 649–662, 2007.

[12] T. Joshi, A. Mukherjee, Y. Yoo, and D. P. Agrawal, "Airtime fairness for IEEE 802.11 multirate networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 4, pp. 513–527, 2008.

[13] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "OAR: an opportunistic auto-rate media access protocol for ad hoc networks," *Wireless Networks*, vol. 11, no. 1-2, pp. 39–53, 2005.

[14] J. del Prado Pavón and S. N. Shankar, "Impact of frame size, number of stations and mobility on the throughput performance of IEEE 802.11e," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC '04)*, vol. 2, pp. 789–795, March 2004.

[15] M. Kim, E.-C. Park, and C.-H. Choi, "Adaptive two-level frame aggregation for fairness and efficiency in IEEE 802.11n wireless LANs," *Mobile Information Systems*, vol. 2015, Article ID 548109, 14 pages, 2015.

[16] A. V. Babu and L. Jacob, "Fairness analysis of IEEE 802.11 multirate wireless LANs," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 3073–3088, 2007.

[17] M. Hui and S. Roy, "Contention window and transmission opportunity adaptation for dense IEEE 802.11 WLAN based on loss differentiation," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 2556–2560, May 2008.

[18] M. Adnan and E.-C. Park, "Assuring per-station fairness in multi-rate WLANs: a hybrid approach of contention window control and frame aggregation," in *Proceedings of the 5th International Conference on Ubiquitous and Future Networks (ICUFN '13)*, pp. 282–287, Da Nang, Vietnam, July 2013.

[19] IEEE 802.11 Working Group, "Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications. Amendment 8: medium access control (MAC) quality of service enhancements," IEEE Std. 802.11e-2005, 2005.

[20] F. Calì, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Transactions on Networking*, vol. 8, no. 6, pp. 785–799, 2000.

[21] D.-M. Chiu and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Computer Networks and ISDN Systems*, vol. 17, no. 1, pp. 1–14, 1989.