

ACCURATE AVAILABLE BANDWIDTH ALLOCATION IN HTTP ADAPTIVE STREAMING

Tan Phan-Xuan and Eiji Kamioka

Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo,
Japan

ABSTRACT

HTTP Adaptive Streaming (HAS) has been becoming a de facto standard for Over-the-top (OTT) video services. Typically, by adapting to network conditions, it provides smoother video quality perceived by the end users. However, when the network condition always fluctuates due to some reasons (e.g. bandwidth competition among HAS player or between HAS player and other applications), the perceived video quality might be deteriorated. This demands an effective approach to maintain specific Quality of Experience (QoE) level for the users. To do so, available bandwidth allocation is chosen as a common QoE control method. However, accurately allocating available bandwidth is still a challenge. In this paper, bandwidth allocation based on the relation between subjective Mean Opinions Score (MOS) and requested bitrate is proposed. The relation is captured by a regression model, which is applied to estimate the needed available bandwidth for the users. As the result of controlling the bandwidth, the users start to request the encoding bitrate equal to target bitrate after several requests, resulting in higher perceived video quality.

KEYWORDS

Quality of Experience (QoE), Quality of Service (QoS), HTTP adaptive streaming, Mean Opinion Score (MOS)

1. INTRODUCTION

OTT video services technically rely on HTTP adaptive streaming (HAS) technology to deliver video content from the server to the client. By adapting to network conditions at client-side, this technology allows the client itself to decide a suitable encoding bitrate for the next requested video data fragment. This results in the optimal server-side scalability and the smoother user experience. However, when the underlying network condition frequently fluctuates for some reasons (e.g. bandwidth competition among HAS players), the HAS players often request to change the encoding bitrate, resulting in QoE deterioration. Therefore, it is necessary to assure a specific QoE level for the user, especially the premium user who pays additional cost for their service. Accordingly, the issue is to accurately generate control action in terms of available bandwidth allocation for the specific user.

Shaping traffic [1-4] is known as the most common control action by which available bandwidth is allocated to the user. This allows the users to experience video with expected encoding bitrate after several requests. To guarantee the accuracy for control action, the systems are required to determine the necessary amount of available bandwidth for the user. The approach [4] calculated the available bandwidth based on a certain pre-defined target bitrate which are referred to expect encoding bitrate of the users. Although this approach assured that the HAS player accurately requests the target bitrate after a specific delay, the mechanism to determine target bitrate has not been mentioned. Instead, they simply choose the target bitrate from the encoding bitrate being

available at the streaming server, resulting in the possibility of an available bandwidth overestimation.

The main contribution of the paper is that the available bandwidth is calculated throughout the target bitrate which is predicted from a regression model between subjective MOS and requested bitrate. As the result, the available bandwidth can be accurately allocated to the users. The rest of the paper is organized as follows: Section 2 will provide an overview of background knowledge. Section 3 will state related works. The proposals will be described in the section 4. Meanwhile, section 5 will show the evaluation results on the proposed method. Section 6 will conclude the paper and future work will also be stated.

2. BACKGROUND KNOWLEDGE

Before reviewing related literature that address the issues in available bandwidth control, the HAS adaption mechanism as well as QoE evaluation model are generally investigated as the backgrounds for our work.

2.1. HAS ADAPTION MECHANISM

Since HTTP adaptive streaming was initially launched by Move Networks in 2006, it has become a critical choice for video service providers to deliver the video content over the Internet. There are remarkable advantages which can be found in HAS technique. By using HTTP/TCP, it becomes cost-effective by leveraging the standard web servers and caches. Moreover, all the edge devices are configured to support HTTP connection, it has no difficulties traversing firewalls and NAT devices. Finally, HTTP adaptive streaming is also known as a pull-based method in which the user fully controls his bitrate selection of the requested video content based on its network condition (e.g. available bandwidth, playback buffer size). Thus, the load on the server will be reduced, resulting in better scalability.

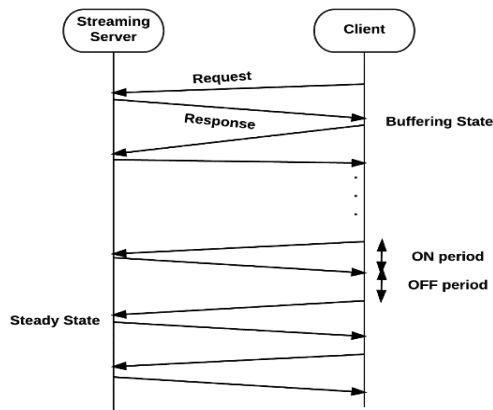


Figure 1. Buffering state and steady state during a streaming session

The concept of adaptive video streaming is based on the idea to adapt the encoding bitrate required by the video stream to the network conditions at client-side. At the server, the video content is divided into fragments. Each fragment is encoded in multiple quality levels, called representations. Based on the estimations of the available throughput and playback buffer size, the user's player will request subsequent fragments at different encoding bitrates in order to deal with varying network conditions. During a streaming session, HAS player performs two main phases: Buffering state and Steady state. In the former one, by continuously downloading video

fragments, HAS player attempts to build up its playback buffer as quickly as possible. In the latter one, HAS player operates as the combination of download and concurrent playback defined by ON and OFF periods, respectively, shown in Fig. 1. The purpose of this phase is to keep playback buffer staying at a stable level (e.g. 30s). Therefore, it is clear to see that both QoS and playback buffer have great influence on the behaviour of HAS adaption. Figure 2 shows a cause-effect relationship which refers to the temporal behaviours of QoS, playback buffer and requested bitrate when bandwidth competition (either among HAS players or among HAS player and other applications) occurs. It is worth noting that the player usually takes time to react to negative changes in network condition [5]. As it does not react based on the latest per-fragment throughput measurements. Instead, it averages those measurements over a long period so that it acts on a smoother estimate of the available bandwidth variations.

2.2. QOE EVALUATION MODEL

Quality of Experience (QoE) is a complicated concept of subjectively perceived quality which can be defined in many ways. According to [6], QoE is "the overall acceptability of an application or service, as perceived subjectively by the end user". In order to measure HAS video QoE, Mean Opinion Score (MOS) has been utilized as the most popular indicator. The opinion score is actually defined as the "value on a predefined scale that a subject assigns to his opinion of the performance of a system". The MOS is the average of these scores across subjects [7]. MOS is not only used to express the results of subjective tests ("subjective MOS") but also used as the output of objective measurement algorithms, which provides an automated alternative to subjective tests. Beyond that, Pseudo-subjective Quality Assessment proposed by Samir Mohamed and Gerardo Rubino [8], provided an effective model to assess QoE, that is QoE hybrid evaluation model. The basic idea is that a Random Neural Network is trained to capture the relationship between QoE influence factors and the users' perceived video quality. As the result, it is capable of evaluating video quality as accurately as human does and then can be applied for real-time QoE estimation in an accurate and automatic way [9-12]. Our research also utilized this method to estimate MOS from according network condition which is defined by QoS parameters (available bandwidth, packet loss, delay and jitter). The process to establish this model is clearly presented in section 4.

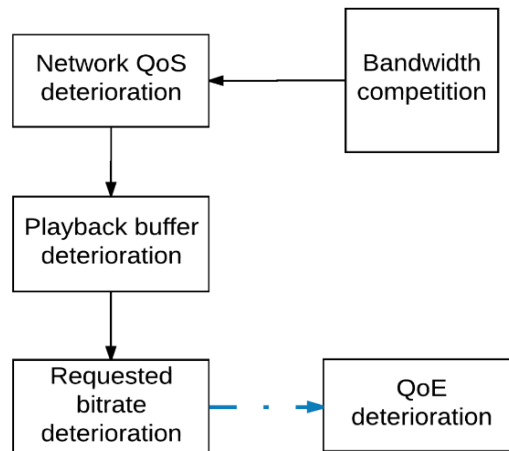


Figure 2. Deteriorations of QoS, playback buffer size and requested bitrate are in turn captured when available bandwidth competition occurs

3. RELATED WORK

QoE management for HAS-based OTT video services refers to two major aspects: QoE monitoring and QoE control. Both of them are crucial for ensuring QoE level for the users when the underlying network conditions fluctuates. If QoE management effectively works to keep the users staying with the service, the revenue of service provider will be maintained or improved. In this paper, we focus on available bandwidth control as a part of QoE control.

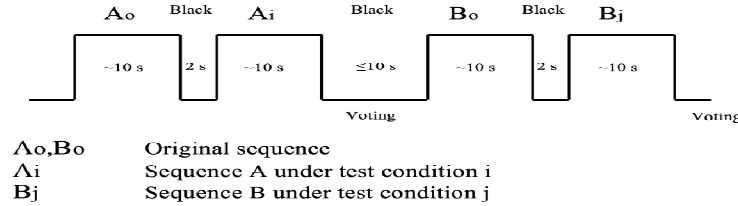


Figure 3. Stimulation presentation timing in DCR method

In general, based on the monitoring results, QoE control decides whether a control action should be performed or not. The final purpose of the control action is to keep QoE at an expected level for the specific users. Accordingly, some existing works proposed new adaption algorithms [13] or introduced a network proxy to select optimal bitrate for the users [15]. In order to enhance QoE, the authors [14] proposed QoE control by adding Forward Error Correction (FEC) packets to the current flow which is capable of compensating for the packet loss. Notably, there are existing studies including our works that focus on shaping traffic as a control action to maintain QoE. When the available bandwidth shrinks due to the bandwidth competition [16], the unfairness and instability problems of the requested bitrate will occur. The traffic shaping method effectively solves these problems [1-4]. Particularly, in [4], the authors proposed a method to identify how much available bandwidth is needed for the traffic shaping based on Eq. (1) (as shown in section 4). Each commercial HAS player actually has their own safety margin to ensure the available bandwidth is enough for the next encoding bitrate (e.g. 20% of Microsoft smooth streaming). The result showed that the target bitrate was accurately requested by the user after assigning the calculated available bandwidth which for the users. However, they did not mention the mechanism how to determine the target bitrate, thus, the easiest way is to simply pick up the highest possible encoding bitrate which is available at the server for the next fragment.

In this paper, the clarified relation between subjective MOS and requested bitrate, which captured by a regression model. By which it allows calculating the target bitrate from the user's expected MOS. Thereby, the needed available bandwidth is accurately assigned to the users. As the result, the users accurately request the encoding bitrate which is equal to calculated target bitrate after several requests.

4. METHODOLOGY

In this section, initially, MOS estimation model is presented as a tool to estimate MOS for our studies. Afterward, the proposal of available bandwidth allocation based on the relation between subjective MOS and requested bitrate are briefly introduced.

4.1. MOS ESTIMATION

In our research, QoE is evaluated throughout a trained Artificial Neural Networks (ANN), which is established by training a dataset comprising of QoS parameters as input and subjective Mean Opinion Scores (MOS) as output. Particularly, input data is a set of various QoS parameters, namely, available bandwidth, packet loss, delay, and jitter. For each selected input parameter, a discrete set of common values is chosen. Each combination of values of QoS parameters is called as a system configuration. There are totally 294 prepared configurations set up at WANEM router (WAN Emulator 3.0). The open source movie, "Big Buck Bunny" has been chosen to be watched by the subjects. The movie is cut into 10-seconds sequences which are available with multiple bitrate at the server. Accordingly, there are totally 294 sequences taken from original movie and are sent throughout the set up network with different configurations. Consequently, the distorted video sequences will be obtained. In order to collect subjective MOS as the outcome of distorted video sequences' evaluations, Degradation Category Rating (DCR) methodology has been chosen [17]. DCR requires that the testing sequences need to be presented in pairs: the first sequence in each pair is always the source reference (original sequence), while the second one is the distorted sequence. The length of original sequence and distorted sequence are equal to 10seconds. There are 17 subjects being asked to watch the video sequences and vote following five-scale Mean Opinion Scores. Finally, the average of those subjective evaluations is obtained.

Before training, dataset with 294 samples is divided into three parts: training data, validation data and testing data corresponding to 70%, 15%, and 15% of the dataset. The training data is presented to the network during training, and the network is adjusted according to its error. The validation data is used to measure the network generalization and to halt the training process when the generalization stops improving. The testing part has no effect on the training process, and thus provides an independent measure of network performance during and after the training process. Apart from the training dataset, the neural network architecture is also considered, which comprises of totally 4 neurons for input layer, 10 neurons for hidden layer and 1 neuron for output layer. For the training algorithm, Levenberg-Marquardt algorithm which requires more memory but less time is used. There are two metrics to evaluate the performance of trained neural networks: Mean squared error (MSE) and Regression R values. MSE is average squared difference between outputs and targets. Lower values are better, whereas, zero means no error. R values show the correlation between outputs and targets. An R of 1 means a close relation. Figure 4 shows the result of training process in which regression R value is equal to 0.9. It means that the predicted data has a well correlation with actual data. Therefore, the trained neural network can be used for estimating MOS.

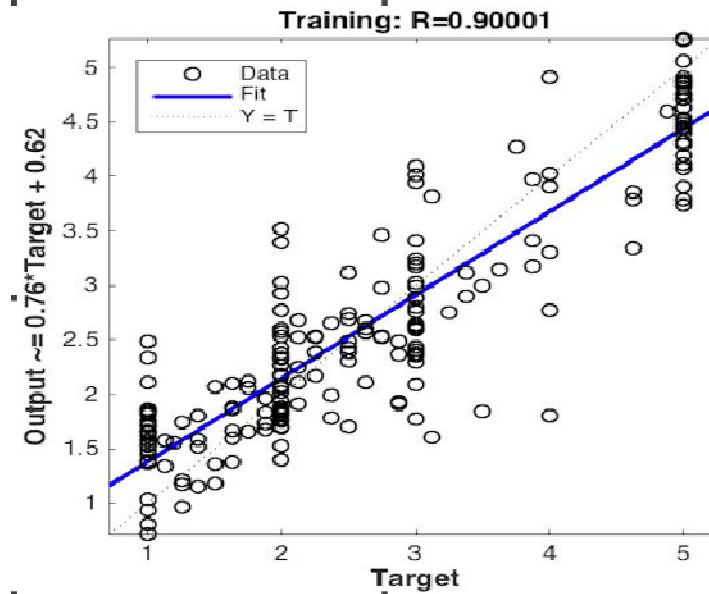


Figure 4. Result of training process with $R=0.9$ referring to a well-fit between actual subjective MOS (horizontal axis) and predicted MOS (vertical axis)

4.2. PROPOSED METHODS

In this section, bandwidth allocation which is based on the relation between subjective MOS and requested bitrate is proposed. The proposal aims at maintaining an expected QoE level in terms of requested bitrate for particular video users, especially the premium user.

As mentioned, shaping traffic has been known as an effective method to maintain the requested bitrate level for the users [1-4]. However, determining an appropriate available bandwidth to allocate to the users is a challenge. According to [4], the authors proposed an equation representing the relation among available bandwidth, target bitrate and margin ration, shown in Eq. (1)

$$\text{margin_ratio} = \frac{\text{avail_BW} - \text{target_BR}}{\text{avail_BW}} \quad (1)$$

where, avail_BW is the available bandwidth, and target_BR is the expected video representation (target bitrate) of the users. Margin ration refers to a conservatism value defined by particular proprietary HAS players. For instance, Microsoft smooth streaming and HLS players apply a conservatism value of 20% and 40%, respectively. Based on this equation, those authors could determine the available bandwidth according to pre-defined target bitrate, resulting in an accurate QoE control. However, they actually obtained the target bitrate based on a list of video encoding bitrate which is available at the server. Particularly, based on current requested bitrate, they picked the highest possible encoding bitrate for the target bitrate. This leads to a suboptimal network resource utilization. Therefore, a relation between subjective MOS and requested bitrate has been studied to determine the target bitrate in QoE control. In general, the regression model which captures this relation is established based on the dataset of subjective MOS and requested bitrate. Accordingly, an experiments were performed, in which 15 students were asked to watch 294 short movies which created from original one under 294 configurations of QoS parameters, the same as in section 4.1, then, gave out their evaluation in term of MOS. The sample dataset collected from experiment is presented in Table 1. Accordingly, the average value of human evaluations was called as subjective MOS. Meanwhile, the requested bitrate is the most frequent

requested encoding bitrate during streaming sessions for each distorted video. An identified trend from raw data that represents the correlation between the subjective MOS and requested bitrate is shown in Fig. 5. It can be seen that the subjective MOS is less than 3 if the requested bitrate is less than 1024Kbps. Moreover, the gaps among the values of those requested bitrates are quite small. This graph also shows a big variation in requested bitrate when the subjective MOS is around about 3. Meanwhile, the requested bitrate is mostly in a range of 1536Kbps and 2048Kbps if the subjective MOS alters between 4 and 5.

Table 1. The most appearance encoding bitrate requested by the subjects and the subjective MOS

Subjective MOS	Requested bitrate (Kbps)
1.87	331
2.75	447
4.63	1427
3.12	991

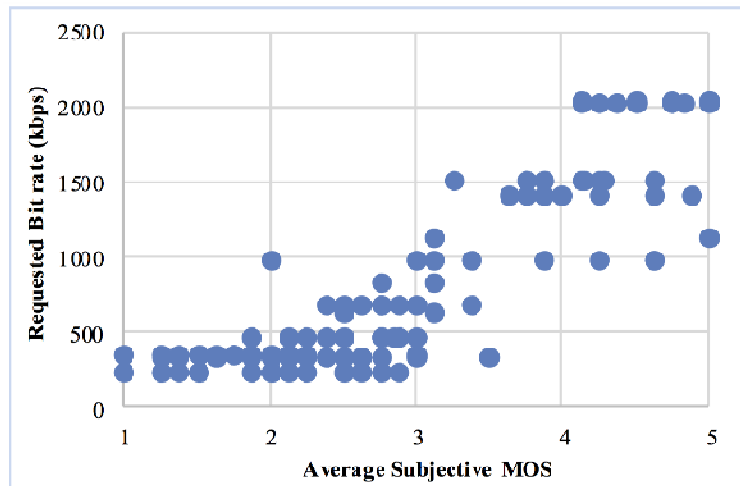


Figure 5. The trend obtained raw data of subjective MOS and requested bitrate

From raw data, a regression model was also established by modelling data with linear combination of basic function in which a set of functions $\phi_0, \phi_1 \dots \phi_P$ was specified along with finding function f in the form of linear combination:

$$f(x) = \sum_{i=0}^P \theta_i \phi_i(x) \tag{2}$$

In this case, Gaussian Radial Basic Functions (RBF) was chosen due to its popular applications. The result of modeling data is plotted in Fig. 6 with high coefficient of determination.

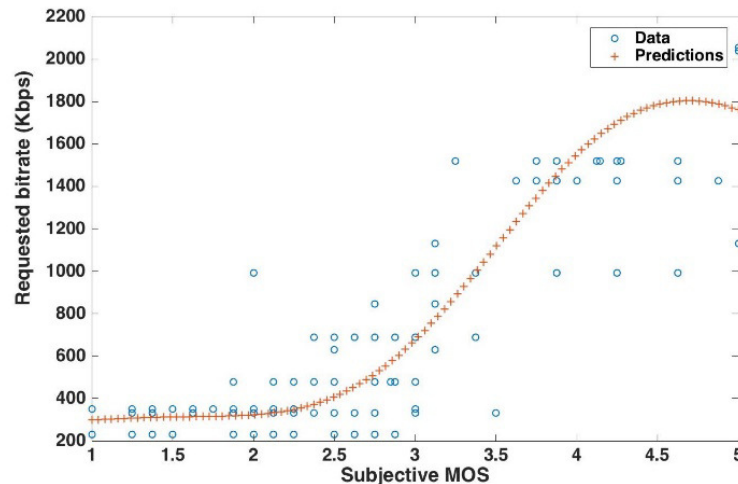


Figure 6. Regression curve represents the relation between subjective MOS and requested bitrate

Once the target bitrate is determined based on a certain expected subjective MOS, the available bandwidth can be calculated via Eq. (1), leading the user to accurately request the target bitrate after some requests.

5. EVALUATION

This section aims to investigate how accurately available bandwidth can control be performed when the perceived quality of the premium user falls to specific lower levels. Note that, since the estimated MOS decreased, QoE control was activated when the estimated MOS was below the threshold. According to [9], a specific estimated MOS threshold of 3 was simply chosen. The evaluation environment was set up as follows: The testbed consisted of a router, a streaming server, and three users. The router was a Linux-based router, namely, WAN Emulator release 3.0 running on a VMware workstation located on a desktop computer with Intel Core i5 3.10 GHz processor and 8 GB RAM. This router works as a controller which is capable of adjusting available bandwidth, packet loss, delay, and jitter. The streaming sever was deployed on a desktop computer with Windows 8.1, Intel Core i5 3.10 Ghz processor and 8 GB RAM. The server published a Microsoft smooth streaming video content of "Big Buck Bunny" which is an open source testing movie. This movie content was encoded with multiple bit rates. Furthermore, a Smooth Streaming-compatible Silverlight player template was installed on the Smooth Streaming enabled streaming server so that Silverlight-based users can play Smooth Streams. The users utilized the laptop computers with MacOS, Core i5 and 8 GB RAM in which the latest version of Microsoft Silverlight add-on was installed. The server and the users' computers were located in different broadcast domains and they were connected via the router. The network topology used for this experiment is shown in Fig.7. By relying on "ping" packets and packets generated by "iperf" tool, a QoS monitoring software deployed at the router monitored the available bandwidth, packet loss, delay and jitter. In addition, Wireshark, which is a network packet analyzer, installed at the router captured the HTTP request from the client.

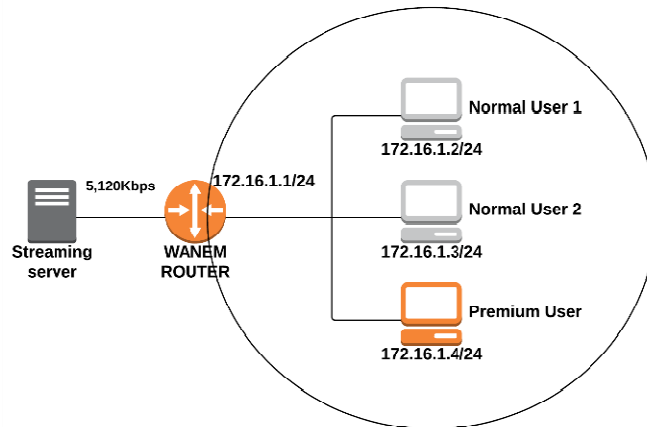


Figure 7. Set up network environment for investigating the accuracy of control action

The experimental procedure for two scenarios of the evaluation are as follows:

- (1) The first user as the premium user starts watching a streaming video content.
- (2) The second and third user respectively stream the video at $t=60s$ and $t=120s$ on purpose to make the network quality of the premium user deteriorated.
- (3) The packet loss, delay and jitter in the network are observed.
- (4) The deterioration is detected by observing the requested bitrate and the estimated MOS.
- (5) The available bandwidth to the premium user is increased to recover the network quality when the deterioration of requested bitrate and estimated MOS are detected.

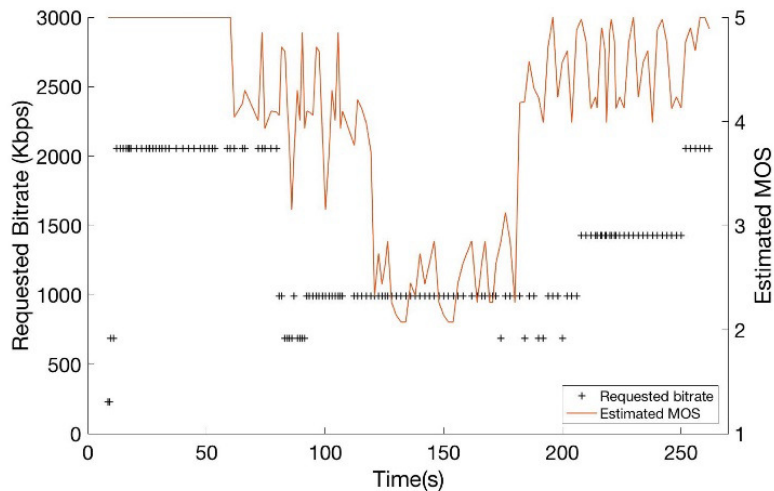


Figure 8. Requested bitrate and estimated MOS of the premium user during his streaming session (before and when bandwidth competition occurs)

Figure 8 shows the changes of the estimated MOS and the requested bitrate of the premium user under different situations. Accordingly, the horizontal axis indicates the time durations, the first

vertical axis is the requested bitrate, and the second vertical axis is the estimated MOS. During the first 60s of the experiment, when the premium user was solely watching the movie, it requested the highest encoding bitrate of 2056Kbps only after several transitions. The player would request the value of 2962Kbps if the users were watching the video in full-screen mode. Meanwhile, the estimated MOS was stable at the highest value of 5. At $t=60s$, once the first normal user started to request the video content, the available bandwidth of the premium user immediately reduced to around 2500Kbps. The estimated MOS decreased to around 4, whereas the player still requested the video content at bitrate of 2056Kbps for 19 seconds. When the second normal user watched the video, the available bandwidth of the premium user shrank to 1578Kbps followed by a fluctuation around the value of 2.5 of the estimated MOS, which is below the set up threshold of 3. At this time, the system defined an expected subjective MOS value of 5 for the premium user. After predicting target bitrate from expected subjective MOS and calculating the available bandwidth based on Eq. (1), at $t=242s$, the player reached to its calculated target bitrate of 2056Kbps. Figure 9 and Fig. 10 show the estimated MOS and the requested bitrate of both two normal users. It can be seen in Fig. 9 that the estimated MOS of the first normal user fluctuated around the acceptable range level of 2 and 3 from time $t=60s$ to $t=262s$. On the other hand, the requested bitrate was stable around 1130Kbps during the time it has to compete for the available bandwidth with only the premium user. This is followed by a slightly decrease to 688Kbps since the second normal user participated in the network. A similar trend of the second normal user also can be seen in Fig. 10.

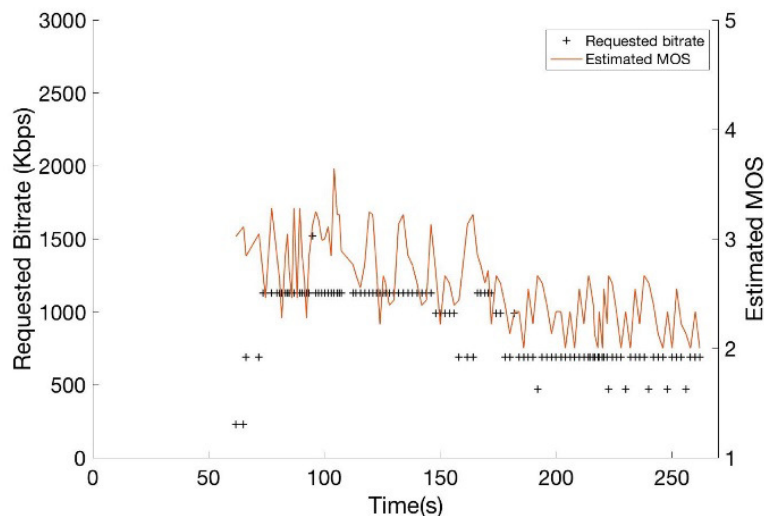


Figure 9. Requested bitrate and estimated MOS of the first normal user during his streaming session

The purpose of this experiment is to confirm the accuracy of available bandwidth control by leveraging the relation of subjective MOS and requested bitrate which captured by a regression model. The result shows that the target bitrate can be accurately predicted from the expected subjective MOS when the perceived quality is less than the threshold. Particularly, when the system set the expected subjective MOS is equal to 5, the regression model predicts the target bitrate between 2056Kbps and 2962Kbps. However, because the users are not watching with the full-screen mode, thus, it is better to pick the former one, which optimizes the network resource utilization. Based on the Eq. (1), the available bandwidth is calculated to be equal to 3343Kbps which is much less than its initial available bandwidth of 5120Kbps. Applying this calculated available bandwidth not only supports the premium user to achieve its expected subjective MOS value, but also optimizes the network utilization and keeps enough available bandwidth for the normal users to request an accepted bitrate level. As shown in Fig. 9 and Fig. 10, even though

there is a serious competition in the network, the estimated MOS of both normal users are still stable at accepted levels. However, this approach is only applied for Microsoft Smooth Streaming technology, thus, it is necessary to confirm with other commercial technologies as well as non-commercial technologies. In addition, the larger number of the premium user also need to be considered.

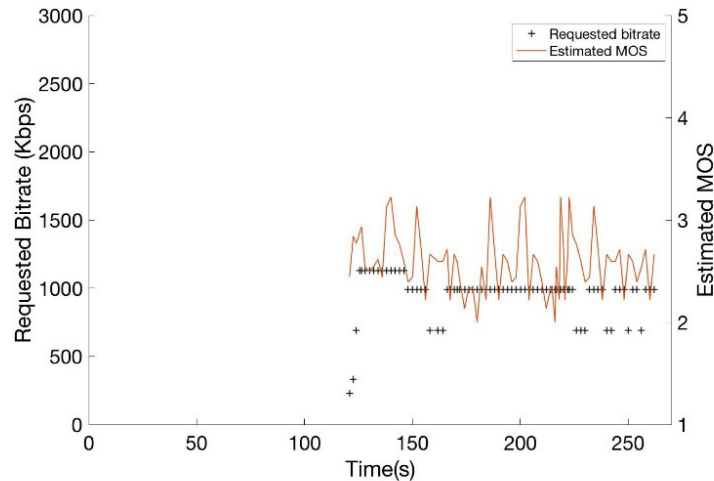


Figure 10. Requested bitrate and estimated MOS of the second normal user during his streaming session

6. CONCLUSION

The fluctuation in network conditions usually requires the HAS players to frequently requests new encoding bitrate, resulting in QoE deterioration. This demands the approach to ensure an expected QoE level for the users. The approach is proposed based on the requirement of accurately generating control action (in terms of available bandwidth allocation). Therefore, a relation between subjective MOS and requested bitrate was taken into account to estimate the target bitrate facilitating available bandwidth allocation process. The experimental result showed that calculated available bandwidth was sufficient for the users to request exact target bitrate that has been pre-defined from expected subjective MOS.

For the future work, the proposed method will be deployed along with QoE monitoring approach in order to clarify its effectiveness in maintaining QoE for HAS. The psychological factors are also considered which improves its performance.

REFERENCES

- [1] S. Akhshabi et al., (2013, February) "Server-based traffic shaping for stabilizing oscillating adaptive streaming players", in Proceeding of the 23rd ACM workshop on network and operating systems support for digital audio and video, pp19-24
- [2] C. B. Ameer et al., (2014) "Shaping HTTP adaptive streams using receive window tuning method in home gateway", in Conference on Performance Computing and Communications (IPCCC), IEEE, pp1-2.
- [3] R. Houdaille and G. Stephane, (2012) "Shaping HTTP adaptive streams for a better user experience", in Proceeding of 3rd Conference on Multimedia Systems, ACM.
- [4] C.T. Guguen et al., (2017) "Improving user experience when HTTP adaptive streaming clients compete for bandwidth", SMPTE Motion Imaging Journal, Vol.126, No.1, pp28-34.

- [5] S. Akhshabi et al., (2012) "An experimental evaluation of rate-adaptive video players over HTTP", *Signal Processing: Image Communication*, Vol.27, No.4, pp271-287.
- [6] ITU-T, R.P. 10/G, (2007) "New Appendix I–Definition of Quality of Experience (QoE)," International Telecommunication Union.
- [7] R. C. Streijl et al., (2016) " Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives", *Multimedia Systems*, Vol.22, No.2, pp213-227.
- [8] S. Mohamed and G. Rubino, (2002) "A study of real-time packet video quality using random neural networks", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.12, No.12, pp1071-1083.
- [9] J. Seppänen et al., (2014) "An autonomous QoE-driven network management", *Journal of Visual Communication and Image Representation*, Vol.25, No.3, pp565-577.
- [10] P. Casas et al., (2013) "A system for on-line monitoring of YouTube QoE in operational 3G networks", *ACM SIGMETRICS Performance Evaluation Review*, Vol.41, No.2, pp44-46.
- [11] K.D. Singh et al., (2012) "Quality of experience estimation for adaptive HTTP/TCP video streaming using H.264/AVC", in *EEE Conference on Consumer Communication and Networking (CCNC)*, pp127-131
- [12] R. K. Mok et al., (2011) "Measuring the quality of experience of HTTP video streaming", in *IFIP/IEEE International Symposium on Integrated Network Management*, pp485-492.
- [13] J. Junchen et al., (2012) "Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive", in *Proceeding of 8th Conference on Emerging Networking Experiments and Technologies*, pp99-108.
- [14] S. Latré, et al., (2009) "An autonomic architecture for optimizing QoE in multimedia access networks," *Computer Networks*, Vol.53, No.10, pp1587-1602.
- [15] K. J. MA, (2012) "Provider-Controlled Bandwidth Management for HTTP-based Video Delivery," PhD Thesis, University of New Hampshire (Department of Computer Science).
- [16] S. Akhshabi, et al., (2012) "What happens when HTTP adaptive streaming players compete for bandwidth", in *Proceeding of 22nd International workshop on Network and Operating System Support for Digital Audio and Video*, pp9-14.
- [17] ITU-T, R.P. 910, (2008) "Subjective video quality assessment methods for multimedia applications".

AUTHORS

Tan Phan-Xuan received the B.E (2009) in Electrical-Electronic Engineering from Military Technical Academy, M.S (2013) in Computer and Communication Engineering from Hanoi University of Science and Technology (HUST). He has worked as a network administrator for BachKhoa Network Information Center (BKNIC) since 2011. Currently he is PhD student at Shibaura Institute of Technology (SIT), Japan. His research interests are Multimedia quality, Networking.



Eiji Kamioka is a professor at Shibaura Institute of Technology. He received his B.S, M.S and D.S degrees in Physics from Aoyama Gakuin University. Before joining SIT, he was working for SHARP Communication Laboratory, Institute of Space and Astronautical Science (ISAS) as JPSP Research Fellow and National Institute of Informatics (NII) as Assistant Professor. His Current Research Interests encompass Mobile Multimedia Communications and Ubiquitous Computing.

