

Application-Level QoS and QoE Assessment of a Cross-Layer Packet Scheduling Scheme for Audio-Video Transmission over Error-Prone IEEE 802.11e HCCA Wireless LANs

Zul Azri BIN MUHAMAD NOH^{†a)}, Student Member, Takahiro SUZUKI^{††b)}, Member, and Shuji TASAKA^{†c)}, Fellow

SUMMARY This paper proposes a cross-layer packet scheduling scheme for QoS support in audio-video transmission with IEEE 802.11e HCCA and assesses application-level QoS and QoE of the scheduling scheme under lossy channel conditions. In the proposed scheme, the *access point (AP)* basically allocates *transmission opportunity (TXOP)* for each station in a *service interval (SI)* like the reference scheduler of the IEEE 802.11e standard, which is referred to as the TGe scheme in this paper. In the proposed scheme, however, the AP calculates the number of *MAC service data units (MSDUs)* arrived in an SI, considering the inter-arrival time of audio samples and that of video frames, which are referred to as *media units (MUs)*, at the application layer. The AP then gives additional TXOP duration in the SI to stations which had audio or video *MAC protocol data units (MPDUs)* in their source buffers at the end of the previous TXOP. In addition, utilizing video frame information from the application layer, we propose video frame skipping at the MAC-level of a source station. If a station fails to transmit a video MPDU, it drops all the following video MPDUs in the source buffer until the next intra-coded frame comes to the head of the buffer. We compare the reference scheduler (TGe scheme), the proposed packet scheduling scheme with and without the video frame skipping at the source in terms of application-level QoS and QoE. We discuss the effectiveness of the proposed packet scheduling scheme from a viewpoint of QoE as well as QoS. Numerical results reveal that the proposed packet scheduling scheme can achieve higher quality than the TGe scheme under lossy channel conditions. We also show that the proposed scheduling scheme can improve the QoS and QoE by using the video frame skipping at the source. Furthermore, we also examine the effect of SI on the QoS and QoE of the proposed packet scheduling scheme and obtain that the appropriate value of SI is equal to the inter-arrival time of video frame.

key words: IEEE 802.11e HCCA, audio-video transmission, packet scheduling, cross-layer, QoE, application-level QoS

1. Introduction

The IEEE 802.11 wireless local area networks (WLANs) play a prominent role in offering ubiquitous connectivity to the Internet. As multimedia applications such as voice over IP (VoIP), video conferencing, and live streaming are becoming prevalent, a demand for support of *Quality of Service (QoS)* in WLANs is increasing. In order to provide

QoS support for multimedia transmission, the IEEE 802.11e *medium access control (MAC)* has been introduced as an extension to the original IEEE 802.11 MAC.

The IEEE 802.11e MAC defines the *hybrid coordination function (HCF)*, which has two access methods: *enhanced distributed channel access (EDCA)* and *HCF controlled channel access (HCCA)* [1]. The former is a contention-based protocol based on *carrier sense multiple access with collision avoidance (CSMA/CA)* and can support priority-based service differentiation. The latter is a polling-based protocol and can support guaranteed media access for real-time transmission. The EDCA, which is based on distributed control, is easy to be implemented, but under heavy load conditions, QoS cannot always be met. For this reason, the centrally controlled HCCA is preferred in providing high assurance of QoS guarantee. Therefore, in this paper, we focus on the audio-video transmission with the HCCA.

For multimedia services over WLANs, we should consider QoS at each level of the protocol stack. Reference [2] identifies six levels of QoS in IP networks: *physical-level*, *node-level*, *network-level*, *end-to-end level*, *application-level*, and *user-level*. In multimedia applications, user-level QoS is the most important since the final goal of multimedia services is to provide high user-level (perceptual) QoS for the end-users; this is also referred to as *Quality of Experience (QoE)* in ITU-T [3]. The ITU-T defines the QoE as the overall acceptability of an application or service, as perceived subjectively by the end-users.

In the HCCA, QoS support is achieved through packet scheduling and admission control. For reference, the IEEE 802.11e standard has presented an example packet scheduler, which is referred to as the *Task Group e (TGe)* scheme in this paper. In the TGe scheme, the *transmission opportunity (TXOP)* duration for a station is calculated by the *access point (AP)* on the basis of the *traffic specification (TSPEC)* information received from the station. The TSPEC consists of a set of parameters that include the mean data rate, delay bound, nominal *MAC service data unit (MSDU)* size, and *maximum service interval (MSI)*.

For simplifying the description of the TGe scheme, let us focus on flow i where ρ_i is the mean data rate in bits per second, L_i is the nominal MSDU size in bits, and R_i is the physical transmission rate of flow i in bits per second.

The AP first determines the *service interval (SI)*. The SI

Manuscript received October 23, 2009.

Manuscript revised January 31, 2010.

[†]The authors are with the Department of Computer Science and Engineering, Nagoya Institute of Technology, Nagoya-shi, 466-8555 Japan.

^{††}The author is with the Faculty of Health Sciences, Nihon Fukushi University, Handa-shi, 475-0012 Japan.

a) E-mail: zulazri@nitech.ac.jp

b) E-mail: suzuki@n-fukushi.ac.jp

c) E-mail: tasaka@nitech.ac.jp

DOI: 10.1587/transcom.E93.B.1384

is a value lower than the smallest value among the MSI of admitted stations and is a submultiple of the beacon interval. Then, the AP calculates the number of MSDU of flow i that arrives at the mean data rate during an SI as

$$N_i^T = \left\lceil \frac{SI \times \rho_i}{L_i} \right\rceil \quad (1)$$

where SI is the duration of the SI. The superscript T means the TGe scheme. Then, the TXOP duration for flow i is computed as

$$TXOP_i^T = \max \left(\frac{N_i^T \times L_i}{R_i} + O, \frac{M}{R_i} + O \right) \quad (2)$$

where M is the maximum allowable size of MSDU, and O is the overhead in time units due to the physical header, MAC header, *inter-frame space* (IFS), acknowledgment frames, and poll frames. The TGe scheme is suitable for transmission of *constant bit rate* (CBR) traffic with an error-free channel.

It should be noted here that the TGe scheme has three disadvantages, which are mentioned below.

First, the TGe scheme cannot guarantee QoS for *variable bit rate* (VBR) traffic because it does not take the data rate and packet size fluctuation into account.

Secondly, in the TGe scheme, the TXOP duration for each station is calculated on the assumption that the channel is error-free. However, transmission errors can occur in WLANs owing to shadowing, multipath fading, and interference. In this case, the TXOP duration becomes insufficient because channel capacity for retransmission traffic is not allocated.

Thirdly, the derived number of arriving MSDUs calculated from Eq. (1) becomes smaller than the real number; this leads to insufficient channel allocation in the TGe scheme [4]. For example, in the case of video flow used in our study, one video frame is generated in every 50 ms, the mean data rate is 800 kbps and nominal MSDU size is 1500 bytes; therefore, the video frame is fragmented into 4 MSDUs on average. Thus, 8 MSDUs are generated in an SI when SI is 100 ms. On the other hand, according to Eq. (1), the number of arriving MSDUs becomes 7 MSDUs if SI is 100 ms.

There are numerous studies trying to improve the inefficiency of the TGe scheme. In [5]–[10], various methods are proposed to improve QoS for transmission of VBR traffic. However, these studies have been focusing only on MAC-level QoS and have not examined QoS at upper levels; that is, the QoS is assessed in terms of the MAC-level throughput and MAC frame delay.

With regard to QoS at upper levels, reference [11] examines the application-level QoS of video transmission with a cross-layer packet scheduling scheme in terms of *peak signal-to-noise ratio* (PSNR). Reference [12] also examines video loss quality at the application-level of another cross-layer packet scheduling scheme for video transmission with the *controlled access phase scheduling* (CAPS) and EDCA.

In the scheduling schemes of these references, a video flow is divided into subflows at the application layer and more important video information are treated as higher priority traffic at the MAC layer. In these references, however, only video traffic is considered. In addition, the authors propose the *multimedia priority dynamic scheduling* (MPDS) scheme for VBR traffic and have assessed application-level QoS and QoE in the case where audio and video are transferred from stations to the AP [13]. In reference [13], however, QoS assessment has been performed assuming an error-free channel, and the MPDS scheme has not cope with the third problem of the TGe scheme.

This paper proposes a cross-layer packet scheduling scheme for audio-video transmission with the IEEE 802.11e HCCA to solve the three problems of the TGe scheme described earlier under lossy channel conditions[†]. In addition, we also introduce video frame skipping at the MAC-level of a source station. We then compare the TGe scheme and the proposed packet scheduling scheme with and without the video frame skipping in terms of application-level QoS and QoE. The novelty of this paper is to show the effectiveness of our proposed scheduling scheme with the video frame skipping from a viewpoint of QoE as well as application-level QoS.

In the proposed scheduling scheme, the AP first calculates basic TXOP for each station in an SI using the inter-arrival time of audio samples and that of video frames at the application layer to overcome the third problem. The AP then allocates additional TXOP in the SI to each station on the basis of the queue length of its source buffer. This TXOP allocation is useful to cope with the first and second problems. We compare the TGe scheme and the proposed packet scheduling scheme with and without the video frame skipping at the source in terms of application-level QoS and QoE for various values of BER and the number of stations. Furthermore, we also examine the effect of SI on the QoS and QoE in the proposed scheduling scheme since they can be highly affected by the polling interval. Application-level QoS assessment is performed through simulation, and QoE assessment is carried out by subjective experiment. Since QoE is directly related to human perception, we utilize a psychometric method referred to as the *method of successive categories* [14].

The rest of the paper is organized as follows. Section 2 explains the mechanism of HCCA and admission control in the TGe scheme. Section 3 describes the proposed packet scheduling scheme. Section 4 specifies simulation conditions and methodology of subjective assessment. Sections 5 and 6 give numerical results of application-level QoS and QoE. Finally, Sect. 7 concludes this paper.

[†]The authors have proposed a packet scheduling scheme for transmission of CBR audio-video traffic and have examined application-level QoS of the scheme in a noisy environment [4]. In the scheme, the number of MSDUs generated by a station in an SI is calculated in the same way as that in this paper. However, in reference [4] channel allocation for VBR traffic is not considered, and QoE assessment has not been carried out.

2. IEEE 802.11e HCCA Mechanism

In this section, we introduce the mechanism of the HCCA and the method of admission control in the TGe scheme.

The HCCA provides polled access to the wireless medium. It is controlled by the *hybrid coordinator (HC)*, which is usually collocated with the AP. In an IEEE 802.11e WLAN with the HCCA, the *contention free period (CFP)* and *contention period (CP)* alternate periodically over time, and a combination of CFP and CP forms a superframe, which starts with a beacon frame. The CFP and CP are used for the HCCA and EDCA, respectively. In the HCCA, however, the AP can poll a station even during the CP as well as the CFP. CFP is called *controlled access phase (CAP)* in IEEE 802.11e.

A station which requires TXOP for a flow must query for a QoS reservation by sending its TSPEC to the AP. Owing to the limitation of the capacity during an SI, the TGe scheduler implements admission control to ensure that all admitted flows have adequate TXOP duration for their QoS. Admission control decides which flow should be admitted and which flow should be dropped from the polling list. When flow $k+1$ issues a QoS reservation, the AP will first check whether the available capacity of the medium exists or not by the following equation:

$$\frac{TXOP_{k+1}^T}{SI} + \sum_{i=1}^k \frac{TXOP_i^T}{SI} \leq \frac{T - T_{CP}}{T} \quad (3)$$

where T is the beacon interval, and T_{CP} is the time for the EDCA. If Eq. (3) is satisfied, the AP admits flow $k+1$ into its polling list and allocates TXOP to the flow on the basis of the TGe scheduling scheme.

3. Proposed Scheduling Scheme

In this section we will explain the proposed scheduling scheme. This scheme is a packet scheduling scheme with cross-layer mechanisms between MAC layer and application layer because it sends MPDUs at the MAC layer using traffic information at the application layer. In this section, we first describe how the AP allocates TXOP to each station in the proposed scheme. We then explain the video frame skipping performed by the scheduling scheme.

3.1 Channel Allocation of the Scheduling Scheme

In the proposed scheduling scheme, the AP first calculates the basic TXOP duration for a flow in a similar way to that of the TGe scheme. However, the TGe scheme has the third problem described in Sect. 1; that is, the derived number of MSDUs calculated from Eq. (1) becomes smaller than the real number. Therefore, we propose a new method for calculating the number of MSDUs arriving in an SI in order to overcome this problem.

In the proposed method, information about *inter-MU*

(*media unit*)[†] time for a flow is required to calculate the number of MSDUs arriving from the flow in an SI. Here, the *inter-MU* time is defined as the interval between the generation of two consecutive video MU or audio MU.

Using the *inter-MU* time together with other TSPEC parameters, the AP first computes the number of MSDUs arriving from flow i within an *inter-MU* time as

$$n_i^P = \left\lceil \frac{\text{inter_MU}_i \times \rho_i}{L_i} \right\rceil \quad (4)$$

where inter_MU_i denotes the inter-MU time for flow i . The superscript P means the proposed scheme. Then, the number of MSDUs that arrives in an SI is computed as

$$N_i^P = \left\lceil \frac{SI}{\text{inter_MU}_i} \times n_i^P \right\rceil \quad (5)$$

Then, the basic TXOP duration for flow i , which is denoted by basicTXOP_i is calculated with Eq. (2) by replacing N_i^T in Eq. (1) with N_i^P in Eq. (5).

Note here that in audio-video transmission with the scheduling scheme, each station has to pass the inter-MU time for audio and that for video from the application layer to the MAC layer and has to exchange the two parameters with the AP as TSPEC parameters. Therefore, extensions of the IEEE 802.11e HCCA MAC are needed to realize this scheme. However, the TXOP calculation for an audio or video flow using the inter-MU time can give an accurate number of MSDUs arriving from the flow.

The AP then calculates additional TXOP for each flow in the SI on the basis of the queue length of the source buffer at the end of the previous TXOP. The QoS control field of the IEEE 802.11e MAC header is utilized to deliver the queue length of the audio buffer and that of the video buffer at the stations. This queue length means the number of *MAC protocol data units (MPDUs)* which could not be transmitted during the previous TXOP because of the insufficient TXOP duration.

After the AP obtains the queue length record during the previous SI and then computes the basic TXOP duration, the AP computes the additional TXOP duration for flow i as follows:

$$\text{addTXOP}_i = \frac{\text{queue}_i \times L_i}{R_i} \quad (6)$$

where queue_i is the number of MPDUs in the (audio or video) queue of flow i . The additional TXOP is required for transmission of MPDUs left in the queue.

As the surplus bandwidth reduces with the increase of admitted flows, the AP might not be able to allocate enough additional TXOP duration. In this case, the AP distributes the surplus bandwidth for each flow in a round-robin basis; that is, additional TXOP duration for transmission of one nominal-sized MSDU will be allocated to each flow until the surplus bandwidth is fully occupied. Note that the

[†]An MU is the unit of information that is delivered from a source station to a destination station at the application-level.

maximum additional TXOP duration for a station in an SI is bounded by $addTXOP_i$.

TXOP duration for flow i in the SI is calculated as the sum of the basic TXOP duration and the additional one for flow i . When a station generates both audio and video flows, TXOP duration for the station is calculated as the sum of the TXOP duration for audio and that for video.

The proposed scheduling scheme is appropriate for transmission of VBR traffic and retransmission owing to transmission errors because the AP allocates additional TXOP to each station after it gives basic TXOP.

3.2 Video Frame Skipping

In this paper, we also propose video frame skipping at the MAC-level of a source station. In the case of compressed video transmission, a raw video stream can be compressed into three kinds of frames: intra-coded frames (I-frames), predictive-coded frames (P-frames), and bidirectionally-predictive-coded frames (B-frames). The group of successive video frames starts with an I-frame, and all frames before the next I-frame are called *group of picture (GOP)*. Note here that if video degradation occurs in a video frame of a GOP owing to transmission error, it will propagate into all the following video frames within the GOP; for example, in the case of video streams with GOP of *IPPPPP* sequences, the loss of the I-frame leads to loss of the entire frames of the GOP because the first P-frame requires the preceding I-frame in order to be decoded, the second P-frame requires the first P-frame, and so on.

In the proposed video frame skipping, frame information of a video flow is delivered from the application layer to the MAC layer at a source station. Therefore, the MAC layer knows the frame types of each MPDU. We consider that this is achieved by utilizing the differentiated services code point (DSCP) field within the IP header. According to the DSCP value, the MAC layer can distinguish the frame type and then maps the value into the TID field within the QoS control field of the MAC header. The value of the DSCP field within the IP header is set depending on the video frame type which can be extracted from the network abstraction layer (NAL) unit header of the video frame. The similar approach has been used in [15] and [16], which have studied cross-layer optimization in IEEE 802.11e EDCA WLAN. When the station fails to send an MPDU of a GOP, the station drops all the following MPDUs of the GOP according to the frame information received from the application layer;

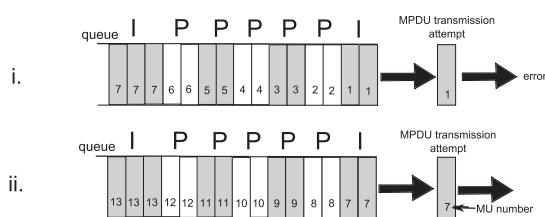


Fig. 1 The concept of video frame skipping.

it then tries to send the first I-frame of the next GOP.

Figure 1 shows an example of the video frame skipping. In this figure, an I-frame and a P-frame are fragmented into 3 and 2 MPDUs, respectively. It should be noted here that one video frame is fragmented into several MPDUs if its length is longer than nominal MSDU size. In Fig. 1(i), a station fails to transmit the first MPDU of I-frame 1 owing to channel transmission error. In this case, the station drops all MPDUs of P-frame 2 to P-frame 6 from its source buffer; it then tries to send the first MPDU of I-frame 7 as shown Fig. 1(ii).

The source video frame skipping can reduce the waste of bandwidth and can increase the channel capacity for transmission of audio-video MPDUs in a noisy environment, though this mechanism is also required an extension of the 802.11e HCCA MAC since the MAC layer needs video frame information at the application-level.

4. Experimental Methodology

In this paper, the application-level QoS is assessed by simulation. Then, the QoE is assessed by subjective experiment. In this section, we first present the simulation conditions used for the application-level QoS assessment. We then elaborate on the methods of subjective experiment.

4.1 Simulation Conditions

Figure 2 illustrates the system configuration used in the simulation. We focus on a single *basic service set (BSS)* which includes an AP and a certain number of multimedia stations. In the simulation, we assumed that channel overlapping problems do not occur since appropriate channel is selected by the AP. The number of multimedia stations are denoted by M . All multimedia stations are located at the same distance (say R) from the AP. We assume the IEEE 802.11b physical layer based on *direct sequence spread spectrum (DSSS)* with a channel data rate of 11 Mb/s [17]. Each multimedia station sends stored audio and video streams to the AP as two separate transport streams using UDP/IP.

Table 1 summarizes media specifications of audio-video flows used in the simulation. We use an audio flow of ITU-T G.711 μ -law and an H.264 video stream. A video MU is defined as a video frame and is transferred as one or more UDP datagrams. An audio MU consists of 1000 audio

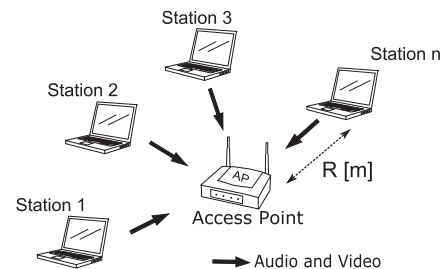


Fig. 2 System configuration.

Table 1 Specifications of audio and video.

	Audio	Video
coding scheme	G.711 μ -law	H.264
image size [pixel]	—	320 × 240
picture pattern	—	IPPPPP
average MU rate [MU/s]	8	20
average inter-MU time [ms]	125	50
average bit rate [kb/s]	64	800
measurement time [s]	60	60

Table 2 Relationship between distance and BER.

Distance, R [m]	BER
Error Free	0
145	1.8×10^{-5}
150	2.4×10^{-5}
155	3.1×10^{-5}
160	4.1×10^{-5}
165	5.6×10^{-5}
170	7.0×10^{-5}

samples, which corresponds to a single UDP datagram.

Two types of the contents are used in the simulation: *Music video* and *Sport*. The *Music video* shows scenes of a Japanese female singer dancing while singing with two men playing guitar. For the *Sport*, scenes of a football match with a commentator's voice have been chosen. Here, we have encoded each video stream with a picture pattern of *IPPPPP* at bit rate of 800 kb/s on average. The audio bit rate is constant at 64 kb/s.

In the simulation, we set the beacon interval to 1000 ms in order to evaluate the effect of SI between 25 ms through 500 ms. The nominal MSDU size for audio and that for video are set to 1000 bytes and 1500 bytes, respectively. Each multimedia station has source buffer for audio MPDUs and that for video ones, separately. Audio transmission is given higher priority than video transmission when a station has both audio MPDUs and video ones. Video frame skipping at the receiver is always performed at the application-level; that is, the AP drops all the following video MUs of the GOP when it fails to receive a video MPDU of a GOP. It should be noted that in the following simulation results, video frame skipping at the source means that the video frame skipping is performed at the MAC-level of the source in addition to that of the application-level of the receiver. Moreover, the CAP ratio is 0.8, which means at most 80% of the bandwidth within SI is allocated for the HCCA. The maximum retry number for MPDU retransmission attempt is set to four. An MPDU will be discarded if this number is exceeded. The SI is set to 100 ms unless otherwise stated.

In the simulation, the *bit error rate (BER)* is set to 1.8×10^{-5} , 2.4×10^{-5} , 3.1×10^{-5} , 4.1×10^{-5} , 5.6×10^{-5} , and 7.0×10^{-5} . As reference, we show the relationship between the distance R and the BER in Table 2; we calculated the distance utilizing the *signal-to-noise ratio (SNR)* based on *Orinoco 802.11b Card* [18] and an empirical curve of BER versus SNR provided by *Intersil WLAN chipset* [19].

In this paper, as application-level QoS parameters for audio-video traffic, we adopt the *average MU delay*, *MU loss ratio*, *coefficient of variation of output interval*, and *mean square error of inter-stream synchronization*. The average MU delay is the average time from the moment an MU is generated at the source station until the moment the MU is output at the receiver. The MU loss ratio is the ratio of the number of MUs not output at the receiver to the number of MUs generated by the source station. The coefficient of variation of output interval, which is defined as the ratio of the standard deviation of the MU output interval to its average, represents the smoothness of the output flow. The mean square error of inter-stream synchronization is an indicator of “lip-sync” and is the average square of the difference between the output time of each video MU and its derived output time obtained from the output time of the corresponding audio MU. The derived output time of a video MU means the output time of the corresponding audio MU plus the difference between the timestamps of the two MUs.

In the following numerical results, we calculated the 95-percent confidence intervals of the simulation results. However, if the interval is smaller than the size of the corresponding simulation symbol in the figure, we do not show it there.

4.2 Methods for Subjective Experiment

The QoE assessment is performed according to the methodology recommended in ITU-T Rec. P.911 [20]. We first made test samples for subjective assessment by actually outputting the audio and video MUs with the output timing obtained from the simulation. We made a 10 s video clip from simulation results for 60 s. In the assessment, we use a PC with a 17 inch-LCD display, and the distance between the assessors and the display is about 50–70 cm. The assessors listen to the audio output using a headphone. The subjective assessment was conducted by 17 students; their ages were 20 s. The assessors were asked to classify the test samples into a certain number of categories each assigned an integer. Here, we use five categories of *impairment* of the *rating-scale method*: “imperceptible” assigned integer 5, “perceptible, but not annoying” 4, “slightly annoying” 3, “annoying” 2, and “very annoying” 1.

5. The Effect of Transmission Error

In this section, we compare the TGe scheme, the proposed scheme with and without the source video frame skipping in terms of application-level QoS and QoE. We discuss the effects of transmission error and the number of stations on the QoS. We also examine how the source video frame skipping can improve the QoS of the proposed scheme. We first study the effect of transmission error on the application-level QoS from simulation results. We then show the QoE results from subjective experiment.

In the case of the TGe scheme, the calculation of arriv-

ing number of MSDUs in an SI can be inaccurate as mentioned in Sect. 1. Therefore, we also introduce a *modified TGe* scheme, where the arriving number of MSDUs in an SI is calculated by Eq. (5), in the same way as the proposed scheme. In the modified TGe scheme, however, additional TXOP duration for retransmission traffic is not allocated. In the following numerical results, TGe' means the modified TGe scheme.

5.1 Application-Level QoS

First, we will show the results of the application-level QoS for *Music video*.

Figures 3 and 4 show the average MU delay for audio and that for video, respectively, as a function of the BER.

In these figures, notation TGe'[3], for instance, refers to the result of the modified TGe scheme when the number of multimedia stations $M=3$. Meanwhile, notation Proposed[5] and Proposed-Skip[5] refer to the results of the proposed scheme with and without the source video frame skipping, respectively, when $M=5$.

These figures show three values of the number of multimedia stations for the proposed scheme: $M=3, 4$, and 5 . On the other hand, the number of multimedia station for the TGe' scheme is fixed at 3. In the TGe' scheme, allocated TXOP duration for a station is almost the same when M is less than or equal to 5. It should be noted that the maximum number of admitted stations is 5, which can obtain from Eq. (3). In the case of the proposed scheme, as the number of stations is larger, channel capacity for additional TXOP decreases even if M is less than or equal to 5.

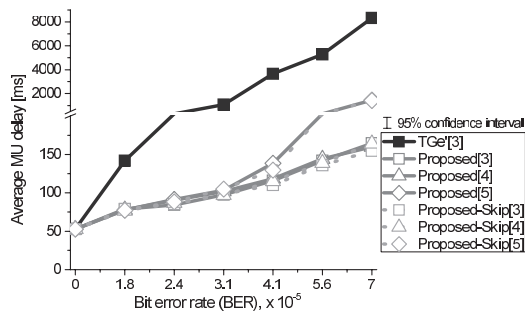


Fig. 3 Average MU delay for audio (Music video).

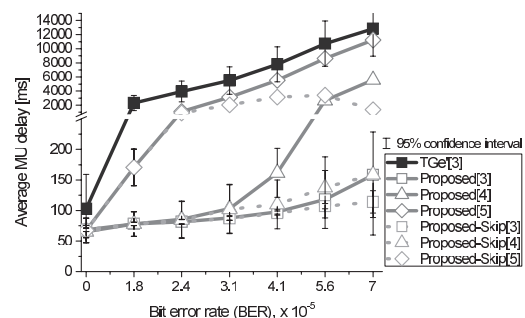


Fig. 4 Average MU delay for video (Music video).

We can find in Figs. 3 and 4 that in the case of TGe' scheme, the average MU delay for audio and that for video increase drastically as BER increases. Under noisy channel environments, stations need to retransmit MPDUs when transmission error occurs. However, in the TGe' scheme, TXOP duration for retransmission traffic is not allocated. Therefore, queue length of the source buffer becomes longer as BER increases.

In contrast, these figures show that the average MU delay for the proposed scheme becomes much smaller than that for the TGe' scheme when $M=3$. This is because the AP allocates additional TXOP duration according to the queue length of the source buffer. However, the average MU delay becomes larger as M increases because the increase of the number of multimedia stations reduces the surplus bandwidth for allocation of additional TXOP duration. In particular, the average MU delay for video becomes very large if BER is larger than 1.8×10^{-5} for Proposed[5] and Proposed-Skip[5], and 4.1×10^{-5} for Proposed[4]. Under these conditions, surplus bandwidth for additional TXOP duration becomes insufficient since many MPDUs are retransmitted.

We next examine the effect of the source video frame skipping on the average MU delay for video. Figure 4 reveals that the average MU delay of video for the Proposed-Skip scheme is smaller than that for the Proposed scheme if BER is larger than 3.1×10^{-5} for $M=4$, and 2.4×10^{-5} for $M=5$, respectively. Therefore, we can say that the source video frame skipping can improve the average MU delay under lossy channel conditions. This is because the source video frame skipping can reduce traffic volume to the wireless channel when the traffic load is heavy owing to retransmission of video MPDUs.

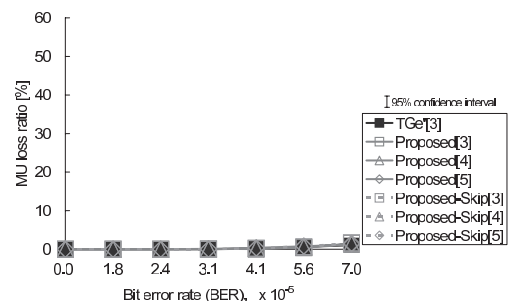


Fig. 5 MU loss ratio for audio (Music video).

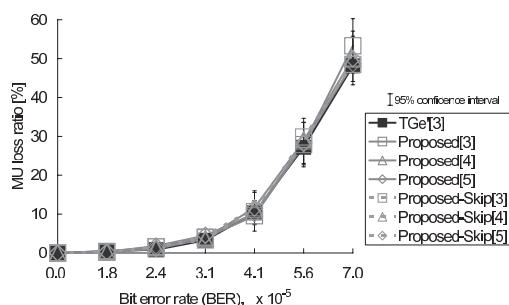


Fig. 6 MU loss ratio for video (Music video).

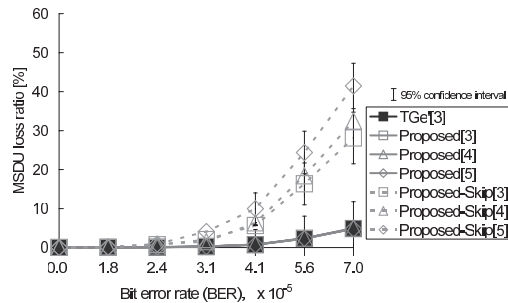


Fig. 7 MSDU loss ratio for video (Music video).

Figures 5 and 6 depict the MU loss ratio for audio and that for video, respectively, as a function of BER. We can observe in Fig. 5 that the MU loss ratio for audio becomes very small because frame error rate is small since the length of an audio MPDU is short under our simulation conditions.

Figure 6 reveals that the MU loss ratio for video increases as BER becomes larger. In addition, Fig. 6 shows that the values of the MU loss ratio for video become almost the same for the TGe', Proposed, and Proposed-Skip schemes. This is because MPDU loss is caused in our simulation only if a station cannot succeed in sending an MPDU within the maximum retry number of retransmission attempt.

Figure 7 shows the MSDU loss ratio for video as a function of BER. We can observe from Figs. 6 and 7 that in the cases of the TGe' scheme and the proposed scheme without the source video frame skipping, the MU loss ratio for video becomes much larger than the MSDU loss ratio when $BER > 3.1 \times 10^{-5}$. A video frame is fragmented into several MSDUs and the loss of an MSDU of a video frame leads to the loss of the corresponding whole video frames. Furthermore, the loss of a video frame leads to the loss of all video frames of a GOP as shown in Sect. 3. Therefore, the values of the MU loss ratio become large even if those of the MSDU loss ratio are small.

We can also find in Fig. 7 that the MSDU loss ratio for the proposed scheme with the source video frame skipping is much larger than that for the TGe' scheme and the Proposed scheme without the video frame skipping if BER is larger than 3.1×10^{-5} . In the case of the Proposed-Skip scheme, many MSDUs are dropped at the source station when BER is larger than 3.1×10^{-5} . This leads to a larger value of MSDU loss ratio. It should be noted that all video frames of the GOP are dropped at the source station if an MSDU of a video frame of a GOP is failed to transmit from the station to the AP.

Figures 8 illustrates the coefficient of variation of output interval for video versus BER. For all the schemes shown in this figure, the coefficient of variation of output interval for video increases when BER becomes larger than 3.1×10^{-5} except the case of Proposed-Skip[5]. This is because MPDU retransmission occurs more frequently as BER increases. In the case of Proposed-Skip[5], it should be noted that the average MU delay for video deteriorates

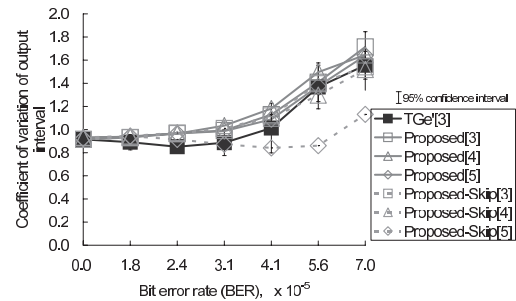


Fig. 8 Coefficient of variation of output interval for video (Music video).

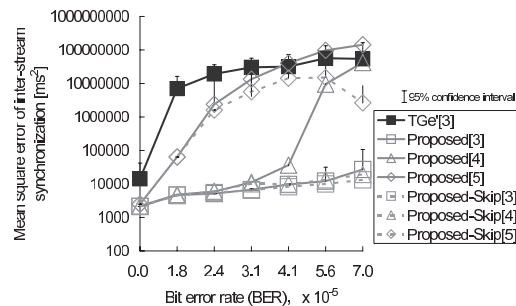


Fig. 9 Mean square error of inter-stream synchronization (Music video).

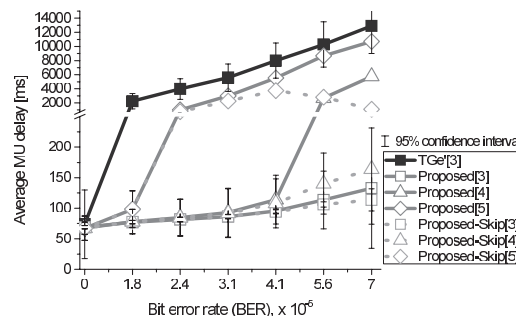


Fig. 10 Average MU delay for video (Sport).

drastically if BER increases beyond 1.8×10^{-5} (see Fig. 4), though the coefficient of variation of output interval is small compared to the other cases.

In Fig. 9, mean square error of inter-stream synchronization is shown as a function of BER. From Figs. 4 and 9, we can say that the inter-media synchronization quality for the TGe' and proposed schemes highly depends on the average MU delay for video.

Then, we will show application-level QoS for *Sport*. Figs. 10 and 11 plot the average MU delay for video, and the mean square error of inter-stream synchronization, respectively, as a function of BER. From Figs. 4 and 9–11, we can find that the values of these application-level QoS parameters for *Music video* become almost the same as those for *Sport*.

5.2 QoE

We then examine the QoE of the TGe' and the proposed

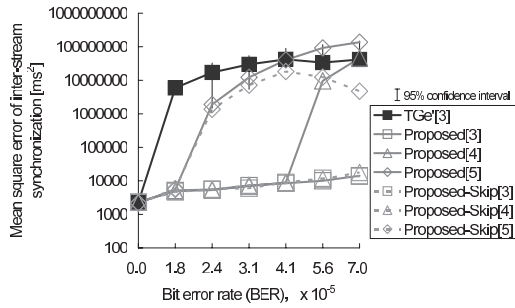


Fig. 11 Mean square error of inter-stream synchronization (Sport).

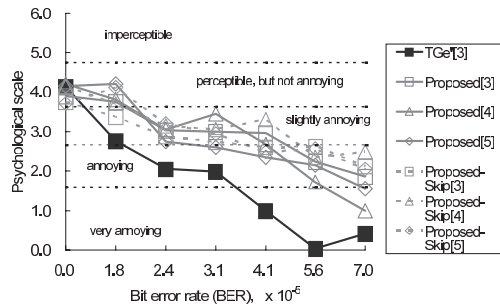


Fig. 12 Psychological scale (Music video).

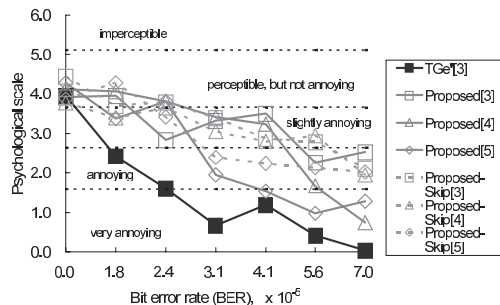


Fig. 13 Psychological scale (Sport).

schemes from subjective experimental results.

In this paper, the results obtained from the rating-scale method are calculated into the *interval scale* with the *law of categorical judgment* [21]. The *mean opinion score (MOS)* is an *ordinal scale*; the integers assigned to the categories only have a greater-than-less-than relation between them. In contrast, in the interval scale, an interval between the scale values means a distance between amounts of the sensory attribute measured [14].

To verify the obtained interval scale, we have performed Mosteller's test [21]. From the Mosteller's test, we cannot reject the hypothesis that the obtained interval scale fits the observed data at a significance level of 0.01. Thus, we refer to the interval scale as the *psychological scale* [14].

Figures 12 and 13 show the psychological scale versus BER for *Music video* and *Sport*, respectively. In these figures, we selected the minimum value of the psychological scale as the origin of the ordinate, and each of four horizontal dotted lines indicates the boundary of a category. We

applied the law of categorical judgment to the measurement results by the rating-scale method for *Music video* and those for *Sport* separately. Therefore, the boundary of each category in Fig. 12 and that in Fig. 13 have become different values. These figures show the values of the psychological scale for the TGe' scheme and the proposed scheme with and without the source video frame skipping.

From Figs. 12 and 13, we can observe that the QoE for the three schemes deteriorates as BER increases. These figures also show that the values of the psychological scale for the TGe' scheme become the lowest of the three schemes. This is because the TGe' scheme allocates constant TXOP duration. Therefore, the TGe' scheme is not flexible for packet size fluctuation of the VBR traffic and cannot allocate surplus bandwidth for MPDU retransmission.

We then discuss the effect of content types on the QoE. We find in these figures that the psychological scale for *Sport* tends to deteriorate more drastically than that for *Music video* as BER becomes larger, though we have found that the values of the application-level QoS parameters become almost the same; for example, Proposed[5] for *Sport* indicates "very annoying" if BER is larger than 3.1×10^{-5} , while Proposed[5] for *Music* indicates around the boundary between "annoying" and "very annoying" even if BER is 7.0×10^{-5} . This implies that the QoE depends on content types.

Furthermore, Fig. 13 also shows that the QoE for Proposed-Skip[5] becomes higher than that for Proposed[5] when BER is larger than 2.4×10^{-5} . This means that the QoE for *Sport* can be improved to some extent by the source video frame skipping. We can also find in Fig. 13 that the QoE for Proposed-Skip[4] becomes higher than that for Proposed[4] when BER is larger than 4.1×10^{-5} .

We then examine which application-level QoS parameters affect the QoE dominantly. For *Music video*, we can see in Fig. 12 that the QoE for the TGe' scheme is lower than that for the Proposed and Proposed-Skip schemes. In addition, the difference in the psychological scale value between the Proposed-Skip scheme and the Proposed scheme is small. We can also make similar observations in Fig. 3. Therefore, we can say that the average MU delay for audio highly affects the QoE. This makes us confirm that *Music video* is audio-dominant. In the case of *Sport*, Fig. 13 shows that the QoE for the proposed scheme can be improved by the source video frame skipping when BER is large to some extent. We can also make similar observations in Fig. 10; therefore, the QoE for *Sport* is affected by the average MU delay for video. This result implies that *Sport* is video-dominant.

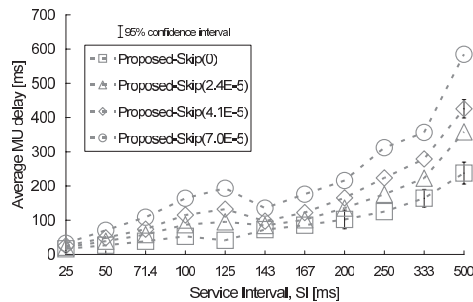
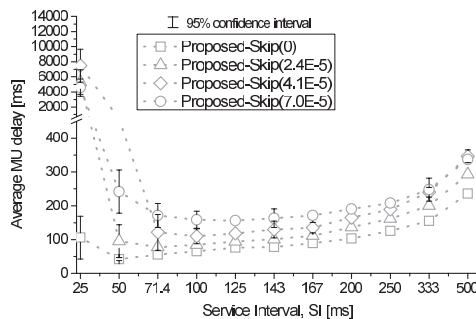
From the above observations, we can say that the source video frame skipping is effective in improving QoE for *Sport* when the channel capacity for retransmission is insufficient owing to a large value of BER.

6. The Effect of SI

In this section, we examine the effect of SI on the

Table 3 Relationship between MSI and SI (ms).

MSI	25	50	75	100	125	150	175	200	250	350	500
SI	25	50	71.4	100	125	143	167	200	250	333	500

**Fig. 14** Average MU delay for audio (Music video).**Fig. 15** Average MU delay for video (Music video).

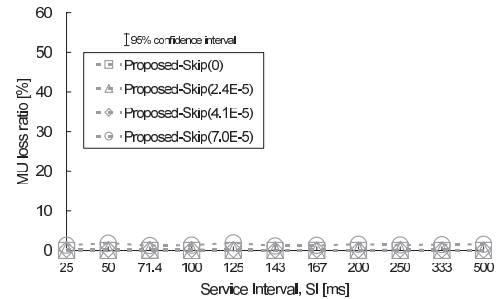
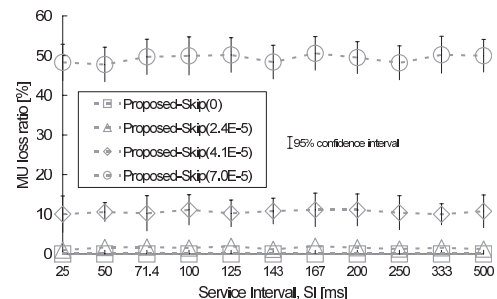
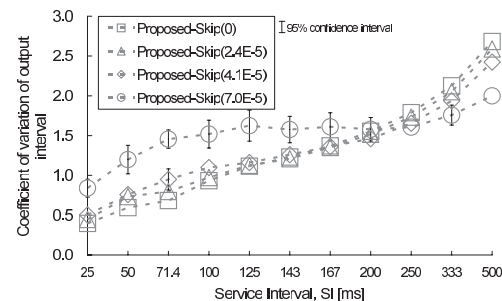
application-level QoS and QoE. We assess here the QoS for the proposed scheme with the source video frame skipping because we found in the previous section that the QoS for the proposed scheme with the source video frame skipping becomes higher than or nearly the same as that for the proposed scheme without the video frame skipping and the TGe' scheme. In the simulation, we set eleven values of MSI, which correspond to SI as shown in Table 3. The number of multimedia stations is set to 4.

In the figures to be shown, notation Proposed-Skip(4.1E-5), for example, in this section refers to the results when BER is 4.1×10^{-5} .

6.1 Application-Level QoS

In the following numerical results, we will show application-level QoS for *Music video*. We have confirmed through simulation that the application-level QoS parameters for *Music video* and those for *Sport* are almost the same.

Figures 14 and 15 show the average MU delay of audio and that of video as a function of the SI. Figure 14 reveals that the average MU delay for audio tends to become larger as the SI increases because the polling interval for each station becomes longer. It should be noted that the average MU delay for audio decreases if SI changes from 125 ms to 143 ms under lossy conditions because N_i^P calculated from Eq. (5) changes from 1 to 2.

**Fig. 16** MU loss ratio for audio (Music video).**Fig. 17** MU loss ratio for video (Music video).**Fig. 18** Coefficient of variation of output interval for video (Music video).

We then find in Fig. 15 that the average MU delay for video also becomes larger as BER increases when $SI > 100$ ms. However, this figure also shows that the average MU delay for video deteriorates if the SI decreases below 71.4 ms for Proposed-Skip(2.4E-5), Proposed-Skip(4.1E-5), and Proposed-Skip(7.0E-5), and 50 ms for Proposed-Skip(0). This is because polling overhead increases if the SI is too small.

Figures 16 and 17 plot the MU loss ratio for audio and that for video as a function of the SI. We can see in these figures that the MU loss ratio for audio and that for video are almost constant regardless the SI.

Figures 18 illustrates the coefficient of variation of output interval for video versus the SI. This figure shows that the coefficient of variation of output interval increases with the increase of SI. This is because a longer polling interval leads to a larger value of MU delay.

Finally, Fig. 19 shows mean square error of inter-stream synchronization as a function of the SI. We can see

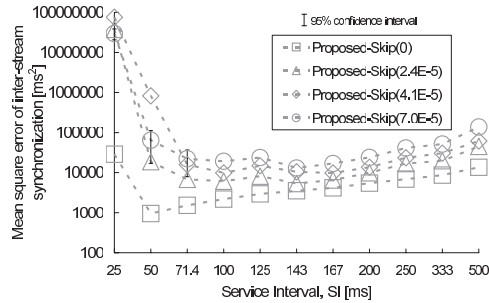


Fig. 19 Mean square error of inter-stream synchronization (Music video).

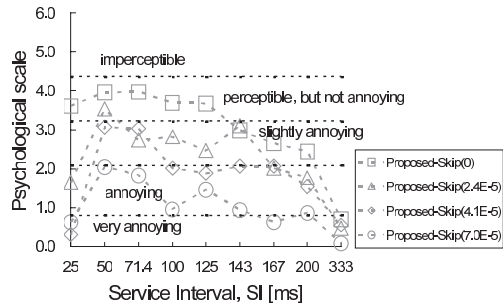


Fig. 20 Psychological scale (Music video).

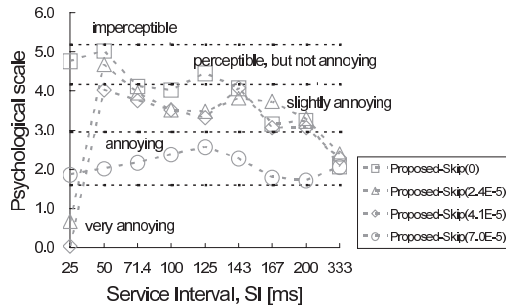


Fig. 21 Psychological scale (Sport).

in this figure that Fig. 19 exhibits similar characteristics to Fig. 15.

6.2 QoE

Figures 20 and 21 show the psychological scale for *Music video* and *Sport*, respectively, as a function of the SI. These figures show that the QoE except Proposed-Skip(7.0E-5) for the two contents deteriorates gradually if the SI increases beyond 100 ms because of longer polling interval. We can also observe in Figs. 14, 15, 18, and 19 that the application-level QoS parameters deteriorate as SI increases beyond 100 ms. In the case of Proposed-Skip(7.0E-5), we find in Figs. 20 and 21 that the QoE is low for all the values of SI. We see in Fig. 17 that the MU loss ratio for video is large when BER is 7.0×10^{-5} .

Figures 20 and 21 also show that the QoE becomes low when the SI is below 50 ms except Proposed-Skip(0) since the fraction of the surplus bandwidth to the SI becomes too

small to allocate enough TXOP duration to stations in a noisy environment. From Fig. 15 and 19, we find that the values of the average MU delay for video and those of mean square error of inter-stream synchronization are very large when SI = 25 ms.

From Figs. 20 and 21, we can say that the appropriate value of the SI for the two contents is 50 ms, which is equal to the inter-MU time for video traffic. Therefore, the inter-MU time for video should be selected as the SI to achieve high QoE under our simulation conditions.

7. Conclusions

In this paper, we have proposed a new cross-layer packet scheduling scheme for audio-video transmission with the IEEE 802.11e HCCA to solve the problems of the TGe scheme. We have also introduced video frame skipping at source stations to reduce the traffic volume sent to the wireless channel in a noisy environment.

We have compared the TGe' scheme, the proposed packet scheduling scheme with and without the source video frame skipping in terms of application-level QoS and QoE. Numerical results have shown that the proposed packet scheduling scheme can achieve higher quality than the TGe' scheme under lossy channel conditions. We have also showed that the proposed scheduling scheme can improve the QoE as well as the application-level QoS by utilizing the source video frame skipping. Furthermore, we also examined the effect of the SI on the QoS and QoE in the proposed packet scheduling scheme. As a result, we have shown that the appropriate value of SI to realize high QoS is equal to the inter-MU time for video traffic.

In the simulation of this paper, the inter-MU time for video and that for audio are set to 50 ms and 125 ms, respectively: that is, the former is shorter than the latter. Our future study includes QoS and QoE assessment using other values of inter-MU time for audio and that for video. We should also investigate the QoS and QoE of the proposed scheme with error concealment.

Acknowledgment

This work was supported by the Grant-In-Aid for Scientific Research of Japan Society for the Promotion of Science under Grant 21360183.

References

- [1] IEEE 802.11 WG, "Specific requirements part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Amendment 8: Medium access control (MAC) quality of service enhancements amendment 7: Medium access control (MAC) quality of service (QoS) enhancements," IEEE 802.11 Std., 2005.
- [2] S. Tasaka and Y. Ishibashi, "Mutually compensatory property of multimedia QoS," Conf. Rec. IEEE ICC'02, pp.1105–1111, April/May 2002.
- [3] ITU-T Rec. G.100/P.10 Amendment 1, "New appendix I—Definition of Quality of Experience (QoE)," Jan. 2007.
- [4] Z. Azri, T. Suzuki, and S. Tasaka, "A packet scheduling scheme

for audio-video transmission over error-prone IEEE 802.11e HCCA wireless LANs," Conf. Rec. IEEE PIMRC'09, Sept. 2009.

- [5] A. Grilo, M. Macedo, and M. Nunes, "A scheduling algorithm for QoS support in IEEE 802.11e networks," *Wireless Communications*, vol.10, no.3, pp.36–43, June 2003.
- [6] P. Ansel, Q. Ni, and T. Tueti, "An efficient scheduling scheme for IEEE 802.11e," *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, WiOpt.*, March 2004.
- [7] I. Inan, F. Keceli, and E. Ayanoglu, "An adaptive multimedia QoS scheduler for 802.11e wireless LANs," *Conf. Rec. IEEE ICC'06*, June 2006.
- [8] D. Xue, Y. Qin, and C.K. Siew, "A service curve based scheduling algorithm for QoS support in 802.11e wireless LANs," *Conf. Rec. IEEE GLOBECOM'06*, Nov. 2006.
- [9] M.M. Rashid, E. Hossain, and V.K. Bhargava, "HCCA scheduler design for guaranteed QoS in IEEE 802.11e based WLANs," *Proc. IEEE WCNC'07*, March 2007.
- [10] A. Foronda, Y. Higuchi, C. Ohta, M. Yoshimoto, and Y. Okada, "Service interval optimization with delay bound guarantee for HCCA in IEEE 802.11e WLANs," *IEICE Trans. Commun.*, vol.E90-B, no.11, pp.3158–3169, Nov. 2007.
- [11] M.V.D. Schaar, Y. Andreopoulos, and Z. Hu, "Optimized scalable video streaming over IEEE 802.11a/e HCCA wireless networks under delay constraints," *IEEE Trans. Mobile Computing*, vol.5, no.6, pp.755–768, June 2006.
- [12] Y.P. Fallah, P. Nasiopoulos, and H. Alnuweiri, "Scheduled and contention access transmission of partitioned H.264 video over WLANs," *Conf. Rec. IEEE GLOBECOM'07*, Nov. 2007.
- [13] Z. Azri, T. Suzuki, and S. Tasaka, "Guaranteeing QoE in audio-video transmission by IEEE 802.11e HCCA," *IEICE Trans. Fundamentals*, vol.E91-A, no.7, pp.1551–1561, July 2008.
- [14] S. Tasaka and Y. Ito, "Psychometric analysis of the mutually compensatory property of multimedia QoS," *Conf. Rec. IEEE ICC'03*, pp.1880–1886, May 2003.
- [15] R. Haywood, S. Mukherjee, and X.H. Peng, "Investigation of H.264 video streaming over an IEEE 802.11e EDCA wireless testbed," *Conf. Rec. IEEE ICC'09*, June 2009.
- [16] N. Cranley and M. Davis, "Video frame differentiation for streamed multimedia over heavily loaded IEEE 802.11e WLAN using TXOP," *Proc. IEEE PIMRC'07*, Sept. 2007.
- [17] IEEE 802.11 WG, "Supplement to part 11: Wireless LAN medium access control (MAC) and physical layer specifications: High-speed physical layer extension in the 2.4 GHz band," *IEEE 802.11 Std.*, 1999.
- [18] W. Xiuchao, "Simulate 802.11b channel within ns2," <http://www.comp.nus.edu.sg/~wuxiucha/research/reactive/publication/Simulate80211ChannelWithNS2.pdf>, 2004.
- [19] HFA3861B; Direct sequence spread spectrum baseband processor, Intersil, Jan. 2000.
- [20] ITU-T Rec. P.911, "Subjective audiovisual quality assessment methods for multimedia applications," Dec. 1998.
- [21] J.P. Guilford, *Psychometric Methods*, McGraw-Hill, New York, 1954.



Zul Azri Bin Muhamad Noh received the B.E. (electrical and computer engineering) and M.E. (computer science and engineering) degrees from Nagoya Institute of Technology, Nagoya, Japan in 2005 and 2007, respectively. He is now a Ph.D. student at Nagoya Institute of Technology. His current research interests include wireless LAN, packet scheduling scheme and multimedia QoS.



Takahiro Suzuki received the B.S., M.S. and Ph.D. degrees in electrical and computer engineering from Nagoya Institute of Technology, Nagoya, Japan, in 1989, 1991 and 1994, respectively. He worked at Nagoya Institute of Technology from April 1994 to March 1995. In April 1995, he joined Faculty of Social and Information Sciences, Nihon Fukushi University, Handa, Aichi, Japan. He is now a Professor in the Health Assist Information Course, Department of Human Care Engineering, Faculty of Health Sciences, Nihon Fukushi University. His current research interests include wireless networks and multimedia communication protocols. Dr. Suzuki is a member of Information Processing Society of Japan.



Shuji Tasaka received the B.S. degree in electrical engineering from Nagoya Institute of Technology, Nagoya, Japan in 1971, and the M.S. and Ph.D. degrees in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1973 and 1976, respectively. Since April 1976, he has been with Nagoya Institute of Technology, where he is now a Professor in the Department of Computer Science and Engineering, Graduate School of Engineering. In the 1984–1985 academic year, he was a Visiting Scholar in the Department of Electrical Engineering at the University of California, Los Angeles. His current research interests include QoE/QoS, wireless networks, and multimedia communication protocols. He is the author of a book entitled *Performance Analysis of Multiple Access Protocols* (MIT Press, Cambridge, MA, 1986). Dr. Tasaka is a member of IEEE, ACM and Information Processing Society of Japan.