Shibaura Institute of Technology

Biological Information Based QoE Management In Adaptive Streaming Services

A DISSERTATION SUBMITTED TO THE
GRADUATE SCHOOL OF ENGINEERING AND SCIENCE OF THE
SHIBAURA INSTITUTE OF TECHNOLOGY

by

PHAN XUAN TAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF ENGINEERING

SEPTEMBER 2018

Abstract

The introduction of adaptive streaming technologies, especially HTTP adaptive streaming (HAS) has significantly improved the video quality perceived by end-user, making video service becomes one of the most dominant services on the Internet. Due to the limitation of network resource supply (e.g., available bandwidth), for profit improvement, service providers have to take into account Quality of Experience (QoE) management by which QoE stands for perceived video quality will be frequently monitored and maintained with optimal network resource utilization. However, with the growth in the availability of multimedia services, coupled with the technological advances in compression and streaming, it is witnessing a great demand for video contents with high quality. Meanwhile, the number of subscribers is also continuously increasing. These situations put more pressure on the existing network infrastructures, requiring an upgrade. However, when the service providers attempt to upgrade their systems, it might reach to the physical limits. Thus, there is a growing need of more efficient QoE management system in adaptive streaming services.

In this dissertation, a biological information based QoE management framework has been proposed. Thereby, a balance between network resource utilization and the resulting QoE is guaranteed. The achieved results are outlined in the following:

- First, one of the requirements of QoE monitoring is to perform early detection of QoE deterioration. The design of QoE monitoring usually comprises of two major steps: Selecting appropriate monitoring factors and selecting suitable monitoring interval. In this research, adaption logic factors comprising of playback buffer, video rate and QoS parameters, have been investigated. Both playback buffer and video rate can only be obtained on a

chunk-by-chunk basic that relies on the timestamp of two successive video requests. In addition, video rate is usually selected based on a throughput estimation over a long time period. Thus, the deteriorations will be perceived by the end-user before control action is generated. Meanwhile, QoS can be monitored with flexible self-defined interval that does not depend on chunk-by-chunk basic, becoming a suitable monitoring factor for early detection purpose. This study aims at determining such the appropriate self-defined interval of QoS monitoring. In adaptive streaming, playback buffer is a situational indicator for video rate adaption. The results of an experiment demonstrated that the first deterioration of playback buffer always provides an accurate prediction of video rate deterioration. Therefore, monitoring QoS with suitable interval can accurately capture the first deterioration of playback buffer, benefiting early detection of video rate or QoE deterioration. The monitoring interval is then proposed to be equal to video chunk size. By using the proposed interval, the balance between computational cost and ratio of video rate deterioration has been achieved.

- Second, in QoE control, threshold plays an important role in deciding when control action should be triggered. However, similar to monitoring interval, it has not been carefully investigated yet. In literature, QoE threshold is usually picked up as the fair level or the middle level in 5-scale Mean Opinion Score (MOS) (the most common QoE indicator) without reasonable explanation. It motivated us to propose a novel method to determine the appropriate value of QoE threshold. In this research, by clarifying the drawback of existing approaches in determining threshold, a novel collaborative approach using psychophysiology and psychophysics was proposed to ascertain an appropriate QoE threshold. Consequently, the experimental results demonstrate that using the proposed threshold can save at least 4.85% of available bandwidth per control compared to the use of fair one.

- Third, in QoE control, bandwidth allocation is commonly used as control action. In order to accurately allocate bandwidth to the users, some existing works calculate the needed bandwidth based on target video rate. However, the determination of target video rate is still a challenge, where contemporary researchers simply pick up the target video rate from a list of available video rate at server. Therefore, we proposed a method to determine the target video rate from expected subjective MOS level by leveraging a regression model which expresses the relation between video rate and subjective QoE. The evaluative results show that once being lower than threshold, estimated QoE will be automatically recovered to the expected level, while more bandwidth can be saved per control.

Acknowledgments

First of all, I would like to express my most heartfelt gratitude and appreciation to my supervisor, Professor **Eiji Kamioka**, who has been a constant source of inspriration for me, professionally and personally. His continued support led me to the right way. Besides guiding my research, he has allowed me to recognize and address some of my weaknesses, and develop into a more effective researcher. I cannot thank him enough for everything that he taught me over three years. In the future, I do hope that I will become as good an supervisor as he is.

I would also like to extend my appreciation to the committee members: Professor **Hiroaki Morino**, Professor **Koichi Gyoda**, Professor **Kenichi Yamazaki** and Professor **Kyoko Yamori** for their advice and comments for my research.

Thanks also to all staff members of Shibaura Institute of Technology. They was always very kind and generous in helping me during my stay in Japan.

I would also like to thank all members of Mobile Multimedia Communication Laboratory who made my life at Shibaura Institute of Technology much easier. The enjoyable parties and gPBL activities are unforgettable memories in Japan.

I would like to acknowledge the financial support from Shibaura Institute of Technology through out the Hybrid Twinning program and Japanese Government Scholarship (MEXT) for my doctoral course and my daily life in Japan.

Finally, I would like to extend my deepest gratitude to my parents, little brother **Phan Tien Dung**, my wife **Ngo Thanh Thuy** and my baby **"Bee"** for their love, encouragement and many sacrifices, making this dream of higher education (PhD degree) come true.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

February of 2018 was the time for the 23rd Winter Olympics Pyeongchang. It was an exhilarating month, with the participants of 92 national teams. Unsurprisingly, the Winter Olympics was one of the most watched worldwide events, and this year was not the exception. Approximately 11.6 million viewers turned into live-streaming coverage on NBCOlympics.com during the first five days of the Games in Pyeongchang in 2018 [3]. Thanks to the recent advances in video streaming and compression techniques, the viewers were able to also watch this event on live-TV streaming services, including Sling TV's Sling Blue package, Hulu with Live TV, YouTube TV, DirecTV Now, Sony's PlayStation Vue, FuboTV, and CenturyLink Stream with high definition quality on their devices. Sporting events like Winter Olympics are not the sole type of attractive content to online viewers, though. The day of having to tune in live or buy expensive discs in order to watch the favorite shows and movie has gone. There are numerous streaming services offering an excellent and abundant selection of TV shows, movies, and original programing, including Netflix, Hulu, Amazon Instant video, etc. The rise of such online services has dramatically altered the media habits of Americans, especially young adults [1]. About six-in-ten of those ages 18 to 29 (61%) say the primary way they watch television now is with streaming services on the internet, compared with 31% who say they monthly watch via a cable or satellite subscription and 5% who mainly watch with a digital antenna.

According to Cisco visual networking index 2017 [4], globally, video traffic is predicted to account for 82% of all consumer Internet traffic by 2021, up from 73% in 2016. Every second, a million minutes of video content will cross the

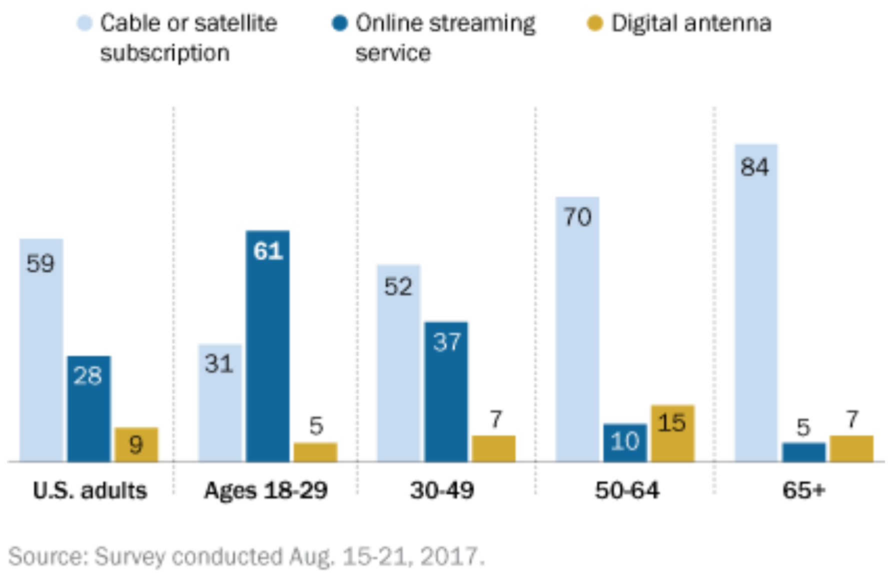

Figure 1.1: Survey on how the US young adults watch television these day [1]

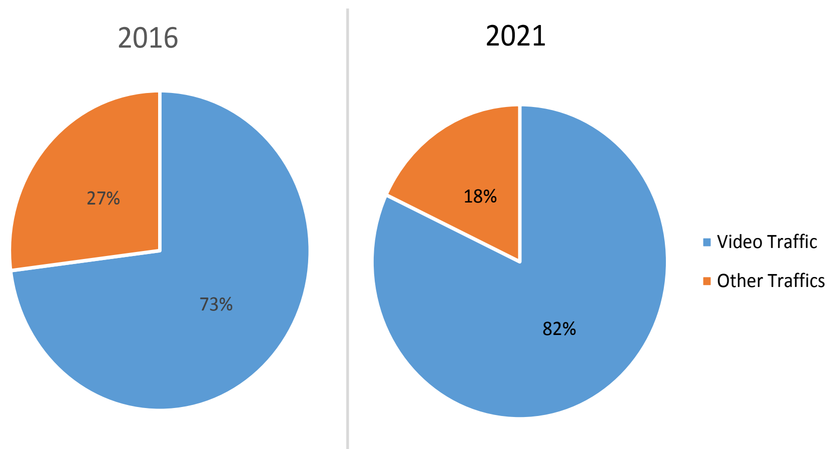

Figure 1.2: The dominance of video services

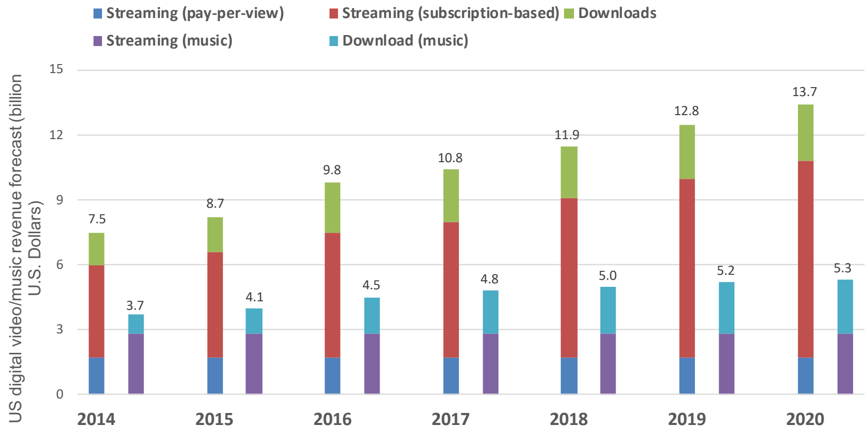

Figure 1.3: Online video services-a billion dollar opportunity

network by 2021. Meanwhile, people around the world watch a billion hours of YouTube video content every single day [5]. In a 2017 report [6], IBM stated that there are two-thirds of adults reported subscribing to video on demand (SVOD) service in a survey of more than one thousand US consumers. The dominance of video services accordingly brings a huge profit to service providers. The revenue in billion US Dollar of video services is increasing year after year and is predicted to reach the highest value of 13.7 by 2020 [7]. Such the huge revenue is attracting more and more service providers to participate in this prominent but competitive market.

1.1 Motivation

In the fierce competition to gain the market share, it is important for service providers to increasingly deliver demanding services with higher quality standard allowing the users′ satisfaction to exceed the expectations. In order to achieve this, in point of view of service providers and the users, it is necessary to adopt adaptive streaming technology, especially HTTP adaptive streaming (HAS) as the major delivery method in addition to the crucial shift from technical quality requirements (QoS) to perceived quality requirements (QoE). In this situation, QoE management for adaptive streaming has been emerged, where the perceived video quality of individual user is frequently monitored and maintained at expected level. Nonetheless, the introduction of high performance end-devices and the mobility thereof put more challenges on the existing network systems. This leads to the frequent upgrades of available resources as well as the evolution of adaptive streaming technologies. Due to physical limits, however, the upgrade of adaptive transmission technologies become more challenge alongside the requirements of more efficient QoE management system.

The adaptive streaming technologies share several critical aspects. First, multiple files from same video content are produced to be distributed to viewers watching on different powered devices via different connection speeds. Second, these files are adaptively streamed, changing the stream that is delivered to adapt to the current network conditions (e.g., throughput, CPU load, etc.). Third, these technologies all operate transparently to the user. It means that all stream switching occurs behind the scenes, resulting in the fact that the viewer might notice a slight change in quality as the streams switch, but no action is required on his part. For years, adaptive streaming technologies has experienced several remarkable transformations in terms of used protocol in order to deliver higher perceived quality video to the users. The transformation is initially discussed with the use of the so-called Real-time Streaming Protocol (RTSP), where its primary mechanism is that the server delivers content at the encoding video rate to match the client's consumption rate [8], resulting in a stable client's buffer levels over time. Thus, the network resource usage will be optimized. In addition, the server via RTP Control Protocol (RTCP) can indirectly monitor the network condition at the client side, thereby, it makes the decision to switch to higher or lower bitrate stream sending to the client's player, resulting in smoother playback at the best possible quality level without pauses or stuttering. However, traversing the edge network devices (e.g., Network Address Translation (NAT) and Firewall) is the noticeable obstacle for this technology because it uses transport protocol (e.g., User Datagram Protocol (UDP)) with uncommon port number to deliver the video content on the Internet. In addition, as a server-based technology, the initial implementation of this technology requires a persistent connection between server and player, which potentially increase the implementation cost and limit deployment scalability. In order to overcome these limitations, HTTP adaptive streaming has been introduced in 2008 [9] as the de-facto standard for adaptive streaming solutions. It requires video content to be divided into small chunks each of which is encoded in different video rate and quality levels. HAS player frequently monitors its underlying network conditions, then, selects which video rate to request for the next video chunks, improving server-side scalability. By doing this way, the HAS player also can control its playback buffer by dynamically adjusting the rate at which new chunks are requested. Consequently, it facilitates the users to watch video with smoother experience without interruption despite the network condition fluctuation. However, during streaming session, each of HAS player strives to optimize their individual quality, which leads to bandwidth competition, causing quality oscillations and buffer starvations. Therefore, the upgrade of adaptive transmission technologies only is not enough.

QoE management has emerged as an alternative solution to converge the requirements of both service providers and the end-user. Typically, with QoE management, perceived video quality in terms of QoE is frequently monitored and maintained at an expected level. QoE is assessed by observing QoE influence factors (e.g., QoS parameters, stalling, rebuffering) and then interpreting them to QoE indicators (e.g., Mean Opinion Score). Meanwhile, QoE control is responsible for maintaining a desirable QoE level as long as possible and it can be done by generating appropriate control action at the right time. Thereby, an efficient QoE management for adaptive video streaming services can guarantee high perceived video quality for the end-user at minimal network resource usage. The existing works increasingly attempt to improve the accuracy of QoE assessment, without seriously considering the other aspects of QoE management such as QoE deterioration detection, control action and especially QoE threshold, where remained challenges are taking place at.

This thesis aims at proposing a novel biological information based QoE management for adaptive streaming services. Thereby, the balance between network resource utilization and the resulting QoE can be efficiently achieved through solving the remained challenges. The existing technology alongside challenges, the overview of the proposed approach as well as the contributions and the organization of the dissertation will be presented in the remainder of this chapter.

1.2 QoE Management and Challenges

The overall goals of QoE management for adaptive streaming services are to match the properties of the video to the expectations of the end-user, while accounting for the available resources and characteristics of the encoding and the transport systems. Figure.1.4 illustrates a block diagram of negative feedback

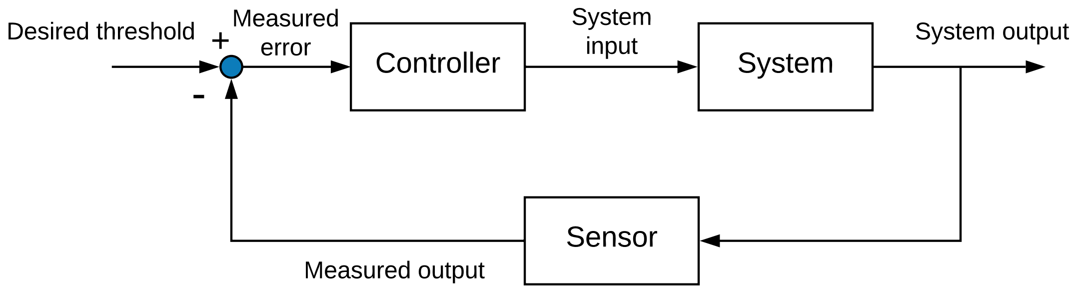

Figure 1.4: Negative feedback control system

control system in control theory, using a feedback loop to control the process variable by comparing it with a desired threshold and applying the difference as an error signal to generate a control output to reduce or eliminate the error [10]. Such the functionalities actually can be found in monitoring layer and control layer within any general QoE management frameworks. Actually, only very few studies on QoE management exist in the public domain because the majority of studies have been mainly sponsored by the players in the telecommunications industry, who view the results of such studies as proprietary. On the contrary, there are numerous proposed solutions for managing QoE of video services, that focus on the improvements of separate monitoring layer and control layers.

The operations of monitoring layer can be broken down into two general steps: (1) QoE assessment, and (2) QoE estimation [11]. QoE assessment aims to model the relationship between different measurable QoE influence factors and QoE indicators. Such models serve the purpose of making QoE estimations, given a set of conditions, corresponding as closely as possible to the QoE as perceived by end-user. There are three types of QoE assessment models - objective models, subjective models and hybrid models. Objective models are defined as the means of estimating subjective quality solely from objective quality measurement or indices. Subjective models are based on psychoacoustic/visual experiments which represent the fundamental and most reliable way to assess users′ QoE, although they are complex and costly. Hybrid models are the most commonly used models in literature, which leverages the advantages of both abovementioned model types to automatically assess QoE in an accurate manner. They rely on pre-trained machine learning model which is established by training a machine learning model with the input (represented by QoE influence factors) and output (represented by QoE indicators). QoE estimation encompasses the acquisition of data of QoE influence factors related to the network environment and conditions, terminal capabilities, users, context and application/service specific information and its quantification. These factors can be frequently obtained via passive or active monitoring methods, and then automatically interpreted to QoE indicators in real-time. As the prerequisite conditions for an implementation of successful QoE monitoring, QoE must be accurately assessed, while QoE deterioration is early detected. Contemporary works often concentrate on the first condition, that is to say, enhancing the accuracy of QoE assessment. Typically, in order to provide accurate QoE assessment, the consideration of only one or two QoE influence factors is generally not sufficient. On the contrary, QoE should be considered in all its dimensions taking into account as many influence factors as possible. In fact, not only network QoS factors (e.g., bandwidth, packet loss, delay and jitter), but also application QoS factors (e.g., initial buffering time, rebuffering frequency, and mean duration of a rebuffering event) [12], biological information [13], and memory-driven factors [14] were separately taken into account in the existing works. However, a joint approach is still a main challenge due to its complexity. As the second condition, early detection of QoE deterioration has not been carefully investigated yet in literature. This condition can be achieved by considering an appropriate set of observable QoE influence factors and suitable monitoring interval. There are factors becoming observable only after the introduction of the other one, thus, they are not suitable for the purpose of early detection of QoE deterioration. For example, rebuffering event only occurs after the drain on playback buffer. In addition to deciding monitoring factors, determination of suitable interval for QoE monitoring also need to be taken into consideration. Monitoring interval actually has a direct impact on the balance of network resource utilization and the resulting QoE. More concretely, if the measurement interval is set small, more computational power is required, but more importantly, a small measurement interval yields inaccurate results [15]. Thereby, the optimal monitoring intervals should be large enough to reduce the computational cost and improve the accuracy, but also small enough to early detect QoE deterioration.

In control layer, estimated QoE of each subsequent stream is compared with a certain QoE threshold. The difference is then sent to the controller for the generation of an appropriate control action/strategy. If the estimated QoE is smaller than a specific threshold, the controller can manage the different system components, allocate necessary resources, execute admission control, or implement other control strategies. Such mechanism yields optimized service delivery by delivering accurate control action at the right time. As the result, the end-users satisfaction will be maximized, while the limited network resources are optimally utilized.

- To generate the control action at the right time, an appropriate QoE threshold must be taken into account. To the best of our knowledge, there are no existing works which carefully perform the investigation of threshold in QoE control. The determination of QoE threshold usually depends on the types of QoE indicator used in assessment model. As mentioned above, either objective models or subjective models or hybrid models are possibly considered as assessment model in QoE monitoring. Among them, the hybrid QoE models which allow QoE estimation to be performed in an automatic and accurate manner in real-time, have gradually been the most common models in literature. In hybrid modeling process, the perceived video quality is subjectively evaluated by using rating approach in which QoE indicator in terms of 5-scale Mean Opinion Score (MOS) is given out. Therefore, in QoE control, the fair level of above indicators scale (the middle value) is simply selected as the threshold. However, the rating approach has significant drawbacks due to its high bias and variability, causing less accuracy in QoE modeling followed by unreliability of the selected threshold.

- Bandwidth allocation is one of the most common control strategies in literature. In general, it is a process of assigning bandwidth to users and applications based on their priorities. When an appropriate amount of bandwidth is allocated to the user, it allows the end-user to watch video with expected quality in addition to the optimization of network resource utilization. The question is how the appropriate bandwidth allocation can be accurately delivered. In [16], the authors proposed a novel method to allocate bandwidth to the end-user based on the target video rate of the next downloading video chunks. Typically, most of commercial video players always maintain a constant gap between the target video rate and the needed bandwidth. Thus, the suitable amount of bandwidth can be obtained if the target video rate is determined. However, the authors simply defined the target video rate based on the list of available video rates of the content at the server. It leads to the fact that the amount of allocated bandwidth was often higher than the need of the users, resulting network resource underutilization. Therefore, it is necessary to propose a new method to determine the target video rate.

1.3 Biological Information based QoE management

In order achieve the research goal through solving abovementioned issues, we propose a biological information based QoE management for adaptive streaming services shown in Fig.1.5. Thereby, there are multiple advantages found in proposed system as compared to the existences. First, a two-phase monitoring layer guarantees QoE deterioration will be early and accurately detected by jointly performing the observations of various QoE influence factors. More concretely, the first phase is concerned with early detection purpose by considering only network QoS factors (e.g., bandwidth, packet loss, delay and jitter). If the estimated QoE is less than the threshold, the controller will perform a suitable control action (e.g., bandwidth allocation). Alternatively, application QoS factors (e.g., rebuffering, initial start delays), biological information (e.g., skin conductance, heart rate and heart rate variability) and memory-driven factors are considered for the second phase. This phase actually provides an alternative QoE estimation in terms of feedback to the controller in order to confirm whether the previous control actions are sufficient. If the estimated QoE is still higher than threshold, there is no extra-operation generated by the controller. If not, the additional action will be triggered to achieve an expected high-quality level for the end-user. Note that the latter estimation is always produced in a longer delay than the former one. The study of second monitoring phase is out of scope of this dissertation, while the whose first phase is about to be presented. Second, relying on a biological information based QoE threshold, controller can trigger control action at the right time. This contributes also towards the research goal, that is to say, guaranteeing expected perceived video quality with minimum network resource. The threshold was derived from a natural logarithmic function which models the relation between human sensation and video quality. Third, by relying on a re-

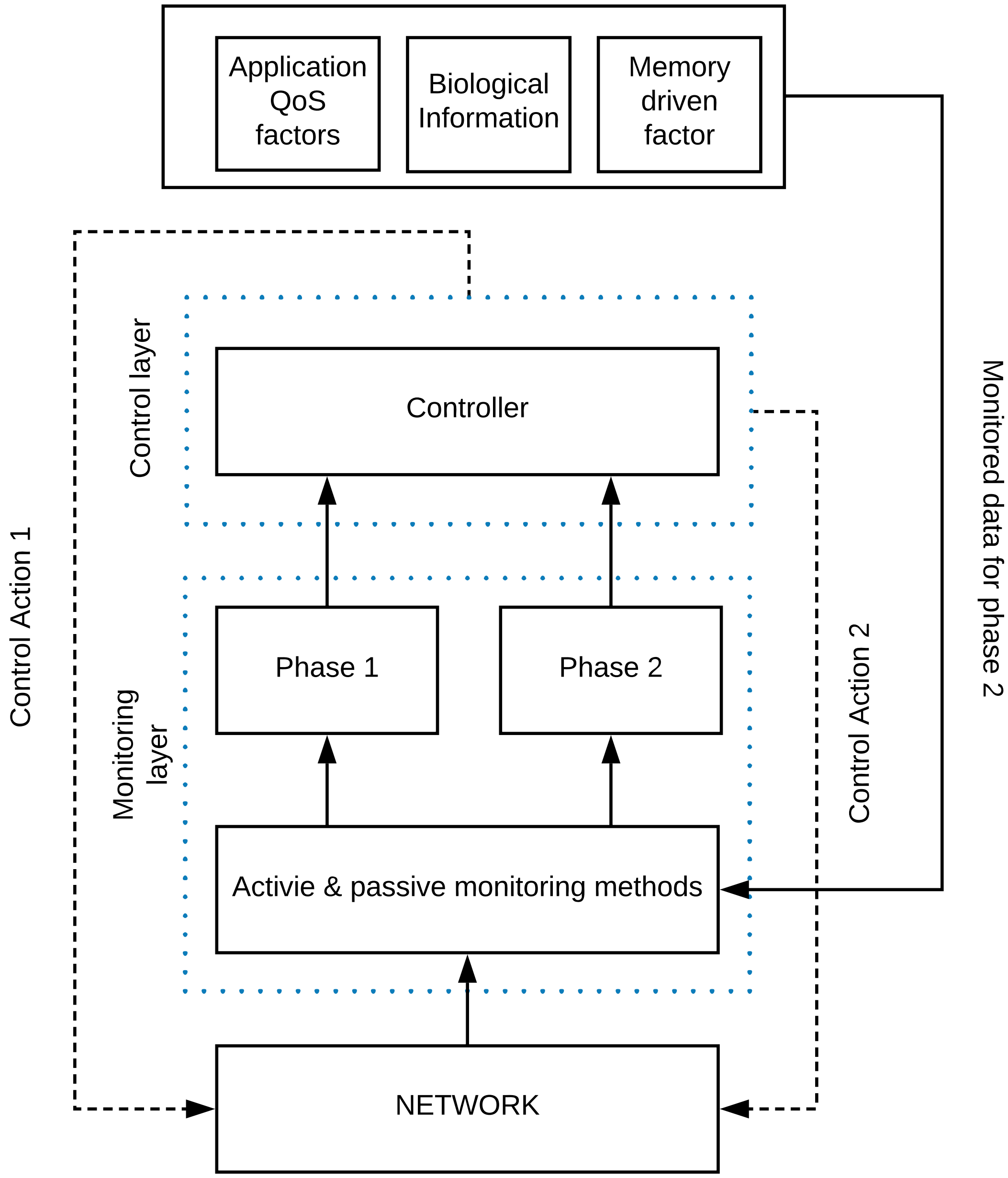

Figure 1.5: Proposed biological information based QoE management framework for adaptive video streaming services

gression model of subjective MOS and video rate, the target video rate can be determined, benefiting accurate bandwidth allocation.

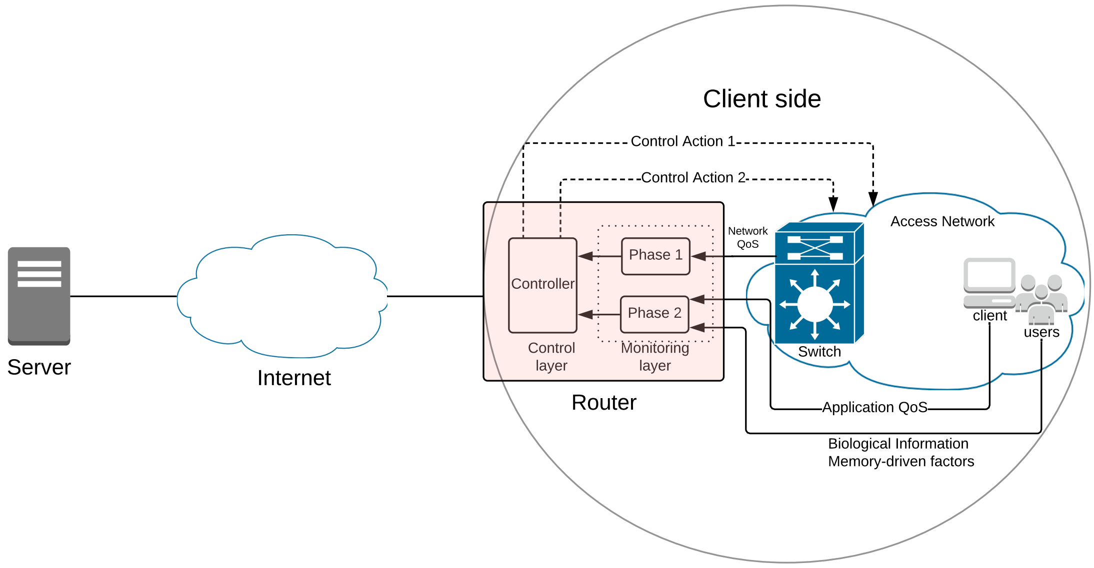

Figure 1.6: Deployment of proposed framework within practical environment

Figure 1.6 illustrates how the proposed framework can be deployed in practical environment. Accordingly, it is expectedly deployed within an open-source router that placed in distribution or core network layer at client side. Apart from basic functions of an edge router such as routing, traffic classification, this device is capable of carrying QoE management functionality. Thanks to active and passive monitoring methods, the data of network QoS, application QoS, biological information and memory-driven factors can be fed from access network to the router. There are two assessment/estimation models are respectively deployed within phase 1 and phase 2 of monitoring layer. While network QoS is used as training input data for the phase 1s assessment model, application QoS, biological information and memory-driven are applied for the one of phase 2. Based on the output of those assessment/estimation models, various decisions can be given out by the controller. The table 3.1 summarizes the possible decisions made by the controller.

The next sub-sections will respectively review the proposed solutions dealing with the abovementioned issues in both monitoring and control layers.

Table 1.1: X: means that estimated QoE is less than threshold. O: means that estimated QoE is higher or equal to threshold. Control action 1 is always produced earlier than control action 2

t_{Phase1}	t_{Phase2}	$t_{Decision}$
X	O	Control Action 1
X	X	Control Action 1 + Control Action 2
O	X	Control Action 2

1.3.1 Early detection of QoE deterioration with appropriate monitoring interval

As stated in previous section 1.2, one of the primary requirements for QoE monitoring is to early detect QoE deterioration, which can be accomplished by an appropriate monitoring interval. This sub-section reviews our proposed method in determining such the monitoring interval in order to achieve the research goal. The proposed method is two-fold. First, the suitable QoE influence factors are selected as monitoring factors. In this work, network QoS factors (bandwidth, packet loss, delay and jitter) were considered for early detection purpose. Second, based on the condition that keep playback buffer stable during streaming session, the monitoring interval is proposed to be equal to video chunks size.

There are numerous QoE influence factors that grouped into perceptual and technical categories, can be promisingly applied for QoE monitoring [11]. Because the end-user always experiences negative changes in video quality once perceptual factors are recognized, thus, for early detection purpose, the factor within technical category are more suitable. In this work, due to the fact that adaption logic plays an important role in adaptive streaming mechanism, the factors that belong to this category are investigated. Among them, only QoS monitoring can be flexibly monitored with arbitrary changeable interval, while the other parameters such as video rate and playback are only obtained in a chunk-by-chunk basic. Therefore, QoS parameters are proposed to be solely monitored for the early detection purpose of the first phase of monitoring layer.

During the streaming session, video player usually experiences two operating states: buffering-state and steady-state. In the first state, the player attempts to build up its playback buffer as quickly as possible and to reach a maximum buffer size. In order to achieve this, the player initially requests new video chunk with low video rate as soon as the previous one is downloaded. In the second state, the player aims to maintain a constant playback buffers size by requesting video chunk with constant interval. Note that the size of playback buffer can be calculated only when the video player requests video chunk to the server. Thus, it becomes a prominent parameter using in prediction of the properties of application QoS factors (e.g., rebuffering, stalling) that have direct impact on perceived video quality. In this research, with the assumption that the other application QoS factors (e.g., rebuffering, stalling) will not occur, video rate is assumed to directly reflect the QoE. Consequently, any obvious changes in playback buffers size will lead to the variation of video rate. Therefore, predicting the negative changes in playback buffer benefits the early detection of QoE deterioration. By studying the condition of a stable playback buffer, it was found that QoS parameters should be monitored with the interval being equal to video chunk size. A series of experiments were conducted to validate the determined monitoring interval. The results demonstrated that using this interval could achieve an expected balance of computational power and the resulting QoE (defined by the ratio of QoE deterioration).

1.3.2 Collaborative approach using psychophysiology and psychophysics for determination of QoE threshold

Traditionally, QoE threshold is simply picked up as the fair quality level of scale of QoE indicators (e.g., MOS) However, such the quality scale is established by human rating which has significant drawbacks due to the high bias and variability. For more precise QoE threshold, the drawbacks of rating approach need to be compensated. Thus, the combination of psychophysics and psychophysiology in determination of the appropriate QoE threshold, was deeply investigated in this research. This is because such the method can jointly explore the internal cognitive and tell us the truth about human perception for a given QoE change. Psychophysics is applied to quantitative evaluation, modeling the relation between a physical stimulus and perception level. According to the classical psychologists, the absolute threshold is the stimulus intensity at which the stimulus intensity changes become detectable. Thus, the desirable threshold is predicted to be equal to the absolute threshold. However, being similar to rating approach, the psychophysics also have drawbacks due to the fact that the methods of perception measurement do not provide sufficient insight into underlying perceptual and cognitive process. Because they rely on assessment scales and open-ended questionnaires. Meanwhile, psychophysiology is concerned with the measurement of physiological signals (e.g. Skin conductance, heart rate, etc.). In other words, it can detect the change of target stimulus through the change of physiological signal. For these reasons, this approach is prominent to be applied in perception measurement of psychophysics. Thereby, the relation between physical stimulus and human perception can be eventually modeled. However, individual difference is recognized as the primary shortcoming of this approach.

Therefore, in this research, a combination of psychophysics and psychophysiology was performed to leverage the advantages of both abovementioned approaches and compensate for their disadvantages. Accordingly, a logarithmic nature function expressing the relation between human perception (extracted from biological information) and stimulus intensity (defined as the deterioration of video rate) was established. As the result, an appropriate constraint of QoE threshold was derived from that modeling process. In order to confirm the hypotheses, a number of experiments were conducted. The experimental results demonstrated that using determined threshold constraint not only produces a high QoE but also saves more bandwidth per control.

1.3.3 User-centric approach to accurate bandwidth allocation

Although being difficult to be performed in complicated practical scenario, bandwidth allocation is still a common control strategy in QoE management proposals. In order to precisely determine allocated bandwidth, the constant gap between target video rate and required bandwidth has been considered in literature. This gap differs by the type of video player. For example, Microsoft Smooth Streaming player requires the bandwidth allocated to the users must be higher than the target video rate about 20%. However, the method to determine the target video rate has not been clearly stated. It was simply picked up from a pre-defined list of available video rate at the server without considering the end-users expectation, leading to over-provisioning bandwidth allocation. Therefore, it is necessary to precisely determine the target video rate in user-centric manner before allocating bandwidth to the end-user. In that situation, a novel method in bandwidth allocation has been proposed. This proposed method is two-fold. First, a pre-defined premium range of subjective expectation in terms of expected MOS is established. This range is from MOS threshold (determined in previous part of dissertation) to the highest MOS value of 5. Second, a regression model of video rate and subjective QoE is considered for interpreting the expected MOS to the target video rate. This model was established by performing function approximations using Gaussian radial basic function. The experimental results demonstrated that by using proposed method, more precise bandwidth allocation can be performed, resulting in more saving bandwidth per control.

1.4 Limitations

The subject of study for this dissertation is QoE management in adaptive streaming services. The focus of this study is to achieve the optimal balance of network resource utilization and the resulting QoE. However, the proposed method was validated in the scenario of experimental setup with several users only. The scenario of many users that share a common bottleneck has not been investigated yet.

In addition, there are numerous issues in maintaining perceived video quality in mobile networks (e.g. bandwidth fluctuation, interference, mobility, etc.), thus, our proposals were initially investigated in wired network only.

As mentioned in previous section, our research also concentrated on the application of biological information in managing QoE in HAS services. In the research, the collaborative approach using psychophysiology and psychophysics for optimal threshold determination in QoE Management was performed. However, the lack of the information associated with Central Nervous System (CNS) is another limitation in the scope of our work.

1.5 Contributions and Thesis Organization

In this dissertation, the proposed biological information based QoE management in adaptive video streaming services is presented. More concretely, three major

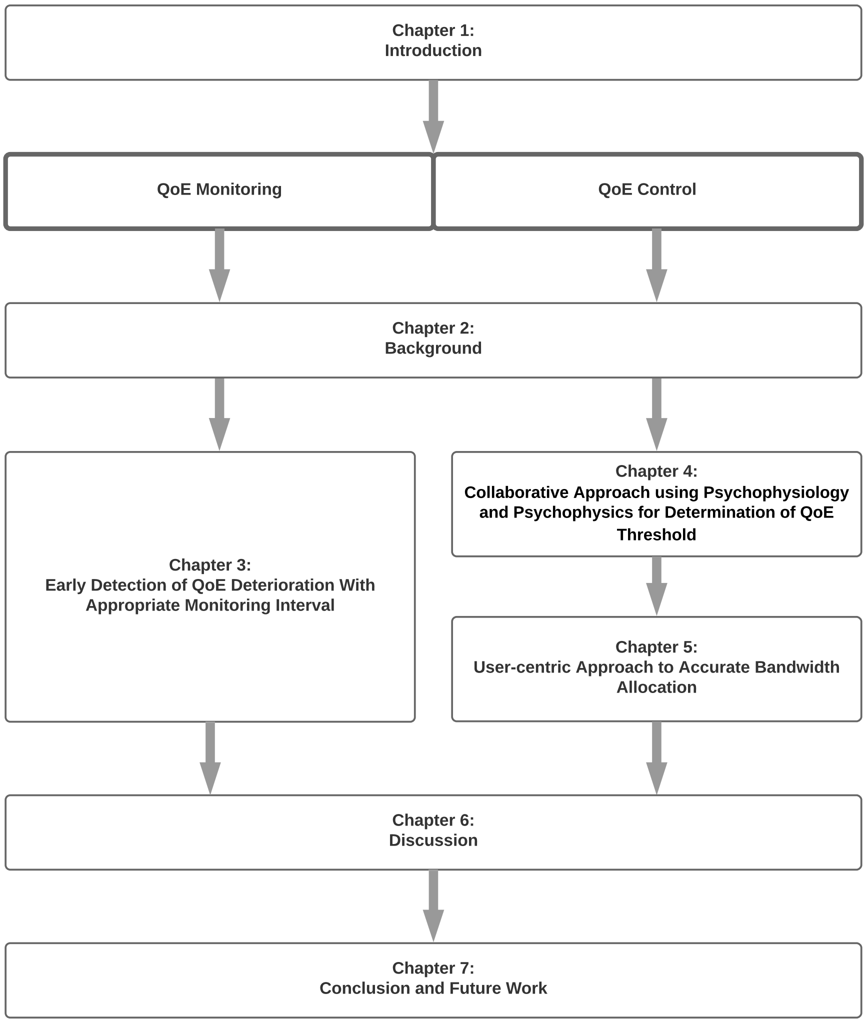

Figure 1.7: Organization of thesis

solutions have been proposed to address several issues related to QoE monitoring and QoE control respectively. The following list summarizes briefly our contributions to the problems stated in section 1.2:

1. A biological information based QoE management in adaptive video streaming services was proposed. Similar to general frameworks, the proposed framework has two main layers including monitoring and control layers. There are two major phases within monitoring layer. The first phase takes into consideration network QoS factors (e.g., bandwidth, packet loss, delay and jitter) as monitoring parameters for early detection of QoE deterioration purpose. Meanwhile, application QoS factors, memory-driven factors, and especially biological information take place within the second phase. Thereby, a feedback that accurately reflects the end-users satisfaction, will decide whether or not control action should be additionally generated. In control layer, an appropriate QoE threshold has been applied supporting an in-time control action. This threshold is determined by relying on modelling process of biological information. For accurate control action, the needed bandwidth will be calculated based on the end-users expectation.

2. An novel method of early detection of QoE deterioration was proposed. In order to achieve a quick detection, it is necessary to determine the monitoring factor and its monitoring interval. Thereby, video rate of playback video is maintained at an expected level. In fact, any major changes in playback buffer will produce an accordingly variation of video rate. Thus, the condition of a stable playback buffer had been investigated. Interestingly, it was found that the monitoring interval should be equal to the video chunk size. Applying the determined interval optimized the computational cost at the controller and eliminated the ratio of QoE deterioration.

3. An approach for determining the appropriate QoE threshold was proposed by taking into account the combination of psychophysiology and psychophysics. More concretely, a general logarithmic nature function which expresses the relation between biological information and the intensity of video rate, was introduced. Thereby, the optimal constraint of QoE threshold was determined and then applied in QoE control. As the result, not only the overall subjective QoE was maintained but also the network resource utilization was improved.

4. A novel approach for bandwidth allocation was proposed. In this approach, target video rate is obtained in a user-centric manner. More concretely, the target

video rate is determined from subjective expectation. Afterward, the needed bandwidth is then calculated and allocated to end-user. As the result, compared to the approach which used self-defined target video rate, this proposed approach not only maintained the expected QoE level for the specific user but also saved the available bandwidth per control.

The organization of this thesis is illustrated in Fig.1.7 and described as follows:

Chapter 1: Introduction. The motivation and background of this research were described in this chapter. In addition, research overview and research limitation also were figured out. The primary contributions of this research were also concretely summarized in this chapter.

Background knowledge is provided in chapter 2. The issues related to two QoE monitoring and QoE control are thoroughly resolved in detail from chapter 3 to chapter 5. Chapter 6 discusses, chapter 7 concretely concludes the work and figures out future work direction.

Chapter 2: Background. This chapter provides a wide range of background knowledge related to QoE, QoE assessment models and HTTP adaptive streaming. In this chapter, the definition of separate phenomenon of "quality" and "experience" will be initially clarified, followed by a unique definition of QoE. The numerous of QoE assessment models which play a key role in QoE management framework, are then presented. Typically, there are three types of assessment models including of objective models, subjective models and hybrid models. Their pros and cons are respectively presented. Finally, the hidden mechanism of HTTP adaptive streaming technology is discussed.

Chapter 3: Early Detection of QoE Deterioration With Appropriate Monitoring Interval. This chapter proposes a method to early detect QoE deterioration in QoE monitoring. Initially, the selection of monitoring factors is considered followed by the determination the most suitable factors for early detection of QoE deterioration purpose. By investigating the condition that keep playback stable during a streaming session, the monitoring interval is determined to be equal as the size of video chunk. The experimental results demonstrate that an optimal trade-off between computational cost and the maintaining QoE is achieved.

Chapter 4: Collaborative Approach using Psychophysiology and Psychophysics for Determination of QoE Threshold. This chapter proposes a new method to determine the optimal constraint of QoE threshold. The proposed approach is to combine psychophysiology and pyschophysics in order to establish a general logarithmic nature function expressing the relation between biological information and stimulus intensity (video rate deterioration). In this chapter, the issues related to the determination of QoE threshold, background of psychophysiology and psychophysics, and proposed method are briefly described. The evaluation results demonstrate that performing QoE management with determined threshold can save more than 4.855% of the bandwidth consumption per control, while ensuring a comparable video quality, in accordance with using the fair threshold.

Chapter 5: User-centric Approach to Accurate Bandwidth Allocation. This chapter proposes a new method to calculate the target video rate in performing precise bandwidth allocation in QoE control. In this chapter, a review on the bandwidth competition in HAS services and the proposed method were clearly presented, emphasizing the crucial role of bandwidth allocation in QoE management in adaptive streaming services. Afterward, the exiting studies in bandwidth allocation is described in order to investigate their pros and cons. Eventually, the proposed method and its evaluation will be clearly stated.

Chapter 6: Discussion. This chapter discusses the work investigated and solutions proposed in this dissertation by which advantages as well as the remaining issues will be summarized.

Chapter 7: Conclusion and Future Work. This chapter concludes the dissertation by which advantages as well as remained difficulties were discussed. Finally, research directions of great interest for the future work were figured out.

Chapter 2

Background

2.1 Overview of QoE

2.1.1 QoE Definition

In order to understand the definition of QoE, the phenomenon of "quality" and "experience" must be initially clarified. Quality is usually connected to the terms of perception. According to [17], quality refers to "the outcome of an individual's comparison and judgment process. It includes perception, reflection about the perception, and the description of the outcome. In contrast to definitions which see quality as "qualitas", i.e. a set of inherent characteristics, we consider quality in terms of the evaluated excellence or goodness, of the degree of need fulfillment, and in terms of a "quality event"". On the other hand, not only involving perception, quality is considered to connect to expectation as well. Some authors strongly indicate the correlation between quality, perception and expectation throughout the following definitions:

- "The feeling of high quality occurs when perception exceeds expectation; the feeling of low quality occurs when perception does not meet expectation" [18].
- "Degree to which a set of inherent characteristics fulfils the requirements", where the requirement is defined as need or expectation [19].

According to [17], experience is defined as follow: "an experience is an individual's stream of perception and interpretation of one or multiple event", where event is considered as "an observable occurrence. An event is determined in space (i.e. where it occurs), time (i.e. when it occurs), and character (i.e. what can be observed)". For instance, an experience might result from an encounter of a human being with a system, service or artifact. An experience does not encompass everything a person has undergone in the past, but this is referred to as a human influence factor on QoE.

After considering the phenomenon of "quality" and "experience, the definition of Quality of Experience (QoE) must be jointly defined. In general, "QoE measures the quality experienced while using a service" [20]. However, in this definition, the terms of perception and expectation have not been taken into consideration. On the other hand, the relation to the subjective perception of the user and its expectation is clearly evident throughout the definition of the European Network on Quality of Experience in Multimedia Systems and Services [17]: "**Quality of Experience (QoE)** is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state".

2.1.2 QoE Influence Factors

Any characteristic of a user, system, service, application, or context whose actual state or setting may have influence on the Quality of Experience for the user [17]. The influence factors (IFs) might be grouped in three categories - Human IF, System IF, and Context IF. More concretely, a human IF is "any variant or invariant property or characteristic of a human user. The characteristic can describe the demographic and socio-economic background, the physical and mental constitution, or the user's emotional state". System IFs refer to "properties and characteristics that determine the technically produced quality of an application or service [21]. Meanwhile, context IFs are factors that embrace any situational property to describe the user's environment in terms of physical, temporal, social, economic, task, and technical characteristics [21][22].

In video services, QoE influence factors can be categorized into technical and perceptual influence factors [11]. While the perceptual influence factors are directly perceived by the user, the technical influence factors are perceived indirectly. On the other hand, the authors in [12] constructed a protocol stack to form a conceptual relationship between QoS and QoE. Throughout this model, it is clear to find that the QoE influence factors can also be grouped into network QoS factors (are indirectly perceived by the user) and application factors (are directly perceived by the user). In that study, the network QoS factors comprising of (Round-trip time (RTT), bandwidth and packet loss) were considered, whereas, initial buffering time, mean rebuffering duration and rebuffering frequency were the focuses as the application QoS factors.

2.2 QoE assessment models

QoE assessment models refers to the translators between a set of technical (QoS) and non-technical (subjective and contextual) key influence factors and user perception, and ultimately, user experience [23]. These models can be categorized into three classes: Objective models, Subjective models and hybrid models. They will be briefly presented in the next subsections.

2.2.1 Objective QoE assessment models

Objective QoE assessment models are concerned with the models that contain objectively collected measurements of factors that affect QoE. By using these methods, the objective factors can be measured automatically at a lower cost than subjective methods. In general, objective quality assessment methodologies can be categorized into five types [24][25]. These are media-layer models [26], [27], parametric packet-layer models [28], parametric planning models [25], bitstream-layer models [29]and hybrid models [30][31][32].

- A media-layer model utilizes speech or video signals to predict QoE. Because it does not require a priori knowledge about the system under testing, such as the codec type or packet loss rate, it can be applied to the evaluation of unknown systems (e.g., codec comparison/optimization). However, by definition, it cannot be used in scenarios in which media signals are not available. For example, it is difficult to obtain media signals at the network mid-point although one can decode the payload of packets. These methods can be further categorized as full-reference (FR) [26][33][34][35], reduced-reference (RR), and no-reference (NR) [36] depending on whether a reference, partial information about a reference, or no reference is used in assessing the quality, respectively. Full- and reduced-reference methods are important for the evaluation of video systems in non-real-time scenarios where both the original video data or a reduced feature data set, and the distorted video data are available. Full-reference visual quality assessment metrics and high-complexity non-real-time RR and NR metrics fall within this class. On the other hand, the in-service methods place strict time constraints on the quality assessment and are performed during streaming applications.

- A parametric packet-layer model predicts QoE solely from packet-header information, enabling very lightweight measurement without handling the media signal itself. However, it was difficulty evaluating the content dependence of QoE, for example, because it does not look at the payload information.

- A Parametric planning model make use of quality planning parameters for networks and terminals to predict the QoE. As a result, it requires a priori knowledge about the system that is being tested.

- A bitstream-layer model occupies a position between media-layer models and parametric packet-layer models. It utilizes encoded bitstream information, in addition to the packet-layer models, so that it can take into account the content-dependent quality evaluation characteristics with a relatively light computational load.

- A hybrid model is a combination of the previously mentioned technologies. It is effective in terms of exploiting as much information as possible to predict QoE.

2.2.2 Subjective QoE assessment models

As the key component in QoE management studies, subjective QoE assessment refers to the quantifying the experienced quality of the users. In general, a panel of assessors (referred to as 'test subjects') is subjected to various quality levels which leads to some form of explicit or implicit response.

- **Rating approach** is the most commonly used method in subjective evaluation. Typically, the information regarding subject′s judgment in the form of rating that describe their perception of the respective quality experienced, is derived from this approach. This method has been standardized by the recommendations like ITU-R BT.500-11 [37], providing detailed guidelines regarding choice of test conditions, rating scales, etc. The most common grading scale for rating process is Mean Opinion Score (MOS) which is based on an ordinal five-point scale (1) bad; (2) poor; (3) fair; (4) good; (5) excellent. Despite being popular method, rating approach has significant shortcomings due to the high bias and variability in the results, leading to the less precise QoE assessment.

- **Psychophysiological approach** provides a measure of implicit rather than explicit responses to physical stimuli and thus overcomes the problem of rating approach [38]. More concretely, this approach utilizes the measurement of biological information in order to detect correlations to psychological responses in humans. The biological information measurements are categorized into the following classes [38] - Central Nervous System (CNS), Eye Measurements and Autonomic Nervous System (ANS).

- **Psychophysical approache** has been introduced to overcome the limitation of rating approach in QoE assessment by attempting to precisely express the relation between perception and physical stimuli. The ideas behind these approaches are either to estimate the parameter value for which the distortion becomes perceptible [39], to scale the relative differences perceived between physical stimuli [40], or to define the smallest detectable different between two stimulus's intensities [41]. Despite being applied in wide range of scientific fields, the accuracy of the estimation of perception

2.2.3 Hybrid QoE assessment models

In QoE management, monitoring component is usually required to be automatically performed in accurate and real-time way. In order to achieve this, the appropriate QoE assessment models have to be carefully taken into consideration. Looking back to previous descriptions of both objective models and subjective ones, it is clear that none of those models can completely fulfil the requirements of

QoE monitoring. Despite of the fact that objective models can be able to provide an automatic way to estimate QoE, the accuracy of such the estimations actually are not as high as those of subjective models. In addition, due to the requirement of reconstruction of video, the application of objective model in real-time is unrealistic. On the other hand, due to the fact that QoE estimation depends on real judgment of human, subjective methods cannot be performed in automatic and real-time fashion. In order to satisfy the original requirements, hybrid QoE assessment models have been introduced by mapping QoE influence factors (e.g. QoS) to the QoE in various ways. These methods leverage the advantages of both objective and subjective methods, while eliminating their drawbacks. QoE influence factors can be mapped to QoE by using the machine learning techniques [42][43][44][32] or by extending the existing objective methods with hybrid no-reference prediction model [45]. As the most common hybrid model, Pseudo Subjective Quality Assessment (PSQA) [31][32] was established by training a Random Neural Network (RNN) to capture the relation between QoS parameters and subjective evaluations by users. As the result, the trained network can be used for QoE estimation in automatic, accurate and real-time manners. For these reasons, hybrid model is sole consideration in QoE assessment in our studied framework.

2.3 HTTP adaptive streaming mechanism

This section provides an overview of HTTP adaptive streaming (HAS) as the most common video delivering techniques. This technique was introduced in 2008 [9] by Move Networks in order to overcome the limitation of existing adaptive streaming techniques as well as HTTP progressive download. Since then, it has quickly become a de-facto standard for adaptive streaming solutions.

The HAS framework [46] between a client and a video server is depicted in Fig.2.2. Initially, the video is partitioned in different small fragments (or chunks), typically a few seconds long. Each chunk is available at multiple video rate. The video chunks are then hosted on one or several media origin servers typically, along with the media presentation description (MPD) an XML metadata file that characterizes the structure and the features of the video presentations, and provides sufficient information to a client to request the appropriate video chunks to the server over HTTP. Actually, there are various information representing the video

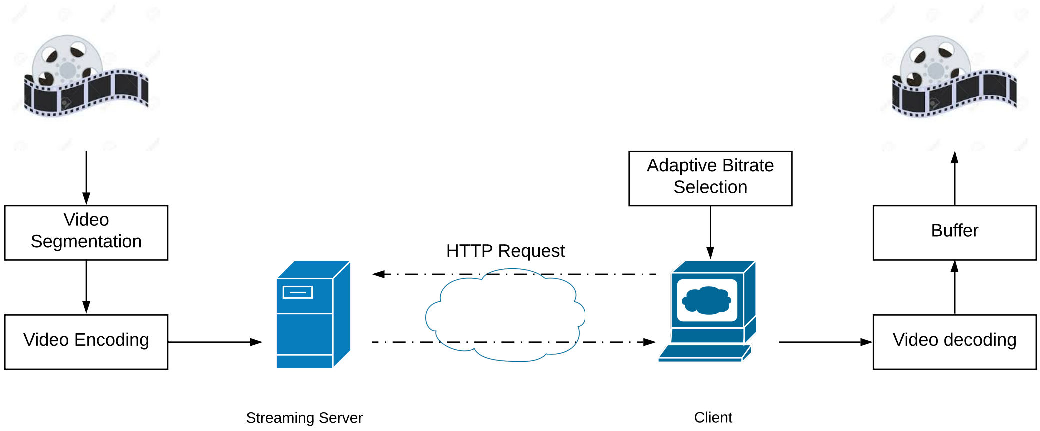

Figure 2.1: HTTP Adaptive Streaming framework

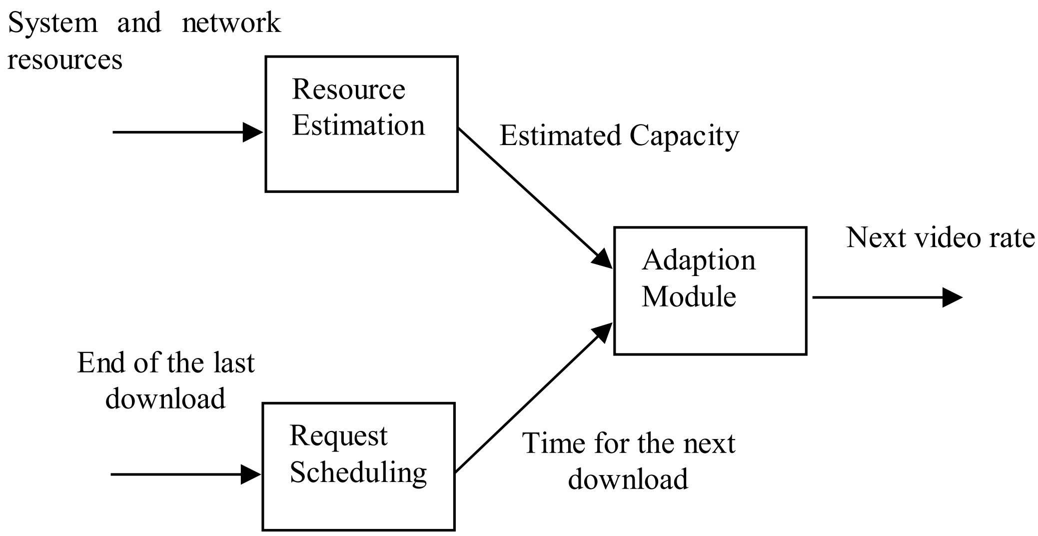

Figure 2.2: ABR framework comprises of three main components: Resource estimation, request scheduling and adaption module

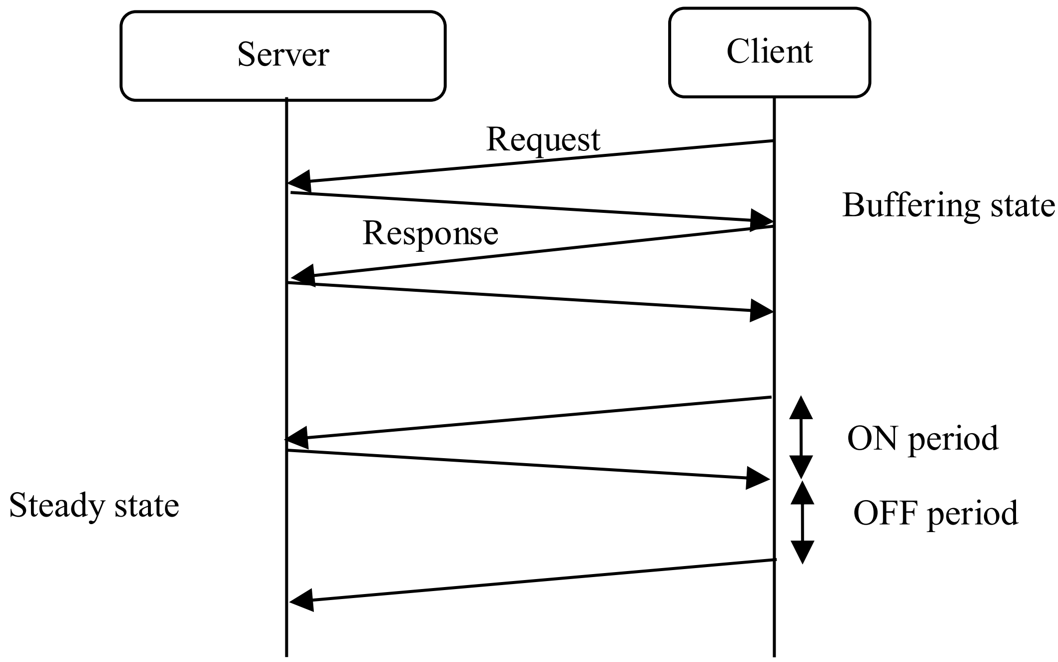

Figure 2.3: Buffering state and steady state in a streaming session

components (video rate, resolutions, the duration of each chunk in seconds, etc.) contained in MPD files. Based on such the information, clients request the video chunks corresponding to their selected representation using HTTP GET or partial GET methods with byte ranges. The most crucial mechanism that behinds this technique is the so-called Adaptive Bitrate Selection (ABR) [47], through which a client determines the profile and schedule of a chunk to download. The general architecture of ABR composes of three subcomponents: resources estimation, chunk request scheduling and adaption as shown in Fig.2.2 . In order to adaptive video quality to a context, network and other system performance parameters such as CPU, display size or battery life are measured. The choice of which parameter becomes a situational indicator depends on the QoE metric that an ABR intends to optimize. The ABR then uses the measurement result in making decision on the schedule and profile of the chunks to be downloaded. Typically, there are two main states in the operation of ABR.: **Buffering** and **Steady** state as shown in Fig.2.3. At the buffering state (or convergence time), HAS player attempts to establish playback buffer as quickly as possible by continuously requesting video chunks from the lowest video rate. A player normally does not wait until the end of the buffering state before a playback begins. This can be achieved when either a certain amount of content is downloaded or the buffer size reaches a predefined target (let say as B_{max}). For example, Microsoft Smooth Streaming has a playback buffer of about 20 seconds but starts playing when the buffer contains just about 10s worth of download [48]. Likewise, Netflix has a buffer size of 300s worth of content but begins playback 13s after receiving the first packet [48]. Afterward, then the steady state (or periodic download) is activated. In this state, HAS player attempts to maximize video rate by keeping playback buffer stable at B_{max}. To do so, the player is required to download a chunk and then pause for a short time before downloading the next chunk. The download period and pause period are called ON and OFF period, respectively. When stimulus occurs (e.g. available bandwidth deterioration), the buffering state will be re-activated.

2.4 Summary

In this chapter, the background of our study has been briefly presented. Initially, the phenomenon of both "Quality" and "Experience" was carefully clarified. After that, the definition of QoE which is in conjunction with the human perception and expectation, was discussed in detail. As the most important component in QoE management framework, QoE assessment models was also taken into consideration in this chapter. Even though only hybrid QoE assessment models have been applied in this research, the other types of models were also investigated in order to emphasize the reliability of hybrid models. Finally, the framework of HAS as well as adaptive bitrate selection (ABR) were the next focuses of this chapter, respectively.

Chapter 3

Early Detection of QoE Deterioration With Appropriate Monitoring Interval

Developing the QoE management process is a non-trivial exercise since capturing QoE is a very subjective process. QoE is subjective because it is driven psychologically as well as technically. Thus, monitoring layer in QoE management system where QoE estimation takes place should be initially discussed. For this reason, in this chapter, the monitoring layer will be thoroughly explained alongside a proposed method for early detection of QoE deterioration. Thereby, the research goal, which achieves the balance between optimizing network resource utilization and maintaining QoE, can be made.

3.1 Introduction

For years, QoE monitoring has become a major concern of contemporary works [49][50][51][52][53][54]. In general, QoE monitoring is performed by observing a wide range of QoE influence factors or quality performance indicators (QPI) in real-time, then by interpreting them as the QoE indicator that is expressed by the Mean Opinion Score (MOS) ranging from 1 ("Bad") to 5 ("Excellent") [55]. In fact, QoE assessment and QoE estimation are extremely important components in any designs of QoE monitoring. While QoE assessment refers to modeling process of QoE influence factors and subjective perception, QoE estimation refers to the use of that model to automatically estimate QoE in real-time. The primary requirements of QoE monitoring are to early and accurately detect QoE deterioration based on reliable estimations during a streaming session, serving the purpose of maximizing QoE with minimum network resource usage. In order to achieve this, a two-phase monitoring layer has been introduced in the proposed biological information based QoE management framework. Fig. 3.1 depicts the proposed QoE monitoring layer comprising of two monitoring phases. The focus of the first phase is to early detect QoE deterioration to avoid the perceivable video quality distortion during a streaming session. Meanwhile, the second phase refers to the validation of the first phase with more reliable QoE estimation followed by additional control if it is necessary. In this dissertation, the first monitoring phase is focused on, whereas the second phase will be considered as the future work.

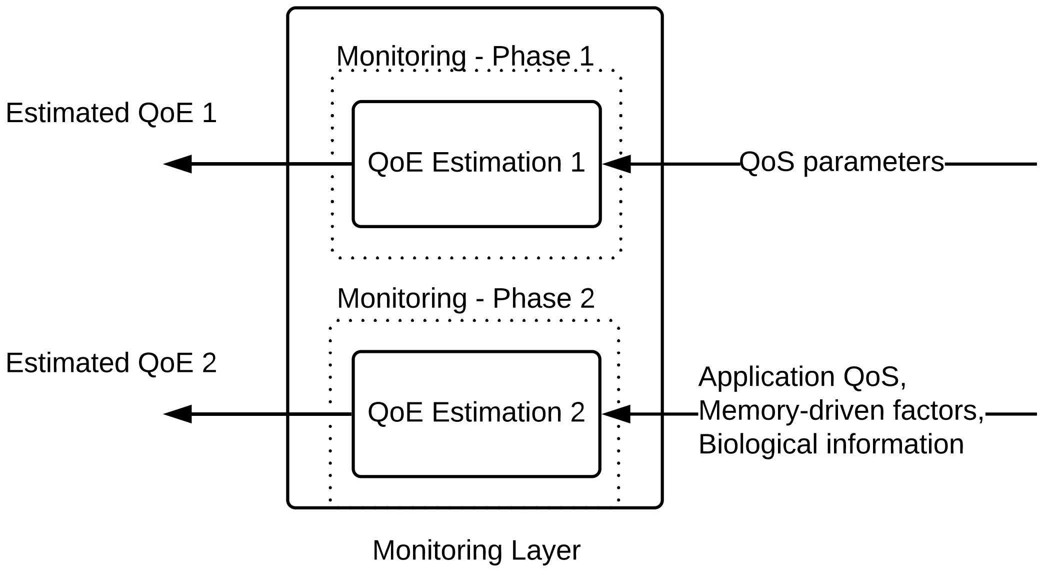

Figure 3.1: Two-phase QoE monitoring layer

In a design of QoE monitoring, the following steps need to be in turn performed: (1) Selection of appropriate monitoring factors and (2) Selection of a suitable monitoring interval. In this study, such a design needs to be aligned with the requirement of the first phase, that is to say, early detection of QoE deterioration.

Monitoring factors can be selected from among numerous QoE influence factors. According to [11], QoE influence factors can be divided into two categories

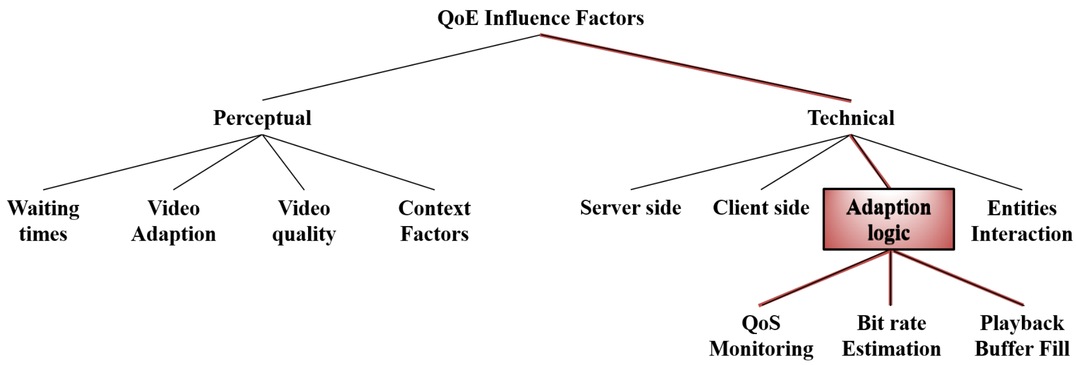

Figure 3.2: Classification of QoE influence factors in adaptive streaming services

which are perceptual factors and technical ones as shown in Fig. 3.2. While the perceptual factors are directly perceived by the users, the technical factors indirectly reflect the perceived video quality. A similar classification approach was also delivered as the protocol stack characterizing a conceptual relation between QoS and QoE in [12][56]. Accordingly, network QoS (e.g. bandwidth, packet loss, delay and jitter) and application QoS (e.g. initial buffering time and rebuffering frequency) are referred to technical and perceptual factors, respectively. For early detection purpose, the technical factors are more suitable. Among them, the factors (playback buffer, video rate and QoS), which belong to adaption logic category have been major concern in this research due to their strong correlations with adaptive video rate mechanism. It is worth noting that both the playback buffer and the video rate are grouped into application QoS category, while the QoS belongs to network QoS category. As mentioned in section 2.3, during **buffering state** and **steady state**, adaptive video player always attempts to quickly establish a sufficient playback buffer and to keeps it stable. Once a drain on playback buffer becomes detectable, video rate will be adaptively varied to avoid video interruption. Therefore, the status of playback buffer potentially determines the variation of video rate and other application QoS parameters. Being the most important application QoS parameter, the video rate expresses the quality of picture frame and directly reflects the video quality perceived by the end-user. With an assumption that playback buffer is sufficient to avoid the distortion caused by the other application QoS parameters (e.g., buffering, stalling, frequency of rebuffering), video rate becomes a sole representative of perceived video quality. Thus, the requirement of the first monitoring phase turns out to be early detection of video rate deterioration. This can be done by observing the status of playback buffer. Because both the playback buffer and the video rate are categorized into application QoS, thus, they are affected by network QoS parameters. In other words, network QoS can be regarded as a situational indicator showing the variation of playback buffer and video rate. Monitoring network QoS provides a prediction of playback buffer, benefiting the purpose of early detection of video rate deterioration.

In adaptive streaming technology, both playback buffer and video rate are typically obtained on a chunk-by-chunk basic [57][58] that relies on the timestamp of two successive requests sent by video player. In other words, the monitoring interval of those parameters are uncontrollable. It leads to the fact that their deteriorations have already been perceived by the end-user before control action is triggered. As a potential monitoring factor, on the contrary, network QoS can be captured with more flexible self-defined interval. To the best of our knowledge, there are no studies that seriously focus on monitoring interval in QoE management. In literature, the existing works usually come up with a very small monitoring interval. However, small interval always causes high computational cost on the network entity in which monitoring component is being deployed in, whereas, long interval leads to high ratio of deterioration of video rate. Therefore, in this chapter, a novel method is proposed to early detect QoE deterioration through the determination of monitoring factor and the appropriate monitoring interval. This proposal is inseparable from an expected balance of the computational cost and the ratio of video rate deterioration. Particularly, the monitoring interval is derived from a considered condition which makes playback buffer stable during a streaming session. Accordingly, the optimal interval is determined as being equal to the size of a video chunk. The experimental results demonstrate that by applying determined interval, a balance between the average CPU load and the ratio of video rate deterioration is achieved at the value of 11.45%.

3.2 Related Work

This section will review the existing works related to QoE monitoring in order to emphasize the necessity of the determination of monitoring factor and an appropriate monitoring interval.

In [59] [60] video rate was monitored to evaluate the performance of the proposed video quality adaption scheme. Thereby, these authors confirmed that their proposal was more advanced than other works in terms of justifying QoE defined by monitored video rate. However, in [48], the authors stated that it always takes time for video rate to adapt to the network condition. Thus, such a control action is meaningless if it relies on video rate monitoring since the video deterioration has already been perceived by the end-user. Both video rate and playback buffer can be obtained on a chunk-by-chunk basic [57][58], thus, the same consequence is also found in playback buffer monitoring [61][62]. In that case, only QoS parameters are prominent for the original purpose of early detection of QoE deterioration.

Network QoS is considered as a monitoring factor in a number of contemporary studies [63][49][31][64][65]. Using the network QoS parameters, the monitoring interval can be flexibly self-defined without depending on chunk-by-chunk basic. It is also well-suited for quickly predicting video rate deterioration [66][67]. If monitoring interval is too small, the computational cost of Controller becomes higher. Additionally, incorrect control action will also be generated due to the spike fluctuation in traffic throughput. Particularly, if the bandwidth instantaneously deteriorates, then recovers, the activation of control action becomes meaningless. If the interval is too large, video rate deterioration might be perceived before the generation of control action. Therefore, an appropriate monitoring interval needs to be determined to optimize computational cost and eliminate the ratio of video rate deterioration.

In this chapter, the proposed method for determining an appropriate monitoring interval will be presented. Along with that a series of experiments have been performed in order to validate the determined interval. The experimental results demonstrate that using determined monitoring interval, a balance between the computational cost and the ratio of video rate deterioration has been achieved.

3.3 Methodology

Typically, the operations of QoE monitoring can be broken down into two steps: QoE assessment and QoE estimation. In this section, the brief description of QoE assessment model that is used in this study will be initially introduced.

Afterward, the method that relates to QoE estimation is proposed for the determination of appropriate monitoring interval.

3.3.1 PSQA approach in QoE assessment

In literature, hybrid QoE assessment has emerged as the most common model in expressing the relation between QoE influence factors and QoE indicators. In this study, Pseudo Subjective Quality Assessment (PSQA) is mainly taken into account due to its advances [32]. This model is capable of estimating QoE in accurate and automatic manner and if necessary in real-time. Particularly, the model was established by training a Random Neural Network (RNN) to map QoS to QoE. The training process and operations of PSQA are depicted in the Fig.3.3 and Fig.3.4, respectively.

In training process, a dataset comprising of QoS parameters (as input) and subjective Mean Opinion Scores (MOS) (as output), has been prepared. Input data for training comprises of various QoS parameters including available bandwidth, packet loss, delay and jitter. For each selected input parameter, a discrete set of common values was chosen. Each combination of values of QoS parameters is called as a system configuration. There were 294 prepared configurations in total set up on WANEM router (WAN Emulator 3.0). The open source movie, Big Buck Bunny had been chosen to be watched by the subjects. The movie was cut into 10-second sequences each of which was available at multiple video rates at the server. Accordingly, there were totally 294 sequences taken from original movie and they were delivered over the set-up network with different configurations. Consequently, the distorted video sequences were also obtained. Output data for training process was subjective perception represented by MOS scores ranging from 1 to 5 given out by subjects for each distorted sequence. MOS was obtained following the instruction of Degradation Category Rating (DCR) methodology [68]. 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 were 17 subjects who were asked to watch 294 sequences and deliver their evaluations in terms of MOS, afterward. Finally, the average of those subjective evaluations was obtained. After preparation step, dataset with 294 samples was 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 was also considered, which comprises of totally 4 neurons for input layer, 10 neurons for hidden layer and 1 neuron for output layer. The training process is shown in Fig.3.3. As the result, correlation coefficient (denoted by R) was calculated, which is equal to 0.91. It means that the predicted data has a well correlation with actual data. Therefore, the trained neural network can be used for MOS estimation.

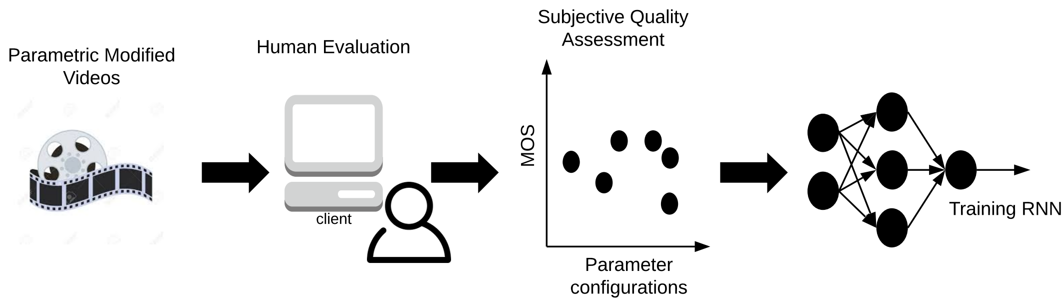

Figure 3.3: Training process of PSQA

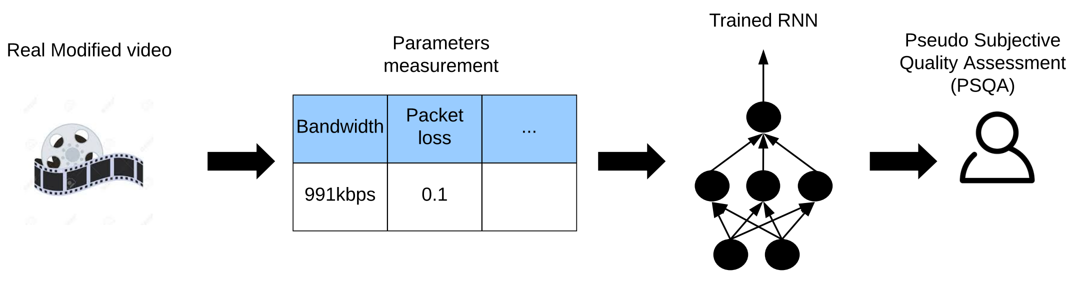

Figure 3.4: Practical usage of PSQA in QoE assessment

3.3.2 Determination of appropriate monitoring interval

The high correlation coefficient obtained from training process allows the use of PSQA assessment model in QoE estimation. Particularly, MOS as QoE indicator is continuously estimated from monitored QoS parameters (bandwidth, packet loss, delay and jitter). For the research purpose of early detection of video rate deterioration, the estimated MOS must precisely reflect the status of video rate in real-time. This can be achieved if MOS is estimated at the right time, depending on QoS monitoring interval. In fact, the smaller QoS monitoring interval is, the earlier detection of video rate deterioration can be guaranteed. However, using too small monitoring interval might cause some consequences. For example, some spike fluctuations of QoS parameters will lead to incorrect estimation, resulting in meaningless control action. In addition, the computational cost of monitoring entity is also a big issue for the large network system. Therefore, monitoring QoS with an appropriate interval is extremely important, especially for precisely capturing the status of video rate. Before going further, it is necessary to understand how video rate varies during a streaming session.

Looking back in adaptive streaming technology, the original video content is always required to be divided into multiple chunks which are available at multiple video rate at the server. Before streaming session is started, the server will initially deliver the signaling metadata or media presentation description (MPD) that contains the characteristics of the video chunks (such as video rate, resolution, etc.) to video player. Based on the MPD and the status of network conditions, the video player makes decisions for video rate selection. The adaptive bitrate selection framework has already been presented in subsection 2.3, describing how the video rate for the next video chunk can be estimated. Accordingly, it can be seen that resource estimation always takes a primary role in adaption process. Therefore, in order to maximize video rate, the video player needs to correctly estimate the resource availability and resource demands [69]. The same indication can be found in an adaptive bitrate selection (ABR) survey [47]. In fact, the resource estimation is performed by considering the estimation of either throughput or buffer occupancy. However, throughput-based ABR is the most commonly used approach in commercial players. This research solely focuses on throughput-based ABR approach. Particularly, throughput-based ABR takes into account the estimated throughput and the alternative video rate which are specified in the metadata. Then, the video rate can be decided as the highest value of the available video rate that is smaller than the estimated throughput. Traditionally, throughput is estimated based on per-chunk mechanism, which uses the throughput of a recently downloaded chunk as a rough estimate of the current network conditions [70][71]. However, the instant throughput derived from a single chunk is hardly used since it is prone to short-term fluctuations as result of, for instance, the time-varying nature of the available bandwidth, or the dynamics of TCP. In order to address this problem, the concept of running average was introduced as follows:

$$\overline{\mathrm{T}} = \begin{cases} \alpha T(i-1) + (1-\alpha)T(i), & i > 1 \\ T(1), & i = 1 \end{cases} \tag{3.1}$$

where $T(i)$ is the throughput of the ith chunk, $\overline{\mathrm{T}}$ is the running average.

However, when video players have to compete for available bandwidth, the operations of ON and OFF period will cause unfairness and instability in video rate selection [72]. It means that sufficient knowledge about the status of video rate cannot be provided only by throughput estimation. As mentioned in subsection 2.3, playback buffer plays a central-role within video rate adaption mechanism. In general, the video players always attempts to maintain playback buffer size at a stable level, resulting in high requested video rate. In addition, playback buffer can be easily obtained in application layer despite bandwidth competition [61]. Therefore, it promisingly provides an accurate prediction of video rate variation. In order to practically confirm this indication, an experiment was conducted. The experimental scenario was as follows: The end-user watches a movie with a high video rate under a good network condition in which available bandwidth is high (around $5000kbps$), whereas, packet loss, delay and jitter are assumed to be negligible. The behaviors of both playback buffer and video rate are continuously observed when:

1. The available bandwidth is dramatically decreased to $1024kbps$ at $t = 20s$ (before playback buffer reaches B_{max}).
2. The available bandwidth is dramatically decreased to $1024kbps$ at $t = 60s$ (after playback buffer reaches B_{max}).

There were two evaluation metrics that were considered in this experiment which are $t_{delay-buffer}$, and $t_{delay-bitrate}$, that is to say, the duration time until the first adaption (either decrease or increase) of both playback buffer and video rate, respectively. The details of experimental setup were as follows: there were three major entities including a client, a streaming server and a router. Microsoft smooth streaming player and a packet sniffer (Wireshark) were deployed at client. Wireshark allows us to capture and analyze the traffic which comes from and to HTTP server offline. The router, namely, WAN Emulator is capable of controlling the available bandwidth of the client. During the experiment, the video rate was derived from HTTP GET packet header, whereas, the playback buffer was calculated through Eq.3.2 as follows:

$$B_{t_k} = B_{t_{k-1}} - \Delta_t + (t_k^{'} - t_{k-1}^{'}) = B_{t_{k-1}} - \Delta_t + V \tag{3.2}$$

where B_{t_k} is playback buffer size at time point t_k, $B_{t_{k-1}}$ is playback buffer size at time point t_{k-1}, whereas, $t_k^{'}$ and $t_{k-1}^{'}$ are timestamp of HTTP video request at time point t_k and t_{k-1}, respectively, Δ_t is the duration time between two successive requests. V is equal to video chunk size (in second). The specific value of V depends on the type of adaptive streaming player.

Table 3.1 shows the sample dataset of experiment with two studied metric $t_{delay-buffer}$ and $t_{delay-bitrate}$. The means of the waiting time until the first negative adaptions of both playback buffer size and video rate are respectively 5.76s and 12.69s, respectively. Interestingly, during the experiment, the 2nd decrease of playback buffer always occurs at the same time with the first decrease of the video rate. Therefore, capturing the first decrease of playback buffer provides a prominent prediction of the decrease of video rate. For this reason, it is necessary to investigate the condition that keeps playback buffer stable during a streaming session. In fact, such the condition can be expressed as Eq. 3.3.

$$B_{t_k} - B_{t_{k-1}} \geq 0 \tag{3.3}$$

Consequently, based on Eq. 3.2, the above condition is transformed as:

$$\Delta_t \leq V \tag{3.4}$$

Accordingly, if the condition in Eq. 3.4 is guaranteed, it will prevent video rate from deteriorating in a streaming session. Particularly, the video player must keep sending requests with the interval ($\Delta_{t(s)}$) which is higher or equal to the video chunk size (V). Therefore, QoS must be monitored with an interval that is at least equal to V to precisely capture of the first decrease of playback buffer.

Table 3.1: Sample dataset with two metrics: $t_{delay-buffer}$, and $t_{delay-bitrate}$

$t_{delay-buffer}$	$t_{delay-bitrate}$
5.37	14.95
4.01	5.37
3.52	17.31
5.73	11.62
4.81	11.56
7.41	12.54
5.80	13.6

3.4 Evaluation

The purpose of this evaluation is to verify how elaborately the proposed monitoring interval facilitates maintaining the video rate level when the network condition is getting worse. More concretely, since the determined interval of MOS monitoring is applied, the following metrics has been evaluated:

- Ratio of video rate deterioration.
- Average CPU load.
- Detection time t_d which represents how quickly video rate deterioration can be detected if compared with method which uses video rate as monitoring indicator.
- Recovery time t_r of video rate which represents the duration time from when control action is generated until video rate is recovered to expected level.

In order to evaluate those criteria, two experiments were performed with environment setup as follow: A TestBed consisted of a router, a streaming server, and a client. Beyond routing and Nat function, the router played a role as a Controller which was installed on a VMware workstation of a desktop computer with Intel Core i5 3.10 GHz processor and 8 GB RAM. The Controller with QoE management algorithm (written in Python) [65] was capable of not only monitoring and controlling QoS data (available bandwidth, packet loss, delay, and jitter), but also calculating MOS based on QoS data. 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 (MSS) video content of "Big Buck Bunny" which is known as an open source testing movie. This movie content was en-coded 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 clients can play Smooth Streams. A video client was a laptop computer with MacOS, Core i5 and 8 GB RAM in which the latest version of Microsoft Silverlight add-on was installed. The server and the client were located in different broadcast domains and they were connected via the router. The network topology used for the experiments is shown in Fig. 3.5. In addition, Wireshark, which is a network packet analyzer, installed on the router captured the HTTP request from the client. Note that MSS applies the value 2s of V during streaming session [48], thus, in this experiment, the optimal interval of 2s was evaluated.

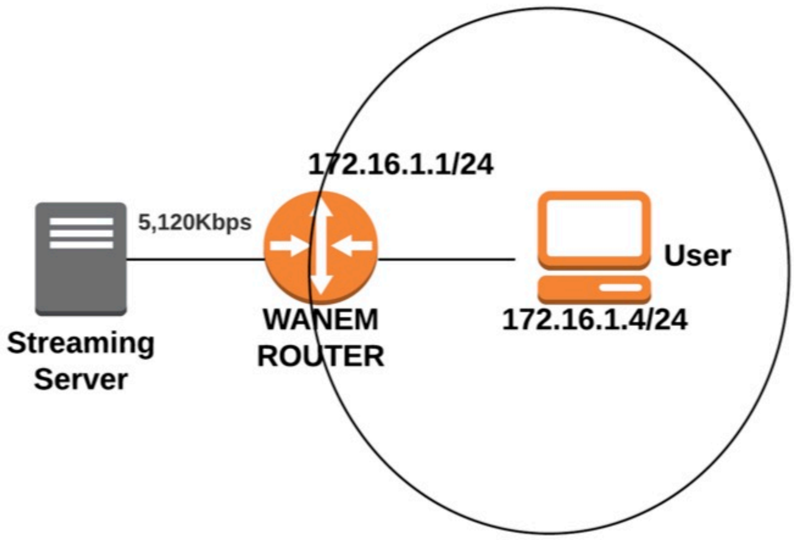

Figure 3.5: Experimental setup for evaluating the optimal monitoring interval throughout three evaluation metrics

For evaluating two first metrics, the experimental scenario was performed as follow: the estimated MOS was monitored with respect to interval $t_{mon} \in \{1, 1.2, 1.5, 1.8, 2, 2.2, 2.5, 2.8, 3, 3.2, 3.5\}$. Meanwhile, the experimental procedure was:

1) A client starts watching a streaming video content.

2) Stimulus is generated in buffering state and steady state by decreasing available bandwidth on purpose to make the network quality deteriorated (from $5000kbps$ to $1024kbps$).

3) The packet loss, delay and jitter in the network and average CPU load in Controller (where QoE monitoring and QoE control are performed) are observed.

4) The deterioration is detected by observing the estimated MOS.

5) The available bandwidth to the user is immediately increased to recover the net-work quality when the deterioration of video rate is detected (from $1024kbps$ to $5000kbps$).

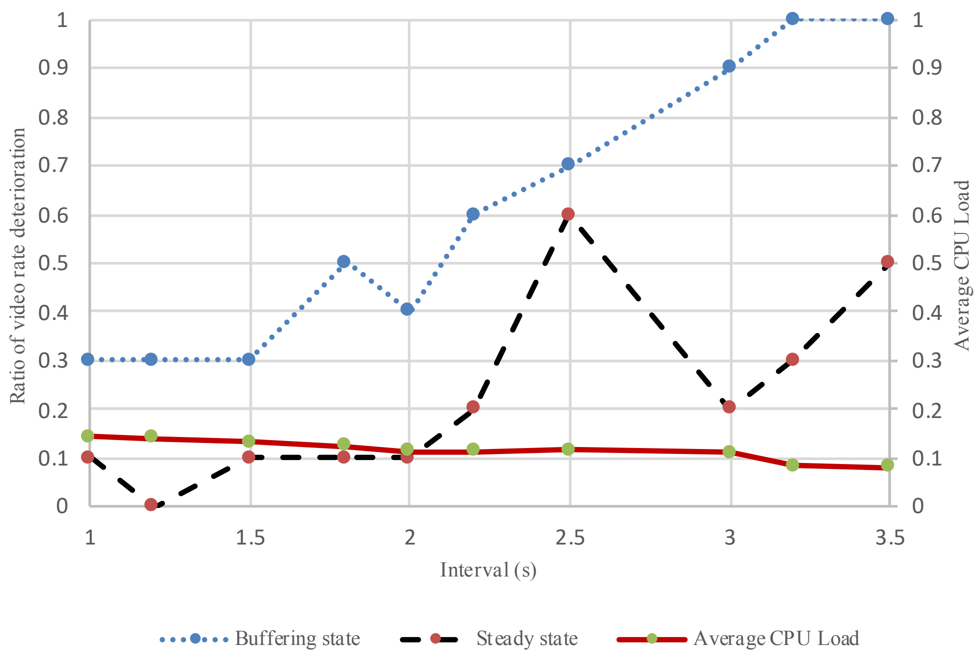

Figure 3.6: Ratio of QoE deterioration and average CPU load in both scenarios

Ratio of video rate deterioration is determined by ratio of the number of times the video rate decreases to the total number of times the experiment is repeated. Meanwhile, average CPU load stands for means of CPU load of the Controller in each experiment's iteration. Particularly, with each value of t_{mon}, the above procedure was repeated 10 times in total. Given that within 10 times, there is n times the video rate decrease $n \leq 10$, even though control action has already been generated. Then, the ratio of video rate deterioration which is the ratio of n to 10 times of total was calculated for each value of t_{mon}. Alternatively, the average CPU load of the Controller for each interval was also recorded.

Figure 3.6 compares the ratio of deterioration of video rate according to the monitoring interval varying from 1s to 3.5s with both buffering state and steady state. It is clear that those ratios significantly increased when $t_{mon} > 2s$. Overall, a much higher percentage of video rate deterioration could be seen in buffering state in comparison with steady state, and buffering state experienced the faster growth of such ratio. As explained in background knowledge section, during the buffering state, HAS player attempts to fill the playback buffer as quickly as possible. Whereas, during the steady state, buffer occupancy is stable at B_{max}. Therefore, video rate becomes more sensitive to stimulus within buffering state than in the steady state. In this figure, during the streaming session, average CPU load showed a clear trend in which it linearly decreased across monitoring interval values from 14.46% to 8.18%.

Particularly, during the buffering state, an increase trend clearly could be seen in ratio of video rate deterioration when the monitoring interval was higher than 2s. A slight fluctuation was found in range of between 1.5s and 2s. However, such fluctuation did not always occur when the whole procedure was repeated several times. Interestingly, the ratio reached to peak of 100% of video rate deterioration when monitoring interval is larger than 3.2s. When monitoring interval was varied from 1s to 2s during steady state, the ratio of video rate deterioration was stable at lowest value of 0.1 of accuracy. However, when the monitoring interval was larger than 2s, the ratio of video rate deterioration quickly rocketed to 0.6 of accuracy before witnessing a large fluctuation in range of between 2.5s and 3.5s. This fluctuation was also explained as the result of limitation of this QoE management algorithm performance. The algorithm frequently called PSQA model (written in Matlab) by which it could generate some "spike" in Controller's processing time. Actually, this abnormal fluctuation could not be seen when the experiment procedure was repeated several times.

The reasonable decrease trend of average CPU load was found from the graph. Interestingly, the line of average CPU load crossed by the line of ratio of video rate deterioration (in the steady state) at the point according to the interval of 2s. At that point, the value of computational cost and the ratio of video rate deterioration are equal to 11.45%.

For the detection time and recovery time criteria, MOS monitoring with defined optimal interval was compared with video rate-based method. The experimental procedure for two scenarios of the evaluation was as follows:

1) A client starts watching a streaming video content.

2) The available bandwidth is reduced on purpose to make the network quality deteriorated.

3) The packet loss, delay and jitter in the network are observed.

4) The deterioration is detected by observing the video rate and the estimated MOS.

5) The available bandwidth to the user is increased to recover the network quality when the deterioration of the video rate (for the first scenario) and estimated MOS (for the second scenario) are detected.

Initially, the capacity of the link from router to server was set to 5000$kbps$. Because there was only one client in the network, thus, the link capacity was equivalent to the available bandwidth of the client. The experiment time was 120 seconds for each scenario. At t=20s, t=60s and t=90s, the available bandwidth of the client was set to a low level of 1024 kbps. During streaming sessions, video rate was continuously captured, whereas, the estimated MOS was monitored in every t_{mon}=2s.

Figure3.7 and Fig.3.8 show the results of experiment in both scenarios. As seen from both graphs, the video rate reached its highest value of 2962$kbps$ at around t=10s. After the available bandwidth was reduced to 1024$kbps$ at t=20s, the video rate decreased to 2056$kbps$ at t=32.46s. Router was immediately controlled to increase the available bandwidth to 5000$kbps$. However, the video rate did not return to 2962$kbps$ within several seconds. It stayed at the value of 2056$kbps$ for 15s. When the available bandwidth was decreased at t=60s, the video rate also took a large delay to react to. It decreased to 2056$kbps$ at 71.99s, and even kept staying at that level, although the router had increased the available bandwidth to 5000$kbps$. To make matters worse, when the available bandwidth was reduced to 1024$kbps$ t=90s, the video rate started decreasing more.

In Fig.3.8, after the available bandwidth was reduced to 1024$kbps$ at t=20s, t=60s, and t=90s, MOS quickly decreased to around 2.75. Those deteriorations were respectively detected at t=23.90s, t=62.90s and t=92.90s, respectively. The

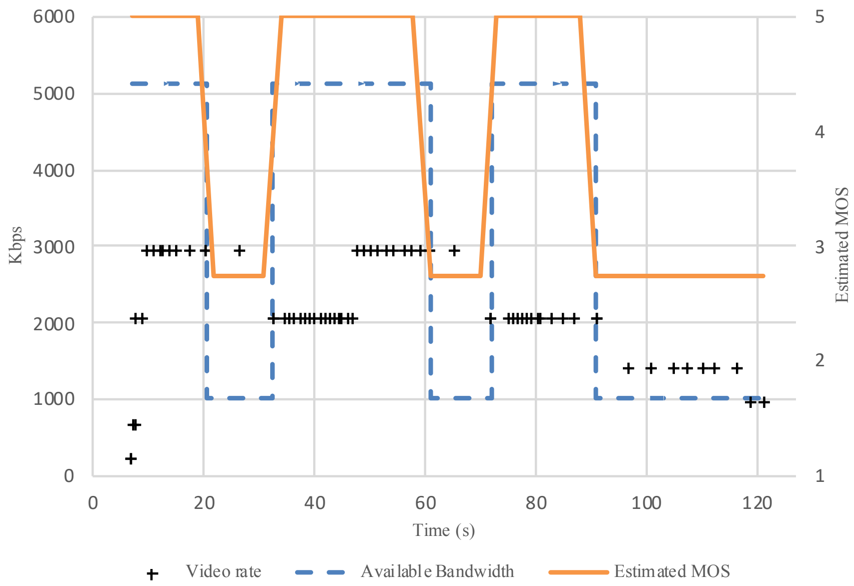

Figure 3.7: Video rate requested by the user, available bandwidth and estimated MOS in the first scenario

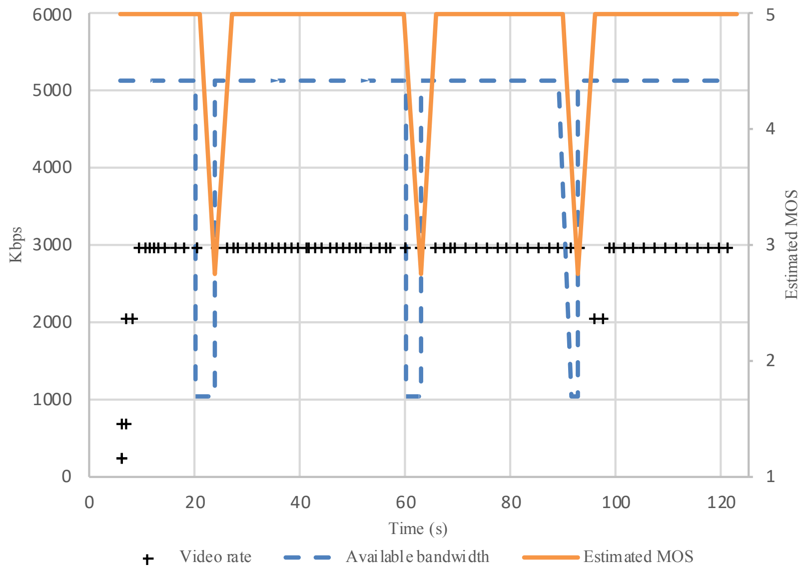

Figure 3.8: Video rate requested by the user, available bandwidth and estimated MOS in the second scenario

router managed to increase the available bandwidth to $5000kbps$, and thus the estimated MOS also quickly returned to 5 at t=26.90s, t=65.90s and t=95.90s. Unlike in the Fig.3.8, any worse deterioration in video rate could not be seen until t=95.69s. But the video rate just deteriorated for a short time from t=95.69s to t=98.49s, then recovered to the original highest value. This is because the estimated MOS detects the network quality change quickly, then the available bandwidth can be adjusted immediately.

It can be seen that that the video rate always takes a large delay to adapt the available bandwidth compared to the estimated MOS. This is because the video rate does not change after detecting the network quality change. In fact, the player reacts, not to the latest fragment download throughput, but to a smoothed estimate of those measurements that can be unrelated to the current available bandwidth conditions. Particularly, in Fig.3.7, when the available bandwidth was decreased, the video rate deterioration could be detected about 12.46s after that. Meanwhile, Fig.3.8 witnessed a short reaction time of estimated MOS. It took only about 3.9s for capturing the deterioration of estimated MOS. It means that by using optimal monitoring interval, MOS-based method can detect video rate deterioration at least $t_d = 8s$ earlier than video rate based method. After controlling available bandwidth, the second scenario witnessed that the video rate remained unchanged or experienced a short-term reduction (observed around T=90s). In contrast, in the first scenario, the video rate did not return to $2962kbps$ within several seconds and it took a large t_r (around 15s) to return or even did not return. This is because the playback buffer size is large enough to compensate for a negative "spike" in the available bandwidth. A small recovering time $t_r \leq 4s$ of video rate which could be seen from the second scenario is meaningful in QoE management. In other words, video rate has been guaranteed to be maximized or to be kept stable at desirable level.

3.5 Summary

In this chapter, a method to early detect QoE deterioration through the determination of monitoring factor and the appropriate QoE monitoring interval was proposed. To sum up, QoS is selected from adaption logic factors as monitoring factor, while its monitoring interval is determined through a condition of stable video rate. Thereby, the monitoring interval was eventually required to be equal to the size of video chunk (in second). By applying this interval, QoE management system could effectively maximize video rate during a streaming session. The effectiveness was represented by early detecting video rate deterioration, short recovery time, low CPU load and low ratio of video rate deterioration.

Chapter 4

Collaborative Approach using Psychophysiology and Psychophysics for Determination of QoE Threshold

4.1 Introduction

Theoretically, QoE control is responsible for making the comparisons between monitored QoE (taken from QoE monitoring) and a specific QoE threshold, and then for triggering a control strategy if it is necessary. In fact, the threshold value will be used to decide whether a control action can be triggered at the right time or not, leading to several potential consequences. If a higher threshold value is set, which means it detects a lower deterioration of video quality, the control action will be generated too early and frequently. As the result, it would bring a high computational cost and a waste of bandwidth, although an expected level of QoE is guaranteed. On the contrary, if a lower threshold value is set, a deterioration of video quality will be subjectively perceived before generating the control action. Consequently, the controlled QoE is not as high as the expectation, although the computational cost and the waste of bandwidth are eliminated. Therefore, the determination of an appropriate QoE threshold is indispensable. However, in literature, such a threshold has not been carefully investigated yet, despite its importance.

As mentioned in section 2.2, hybrid QoE models have become the most commonly used methodology for contemporary studies due to their advances. These models actually consist of a combination between the subjective models and the objective models. Particularly, the relation between subjective perception and QoE influence factors are modeled through a training process in machine learning approach. For these reasons, these models can deliver the assessments as precisely as human does in an automatic and real-time manner. Therefore, they have been increasingly applied in a wide range of studies. In this study, the hybrid models are also the sole option for the design of QoE monitoring within the proposed QoE management framework.

In the modeling process of hybrid models, the subjective perception is obtained by subjective assessment methods (e.g., rating approach, psychophysics) as described in section 2.2. In fact, the rating approach is the most commonly used method in subjective assessment, where the subjects are asked to provide their evaluation in terms of 5-scale Mean Opinion Score (MOS) for given video sequences. Therefore, the fair level in 5-scale MOS (middle value of the 5-scale) is simply selected as the threshold for QoE control [67][73][74]. However, this approach is inheritably biased due to the qualitative nature of the scale (MOS scale) and brings a large variance in the result. This demands a new method to determine a more reliable threshold for QoE control in adaptive streaming services.

Alternatively, psychophysics approach has been potentially considered in the determination of QoE threshold. This approach provides a tool for measuring the perceptual performance of subjects, introducing visibility thresholds and just noticeable differences (JNDs), which are the humans perception levels suitably connecting to physical values. In other words, the psychophysics quantitatively clarifies the relation between physical stimuli and a level of humans perception. Actually, such a relation was successfully modeled as a general logarithmic nature in Weber-Fechner Law, which has already been applied in a wide range of QoE assessment studies [75][76][77][78]. In this approach, the threshold level of stimuli which refers to an absolute threshold is determined by introducing either a gradual increment until stimulus becomes detectable or a gradual decrement until stimulus becomes undetectable. In this study, such a threshold can be considered as the desirable QoE threshold. Nonetheless, in psychophysics, assessment scales like MOS and open-ended questionnaires are still typically used in order to

Table 4.1: The pros and cons of rating approach, psychophysical approach, and psychophysiological approach

Characteristics	Rating approach	Psychophysical approach	Psychophysiological approach
Perception measurement	Subjective evaluation in terms of Mean Opinion Score (MOS)	Biological information (e.g., ECG, EEG, EDA, etc.)	Subjective evaluation in terms of MOS, open-ended questionnaires
Pros	Environmental setup for perception measurement is simple	High accurate assessment without biased and variability	Do not depend on the individual difference
Cons	High biased and variability	Individual difference	High biased and variability

quantitatively and qualitatively evaluate the user's perception to media content. These methods depend on humans conscious responses and often do not provide sufficient insight into underlying perceptual and cognitive process. Therefore, the psychophysics approach alone cannot provide sufficient information for QoE assessment in adaptive streaming services.

In order to address the drawbacks of psychophysics, the approach of psychophysiology has been alternatively taken into account in this study. The psychophysiology refers to physiological signals when stimuli are given, and then the correlations between the characteristics of the physiological signals and the given stimuli are discussed. The physiological measurements are categorized into the following classes [38] - Central Nervous System (CNS), Eye Measurements and Autonomic Nervous System (ANS). The psychophysiology detects the change of target stimulus through the change of corresponding physiological signal. It means that since the user's perception to the stimulus is obtained by the physiological signal, it achieves more precise QoE assessment than the previous approaches do. However, there is a significant limitation in the psychophysiology, which is the individual physiological difference that may produce systematic errors among subjects or groups thereof. Hence, a new approach which is applicable to a general population is required.

The pros and cons of the above approaches are summarized as in table 4.1. Accordingly, the abovementioned issues can be solved when the combination of the psychophysics and the psychophysiology is taken into consideration, compensating for the disadvantages of psychophysics with the advantages of psychophysiology, and vice versa. As the result, more reliable QoE assessments without depending on individual physiological difference will be achieved. Accordingly, a more reliable QoE threshold will be determined.

This chapter proposes a new method to determine an appropriate QoE threshold to ensure that the perceived video quality stays at high and stable level with the minimal network resource utilization by taking into account the combination of psychophysiology and psychophysics in QoE assessment. The contributions of this study are as follows:

(1) Establishing a general logarithmic function expressing the relation between human perception and stimulus intensity. Based on the function, the human perception is estimated through physiological measurements (Note that in this work, the term of "biological information" will be used instead of "physiological information", and stimulus will be referred to as a QoE influence factor).

(2) Defining an appropriate QoE threshold by determining the absolute threshold or its constraint.

4.2 Related work

Determination of QoE threshold plays an important role in balancing network resource utilization and the resulting QoE. The existing threshold is not as reliable as the expectation since each individual has different interests and expectations for video quality [79]. For determining a more reliable QoE threshold, a novel method which is the combination of psychophysiology and psychophysics has been proposed. This section will review the existing works that related to QoE assessment using either psychophysiology or psychophysics.

Estimating human's perception using the intensity of a specific stimulus has been a major concern of many scientists for years. The initial cornerstone was marked by the breakthroughs of Weber and Fechner, where the relation between human perception and stimulus intensity was modeled as a logarithmic nature expressed by Weber-Fechner Law. Consequently, the Weber-Fechner Law (WFL) has been widely applied in various research areas. In [80] the logarithmic nature of QoE for a given QoS parameter is discussed because of the underlying WFL. Meanwhile, in an effort to derive a specific the so-called Web QoE model, the authors in [75] also successfully took into account WFL for modeling the relation between QoE and page load time. Based on WFL, the utility functions between QoE and QoS was derived in [76]. In this research, the authors considered the QoE as the human perceptual intensity of the stimulus and the reciprocal of the QoS parameter (time delay) as the physical magnitude of the stimulus. The utility functions show the relation between QoE and allocated bandwidth of each user types were also represented by WFL in [78]. In our research, the appropriate QoE threshold is also expected to be determined by modeling the impact of stimulus intensity on human perception based on the logarithmic nature of WFL.

Psychophysiology has been increasingly applied in QoE assessment for multimedia services, including adaptive streaming services [38]. There are numerous studies which focus on QoE assessment by using biological measurements. [81] presented a significant correlation between EEG/ECG and video quality levels, [82] illustrated that Electrodermal Activity (EDA) is effective in measuring human perception to given visual fatigue, whereas [83] concluded that the same result was found from both EEG and EDA. Particularly, in an effort to directly measure the changes of perceived video quality using EEG, the authors in [13] showed that abrupt changes of video quality give rise to specific components in the EEG that can be detected in a single-trial basic. Potentially, a neurotechnological approach to video assessment can lead to more objective quantifications of quality change detection, overcoming the limitation of subjective approaches (such as subjective bias and the requirement of overt response). Those contributions promisingly motivate applying psychophysiology in QoE management in adaptive streaming services. In our research, the biological information (EDA, heart rate and heart rate variability) associated with Autonomic Nervous System (ANS) is initially used to determine the appropriate QoE threshold.

4.3 Background

Section 4.2 clarifies the applications of biological measurements in QoE assessment, especially ECG and EDA which are associated with ANS. Considering the measurement efficiency and the convenience for the users, EDA, heart rate (HR) and heart rate variability (HRV), are eventually selected as measured signals in this research. This section investigates the background knowledge related to the terms of biological signals (EDA, HR and HRV) and psychophysics (Sensation, perception and Weber-Fechner Law).

4.3.1 Electrodermal Activity

In human's perceptual process, sensory receptors convert sensations into electrical impulses [84]. The electrical impulses are relayed to the brain and the responses to the stimulations are induced. Electrodermal activity refers to changes in electrical conductance of the skin, which is associated with eccrine sweat gland activity innervated by the sympathetic branch of the autonomic nervous system. Measuring electrodermal activity (EDA) is a promising way to determine the amount of an individual's response to stimuli [85][86]. Thus, it can be a useful index of changes in sympathetic arousal that is tractable to emotional and cognitive states. Skin conductance (SC) is the most widely studied property of EDA, representing autonomic changes in the electrical properties of the skin. SC comprises of two major components - tonic and phasic. Tonic component is represented by skin conductance level (SCL) which is the baseline level of SC, in the absence of any particular discrete environmental event. On the other hand, phasic component refers to skin conductance responses (SCRs) reflecting the short-time response to a given stimulus. SCRs are also the results from sympathetic neuronal activity. In SC qualitative modelling or poral valve modelling [87], SC is the result of a numerous activities related to sweat ducts, pore, sweat glands, etc. In the initial condition, the distal part of the sweat ducts is collapsed by the external pressure of hydrated surrounding corneum. As the result, most of the pores are closed. When the sweat fills the ducts to their limitation of capacity, the intraductal pressure will cause a hydraulic driven diffusion of sweat into the corneum. The rising of SC is caused by the increasing the hydration in deeper levels of the corneum. As the sweat is reabsorbed into the dermis or diffuses away from the periductal area, SC will slowly recover, resulting in a rather flat SCR. When the secretion of the sweat is enough and the intraductal pressure becomes stronger than the tissue pressure of the corneum, the pore will eventually open. The sweat will now be forced out through the pore. Consequently, SC will drastically increase. The amount of sweat which is pushed out through the pore is substantial. Therefore, after a short time, the secretory rate cannot keep up with the loss of sweat. The

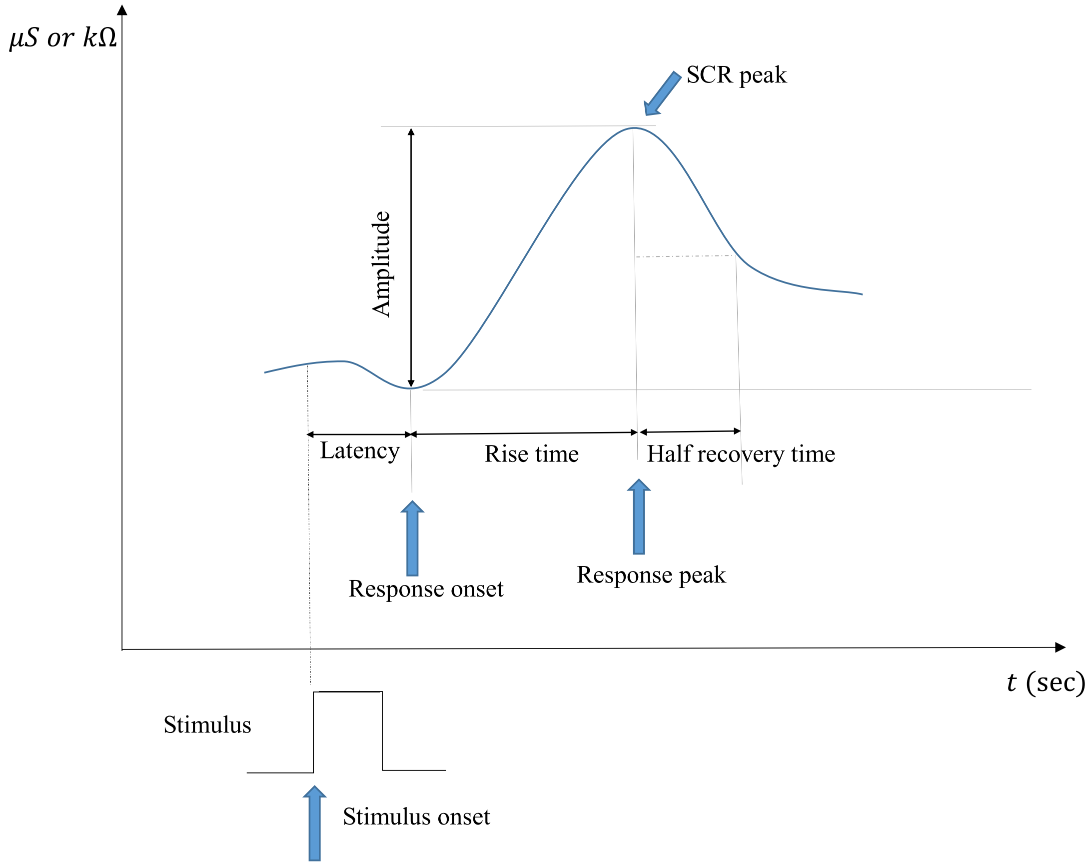

Figure 4.1: The graphical representation of principal SCR related components [2]

intraductal pressure will soon fall below the tissue pressure, and the pores will be collapsed again followed by a rapid fall in SC. Basically, SCR shape can primarily be ascribed to two different underlying processes: one is an unconditional diffusion process which causes a rather flat SCR, and the other is an optional opening of pores which will add a steep peak to the basic SCR shape. Thus, these processes increase the variability of the SCR shape. Figure 4.1 depicts an example of SCR shape which comprises the following components - Latency (time from stimulus onset to SCR onset), Rise time (time from SCR onset to SCR peak) and Half recovery time (time from SCR peak to 50% recovery of SCR amplitude). In the figure, the vertical and horizontal axes represent skin conductance value and duration time, respectively.

4.3.2 Heart Rate

Heart rate (HR) is the speed of the heart beat measured by the number of contractions of the heart per minute (bpm)[88]. The heart rate can vary according to the bodys physical needs, including the need to absorb oxygen and excrete carbon dioxide. It is usually equal or close to the pulse measured at any peripheral point. In general, heart rate is regulated by sympathetic and parasympathetic input to the sinoatrial node. The accelerans nerve provides sympathetic input to the heart by releasing norepinephrine onto the cells of the sinoatrial node (SA node), and the vagus nerve provides parasympathetic input to the heart by releasing acetylcholine onto sinoatrial node cells. Therefore, stimulation of the accelerans nerve increases heart rate, while stimulation of the vagus nerve decreases it. The balancing action of the sympathetic nervous system and parasympathetic nervous system controls the HR. The degree of variability in the HR provides information about functioning of the nervous control on the HR and the heart's ability to respond. Thus, investigating the behavior of HR can express the impact of stimulus intensity on human reaction.HR can be measured by finding the pulse of the heart. The pulse rate can be found at any point on the body where the arterys pulsation is transmitted to the surface by pressuring it with the index and middle fingers; often it is compressed against an underlying structure like bone.

4.3.3 Heart Rate Variability

Heart rate variability (HRV) is a popular noninvasive tool to estimate cardiac autonomic modulation because it is easily measured without interfering with neural functions of the control mechanisms. Analysis of these fluctuations has been useful in providing information regarding the physiology of the active autonomic control branches [89][90][91]. In fact, HRV is the degree of fluctuation in the length of the intervals between heart beats. It is mirroring the regularity of hearth beats: bigger regularity lower HRV and vice versa. Regularity of heartbeats is derived from a quantity of numbers; equal to the times elapsed between successive heartbeats, namely R-R intervals which are measured in millisecond (ms).

4.3.4 Sensation, perception, Weber Law, and Fechner Laws

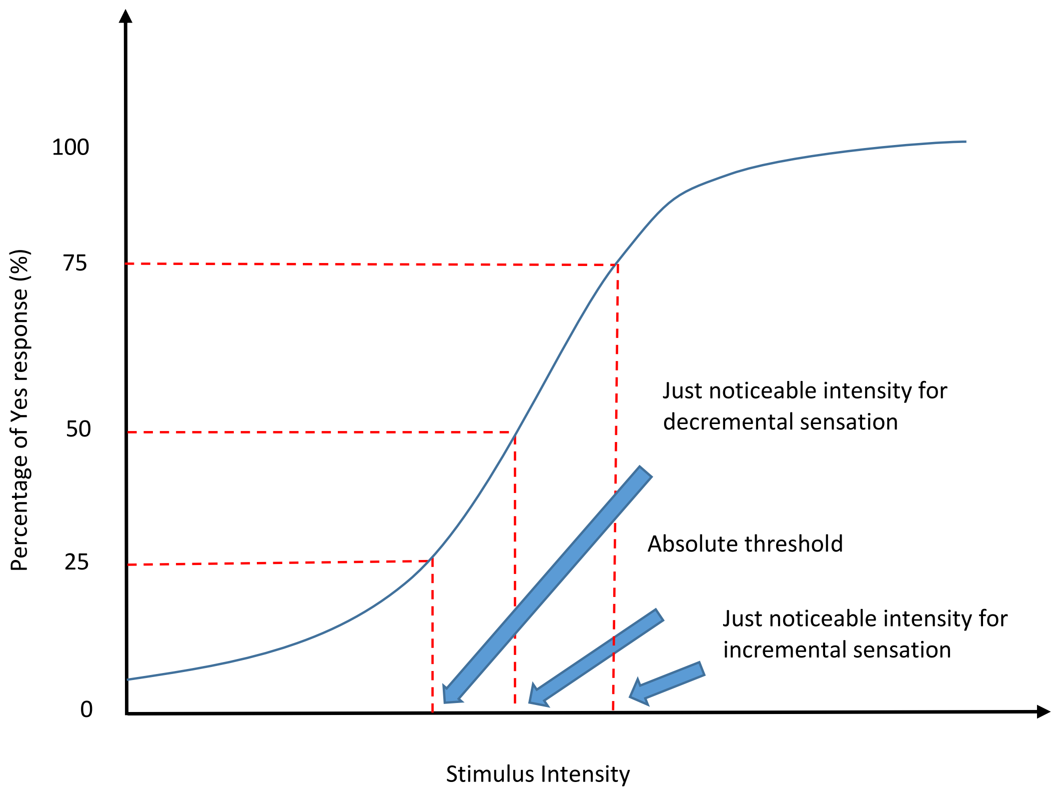

Figure 4.2: Absolute threshold obtained from psychometric function

Although the terms of sensation and perception are often used interchangeably, it is important to clarify the difference between them. The sensation is the process in which the sensory receptors and nervous system receive stimulus energy from the outside environment. It also represents the amount of the stimulus energy. Meanwhile, the perception is the process of organizing and interpreting sensory information, enabling human to recognize meaningful objects and events. More concretely, the sensory process captures information from the outside world and transforms it into biological signals that are interpreted by the brain [92][93]. Afterward, the brain produces a perceptual representation that allows human to appreciate the outside world. In other words, the perception represents a single unified awareness of a stimulus that in turn arises from the sensation produced by the sensory system.

As mentioned in section 4.1, human perception is an important role in QoE assessment. Typically, the good way to understand this concept is to establish a quantifiable relation between the physical stimulus and the perception. Consequently, the perceptual quality of a stimulus can be represented in numerical terms, enabling the comparison with other stimuli [92]. Weber and Fechner were experimental psychologists of the 19th century who made efforts in establishing such a relation. Before going further, the definitions of both "absolute threshold" and "difference threshold" which are of crucial in this research, need to be clarified.

As mentioned in [92][93][94][95], the absolute threshold is defined as the minimum level of stimulus intensity that can be recognized as a sensory event by the brain. The stimulus intensity in the range of lower than the absolute threshold is called as subthreshold intensity. No detectable sensation is produced from subthreshold intensity. On the other hand, the stimulus intensity higher than the absolute threshold is called supra-threshold intensity. Sensation takes place from the supra-threshold. In the supra-threshold intensity, a certain level of intensity change either increment or decrement intensity can be detected as perception. The difference threshold concept came up to determine how much change in stimulus intensity is needed to recognize the difference in sensation, which is called Just Noticeable Difference (JND).

Fechner tried to understand the relation between stimulus intensity and perception. To achieve this, the absolute threshold was determined by a series of experimental approach. In this experiment, the reference stimulus with a constant intensity was prepared. Besides the reference stimulus, the subjects were given

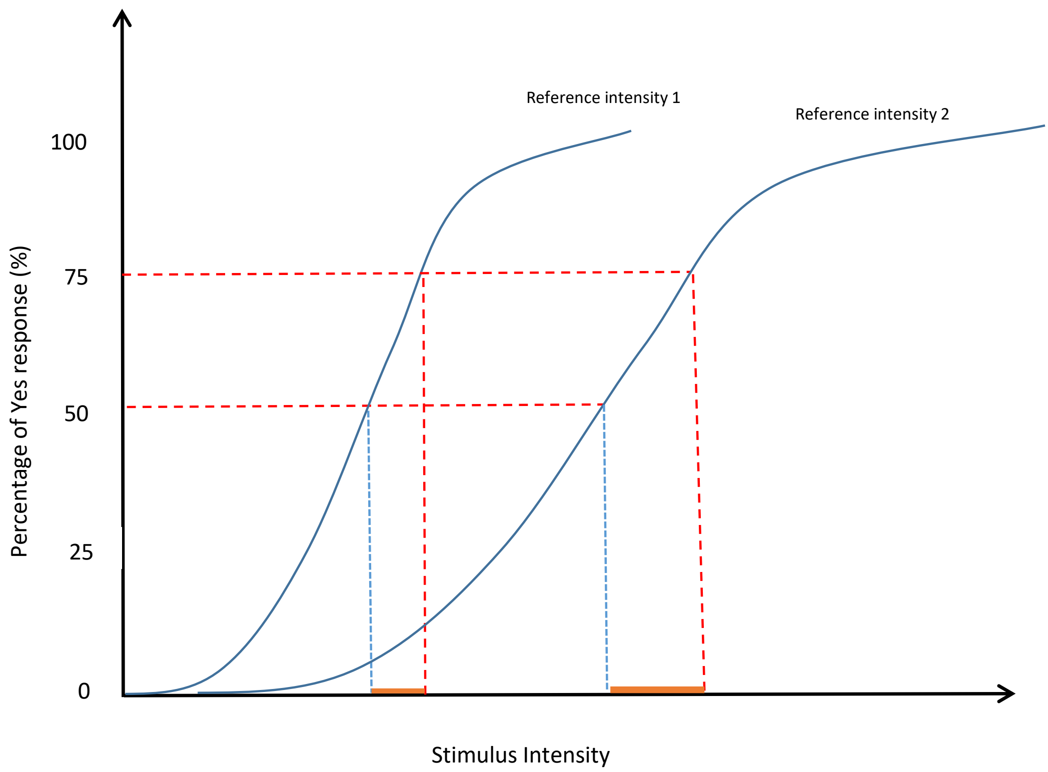

Figure 4.3: The greater change in absolute threshold, just noticeable intensity when the reference intensity progressively increases.

different stimulus with intensity randomly varied in pre-defined range. They compared each given stimulus with the reference stimulus and answered the "Yes" or "NO" which corresponds to whether the stimulus change was perceived or not. As illustrated in Fig.4.2, the absolute threshold is determined as a stimulus intensity with which the stimulus was perceived by the 50% of the subjects. Whereas, the intensities at which the 25% and 75% of "yes" response were obtained are considered as just noticeable intensities for incremental and decremental sensation, respectively, according to the definition of JND.

The Weber's Law answers the question how the absolute threshold, just noticeable intensity and JND change when the reference intensity varies. After conducting the series of experiments, Weber found that the difference threshold increases in a linear fashion with stimulus intensity as shown in Eq.4.1 and Fig.4.3. It means that higher supra-threshold intensity requires larger change in intensity (ΔI) needed to produce a change in sensation, namely, JND.

$$\Delta I = k * I \tag{4.1}$$

After the introduction of the Weber's Law, Fechner eventually figured out the relations between stimulus intensity and perception by integrating his bold assumption with Weber's Law. The assumption was stated as: the subjective size of the JNDs to be constant, irrespective of sensation magnitude. This assumption can be mathematically achieved in a logarithmic function. The Eq.4.2 completely expresses his idea:

$$S = k^{'} \log I \tag{4.2}$$

Where constant $k^{'}$ is related to, but not identical to, the constant k in Weber's Law. S is the magnitude of sensation, whereas I stands for stimulus intensity, and $\log I$ indicates natural logarithm of I. It can be concluded that both data points whose intensities are equal to the absolute threshold and the just noticeable intensity must belong to the curve of Eq.4.2.

4.4 Methodology

In this section, the hypotheses of this research are introduced in detail, explaining how they connect to the proposal.

4.4.1 Hypotheses

Looking back to section 4.1, in this research, a new method to ascertain the optimal QoE threshold for control component in QoE management is proposed. More concretely, the collaborative approach using the psychophysiology and the psychophysics to clarify a general logarithmic nature function between human perception and stimulus intensity is thoroughly discussed. In this research, the perception has been estimated via biological information (including SC, HR, and HRV) measurements. Afterward, either absolute threshold or just noticeable intensity is expectedly obtained from the general logarithmic nature function, which facilitates the determination of optimal QoE threshold.

In order to achieve this, the following hypotheses must be justified:

- Biological reaction to a stimulus change varies from person to person and depends on the types of biological information. It can be detected by an increment or decrement of the amplitude.
- Biological information, which shows the level of human perception to stimulus intensity, has a logarithmic nature relation with the stimulus intensity.
- The absolute threshold which is the minimum stimulus intensity perceived by 50% of subjects is ideally regarded as the stimulus threshold.

4.4.2 Modelling of biological information

In order to model a general logarithmic nature between human perception and stimulus intensity, initially, the type of stimulus given to the human needs to be clarified. This can be done by investigating the prominent QoE influence factors. According to [11], there are numerous influence factors categorized into perceptual and technical factors. Video rate actually influences the quality of video frame; hence, it can be regarded as one of the factors directly perceived by subjects. Therefore, it was chosen as the stimulus type in this research. There are copious existing models which have a capability to interpret the video rate to QoE (indicated by Mean Opinion Score) and vice versa [96][97], and thus, the video rate threshold can be used for the trigger of control action in QoE management, instead of QoE threshold.

In section 4.2, the wide range of Weber-Fechner Law (WFL) applications has been introduced, especially in QoE assessment throughout modelling the relation between QoE and QoS. However, subjective QoE is insufficient to characterize the human perception in QoE assessment. On the other hand, the human perception introduced by biological information promisingly tells us the truth on how stimulus intensity is perceived. Additionally, since more precise perception can be derived from the combination of different types of biological information [38], it is necessary to jointly investigate the impact of stimulus on multiple biological signals. In this subsection, by applying WFL, the impact of the video rate on the human perception estimated by multiple biological information, which are SC, HR, and HRV, is modeled. Therefore, the Eq.4.2 in subsection 4.3.4 can be re-written as follows:

$$y_i = k_i^{'} \log x \tag{4.3}$$

Where $i \in \{1, 2, \ldots, n\}$ is the index indicating each subject, and y_i is the perception level of subject i to the stimulus of SC or HR or HRV, respectively. Let $k_i^{'}$ be the sense-specific constant depending on the sense (SC, HR and HRV) and type of stimuli.

In order to achieve the research purpose, the following tasks need to be covered:

- Establish the collection of n regression curves of Eq.4.3 by fitting data obtained from each subject.
- Investigate the existence of the data point which is crossed by at least 50% of regression cruves.
- Determine the general logarithmic function representing for all subjects from the data point which is determined in the second step.
- Determine the absolute threshold.

To clarify the above tasks, more explanations are needed. As mentioned in hypotheses, the perception varies from person to person and depends on the type of stimulus, thus, the shape of regression curve of Eq.4.3 is predictably different in each subject, only meaning that each subject separately perceives the stimulus intensity. If 50% and more than 75% of subjects have the same perception magnitude to the same intensities, they are considered as the absolute threshold and the just noticeable intensity, respectively. Then, those regression curves are expected to intersect at the same data point denoted by $P(x_0,y_0)$. According to the conclusion of subsection 4.3.4, if this data point exists, it must be the crossed point by the curve of the general logarithmic functions y. Data point $P(x_0,y_0)$ can be determined by minimizing the sum of squared residuals, defined as the square of the difference between y and y_i:

$$f(x,y) = \sum_{i=1}^{\infty}(y - y_i)^2 \tag{4.4}$$

Assume that $f(x,y)$ is a continuous function, then x_0 and y_0 are the extremums that satisfy the first derivative:

$$\frac{\partial f}{\partial x}(x_0;y_0) = \frac{\partial f}{\partial y}(x_0;y_0) = 0 \tag{4.5}$$

Afterward, the constant $k^{'}$ in Eq.4.2 can be calculated as:

$$k^{'} = \frac{y_0}{\log x_0} \tag{4.6}$$

Therefore, the general logarithmic function can be determined as follow:

$$S = \frac{y_0}{\log x_0}\log I \tag{4.7}$$

In this research, some experiments were performed following the "method of limits" [92] by asking the subjects to watch a movie with gradually decreasing the video rate. Thereby, the video rate was deteriorated from the highest to the lowest level within a pre-defined range of j levels. The subtraction of the current video rate level from the highest level was the stimulus intensity. Thus, the stimulus intensity was gradually increased respect for the decrease of the video rate. In the experiment, the duration time between each decrement was about

5s, thus, the current level of intensity was considered as the reference intensity for a judgment of the next higher intensity level.

The following assumptions need to be subsequently made:

- Data point $P(x_0,y_0)$ exists.
- x_0 is either absolute threshold or just noticeable intensity.
- Let X_m be the m^{th} reference intensity for a judgment of $(m+1)^{th}$ intensity with $1 \leq m \leq j$.
- Let $X_{threshold}$ be the stimulus threshold that needs to be determined.

If x_0 is equal to the absolute threshold, the stimulus threshold is ideally determined as: $X_{threshold} = x_0$

If x_0 is recognized as the just noticeable intensity, then, the absolute threshold is need to be determined. Actually, only the constraint of absolute threshold can be determined in this case. Because the constant k in Weber's Law (Eq.4.1) is unknown, thus, the next higher intensity cannot be ascertained from the current one. Accordingly, the absolute threshold is derived from the following constraint:

$$X_n \leq X_{threshold} \leq x_0 \tag{4.8}$$

4.5 Evaluation

In this section, the hypotheses of this research are practically confirmed through some experiments. In the experiments, subjects watched short video clips. Video quality of clips was gradually changed from the highest to the lowest level while SC, HR and HRV data were being continuously recorded.

4.5.1 Experimental environment

Figure 4.4 shows the experimental setup which comprises of a screen, a Zoom watch (for measuring heart rate and heart rate variability) connected respectively to Elite HRV app on iPhone 6 via Bluetooth and a Grove - GSR sensor [98] connect to Arduino UNO board (for measuring skin conductance). The sampling

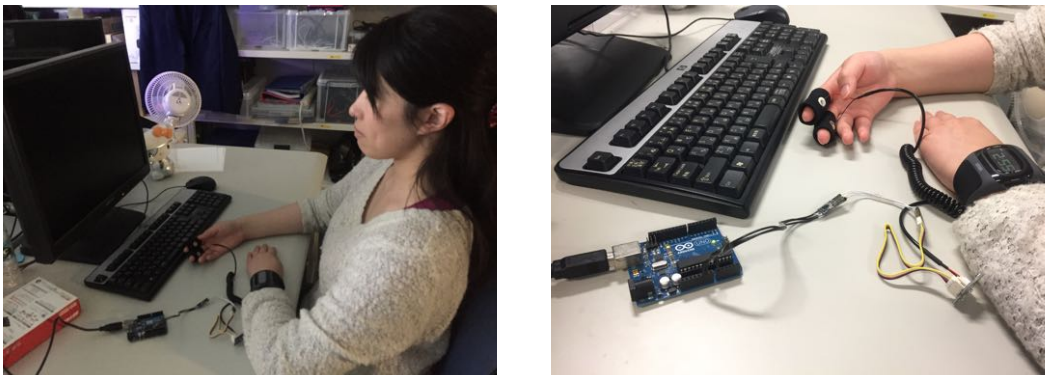

Figure 4.4: Environmental Setup with Zoom watch and Grove-GSR sensor

frequency for HR and SC monitoring were 1Hz and 20Hz, respectively. Note that the sampling frequency of HR was fixed by the device vendor, whereas the one of SC was flexibly changeable.

In order to maximize the reliability of the experiments, the video clips should satisfy some requirements. The video material utilized in these experiments should not be a semantically important content or a salient content. As the result, the influences due to high-level image understanding were eliminated. Furthermore, the duration of watching the video should not be too long in order to avoid the distraction of the subjects. Also, the video rate must have been decreased level by level.

To meet these requirements, a video clip generated from an open source 4k movies was prepared as follows:

- The utilized video clip had an abstract content; a view of sky at night which slowly moves from the left to the right.
- The original video clip was encoded to Internet Information Services (IIS) smooth streaming [99] - 720 CBR (constant bitrate) by using Microsoft Encoder pro 4.0 which allows it to be available with multiple video rates ($kbps$): 2962, 2056, 1427, 991, 688, 477, 331 and 230$kbps$. The reason to encode the video with constant bitrate instead of variable bitrate is to ensure that the video clips always keeps the same pre-defined bitrate level. The stimulus intensities were calculated by doing the subtraction of current video rate level from the highest level of 2962$kbps$. The stimulus intensities

were: 906, 1535, 1971, 2274, 2485, 2631, and 2732*kbps*. (Note that by default the video clip initially started from the highest video rate of 2962*kbps* and the first decrease started at 2056*kbps*)

- A seamless video clip of 40 seconds which comprises of eight 5-second-chunks was generated. Each of chunk is respectively available with each video rate mentioned above. Therefore, this video clip contains seven stimulus intensities. In addition, the current stimulus intensity was considered as the reference intensity of the judgment for the next higher intensity.

The video clip was displayed on a 22-inch screen Samsung SyncMaster 2243BW with a native resolution of 1680 x 1050 The video resolution was 1276 x 660 pixels or 32.4 x 16.8 cm and the viewing distance was 67.2 cm (four times of the video height on the screen) in compliance with the specifications in [37]. Ten subjects (4 females and 6 males in the age group of 20-27) participated in the experiment. They were doing the researches not related to this research topic. All subjects had normal or corrected-to-normal vision. The subjects sat in front of the display in a dark and quiet room. They were also asked to leave all their mobile phones or any noise-making devices which could distract them during the experiment. After attaching the measurement devices, the subjects had five minutes at rest to be familiar with the experimental environment that made them achieve the highest comfortable status. When the subjects were ready, they clicked a button on the screen to play the video, whereas SC, HR, and HRV recordings were also started. The experiment was repeated two times for each subject.

4.5.2 Data acquisition and data transformation

Skin conductance raw data was originally read from the Serial port on the computer using a simple program written in Python. Then, it was exported to CSV format followed by its transformation to "txt" file which is suitable input form for Ledalab - an open source SC data analysis tool [100][101]. According to [102], there are individual differences in amplitude of SC. In other words, the overall response of a subject is not the same as the one of others. Therefore, it is important to normalize and standardize the raw data collected from the above experiment.

In this research, the skin conductance data was transformed following the two-phase: Normalization and standardizations [103]. Firstly, the raw data was normalized based on the Eq.4.9. Afterward, the data was standardized by calculating the ratio of normalized data and its mean with Eq.4.10.

$$SC_{nor} = \frac{SC - SC_{min}}{SC_{max} - SC_{min}} \tag{4.9}$$

$$SC_{sta} = \frac{SC_{nor}}{EX(SC_{nor})} \tag{4.10}$$

Here, SC_{nor} is normalized SC data, SC_{min} and SC_{max} are the minimum and maximum SC values, respectively. Meanwhile, SC_{sta} is standardized SC data, $EX(SC_{nor})$ is the mean of the normalized SC data.

After normalization and standardization, SC data needs to be analyzed to confirm the hypotheses throughout the following criteria:

(1) Event-related skin conductance response (ER-SCRs) of each subject

(2) Regression model which represents the relation between SC and stimulus's level

Firstly, ER-SCRs presents the responses of subjects for a given stimulus. The number of responses as well as the amplitude of responses play a critical role in investigating how each subject reacts to the stimulus. As mentioned in section 4.3.1, SC data comprises of tonic and phasic components. Thus, to extract ER-SCRs, the SC data is decomposed into its tonic and phasic components by Ledalab tool. The decomposition results in the extraction of un-superposed response components and thus allows for an unbiased quantification of SCR characteristics (e.g., SCR amplitude). The following outputs should be derived from the decomposition: The number of response and SCR amplitude. SC theoretically reacts to a stimulus by skin conductance response and the SCR usually occurs about one to five seconds after stimulus's onset [2][104]. Therefore, the outputs were obtained within the pre-defined response window (from 1 to 5 seconds).

Secondly, to model the impact of stimuli on skin conductance, the average value of SC_{sta} data was calculated within the response window. Therefore, for each trial of experiment, there were totally seven average values according to seven setup-stimulus. By performing the logarithmic approximation, the above-mentioned relation was established.

Heart rate raw data was initially stored on Elite HRV app where both the average value and the general trend were clearly shown off. The raw data can be extracted by email as the attached files. The following criteria were used to confirm the hypotheses of this research:

(1) The variation of HR within the pre-defined response window

(2) Logarithmic approximation obtained from the normalized data.

The variation of HR stands for the standard deviation which presents the fluctuation of HR data from the subject's resting HR baseline. Before modeling the relation between HR and stimulus, the raw data must be normalized by applying the following equation:

$$HR_{nor} = \frac{HR - HR_{min}}{HR_{max} - HR_{min}} \tag{4.11}$$

Here, HR_{nor} is normalized HR data, HR_{min} and HR_{max} are the minimum and maximum HR values, respectively.

Heart rate variability raw data was also stored on Elite HRV app before being extracted by email as the attached files. The output data was R-R intervals in milliseconds. HRV can be assessed in two ways, either as a time domain analysis or in the frequency domain as a power spectral density (PSD). The purpose of the experiment is to investigate the variation of R-R intervals for given stimulus intensity, thus, time domain analysis is more suitable. Particularly, the square root of the mean squared differences (RMSSD) of successive R-R intervals was taken into account in this research. Eventually, the following criteria were used to confirm the hypotheses:

(1) RMSSD calculated from each pre-defined response window

(2) Logarithmic approximation obtained from RMSSD

4.5.3 Numerical results

This subsection shows the numerical results of SC, HR, and HRV measurements followed by analyzing criteria mentioned in subsection 4.5.2. Consequently, the

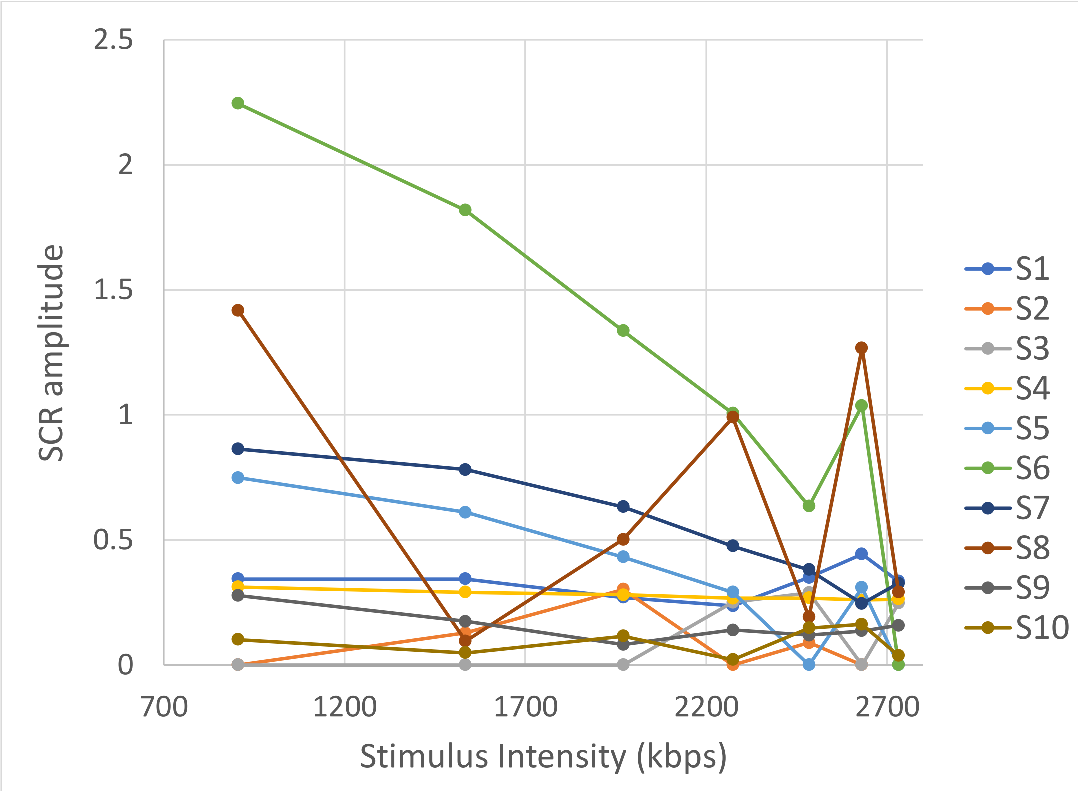

Figure 4.5: The SCR-amplitude of significant SCR re-convolved from corresponding phasic driver-peaks

hypotheses were practically justified.

Skin conductance data, in general, shows that the amplitudes of responses varies from person to person due to the difference of their skin properties. Table 4.2 presents an example of the experimental result, which shows the discrete decomposition analysis (DDA) of a subject. DDA is a method to decomposes the SC data into the tonic and discrete phasic components. The result implies how the particular subject perceives the stimulus's intensities through the number of significant SCRs and the amplitude of significant SCRs. The stimulus intensities are shown in the first column "Event.Name". Whereas, the number of significant SCRs (ER-SCRs) and the amplitude of significant SCRs are presented in the column of "DDA.nSCR" and "DDA.AmpSum". Data in column "DDA.Latency" presents the response latency of the first significant SCR. Lastly, "DDA.Tonic" column shows the mean tonic activity of decomposed tonic component. The numbers of responses and the amplitudes decrease (sometimes decrease to zero) when stimulus intensity increases. Fig.4.5 illustrates the amplitude of SCR ob-

Table 4.2: NS-SCR analysis in SC data. The results were the output of Discrete Decomposition Analysis (DDA) done by lelalab tool

Event.Name	DDA.nSCR	DDA.Latency	DDA.AmpSum	DDA.Tonic
906	1	4.075	0.749	5.464
1535	1	3.775	0.610	5.464
1971	1	4.375	0.431	5.464
2274	1	3.875	0.290	5.464
2485	0	NaN	0	5.464
2631	1	1.575	0.310	5.464
2732	0	NaN	0	5.464

tained from each subject. It can be seen that the amplitude varies from subject to subject according to the stimulus intensity increment

Heart rate and **heart rate variability** data presents the same implications as SC data when subjects differently perceive the stimuli. As mentioned in subsection 4.5.2, the standard deviation as the variation of HR from the resting HR baseline was a crucial criterion. Fig.4.6 illustrates the standard deviation of HR in each video rate within pre-defined response window. In the figure, each color shows each subject. In general, there is no conspicuous consistency among subjects. Some subjects, e.g. subject 1, 2, 8 and 9, expose high variations when stimulus intensity increases to the values of $1535kbps$ and $2274kbps$. Meanwhile, RMSSD calculated within pre-defined response window, is used for a reliable measure of HRV and parasympathetic activity. According to Fig. 4.7, RMSSD data also varies from subject to subject.

After confirming the first hypothesis, the tasks mentioned in subsection 4.4.2 need to be accomplished. Initially, the three series of logarithmic nature regression curves of Eq.4.3 were established by respectively fitting SC data, HR data, and HRV data obtained from all subjects. As the result, each series of either SC data or HR data or HRV data comprised of ten curves of ten subjects. The accuracy of those logarithmic approximation was represented by correlation of determination denoted by R-squared. The according R-squared of thirty regression curves are shown in Table 4.3. It is clear to see that SC data produces a better approximation with R-squared of 0.78892 and higher, except subject 3 and

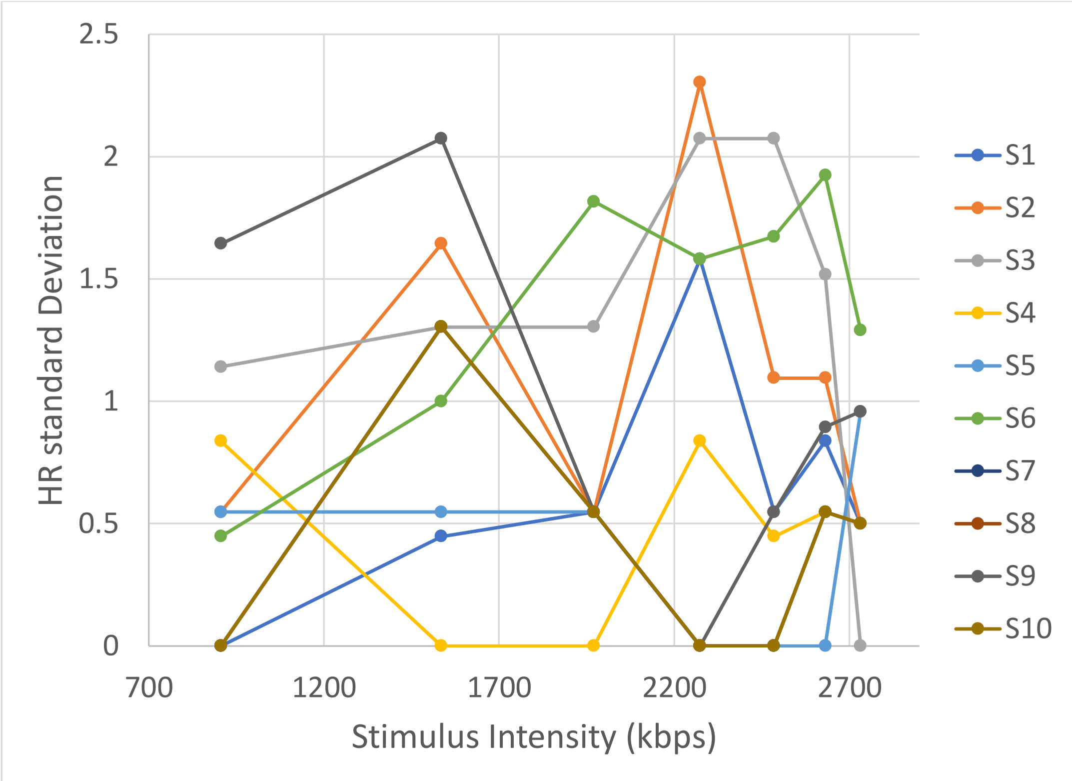

Figure 4.6: Standard deviation of Heart Rate data obtained from particular subject

4. Meanwhile, oppositely, R-squared of HR and HRV are extremely low. It means that the relations between either heart rate or heart rate variability and stimulus were not well fitted.

Investigating the existence and the determination of data point $P(x_0,y_0)$ are the next tasks in this subsection. The functions of regression curves taken from SC data have the following form:

$$y_i = a_i \log x + b_i \tag{4.12}$$

Where $i \in \{1, 2, \ldots, n\}$ is the index indicating each subject, and y_i is the perception level of subject i to the stimulus of SC. a_i and b_i are constants of those functions. Let X denote $\log x$ in Eq. 4.12 for simplicity.

As mentioned in subsection 4.4.1, to ensure that x_0 is the absolute threshold or the just noticeable intensity, the data point $P(x_0,y_0)$ must be crossed by at least 50% of regression curves. Fig. 4.8 depicts the regression curves of SC data

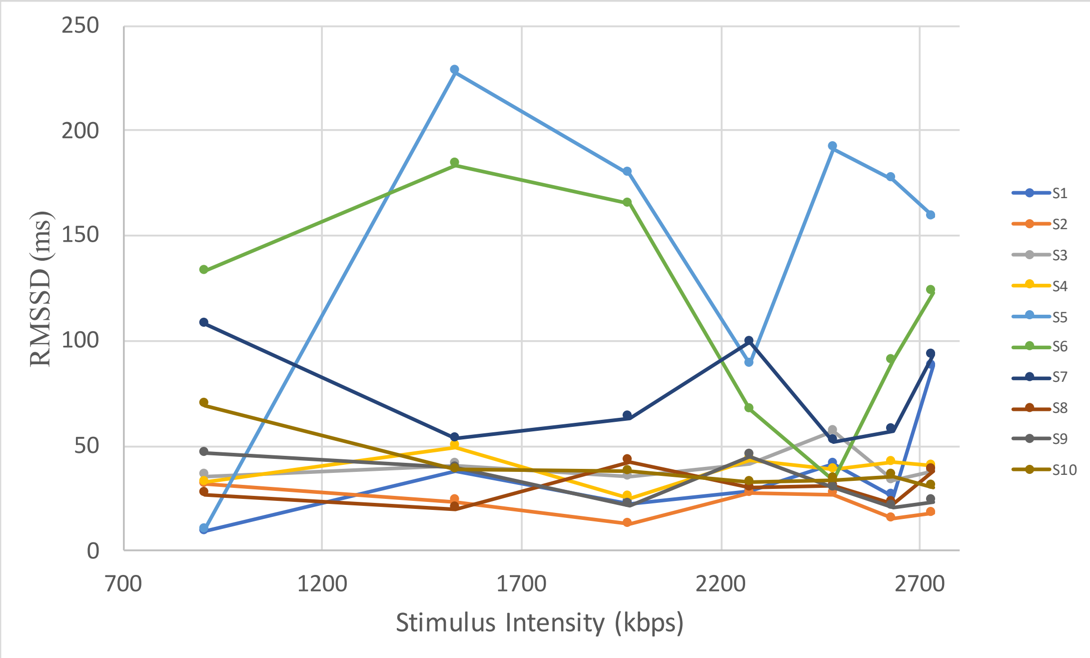

Figure 4.7: Square root of the mean squared differences of successive R-R intervals obtained from particular subject

obtained from 10 subjects. Visually, it is clear that the curves of subject 3 and 4 do not intersect with the rest of curves at the same data point. As the initial prediction, eight curves obtained from the rest of subjects seems to intersect at one data point. Thus, the intensity of this data point is predictably perceived by about 80% of subjects (except subject 3 and 4). In other words, it turns out to be a just noticeable intensity. To confirm this, the method of least square in Eq.4.4 was performed determining the data point $P(x_0,y_0)$. The requirement is to find x_0 and y_0 satisfying the minimum sum of squared residuals, defined as the square of the difference between y and y_i.

Then x_0 and y_0 are the extremums that respectively satisfies the first derivative of Eq. (5). The values of x_0 and y_0 were eventually calculated as $2113.62kbps$ and 1,0079, respectively. This means that there is the existence of one data point which is crossed by eight regression curves in the investigating range of stimulus intensity, that is to say, from $906kbps$ to $2732kbps$. This leads to the fact that x_0 was confirmed as the just noticeable intensity. Therefore, based on the implication in subsection 4.3.4, the data point $P(x_0,y_0)$ must belong to the general logarithmic curve of Eq.4.2. Then, from the value of x_0,y_0, the constant k' in general logarithmic function is calculated as: $k' = \frac{y_0}{\log x_0} = 0.13$

Table 4.3: The correlation of determination denoted by R-squared obtained from each subject in both SC data, HR data, and HRV data

Subject	$R - squared_{SC}$	$R - squared_{HR}$	$R - squared_{HRV}$
1	0.99084	0.83597	0.33891
2	0.89269	0.25126	0.71584
3	0.18767	0.38341	0.06240
4	0.16698	0.36150	0.04266
5	0.87721	0.81180	0.33354
6	0.90581	0.00350	0.23698
7	0.92398	0.58030	0.13811
8	0.91765	0.27501	0.44766
9	0.99920	0.79140	0.65597
10	0.78892	0.21132	0.86912

Consequently, the general logarithmic function which expresses the relation between SC and stimulus intensity will be: $y = 0.13 \log x$

According to the subsection 4.4.2, due to the unknown of constant k in Weber's Law (Eq.4.1), it is impossible to determine the absolute threshold. Instead of this, the constraint of the absolute threshold is established. Because $x_0 = 2113.62kbps$ and being in the range intensity of (1971, 2274), thus, the reference intensity is $X_m = 1971kbps$. According to Eq.4.8, the constraint of the absolute threshold is determined as: $1971kbps \leq X_{threshold} \leq 2113.62kbps$

By subtracting the reference intensity and just noticeable intensity from the highest video rate level of $2962kbps$, the video rate threshold $BR_{threshold}$ is accordingly defined:$848.38kbps \leq BR_{threshold} \leq 991kbps$

For HR and HRV data, the same method was expected to perform for determination of either absolute threshold or its constraint. However, according to Table 4.3, each type of data has only three regression curves that have high accuracy in terms of correlation of determination R-squared. It means that the intensity x_0 of data point $P(x_0,y_0)$ was perceived with the same amplitude of perception by only 30% of subjects. Thus, the existence of both absolute threshold and just noticeable intensity, that is to say, perceived by at least 50% of subjects is ex-

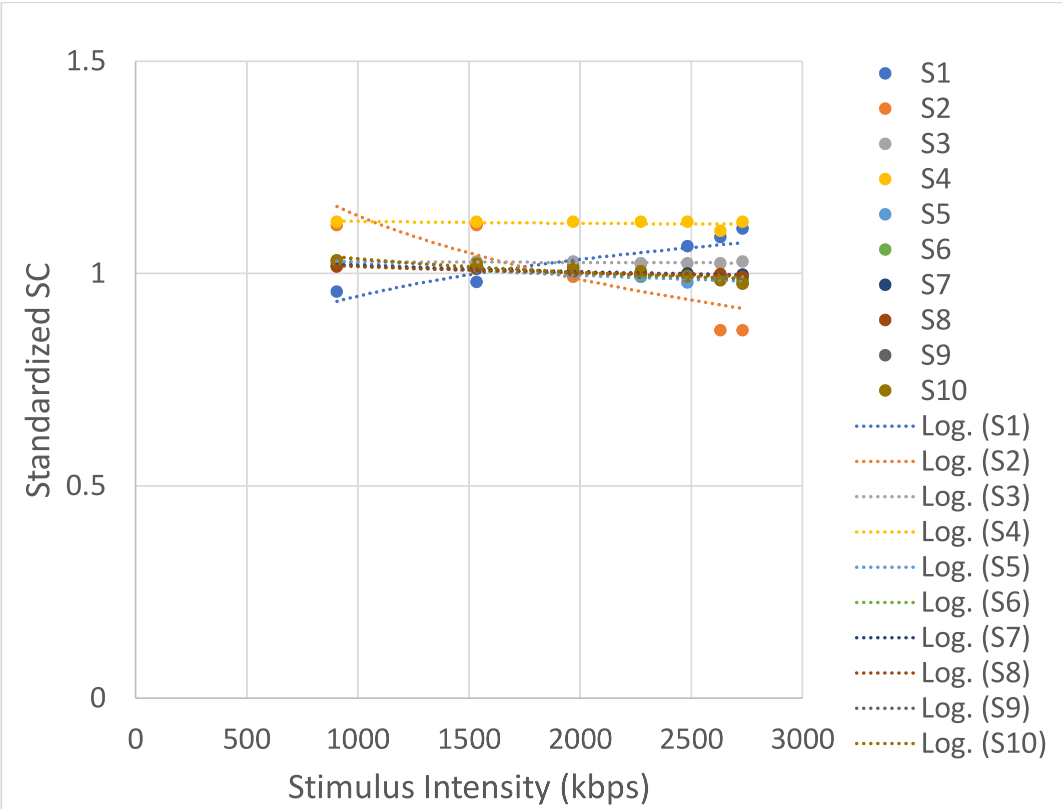

Figure 4.8: The logarithmic nature regression curves of SC data obtained from 10 subjects. $\log S_i$ means the logarithmic nature curves, whereas S_i is the data point of each subject

cluded. Eventually, the only $BR_{threshold}$ can be derived from SC data. According to [16], video player always attempts to maintain a constant gap between the target video rate and the needed bandwidth. Such a constant gap is determined by the following equation (similar to Eq 3.3 in subsection 3.3.2):

$$Cons = \frac{BW - targetBR}{BW} \tag{4.13}$$

where, BW is the available bandwidth, and $targetBR$ is the expected video representation (target video rate) of the users. $Cons$ refers to a conservatism value defined by particular proprietary video players. In fact, Microsoft smooth streaming applies a conservatism value of 20% [16].

Accordingly, the required bandwidth BW for $BR_{threshold}$ is calculated as: $1060.48kbps \leq BW \leq 1238.75kbps$. On the other hand, according to [32], by using PSQA model, the values of QoS parameters (bandwidth, packet loss, delay and jitter) can be interpreted to MOS, and vice versa. Therefore, this constraint is then converted into optimal MOS constraint:$2.78 \leq MOS_{threshold} \leq 2.91$ with the assumption that packet loss, delay and jitter are negligible.

4.6 Discussion

Based on the numerical results, the hypotheses have been successfully confirmed followed by the determination of an optimal constraint of video rate threshold. In this section, it is important to discuss the validation of this optimal constraint of MOS, and thus, a simple experiment was conducted. In this experiment, ten subjects (they are different from the subjects in the experiment in subsection 4.5.1) were asked to watch a short movie with the duration of 2 minutes, then, to provide their subjective evaluation following the 5-scale MOS. There were two scenarios in this experiment in order to validate the performance of optimal $MOS_{threshold}$ in comparison with the fair $MOS_{threshold}$. In both scenarios, the movie's video rate was gradually decreased in response to the negatively changing the network QoS parameters (bandwidth, packet loss, delay and jitter), while the estimated MOS was being calculated by PSQA model [32][105][73]. When the estimated MOS decreases to $MOS_{threshold}$, the control action will be triggered (control action is activated by allocating the higher bandwidth to the user [106]). In this case, the optimal $MOS_{threshold}$ was equal to 2.78 which is the lowest value of optimal constraint. The fair $MOS_{threshold}$ was equal to 3 which is defined as the fair value in 5-scale MOS [67][73][74]. The subjective MOS collected from subjects in the first scenario was compared with the one obtained in the second scenario. The validation criteria were the similarity of subjective MOS and network resource utilization in both scenarios.

To investigate the similarity of subjective MOS, the analysis of variance (ANOVA) approach [107][108], especially one-way ANOVA which is a technique for comparison means of two or more samples (using F distribution), was accomplished. The result of ANOVA indicates that the overall subjective MOS which were obtained from all the subjects in both scenarios were equal.

Now resource utilization is in turn considered. Thus, the required bandwidth allocations, which are sufficient for the users to experience the video quality with MOS higher than optimal and fair $MOS_{threshold}$, were investigated. As mentioned in section 4.5.3, the required bandwidth for $MOS_{threshold}$ of 2.78 was equal to $1060.48kbps$. Assume that the values of QoS parameters, except bandwidth, were negligible during the experiments. By using the PSQA model, the required bandwidth for fair $MOS_{threshold}$ of 3 was calculated to be equal to $1303.23kbps$. The setup of this experiment was the same as those of our previous works in [96], where the link capacity was limited to $5000kbps$. Therefore, it is clear to see that in terms of resource utilization, using the optimal $MOS_{threshold}$ can save at least $\frac{1303.23-1060.48}{5000} * 100 = 4.855\%$ of bandwidth allocation per control compared to the fair $MOS_{threshold}$. In terms of resource utilization, this result might be modest. However, it will be significantly meaningful for the practical system which is much larger and more complicated than the one in this research.

The validation result shows that the constraint of QoE threshold optimally satisfies the research purpose of QoE management, that is to say, ensuring that QoE is stable at an expected level with minimal network resource usage. Therefore, using biological information produces a prominent result in modeling a logarithmic relation of human perception and stimulus intensity.

However, there are some other issues also need to be discussed. Firstly, the particular regression analysis of either SC data or HR data or HRV data obtained from each subject did not always offer the expected result with sufficient accuracy. For SC data, the results from some subjects surprisingly produced very low correlation of determination, especially subject 4. According to the result of DDA in subsection 4.5.3, subject 4 also did not produce the significant SCRs for most of intensities. Thus, during the experiment, this subject maybe did not concentrate or got some invisible distractions. For HR and HRV data, although using HR/HRV monitoring device can avoid the intrusiveness which was recognized as one of the major limitation of psychophysiology in [38], only 30% of subjects produced high accuracy regression curves leading to their exclusion of this research. As stated in subsection 4.5.1, the sampling frequency of HR was only 1Hz and was fixed by the device vendor. In addition, the value of HR obtained from monitoring device is the average value. Therefore, HR measurement cannot capture significant biological characteristics, resulting in a miserable regression analysis. In addition, the heart response time to sympathetic stimulation is relatively slow. It takes about 5 seconds to increase HR after the actual onset of sympathetic stimulation and almost 30 seconds to reach its peak steady level [89][109]. Therefore, analyzing short-term HR and HRV (within 5-second time window) provide insufficient accurate results, requiring a new experimental scenario with longer video sequences.

Human characteristic is another discussion point. In the perceptual process, sensation refers to the initial steps - converting physical features of the environment into electrochemical signals within specialized nerve cells and sending those signals to the brain for processing. Meanwhile, perception refers to the last steps, whereby the initial sensory signals are used to form mental representations of the objects and events in a scene so that they can be recognized. Therefore, the electrical activity of the brain measured by Electroencephalography (EEG) should be taken into account in this research. The combination of EEG associated with Central Nervous System (CNS) and measurement methods associated with Autonomic Nervous System (ANS) will promisingly produce a better result in modeling the relation between perception and stimulus intensity applying in QoE management research.

4.7 Summary

In this chapter the relation between biological information (SC data) and stimulus intensity has been modeled as a general logarithmic nature function. Thereby, the optimal constraint of QoE threshold has been determined and validated. The results of validation show that by using the determined threshold, the overall QoE is guaranteed to be stable at high level, while the network resource utilization is impressively improved. The obtained threshold constraint is suitable for only the scenario in this research, but the approach can be applied in more general cases. This research also confirms the feasibility of applying biological information in QoE management.

Chapter 5

User-centric Approach to Accurate Bandwidth Allocation

5.1 Introduction

The proposals presented in chapter 3 and chapter 4 have efficiently solved the issues of early detection of QoE deterioration and generating control action at the right time. This chapter introduces a proposed method for the accurate generation of control action, towards the research goal, that is to say, the balance between network resource utilization and the resulting QoE.

In adaptive streaming services, based on the metadata and status of terminal/networks, the decision engine at the video player makes decisions on which/when video chunks are requested and downloaded. As the results, it optimizes the server-side scalability and provides smoother user experience. However, when the underlying network condition fluctuates for some reasons (e.g., bandwidth competition among video players [72]), video rate will vary more frequently, resulting in QoE deterioration. In this situation, it is necessary to guarantee a specific video rate level for the end-user, especially the premium user who pays additional cost for their service. To do so, accurately triggering a control strategy is crucial. Eventually, not only the perceived video quality is guaranteed, but also the network resource is saved.

Shaping traffic [59][110][106][16] is known as the most common control action type by which available bandwidth is allocated to the user. This allows the

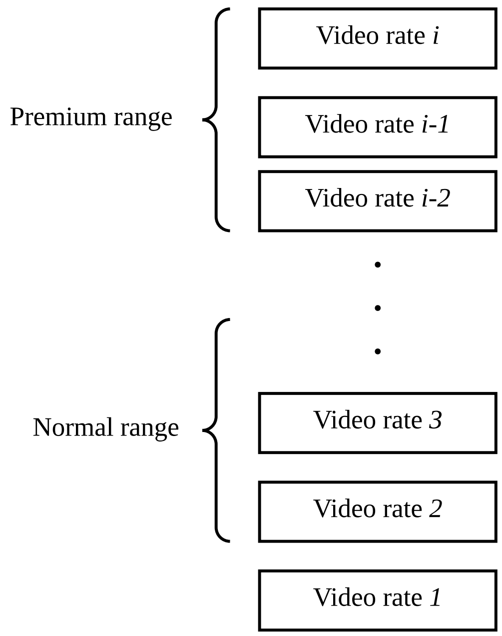

Figure 5.1: The pre-defined range of video rate. The establishment of this range is relied on the MPD

end-user to experience video with an expected video rate after several requests, improving his/her QoE. In order to achieve the optimal trade-off between network resource utilization and maintaining QoE, the systems need to precisely determine the allocated bandwidth. The authors in [16] proposed a method to determine the allocated bandwidth based on a certain target video rate, i.e., expected encoding video rate of the end-user. As the result, the target video rate was successfully requested by video player, after a specific delay. However, the method to determine the target video rate, has not been clearly stated. In fact, the target video rate can be randomly taken from pre-defined ranges in accordance with the classification of the end-users, as shown in Fig. 5.1. However, without taking into consideration the other QoE influence factors, such the predefine ranges of video rate do not precisely reflect the end-users expectations (or expected QoE), resulting in some consequences. For example, regardless of terminal display screen, the waste of bandwidth is introduced since the movie is played on a small display screen, whereas, playing the movie on a bigger display screen causes the deterioration of perceived video quality. Therefore, in this study, a novel method is proposed in order to determine the target video rate that closes to human expectation. Thereby, more precise bandwidth allocation will be performed in QoE control. The proposed method is two-fold. First, the numerical ranges of subjective expectation are established for different types of the end-users (the premium users will be focused in this study). Accordingly, the numerical expectation of the premium user will ranges from determined MOS threshold (taken from the studies in the previous chapter) to the highest value of 5. Second, a regression model of video rate and subjective perception (refers to MOS values) is applied to interpret the expected QoE to the target video rate.

In this chapter, the main contribution is to propose an novel bandwidth allocation approach by establishing a numerical range of the user's expectation for determination of target video rate, by applying a regression model of video rate and MOS for calculating the needed bandwidth based on the determined target video rate. Consequently, the proposed approach not only keeps the QoE of the premium user stable at an expected level, but also more bandwidth is saved.

5.2 Related work

In general framework, QoE control is usually considered as control decision, which guarantees QoE not to be deteriorated under a specific threshold. This section provides an overview of control strategies in QoE control. Eventually, shaping traffic will be emphasized as the most common and effective control action in literature.

Accurately triggering control action plays an important role in maintaining the optimal trade-off between network resource utilization and the QoE. Therefore, the contemporary studies are attempting to propose various control approaches and to improve their accuracy. In order to maximize the QoE, some authors proposed a new adaption algorithm [111] or introduced a network proxy to select optimal video rate for the end-user [112]. Meanwhile the authors in [113] proposed QoE control by adding Forward Error Correction (FEC) packets to the current flow which is capable of compensating for the packet loss. In addition, a new control approach was also introduced in this study. Particularly, a trained neural network was used to decide the most suitable control action for the current contexts. Notably, there are existing studies including our works, indicate shaping traffic as an effective control strategy to maintain QoE. When the available bandwidth shrinks due to the bandwidth competition [72], the unfairness and instability problems of the requested video rate will occur. The traffic shaping method effectively solves these problems [59][110][106][16]. Particularly, in [16], the authors proposed a method to identify how much available bandwidth is needed for the traffic shaping. Each commercial HAS player actually has their own safety margin to ensure the available bandwidth is enough for the next encoding video rate (e.g., 20% of Microsoft smooth streaming). The result showed that the target video rate was accurately requested by the end-user. However, they did not mention the mechanism determining the target video rate, thus, the easiest way is to simply pick up the highest possible encoding video rate in a pre-defined range.

In this chapter's proposal, the clarified relation between subjective MOS and requested video rate, which captured by a regression model, has been applied. By which it allows calculating the target video rate 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 video rate which is equal to calculated target video rate after several requests.

5.3 Bandwidth Competition in Adaptive Streaming Services

One of the reasons causing the QoE degradations in a shared network during a streaming session is bandwidth competition [106][48][59][114][115]. Typically, the bandwidth competition is expressed as either the interplay among several video players or the interplay between a video player and other applications. When two or more video players start to compete for the limited network bandwidth, a series of performance problem such as unfair sharing and video rate fluctuation will consequently occur. In other words, the major performance issues for competing video players within a network are the instability and the unfairness. The instability refers to the frequent switching video rate of the video player. Several studies show that more frequent quality switching is invoked when more than one instance of video players compete for bottleneck bandwidth [11][106]. As sharing a bottleneck link, multiple competing players converges to an unequitable allocation of the network resources, which causes the unfairness in adaptive streaming services.

In order to confirm the abovementioned problems, a simple experiment was conducted. In a certain time, two users were required to watch the sample movie. In this experiment, Microsoft smooth streaming player which is an IIS Media Services extension, enable adaptive streaming over HTTP, will be used. Wireshark- a monitoring software will capture the requested video rate through analyzing HTTP header. The experiment was repeated for several times to indicate the stochastic nature of the issues. Figure 5.2 illustrates the unfair sharing of bandwidth among clients. Specifically, the playback rate of the user 2 were 630$kbps$ constantly, whereas the user 1 experienced a happy time with high video rate (from 1500$kbps$ to 2500$kbps$). If the user 2 is a premium user, then this is a big problem for the content provider.

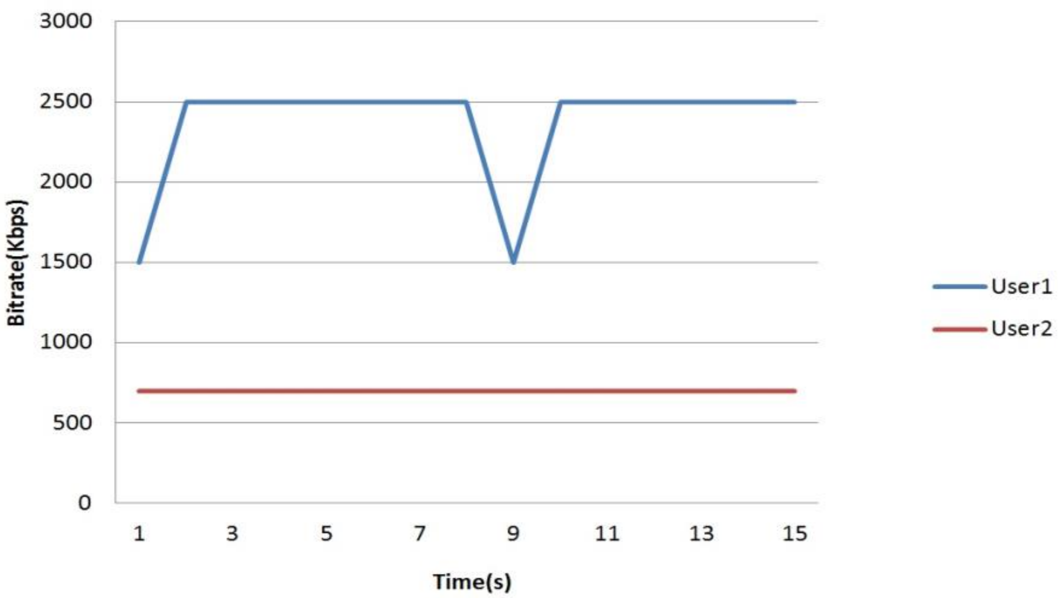

Figure 5.2: The unfair sharing of bandwidth among clients

The instability of video rate during a streaming session can be seen in Fig.5.3. After gradually increasing to the highest value of 2500$kbps$, the playback rate of the user 2 rapidly decreased to the lowest value of 630$kbps$ after several seconds. Such high video rate fluctuations may result in visible variations in the content quality for this user. The positive point here is that two users shared the bandwidth equally on average. Actually, the unfairness and frequent adaptation are caused not by TCPs congestion control but by the offered load that each video player requests. For instance, in the Fig.5.2, the user 2 estimates the

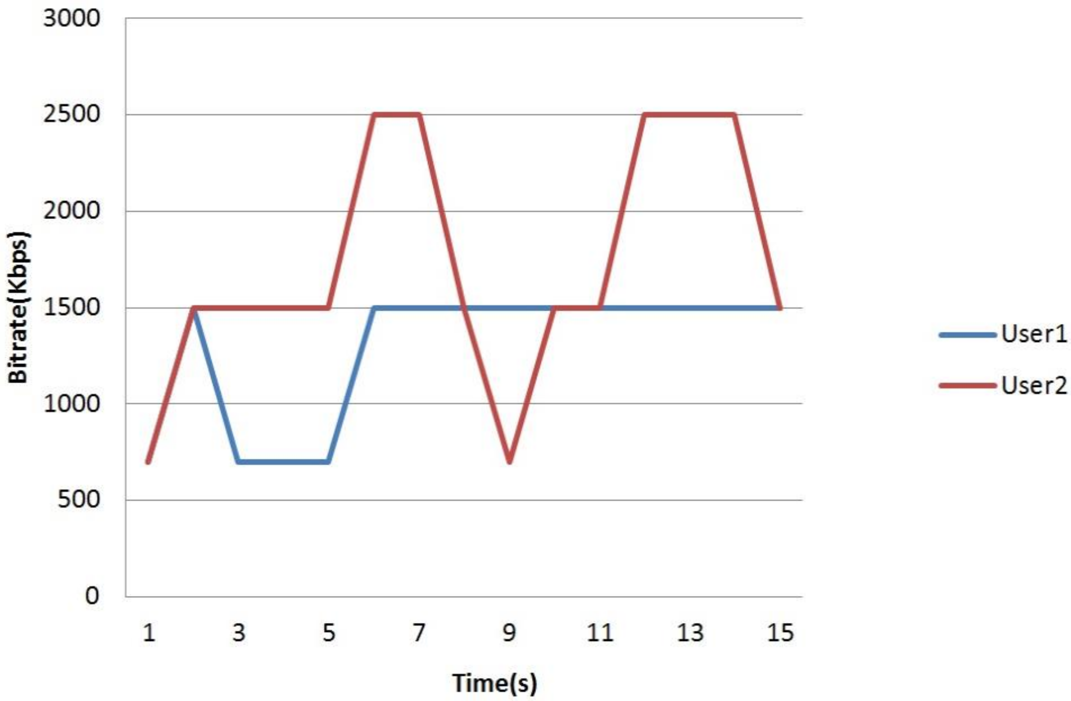

Figure 5.3: The instability of video rate during a streaming session

available bandwidth to be much lower, and it does not even try to increase its requested video rate. [72] shows that the root reasons of the three performance issues (instability, unfairness, and underutilization) caused by the temporal overlap requests among video players. As stated in section 2.3, during the streaming session, video player usually experiences two state: Buffering-state and Steady-state. In buffering-state, video players start to build up its playback buffer as quickly as possible. After playback buffer size reaches a certain specified value, Steady-state is triggered. The Steady-state includes ON and OFF periods. Video players download the next fragment of content based on the available bandwidth estimation in the ON period. Video players stay in idle mode in the OFF period. In other words, the ON period is the downloading time and OFF period is the playback time. When two or more users started to watch a movie, the temporal overlap of the ON-OFF periods among video players can cause the incorrect bandwidth estimation.

The impact of other applications on the quality of adaptive streaming services, has not been widely addressed yet. In the presence of competing flow of other applications clients, the video player often suffers a dramatic anomalous drop in the video playback rate. As described in the previous discussion, after the buffer becomes full, the video player enters a periodic ON-OFF sequence. When there is a competing flow, this flow will fill the buffer during the OFF period of the HAS player, and thus, the video flow detects very high packet loss rate, and low available bandwidth. This problem will repeat for every ON- OFF period resulting in low QoE of the HAS player. Therefore, performing appropriate bandwidth allocation to the users promisingly eliminate of negative effect of bandwidth competition.

5.4 Methodology

In this section, initially, QoE management algorithm is presented in order to emphasize the role of QoE control in maintaining the balance between network resource utilization and the QoE. Afterward, the proposed method for accurate bandwidth allocation will be introduced alongside the brief descriptions about its operations in QoE control.

5.4.1 QoE control in adaptive streaming services

QoE management algorithm in adaptive streaming services is depicted in Fig.5.4, where QoE control takes a key role. Initially, all incoming traffic from the ISPs network will be classified in order to identify the application behinds each relevant traffic flow. By knowing the application of the flow, the system can perform the suitable admission control for that flow. In fact, the admission control operates on a flow-basic. Since a new flow is about to enter the QoE management entity, the system decides whether theres enough capacity to fully support this flow. If it beyond of system capacity, allowing a new flow to enter would cause the quality degradations of the existing flows. Afterward, the perceptual quality is estimated for each flow, indicating the situation of the flow. In this case, MOS is used as a way to express the estimated quality quantitatively as the average opinion of a group of the users. The estimated MOS is determined as output of PSQA [32][31] - a machine learning model. In a simple scenario, assume that there is only one class of the end-users, that is to say, the premium class. In this class, the priorities of the end-users are equal. Thus, only one threshold has been used to judge the current situation of estimated QoE. In a more complex scenario, it is

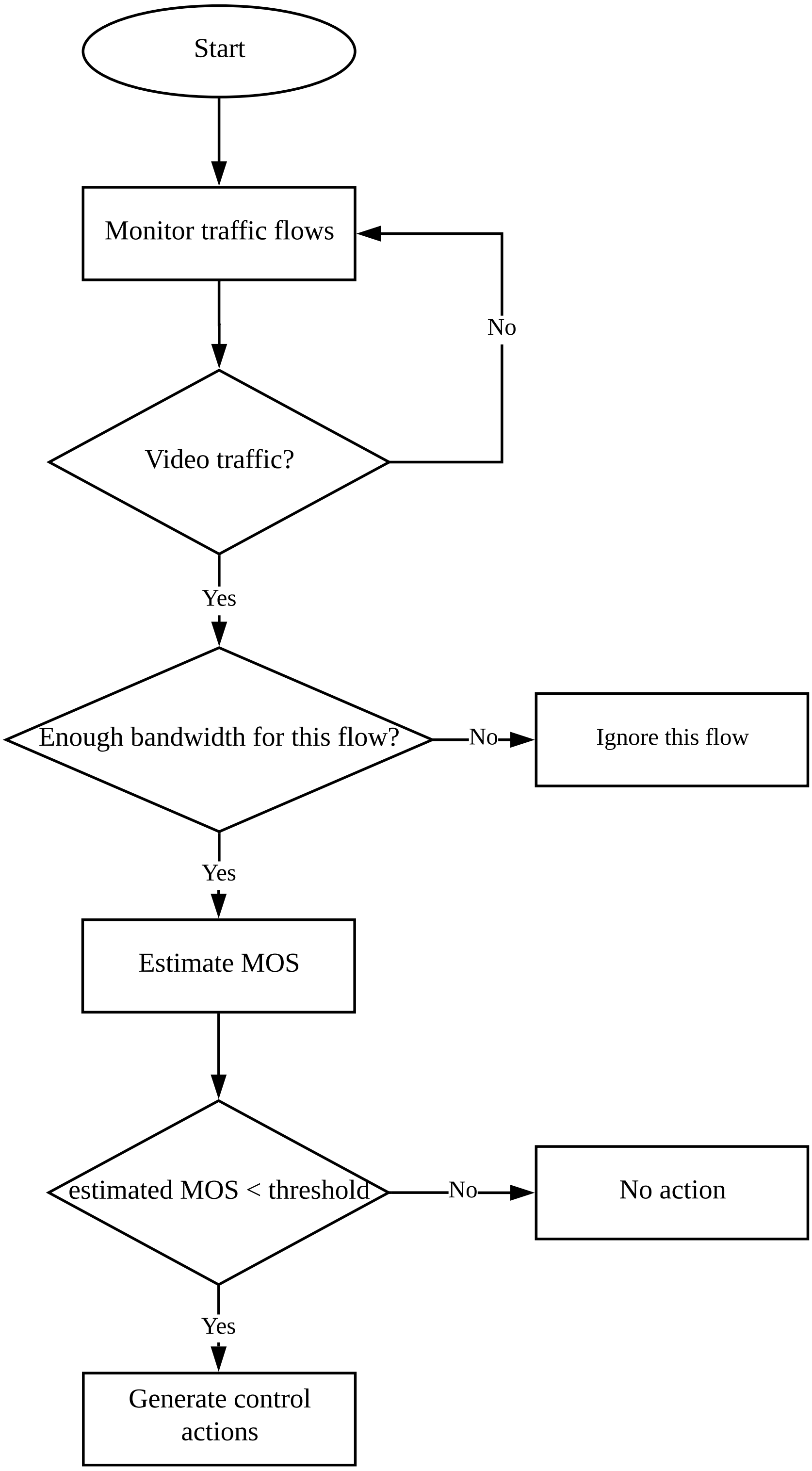

Figure 5.4: QoE management algorithm. MOS is always kept stable at a level which is higher than threshold

necessary to classify the users into premium class and normal class. Each class has the different priority. Particularly, the premium class takes the highest priority, meanwhile the default priority is given to the normal class. Therefore, at least two types of threshold should be considered. While a high threshold is for classifying the premium class, the normal class is judged by lower threshold. In this research, only algorithm for the simple scenario has been taken into consideration. If the estimated MOS decreases below the high threshold, the system generates the suitable control actions to guarantee an expected MOS for that flow. As discussed in previous sections, bandwidth allocation is used as the major control action in this study. The focus of this action is to assign a specific bandwidth to certain users or services. It can also be used to throttle down other traffic in order to give better performance to certain users or services. In contemporary studies, by relying on the target video rate, the necessary bandwidth can be calculated and allocated to the end-user. Technically, bandwidth allocation as a control action can be performed through two steps: (1) determine the target video rate and (2) calculate and allocate the needed bandwidth.

- Target video rate is randomly taken from the premium range of available video rate, which is depicted in Fig. 5.5. There are some significant drawbacks found in this approach. First, each video content might not be encoded in the same way, resulting in different range of available video rate at the server. In other words, the premium range might vary among video contents, resulting in inaccurate determination of target video rate. Second, the range of video rate does not accurately reflect the subjective expectation.

- The needed bandwidth can be calculated directly based on the constant gap presented in [16].

5.4.2 Proposed method

In this subsection, the proposed method for accurate bandwidth allocation in QoE control will be presented. Similarly to the existing QoE control's mechanism, the proposed method encompasses two steps: (1) Determine the target video rate and (2) calculate and allocate the needed bandwidth to end-user. While, the second

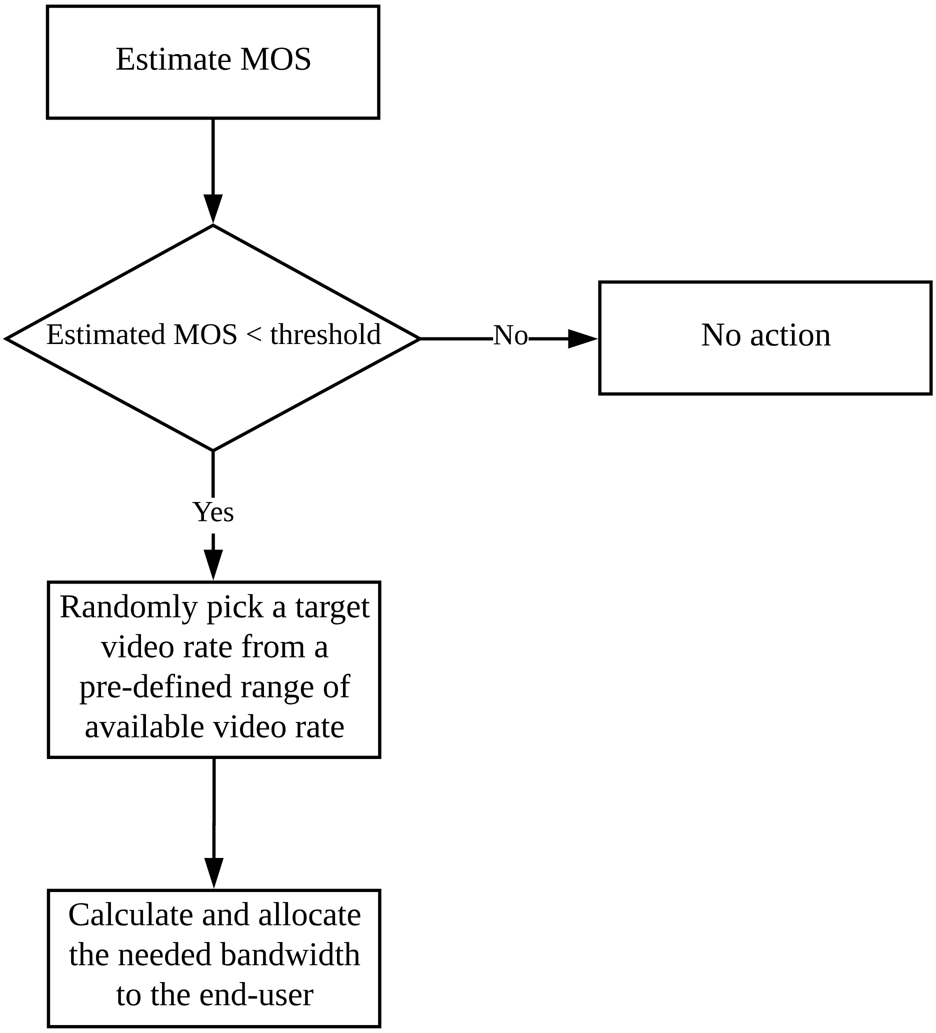

Figure 5.5: The exiting QoE control approach, where the target video rate is randomly taken from a pre-defined premium range of available video rate

step is done in the same way as the existing approach, the novelty of this study will be found in the first step.

1. In this approach, the target video rate will be determined through the following two-phase procedure (as shown in Fig. 5.6):

    - Define the range of the premium users expectation
    - Interpret the subjective expectation to target video rate

    The first phase encompasses the determination the range of subjective expectation which is numerically referred to MOS (as QoE indicator). This study is actually assumed to solely focus on the premium user, therefore, the ideal expectation values is ranged from MOS threshold to the highest MOS value of 5. In chapter 4, the relation between human perception and biological information has already been modeled as a nature logarithmic function. Thereby, a precise QoE assessment was achieved regardless of individual different as well as the biased and variability. Accordingly, an optimal constraint of MOS threshold has been established. In this study, such the constraint is then used to clarify the range of expected MOS of the premium user, as depicted in Fig. 5.7. Therefore, for any value of MOS ranging from MOS threshold (derived from optimal constraint) to the highest value of 5 is consider as the expected perceived video quality of the premium user. As the result, once the estimated MOS of a premium user falls below the threshold, a random value in that range will be taken as the subjective expectation for that user. This value will be pushed to the second phase for its in turn operations.

    The second phase is responsible for interpreting the expected MOS to the target video rate which is in turn used for calculating and allocating bandwidth to the end-users. To do so, this study proposes to use a regression model that expresses the relation between video rate and subjective perceived quality. By using this model, the expected MOS can be automatically interpreted to the according target video rate in real-time. In general, the regression model is established based on the dataset of subjective MOS and requested video rate derived from the experiment presented in subsection 3.3.1 for the establishment of PSQA. In this experiment, network conditions are varied by adjusting QoS parameters, producing distorted videos.

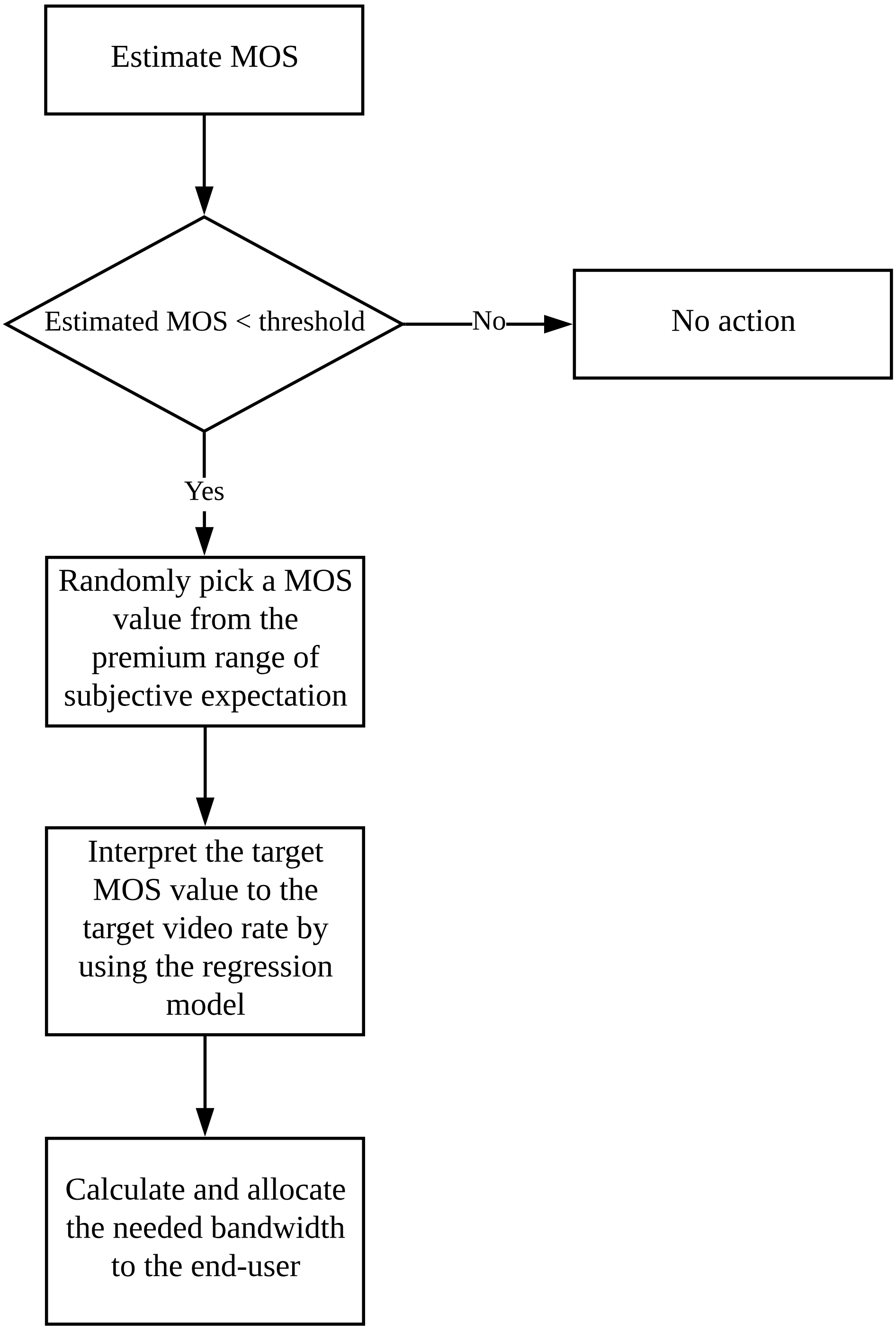

Figure 5.6: The proposed method for precise bandwidth allocation

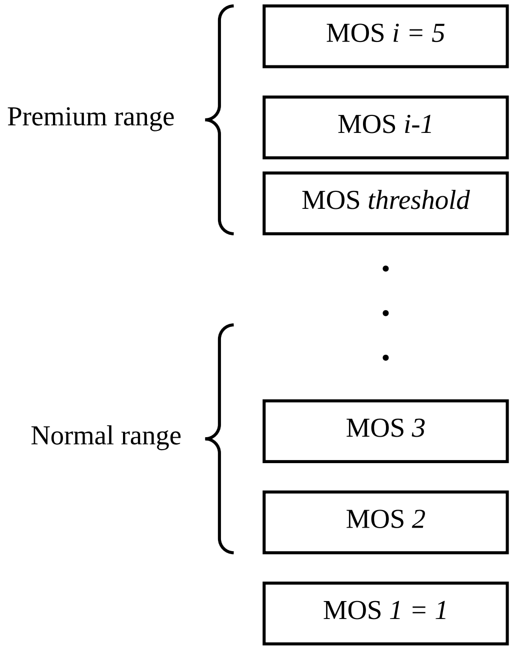

Figure 5.7: A pre-define range of subjective expectation in terms of expected MOS

Subjects who are asked to watch these videos, provide their evaluation in terms of subjective MOS. Alternatively, the video rate which is requested by video player, is continuously monitored. Assume that there is only one type of terminal display screen size used in this experiment. The sample dataset which is obtained from that experiment, is illustrated in Table 5.1. Accordingly, the average value of all subjects evaluations refers to subjective MOS. Meanwhile, the requested video rate is the most frequently requested one during streaming sessions for each distorted video.

The tendency of the correlation between subjective MOS and requested video rate is depicted in Fig.5.8. In general, the MOS linearly increases in accordance with the requested video rate since video rate has a strong correlation with perceived video quality. Particularly, once MOS value increases above the threshold, its tendency becomes more obvious. Before that it witnesses a frequent variation of MOS for the video rate that is below $1024kbps$. Visually, the MOS mostly ranges from 4 to 5 if the requested video rate is varied from $1536kbps$ to $2048kbps$. It is worth noting that the video contents do not always have the same encoding way, thus, the

Table 5.1: The most frequent requested video rate and the subjective MOS

Available bandwidth (kbps)	Packet loss (%)	Latency (ms)	Jitter (ms)	Subjective MOS	Requested video rate (kbps)
700	0	10	5	1.88	331
1000	3	250	5	2.75	447
2000	0	1	5	4.63	1427
4000	3	250	5	3.12	991

according requested video rate for particular MOS value, might be varied.

From raw data, a regression model was also established by modelling data with linear combination of basic function in which a set of functions ϕ_0, ϕ_1,, ϕ_P was specified along with finding function f in the form of linear combination:

$$f(x) = \sum_{i=0}^{P} \theta_i \phi_i(x) \tag{5.1}$$

In this case, the target function for the regression model is unknown, thus, the Gaussian Radial Basic Functions (RBF) was chosen. The result of modeling data is plotted in 5.9 with high coefficient of determination denoted as R-squared of 0.8531. Once the target video rate is determined based on a certain expected subjective MOS, the available bandwidth can be calculated and allocated to the end-user.

2. Based on determined target video rate, the bandwidth can be calculated through Eq. 4.13. This equation presents a constant gap between the target video rate and the needed bandwidth. Typically, *Cons* refers to a conservatism value (constant gap) defined by particular proprietary video players. For instance, Microsoft smooth streaming (MSS) and HLS players apply a conservatism value of 20% and 40%, respectively. In this study, MSS player with the constant gap of 20% is practically used.

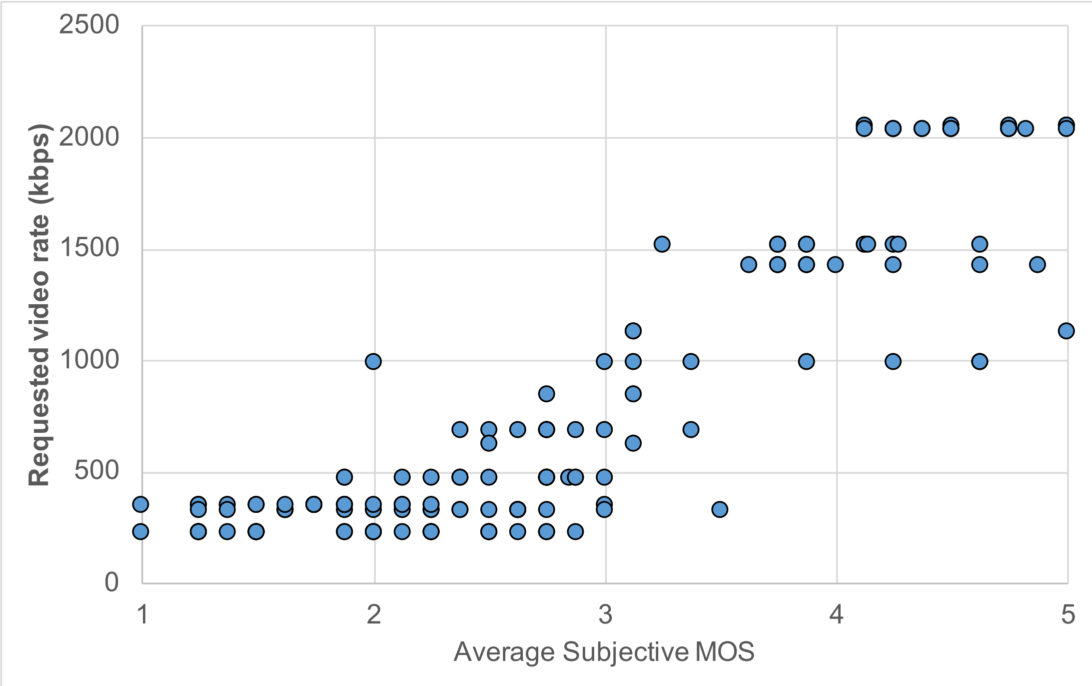

Figure 5.8: The general trend obtained from subjective MOS and requested video rate

5.5 Evaluation

This section aims to investigate how accurately can bandwidth allocation be performed when the perceived quality of the premium user falls to specific lower levels. Note that, in the scenarios that estimated MOS was gradually decreased, QoE control was activated only when the estimated MOS was below the threshold. In this case, the MOS threshold of 2.78 was derived from previous chapter. 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 3.10GHz processor and 8GB 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 3.10Ghz processor and 8GB RAM. The server published a Microsoft smooth streaming video content of "Big Buck Bunny" which is known as an open source testing movie. This movie content was encoded with multiple video rates. Furthermore, a Smooth Streaming-compatible Silverlight player template was installed on the Smooth Streaming en-

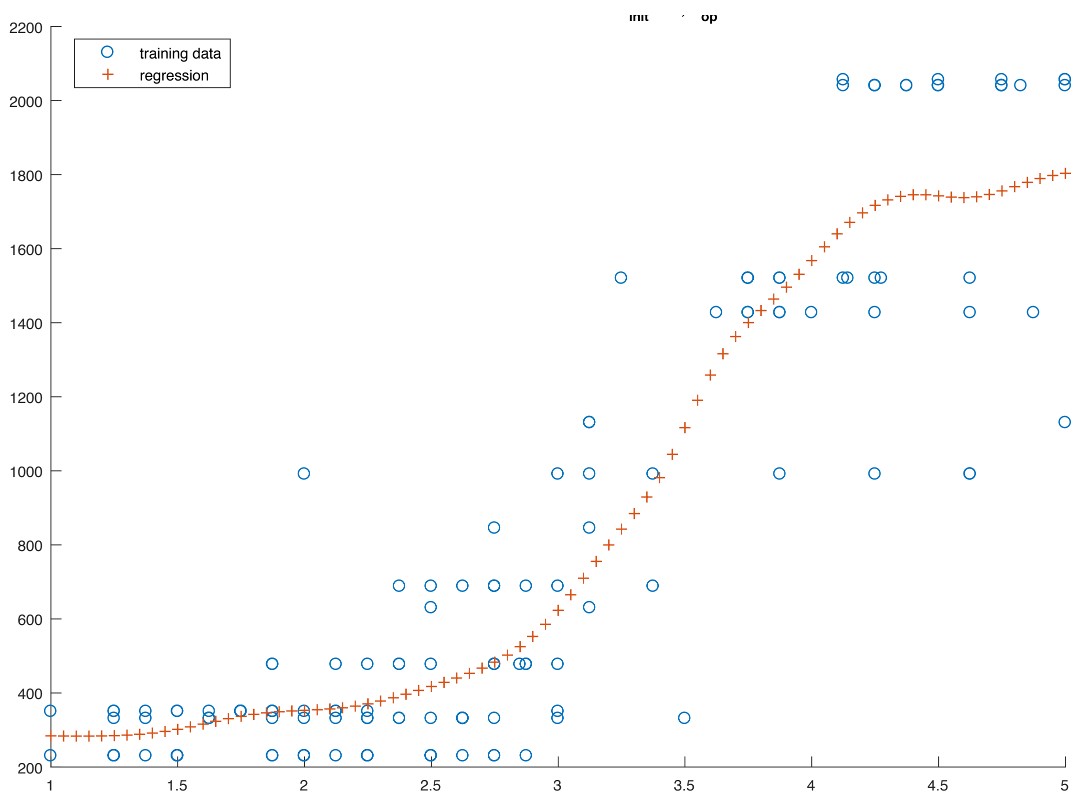

Figure 5.9: Regression approximation modeling the relation between requested video rate and subjective MOS. The horizontal axis presents the subjective MOS. The vertical axis shows the requested video rate (in kbps)

abled streaming server so that Silverlight-based users can play Smooth Streams. The users utilized the laptop computers with MacOS, Core i5 and 8GB 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.5.10. 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.

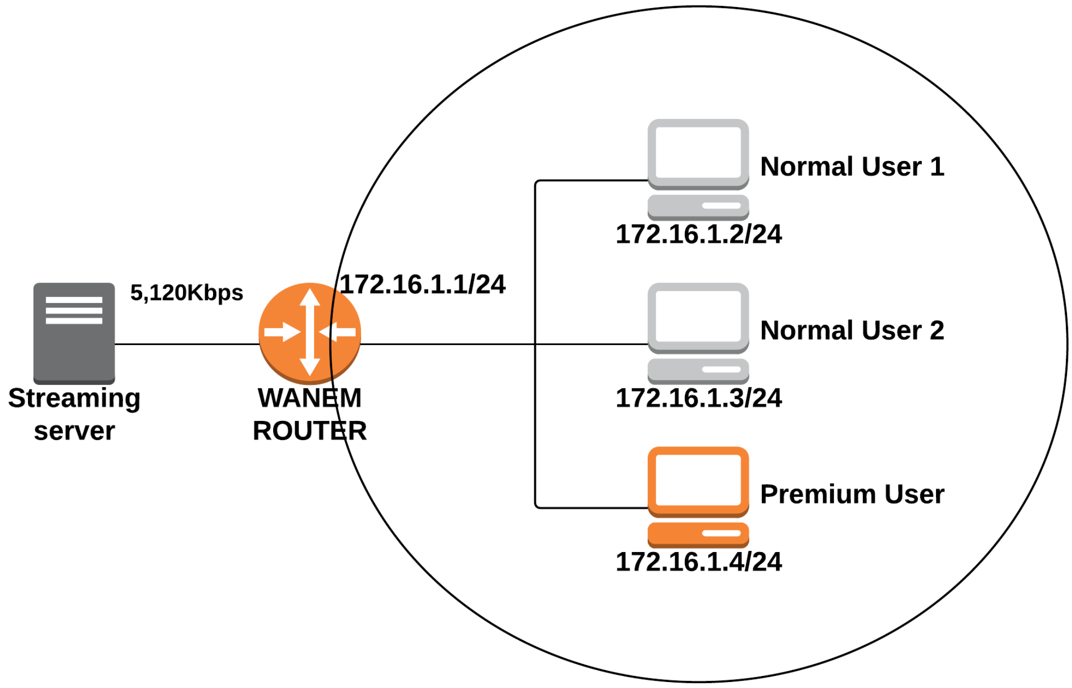

Figure 5.10: Set up network environment for investigating the accuracy of control action

The experimental procedure for two scenarios of the evaluation is 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 video rate and the estimated MOS.

(5) The available bandwidth to the premium user is increased to recover the network quality when the deterioration of requested video rate and estimated MOS are detected.

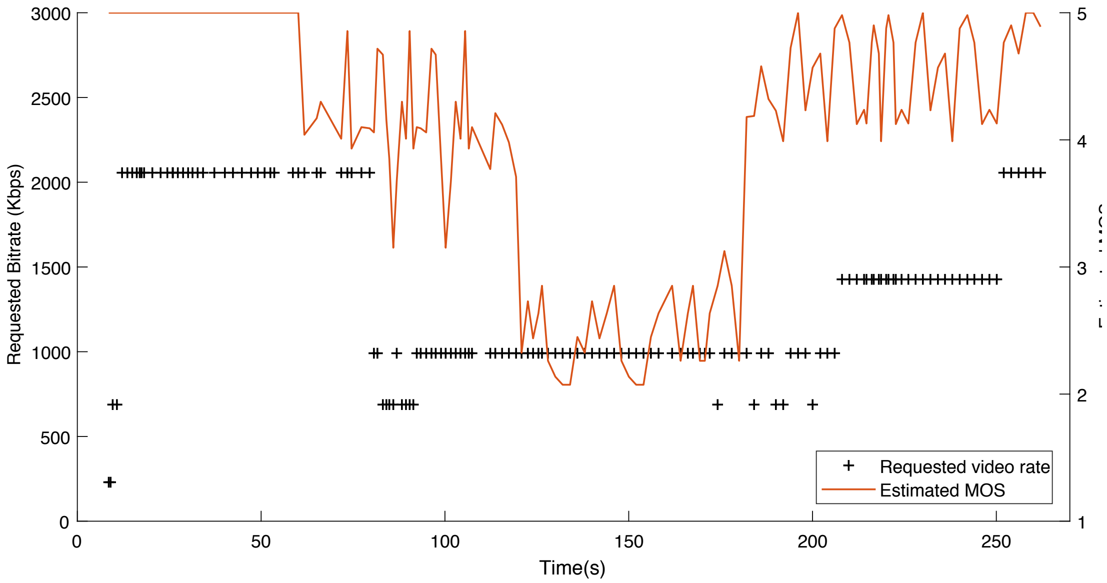

Figure 5.11: Requested video rate and estimated MOS of the premium user during his streaming session (before and when bandwidth competition occurs)

Figure 5.11 shows the changes of the estimated MOS and the requested video rate of the premium user under different situations. Accordingly, the horizontal axis indicates the time durations, the first vertical axis is the requested video rate, 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 took several transitions to reach the highest encoding video rate of 2056$kbps$. The player would have requested the value of 2962$kbps$ if the users were watching the video in full-screen mode. Meanwhile, the estimated MOS was stable at the highest value of 5 since the available bandwidth was equal to the link capacity of about 5000kbps. 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 2500$kbps$. The estimated MOS decreased to around 4, whereas the player still requested the video content at video rate of 2056$kbps$ for 19 seconds. When the second normal user watched the video, the available bandwidth of the premium user shrank to 1578$kbps$ followed by a fluctuation around the value of 2.5 of the estimated MOS, which is below the set-up threshold of 2.78. At this time, the system took the highest MOS value of 5 for the premium users expectation. Note that, the exist studies directly take the highest video rate value of 2962$kbps$ for the premium user instead of consider the subjective expectation. Afterward, by using the regression model, the target video rate can be calculated from the expected subjective MOS followed by the calculation of the available bandwidth based on Eq. 4.13. The determined target video rate was actually equal to 2056$kbps$. Interestingly, at t=242s, the player successfully reached to its calculated target video rate of 2056$kbps$.

Fig.5.12 and Fig.5.13 show the estimated MOS and the requested video rate of both two normal users. Particularly, in Fig.5.12, the estimated MOS of the first normal user fluctuated around the tolerable levels of 2 and 2.5 from t=60s to t=262s. On the other hand, the requested video rate was stable around 1130$kbps$ during the time it has to compete for the available bandwidth with only the premium user. This is followed by a slightly decrease to 688$kbps$ since the second normal user participated in the network. A similar tendency of the second normal user also can be seen in Fig.5.13.

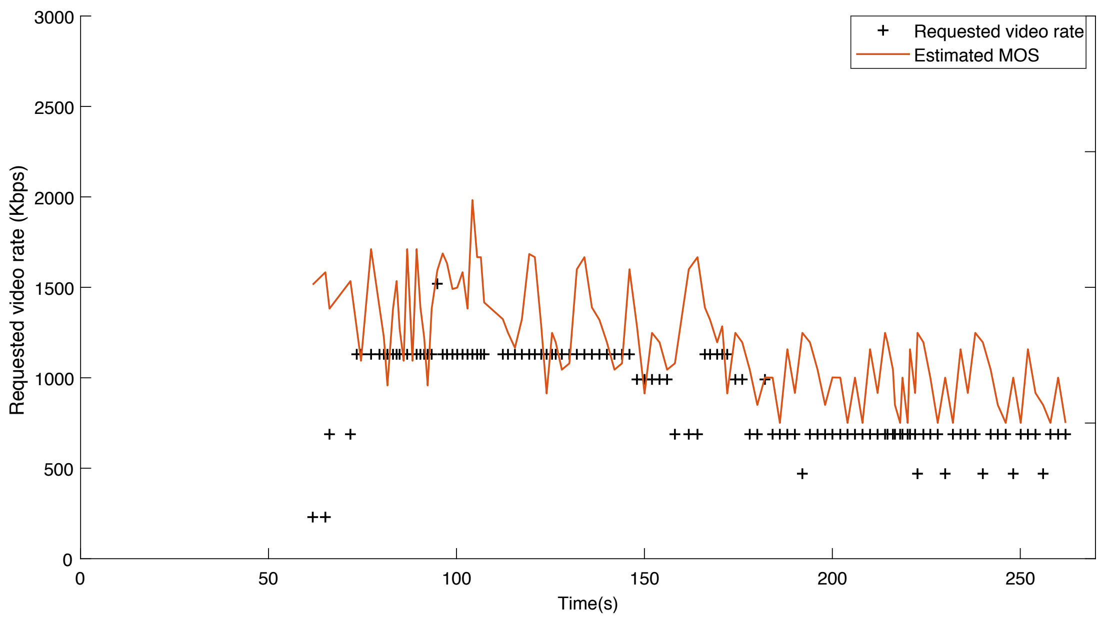

Figure 5.12: Requested video rate and estimated MOS of the first normal user during his streaming session

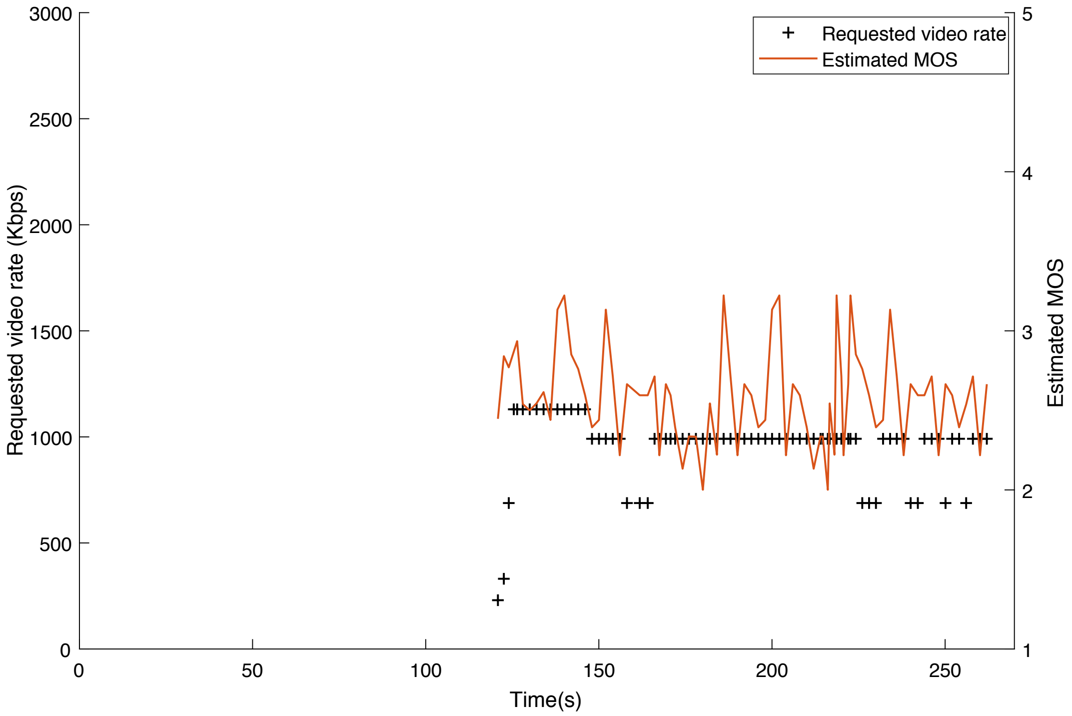

Figure 5.13: Requested video rate and estimated MOS of the second normal user during his streaming session

5.6 Summary

The purpose of the research in this chapter is to propose a user-centric manner in performing accurate bandwidth allocation. The experimental results show that the appropriate target video rate can be accurately predicted from the expected MOS when the estimated MOS is less than the threshold. As the result, the needed bandwidth can be efficiently assigned to the premium user. Applying the proposed method not only supports the premium user to achieve its expected perceived video quality, but also optimizes the network utilization and provides sufficient available bandwidth to the normal users for tolerable quality. As shown in Fig.5.12 and Fig.5.13, although there is a serious competition in the network, the estimated MOS of both normal users are still stable at tolerable levels. However, this approach is currently applied for Microsoft Smooth Streaming technology only, thus, it is necessary to be re-confirmed with other commercial technologies as well as non-commercial ones. In addition, the larger number of the premium user also need to be considered.

Chapter 6

Discussion

This chapter discusses the works investigated and solutions proposed in this dissertation by which advantages as well as the remaining issues will be summarized.

As mentioned in this dissertation, video services' market witnesses a remarkable shift from technical quality requirements to perceived video quality which is defined as QoE. Thus, in order to expand the market and improve the profit, adaptive video service providers must take QoE management into consideration as the indispensable concern. Currently, adaptive streaming technique has attracted a large attention from academic communities, resulting in numerous publications. These studies mostly focus on optimize the balance between network resource utilization and the resulting QoE. However, besides advantages utilized from exiting works, several issues remain waiting to be solved. These issues are separately associated with two major components of QoE management - QoE monitoring and QoE control. The issues and according proposed solutions are summarized and concretely discussed as follows:

6.1 Early detection of QoE deterioration with appropriate monitoring interval

In the first proposal, the monitoring interval in QoE monitoring was determined by taking into consideration playback buffer size obtained in video player. The balance between the computational cost and the ratio of QoE deterioration was achieved when the monitoring interval was equal to video chunk size of 2 seconds.

However, in this research, video rate was solely assumed as perceived video quality, excluding other perceptual factors such as rebuffering, frequency of rebuffering, and initial starting time. In addition, the proposal has been evaluated with Microsoft smooth streaming decoder only. The confirmation with the other decoders (e.g. YouTube, HTTP Live Streaming, etc.) is needed. Moreover, in the experimental evaluation, the effectiveness of the determined interval has been confirmed with only one client as premium user. The impact of fluctuation of network condition is excluded making the experimental scenario become easier but not realistic.

6.2 Collaborative approach using psychophysiology and psychophysics for determination of QoE threshold

As the most important component in QoE management framework, QoE control is demanded to precisely generate control actions at the right time in order to ensure that the QoE is stable at an expected level with minimal network resource usage. To achieve this, the determination of QoE threshold is needed. In literature, the QoE threshold is simply determined by selecting the fair quality in 5-scale MOS which is common QoE indicator for rating assessment approach. However, the rating approach has significant drawbacks due to high bias and variability. In that situation, psychophysiology provides a better way to quantify human perception by relying on biological information. The significant drawback of this approach is the individual difference that can be compensate by psychophysics. However, psychophysics uses the same perception methods as the rating approach. It leads to a demand of an approach that can harmonize those things. Therefore, a collaborative approach using psychophysiology and psychophysics was proposed in determining the QoE threshold. As the result, the logarithmic nature function representing the relation between human perception (measured by skin conductance) and physical stimulus (defined by video rate deterioration) has been successfully established. The optimal QoE threshold was then determined through an optimal constraint of absolute threshold derived from the above logarithmic nature function. The effectiveness of optimal threshold has already been confirmed through a series of experiments. Whereby, at least 4.855% of bandwidth allocation can be saved for each control, meanwhile the overall subjective QoE was kept at an expected level.

The original intention in this research is to investigate the combination of multiple biological signals, that is to say, SC, HR, and HRV. However, HR and HRV data produce the regression results with very low accuracy, arising a question whether or not HR and HRV are suitable for the estimation of human perception to stimulus intensity. Small sampling frequency is predictably the main reason for this case. While SC data was obtained with a flexible sampling frequency (in this case, it is set to 20Hz), HR data's sampling frequency was fixed by device vendor (it was equal to 1Hz). In addition, the response time to sympathetic stimulation is relatively slow. It takes about 5s to increase HR after the actual onset of sympathetic stimulation and almost 30s to reach its peak steady level. Meanwhile, the duration of reference video is about 40s. Therefore, the inaccurate characteristics of HRV and HR are potentially captured.

6.3 User-centric approach to accurate bandwidth allocation

Bandwidth allocation is one of the common actions in QoE control. Although the introduction of adaptive streaming technology produces smoother perceived video quality, the bandwidth competition during streaming sessions still pose the challenges. In that situation, the accurate bandwidth allocation in QoE control will extremely strengthen network utilization. The saved bandwidth is in turns used by other users, improving the overall subjective perception to video quality. In adaptive streaming services, the allocated bandwidth can be calculated based the determination of target video rate due to the fact that video players usually maintain a constant gap between the target video rate and the needed bandwidth. Thus, accurate target video rate prediction will facilitate a more precise bandwidth allocation. In this research, the target video rate is determined based on: (1) the establishment of a pre-defined subjective expectation and (2) the use of regression model of video rate and QoE. The experimental results demonstrated that by using this method, bandwidth utilization has been much more improved, while the QoE of the premium user was recovered to the expected level.

However, this proposal was confirmed in a simple scenario with a small number of the users. Thus, more complicated scenarios must be considered.

Chapter 7

Conclusion and Future Work

This chapter concludes the dissertation and figured out directions for future work

7.1 Conclusion

As analyzed in the chapter 1, the balance of network resource utilization and the resulting QoE must be achieved in QoE management in adaptive streaming services. This has become the motivation for numerous contemporary studies for years. In this research, such the balance has been achieved by solving three major issues associated with QoE monitoring and QoE control. The conclusions of this dissertation are summarized as follows:

Chapter 3 focused on proposing a method to determine the appropriate interval for QoE monitoring. In this chapter, the behavior of playback buffer during the streaming session was clarified. Thereby, this parameter can be used to predict the deterioration of video rate which was directly considered as QoE. By investigating the condition which keeps the playback buffer stable, an appropriate monitoring interval was determined as being equal to the size of video chunk. The determined monitoring interval was then evaluated by comparing with other values of interval in terms of computational cost and ratio of video rate deterioration. The experimental result demonstrated that with the appropriate interval, a low ratio of video rate deterioration (around 11.45% for buffering state and 40% for steady state) and small average CPU Load (about 11.45%) were derived, resulting in an expected balance at the value of 11.45%.

Chapter 4 aimed at proposing a collaborative approach using psychophysiology and psychophysics for the determination of the appropriate threshold in QoE control. More concretely, the QoE threshold was determined by means of clarifying the logarithmic nature function which expresses the relation between human perception and stimulus intensity (defined by video rate deterioration). As the biological information, Skin Conductance, Heart Rate and Heart Rate Variability associated with Autonomic Nervous System (ANS) have been investigated. The evaluation results demonstrate that the QoE management by using the determined threshold can save more than 4.855% of the bandwidth consumption per control, while QoE is guaranteed to be stable at an expected level

Chapter 5 focused on proposing a method to determine the target video rate for precise bandwidth allocation. Basically, bandwidth allocation can be accurately performed based on the target video rate taken from a pre-defined premium range of available video rate. However, the video rate range does not accurately reflect the human expectation. Therefore, in this chapter, a user-centric approach is proposed for a precise bandwidth allocation. This approach initially defines a premium range of subjective expectation in terms of MOS, then establishes a regression model of video rate and MOS to convert the expected MOS to the target video rate. Eventually, the needed bandwidth can be easily calculated from determined target video rate. As the result, by applying the proposed approach, more bandwidth can be saved per control, while the perceived video quality is guaranteed.

7.2 Future Work

For years, the significant evolution of mobile networks has brought not only chances but also challenges for QoE management in adaptive streaming services. Using mobile networks, the users now can enjoy video services with the various of content and mobility. Therefore, QoE management in mobile networks has increasingly become hotter and hotter than ever in academic and industrial field. It creates motivations to consider an extension towards the mobile networks for the future works.

The appropriate QoE monitoring interval in this research was actually validated with Microsoft Smooth Streaming decoder only. Thus, in the next research, it will be re-confirmed with more complicated scenarios and for various types of commercial and open source decoders.

In this research, the appropriate threshold of QoE control was determined by considering the combination of psychophysiology and psychophysics. This combination produces a logarithmic nature function expressing the relation between skin conductance and stimulus intensity. However, the impact of other biological information, especially EEG and ECG has not been investigated yet. It opens the new direction of applying EEG and ECG for our current research.

Bibliography

[1] Survey young adults watching tv. http://www.pewresearch.org/fact-tank/2017/09/13/about-6-in-10-young-adults-in-u-s-primarily-use-online-streaming-to-watch Accessed: 2018-01-15.

[2] Wolfram Boucsein. *Electrodermal activity.* Springer Science & Business Media, 2012.

[3] Olympics winter 2018. https://www.statista.com/statistics/809363/winter-olympics-streaming-viewers. Accessed: 2018-01-15.

[4] Cisco white paper. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html. Accessed: 2018-01-15.

[5] YouTube statistics. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html. Accessed: 2018-01-15.

[6] IBM - the state of streaming 2017 report. https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=KUW12563USEN. Accessed: 2018-01-15.

[7] Statistas online video - a billion-dollar opportunity. https://goo.gl/eVyCwQ. Accessed: 2018-06-15.

[8] Ali Begen, Tankut Akgul, and Mark Baugher. Watching video over the web: Part 1: Streaming protocols. *IEEE Internet Computing*, 15(2):54–63, 2011.

[9] Move Networks the introduction of adaptive video streaming. https://www.businesswire.com/news/home/20100915007073/en/Move-Networks-Awarded-Fundamental-Patent-Adaptive-Video. Accessed: 2018-01-15.

[10] Control Theory negative feedback control. https://en.wikipedia.org/wiki/Control_theory. Accessed: 2018-06-20.

[11] Michael Seufert, Sebastian Egger, Martin Slanina, Thomas Zinner, Tobias Hossfeld, and Phuoc Tran-Gia. A survey on quality of experience of http adaptive streaming. *IEEE Communications Surveys & Tutorials*, 17(1):469–492, 2015.

[12] Ricky KP Mok, Edmond WW Chan, and Rocky KC Chang. Measuring the quality of experience of http video streaming. In *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on*, pages 485–492. IEEE, 2011.

[13] Simon Scholler, Sebastian Bosse, Matthias Sebastian Treder, Benjamin Blankertz, Gabriel Curio, Klaus-Robert Muller, and Thomas Wiegand. Toward a direct measure of video quality perception using eeg. *IEEE transactions on Image Processing*, 21(5):2619–2629, 2012.

[14] Christos G Bampis and Alan C Bovik. Learning to predict streaming video qoe: Distortions, rebuffering and memory. *arXiv preprint arXiv:1703.00633*, 2017.

[15] Janne Seppänen and Martin Varela. Qoe-driven network management for real-time over-the-top multimedia services. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 1621–1626. IEEE, 2013.

[16] Charline Taibi Guguen, Francoise Le Bolzer, and Remi Houdaille. Improving user experience when http adaptive streaming clients compete for bandwidth. *SMPTE Motion Imaging Journal*, 126(1):28–34, 2017.

[17] Kjell Brunnström, Sergio Ariel Beker, Katrien De Moor, Ann Dooms, Sebastian Egger, Marie-Neige Garcia, Tobias Hossfeld, Satu Jumisko-Pyykkö, Christian Keimel, Mohamed-Chaker Larabi, et al. Qualinet white paper on definitions of quality of experience. 2013.

[18] Noriaki Kano. Attractive quality and must-be quality. *Hinshitsu (Quality, The Journal of Japanese Society for Quality Control)*, 14:39–48, 1984.

[19] David Hoyle. *ISO 9000 Quality Systems Handbook-updated for the ISO 9001: 2015 standard: Increasing the Quality of an Organizations Outputs.* Routledge, 2017.

[20] Lalit P Jain, Walter J Scheirer, and Terrance E Boult. Quality of experience. In *IEEE multimedia*. Citeseer, 2004.

[21] Satu Jumisko-Pyykkö. User-centered quality of experience and its evaluation methods for mobile television. *Tampere University of Technology*, page 12, 2011.

[22] Satu Jumisko-Pyykkö, Dominik Strohmeier, Timo Utriainen, and Kristina Kunze. Descriptive quality of experience for mobile 3d video. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pages 266–275. ACM, 2010.

[23] Raimund Schatz, Tobias Hoßfeld, Lucjan Janowski, and Sebastian Egger. From packets to people: Quality of experience as a new measurement challenge. In *Data traffic monitoring and analysis*, pages 219–263. Springer, 2013.

[24] Akira Takahashi, David Hands, and Vincent Barriac. Standardization activities in the itu for a qoe assessment of iptv. *IEEE Communications Magazine*, 46(2), 2008.

[25] Jiarun Song, Fuzheng Yang, Yicong Zhou, and Shan Gao. Parametric planning model for video quality evaluation of iptv services combining channel and video characteristics. *IEEE Trans. Multimedia*, 19(5):1015–1029, 2017.

[26] Quan Huynh-Thu and Mohammed Ghanbari. The accuracy of psnr in predicting video quality for different video scenes and frame rates. *Telecommunication Systems*, 49(1):35–48, 2012.

[27] Zhou Wang, Ligang Lu, and Alan C Bovik. Video quality assessment based on structural distortion measurement. *Signal processing: Image communication*, 19(2):121–132, 2004.

[28] Kazuhisa Yamagishi and Takanori Hayashi. Parametric packet-layer model for monitoring video quality of iptv services. In *Communications, 2008. ICC'08. IEEE International Conference on*, pages 110–114. IEEE, 2008.

[29] Fuzheng Yang and Shuai Wan. Bitstream-based quality assessment for networked video: a review. *IEEE Communications Magazine*, 50(11), 2012.

[30] Majd Ghareeb and Cesar Viho. Hybrid qoe assessment is well-suited for multiple description coding video streaming in overlay networks. In *Communication Networks and Services Research Conference (CNSR), 2010 Eighth Annual*, pages 327–333. IEEE, 2010.

[31] Samir Mohamed and Gerardo Rubino. A study of real-time packet video quality using random neural networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(12):1071–1083, 2002.

[32] Gerardo Rubino. Quantifying the quality of audio and video transmissions over the internet: the psqa approach. In *Communication Networks And Computer Systems: A Tribute to Professor Erol Gelenbe*, pages 235–250. World Scientific, 2006.

[33] Zhou Wang, Alan C Bovik, and Ligang Lu. Why is image quality assessment so difficult? In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 4, pages IV–3313. IEEE, 2002.

[34] Olivia Nemethova, Michal Ries, Eduard Siffel, and Markus Rupp. Quality assessment for h. 264 coded low-rate and low-resolution video sequences. In *Communications, Internet, and Information Technology*, pages 508–512, 2004.

[35] Stephen Wolf, Margaret H Pinson, Arthur A Webster, Gregory W Cermak, and E Paterson Tweedy. Objective and subjective measures of mpeg video quality. *Society of Motion Picture and Television Engineers*, pages 160–178, 1997.

[36] Muhammad Shahid, Andreas Rossholm, Benny Lövström, and Hans-Jürgen Zepernick. No-reference image and video quality assessment: a classification and review of recent approaches. *EURASIP Journal on Image and Video Processing*, 2014(1):40, 2014.

[37] RECOMMENDATION ITU-R BT. Methodology for the subjective assessment of the quality of television pictures. 2002.

[38] Ulrich Engelke, Daniel P Darcy, Grant H Mulliken, Sebastian Bosse, Maria G Martini, Sebastian Arndt, Jan-Niklas Antons, Kit Yan Chan, Naeem Ramzan, and Kjell Brunnström. Psychophysiology-based qoe assessment: a survey. *IEEE Journal of Selected Topics in Signal Processing*, 11(1):6–21, 2017.

[39] Gustav Fechner. Elements of psychophysics. vol. i. 1966.

[40] Laurence T Maloney and Joong Nam Yang. Maximum likelihood difference scaling. *Journal of Vision*, 3(8):5–5, 2003.

[41] Ernst Heinrich Weber. *EH Weber on the tactile senses*. Psychology Press, 1996.

[42] Mohammed Alreshoodi and John Woods. Survey on qoe\qos correlation models for multimedia services. *arXiv preprint arXiv:1306.0221*, 2013.

[43] Vlado Menkovski, Georgios Exarchakos, and Antonio Liotta. Machine learning approach for quality of experience aware networks. In *Intelligent Networking and Collaborative Systems (INCOS), 2010 2nd International Conference on*, pages 461–466. IEEE, 2010.

[44] Vlado Menkovski, Georgios Exarchakos, Antonio Liotta, and Antonio Cuadra Sánchez. Measuring quality of experience on a commercial mobile tv platform. In *Advances in Multimedia (MMEDIA), 2010 Second International Conferences on*, pages 33–38. IEEE, 2010.

[45] Emad Danish, Anil Fernando, Mohammed Alreshoodi, and John Woods. A hybrid prediction model for video quality by qos/qoe mapping in wireless streaming. In *Communication Workshop (ICCW), 2015 IEEE International Conference on*, pages 1723–1728. IEEE, 2015.

[46] Ozgur Oyman and Sarabjot Singh. Quality of experience for http adaptive streaming services. *IEEE Communications Magazine*, 50(4), 2012.

[47] Yusuf Sani, Andreas Mauthe, and Christopher Edwards. Adaptive bitrate selection: a survey. *IEEE Communications Surveys & Tutorials*, 19(4):2985–3014, 2017.

[48] Saamer Akhshabi, Sethumadhavan Narayanaswamy, Ali C Begen, and Constantine Dovrolis. An experimental evaluation of rate-adaptive video players over http. *Signal Processing: Image Communication*, 27(4):271–287, 2012.

[49] Janne Seppänen, Martín Varela, and Aggeliki Sgora. An autonomous qoe-driven network management framework. *Journal of Visual Communication and Image Representation*, 25(3):565–577, 2014.

[50] Florence Agboma and Antonio Liotta. Addressing user expectations in mobile content delivery. *Mobile Information Systems*, 3(3-4):153–164, 2007.

[51] Florence Agboma and Antonio Liotta. Quality of experience management in mobile content delivery systems. *Telecommunication Systems*, 49(1):85–98, 2012.

[52] Vlado Menkovski, Georgios Exarchakos, Antonio Liotta, and Antonio Cuadra Sánchez. Quality of experience models for multimedia streaming. *Advancing the Next-Generation of Mobile Computing: Emerging Technologies: Emerging Technologies*, 112, 2012.

[53] Kamal Deep Singh, Yassine Hadjadj-Aoul, and Gerardo Rubino. Quality of experience estimation for adaptive http/tcp video streaming using h. 264/avc. In *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, pages 127–131. IEEE, 2012.

[54] Pedro Casas, Michael Seufert, and Raimund Schatz. Youqmon: a system for on-line monitoring of youtube qoe in operational 3g networks. *ACM SIGMETRICS Performance Evaluation Review*, 41(2):44–46, 2013.

[55] ITUT Rec. P. 800: Methods for subjective determination of transmission quality. *International Telecommunication Union, Geneva*, 1996.

[56] Wanmin Wu, Ahsan Arefin, Raoul Rivas, Klara Nahrstedt, Renata Sheppard, and Zhenyu Yang. Quality of experience in distributed interactive multimedia environments: toward a theoretical framework. In *Proceedings of the 17th ACM international conference on Multimedia*, pages 481–490. ACM, 2009.

[57] Te-Yuan Huang, Ramesh Johari, and Nick McKeown. Downton abbey without the hiccups: Buffer-based rate adaptation for http video streaming. In *Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking*, pages 9–14. ACM, 2013.

[58] Dmitri Jarnikov and Tanır Özçelebi. Client intelligence for adaptive streaming solutions. *Signal Processing: Image Communication*, 26(7):378–389, 2011.

[59] Saamer Akhshabi, Lakshmi Anantakrishnan, Constantine Dovrolis, and Ali C Begen. 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*, pages 19–24. ACM, 2013.

[60] Sangwook Kim, Dooyeol Yun, and Kwangsue Chung. Video quality adaptation scheme for improving qoe in http adaptive streaming. In *Information Networking (ICOIN), 2016 International Conference on*, pages 201–205. IEEE, 2016.

[61] Te-Yuan Huang, Ramesh Johari, Nick McKeown, Matthew Trunnell, and Mark Watson. A buffer-based approach to rate adaptation: Evidence from a large video streaming service. *ACM SIGCOMM Computer Communication Review*, 44(4):187–198, 2015.

[62] Raimund Schatz, Tobias Hoßfeld, and Pedro Casas. Passive youtube qoe monitoring for isps. In *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on*, pages 358–364. IEEE, 2012.

[63] Hyun Jong Kim and Seong Gon Choi. A study on a qos/qoe correlation model for qoe evaluation on iptv service. In *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*, volume 2, pages 1377–1382. IEEE, 2010.

[64] Tobias Hoßfeld, Phuoc Tran-Gia, and Markus Fiedler. Quantification of quality of experience for edge-based applications. In *Managing Traffic Performance in Converged Networks*, pages 361–373. Springer, 2007.

[65] Tan Phan-Xuan and Eiji Kamioka. Efficiency of qoe-driven network management in adaptive streaming over http. In *Communications (APCC), 2016 22nd Asia-Pacific Conference on*, pages 517–522. IEEE, 2016.

[66] Tobias Hoβfeld, Raimund Schatz, and Sebastian Egger. Sos: The mos is not enough! In *Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on*, pages 131–136. IEEE, 2011.

[67] Robert C Streijl, Stefan Winkler, and David S Hands. Mean opinion score (mos) revisited: methods and applications, limitations and alternatives. *Multimedia Systems*, 22(2):213–227, 2016.

[68] P ITU-T RECOMMENDATION. 910. *Subjective video quality assessment methods for multimedia applications*, pages 910–200804, 2008.

[69] Truong Cong Thang, Hung T Le, Hoc X Nguyen, Anh T Pham, Jung Won Kang, and Yong Man Ro. Adaptive video streaming over http with dynamic resource estimation. *Journal of Communications and Networks*, 15(6):635–644, 2013.

[70] Truong Cong Thang, Quang-Dung Ho, Jung Won Kang, and Anh T Pham. Adaptive streaming of audiovisual content using mpeg dash. *IEEE Transactions on Consumer Electronics*, 58(1), 2012.

[71] Luciano Rubio Romero. A dynamic adaptive http streaming video service for google android, 2011.

[72] Saamer Akhshabi, Lakshmi Anantakrishnan, Ali C Begen, and Constantine Dovrolis. What happens when http adaptive streaming players compete for bandwidth? In *Proceedings of the 22nd international workshop on Network and Operating System Support for Digital Audio and Video*, pages 9–14. ACM, 2012.

[73] Kandaraj Piamrat, Adlen Ksentini, César Viho, and Jean-Marie Bonnin. Qoe-aware admission control for multimedia applications in ieee 802.11 wireless networks. In *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, pages 1–5. IEEE, 2008.

[74] Eirini Liotou, Dimitris Tsolkas, Nikos Passas, and Lazaros Merakos. Quality of experience management in mobile cellular networks: key issues and design challenges. *IEEE Communications Magazine*, 53(7):145–153, 2015.

[75] Tobias Hoßfeld, Sebastian Biedermann, Raimund Schatz, Alexander Platzer, Sebastian Egger, and Markus Fiedler. The memory effect and its implications on web qoe modeling. In *Teletraffic Congress (ITC), 2011 23rd International*, pages 103–110. IEEE, 2011.

[76] Tatsuya Yamazaki, Takumi Miyoshi, Masato Eguchi, and Kyoko Yamori. Quality of experience modeling with psychological effect for interactive web services. In *Network Operations and Management Symposium (NOMS), 2014 IEEE*, pages 1–4. IEEE, 2014.

[77] Vasilios A Siris, Konstantinos Balampekos, and Mahesh K Marina. Mobile quality of experience: Recent advances and challenges. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*, pages 425–430. IEEE, 2014.

[78] Huong Pham-Thi and Takumi Miyoshi. User-centric approach for bandwidth allocation method based on quality of experience. *IEICE Transactions on Communications*, 99(6):1282–1290, 2016.

[79] Stefan Winkler and Praveen Mohandas. The evolution of video quality measurement: From psnr to hybrid metrics. *IEEE Transactions on Broadcasting*, 54(3):660–668, 2008.

[80] Peter Reichl, Sebastian Egger, Raimund Schatz, and Alessandro D'Alconzo. The logarithmic nature of qoe and the role of the weber-fechner law in qoe assessment. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5. IEEE, 2010.

[81] Naeem Ramzan, Sebastian Palke, Thomas Cuntz, Ryan Gibson, and Abbes Amira. Emotion recognition by physiological signals. *Electronic Imaging*, 2016(16):1–6, 2016.

[82] Chi Jung Kim, Sangin Park, Myeung Ju Won, Mincheol Whang, and Eui Chul Lee. Autonomic nervous system responses can reveal visual fatigue induced by 3d displays. *Sensors*, 13(10):13054–13062, 2013.

[83] Mojtaba Khomami Abadi, Jacopo Staiano, Alessandro Cappelletti, Massimo Zancanaro, and Nicu Sebe. Multimodal engagement classification for affective cinema. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 411–416. IEEE, 2013.

[84] Grace Iarocci and John McDonald. Sensory integration and the perceptual experience of persons with autism. *Journal of autism and developmental disorders*, 36(1):77–90, 2006.

[85] Don C Fowles. The eccrine system and electrodermal activity. *Psychophysiology: Systems, processes, and applications*, 1:51–96, 1986.

[86] Daniel N McIntosh, Lucy Jane Miller, Vivian Shyu, and Randi J Hagerman. Sensory-modulation disruption, electrodermal responses, and functional behaviors. *Developmental medicine and child neurology*, 41(9):608–615, 1999.

[87] Mathias Benedek and Christian Kaernbach. Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4):647–658, 2010.

[88] Heart Rate the introduction. `https://en.wikipedia.org/wiki/Heart_rate`. Accessed: 2018-06-17.

[89] Solange Akselrod, David Gordon, F Andrew Ubel, Daniel C Shannon, AC Berger, and Richard J Cohen. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *science*, 213(4504):220–222, 1981.

[90] Bruce Pomeranz, RJ Macaulay, Margaret A Caudill, Ilan Kutz, Dan Adam, David Gordon, Kenneth M Kilborn, A Clifford Barger, Daniel C Shannon, and Richard J Cohen. Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology*, 248(1):H151–H153, 1985.

[91] U Rajendra Acharya, K Paul Joseph, Natarajan Kannathal, Choo Min Lim, and Jasjit S Suri. Heart rate variability: a review. *Medical and biological engineering and computing*, 44(12):1031–1051, 2006.

[92] Avijit Chaudhuri. *Fundamentals of sensory perception*. Oxford University Press, 2011.

[93] Steven Yantis. *Sensation and perception*. Palgrave Macmillan, 2013.

[94] Andrew M Colman. *A dictionary of psychology*. Oxford University Press, USA, 2015.

[95] George A Gescheider. *Psychophysics: the fundamentals*. Psychology Press, 2013.

[96] Tan Phan-Xuan and Eiji Kamioka. Accurate available bandwidth allocation in http adaptive streaming. *International Journal of Computer Networks and Communications*, 9(5):83–94, 2017.

[97] G Cermak, M Pinson, and Stephen Wolf. The relationship among video quality, screen resolution, and bit rate. *IEEE Transactions on Broadcasting*, 57(2):258–262, 2011.

[98] GSR sensor available:. `http://wiki.seeedstudio.com/Grove-GSR_Sensor/`. Accessed: 2018-01-15.

[99] IIS Smooth Streaming available:. `https://www.iis.net/downloads/microsoft/smooth-streaming/`. Accessed: 2018-01-15.

[100] C Karenbach. Ledalab-a software package for the analysis of phasic electrodermal activity. *Internal report, Allgemeine Psychologie, Institut für Psychologie, Institut ftir Psychologie, Tech. Rep.*, 2005.

[101] Mathias Benedek. Analysis of eda using ledalab. 2016.

[102] Jason J Braithwaite, Derrick G Watson, Robert Jones, and Mickey Rowe. A guide for analysing electrodermal activity (eda) & skin conductance responses (scrs) for psychological experiments. *Psychophysiology*, 49(1):1017–1034, 2013.

[103] Jason J Braithwaite and Derrick G Watson. Issues surrounding the normalization and standardisation of skin conductance responses (scrs)., 2015.

[104] Douglas F Levinson and Robert Edelberg. Scoring criteria for response latency and habituation in electrodermal research: A critique. *Psychophysiology*, 22(4):417–426, 1985.

[105] Kandaraj Piamrat, Cesar Viho, Jean-Marie Bonnin, and Adlen Ksentini. Quality of experience measurements for video streaming over wireless networks. In *Information Technology: New Generations, 2009. ITNG'09. Sixth International Conference on*, pages 1184–1189. IEEE, 2009.

[106] Rémi Houdaille and Stéphane Gouache. Shaping http adaptive streams for a better user experience. In *Proceedings of the 3rd Multimedia Systems Conference*, pages 1–9. ACM, 2012.

[107] Donald F Morrison. *Multivariate analysis of variance*. Wiley Online Library, 2005.

[108] Richard M Heiberger and Erich Neuwirth. One-way anova. In *R through excel*, pages 165–191. Springer, 2009.

[109] HRV analysis system clinic information. `http://www.medicore.com/download/HRV_clinical_manual_ver3.0.pdf`. Accessed: 2018-06-19.

[110] Chiheb Ben Ameur, Emmanuel Mory, and Bernard Cousin. Shaping http adaptive streams using receive window tuning method in home gateway. In *Performance Computing and Communications Conference (IPCCC), 2014 IEEE International*, pages 1–2. IEEE, 2014.

[111] Junchen Jiang, Vyas Sekar, and Hui Zhang. Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive. *IEEE/ACM Transactions on Networking (TON)*, 22(1):326–340, 2014.

[112] Kevin J Ma. *Provider-Controlled Bandwidth Management for HTTP-based Video Delivery*. PhD thesis, University of New Hampshire (Department of Computer Science), 2012.

[113] Steven Latré, Pieter Simoens, Bart De Vleeschauwer, Wim Van de Meerssche, Filip De Turck, Bart Dhoedt, Piet Demeester, Steven Van den Berghe, and Edith Gilon de Lumley. An autonomic architecture for optimizing qoe in multimedia access networks. *Computer Networks*, 53(10):1587–1602, 2009.

[114] Stefan Lederer, Christopher Müller, and Christian Timmerer. Dynamic adaptive streaming over http dataset. In *Proceedings of the 3rd Multimedia Systems Conference*, pages 89–94. ACM, 2012.

[115] Panagiotis Georgopoulos, Yehia Elkhatib, Matthew Broadbent, Mu Mu, and Nicholas Race. Towards network-wide qoe fairness using openflow-assisted adaptive video streaming. In *Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking*, pages 15–20. ACM, 2013.

Publication List

Journal Articles

[P.1] Tan Phan-Xuan and Eiji Kamioka, "Early Detection of QoE Deterioration in HTTP Adaptive Streaming," Journal of Computer and Communications (JCC), Vol.5, pp.14-27, No.14, 2017.

[P.2] Tan Phan-Xuan and Eiji Kamioka, "ACCURATE AVAILABLE BANDWIDTH ALLOCATION IN HTTP ADAPTIVE STREAMING," International Journal of Computer Networks and Communications (IJCNC), Vol.9, No.5, pp.83-94, September, 2017.

[P.3] Tan Phan-Xuan and Eiji Kamioka, "Collaborative Approach using Psychophysiology and Psychophysics for Optimal Threshold Determination in HAS Service QoE Management," Journal of Computer and Communications (JCC), Vol.6, pp.57-81, No.8, 2018.

Conference Papers

[P.1] Tan Phan-Xuan and Eiji Kamioka, "A MACHINE LEARNING BASED NETWORK MONITORING AND MANAGEMENT SYSTEM", Proceedings of the 10th South East Asia Technical University Consortium Symposium (SEATUC2016), Shibaura Institute of Technology, Tokyo, Japan, February 23, 2016.

[P.2] Tan Phan-Xuan and Eiji Kamioka, "Efficiency of QoE-Driven Network Mnagement in Adaptive Streaming Over HTTP", Proceedings of the 22th Asia-Pacific Conference on Communications (APCC2016), Yogyakarta, Indonesia, pp.517-522, August 26.

[P.3] Tan Phan-Xuan, Eiji Kamioka, "QoE Based Network Management: Tackling QoE Reduction in Adaptive Streaming over HTTP", Proceedings of IEICE Society Conference, BS-5-14, September 21, 2016, pp.S78-S79.

[P.4] Tan Phan-Xuan and Eiji Kamioka, "Network Management Based QoE Estimation for Adaptive Streaming over HTTP", Proceedings of the 17th International Network Strategy and Planning Symposium (Networks2016), Montreal, Canada, pp.261-266, Septembrt 27, 2016.

[P.5] Phan Xuan Tan and Eiji Kamioka, "CONTROLLING NETWORK BASED QOE IMPROVEMENT IN HTTP ADAPTIVE STREAMING," Proceedings of the 11th South East Asia Technical University Consortium Symposium (SEATUC2017), Ho Chi Minh City University of Technology, Ho Chi Minh, Vietnam, pp.1-5, March 14, 2017.

[P.6] Tan Phan-Xuan, Eiji Kamioka, "Two-phase QoE Monitoring in HTTP Adaptive Streaming", Proceedings of IEICE Society Conference, BS-7-13, September 12, 2018