QoE Assessment of Audio-Video Transmission over IEEE 802.11aa GCR with Off-the-Shelf Devices

Daiki Teranishi and Toshiro Nunome
Department of Computer Science and Engineering, Graduate School of Engineering,
Nagoya Institute of Technology, Nagoya 466–8555, Japan
teranisi@inl.nitech.ac.jp, nunome@nitech.ac.jp

***Abstract*—This paper evaluates the effect of reliable groupcast on Quality of Experience (QoE) of audio-video IP transmission over a wireless LAN with off-the-shelf devices. We utilize IEEE 802.11aa GroupCast with Retries (GCR) Unsolicited Retry as a reliable groupcast scheme. We think that a wireless access point groupcasts audio and video streams to several terminals connected to the access point. We use GCR Unsolicited Retry as a transmission method of audio-video data and evaluate application-level Quality of Service (QoS) on several experimental conditions. In addition, we assess QoE by a subjective experiment. As a result, we find that GCR Unsolicited Retry presents higher QoE than a conventional transmission method.**

***Index Terms*—wireless LAN, IEEE 802.11aa, groupcast, audio and video transmission, QoE**

I. Introduction

Owing to the advancement of wireless communication technologies, mobile terminals such as tablet PCs and smartphones have become popular rapidly. The terminals are equipped with wireless LAN devices. People can use wireless communications easily.

The IEEE 802.11 committee has been standardizing wireless LANs and is still revising the standards. The maximum transmission rate of the first standard of IEEE 802.11 is 2 Mbps; it is slower than the rate in the wired networks. However, the latest standard has the transmission rate of Gbps order. The wireless LANs can be alternative of wired LANs. We can select the wired and wireless connections according to objectives and conditions.

IP networks are generally best-effort; packets can be lost during transmission and can be affected by network delay jitter. These impairments deteriorate the output quality of continuous media such as audio and video; then, *Quality of Service (QoS)* becomes lower. It leads to deterioration of *Quality of Experience (QoE)* [1] in many cases. Enhancement of QoE in best-effort networks is essential for many network services.

We can use groupcast for efficient transmission of the continuous media. It can transmit the same data to multiple receivers. Multicast in the network layer utilizes groupcast in the data-link layer. There is no reliable transmission mechanism for groupcast in previous IEEE 802.11 wireless LANs [2].

Thus, IEEE 802.11aa has been standardized for reliable groupcast communications [3]. It enhances QoS provisioning mechanisms of IEEE 802.11e. It also introduces GroupCast with Retries (GCR).

There are many studies on audiovisual multicast/groupcast transmission by means of IEEE 802.11aa GCR. Most of them employ computer simulation for performance evaluation. We have also evaluated QoE of audiovisual groupcast with IEEE 802.11aa GCR [4]. The computer simulation models real situation. However, the model is not necessarily faithful to the actual situation because of some abstraction.

Several pieces of research on multimedia wireless communications with actual devices have been performed. Reference [5] evaluates the QoE of several VoIP applications through a campus wireless LAN; the paper does not employ video. In addition, the study does not treat multicast/groupcast communications.

Reference [6] performs an experiment to evaluate the groupcast mechanisms. The authors have implemented the mechanisms over commodity hardware and have assessed their performance under a variety of real-life scenarios. However, the paper only evaluates the delivery rate and throughput.

In this paper, we assess QoE of IEEE 802.11aa with real devices; it is a challenging issue. We propose a methodology to assess QoE of audio and video reliable groupcast by an experiment with off-the-shelf devices. As a first step of the study, we deal with IEEE 802.11aa GCR Unsolicited Retry. We firstly evaluate application-level QoS and gain trace files of output timing under several experimental conditions. From the files, a representative file for each experimental condition is picked out. We then perform a subjective experiment for obtaining multidimensional QoE.

The rest of the paper is structured as follows. Section II outlines the data transmission methods. Section III explains the driver and firmware employed in the experiment. Sections IV and V describe the method of the experiment and that of the QoE assessment, respectively. We present experimental results in Section VI, and Section VII concludes this paper.

II. Data transmission methods

This section explains GCR Unsolicited Retry. We depict a transmission sequence of GCR Unsolicited Retry in Figure 1. The scheme transmits a MAC frame multiple times. The number of transmissions in Figure 1 is two. The access way to the wireless channel in the scheme is the same as that

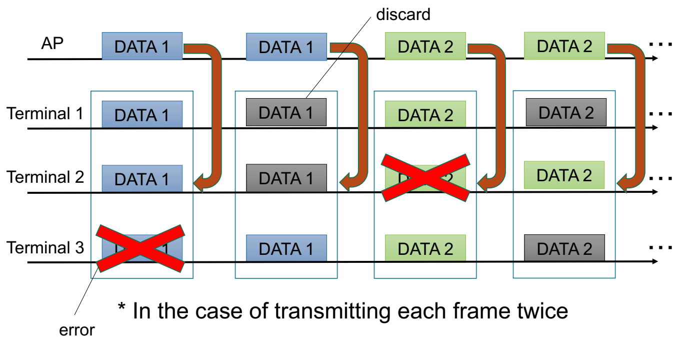

Fig. 1. GCR Unsolicited Retry (transmit each frame twice)

in Groupcast without retry. The retry flag is attached to a frame which is transmitted second or more times. The terminal discards the duplicate MAC frames when the terminal receives the same MAC frame two or more times.

This scheme groupcasts each MAC frame predetermined times. It can provide soft reliability irrespective of the receiver's condition.

III. Driver and Firmware

In order to evaluate the GCR mechanisms on the real environment, we need to implement the access control mechanisms at the MAC layer into off-the-shelf devices. However, there are few commercial wireless LAN devices in which we can modify the firmware. In addition, source codes of the wireless LAN devices are usually secret.

In this study, we employ OpenFWWF [7] to customize the MAC layer of wireless LANs. OpenFWWF is a wireless LAN firmware for Broadcom's wireless LAN chipsets developed by the UniBS NTW group at the University of Brescia. We have implemented IEEE 802.11aa GCR Unsolicited Retry according to the technical report [8]. With the firmware, the b43 driver, and the BCM4318 wireless LAN chipset, we realize the function of GCR Unsolicited Retry.

IV. Experimental Condition

Figure 2 shows the network configuration in the experiment. Media Server transmits audio and video to two Media Receivers. We assume Media Server as a wireless LAN access point; in actual, the ad-hoc mode is employed owing to the simplicity of implementation. Media Server is DELL Dimension 9200, and Media Receivers are DELL Inspiron 530. The terminals are equipped with Buffalo's WLI2-PCI-G54S, which has the BCM4318 wireless LAN chipset. The Linux kernel version is 2.6.36.4. We employ a modified b43 driver and OpenFWWF.

The MAC protocol is IEEE 802.11 DCF. The transmission rate is 12 Mbps. The wireless channel is 11 on the 2.4 GHz band.

Media Server and Media Receivers are placed into a double-layer RFI shielded tent (Toyo's DST-S2-NR). The size of the tent is 115 cm^3. The shielding performance is 50 dB at 50 kHz

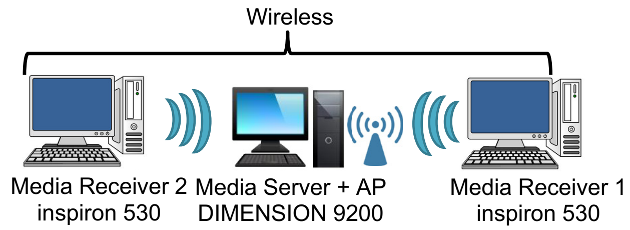

Fig. 2. Experimental system

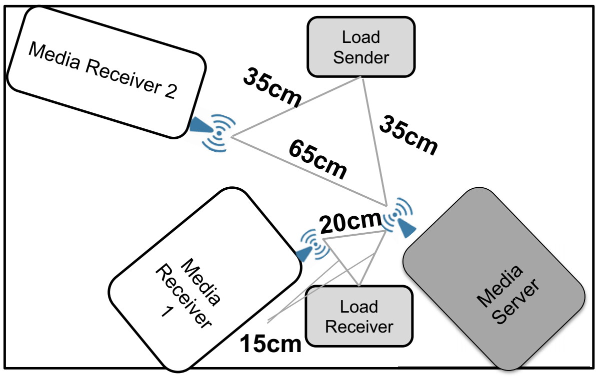

Fig. 3. Layout

$\sim$ 18 GHz. Figure 3 shows the layout inside of the tent. The distance between the antenna of Media Server and that of Media Receiver 1 is 20 cm. The antenna of Media Receiver 2 is placed 65 cm away from that of Media Server.

We employ RTP/UDP/IP for audio and video transmission. The specification of audio and video are shown in Table I. A *Media Unit (MU)* is a unit for media synchronization control. We define a video frame and a constant number of audio samples as a video MU and an audio MU, respectively. A video frame consists of 9 video slices, and an RTP/UDP packet transmits each slice. An audio MU is transmitted by an RTP/UDP packet. In the IP layer, the packets can be fragmented according to Maximum Transmission Unit (MTU). We employ aac for audio encoding and JM [9] version 18.0 for video encoding.

We employ a simple playout buffering control for each media as a media synchronization control. The playout buffering time is set to 500 ms. When a slice consisting of a video MU is dropped, we employ error concealment and output the MU.

The contents of audio and video in the experiment are *news* and *sport*. The content of sport is a baseball game, and that of news is an opening speech of a news TV program.

As an interference wireless communication for the audio and video streams, Load Sender generates UDP datagrams of 1480 bytes each with exponentially distributed interval and sends them to corresponding Load Receiver. Both load terminals are laptop PCs with internal wireless LAN devices. The average bitrate of the interference load traffic is 500 kbps.

TABLE I
SPECIFICATION OF AUDIO AND VIDEO

item	audio	video
coding method	MPEG-4 AAC-LC	H.264/AVC
image size [pixels]	-	768 × 432
number of slices	-	9
GOP	-	I+4P's
average MU rate [MU/s]	46.875	30.0
average bitrate [kb/s]	128.0	(news) 1976.71 (sport) 2171.22
duration [sec]	10.0	10.0

TABLE II
PAIRS OF POLAR TERMS

category	pair of polar terms
Video temporal	The video is smooth – The video is rough
Video spatial	The video is sharp – The video is blurred
Audio	The audio is natural – The audio is artificial
Synchronization	The audio and video are in synchronization – The audio and video are out of synchronization
Overall satisfaction	Excellent – Bad

The wireless channel is the same as that for the audio and video groupcast, but interference traffic uses different ESSID.

The number of experimental conditions for comparison is eight; they are combinations of the number of data transmissions in GCR unsolicited retry and with or without interference traffic.

The experiment started at around 23:00 on a weekday. We checked wireless RF conditions around the experimental system by Chanalyzer before beginning the experiment.

V. QoE ASSESSMENT METHOD

In the experiment, we ask the assessors to evaluate the audio and video streams output at the media receiver terminal. We then obtain subjective assessment results.

In the subjective assessment, we employ trace files which record the receive timing of video slices and audio MUs. It becomes easy to reproduce the experimental conditions using the files. We select the trace file which has the nearest application-level QoS parameter values from the average values for each terminal and condition.

The total assessment time for each assessor is about 15 minutes. The assessors are 15 male students of our university who major in computer science.

In the assessment, we employ five pairs of polar terms. Table II shows the pairs of polar terms in the subjective experiment. For each pair, a subjective score is measured by the *rating scale method* [10]. In the method, an assessor classifies the stimuli into a certain number of categories; here, each criterion is evaluated to be one of five grades (score 5 is the best, and score 1 is the worst). Finally, we calculate the mean opinion score (MOS), which is an average of the rating scale scores for all the users.

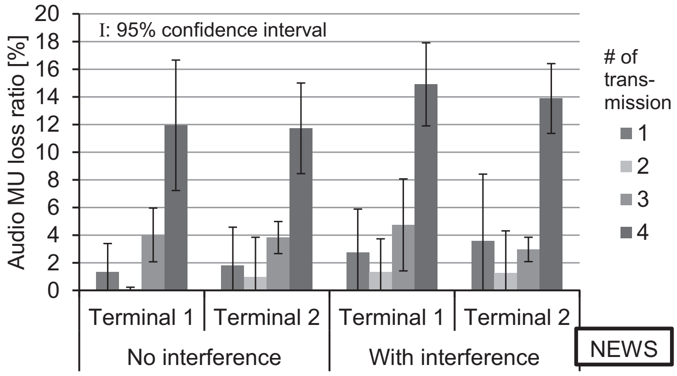

Fig. 4. Audio MU loss ratio (news)

VI. EXPERIMENTAL RESULTS

A. Application-level QoS

In this section, we discuss the application-level QoS assessment results. The application-level QoS is adjacent to QoE at the layered network model. Hence, the application-level QoS is closely related to QoE.

This paper deals with the audio MU loss ratio, the video MU loss ratio, and the error concealment ratio for video as the application-level QoS parameters. The error concealment ratio represents the percentage of slices error-concealed (i.e., lost slices) in a frame; it shows the image quality of video stream. The MU loss ratio is the ratio of the number of MUs not output at the recipient to the number of MUs transmitted by the sender.

On the application-level QoS assessment, there are small differences between the two contents. Thus, we focus on the results for news here.

Figure 4 depicts the MU loss ratio of audio. Figure 5 represents the error concealment ratio of video. The video MU loss ratio is shown in Figure 6. Each parameter value is the average of 15 experimental runs. We also show 95 % confidence intervals in these figures. In the abscissa, "no interference traffic" means that there is no interference data transmission. On the other hand, "with interference traffic" has the interference load traffic. "Terminal 1" shows the results in Media Receiver 1, and "Terminal 2" represents the results in Media Receiver 2. In the legends of the figures, the notations '1', '2', '3', and '4' mean that the number of transmissions is 1, 2, 3, and 4, respectively.

From the results in Figures 4, 5 and 6, we discuss the effect of the interference traffic.

When the number of transmissions is one, "with interference traffic" has higher video error concealment ratio and video MU loss ratio than "no interference traffic". This is because frame collisions and transmission postponements increase owing to the interference traffic.

For transmitting each frame twice, the error concealment ratio and the video MU loss ratio are smaller than those for

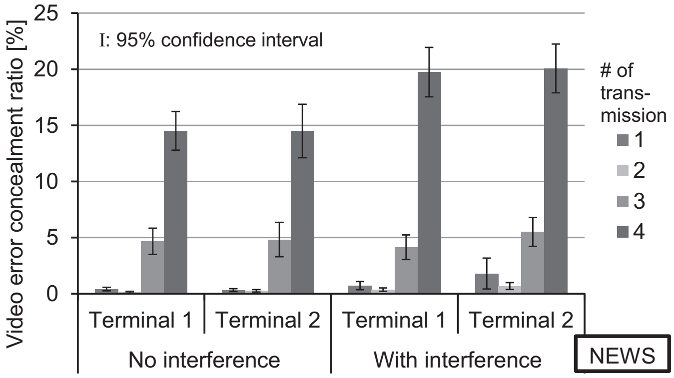

Fig. 5. Video error concealment ratio (news)

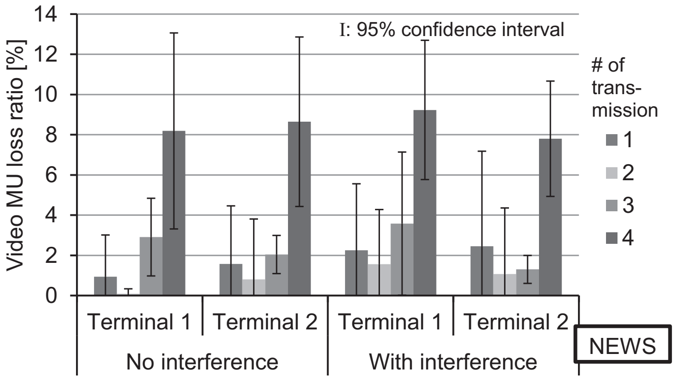

Fig. 6. Video MU loss ratio (news)

transmitting each frame once. The duplicate transmission can enhance the reliability of frame transmission.

When the number of transmissions is three, the difference between "no interference traffic" and "with interference traffic" on the video MU loss ratio and the error concealment ratio is small. In addition, "with interference traffic" in terminal 2, the MU loss ratio for transmitting each frame three times is smaller than that for transmitting once. In this experimental condition, the frame loss and propagation error can cause easily because of interference traffic. When the number of transmissions is three, the repetition of transmission can recover the frame loss. Terminal 1 (Media Receiver 1) is nearer from AP and the load terminals than terminal 2 (Media Receiver 2). Thus, the effect of channel interference in terminal 1 is larger than terminal 2.

When the number of transmissions is four, the MU loss ratio is larger than the other number of transmissions. This is because of channel congestion due to multiple frame transmission.

B. QoE assessment results

We show the QoE assessment results in Figures 7 through 16. Figures 7 through 11 are the results for news, while Figures 12 through 16 depict the MOS values for sport.

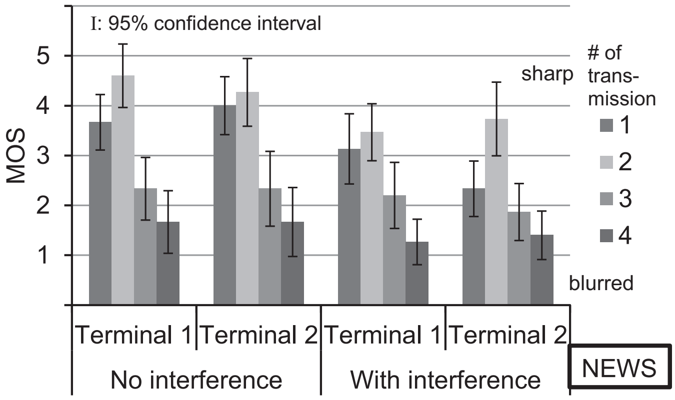

Fig. 7. The video is sharp – blurred (news)

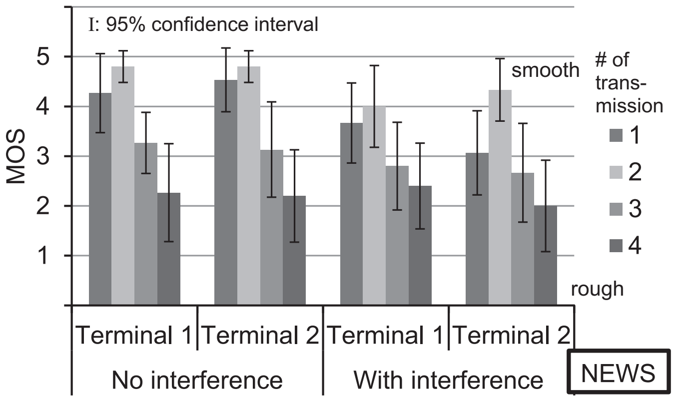

Fig. 8. The video is smooth – rough (news)

1) The video is sharp – blurred: We see in Figures 7 and 12 that the transmission twice achieves the highest MOS value among the four candidates of transmission times irrespective of the interference traffic. The result relates to the error concealment ratio in Figure 5.

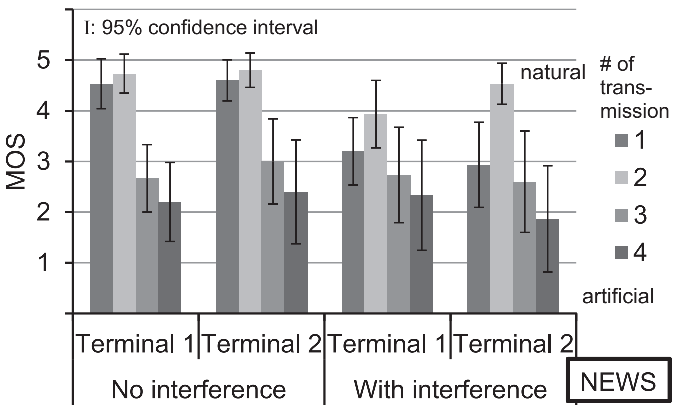

Fig. 9. The audio is natural – artificial (news)

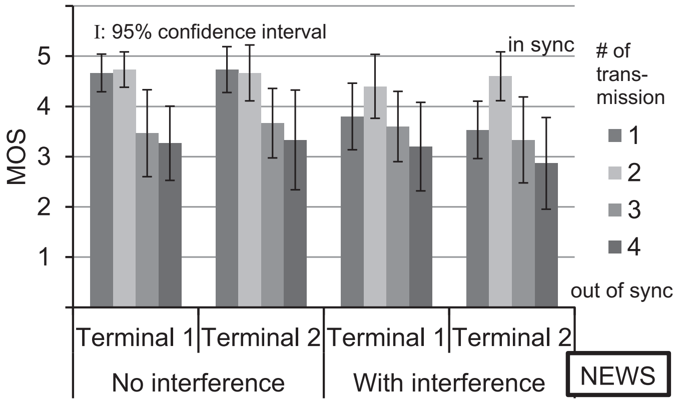

Fig. 10. The audio and video are in synchronization – out of synchronization (news)

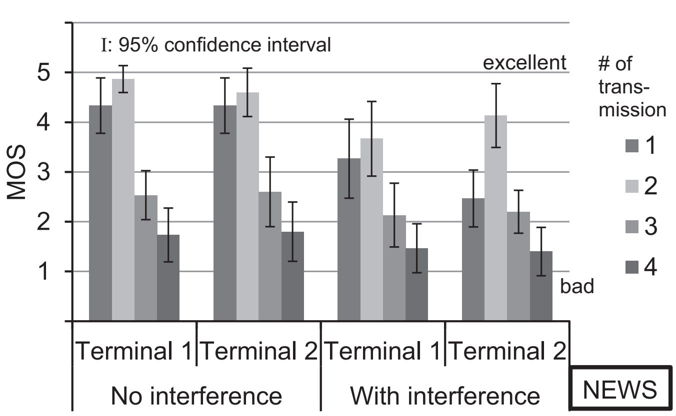

Fig. 11. Excellent – Bad (news)

2) The video is smooth – rough: In Figures 8 and 13, we notice that the MOS value for transmission twice is the highest, that for transmission once is the second highest, and that for transmission three times is the third. It strongly correlates to the video MU loss ratio shown in Figure 6.

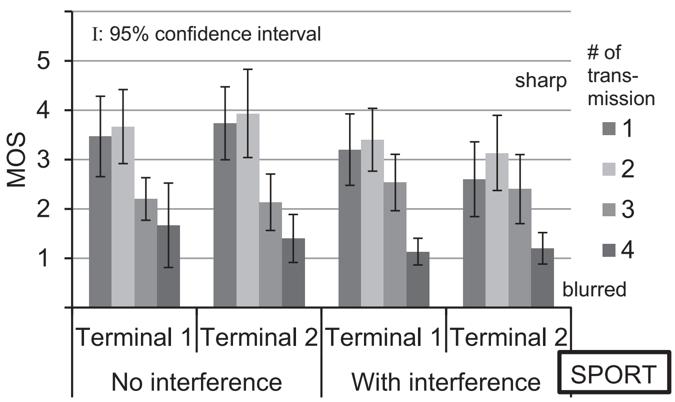

Fig. 12. The video is sharp – blurred (sport)

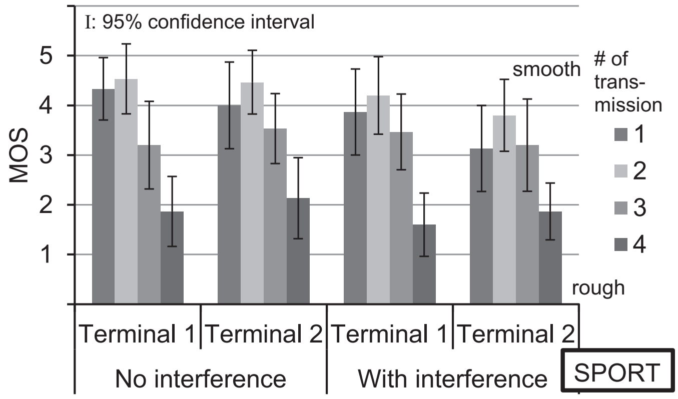

Fig. 13. The video is smooth – rough (sport)

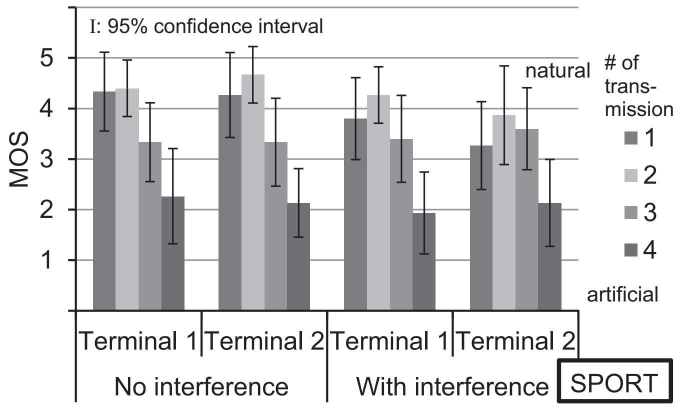

Fig. 14. The audio is natural – artificial (sport)

3) The audio is natural – artificial: We find in Figures 9 and 14 that the MOS value is the largest when the number of transmissions is two. It is reflected by the results of the audio MU loss ratio in Figure 4.

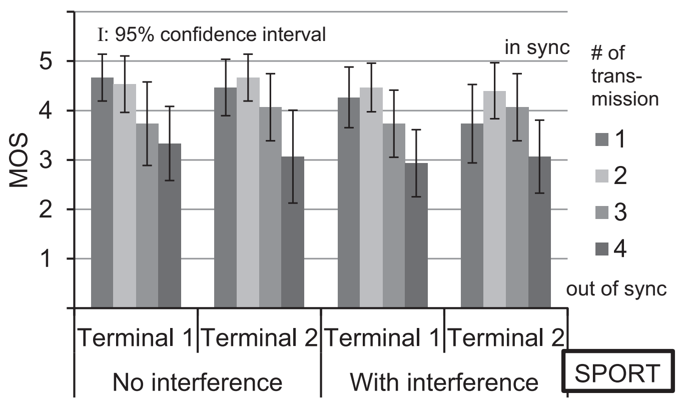

Fig. 15. The audio and video are in synchronization – out of synchronization (sport)

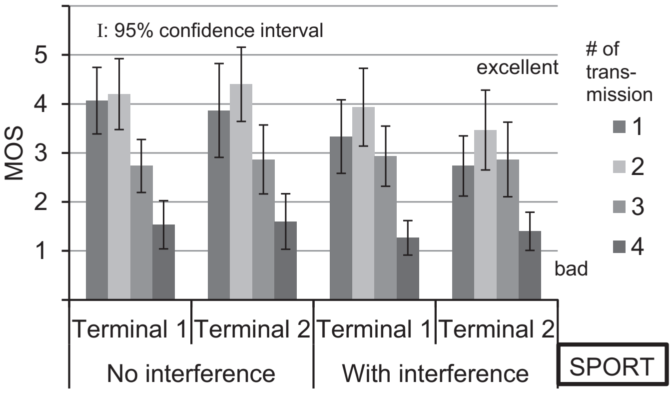

Fig. 16. Excellent – Bad (sport)

4) The audio and video are in synchronization – out of synchronization: We see in Figures 10 and 15 that transmission of each frame twice achieves the highest MOS value with interference traffic, while there is a small difference between once and twice without interference traffic. The MU loss ratio with interference traffic is larger than that without interference traffic. The assessors feel out of synchronization when the MU loss occurs. When the number of transmissions is two, the lost MUs can be recovered efficiently. Thus, the MOS value in transmission twice with interference traffic becomes large.

5) Excellent – Bad: In Figures 11 and 16, we find that the MOS value for transmitting twice is the highest, once is the second, three times is the third, and four times is the worst irrespective of the interference traffic. Thus, transmission of each frame twice is well-balanced between the ability of recovery and overhead.

In addition, when the number of transmissions is one, the MOS value with the interference traffic differs from that without the interference traffic. However, the difference becomes small when the number of transmissions is three.

In our previous simulation results, transmission twice tends to have high QoE in GCR Unsolicited Retry. Thus, the results in the off-the-shelf devices agree with those in the computer simulation.

VII. Conclusions

In this paper, we evaluated the effectiveness of IEEE 802.11aa GCR Unsolicited Retry on the real environment with the off-the-shelf devices from a QoE point of view. In the experiment, we varied the number of transmissions for each frame in GCR Unsolicited Retry. From the assessment results, we noticed that QoE enhances when the number of transmissions is two. In addition, as the number of transmissions increases more than two, QoE degrades because of congestion owing to multiple transmission.

In future work, we need to assess the effect of burst loss. In addition, we have to assess under various conditions.

Acknowledgment

We thank Dr. Francesco Gringoli at the University of Brescia for his advice on implementation.

References

[1] ITU-T Rec. P.10/G.100, "Vocabulary for performance and quality of service, Amendment 5: New definitions for inclusion in Recommendation", Nov. 2016.

[2] K. Maraslis, P. Chatzimisios and A. Boucouvalas, "IEEE 802.11aa: Improvements on video transmission over wireless LANs," Proc. IEEE ICC 2012, pp. 115-119, June 2012.

[3] IEEE Standard 802.11aa-2012,"Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications - Amendment 2: MAC enhancements for robust audio video streaming," *IEEE802.11 Std*., May 2012.

[4] T. Nunome and T. Komatsu, "QoE enhancement of audio-video reliable groupcast with IEEE 802.11aa," to appear in IEICE Transactions on Communications, vol. E101-B, no. 7, July 2018.

[5] M. -D. Cano and F. Cerdan, "Subjective QoE analysis of VoIP applications in a wireless campus environment," Telecommunication Syst. (2012) 49, pp. 5-15, 2012.

[6] P. Salvador, L. Cominardi, F. Gringoli and P. Serrano, "A first implementation and evaluation of the IEEE 802.11aa group addressed transmission service," ACM SIGCOMM Compt. Comm. Review, vol. 44, no. 1, pp. 36-41, Jan. 2014.

[7] F. Gringoli and L. Nava, "OpenFWWF - Open Firm Ware for WiFi networks," (http://www.ing.unibs.it/open-fwwf/).

[8] P. Salvador, L. Cominardi, F. Gringoli, and P. Serrano, "Implementations details of the IEEE 802.11aa Group Addressed Transmission Service", Technical report, Feb. 2013.

[9] "H.264/AVC JM reference software", (http://iphome.hhi.de/suehring/tml/).

[10] J. P. Guilford, Psychometric Methods, McGraw-Hill, N. Y., 1954.