Enhancing QoS of Voice over WLANs

Byoungjin Kim, Hyewon Lee, Seongho Byeon, Kwang Bok Lee, and Sunghyun Choi
School of Electrical & Computer Engineering and INMC, Seoul National University, Seoul, Korea
Email: {bjkim, hwlee, shbyeon}@mwnl.snu.ac.kr, {klee, choi}@snu.ac.kr

Abstract—Voice over IP (VoIP) service is emerging as a popular application in wireless local area networks (WLANs). However, the existing rate adaptation (or link adaptation) schemes for WLANs have been developed to improve the throughput of data traffic, not the quality of service (QoS) of VoIP traffic. In this paper, we investigate the characteristics of VoIP traffic and the limitations of state-of-the-art rate adaptation schemes, and then enhance the QoS of voice over WLANs (VoWLANs) by ameliorating the existing rate adaptation schemes. Specifically, we design fast decrease to control the transmission rate of retransmissions, and retry scheduling to avoid the deep fading of the wireless channel as well as hidden terminal interference. We comparatively evaluate the QoS of the revised rate adaptation schemes via ns-3 simulations and MadWiFi implementations in various communication environments, and demonstrate that the proposed scheme improves the R-score performance by up to 80% depending on the network scenarios.

Index Terms—IEEE 802.11, WLAN, voice over IP (VoIP), quality of service (QoS), rate adaptation

I. INTRODUCTION

Today’s numerous mobile devices such as laptops, smartphones, and mobile tablet personal computers have been adopting wireless local area network (WLAN) technology as their wireless access interfaces. The IEEE 802.11 standard, often referred to as WiFi [1], has been evolving accordingly in order to comply with the increased needs of users, and various types of services over WLAN are being provided today. One of the most popular services is voice over WLAN (VoWLAN) such as Internet phones and Skype, and numerous researches have studied provisioning quality of service (QoS) for VoWLANs [2–6].

Currently, IEEE 802.11a/b/g physical (PHY) layer supports a number of PHY transmission rates (or PHY rates) with different throughput and error rate performances [7, 8]. It is generally known that low transmission rates are more robust than high transmission rates, while high transmission rates can provide WLAN with higher throughput. Hence, rate adaptation or link adaptation, which adaptively selects a PHY transmission rate with a specific objective in a given communication environment, has been an important research topic in WLANs [9–15]. The objective of a rate adaptation scheme actually can vary with communication environments, e.g., traffic type to be served, number of users, metric for network performance, etc. It should be noted that, however, most of up-to-date researches on rate adaptation focus on maximizing the throughput of WLAN, not on guaranteeing the QoS of VoWLAN.

IEEE 802.11 medium access control (MAC) standard defines retransmissions of MAC frames in order to provide a better frame delivery probability against channel errors and/or frame collisions. Upon a frame transmission failure, which is recognized by an absence of corresponding acknowledgement (ACK) frame response, MAC frames can be retransmitted up to retry limit times. However, continuous retransmissions without care could be detrimental in some cases, as the QoS of voice traffic is known to be very sensitive to frame losses. For example, let us assume that a frame transmission fails due to the deep fading of the wireless channel and/or severe hidden terminal interference, where both of these events can frequently happen in practical WLAN environments. If a WLAN station continues to retransmit the failed frame within a short time interval in which the channel status remains poor, then the frame transmissions will continue to fail, and the frame will be dropped eventually after a finite number of retransmissions. Thus, considering that the voice traffic is sensitive to frame losses, continual frame losses at the MAC layer can severely degrade the QoS of VoWLAN.

In this paper, we investigate state-of-the-art rate adaptation schemes and discuss their QoS provisioning capabilities for VoWLAN. We also propose simple yet effective modifications to the existing rate adaptation schemes to improve the QoS of VoWLAN. We further enhance the QoS of VoWLAN by scheduling the frame retransmissions, so that WLAN stations can cope with the deep fading and hidden terminal interference, and provide a better delivery ratio of VoWLAN frames.

The rest of the paper is organized as follows. We first introduce the VoWLAN and investigate the existing rate adaptation schemes of the WLAN in Section II. Section III proposes two novel features for VoWLAN, namely, fast decrease and retry scheduling, to improve the QoS. In Section IV, we comparatively evaluate the QoS of the proposed schemes in various communication environments via network simulator (ns-3) and MadWiFi implementation result [16, 17]. Finally, Section V discusses the effectiveness of the proposed schemes on enhancing the QoS of VoWLAN as well as further extensions, and then concludes the paper.

II. REVISIT OF VOICE OVER WLANS

As the VoWLAN service is getting popular, there have been intensive researches on the VoWLAN, e.g., the call admission control of VoIP sessions [2–4], and the capacity enhancement of VoWLANs [5, 6]. In this paper, we specifically focus on the rate adaptation issue of the VoWLAN in order to enhance the QoS.

We first study the characteristics of VoIP traffic as well as its performance metric, R-score. Then we investigate the existing rate adaptation algorithms for WLANs, and discuss their operations and limitations. We also explain the retransmission policy of the IEEE 802.11 MAC and how this affects the QoS of VoWLAN.
A. Characteristics of VoIP Traffic

Unlike the typical objective of non-real-time data traffic, i.e., maximizing its throughput, real-time VoIP traffic requires guaranteed QoS. Aiming at provisioning high QoS VoWLAN service, we first study the characteristics of VoIP traffic, and then discuss the means to evaluate its QoS.

1) Voice codec and mouth-to-ear delay: For VoIP, an analog voice signal is sampled and encoded using a voice codec such as ITU-T G.711/729 [18, 19], or Internet low bit-rate codec (iLBC) [20], into a digital bit stream. The encoded signal is then packetized periodically and then is delivered to the MAC layer as a frame payload. Without silence suppression, which detects silent durations of human speech and does not generate any voice packet during that interval, the voice traffic is basically constant bit rate (CBR) traffic, i.e., voice packets of a fixed size are generated and transmitted periodically. After packet deliveries through a network, however, each voice packet experiences a different latency, and hence, the inter-packet interval of the received voice traffic varies over time, referred to as jitter. In order to compensate such jitters, a VoIP receiver uses a de-jitter buffer, which first enqueues a number of voice packets and then starts outputting them with a constant inter-packet interval. Then the VoIP packets are de-packetized, converted to an analog voice signal, and finally played back to a listener. In consequence, a voice signal experiences mouth-to-ear delay, which denotes the latency that a voice signal takes from a speaker’s mouth to a listener’s ear, consisting of packetization, processing, de-jitter buffer, wireline/wireless transmission delays, etc.

Fig. 1 illustrates how voice packets are generated, and what a mouth-to-ear delay consists of. It should be noted that each voice packet eventually experiences a fixed delay, which is the delay which the very first voice packet experiences, due to the de-jitter buffer. A voice packet experiencing an excessive delay, i.e., a delay longer than the fixed delay, cannot be played back to the listener, and is equivalent to a packet loss. G.711 codec, which is the simplest yet widely used voice codec, generates 64 kbps voice traffic with a constant packet size and interval [18]. In this paper, we assume a typical setting for G.711, i.e., 208-byte MAC frame payload size with 20 ms inter-packet interval and the de-jitter buffer size of 3 packets. We do not consider the silence suppression as it is known to deteriorate the quality of experience at the user side. Accordingly, the packetization and de-jitter buffer delays are assumed to be 20 ms and 60 ms, respectively, and we ignore the processing delay which could be minimized depending on the computing power of VoIP devices.

2) Performance metric of VoIP: ITU-T Recommendation G.107 specifies E-model, which provides a convenient and objective quality metric, R-score. In the case of G.711 voice codec, R-score is expressed by the following simplified equa-

\[ R = 93.2 - I_{\text{delay}} - I_{\text{loss}}, \]

where \( I_{\text{delay}} \) and \( I_{\text{loss}} \) present impairment factors due to mouth-to-ear delay and packet losses of VoIP traffic, respectively [22], and R-score greater than or equal to 80 is generally regarded as good voice quality. In this paper, we adopt E-model to evaluate the quality of VoIP traffic.

According to E-model given in Eq. (1), voice quality is mainly affected by the packet loss rate rather than the mouth-to-ear delay as long as the delay is under 150 ms. Since today’s Internet can normally provide negligible packet losses and the latency less than 35 ms [23, 24], thus making it feasible to achieve the mouth-to-ear delay under 150 ms, we conclude that managing the packet losses over the wireless access network, i.e., WLAN, is very important to guarantee high QoS for VoWLAN.

B. Existing Rate Adaptation Schemes

Today’s IEEE 802.11 PHY standards provide various transmission rates, e.g., the IEEE 802.11a PHY supports 8 transmission rates (6, 9, 12, 18, 24, 36, 48, and 54 Mbps), with different throughput and error rate. Based on this multi-rate support, numerous researchers have been studying the rate adaptation of WLANs. We now introduce state-of-the-art rate adaptation schemes and their operations.

1) Automatic Rate Fallback (ARF) and Adaptive ARF (AARF): Automatic Rate Fallback (ARF) is developed for WaveLan II devices, and adopts a simple history-based rate adaptation algorithm [9]. Whenever \( N \) consecutive frames (referred to as success threshold, which is 10 frames by default) are successfully transmitted, or a timer to probe the next higher PHY rate expires, the transmitter increases the PHY rate by one step. If consecutive frames (2 frames by default) are lost, or the very first frame transmitted with the increased PHY rate (referred to as a probing frame) fails, then the transmitter decreases the PHY rate by one step.

M. Lacage et al. design adaptive ARF (AARF) that adaptively manages the success threshold [10]. Based on ARF, AARF mitigates the probing overhead by exponentially increasing the success threshold whenever a probing frame fails.

2) Collision-Aware Rate Adaptation (CARA): Based on the operation of ARF, J. Kim et al. propose collision-aware rate adaptation (CARA) algorithm that can empirically distinguish the channel and collision errors, so that the PHY rate can be
decreased only after a channel error [11]. In CARA, a WLAN station enables request-to-send (RTS) transmission after a frame transmission failure. Based on the transmission results of the RTS and the frame following the RTS transmission, the WLAN station can tell whether the previous frame loss could be due to a channel error or collision.

CARA is further improved to quickly respond to the channel dynamics in [12], by modifying the resetting rule of the success counter of ARF. In ARF and CARA, the success counter, which denotes the number of consecutive successful transmissions, is reset to zero whenever a frame transmission fails. In CARA-RI, however, the success counter is reset after two consecutive failures, i.e., only after the frame loss is confirmed to be due to a channel error.

3) Robust Rate Adaptation Algorithm (RRAA): Above-explained ARF-based rate adaptation algorithms, i.e., ARF, AARF, CARA, and CARA-RI, use the history of consecutive transmission trials. On the other hand, robust rate adaptation algorithm (RRAA) proposed in [13] facilitates frame error rate (FER) statistics. For each PHY rate, RRAA builds a table that contains (1) rate increasing/decreasing thresholds in terms of FER assuming 1300 bytes frame length, and (2) estimation windows (ewnd) in terms of the number of frames. After transmitting as many as ewnd frames with a specific PHY rate, a WLAN station evaluates the FER and compares it with the rate increasing/decreasing thresholds; if the evaluated FER is smaller (greater) than the rate increasing (decreasing) threshold, then the WLAN station increases (decreases) the PHY rate by one step. RRAA also improves the performance of WLAN by using an adaptive usage of RTS (A-RTS). A-RTS of RRAA is more sophisticated than CARA’s adaptive usage of RTS, and turns out to be more powerful, especially, when there exist hidden terminals.

4) SampleRate and Minstrel: SampleRate also utilizes the FER statistics to select a PHY rate [14]. Based on the statistics of the average FER of each PHY rate, SampleRate estimates the expected transmission times including retransmissions with the given PHY rate and frame size. SampleRate then selects the PHY rate with the minimum expected transmission time, i.e., the PHY rate that maximizes the expected throughput. In order to collect an enough number of samples with various PHY rates, SampleRate probes a randomly selected PHY rate whenever a probing timer expires.

In an open-source WLAN device driver, MadWiFi [17], a revised version of SampleRate, called Minstrel [15], is implemented. Compared with SampleRate which does not specify the PHY rates that are used in retransmissions in its original proposal [14], Minstrel defines retry chain that specifies the PHY rates of retry frames based on the retransmission procedure of MadWiFi. Minstrel also adopts the moving average with a 100 ms window size in order to exclude the stale samples.

5) PHY-assisted Rate Adaptations: PHY layer information can also aid the selection of PHY rates. SoftRate [25] fetches bit error rate values, and AccuRate [26] captures symbol-level signal dispersions of constellation points from PHY. The weighted sum of each subcarrier’s SNR is used instead of packet-level SNR in [27]. These solutions, however, rely on the PHY information which might not be easy to access with most commercial WLAN devices, thus making them mostly unpractical today.

C. IEEE 802.11 MAC and QoS Provisioning

The IEEE 802.11 MAC standard defines how WLAN stations share the wireless medium, and deliver MAC frames to recipients. We briefly examine the MAC layer operation of WLANs and QoS provisioning features of the IEEE 802.11 standard [7], and discuss their influences on the QoS of VoWLAN.

1) Retransmission Policy of IEEE 802.11 MAC: The IEEE 802.11 enhanced distributed channel access (EDCA) MAC is based on the carrier sensing multiple access (CSMA) [7]. A WLAN station that attempts to transmit a frame senses the wireless medium to determine whether it is idle or not. If the wireless medium has been idle for longer than arbitrary inter frame space (AIFS), then the WLAN station begins a backoff process with a backoff counter which is randomly chosen from [0, CW], where CW is a contention window size. Only after the medium remains idle until the backoff counter expires, the WLAN station transmits a frame. If the transmission is not acknowledged with an ACK frame, i.e., the frame is determined to be lost, then the WLAN station retransmits the previous frame after another backoff. Beginning from the minimum CW size (CWmin), the CW value is doubled for every frame transmission failure, until the CW reaches the maximum CW size (CWmax).

Continuous retransmissions, however, could increase the frame loss rate in the following scenario. Suppose that a frame transmission fails due to a channel error. Depending on the random backoff counters, a WLAN station has a chance to retransmit the frame within a short time interval, and such a retransmission is likely to fail again if the channel status remains bad and the PHY rate is the same. This problem becomes even worse for the VoWLAN with the differentiated service of the EDCA MAC.

2) Access Category (AC) in IEEE 802.11: The IEEE 802.11 EDCA MAC provides differentiated services for real-time traffic types such as voice and video applications by classifying the types of traffic into 4 access categories (ACs), namely, voice (AC_VO), video (AC_VI), best effort (AC_BE), and background (AC_BK) [7]. The AC that requires high priority, e.g., AC_VO, has small AIFS/CWmin/CWmax values, so that the AC_VO traffic can more aggressively access the wireless medium than the other ACs.

Therefore, the time gaps between two consecutive frame retransmissions of the VoIP traffic with the IEEE 802.11 EDCA MAC could be shorter than a few milliseconds. The typical coherence time, which denotes the time duration that the channel state does not change, of the WLAN is known to be at least a few tens of milliseconds [28, 29]; so this few milliseconds time gap is not long enough for VoWLAN stations to exploit the time diversity. Hence, we conclude that we need not only a sophisticated PHY rate control for reducing the number of retransmissions, but also an appropriate retransmission scheduling for the reliable deliveries of retry frames.
III. ENHANCING QoS OF VoWLAN

In this section, we discuss the limitations of the existing rate adaptation schemes introduced in Section II, and propose two novel features, namely, fast decrease (FD) and retry scheduling (RS) for VoWLAN. FD tries to reduce the number of retransmissions by controlling the PHY rates of retry frames, and RS postpones excessive retransmissions in order to overcome the deep fading of the wireless channel.

A. Limitations of the Existing Rate Adaptation Schemes

ARF-based rate adaptation schemes, i.e., ARF, AARF, CARA, and CARA-RI, necessitate the results of consecutive frame transmissions, and this is known to be the main reason of their slow response, which can be more problematic for VoWLAN. Suppose the G.711 voice codec with 20 ms packet generation interval. In order for a WLAN station to increase the PHY rate by one step, the WLAN station has to transmit 10 consecutive frames successfully, which might take longer than 200 ms including contentions and retransmissions. For the rate decreasing operation, the PHY rate is decreased by one step for every two consecutive failures, which might not provide robust PHY rates in the deep fading channel. Hence we need more agile rate increasing/decreasing rules.

Fig. 2(a) presents a simulation example (ns-3) of the PHY rate increasing/decreasing process of CARA-RI, which works similar to ARF. The solid line depicts the SNR variation, and O and X signs denote the frame transmission successes and failures, respectively, where the inner figure shows the detailed events in the deep fading near \( t = 30.45 \) seconds. Note that the deep fading channel can barely support 6 Mb/s PHY rate, while CARA-RI sequentially tries 18, 12, 9, and 6 Mb/s PHY rates, and eventually drops the frame due to the retry limit expiration.

In RRAA, the FER of the current PHY rate is estimated by transmitting as many as ewnd frames in order to exclude occasional frame losses and gather statistically-meaningful samples. Therefore, RRAA cannot be fast enough to trace the fast-fading channel, especially due to the ewnd sizes which range from 6 to 40 frames depending on the PHY rate [13]. Note that in the case of 24, 36, 48, and 54 Mb/s of 802.11a, the corresponding ewnd size is 40 frames, which are equivalent to approximately 1 second for the VoIP traffic with 20 ms packet generation interval; this is too long time to adjust the PHY rate. Moreover, the default rate increase/decrease thresholds in [13] are designed for the data traffic with a 1300 byte frame, and these thresholds are not valid for the short VoIP frames.

Fig. 2(b) demonstrates an example of the PHY rate selection of RRAA. The inner figure magnifies the consecutive transmission failures from \( t = 30.43 \) to 30.47 seconds. Compared with CARA-RI, RRAA tries to steadily use 12 Mb/s PHY rate in the deep fading channel, and then the frame is dropped due to the expiration of the retry limit. Even after this frame drop, the same PHY rate, i.e., 12 Mb/s, is continually used for the next frame. The inefficiency of CARA-RI and RRAA can be alleviated by adopting a rate-decreasing rule for retransmissions as introduced in the following.

B. Rate Adaptation: Fast Decrease (FD)

Regarding the retransmission procedure of the IEEE 802.11 MAC, (re)transmission failures greater than the retry limit result in a frame drop. Therefore, it is highly recommended for the VoWLAN to retransmit using robust PHY rates, i.e., low PHY rates, in order to prevent VoIP frame drops and guarantee high voice quality.

We propose fast decrease (FD) that defines the PHY rates for the retransmissions of VoIP frames. FD basically decreases the PHY rate for every other retransmission failures, as the frame loss could be due to both collision and channel errors. For instance, in an 802.11a PHY rate set, if a frame is initially transmitted at 48 Mb/s PHY rate, then the first retransmission uses 48 Mb/s considering the possibility of collision errors, and the second and third retransmissions use 36 Mb/s. In order to provide the best frame delivery probability, FD uses the lowest PHY rate beginning the fourth retransmissions, e.g., 6 Mb/s in the previous example. Table I summarizes the PHY rate use of FD with an example of 802.11a PHY, assuming retry limit of 7.

For ARF-based rate adaptation algorithms, FD needs to quickly increase the PHY rate after the lowest PHY rate transmission. Suppose that a frame is successfully delivered at the fifth retransmission using 6 Mb/s, assuming that the actual wireless channel can accommodate 24 Mb/s. If one selects 6 Mb/s for the next frame, then it loses a chance to use higher PHY rates. Therefore, we define reference rate that is referred at the frame transmission after the successful retransmissions with the lowest PHY rate. According to the reference rate, one can try 24 Mb/s at the next frame transmission in the previous example. The reference rates are denoted within brackets in Table I.

In the case of FD applied to RRAA, referred to as RRAA-FD, the transmission results of retries are reflected on the FER calculation as follows. The transmission failures of retries with low PHY rates are recorded as the failures of high PHY rates, while it is not the case for the transmission successes For example, suppose that RRAA-FD initially fails to transmit at 48 Mb/s, and the fifth retransmission at 6 Mb/s by applying FD succeeds. RRAA-FD then assumes that if four retransmissions
were using the initial 48 Mb/s, then all of those would fail, and hence, increases the failure counter of 48 Mb/s by 5, while its success counter remains unchanged.

Based on FD, we design Agile ARF (AgARF), which is an advanced version of AARF. AgARF quickly decreases the PHY rate during the FD rule upon a frame transmission failure. In order to compensate the fast PHY rate decrease and opportunistically increase it back, we reduce the minimum success threshold to 5 from 10, and resets the success counter in the same manner as CARA-RI. AgARF also adopts the A-RTS of RRAA and enables RTS transmissions, which could be effective in high-contention environments.

The transmissions with the lowest PHY rate in FD can reduce the transmission errors against bad channel status. Nonetheless, if the channel status becomes too bad due to the deep fading and/or severe hidden terminal interference, even the retransmissions with the lowest PHY rate may continue to fail. Hence, we develop a scheduling policy for the retransmissions in the following so that a WLAN station can overcome this problem.

C. Retransmissions: Retry Scheduling (RS)

If the retransmissions with the lowest PHY rate keep failing, then the VoWLAN station would better wait until the channel status becomes favorable, e.g., the wireless channel escapes the deep fading or the transmission of hidden terminal interference terminates. In this case, the question is until when the VoWLAN station waits for the retransmissions. Relying on the periodicity of the VoIP traffic, one possibility is postponing the retransmission until the next frame arrives at the MAC layer, so that the time diversity effect can be maximized while the next frame transmission is not much delayed by the delayed retransmissions of the preceding frame. Few tens of milliseconds coherence time of the WLAN also justifies delaying the retransmission with G.711 VoIP packet generation interval (20 msec), as the channel state could be changed in the delayed retransmission.

We refer to this delayed retransmissions as retry scheduling (RS). More in detail, the VoWLAN station first attempts retransmissions up to three times with the PHY rates based on FD, while FD with RS is slightly modified so that the VoWLAN station uses the lowest PHY rate at the third retransmission, and the PHY rate which is supposed to be used based on the original FD rule (i.e., the same PHY rate as the second retransmission) remains as the reference rate. If the third retransmission fails, then the VoWLAN station postpones the fourth retransmission until the next VoIP frame arrives at the MAC layer. For the fourth retransmission after RS, i.e., the delayed retransmission, the VoWLAN station tries the PHY rate that is one step lower than the reference rate because the third retransmission failed. If the third retransmission with the lowest PHY rate succeeds, however, then the VoWLAN station transmits the next VoIP frame with the reference rate. Between the third and fourth retransmissions, the VoWLAN station may transmit other pending frames in order to prevent a queue blocking event. Table II summarizes the example of the PHY rate use of FD combined with RS including reference rates in brackets.

Fig. 3 illustrates the operation of both FD and RS applied to CARA-RI from ns-3 simulation. The third retransmission of frame $n$ uses the lowest PHY rate and succeeds. In the case of frame $(n + 1)$, the initial transmission with the reference rate (12 Mb/s) fails, and the third retransmission with 6 Mb/s (where the reference rate becomes 9 Mb/s) fails as well. Hence, the fourth retransmission is delayed until frame $(n + 2)$ arrives. As the channel status gets better, the fourth retransmission of frame $(n + 1)$ with 6 Mb/s, which is one-step lower than the reference rate 9 Mb/s, and the following initial transmission of frame $(n+2)$ are successfully delivered.

One might argue that the proposed RS delays the delivery of a VoIP frame, thus degrading the QoS eventually. However, the extra delays of the VoIP frames can be absorbed by the de-jitter buffer. Note that with the proposed RS, a frame might experience an extra delay of about 20 ms, and the assumed de-jitter buffer (whose size is 3 packets) can mostly manage this extra delay. Moreover, just before the delayed retransmission, RS combined with FD tries the lowest PHY rate, which is supposed to be the most robust. The erroneous transmission at the lowest PHY rate implies that the other PHY rates are likely to fail too, and hence, a WLAN station has no reason to waste the wireless channel with erroneous transmissions; the wireless channel would rather be used for frames destined to or transmitted by other WLAN stations. Therefore, we claim that RS can enhance the QoS of VoIP traffic as well as the efficiency of the wireless channel usage.

Another technical issue of RS is its standard compliance. Strictly speaking, RS is not a standard-compliant operation.

### Table I

<table>
<thead>
<tr>
<th>Transmission order</th>
<th>PHY rate use of Fast Decrease</th>
<th>802.11a example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial try</td>
<td>Initial rate</td>
<td>48 Mb/s</td>
</tr>
<tr>
<td>1st retry</td>
<td>Unchanged</td>
<td>48 Mb/s</td>
</tr>
<tr>
<td>2nd retry</td>
<td>One step down</td>
<td>36 Mb/s</td>
</tr>
<tr>
<td>3rd retry</td>
<td>Unchanged</td>
<td>36 Mb/s</td>
</tr>
<tr>
<td>4th retry</td>
<td>Lowest [One step down]</td>
<td>6 Mb/s [24 Mb/s]</td>
</tr>
<tr>
<td>5th retry</td>
<td>Lowest [Unchanged]</td>
<td>6 Mb/s [24 Mb/s]</td>
</tr>
<tr>
<td>6th retry</td>
<td>Lowest [One step down]</td>
<td>6 Mb/s [18 Mb/s]</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>Transmission order</th>
<th>PHY rate use of Fast Decrease with Retry Scheduling</th>
<th>802.11a example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial try</td>
<td>Initial rate</td>
<td>48 Mb/s</td>
</tr>
<tr>
<td>1st retry</td>
<td>Unchanged</td>
<td>48 Mb/s</td>
</tr>
<tr>
<td>2nd retry</td>
<td>One step down</td>
<td>36 Mb/s</td>
</tr>
<tr>
<td>3rd retry</td>
<td>Lowest [Unchanged (*)]</td>
<td>6 Mb/s [36 Mb/s]</td>
</tr>
<tr>
<td>4th retry</td>
<td>One step down from (*)</td>
<td>24 Mb/s</td>
</tr>
<tr>
<td>5th retry</td>
<td>Lowest [Unchanged]</td>
<td>6 Mb/s [24 Mb/s]</td>
</tr>
<tr>
<td>6th retry</td>
<td>Lowest [One step down]</td>
<td>6 Mb/s [18 Mb/s]</td>
</tr>
</tbody>
</table>

Fig. 3. An example of FD and RS applied to CARA-RI.
since RS intentionally delays the frame retransmissions at the MAC layer. Note that according to the standard, frame retransmission instants are determined by the random backoff process and exponentially increasing CW values for retransmissions. However, it should be noted that RS does not incur any interoperability problem of existing VoWLAN devices as the medium access function via backoff processes is the same as the original EDCA MAC. As RS can substantially reduce the VoIP frame losses, we insist that RS is an effective operation for improving the QoS of VoWLAN.

As an extension of RS, the retransmissions can be delayed further than the next frame arrival, where the specific design is dependent on the other delay components of the network, e.g., the wireline delay and de-jitter buffer size. The delayed retransmission frames could also be transmitted back-to-back with the new frames by using frame aggregations defined in [30], in order to enhance the resource utilization. We leave the detailed extensions of the RS design as a future work, and we focus on evaluating the proposed RS design in this paper.

IV. PERFORMANCE EVALUATION

Based on the simulation results of ns-3 simulator [16], we evaluate the QoS of VoWLAN with various rate adaptation schemes. The VoIP traffic is modeled as CBR traffic with 208 bytes frame payload size and 20 ms packet generation interval assuming G.711 voice codec without silence suppression. We suppose the VoIP communication scenario demonstrated in Fig. 1, where each VoIP terminal accesses the network through a WLAN AP connected to the Internet via a wireline backhaul, and the other VoIP terminal resides beyond the Internet. For the WLAN access, we assume IEEE 802.11a PHY and 802.11 EDCA MAC, where VoIP and TCP data traffic are mapped into the access category of AC_VO and AC_BE, respectively, with default channel access parameters given in [7]. The delays of the wireline Internet and de-jitter buffer are assumed to be 35 and 60 ms, respectively [23, 24]. For each set of simulation scenarios, we average the results of 50 simulation runs where each run for 100 second simulation time, and the R-score is basically used for the performance metric of VoIP. We also adopt the path-loss model with path-loss exponent of 4 for a typical indoor environment, and the Rayleigh fading channel model with the Doppler velocity of 0.1 m/s for static stations and 1 m/s for moving stations [29].

A. Comparing Schemes

We compare the performance of the following well-known rate adaptation schemes, while some of those are modified for VoWLAN support as explained below.

1) ARF-based families: We basically consider ARF and AgARF, and the revision of AgARF by using RS (AgARF w/ RS). CARA-RI is also revised by using FD (CARA-RI-FD), and both FD and RS (CARA-RI-FD w/ RS).

2) FER-statistics-based families: Various versions of RRAA are considered. Basically the original RRAA adopts the FER table which is calculated assuming 1300 bytes frame payload size and 40 $\text{ewnd}$ size, which are not suitable for the VoIP traffic. Hence, we modify the FER table using 208 byte frame payload size and use the $\text{ewnd}$ size of 10, and then apply FD (RRAA-e10-FD). Finally RRAA-e10-FD is further revised with RS (RRAA-e10-FD w/ RS). Minstrel, another representative FER-statistics-based scheme, is also compared.

3) Reference schemes: We implement a reference rate adaptation scheme, referred to as Genie, that ideally determines the best PHY rate without any overhead. Specifically, Genie selects the highest PHY rate that yields FER less than 1% for a given wireless channel based on the SNR. We also consider the 6 Mb/s fixed rate case as a baseline reference.

B. Static Topology

First, we study the QoS provisioning capabilities of the existing rate adaptation schemes in a static environment. We consider a WLAN composed of an AP and $N$ static VoIP stations, where the stations are equally and uniformly apart from the AP by 20 m, as shown in Fig. 4(a). At this distance, the time-varying wireless channel supports 18 and 24 Mb/s PHY rates in average. We present the lowest R-score out of 2N VoIP sessions (including uplink and downlink sessions) in the following figures, so that we can check whether all the VoIP sessions can be provided the satisfactory R-score.

Fig. 5 demonstrates the QoS of ARF, CARA-RI, RRAA, Minstrel, 6 Mb/s fixed rate, and Genie rate adaptation schemes. The upper histogram shows the R-score values, and the lower one shows the average channel occupancy time used to transmit a single VoIP frame including RTS/CTS and ACK overheads with various rate adaptation schemes. We verify that with small numbers of VoIP stations, RRAA performs even worse than ARF due to the frame drops caused by its
slow adaptation and inappropriate FER table reference; RRAA steadily uses high PHY rates so that the retransmissions could fail frequently. On the other hand, ARF and CARA-RI decrease the PHY rate by one step after two consecutive failures, and hence, the retransmissions have higher probabilities to be successfully delivered.

One might argue that using the lowest PHY rate in a fixed manner, i.e., the 6 Mb/s case, provides the best QoS thanks to its robustness. This is true when there exists only a single station; the R-score of 6 Mb/s is the best. However, as shown in Fig. 5, it is not the case for the high contention environment with 20 VoIP stations as the lowest PHY rate consumes too much of the wireless channel, so none of the stations can be supported in this environment. It is also known that ARF mostly uses the lowest PHY rate in high contention environments [11], and we confirm that by observing the R-score and channel occupancy time of ARF and 6 Mb/s. Accordingly, the R-score of ARF with 20 VoIP stations is as low as that of 6 Mb/s case, and we reconfirm that it is required to mitigate the collision problem by using RTS adaptively as CARA-RI and RRAA do.

We also verify the employed PHY rate distributions of each rate adaptation scheme in Fig. 6 assuming one VoIP station in the network. Each stacked bar represents the number of frames that were transmitted with a specific PHY rate, where the left-hand-side and right-hand-side present erroneous and successful PHY rate uses, respectively. We observe that Minstrel wastes lots of resources due to erroneous transmissions with unnecessarily high PHY rates, which is caused by the random probing. Thanks to the retry chain of Minstrel, however, more robust PHY rates can be used for these transmissions, so the R-score performance is fair in low contention environments. RRAA tends to steadily use a specific PHY rate even in bad channel conditions, so that the number of erroneous frame transmissions is greater than that of ARF, CARA-RI, or Minstrel.

Now, we confirm the performance gain of the FD and RS designs in Fig. 7. Due to the lack of space, we only present ARF, CARA-RI, RRAA, and Genie rate adaptation schemes with one and 15 VoIP stations; for the 20 VoIP stations case, none of these rate adaptation algorithm can support R-score greater than 80, which is out of our interest. All the above schemes are revised by applying (1) FD or (2) both FD and RS. The R-score is enhanced by applying FD, while RS further enhances the R-score above 80. Especially for RRAA, the performance gain of FD is significant since the original rate selection of RRAA is too static. RS improves the R-score for all of the rate adaptation schemes including Genie by removing the continuous retransmission failures with the 6 Mb/s PHY rate. We also observe that the RS gain in 15 contending station case is slightly decreased compared with the one station case, as the high contention environment already exploits the time diversity of retransmissions due to the severe contention. From Fig. 7, we conclude that the FD and RS features improve the existing rate adaptation schemes, where the gain varies from 11% (in the case of ARF and CARA-RI) to 80% (in the case of RRAA). In the following discussions, we present ARF and RRAA, as well as their variants, as representative schemes of ARF-based families and FER-statistics-based families, respectively.

C. Random Topology with Heterogeneous Traffic Types

In this scenario, five static VoIP stations and five static TCP data stations are randomly deployed within a circle of 20 m radius, where an AP resides at the center of the circle as depicted in Fig. 4(b). Two TCP data stations have uplink flows and three TCP data stations have downlink flows as downlink TCP traffic is more general in practical WLAN environments, where each TCP flow has 5 Mb/s traffic load from the application layer with 1508 byte frame payload size. Both VoIP and TCP data stations use the same rate adaptation scheme, except the RS feature which is used only by VoIP stations.

Fig. 8 demonstrates the empirical cumulative distribution functions (CDFs) of the R-scores of ARF, RRAA, and Genie rate adaptation schemes including their variants. An interesting observation is that the average R-score of this scenario with both VoIP and TCP data stations is slightly higher than that of the previous scenario with only VoIP stations. This is due to the random topology which has high chances that VoIP stations get closer to the AP, as well as the advantage of the time diversity effect in retransmissions of the VoIP frames. That is, the retransmissions with the same PHY rate can succeed after the channel contention with the TCP data stations, and the diversity effect of the heterogeneous network is more advantageous than in the VoIP-only case since TCP data transmissions take longer time than short VoIP frames. From Fig. 8, it is shown that RRAA cannot support 80% of VoIP stations with the R-score higher than 80. ARF and Genie perform better than RRAA, but their outage probability,
All the WLAN stations use the same rate adaptation scheme, i.e., the probability that the R-score is below 80, is 1%. By applying FD and RS features, however, the revised RRAA can reduce the outage probability to less than 0.5%, and Agile ARF w/ RS can perfectly support all the VoIP stations with R-score greater than 80 since the frame losses are significantly reduced.

D. Mobile Topology

A rate adaptation is required to be highly responsive in mobile environments in order to trace the channel dynamics. We now evaluate the QoS of the rate adaptation schemes considering their responsiveness in the following mobile scenario.

We consider a simple mobile scenario with a single VoIP station which moves back-and-forth along a straight line at 1 m/s velocity as shown in Fig. IV-D. Once the VoIP station reaches the end of the line, where the VoIP station can marginally use the 6 Mb/s PHY rate, it turns around and goes back. We evaluate the R-scores every 5 seconds and record them as a function of time.

Fig. 10 shows the time history of the R-score performance in this mobile scenario. Note that ARF that cannot quickly adapt to the channel dynamics performs badly when the VoIP station moves far from the AP, and RRAA does even worse. This is the same for Genie as well, since the 6 Mb/s transmissions could fail due to the deep fading. On the other hand, the rate adaptation schemes revised by using the FD and RS features dramatically improve the R-score performance, since using the robust PHY rate as well as delaying the retransmissions can increase the frame delivery ratio in long-distance communications.

E. Hidden Terminal Environment

Finally, we investigate the performance of the rate adaptation schemes in the WLAN environments with hidden terminals. Fig. 11(a) illustrates the experimental deployment of a VoIP station and a TCP data station in opposite directions. All the WLAN stations use the same rate adaptation scheme, except the RS feature which is only applied for the VoIP traffic. Since the WLAN stations are far enough, these two cannot sense the transmission of each other. The TCP data station generates an uplink TCP flow with 2 Mb/s traffic load. Fig. 11(b) illustrates the case that hidden terminals are associated with different APs. Two APs can sense the transmission of each other, but the WLAN stations cannot sense their neighboring AP that they are not associated with. The TCP data station receives a downlink TCP flow with 2 Mb/s traffic load.

The R-score performances of both uplink and downlink VoIP sessions in hidden terminal environments are presented in Fig. 12 when both VoIP and TCP data stations use the same rate adaptation scheme. Due to the settings of hidden terminals, the uplink VoIP session suffers from the hidden terminal interference. Note that in the single-cell hidden scenario, the uplink sessions of AgARF and RRAA-e10-FD performs better than those in the multi-cell hidden scenario. In the single-cell hidden scenario, the hidden terminal problem can be resolved by using the RTS/CTS exchanges, while it is not the case in the multi-cell hidden scenario. Therefore, AgARF and RRAA-e10-FD that adaptively use the RTS/CTS exchanges can perform better in the single-cell hidden scenario.

From the R-score performance of the uplink VoIP sessions, we observe that ARF performs better than RRAA, which is supposed to use higher PHY rates than ARF. This is due to the robustness of low PHY rates to the hidden terminal interference. Thanks to the capture effect, the transmission with a low PHY rate can succeed even though it encounters hidden terminal interference. For the same reason, the rate adaptation schemes with the FD feature can perform better than their original versions. Relying on the delayed transmissions of RS, the revised rate adaptation schemes have higher chances to escape both deep fading and hidden terminal interference. Therefore, the R-score performance is clearly improved by applying RS.

F. Preliminary MadWiFi Implementations

We have implemented the proposed Agile ARF with RS in MadWiFi [17] using Lenovo X41 laptop equipped with Cisco Air-CB21AG WLAN card. Due to the limitations of
MadWiFi, RTS/CTS exchanges cannot be enabled/disabled for every frame (re)transmission, and hence, we exclude the A-RTS feature in the current implementation. Here, we present a representative result due to the space limit. We experiment an one-to-one VoIP communication scenario with IEEE 802.11a PHY in a static office environment, where the wireless channel supports 18 Mb/s PHY rate in average. IxChariot is used to generate bi-directional G.711 VoIP traffic and evaluate the representative result. We experiment ages the timing for the retransmissions in order to avoid both R-score [31]. We observe that Agile ARF with RS improves RTS feature in the current implementation. Here, we present a approximately 12%, and this demonstrates that the proposed algorithm works in practice.

V. CONCLUSION AND FUTURE WORK

In this paper, we investigate the characteristics of VoWLAN systems, and discuss the pros and cons of the existing rate adaptation schemes regarding the VoIP traffic. Based on the loss-sensitivity of the VoIP traffic, we design two novel features for the rate adaptation of VoWLANs. Fast decrease (FD) guides the PHY rates of retransmission frames to low(est) PHY rates so that those frames can be successfully delivered before the retry counter expires. Retry scheduling (RS) manages the timing for the retransmissions in order to avoid both deep channel fading and hidden terminal interference. Through extensive ns-3 simulations and prototype measurements, we demonstrate that applying these two features to the existing rate adaptation schemes can dramatically improve the QoS of VoWLAN.

We are currently expanding our experimental study for comparison with other rate adaptation algorithms, and we are also considering the 802.11g PHY extension. Through the future work on refining the RS design, i.e., jointly optimizing the delayed transmission and de-jitter buffer size, we expect to further enhance the QoS of VoWLAN.

ACKNOWLEDGMENT

This research was supported by LG U+ project, “Research on WLAN management and operation technologies for real-time service in condensed WLAN environments,” and the National Research Foundation of Korea (NRF) grand funded by the Korea government (MEST) (No. 2001-0017027).

REFERENCES


