Non-intrusive QoS Monitoring Method for Realtime Telecommunication Services

Takanori Hayashi†*, Masataka Masuda, Toshiko Tominaga, and Kazuhisa Yamagishi*

Abstract

This article describes a non-intrusive QoS (quality of service) monitoring method for realtime telecommunication services. The key point of this method is that the invalid packet ratio and the invalid frame ratio are introduced to represent the factors affecting the speech and video qualities, respectively. Moreover, we can estimate multimodal quality by considering the individual qualities and their interaction. The results of subjective tests showed that the estimation accuracy of these factors was sufficient for practical use. This method enables us to manage QoS on a call-by-call basis for interconnections among multiple network service providers including a variety of terminals and applications.

1. Introduction

With recent advances in IP (Internet protocol) and multimedia technologies, various realtime telecommunication services such as IP telephony, videophones, and video-streaming services are being developed and provided. It is important to properly evaluate and manage the end-to-end quality of service (QoS) to provide comfortable services because their performances are generally not guaranteed in present IP networks. In the next-generation IP network, moreover, commercial realtime telecommunication services must have an acceptable QoS. The prime criterion for QoS is subjective quality, i.e., the users' perceptions of service quality. This can be measured through subjective quality assessment. However, while subjective quality assessment is the most reliable method, it is time-consuming and expensive. In particular, it is highly impracticable as a method for monitoring/managing quality in real time. Therefore, we need an objective method that is based solely on physical measurement but produces results comparable with those of subjective testing. For managing the end-to-end QoS on a call-by-call basis, passive (non-intrusive) QoS monitoring is more suitable than active (intrusive) QoS monitoring. Non-intrusive methods use either speech/video signal or network performance metrics. We focus on the latter approach because it does not require signal recording equipment and can be implemented easily. This article describes a non-intrusive QoS monitoring method for realtime telecommunication services that can evaluate and manage the end-to-end QoS on a call-by-call basis by passive measurement of network performance. First, we introduce the framework of the non-intrusive QoS monitoring method and explain the quality estimation factors for speech, video, and multimodal qualities. Then, we show the effectiveness of this method through the results of subjective quality assessment tests.

2. Framework of non-intrusive QoS monitoring

The framework of non-intrusive monitoring to estimate and manage QoS for realtime telecommunication services is outlined in **Fig. 1**. Three key techniques are required: intermediate quality parameter estimation, intermediate quality parameter reporting, and end-to-end QoS estimation.

(1) Intermediate quality parameter estimation

It is difficult to estimate end-to-end QoS of realtime telecommunication services by measuring only net-

† NTT Service Integration Laboratories
Musashino-shi, 180-8585 Japan
E-mail: t.hayashi@lab.ntt.co.jp

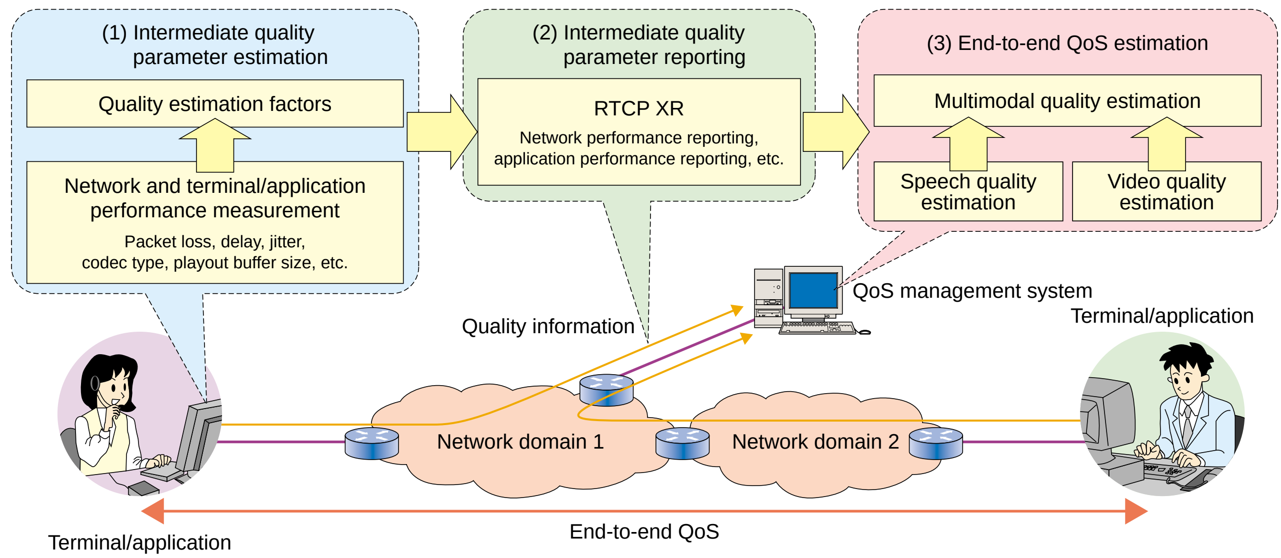

Fig. 1. Framework of non-intrusive QoS monitoring and estimation.

work performance parameters such as IP packet loss and transmission delay because it depends not only on network performance but also on terminal/application implementation such as codec type and jitter buffer size. For example, even if no IP packets are lost in the network, data can still be lost by jitter buffer overflow/underflow due to variations in network delay. On the other hand, packet-loss concealment techniques used at a receiving terminal can ensure that the user does not perceive the quality degradation even if some IP packets are lost in the network. Therefore, it is necessary to estimate intermediate parameters by measuring both network and terminal/application performances in order to estimate the end-to-end user's perceptual QoS at the receiving terminal.

(2) Intermediate quality parameters reporting

For integrated QoS management for realtime telecommunication services, information about the intermediate quality parameters should be collected in a QoS management system. The Internet Engineering Task Force (IETF) defined RTCP XR (realtime protocol control protocol extended report) in RFC3611 [1]. It defines the extended report packet type for RTCP and a container for the information obtained by intermediate quality parameter estimation. This mechanism gives quality information about interconnection calls in environments with multiple network service providers or a variety of terminals/applications. It should be noted that RTCP XR itself does not have functions for estimating quality. A quality estimation metric is not recommended in this RFC but is for further study.

(3) End-to-end QoS estimation

Speech and video qualities are estimated individually using reported intermediate quality parameters. Finally, the quality of realtime telecommunication services should be managed by estimating multimodal quality considering the interaction among the individual speech and video qualities. In Fig. 1, the QoS management system has functions for estimating individual and multimodal qualities, but these functions could be implemented in terminals/applications.

3. Intermediate quality parameter estimation

We use an intermediate quality parameter metric for estimating the end-to-end QoS from network performance metrics [2], [3]. There is a strong correlation between the end-to-end QoS and the total data loss in the network and terminals/applications. Therefore, the key point in properly evaluating users' perceptual QoS is how to measure the data loss at the jitter buffer in a terminal/application. Although the data loss in a terminal/application generally cannot be measured by network managers, our method can estimate it from network performance measurements.

We chose to use the *invalid packet ratio*[*1] as the

*1 Invalid packets: These are packets lost in the network or packets that arrive late at a terminal/application. The number of invalid packets divided by the total number of sent packets is defined as the invalid packet ratio.

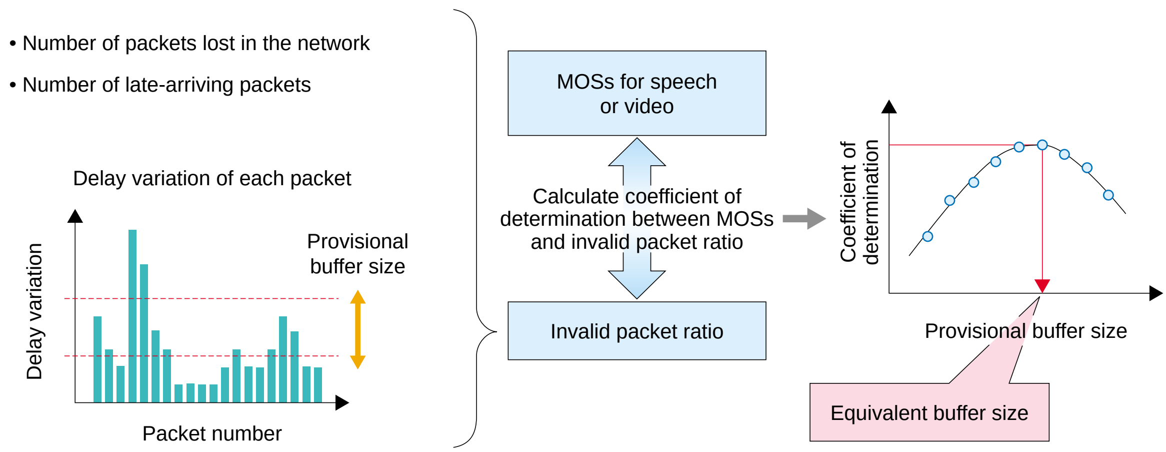

Fig. 2. Method of deriving an equivalent buffer size.

main quality estimation factor. In most cases, subjective quality is represented as a mean opinion score (MOS) calculated by averaging the subjects' scores, which can be estimated from the invalid packet ratio in our method. The end-to-end QoS can thus be estimated by determining the relationship between the MOS and the invalid packet ratio. We define the invalid packet ratio as the ratio of sent packets to the sum of the lost and late-arriving packets[*2], i.e., those that do not enter the terminal's jitter buffer in the required time and are treated as lost data. The invalid packet ratio is calculated as a function of the jitter buffer, packet loss, and delay variation[*3] of packets. The feature of this approach is that these network and terminal performances can be converted into a single metric, which is calculated in the following way. Note that this approach does not take into account the effects of FEC (forward error correction), ARQ (automatic repeat request), etc.

When the size of the terminal's jitter buffer is not known and is not constant, i.e., the buffer is dynamic, we can derive an equivalent jitter buffer size based on the subjective quality evaluation, as illustrated in **Fig. 2**. Invalid packet ratios are calculated for various provisional buffer sizes, and a regression curve of the MOS versus invalid packet ratios is calculated. The buffer size for which the coefficient of determination is maximum is defined as the equivalent buffer size. The regression curve calculated with the equivalent buffer size can then be used as the quality estimation function. If the equivalent buffer size and this regression curve are identified beforehand, then the MOS can be estimated simply by measuring the network performance for the same terminals/applications. Therefore, our method enables efficient network performance management considering the end-to-end QoS because it is not necessary for network managers to perform subjective quality evaluation trials in each network.

*2 Late-arriving packets: The number of these packets is calculated using the delay variation and the jitter buffer size. When the delay variation for each packet is greater than an equivalent jitter buffer size (described later), the packet is defined as a late-arriving packet.

*3 Delay variation: We can use the packet-receiving interval or the delay variation metrics recommended by ITU-T Y.1540 [4] or IETF RFC3393 [5] as the delay variation metrics.

4. End-to-end QoS estimation

A subjective quality experiment was conducted using a point-to-point audiovisual communication service to evaluate the efficiency of our method. The experimental parameters for network performance were packet loss and delay variation.

4.1 Speech quality estimation

The relationship between speech quality and packet loss ratio is shown in **Fig. 3(a)**. **Figure 3(b)** shows the result of converting the packet loss ratio in Fig. 3(a) into an invalid packet ratio derived by the intermediate quality parameter metric. Speech quality was evaluated using a five-grade quality rating scale (5: excellent, 4: good, 3: fair, 2: poor, 1: bad) and represented as a MOS calculated by averaging the subjects' scores. The results in Fig. 3 show that the invalid packet ratio enables the accuracy of subjective speech quality estimation to be dramatically improved compared with using the packet loss ratio.

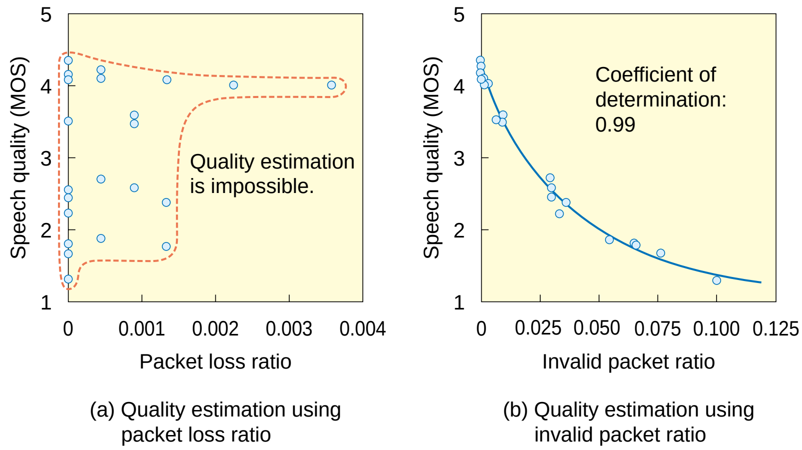

Fig. 3. Speech quality estimation using two quality estimation factors.

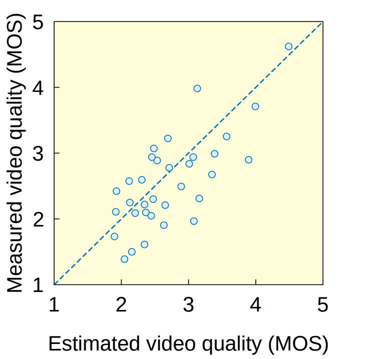

Fig. 4. Video quality estimation using invalid packet ratio.

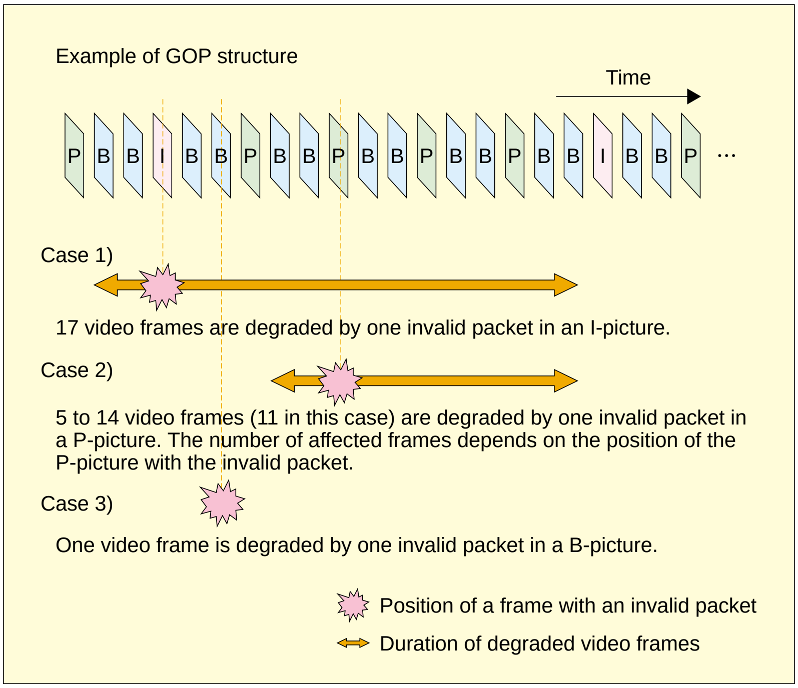

Fig. 5. Examples of how an invalid packet can affect video frames.

4.2 Video quality estimation

The relationship between measured video quality and video quality estimated using the invalid packet ratio is shown in **Fig. 4**. Although the coefficient of determination is sufficient, it is lower than that in Fig. 3(b) because the video coding method used in the subjective test was MPEG2 based on inter-frame coding. There are some cases in which one invalid packet leads to quality degradation in multiple video frames whose duration depends on the type of picture*4 with the invalid packet and on the structure of a group of pictures (GOP), as shown in **Fig. 5**. Therefore, we used a new quality estimation factor, the *invalid frame ratio*. It is calculated as the number of degraded frames divided by the total number of frames. Frames are judged to be degraded from the header information in invalid packets. Quality estimation using the invalid frame ratio is shown in

*4 An MPEG2 video sequence is constructed from three types of pictures. I-pictures (intra-coded pictures) are coded without reference to other pictures. P-pictures (predictive-coded pictures) are coded with a past intra- or predictive-coded picture and are generally used as a reference for further prediction. B-pictures (bidirectionally predictive-coded pictures) are coded with both past and future reference pictures and are not used as references for prediction.

Fig. 6. The quality estimation accuracy in Fig. 6 is much better than that in Fig. 4.

4.3 Multimodal quality estimation

Finally, we consider a model that can accurately evaluate multimodal quality from the individual media qualities and their interaction [6], [7]. The relationship between subjectively measured and objectively estimated multimodal quality is shown in **Fig. 7**. We found that the model's evaluation error (root mean square error) was less than the mean of the 95% confidence interval for the subjective MOS, so the quality evaluation accuracy of this model is sufficient for practical use.

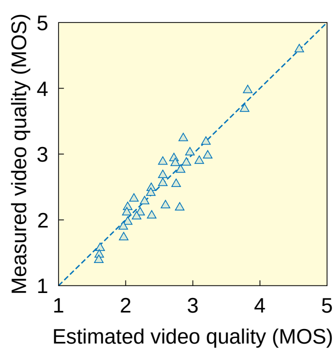

Fig. 6. Video quality estimation using invalid frame ratio.

5. Conclusion

We described a non-intrusive QoS monitoring method for realtime telecommunication services. The invalid packet ratio and the invalid frame ratio are effective intermediate quality parameters for estimating speech and video qualities, respectively. We verified that the method's quality evaluation accuracy was sufficient not only for the individual speech and video qualities but also for multimodal quality.

In the next-generation IP network, proper management of the end-to-end QoS will be an important function in the network. We believe that non-intrusive measurement based on passive network measurement is the most promising way of managing quality on a call-by-call basis.

References

[1] RFC3611, "RTP Control Protocol Extended Reports (RTCP XR)," Nov. 2003.

[2] M. Masuda and K. Ori, "Network performance metrics in estimating the speech quality of VoIP," APSITT 2001, pp. 333-337, Nov. 2001.

[3] M. Masuda and T. Hayashi, "Non-intrusive Quality Monitoring Method of VoIP Speech Based on Network Performance Metrics," IEICE Trans. Commun., Vol. E89-B, No. 2, pp. 304-312, Feb. 2006.

[4] ITU-T Rec. Y.1540, "Internet protocol data communication service—IP packet transfer and availability performance parameters," Dec. 2002.

[5] RFC3393, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)," Nov. 2002.

[6] T. Hayashi and K. Yamagishi, "Multi-Modal Quality Estimation Model based on Interaction of Audio and Video Quality," Proc. The 2004 IEICE General Conference, B-11-12, p. 505, March 2004 (in Japanese).

[7] K. Yamagishi and T. Hayashi, "Opinion Model using Psychological Factors for Interactive Multimodal Services," IEICE Trans. Commun., Vol. E89-B, No. 2, pp. 281-288, Feb. 2006.

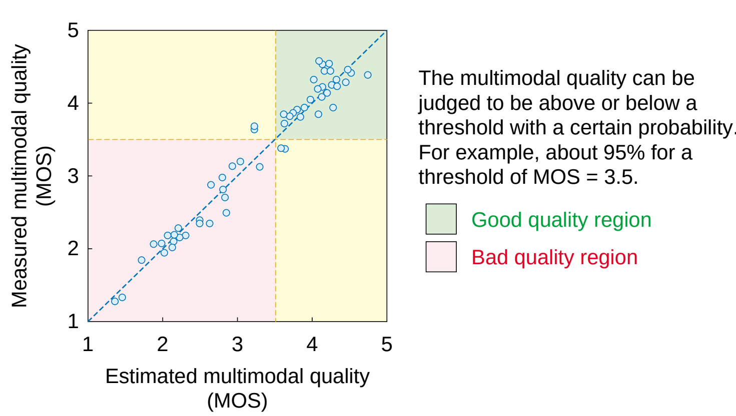

Fig. 7. Estimation of multimodal quality from individual media qualities.

Takanori Hayashi
Senior Research Engineer, Communication Traffic & Service Quality Project, NTT Service Integration Laboratories.

He received the B.E. and M.E. degrees in physical engineering from the University of Tsukuba, Ibaragi in 1988 and 1990, respectively. Since joining NTT Laboratories in 1990, he has been engaged in the subjective quality assessment of multimedia telecommunications and network performance measurement methods. He is currently working on a method of assessing multimodal quality on IP networks. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Toshiko Tominaga
Research Engineer, Communication Traffic & Service Quality Project, NTT Service Integration Laboratories.

She received the B.E. and M.E. degrees in industrial engineering from the University of Electro-Communications, Tokyo in 1987 and 1989, respectively. Since joining NTT Laboratories in 1989, she has been studying quality management methods for telecommunication services. She is currently working on quality management methods for video communications over IP networks. She is a member of IEICE.

Masataka Masuda
Research Engineer, Communication Traffic & Service Quality Project, NTT Service Integration Laboratories.

He received the B.E. and M.E. degrees in electrical engineering from Shibaura Institute of Technology, Tokyo in 1997 and 1999, respectively. Since joining NTT Service Integration Laboratories in 1999, he has been working on performance evaluation and management in IP networks and speech quality evaluation of VoIP. He is a member of IEICE.

Kazuhisa Yamagishi
Research Engineer, Communication Traffic & Service Quality Project, NTT Service Integration Laboratories.

He received the B.E. degree in electrical engineering from Tokyo University of Science, Chiba in 2001 and the M.E. degree in electronics, information and communication engineering from Waseda University, Tokyo in 2003. Since joining NTT Service Integration Laboratories in 2003, he has been engaged in subjective quality assessment of multimedia telecommunications and image coding. Currently, he is working on the quality assessment of interactive multimodal services over IP networks. He is a member of IEICE.