INVITED PAPER *Special Issue on the Past, Present, and Future of Communications Technologies in the IEICE*

Collaborative Quality Framework: QoE-Centric Service Operation in Collaboration with Users, Service Providers, and Network Operators

Akira TAKAHASHI[†a)], ***Senior Member and*** **Takanori HAYASHI**[†], ***Member***

SUMMARY We propose a framework called "QoE-centric Service Operation," with which we attempt to implement a means to enable the collaboration of end-users, service providers, and network providers to achieve better QoE of telecommunication services. First, we give an overview of the transition in the quality factors of voice, video, and web-browsing applications. Then, taking into account the fact that many quality factors exist not only in networks, but also in servers and terminals, we discuss how to measure, assess, analyze, and control QoE and the technical requirements in each component. We also propose approaches to meet these requirements: packet- and KPI-based QoE estimation, compensation of sparse measurement, and quality prediction based on human behavior and traffic estimation. Finally, we explain the results of our proof-of-concept study using an actual video delivery service in Japan.

key words: *QoE, operation, quality, QoS*

1. Introduction

The business environment of telecommunications services has changed drastically not only in Japan but worldwide in the 1980s. In the United States, AT&T was broken up in 1982. In Japan, NTT was privatized in 1985 and later broken up into local, long-distance, and mobile companies. After such restructuring, users can choose their preferred terminals, such as personal computers, as well as telephone sets, and connect them to networks. Many service providers have ventured into telecommunications business, offering various attractive applications by using carrier networks.

Telecommunication carriers have been making every effort to improve user quality of experience (QoE) [1] since the telephone era, which forms an important part of user satisfaction for telecommunication services. However, the above-mentioned situation makes it difficult for telecommunication carriers to have full control of QoE. Thus, improving QoE requires collaboration with other players in the ecosystem.

Another aspect that has recently become important to improve QoE is "flexible quality management." Radio access and software defined network/network function virtualization (SDN/NFV) are the key network technologies used in current and future telecommunication services, and one of the common features of these technologies is "time-space variability" of network performance. Therefore, dynamic quality management is more important than pre-service quality planning of networks.

From these viewpoints, we propose a framework in which all the players in the ecosystem, i.e., end users, service providers and network providers, cooperate and collaborate to improve the QoE of telecommunication services. We call this framework "*QoE-centric Service Operation*."

In this paper, we first review the history of QoE studies in telecommunications in terms of voice, video, and web-browsing applications to help in understanding the fundamentals behind the discussion in the following sections. We start discussion by defining the requirements on the technologies composing the proposed framework. Then, we discuss possible approaches to meet them. Finally, we introduce our state-of-the-art investigation on the proposed framework, applying it to one of the most popular video delivery services in Japan.

2. QoE Dimensions

2.1 Voice Communications

The study of telecommunication quality began by focusing on designing and maintaining telephone services. First, the issues were transmission loss, frequency distortion, circuit noise, and circuit echo because all the transmission equipment processed the speech signals in an analogue manner. As a result, the primary quality assessment dimensions were loudness, due to transmission loss, and intelligibility/annoyance due to distortion, noise, and circuit echo [2]. In addition to such speech transmission quality, the availability and connectivity of telephone services were independently studied because transmission quality was considered only if the service was available and the telephone connection was established.

By the digitization of transmission equipment and links, most of the above analogue degradation disappeared. The primary quality factors in telephone services were quantization distortion due to PCM coding, delay, and circuit echo. At this stage, the quality dimension was usually "overall satisfaction," rather than loudness and intelligibility/annoyance of speech because telephone quality was high enough, especially in Japan.

In the 1990's, cellular phone systems and the personal

Manuscript received October 14, 2016.
Manuscript revised January 26, 2017.
Manuscript publicized March 22, 2017.
[†]The authors are with NTT, Musashino-shi, 180-8585 Japan.
a) E-mail: takahashi.akira@lab.ntt.co.jp
DOI: 10.1587/transcom.2016PFI0013

handy-phone system (PHS) penetrated the telephone market. In these systems, low-bitrate codecs were used to make effective use of radio resources, introducing non-negligible speech distortion. Due to the instability of radio links, speech quality degradation due to transmission error also became important. At the same time, due to the nature of radio communications, it became difficult to separately discuss the three quality dimensions of availability, connectivity, and transmission quality when evaluating the QoE of cellular services.

In the early 2000's, the PSTN service started migrating to Internet Protocol (IP)-based networks. This is called "IP telephony," in which delay due to playout buffering is inevitable in addition to packet loss in both networks and terminal playout buffer. At this point, the evaluation of non-stationary or discrete distortion events became important, even in wired networks [3].

One of the most recent public telephony services is Voice over LTE (VoLTE), which expands the signal bandwidth of speech from 3.4 to 7 kHz, resulting in more natural voice quality. The quality factors of VoLTE have mixed characteristics of traditional cellular phones and IP telephony since it is based on VoIP technology over radio access networks.

2.2 Web-Browsing

Web-browsing is one of the most popular applications of recent telecommunications and often used as a means to access public services as well as on-line shopping.

At the beginning of web-browsing in the 1990's, due to the limitations on network and terminal performance, the content was not so rich; therefore, users' expectations on QoE were not very high. It is often said that the objectives of the page load time, which represents the time used to show the next content after a user requires (or clicks) the original page, was about eight seconds during this time [4]. Recently, as network speed and terminal capability increased, users' expectations seem to have increased to a couple of seconds.

Since the structure and mechanism of traditional web content are rather simple, the QoE of web-browsing simply depends on the average throughput of the network to download the html and media files. However, web technologies, such as JAVA, have dramatically evolved in this century, and the characteristics of web content changed and became complex. Due to such changes, it is no longer a simple file download, but combinations of program download/execution and still image and video download/playout. This makes the relationship between network performance and QoE of web-browsing much more complicated [5].

2.3 Video Streaming

With the advances in broadband network and multimedia technologies such as encoders and decoders (codecs), various video services such as videophones, videoconferencing, video on demand (VoD), and IPTV have been developed and provided. Recently, video streaming service providers can deliver ultra-high definition (UHD) video content over IP networks.

In analog video, such as conventional TV broadcasts, quality degradation occurs in a continuous fashion with gradual degradation in quality. However, in digital video, quality degradation is both spatially and temporally discontinuous and characterized by a very large amount of degradation occurring once the digital signal is disrupted, interrupted, or stopped. The factors that cause degradation in the quality of video services can be divided into spatial distortion and temporal distortion. Spatial distortion causes the picture quality and/or screen resolution to decrease, and its typical/characteristic effect on video quality is causing mosaic distortion (block distortion) of the video. Temporal distortion typically manifests as a reduction in frame rate or a frozen picture, resulting in a jerky picture and a loss of smoothness in moving objects. There are also spatio-temporal distortion effects including disruption in the video signal [6], [7].

Recent video streaming services can be classified as Real-Time Protocol (RTP)-based (i.e., linear TV) or Hypertext Transfer Protocol (HTTP)-based streaming (i.e., adaptive bitrate streaming) [8]. In RTP-based video streaming services, typical quality degradation factors are block distortion due to video compression and disruption of the video signal due to network performance degradation such as IP packet loss. On the other hand, in HTTP-based video streaming services, typical quality degradation factors are quality level change and video stalling due to network throughput fluctuation. Thus, perceived video quality factors have changed quite a bit within a couple of decades.

3. QoE-Centric Service Operation and Its Requirements

Various recent QoE investigations have revealed that the relationship "QoE $\approx$ network performance" no longer holds [9]. What makes the situation more complicated is that even network operation has become difficult in an end-to-end sense due to the multi-provider environment. Thus, there are quite a few issues that cannot be found and solved only through conventional network operation. Figure 1 explains the quality factors in video streaming services.

Even if the network performance is the same, the resultant QoE heavily depends on what application an end user uses. For example, 500 kb/s is sufficient for VoIP users but not for users enjoying high-resolution video streaming. The QoE also depends on server congestion, which does not necessarily correspond to network congestion, for client-server-type applications. By taking into account such factors affecting QoE, our QoE-centric Service Operation attempts to harmonize the three main players in the service chain; end-users, service providers, and network providers.

Figure 2 illustrates the main concept of QoE-centric Service Operation, which consists of three parts. The first part is network operation, which is similar to conventional

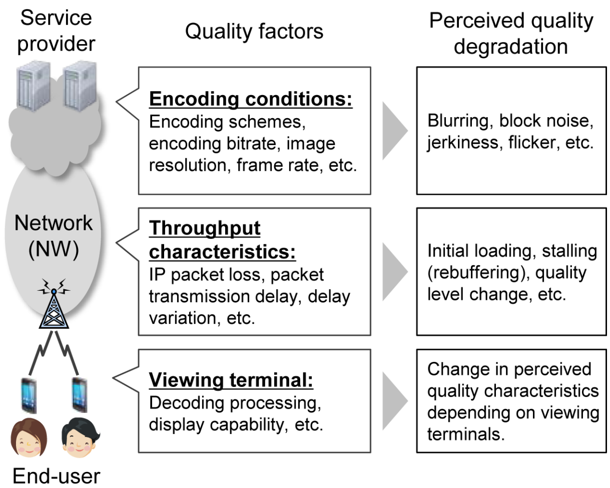

Fig. 1 Quality factors for video streaming services.

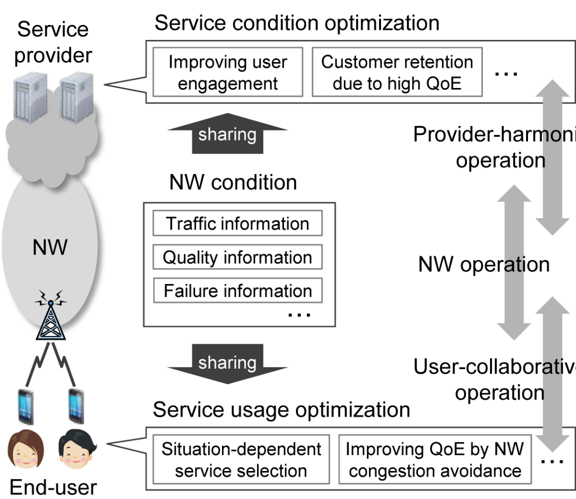

Fig. 2 QoE-centric service operation.

network operation but incorporates QoE-centric functions. The second part is what we call provider-harmonic operation, in which service and network providers share the information related to QoE to improve it in a collaborative manner. The third part is user-collaborative operation, in which a network provider provides end-users with the information that can be used to optimize their behavior in terms of QoE. Figure 3 summarizes the functions that are essential to achieve QoE-centric Service Operation. The following sections explain the requirements on these functions.

3.1 QoE Assessment

First, we need to have a means to quantify the goodness or poorness of QoE. The International Telecommunication Union (ITU) has standardized many methodologies for subjectively measuring the perceptual quality of voice, audio, video, and other applications such as web-browsing [10]–[13]. However, subjective assessment methodologies require a psycho-physical experiment, which is not applicable

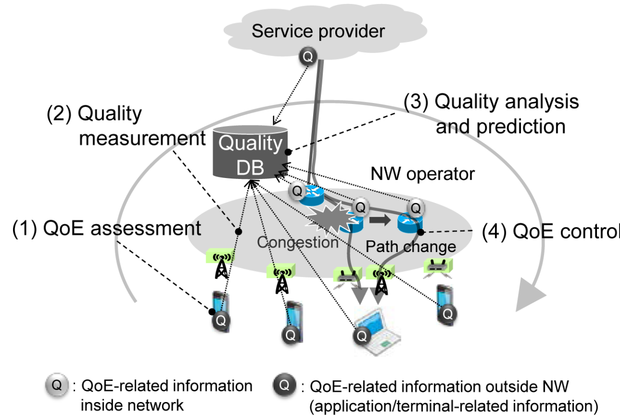

Fig. 3 Functions of QoE-centric service operation.

to in-service operation. Therefore, to take the QoE in actual services into account, objective quality-assessment methodologies for estimating subjective quality from physical and measureable parameters are indispensable.

3.2 Quality Measurement

To obtain the data used in objective quality assessment, the mechanism by which the input data used in the quality-estimation model can be recorded, measured, or extracted is necessary. Traditionally, we used to use network-performance measurement devices to monitor fundamental Key Performance Indicators (KPIs) such as packet-loss rate and round-trip delay. Recently, the deep packet inspection (DPI) scheme, which enables a detailed look at packet information, has expanded the variety and scalability of KPI measurement inside a network [14].

On the other hand, the more the quality factors exist outside networks, the more the importance of collecting data at servers and/or terminals increases. From this viewpoint, the crowd-sourcing scheme [15], in which one requests end-users to conduct measurement, for example, by application software implemented in smartphones and to provide the measurement results, is of great interest.

Although the DPI and crowd-sourcing schemes drastically enhance the scale of measurement, the obtained data are not dense enough to evaluate the entire service areas. Therefore, we need a means to compensate for the measurement data, which is not simple due to the variability of radio transmission.

3.3 Quality Analysis and Prediction

The QoE assessment based on the collected measurement data is followed by visualizing the current QoE achieved by the network and identifying the problems. To solve the identified problems, cause analysis of QoE degradation is also important.

The effectiveness of the DPI and crowd-sourcing measurement schemes introduced in the previous section depends on the correlation between measured KPIs and QoE, and the "cleanness" of the obtained data, respectively.

Although we can expect accurate results in the DPI measurement, the issue is what to measure. Since the measurement is conducted inside the network, it is often difficult to analyze the payload information due to encryption. This may result in observing more generic KPIs, e.g., the IP packet-loss rate and round trip time, which do not necessarily represent the resultant QoE. Therefore, establishing well-balanced KPIs in terms of both their availability and correlation with QoE is required.

In the crowd-sourcing measurement, using user devices results in certain "noise" which is not negligible. Therefore, establishing analysis methodologies for extracting the essential features from such noisy data is also needed.

In addition to the analysis of the current QoE, prediction in the time domain is also important for effective quality control, which is introduced in Sect. 3.4. This is because, in time varying circumstances, traffic control or user navigation based only on the current observation will result in poor QoE improvement.

3.4 Quality Control

Finally, we need quality control methodologies for taking into account the analysis of the current QoE distribution and its prediction in the time domain. Quality control includes traffic and resource control making use of the SDN/NFV functions. In addition to these network control mechanisms, we believe quality control should also be done in collaboration with service providers and end-users since, especially in mobile networks, the network resources are limited and it is sometimes difficult to allocate more resources to improve QoE.

3.4.1 Provider-Harmonic Operation

Allocating more resources to a certain service or user may improve QoE. However, in a very competitive environment, such as at a crowded train station, it is not always possible. In such a case, optimization between service and network providers plays an important role.

We consider the QoE of progressive download video. Figure 4 illustrates an example of QoE characteristics for progressive download video streaming services. The horizontal axis shows the video encoding bitrate or delivery rate, and the vertical axis shows the video quality that end-users perceive. It is easy to understand that the higher the encoding video bitrate is, the richer the video quality is when the available throughput is greater than the encoding bitrate (the gray solid line in Fig. 4). If the available network throughput is 500 kb/s (the dashed line in Fig. 4), however, video encoded at 1 Mb/s cannot be played out properly and stalls due to the underflow of the playout buffer. Apparently, we can avoid such stalling by reducing the encoding bitrate to below 500 kb/s. That is, there is an optimal bitrate in terms of QoE for a given network quality.

Therefore, if a network provider informs a service provider of the prediction of network quality, the service provider can optimize their customers' QoE without requiring any additional resources from the network provider. From the network provider's viewpoint, they can offer better QoE to their end users without any additional investment to enhance network capacity. This is win-win relationship for both players.

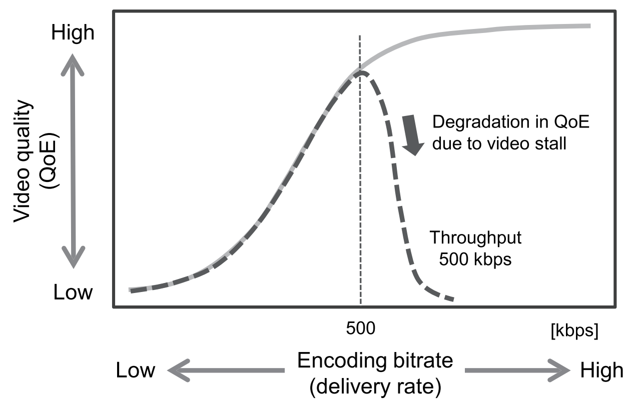

Fig. 4 QoE characteristics for video streaming services.

3.4.2 User-Collaborative Operation

In provider-harmonic operation, we try to optimize the QoE of an end user at a certain location and time. However, in a very crowded situation, the network performance, e.g., throughput, is not high enough, and the optimized choice of bitrate at the server cannot result in sufficient QoE.

Those users who inevitably need to use the network at that location at that time must accept that QoE. However, if they accept changing the location or waiting for some time, they may acquire better QoE since the network performance may improve.

The quality of conventional telecommunication services is something that is given by a network provider. On the other hand, our framework provides end users with a means to obtain higher quality. Again, this scheme does not require any additional investment to either network/service providers or end users for improving QoE.

This paradigm shift is similar to the relationship between public transportation and private vehicles. For those who use public transportation, the most important thing is that the transportation company adheres to a timetable. On the other hand, drivers need information such as road and parking lot congestion and their prediction. Since ICT literacy is drastically increasing, we need to meet users' expectation in the same manner.

To this end, users need to know the QoE level expected at a certain location in the near future, e.g., 10 minutes later. This is one of the motivations to study the quality prediction pointed out in Sect. 3.3.

4. Technical Approaches

This section discusses the technical approaches to meet the

Table 1 Objective quality assessment models.

Category	Media/bitstream-layer model	Packet-layer model	KPI-layer model
Input information	Media signal, bitstream data	Packet header information	KPI values
Primary application	Quality benchmarking, In-service non-intrusive monitoring	In-service nonintrusive monitoring	
Existing standards and ongoing projects in ITU			
Speech	ITU-T P.862, P.863	ITU-T P.564	ITU-T G.107, G.107.1
Audio	ITU-R BS1387	ITU-T P.1201, P.NATS	ITU-T G.1070 [Videophone], G.1071 [IPTV]
Video	ITU-T J.144 [SDTV], J.247 [PC], J.341 [HDTV] etc.		
Multimedia	ITU-T J.148		

following requirements for implementing QoE-centric Service Operation:
A) Objectively estimating subjective quality (QoE),
B) Compensating for sparse measurement data,
C) Defining KPIs corresponding well to QoE,
D) Predicting QoE in the time domain.

4.1 QoE Estimation

As pointed out in Sect. 3.1, QoE estimation models for objective quality assessment are key to quantitatively evaluating QoE. Such models fall into one of the following three categories: (1) media/bitstream-layer models, (2) packet-layer models, or (3) KPI-layer models (Table 1).

These models use media signals (e.g., speech waveform data, and video pixel data) or encoded bitstream (e.g., MPEG Elementary Stream), packet-header information (e.g., MPEG Transport Packet Header), and KPIs (e.g., codec-type, packet-loss rate, and delay), respectively† They are selectively used depending on what information is available in each operation scenario. For example, when monitoring the encoding quality at the headend of video streaming services, one uses a media-layer model, which provides the most precise estimation of audiovisual perceptual quality.

For the two important schemes of DPI and crowd-sourcing measurement introduced in Sect. 3.2, packet-layer and KPI-layer models are suitable.

4.1.1 KPI-Layer Model

Several KPI-layer models for estimating QoE from KPIs of application-setting parameters and network performance parameters had been standardized as ITU-T recommendations. The ITU-T recommendations G.107 [16], G.1070 [17], and G.1071 [18] provide KPI-layer models for IP telephony, videophone, and IPTV, respectively. For example, in ITU-T recommendation G.1071, application-setting parameters, such as codec, coded bitrate, video size, and video frame rate, and network performance parameters, such as

†Combination of such models is also possible, and called a "hybrid model."

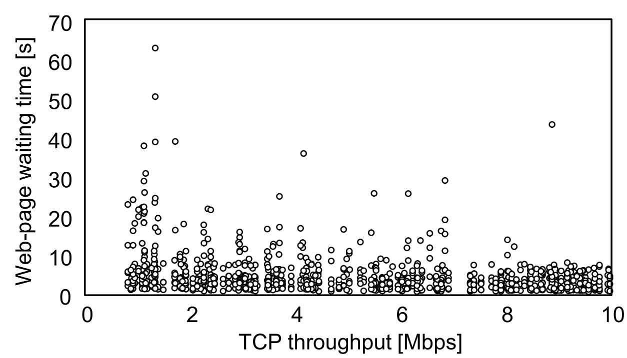

Fig. 5 Relationship between download throughput and web-page waiting time.

packet loss rate, are input to the KPI-layer model to estimate QoE for IPTV services. These KPIs can be measured using DPI or crowd-sourcing. Because time fluctuation of network performance is not considered, KPI-layer models estimate QoE from the viewpoint of the averaged quality characteristics.

There is no KPI-layer model for Web browsing because it is difficult to correlate web-page waiting time with network performance parameters such as Transmission Control Protocol (TCP) throughput. Figure 5 shows an example of the relationship between TCP throughput and web-page waiting time for a specific web-page. This result shows that TCP throughput does not represent the QoE of web-browsing. Since web content often consists of many objects with small volume, web-page waiting time does not match TCP throughput derived as time when a large-volume file was transferred.

4.1.2 Packet-Layer Model

The KPI-layer models estimate QoE based on the averaged quality characteristics. On the other hand, packet-layer models estimate QoE for each user by taking into account the time fluctuation of network performance. For example, ITU-T recommendation P.1201 [19] provides packet-layer models for IPTV services. By measuring the behavior of packets using packet header information, the effect of QoE degra-

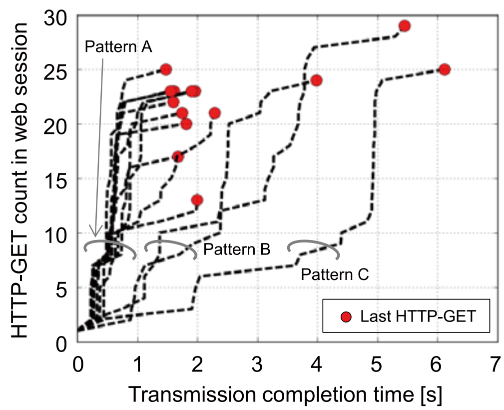

Fig. 6 Relationship between HTTP-GET count and transmission completion time for each object.

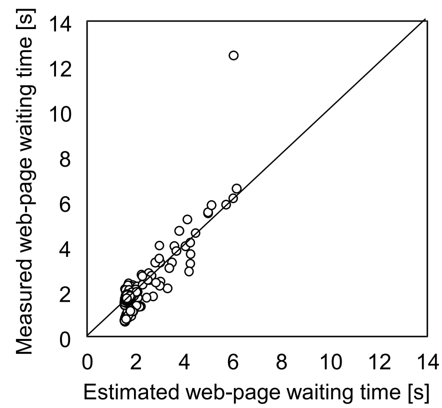

Fig. 7 Estimation accuracy of web-page waiting time.

dation due to packet losses can be considered. Currently, ITU-T Study Group (SG) 12 is discussing the standardization of a packet-layer model for progressive download video streaming services and plans to release a recommendation in 2017. This recommendation is provisionally called "Parametric non-intrusive assessment of TCP-based multimedia streaming quality, considering adaptive streaming" (P.NATS).

It may be possible to estimate QoE for web browsing by adopting the packet-layer-model approach. Figure 6 shows the relationship between the number of transmitted objects (HTTP-GET count) and transmission completion time for each object comprising a specific web page. Even if a web page's address is the same, the number of objects sometimes changes when content is renewed. For example, popular Japanese web-pages such as Yahoo!, Amazon, Rakuten, and goo get renewed in a short span of time, so a unified model is needed in order to estimate the web-page waiting time even when the number of objects is changed in a small range. To do this, we first get the approximate number of objects from a specific web-page from the data of both the number of transmitted objects and real web-page waiting time in a user terminal. We then estimate the web-page waiting time from the features in the beginning part of the transmission patterns by using a support vector machine (SVM) [20]. The relationship between the web-page waiting time that can be measured in a user terminal and the features in the beginning part of the transmission patterns of objects is learned beforehand. Figure 7 shows the estimation accuracy of this approach. We found that this approach had sufficient accuracy for almost all plots. However, the estimation error of one plot was large because the abrupt changes of the feature of the transmission patterns cannot be taken into account with this approach. This is for further study.

4.2 Measurement and Compensation

Exploiting schemes such as DPI and crowd-sourcing is effective to visualize the QoE characteristics in geographical areas, which is important, for example, in the area quality management in mobile access networks. However, to geographically cover all the area is not so straightforward because the radio transmission characteristics in a real environment is not monotonically distributed, and simple compensation between multiple measurement locations does not provide a correct estimate. On the other hand, simulation technologies of radio propagation characteristics have become sufficiently mature [21]. Therefore, a combination of actual measurement and such simulation can provide estimates of KPIs in a dense manner, even if the measurement data are geographically sparse.

Figure 8 illustrates this idea. It is usually difficult to estimate the continuous quality map of upper-layer KPIs (e.g., TCP throughput) because the correlation of the adjacent locations is not very high, depending on the radio transmission quality, congestion in the mobile front haul, and so on. That is, the KPI value of a certain location does not necessarily represent that of neighboring locations.

However, we may be able to assume that a KPI value, such as TCP throughput, is consistent if two locations are in the area of the same beam of a base station (i.e., the same congestion condition of upper networks) and have the same radio condition. If this assumption holds, we can estimate the KPI values in the entire region (in the same beam range) based on the actual measurement data of the radio condition and KPI value in conjunction with the radio propagation simulation. More concretely, if we have a measured radio condition such as Reference Signal Received Power (RSRP) and the associated TCP throughput, we can expect the same TCP throughput in other locations where we obtain equivalent RSRP through the radio propagation simulation.

To confirm the validity of estimating KPIs based on the radio condition, we measured the RSRP and TCP throughput in a real LTE environment. Ideally, it is better to measure all the data at the same time to exclude the effect of the traffic condition, which may change the relationship between the RSRP and TCP throughput. Due to the limitation of measurement devices, the data were obtained sequentially in our experiment. Since we carried out our experiment in a quiet residential area, the relationship between the RSRP and TCP throughput was expected to be preserved during the

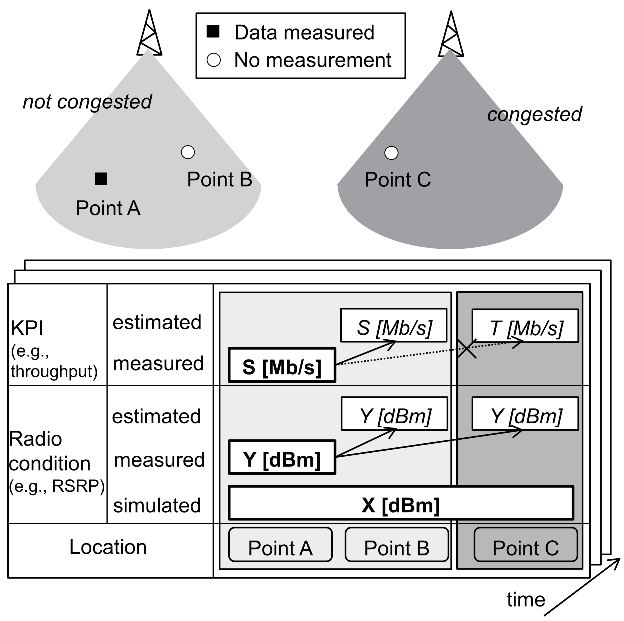

Fig. 8 Estimation of KPIs based on sparsely measured data and radio propagation simulation.

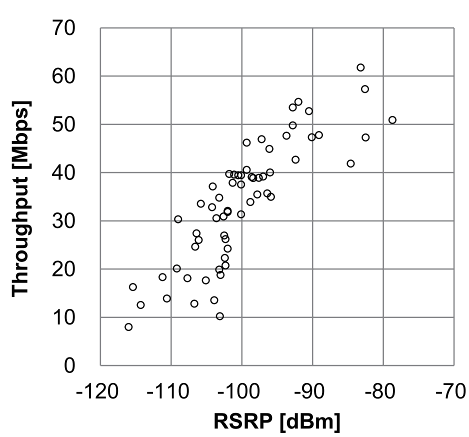

Fig. 9 Relationship between RSRP and TCP Throughput.

measurement time frame. Figure 9 illustrates this relationship between the RSRP and TCP throughput, which were actually measured at the same time using a smartphone. The measurement was carried out within the same beam area with a fixed transmission radio band. We observed good consistency between them.

This result implies that we may be able to estimate the KPIs that represent QoE, based on the actual measurement values of radio transmission performance and associated KPIs, which can be collected using the crowd-sourcing scheme, and the simulation of the radio transmission performance for the entire region of interest. This can result in drawing a QoE map of individual services by mapping the KPIs to QoE based on respective mapping techniques.

It has been reported that the radio propagation simulation provides very close estimates of the actual radio transmission measurement results, such as received field strength

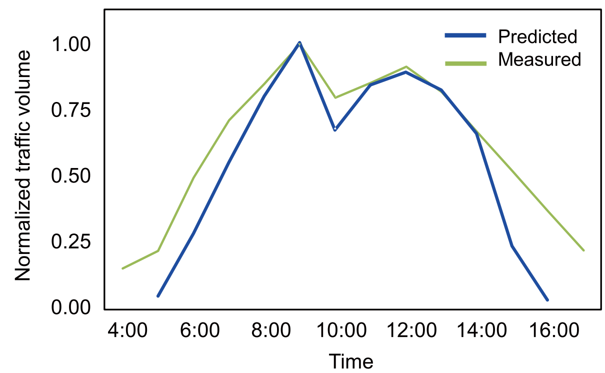

Fig. 10 Time variation of predicted and measured traffic.

[22]. However, if we assume crowd-sourcing data collection, we need to investigate the stability of radio transmission measurement, such as RSRP, when measured under realistic usage conditions of smartphones. This is for further study.

4.3 Quality Prediction

To avoid quality deterioration, mobile network operators need to estimate the appropriate communication facilities. It is challenging to efficiently deploy communication facilities, such as Wi-Fi access points and mobile base stations, to transfer the huge mobile traffic during a large-scale event. Therefore, we assume that human behavior during an event can be categorized into typical behavior patterns determined by the event content, and propose an approach for predicting the mobile traffic demand of an event venue through a multi-agent simulation given the typical behavior patterns of mobile users during an event. The behavior rules used for the simulation are first constructed from the data of past events that are similar to the predicted events. We also evaluated the reproducibility of group behavior during an indoor event through a simulation using individual behavior rules. In the simulation, user-behavior models and environmental parameters are used as inputs. User-behavior models consist of movement rule and traffic occurrence condition, which are based on each user's purpose. Environmental parameters consist of the number of users, range of user movement, and exhibition locations where the user can achieve his/her purpose. Moreover, the limit of the number of users in the simulation is determined from the movement range and exhibition locations due to reconstructing of a crowd or a waiting line.

An example of predicted results and measurement values of mobile traffic at a particular indoor event are shown in Figure 10. To compare the occurrence times of the traffic peaks, the measured and predicted mobile traffic volume are normalized by the maximum value. This shows that our approach can be used to predict peak hours of event traffic. Prediction of traffic peaks at fine spatial granularity and improving prediction accuracy regarding the absolute value of traffic volume are currently being investigated.

5. Proof of Concept

5.1 Overview

We give an example of QoE-centric Service Operation targeting a progressive-download video streaming service.

As shown in Fig. 1, the important factors affecting QoE in video streaming services exist both inside and outside the carrier network. Therefore, even when higher video quality is achieved with a higher encoding bitrate, video playback may frequently stall in a congested mobile network due to insufficient throughput. Such video stalling not only degrades the QoE but also greatly affects the length of video watching time i.e., video viewing behavior [23]. Therefore, QoE control carried out by the network operator and video streaming service provider in a collaborative manner is effective in optimizing QoE. An effective response to such a situation would be to deliver the video at a coding bit rate for which QoE could be optimized by taking network congestion into account. We define the interface between the network operator and video streaming service provider as the "quality application programming interface (API)" for exchanging quality-related information.

For video streaming service providers and end users, this form of control improves QoE, while for the network operator, it can reduce traffic that does not contribute to improving QoE while reducing network load (i.e., a reduction in the required facilities). In short, controlling the network in this manner can result in a win-win relationship.

5.2 Quality API

Our developed quality API exchanges quality-related information to optimize QoE. The quality API provides recommended bit-rates to video streaming service providers considering network congestion. Figure 11 shows the sequence diagram that explains how videos are played back using the quality API.

When a user sends a request to a video service provider to start watching a video, the user's attribute information including his/her location and network environment (e.g., LTE, 3G, or WiFi) is sent along with the request (step 1). The video streaming service provider then sends the information of the user's attributes and the list of encoding conditions (e.g., encoding bit-rate, resolution, and frame rate) for each of the several available video quality indicators to the quality API (step 2). To obtain the optimal encoding conditions for each video request, the quality database (DB) derives the encoding conditions with the highest QoE in consideration of predicted network quality (step 3). Then the optimum encoding conditions are sent back to the video streaming service provider in response to its request (step 4). According to this recommendation, the video streaming service provider distributes the video to the user (step 5). After the user has finished watching the video (step 6), information on network throughput is sent to the quality API for future reference (step 7). This allows the latest network quality information to be continually updated and high-precision estimates to be made (step 8).

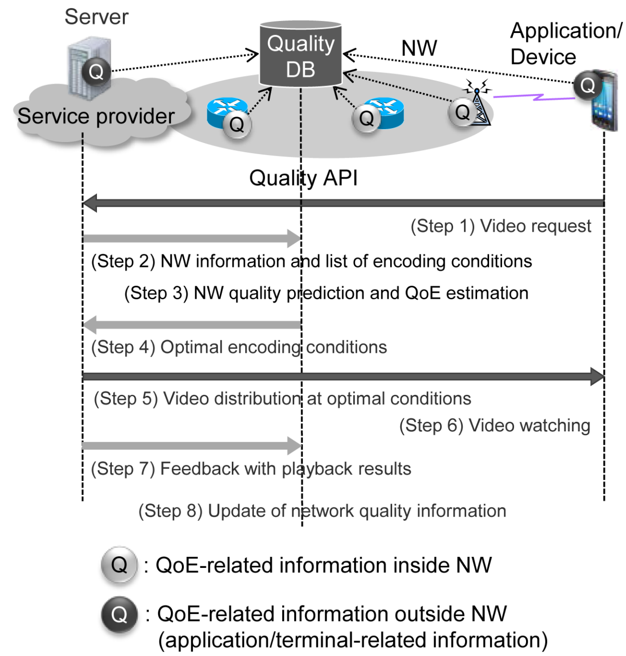

Fig. 11 Video streaming using quality API.

5.3 Results and Discussion

Joint experiments were conducted with Dwango, the provider of the Niconico video streaming service, to test the effectiveness of the quality API.

The experiments were conducted to determine to what degree bit rates optimized based on recommendations were able to improve QoE. In the experiments, video playback information was sent to our quality API for users of Android devices on the NTT DOCOMO LTE network on July 3 and 4, 2014. The total number of playback times was about three-hundred thousand. The results indicate that we can expect the quality API to significantly increase the QoE for viewers of Niconico videos. Specifically, we found:

(A) Lower incidence of video stalling

We confirmed that the quality API was able to reduce the number of users affected by video stalling from 33% to just 2% during peak hours, as shown in Fig. 12.

(B) Improved QoE

Figure 13 shows the QoE improvement during peak hours for 1,900 videos selected at random. The horizontal axis shows the amount of QoE improvement due to the quality API, and the vertical axis shows the cumulative probability density. As shown in Fig. 13, QoE improved 0.6 on average for 40% of the sample videos on a 5-grade quality scale defined by ITU-T Rec. P.910 [11]. On the other hand, QoE degraded 0.35 on average for 10% of the sample videos. Thus, the overall QoE improved about 0.2 on average by using quality API.

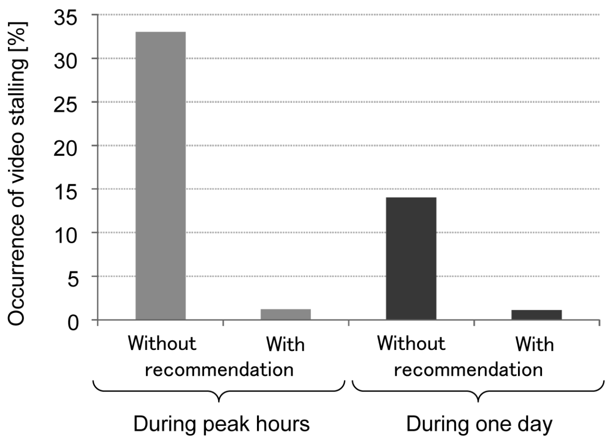

Fig. 12 Efficiency of quality API.

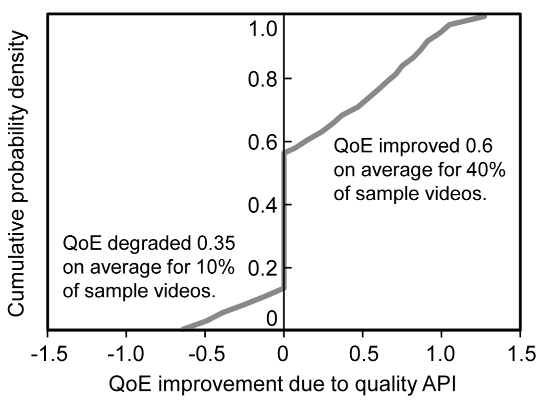

Fig. 13 QoE improvement due to quality API.

(C) Lower overall volume of transmitted data

A secondary effect of optimizing bit-rates with this technology was that we can confirm a decrease in transmitted data volume by about 17%.

6. Conclusion

We proposed a framework called "QoE-centric Service Operation," which accelerates the collaboration of end-users, service providers, and network providers to achieve better QoE of telecommunications services. We discussed the necessary functions of this framework and their requirements. Based on these considerations, we proposed possible approaches to meet the requirements and provided evidence of the validity of these approaches. We argue that this paper will evoke further studies in this area, including new approaches to improve QoE based on the proposed framework.

References

[1] ITU-T Recommendation P.10/G.100, "Vocabulary for performance and quality of service, Amendment 5," July 2016.
[2] S. Moeller, Assessment and Prediction of Speech Quality in Telecommunications, Kluwer Academic Publishers, 2000.
[3] A. Raake, Speech Quality of VoIP: Assessment and Prediction, John Wiley & Sons, New York, 2006.
[4] Zona Research, Inc., "The economic impacts of unacceptable web-site download speeds," http://www.webperf.net/info/wp_downloadspeed.pdf, April 1999, accessed Oct. 11. 2016.
[5] H. Yamamoto, S. Nakamura, H. Honda, and A. Takahashi, "Performance estimation techniques for browser-based applications," NTT Technical Review, vol.11, no.5, May 2013.
[6] S. Winkler, Digital Video Quality: Vision Models and Metrics, Wiley, Hoboken, 2005.
[7] A. Punchihewa, Video Compression, InTech, Rijeka, 2012.
[8] D. Minoli, Linear and Non-Linear Video and TV Applications: Using IPv6 and IPv6 Multicast, Wiley, Hoboken, 2012.
[9] A. Takahashi and T. Hayashi, "Toward QoE-centric operation of telecommunication services," NTT Technical Review, vol.11, no.5, May 2013.
[10] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," Aug. 1996.
[11] ITU-T Recommendation P.910, "Subjective video quality assessment methods for multimedia applications," April 2008.
[12] ITU-R Recommendation BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," Jan. 2012.
[13] ITU-T Recommendation P.1501, "Subjective testing methodology for web browsing," Feb. 2014.
[14] S. Kumar, S. Dharmapurikar, F. Yu, P. Crowley, and J. Turner, "Algorithms to accelerate multiple regular expressions matching for deep packet inspection," SIGCOMM'06 Proc. 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp.339–350, Sept. 2006.
[15] S. Rosen, S. Lee, J. Lee, P. Congdon, Z. Mao, and K. Burden, "MCNet: Crowdsourcing wireless performance measurements through the eyes of mobile devices," IEEE Commun. Mag., vol.52, no.10, pp.86–91, Oct. 2014.
[16] ITU-T Recommendation G.107, "The E-model: A computational model for use in transmission planning," June 2015.
[17] ITU-T Recommendation G.1070, "Opinion model for videotelephony applications," July 2012.
[18] ITU-T Recommendation G.1071, "Opinion model for network planning of video and audio streaming applications," June 2015.
[19] ITU-T Recommendation P.1201, "Parametric non-intrusive assessment of audiovisual media streaming quality," Oct. 2012.
[20] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol.20, no.3, pp.273–297, Sept. 1995.
[21] K. Ng, E. Tameh, A. Doufexi, M. Hunukumbure, and A. Nix, "Efficient multielement ray tracing with site-specific comparisons using measured MIMO channel data," IEEE Trans. Veh. Technol., vol.56, no.3, pp.1019–1032, May 2007.
[22] E.M. Vitucci, V. Degli-Esposti, F. Fuschini, J.S. Lu, M. Barbiroli, J.N. Wu, M. Zoli, J.J. Zhu, and H.L. Bertoni, "Ray tracing RF field prediction: An unforgiving validation," International Journal of Antennas and Propagation, vol.2015, Article ID 184608, 2015.
[23] F. Dobrian, V. Sekar, A. Awan, I. Stoica, D.A. Joseph, A. Ganjam, J. Zhan, and H. Zhang, "Understanding the impact of video quality on user engagement," Proc. ACM SIGCOMM'11, pp.362–373, Aug. 2011.

Akira Takahashi received his B.S. degree in mathematics from Hokkaido University in Japan in 1988, M.S. degree in electrical engineering from the California Institute of Technology in the U.S. in 1993, and Ph.D degree in engineering from the University of Tsukuba in Japan in 2007. He joined NTT Laboratories in 1988 and has been engaged in the quality assessment of audio and visual communications. Currently, he is the Project Manager of NTT Network Technology Laboratories. He has been contributing to ITU-T Study Group 12 (SG12) on QoS, QoE, and Performance since 1994. He is a Vice-Chairman of ITU-T SG12. He received the Telecommunication Technology Committee Award in Japan in 2004 and 2016 and the ITU-AJ Award in Japan in 2005 and 2013. He also received the Best Tutorial Paper Award from IEICE in Japan in 2006, Communications Society: Outstanding Contributions Award from IEICE in Japan in 2015, and the Telecommunications Advancement Foundation Awards in Japan in 2007 and 2008.

Takanori Hayashi received his B.E., M.E., and Ph.D degrees in engineering from the University of Tsukuba, Ibaraki, in 1988, 1990, and 2007. He joined NTT Laboratories in 1990 and has been engaged in the quality assessment of multimedia telecommunication and network performance measurement methods. Currently, he is the manager of the Communication Quality Group at NTT Network Technology Laboratories. He received the Telecommunication Advancement Foundation Award in Japan in 2008 and the Telecommunication Technology Committee Award in Japan in 2012.