Model for Mobile Online Video viewed on Samsung Galaxy Note 5

Debajyoti Pal and Vajirasak Vanijja
IP Communications Laboratory, School of Information Technology,
King Mongkut's University of Technology Thonburi
Bangkok, 10140 - Thailand
[e-mail : debajyoti.pal@mail.kmutt.ac.th]

Abstract

The primary aim of this paper is to propose a non-linear regression based technique for mapping different network Quality of Service (QoS) factors to an integrated end-user Quality of Experience (QoE) or Mean Opinion Score (MOS) value for an online video streaming service on a mobile phone. We use six network QoS factors for finding out the user QoE. The contribution of this paper is threefold. First, we investigate the impact of the network QoS factors on the perceived video quality. Next, we perform an individual mapping of the significant network QoS parameters obtained in stage 1 to the user QoE based upon a non-linear regression method. The optimal QoS to QoE mapping function is chosen based upon a decision variable. In the final stage, we evaluate the integrated QoE of the system by taking the combined effect of all the QoS factors considered. Extensive subjective tests comprising of over 50 people across a wide variety of video contents encoded with H.265/HEVC and VP9 codec have been conducted in order to gather the actual MOS data for the purpose of QoS to QoE mapping. Our proposed hybrid model has been validated against unseen data and reveals good prediction accuracy.

Keywords: QoS, QoE, MOS, H.265, VP9, Subjective Test, Regression Analysis

1. Introduction

Recently as per a report published in [1], video traffic constitutes more than 55% of the overall Internet traffic and it is predicted to grow at a very fast pace. Online video streaming is gaining popularity on devices having small form factor screens ranging from 4 inches to 6 inches [2]. This rapid growth in the demand for network bandwidth has resulted in the emergence of newer codecs like H.265/HEVC and VP9 that provides a good quality to compression ratio. For streaming services like video there should be a continuous assessment of the end users perceived video quality as the network over which data is being transmitted is inherently unreliable. This leads us to the concept of Quality of Experience (QoE) which is a result of multidimensional factors being influenced by the system, user or some other contextual factors [3]. The International Telegraph Union Telecommunication Standardization Sector (ITU-T) defines QoE as the overall acceptability of an application or service, as perceived subjectively by the end user [4]. It includes the complete end-to-end system effects that can be due to the client, terminals, network, service infrastructure or some other factors. From an Internet Service Provider's (ISP's) point of view however, it is more important to understand the fundamental relationship between the various network and application level Quality of Service (QoS) factors and the underlying QoE. This is an extremely important requirement for any effective QoE management scenario. Hence, the QoS parameters are the most important business relevant parameters for the ISP's [5]. Therefore; there is a sharp demarcation between QoS and QoE- the former being easier to measure and hence more technical, while the latter being a more psychological state of mind of the end users and hence relatively complex to evaluate mathematically.

QoE is generally evaluated using three different approaches such as subjective tests, objective tests and pseudo-subjective or hybrid tests [6]. Subjective tests are considered to be the most accurate amongst the three mentioned, wherein the people need to watch video samples and then rate those. Results of the tests are generally expressed in the form of Mean Opinion Score (MOS) or some variation of it depending upon the method used [7-10]. Such tests though accurate are time consuming and costly to conduct. Objective tests on the other hand are extremely effective for in-service quality monitoring but require complex algorithms or specialized equipment for evaluation [11]. They are extremely fast and easy to deploy, but the accuracy as compared to the subjective tests is questionable and depends upon the assessing algorithm. Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM) and Video Quality Metric (VQM) are some of the well-known objective algorithms [12-14]. These methods are full reference (FR) methods that require the presence of the original signal and hence require an additional bandwidth overhead. As such, they are not suitable for streaming applications for which the hybrid tests are preferred. These methods use a combination of the subjective and objective methods to predict the video quality and do not require the source signal to be present. ITU-T G.1070 model and Pseudo Subjective Quality Assessment (PSQA) are some of the well-known examples [15].

In this paper, we are proposing a new regression based hybrid and integrated QoS/QoE mapping model in order to evaluate the quality of streaming videos, impaired solely due to the network QoS parameters. The model is a hybrid one as it combines the subjective results with relevant mathematical techniques to predict the video quality. We begin by conducting a

subjective test where the videos are impaired by six different network QoS parameters x_1 to x_6 (three Key Performance Indicators or primary network factors + three secondary network factors). Packet loss, jitter and throughput are the 3 Key Performance Indicators (KPI's), whereas variable initial delay, buffering delay and auto scale resolution are the 3 secondary factors taken into consideration. We use only H.265/HEVC and VP9 as the video codecs as they are representative of the modern day streaming scenario. After gathering the subjective MOS from the experiment, we investigate into the effectiveness of the network parameters that we have chosen for the purpose of QoS to QoE mapping. This is done by carrying out an Analysis of Variance (ANOVA) measure over the 6 QoS parameters considered and Subjective MOS as the target output i.e. dependent variable. Next, a non-linear regression technique is used to map each individual QoS parameter (x_1 to x_6) its QoE counterparts $Q_1(x_1)$ to $Q_6(x_6)$. At this point, we also consider other popular mapping functions used by different authors and select the optimal one based upon a decision variable DV that has been explained in a later section. The individual mapping functions that we obtain are quite different from the ones found in papers [16-18].

After deriving the individual quality models $Q_1(x_1)$ to $Q_6(x_6)$ empirically, we try to find out the generic structure of the integrated model $Q(x_1, x_2,..., x_6)$. This is a challenging task, where the QoE is affected by several QoS parameters simultaneously. The most common approach adopted by many researchers in papers [17-19], is based upon an additive/weighted-sum approach that is not only less accurate, but also suffers from some serious drawbacks that we have proved in this research. The weighted sum approach that we have undertaken is based upon the Analytical Hierarchical Process (AHP) algorithm. In this paper, we propose a generic QoE equation in the multiplicative format keeping in mind the drawbacks of the additive format. The final integrated multivariate QoE estimation function is obtained by carrying out a liner regression over the additive and multiplicative QoE's found already. Finally, we conduct another subjective test with videos being impaired by multiple factors and evaluate our model performance.

Rest of the paper is organized as follows. In Section 2, other related works are reviewed. Section 3 presents the various QoS factors that we select and the subjective tests that are carried out in detail. Section 4 provides the analytical approach for individual QoS to QoE mapping functions along with the decision variable. In Section 5, we evaluate the integrated QoE for the various QoS factors and estimate the accuracy of the model. Finally, Section 6 provides the conclusion and scope of future work.

2. Related Work

In this section we present a concise information about the work that has been done till date to estimate the video quality, correlation between QoS and QoE and evaluating QoE from QoS parameters.

2.1 Video Quality Estimation

Subjective measurement techniques are the most accurate video quality estimation tools available till date. Standard procedure to conduct subjective tests has been laid down by ITU-T in its various Recommendations. Different techniques can be used depending upon the application requirements. Absolute Category Rating (ACR), Absolute Category Rating with Hidden Reference (ACR-HR), Degradation Category Rating (DCR) and Pair Comparison (PC)

are some of the most commonly used experimental techniques. Both ACR and ACR-HR are examples of single stimulus method where only one video sequence is shown at a time, while DCR and PC are examples of double stimulus method where the original and degraded video sequences are presented in pairs to the user. The Recommendations mandate the video sequences to be under 10s duration; a criterion which cannot represent a real life video streaming scenario [20]. A detailed comparison between the different subjective techniques can be found in papers [21, 22]. As subjective methods are very expensive and time consuming to conduct, hence we need to go for other objective techniques.

Objective models are based upon certain algorithms or mathematical equations that try to predict the subjective video quality. They can be classified based upon the amount of reference information presented to the assessing software/model into three main types viz. Full Reference (FR), Reduced reference (RR) and No Reference (NR) methods [23]. In FR technique, both the original video sequence and the impaired one must be presented to the assessing software/model. PSNR, SSIM and VQM are examples of such technique. RR schemes extract certain features of interest from the original video and use that for comparison against the impaired ones. Authors in paper [24] discuss about one such approach. In case of NR technique, no reference video is needed for quality estimation. ITU-T G.1070 and the work done by authors in [25] and [26] represent such a method.

The metrics that have been taken for video quality prediction can be of different types. Generally, they are classified into three main types as the network QoS, application QoS and the user QoS factors [27]. Network QoS refers to the network performance that includes the effects of packet loss, jitter and bandwidth. Application QoS is mainly related with the codec related factors like type of codec used, video resolution, type of video content and the Group of Pictures (GOP) structure. User QoS is equivalent to the MOS.

2.2. Correlation between QoS and QoE

Due to the high cost of the subjective tests, several authors have attempted to find the QoE from various QoS factors. Mathematically QoE can be defined by a mapping function being affected by a number of influence parameters on the visual quality being perceived by the end user. If x_i denotes a general influence parameter; then the QoE can be represented as QoE = $f(x_1, x_2, x_i\dots x_n)$ for n number of factors.

Authors in paper [28] investigated the relationship between voice transmission conditions and the human perception quality and they concluded that the relationship is not linear. However, it needs to be seen if the same relationship holds true for video traffic. Authors propose similar models in [29] and [30]. However, they consider the effect of only one QoS factor on the user QoE that is not representative of a practical video streaming scenario. With rapid growth of the Internet, apart from packet loss several other factors need to be considered while evaluating the perceived quality. Another study in [31] finds the relationship between bit-rate, packet-loss rate and jitter on QoE and proposes a logarithmic relationship. This is the IQX hypothesis, which is an extension of the Weber-Fechner law, but gives the QoE for a single impairment factor only at any point of time. Although, the IQX hypothesis is good at estimating the QoE of any network parameter, the authors tested it with voice traffic only that differs substantially from video traffic. Thus, the generic exponential relationship between the network QoS and user QoE might need some fine-tuning which we explore in this paper. In addition, since the IQX hypothesis considers a single impairment factor only, we propose a

regression based approach for QoE determination from multiple impairment factors. Authors in paper [32] have proposed an adaptive QoE measurement scheme for IPTV services. However, they measure the QoE based upon the quality of images contained in the video stream and do not consider any network related impairment factor. They just give a proof of concept without analyzing the performance of their proposed method. Similarly, authors in paper [33] try to find a general approach towards the identification of the QoE influence factors where they categorize the QoE into multidimensional IF spaces. The different influence factors are given different weights that depend upon the application context while evaluating the QoE. Their proposed ARCU model does not address the issue of modeling QoE in the case of multimedia services composed of multiple media components.

A Dynamic Adaptive Streaming over HTTP (DASH) based multi-view video streaming system that can minimize the view-switching delay by employing proper buffer control, parallel streaming and server push schemes has been presented by authors in [34]. Similar HTTP based video streaming for long-term evolution (LTE) cellular networks has been proposed in [35]. Authors in papers [36] and [37] try to predict the video QoE for a DASH based video streaming scenario. Similarly, in paper [38] the authors present a general framework for robust video adaptation by several commercial players by proposing a new algorithm. Papers [39], [40] and [41] provide an excellent survey on the QoE estimation techniques in place for a video streaming scenario in general.

The main challenge is to understand the fundamental relationship between the various QoS factors and QoE first, and then go for the multidimensional relationship between the two. Authors in paper [42] propose a multidimensional QoE model for a mobile web-browsing experience in which they use an additive approach. However, their basic assumption in following an additive approach is a linear relationship between the QoS and QoE factors, which is not true in our case, as experiments will reveal in later part of the paper. Thus, following a weighted sum approach is questionable in our case. In paper [43], authors try to evaluate the QoS/QoE of an IPTV service based upon network impairments only. Instead of taking the MOS scores, they consider the Structural Similarity (SSIM) scores, for the purpose of QoS to QoE mapping. After evaluating the final QoE, they do not provide any sort of model validation by carrying out a subjective test. A weighted sum approach is again followed which has certain drawbacks as will be explained in later sections. Authors in paper [44] determine the audiovisual quality of a service where the combined audio and video quality predicted by the model is given by:

$$MOS_{AV} = \alpha MOS_A + \beta MOS_V + \gamma MOS_A \times MOS_V + \zeta, \text{ where}$$

MOS_{AV} refers to the overall audiovisual quality, MOS_A and MOS_B refers to the audio and video quality respectively. The main drawback of this study is the lack of any subjective tests to validate the results. Also, the primary emphasis of the paper was to provide an effective audio video synchronization rather than modeling the QoE in terms of the QoS factors. Extensive literature review that has been done shows that although QoE estimation is not a new technique, yet there are a number of loopholes in the process. Most of the papers have tried to provide an individual QoS to QoE mapping by following an IQX hypothesis for the network parameters. Only, the effect of packet loss, jitter, throughput and delay has been taken into account while estimating the QoE. However, with the rapid growth in video streaming services especially over mobile devices that has limited hardware capability in terms of

storage buffer, the effects of buffering delay need to be considered also. Recently introduced adaptive playout strategies by services like YouTube, wherein depending upon the network condition resolution of the streamed videos can change also need to be accounted for. For this reason, we include three new parameters for this research and justify their inclusion also by carrying out an ANOVA analysis. Variable Initial Delay, Buffering Delay and Auto Scale Resolution are the three new parameters that we introduce in this work. A weighted sum approach is generally followed for calculating the QoE from multiple QoS parameters. However, such a technique has its own drawback that we will explain later. We follow an alternative approach for QoE estimation and validate our model with unseen results obtained from the subjective test.

3. QoS Metric Selection and Subjective Evaluation

In this section, we discuss about the subjective experiment carried out, video sequences used and the network QoS parameters that have been selected for this purpose.

3.1 Subjective Evaluation

We carry out two subjective tests as per the ITU-T Recommendation P.910. Fifty-nine participants were involved in each test and mixed in gender. The age range varied from 18-70 years. **Fig. 1** shows the breakdown of the participant ages.
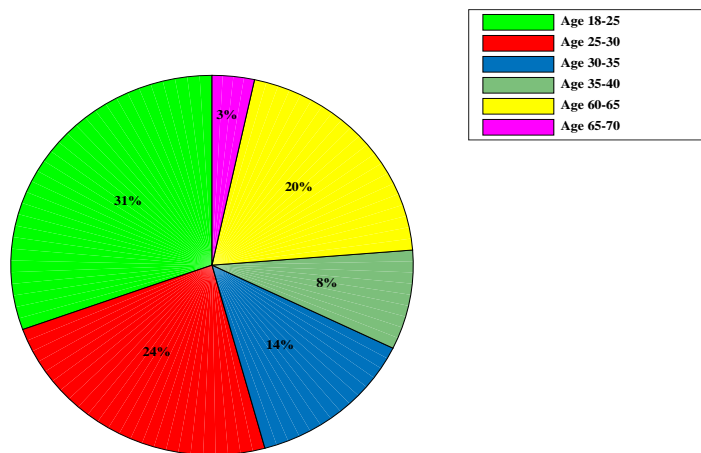

Fig. 1. Breakdown of the participant ages.

In the first test, the subjects are shown videos that are impaired with only one network QoS factor. Since, we have considered six different network QoS factors; therefore each of the videos have been impaired by exactly one factor for the first test. The experimental details have been provided in Section 3.3 later. For the second test; the same set of master video sequences are impaired by more than one network QoS factor. The details are given later in Section 5.3. We use eleven different video sequences for the purpose of this research. While eight of the video sequences are used for building our prediction model, the remaining three are used for the purpose of model validation only. Video sequences 1, 2, 4, 5, 6, 7, 9 and 11 are

used for model building, while the remaining ones are used for validation. This has been shown in **Table 1.** All the volunteers have been screened for any visual defects specifically color blindness or myopia before conducting the test. None was discarded in this process. Before conducting the experiment approval has been obtained from the ethical committee of the university as it involves human subjects. The entire test is conducted in a controlled laboratory environment over a period of 8 weeks. Training sessions have been conducted using demo videos in order to familiarize the participants to the actual test conditions. We used the Samsung Galaxy Note 5 for showing the videos as it supports resolutions of up to 2K and has inherent support for the modern codecs H.265 and VP9. Each video sequence is of 10s duration and of Full HD resolution. During the assessment, the participants are left alone in order to minimize the unwanted effects of being supervised [45]. Details about the video sequences used have been provided in the next sub-section. The subjective opinion is gathered using the single stimulus ACR technique in which the participants were asked to rate each video presented to them on a scale of 1 to 5; where 1 represents the worst and 5 the best quality. The subjects are provided with scoring sheets to record their opinions. As subjective assessment is a tedious and high concentration task, we divided the viewing into multiple sessions with each session lasting for 15 minutes duration. After the test, all the offline scores are manually entered into a computer for the purpose of data analysis. The subjective test is done by strictly adhering to the standards mentioned in ITU-T Recommendation P.910.

3.2 Video Selection

We use the publicly available SVT High Definition Multi Format Test Set database maintained by the Video Quality Experts Group (VQEG) for selecting the reference videos [46]. Eleven reference videos were chosen; each having different levels of spatial and temporal information i.e. SI and TI values. This has been shown in **Fig. 2** for easy reference. A snapshot of all the video sequences that have been considered are given in **Fig. 3.**

The spatial perceptual information (SI) is based upon the Sobel filter; whereas the temporal perceptual information (TI) is based upon the motion difference feature. These values have been calculated for each video sequence based upon the outline given in ITU-T Recommendation P.910. Perceived video quality depends heavily upon the type of video content, which has been established by researchers in [47] and [48]. Since, SI and TI values are a direct indication of the extent of video content complexity; hence, we chose those videos where these values vary over a wide range in order to cover the entire gamut possible. This fact is evident from **Fig. 2** shown in next page.

Further details about the eleven clips have been provided in **Table 1.** All the clips are of 10 seconds duration and in the native YUV 4.2.0 format. Resolutions of all the videos are fixed at Full HD (1920×1080). We did not take resolutions higher than Full HD since video streaming is normally not done at higher resolutions and due to the lack of reference video contents. Similarly, lower resolutions were also not chosen, as they are not representative of present day video streaming scenario.
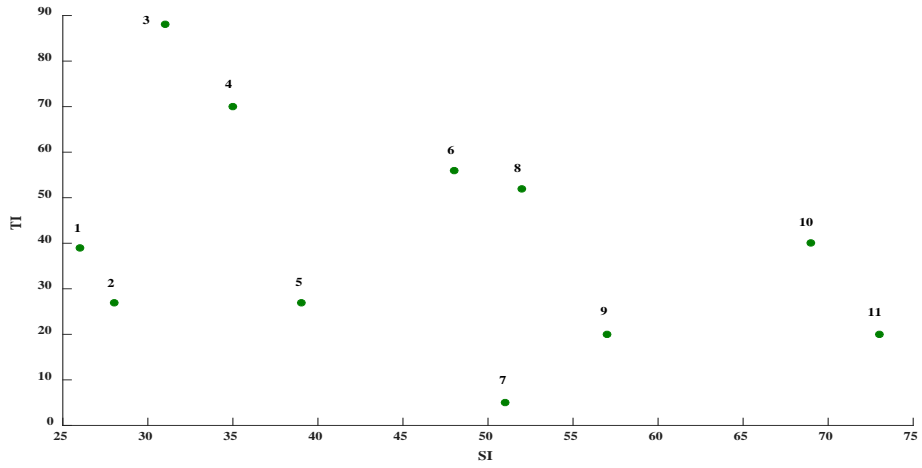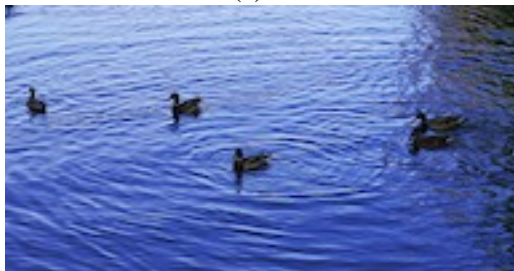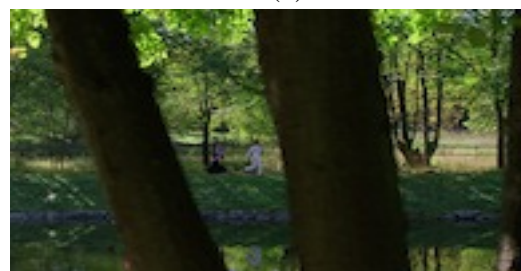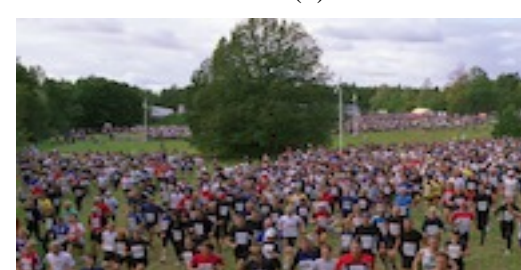
Fig. 2. SI and TI values for selected video clips

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

Fig. 3. Snapshot of used video sequences (a) Harbor (b) Ice (c) DucksTakeOff (d) ParkJoy (e) Crew (f) CrowdRun (g) Akiyo (h) Soccer (i) Foreman (j) Football (k) News

Table 1. Video Clip Details

Seq. Number	Seq. Name	Frame Rate	Resolution	Chroma Format	Content Complexity
1	Harbor	60 fps	1920 × 1080	4.2.0	1014
2	Ice	60 fps	1920 × 1080	4.2.0	756
3	DucksTakeOff	50 fps	1920 × 1080	4.2.0	2728
4	ParkJoy	50 fps	1920 × 1080	4.2.0	2450
5	Crew	60 fps	1920 × 1080	4.2.0	1053
6	CrowdRun	50 fps	1920 × 1080	4.2.0	2688
7	Akiyo	30 fps	1920 × 1080	4.2.0	255
8	Soccer	60 fps	1920 × 1080	4.2.0	2704
9	Foreman	30 fps	1920 × 1080	4.2.0	1140
10	Football	30 fps	1920 × 1080	4.2.0	2760
11	News	30 fps	1920 × 1080	4.2.0	1470

We encode the reference videos with both H.265 and VP9 codecs. The same encoder configuration has been used for both the reference and impaired videos to get the best quality.

In this work, we are interested in finding out the effect of the network QoS factors only on the overall video QoE, due to which we keep the distortions arising from the codec the same for both the original and impaired sequences.

Next we present the network QoS factors that have been used to create the impaired videos and evaluate the QoE.

3.3 Network QoS Factors

In this work, we select six different network QoS metrics. As most of the authors have taken into account the effect of packet-loss, jitter and throughput to be the main factors for QoS to QoE mapping as discussed earlier, we consider these to be the primary factors or Key Performance Indicators (KPI's). In addition, we introduce three more factors variable initial delay, buffering delay and auto scale resolution in our experiment to calculate the QoE. This has been shown in **Fig. 4.**
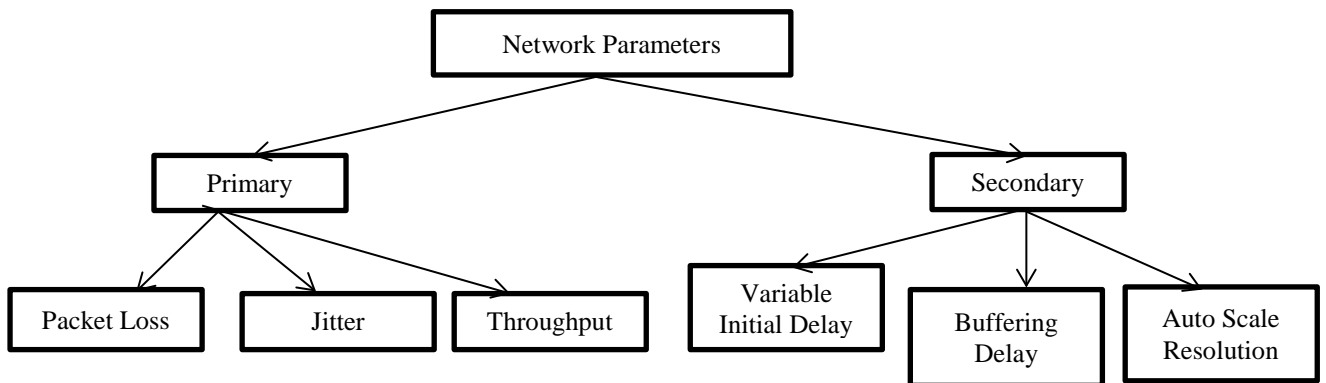

Fig. 4. Network QoS factors

Next, we present a description of the factors included.

- **Packet loss (P)**: IP packets may be discarded during their transit over the network or dropped at any intermediate nodes due to network congestion or buffer overflow. Here, we consider a random packet loss pattern as it has a significant detrimental effect on the video stream quality as compared to other types of packet losses [49].
- **Jitter (J):** It is defined as the variable delay in receiving packets at the receiver. It can occur due to network congestion, improper queuing or several other factors.
- **Throughput (T):** It refers to the amount of data that is successfully transferred from one place to another in a given time period. Its influence towards the video QoE has been well accepted by the research community.
- **Variable Initial Delay (VID):** In order to overcome the effect of jitter, every client is equipped with a play- out buffer. A streaming video will start to play only after the buffer has been filled up to a certain threshold value [50] [51]. VID is defined as the time-gap between the arrival of the first frame in the buffer and its subsequent play-out on the client screen at the start of video playback. i.e. $VID = T_{Playout} - T_{Arrival}$ (1st Frame)
- **Buffering Delay (BD):** During video playback, there can be a situation where the play-out buffer at the client is temporarily empty. In such cases, the video playback

would be intermittent and the quality will be degraded seriously. The video playback would stall. The total playback time is increased by each buffering delay interval. The viewers will encounter frequent buffering delays when the player at the client side requests a higher bitrate than what is actually available in the network.

- **Auto Scale Resolution (ASR):** In an adaptive video streaming scenario the videos are encoded at multiple discreet bitrates i.e. at different resolutions. For example, the most commonly used video resolutions by YouTube are at 144p, 240p, 360p, 480p, 720p or 1080p Depending upon the available network bandwidth and other factors a particular bitrate stream is broken into multiple segments or chunks, with each segment lasting between 2 to 10 seconds. For the sake of this research the resolution combinations that we choose are (360p+480p), (720p+360p), (720p+480p), (360p+1080p) and (1080p+720p) combination. The duration of the video, sequences that we use in our experiment are 10 seconds each. Considering the fact that the duration of each fragmented segment should be between 2 to 10 seconds in case of a resolution switch, we take into account only those cases where the number of resolution switches is two. Higher number of resolution switches has not been considered keeping in mind the total length of the original video sequences. We express the ASR factor as the ratio of a particular resolution combination to the minimum resolution combination of the videos that we use. For, example the ASR factor for (720p+360p) combination is $(1280 \times 720 + 640 \times 360) / (640 \times 360 + 854 \times 480) = 1.8$. Similarly, for (360p+480p), (720p+480p), (360p+1080p) and (1080p+720p) the ASR factor is 1, 2.1, 3.6 and 4.7 respectively. This has been shown in **Table 2.**

As it is evident from the description of the network QoS factors considered above, we find that the secondary factors are dependent upon the three primary factors. For example, variable initial delay is a direct effect of jitter. Due to the effect of jitter, the different video packets can reach at the destination side at different points of time. In order to improve the viewing experience, every client is equipped with a buffer to ensure a smooth playback. Therefore, unless the buffer is filled up to a certain extent video playback will not begin. Having intermittent delays during video startup can reduce the viewing experience, which arises due to the effect of jitter. Similarly, buffering delays are caused when the player at the client side requests a higher bitrate than what is actually available in the network. The video players are normally equipped with a throughput estimator with the help of which they can monitor the available network throughput during a download. Thus, if the network throughput is low and the playout buffer at the client end not entirely full, then there can be a situation where momentarily the playout buffer becomes empty, thereby resulting in an intermittent video playback. Similarly, the player can also experience a low frame rate when a significant number of packets are lost during the video rendering process. Hence, both packet loss and network throughput has a direct impact on buffering delay, which in turn degrades the QoE of the viewers. Auto scale resolution is a type of adaptive bitrate streaming technique that is used by the video content providers with an aim to improve the viewing QoE. The video content provider stores the same video contents in multiple resolutions and then depending on various network factors like the available network bandwidth, extent of jitter present and the overall network loading conditions selects a particular resolution for showing to the users. Automatic switching to lower or higher resolutions than what is current being played happens depending upon the network conditions and other factors like amount of playout buffer left, video rendering capability of the viewer's device, etc. Hence, all the secondary factors that we have considered are a consequence of the primary ones.

Table 2 shows the parameter settings that are used to create the impaired videos for the first subjective test. For every set of parameter, the range of values has been carefully selected so that they match with those used by authors in other papers or beyond which the perceived video quality saturates (either excellent or poor).

Table 3 and **Table 4** show the encoder configuration details for H.265/HEVC and the VP9 codec respectively. All these video samples were prepared and preloaded into a Samsung Galaxy Note 5 having 4 GB of RAM, 64 GB of internal storage and running Android 6.0.1 before being shown to the subjects.

Table 2. Experimental Setup

Parameter	Details
Video Codec	H.265, VP9
Encoder Version	Ffmpeg version 3.1.3
Video Format	Full HD progressive (1080p)
Packet Loss in %	0.1, 0.5, 1, 3, 5, 10
Jitter in milliseconds	1, 2, 3, 4, 5
Throughput in kbps	500, 1000, 2000, 3000, 5000
Variable Initial Delay in seconds	5, 15, 25, 35, 45
Buffering Delay in seconds	2, 4, 6, 8, 10
Auto Scale Resolution	1, 1.8, 2.1, 3.6, 4.7

Table 3. H.265/HEVC Encoder Configuration

Parameter	Details
Encoder Version	HM 16.6
Profile	Main
Reference Frames	4
R/D Optimization	Enabled
GOP	8
Coding Unit Size/Depth	64/4
Fast Encoding	Enabled
Rate Control	Disabled
Internal Bit Depth	8

Table 4 VP9 Encoder Configuration

Parameter	Details
Encoder Version	Ffmpeg 3.1.3
Encoding Quality	Best
No Of Passes	2
Bit Rate Control Mode	Variable Bit Rate (VBR defined by target bit rate)
Constrained Quality (CQ) Level	Kept same as Quantization Parameter QP
Initial, optimal and maximum buffer level	4000 ms, 5000 ms, 6000 ms
GOP Size	Auto
GOP Length (Intra Period)	320
Internal Bit Depth	8

Since, the scope of this work is limited to a mobile video viewing experience only, we chose to use a Samsung Galaxy Note 5 as it supports inherent (hardware level) decoding of the

videos encoded with H.265 and VP9 codec. At the time when the experiment was performed, there was no other mobile device that supported hardware level decoding of both these codecs to the best of knowledge of the authors. To create the impaired videos we used the Network Emulator for Windows Toolkit (NEWT) provided by Microsoft to simulate the various QoS parameters. The overall methodology for creating the impaired videos and measuring the QoE/MOS has been shown in **Fig. 5.**
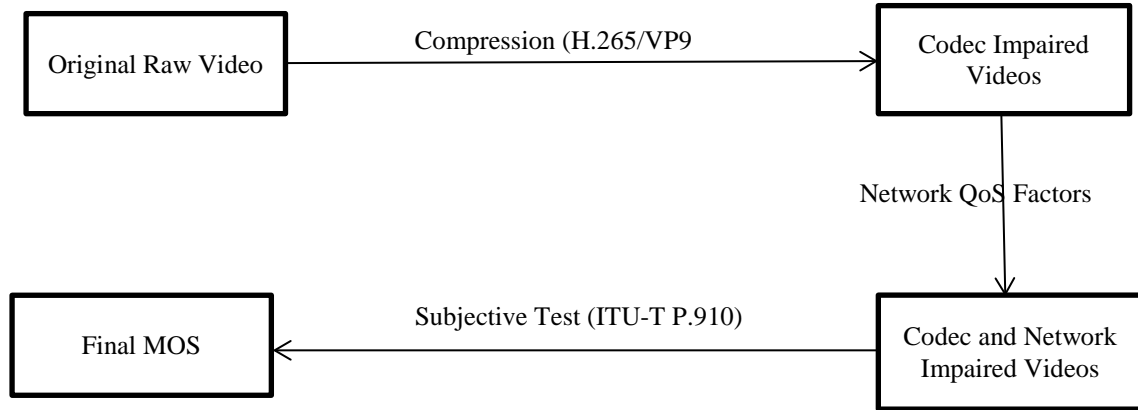

Fig. 5. Process of measuring MOS.

In the next section, we present the subjective test results and the QoS to QoE mapping approach.

4. Subjective Results and Individual QoS to QoE Mapping

4.1 Subjective Results

For the first subjective test we recorded 40,238 scores (341 impaired videos × 59 subjects × 2 codecs). **Tables 5 to 10** shows the summary of the results that we obtain for the six different network QoS factors. The standard deviation (σ) and mean MOS (with 95% CI) are shown for both the codecs H.265 and VP9 for a sample size of 59. For reasons of simplicity, results across the 11 different video sequences are consolidated into one for each factor.

To begin with, the process of outlier detection was performed in order to remove any sort of data inconsistency. If S_{ij} represents the score obtained by the jth subject for the ith test sequence, then S_{ij} would be considered as an outlier if $S_{ij} > q_3 + 1.5(q_3 - q_1)$ OR $S_{ij} < q_1 - 1.5(q_3 - q_1)$, q_1 and q_3 being the 25th percentile and 75th percentile of the score distribution respectively [52]. This range is approximately equal to 99.3% of the normally distributed data. A subject will be removed from any further consideration if more than 20% of his/her scores are outliers. Following this method in our experiment, we did not find any outliers. For a sample size N, the MOS has been calculated as:

$$MOS_i = \left(\sum_{j=1}^{N} S_{ij}\right)/N \tag{1}$$

Table 5. Consolidated Report for Packet Loss (PL) %

PL	σ(H.265/VP9)	Mean(95%CI)(H.265/VP9)
0.1	1.04/0.87	3.71±0.16/3.63±0.13
0.5	0.84/1.03	3.07±0.13/3.29±0.16
1	0.77/0.81	2.42±0.12/2.72±0.13
3	0.67/0.65	1.88±0.10/2.15±0.10
5	0.77/0.65	1.78±0.12/1.57±0.10
10	0.60/0.64	1.33±0.09/1.44±0.10

Table 6. Consolidated Report for Jitter (J) in milliseconds

J	σ(H.265/VP9)	Mean(95%CI)(H.265/VP9)
1	0.84/0.86	3.12±0.13/3.81±0.13
2	0.56/0.62	1.40±0.09/1.41±0.10
3	0.51/0.58	1.16±0.08/1.35±0.10
4	0.23/0.38	1.06±0.04/1.11±0.06
5	0.14/0.37	1.02±0.02/1.07±0.06

Table 7. Consolidated Report for Throughput (T) in kbps

T	σ(H.265/VP9)	Mean(95%CI)(H.265/VP9)
500	0.27/0.13	1.06±0.05/1.02±0.02
1000	0.67/0.63	1.76±0.12/1.67±0.11
2000	1.02/1.11	3.82±0.18/4.00±0.20
3000	0.97/0.97	3.87±0.17/4.09±0.17
5000	0.98/1.06	3.87±0.18/4.12±0.19

Table 8. Consolidated Report for Variable Initial Delay (VID) in seconds

VID	σ(H.265/VP9)	Mean(95%CI)(H.265/VP9)
5	0.80/0.73	4.27±0.12/4.36±0.11
15	1.11/0.99	3.24±0.17/3.98±0.15
25	0.89/1.13	2.03±0.14/3.14±0.18
35	0.76/0.89	1.68±0.12/2.43±0.14
45	0.44/0.83	0.75±0.07/1.67±0.13

Table 9. Consolidated Report for Buffering Delay (BD) in seconds

BD	σ(H.265/VP9)	Mean(95%CI)(H.265/VP9)
5	1.01/0.77	3.69±0.16/4.06±0.12
10	1.02/0.79	2.83±0.16/3.21±0.12
20	1.25/1.26	1.57±0.19/1.93±0.19
30	1.04/1.02	1.09±0.16/1.56±0.16
40	0.72/0.77	0.68±0.11/1.01±0.12

Table 10. Consolidated Report for Auto Scale Resolution (ASR)

ASR	σ(H.265/VP9)	Mean(95%CI)(H.265/VP9)
1	0.63/0.64	3.91±0.09/3.86±0.10
1.8	0.86/0.86	3.11±0.13/3.14±0.13
2.1	0.70/0.69	3.92±0.11/3.97±0.11
3.6	0.70/0.62	4.12±0.11/4.17±0.09
4.7	0.88/0.88	3.29±0.14/3.29±0.14

4.2 Individual QoS to QoE Mapping

Here, we go for the individual QoS to QoE mapping by using a non-linear regression approach. We select the optimal model based upon a decision variable DV. The overall goodness of fit statistics is generally expressed in terms of the sum of squared error (SSE), root mean square error (RMSE), R^2 change or the $adjusted - R^2$ change values. For SSE and RMSE; values closer to 0 indicate that the model has a smaller random error component, and that the fit will be more useful for prediction. Similarly, R^2 and $adjusted - R^2$ values closer to 1 indicate that a greater proportion of variance is accounted for by the model. When describing the modeling accuracy; different authors have used these different metrics. Therefore, in this research we use a combined decision variable DV that is defined as:

$$DV = \frac{(r^2 \times Adjusted - r^2)}{(SSE \times RMSE)} \tag{2}$$

Equation 2 suggests that a higher value of DV is always desirable. So, our model is optimized to get the highest value of DV possible. Apart from building our own model prediction equations, we also considered mathematical relationships between the QoS and QoE as outlined by other authors that have been pointed out in the related work section. In particular, we consider the following popular equation types as given from (3) to (8).

$QoE = a + bx$ (Linear mapping) $\tag{3}$
$QoE = a + bx + cx^2 + dx^3$ (Cubic polynomial [53]) $\tag{4}$
$QoE = \frac{a}{1+exp^{[-b(x-c)]}}$ (Logistic function [53][54]) $\tag{5}$
$QoE = aexp(-bx) + c$ (IQX hypothesis) $\tag{6}$
$QoE = -alog(x) + b$ (Logarithmic function [54]) $\tag{7}$
$QoE = a + \frac{d}{\left(1+\left(\frac{x}{c}\right)^b\right)^a}$ (Logistic function [55]) $\tag{8}$

The above equations are generic QoS to QoE mapping equations for which we need to find out the different coefficients as per the application scenario. Our regression based fitting functions are discussed next. We obtained the maximum value of the decision variable DV for our set of equations that has been shown in **Table 11.**

Table 11. Value of Decision Variable DV

Network QoS Factors	Decision Variable DV (H.265/VP9)	Generic Relationship
Packet Loss	1132.51/2407.75	Exponential Variation
Jitter	8328.25/9687.12	Exponential Variation
Throughput	60.13/54.66	Logarithmic
Variable Initial Delay	37.08/77.60	Exponential Variation
Buffering Delay	871.13/1052.12	Exponential Variation
Auto Scale Resolution	1235.68/915.54	Exponential Variation

.

Fig. 6(a) shows the relation between packet loss and MOS. Both the codecs show a similar performance trend. However, the performance of VP9 is slightly better especially at higher values of packet loss. There is a sharp decline in MOS for packet loss greater than 0.5%. The relationship between the two is expressed by the two-factor exponential function as given by equation 9. The coefficients a, b, c and d are found from our experiment i.e. from the

regression analysis that we carry out and presented in **Table 12.**

$$MOS_{PL} = a \times exp^{(b \times PL)} + c \times exp^{(d \times PL)} \tag{9}$$

Fig. 6(b) shows the graph of jitter vs. MOS. With an increase in jitter, the viewing quality considerably drops for both the codecs. However, the performance of VP9 is slightly better than H.265 for all values of jitter. Equation 10 shows the relationship between them. Coefficients a, b, c and d are given in **Table 12.**

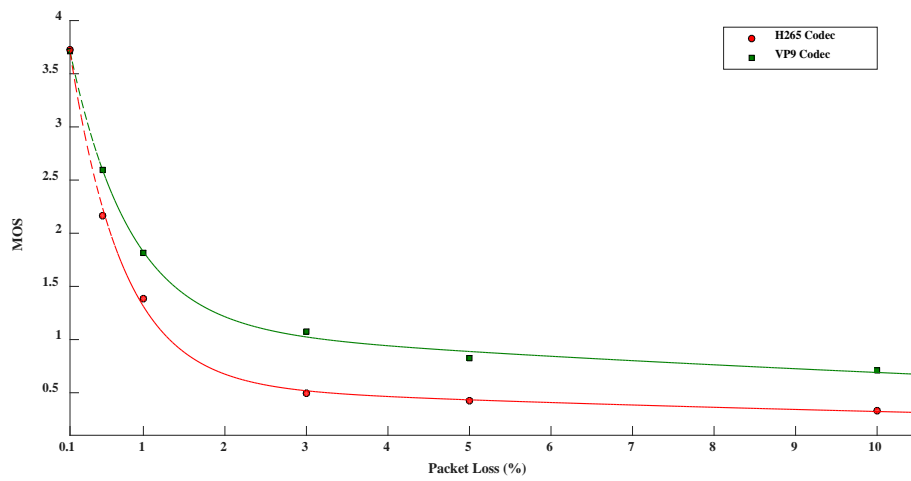$$MOS_J = a \times exp^{(b \times J)} + c \times exp^{(d \times J)} \tag{10}$$
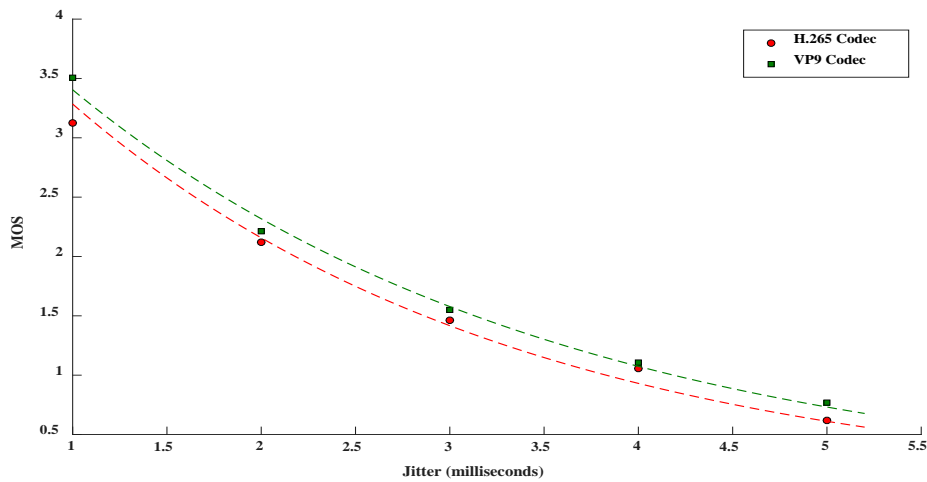

Fig. 6(a). Packet Loss vs. MOS

Fig. 6(b). Jitter vs. MOS

The relationship between MOS and the network throughput has been depicted in **Fig. 6(c).** In this case, we get a logarithmic relationship between the two as the best-fit model and has been shown by equation 11. Coefficient values are presented in **Table 12.** Performance of H.265 is

seen to be marginally better than VP9 for lower values of network throughput up to 2000 Kbps.

$$MOS_T = a \times log(T) + b \tag{11}$$

Variable initial delay was the new factor that we had introduced in this paper. **Fig. 6(d)** suggests that it has a significant impact on the viewing quality. As we increase the delay before the start of video playback, the MOS deteriorates rapidly. In fact, for initial delays greater than 15s the viewing quality is very poor. Their relationship has been shown by equation 12.

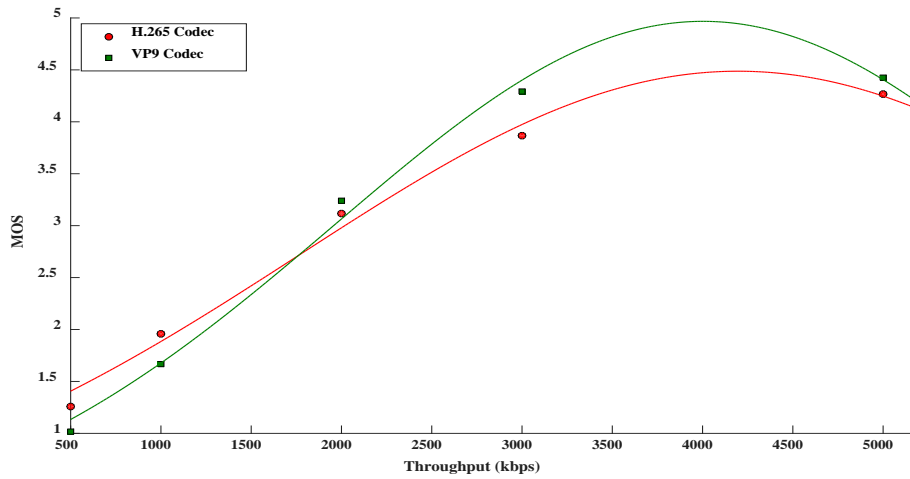$$MOS_{VID} = a \times exp^{(-b \times VID)} + c \tag{12}$$

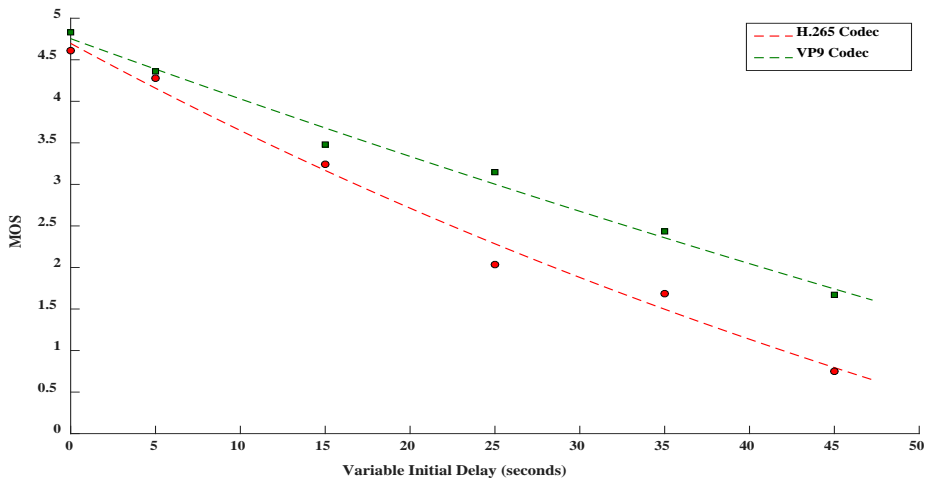
Fig. 6(c). Network Throughput vs. MOS

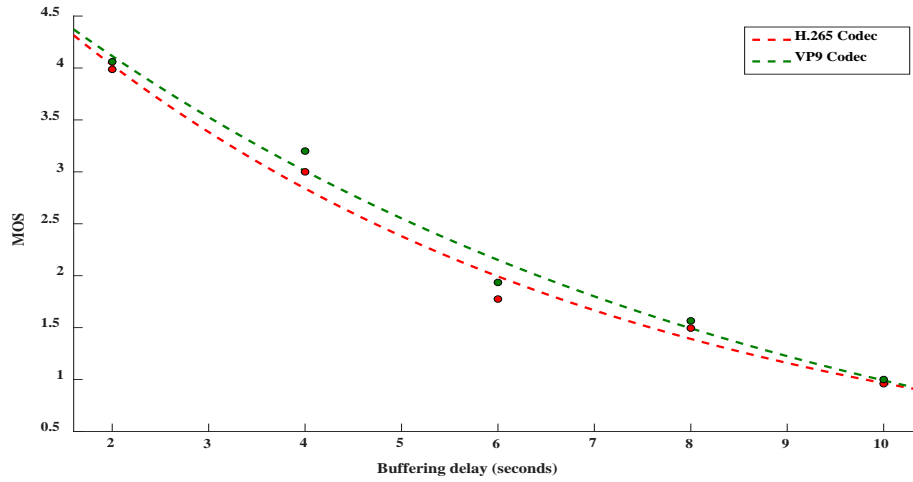**Fig. 6(d).** Variable Initial Delay Delay vs. MOS
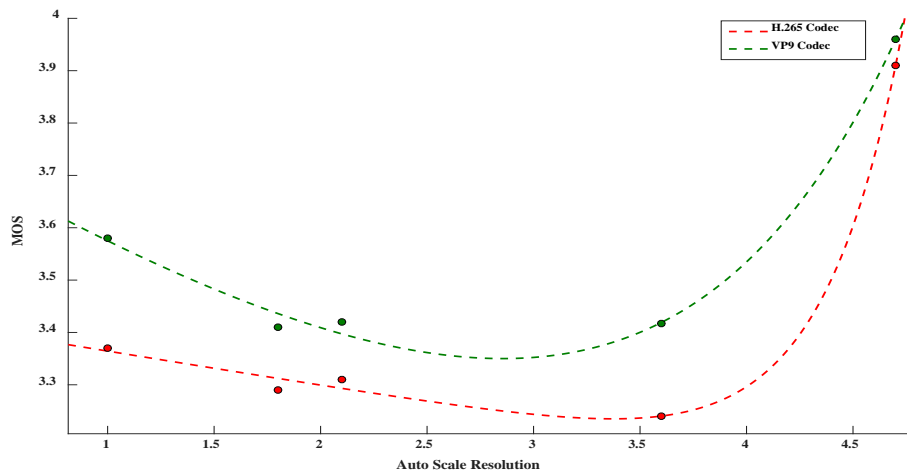
Fig. 6(e). Buffering Delay vs. MOS

Fig. 6(f). Auto Scale Resolution vs. MOS

Fig. 6(e) shows the effect of buffering delay on the viewing MOS. Even for slight buffering delays in the range of 5 to 10 seconds, there is a sharp fall in the MOS value. In fact, the viewing quality is severely affected as the value of buffering delay increases. Though marginally, the performance of VP9 seems to be a little better. Equation 13 depicts the relationship between the two. **Table 12** gives the coefficients.

$$MOS_{BD} = a \times exp^{(b \times BD)} + c \times exp^{(d \times BD)} \qquad (13)$$

Finally, **Fig. 6(f)** shows the graph of MOS vs. auto scale resolution. MOS is maximum for HD resolution combination. For other combinations too, this technique gives a satisfactory result. Thus, definitely such methods as employed by services like YouTube can improve the viewing quality. The relationship has been shown by equation 14.

$$MOS_{ASR} = a \times exp^{(b \times ASR)} + c \times exp^{(d \times ASR)} \qquad (14)$$

Table 12. Coefficient Values obtained from our Experiment

Codec	Parameter	a (95%CI)	b (95%CI)	c (95%CI)	d (95%CI)
H.265	Packet loss	3.66	-1.56	0.57	-0.06
H.265	Jitter	4.51	-0.37	-2.09×10^{-16}	6.73
H.265	Throughput	-1.39	-7.44	-	-
H.265	Variable Initial Delay	9.69	0.01	-4.99	-
H.265	Buffering Delay	4.26	-0.07	0.71	-0.01
H.265	Auto Scale Resolution	3.47	-4.46×10^{-8}	8.65×10^{-16}	1.15×10^{-5}
VP9	Packet loss	2.96	-1.38	1.13	-0.05
VP9	Jitter	11.62	-3.39	4.41	-0.35
VP9	Throughput	-1.65	-9.40	-	-
VP9	Variable Initial Delay	15.9	0.0047	-11.14	-
VP9	Buffering Delay	2.44	-0.11	3.04	-0.03
VP9	Auto Scale Resolution	3.38	-3.72×10^{-7}	0.69	4.66×10^{-7}

Next, we study the impact of the QoS factors considered on the video quality. For this, we perform an ANOVA (analysis of variance) on the MOS dataset that we have. The results have been shown in **Table 13.** The second column shows the Sum of Squares, third column is the Degrees of Freedom associated with the model, fourth column is the Mean Squares i.e. the ratio of the Sum of Squares to the Degrees of Freedom, fifth column shows the F-statistic value and the sixth column represents the p-value. We observe from the table that for both the 3 KPI's as well as the 3 factors that we have introduced in this paper; all the predictors are statistically significant (p-value of less than 0.01). Hence, we justify their inclusion in this research.

We also calculate the PCC (Pearson Correlation Coefficient) of our model for the various factors considered. This has been shown in **Table 14.** We see that the MOS values which our model predicts has a high degree of correlation with the actual subjective scores.

Table 13. ANOVA Result for the predictors

Parameter	Sum of Squares	Degrees of Freedom	Mean Squares	F Statistic	p- Value
PL	15.747	5	3.149	37.382	1.91×10^{-4}
J	8.505	4	2.126	114.556	4.2×10^{-5}
T	15.695	4	3.923	109.568	4.7×10^{-3}
VID	18.561	5	3.712	13.798	3×10^{-3}
BD	12.51	4	3.128	43.038	4.59×10^{-4}
ASR	1.433	4	0.358	59.70	2.08×10^{-4}

Table 14. Correlational Analysis of the Network QoS Factors

Parameter	Pearson Correlation
Packet Loss	0.952
Jitter	0.978
Throughput	0.874
Variable Initial Delay	0.885
Buffering Delay	0.914
Auto Scale Resolution	0.924

5. Integrated QoE Measurement

In this section we try to evaluate the integrated QoE of all the network QoS factors that we have introduced here. A subjective test containing multi-factor impaired videos is also conducted in the last to establish the validity of the proposed model. To the best of our knowledge, all the QoE assessment models in literature follow the weighted sum approach. We try to obtain the integrated QoE by adopting the following procedure:

- In step 1, the overall QoE due to all the network QoS factors is found out by using an additive approach. We use the Analytic Hierarchy Process (AHP) technique for this purpose and denote the QoE obtained as QoE_{Add}.
- In step 2, we propose a method to calculate the QoE using a multiplicative approach and denote it by QoE_{Mul}.
- Finally, in step 3 we take into account the interaction between QoE_{Add} and QoE_{Mul} to propose the integrated QoE using a linear regression approach and denote it by $QoE_{Integrated}$.

5.1 Finding the Additive QoE

AHP algorithm has been widely used in the past to solve multi-index problems. In the additive approach, every QoS factor has to be assigned a weight depending upon the extent of its effect on the viewing experience. The general form for additive QoE is shown in equation 15.

$$QoE_{Add} = w_1 MOS_1 + w_2 MOS_2 + \cdots w_n MOS_n \tag{15}$$

The key here is to calculate the weights w_1, w_2, etc., for which we use the AHP technique. **Table 15** shows the normalized weight of the various QoS parameters as calculated by this method. Thus, for our case equation 15 reduces to:

$$QoE_{Add} = 0.26 MOS_{PL} + 0.30 MOS_J + 0.04 MOS_T + 0.09 MOS_{VID} + 0.10 MOS_{BD} + 0.20 MOS_{ASR} \tag{16}$$

Table 15. Normalized Weight of all the Network QoS Factors

Network QoS Parameter	Normalized Weight
Packet Loss	0.26
Jitter	0.30
Throughput	0.04
Variable Initial Delay	0.09
Buffering Delay	0.10
Auto Scale Resolution	0.20

QoE, which is obtained by this weighted sum approach, has a disadvantage. A video that has been distorted by two QoS metrics should not have a better QoE than the video which has been distorted by only one of the two QoS metrics. This statement should be valid for any multiple QoS factor impaired videos. We illustrate this with an example given in **Table 16.** The MOS values are calculated from the individual QoS to QoE mapping functions that we presented previously in equations 9 to 14. The normalized MOS is calculated by multiplying each individual network QoE factor by its weight. Finally, QoE_{Add} is obtained by adding the

corresponding impairment terms. In this particular case, the range of the MOS values is from 0.62-4.07. The normalized MOS value obtained by the additive approach is 2.028, which is within the range. However, it contradicts the fact that the MOS should not be greater than 0.62, which is the minimum MOS obtained in this case. Thus, clearly there is an anomaly while calculating the MOS using the weighted sum approach. We overcome this problem in the next section.

Table 16. Calculation of QoE_{Add}

Network Factor	QoS Value	MOS	Normalized MOS
Packet Loss	1%	1.31	0.34
Jitter	5 ms	0.62	0.18
Throughput	2000 Kbps	3.10	0.12
Variable Initial Delay	15 s	3.17	0.28
Buffering Delay	10 s	2.78	0.27
Auto Scale Resolution	1	4.07	0.81
QoE_{Add}			2.03

5.2 Finding the Multiplicative QoE

In order to solve the problem of the weighted sum approach, we introduce the multiplicative form in this section. Since, the MOS rating is given on a scale of 1 to 5 hence; we normalize the effect of each QoS factor by dividing it by 5. This has been shown in equation 17.

$$QoE_{Mul} = 5 \times \left(\frac{MOS_{PL}}{5}\right) \times \left(\frac{MOS_{J}}{5}\right) \times \left(\frac{MOS_{T}}{5}\right) \times \left(\frac{MOS_{VID}}{5}\right) \times \left(\frac{MOS_{BD}}{5}\right) \times \left(\frac{MOS_{ASR}}{5}\right) \qquad (17)$$

We find that each individual QoE factor has been normalized on a scale of 5, while evaluating its contribution towards the final multiplicative QoE_{Mul}. This can eliminate the problem that was associated with the weighted sum approach. **Table 17** shows the sample calculation for QoE_{Mul} using equation 17. The same set of values for the various network QoS factors have been used that we had considered while calculating QoE_{Add}.

We observe that the QoE_{Mul} value is 0.03 for this particular set of condition that is considerably lesser than the minimum MOS value of the network QoS factors considered (0.62 for jitter in this case). Thus, the problem associated with the weighted sum approach has been solved.

Table 17. Calculation of QoE_{Mul}

Network Factor	QoS Value	MOS	Normalized MOS
Packet Loss	1%	1.31	0.26
Jitter	5 ms	0.62	0.12
Throughput	2000 Kbps	3.10	0.62
Variable Initial Delay	15 s	3.17	0.63
Buffering Delay	10 s	2.78	0.56
Auto Scale Resolution	1	4.07	0.81
QoE_{Mul}			0.03

Comparing the values of QoE_{Add} and QoE_{Mul} for the same set of network QoS values reveal that the additive approach tends to over-predict the actual viewing quality, while the multiplicative approach tends to under-predict the same. Hence, we must also consider the effect due to the interaction between these two factors i.e. QoE_{Add} and QoE_{Mul}. Therefore, in

order to find the final integrated QoE of the video streaming service, we use a linear regression approach to include the effect of both.

5.3 Towards the Final Integrated QoE

In order to evaluate and validate our integrated QoE function for multiple network QoS impairments, we carry out a second subjective test which contain videos distorted by more than one QoS metrics. The same 59 subjects who participated in the first test were considered here too in order to remove any chances of subject bias. We created 88 multi-factor impaired video sequences for obtaining the subjective results across both the codecs (44 for each codec). The impairment detail has been shown in **Table 18.**

Table 18. Details of Second Subjective Test₁

No of Impairment Factors	Impairment Factor Details	No of Impaired Video Sequences
2	Packet Loss + Jitter	8
2	Jitter + Throughput	6
3	Packet Loss + Variable Initial Delay + Auto Scale Resolution	6
3	Jitter + Buffering Delay + Throughput	6
4	Packet Loss + Throughput + Buffering Delay + Auto Scale Resolution	6
4	Jitter + Auto Scale Resolution + Variable Initial Delay + Throughput	6
6	Packet Loss + Jitter + Throughput + Variable Initial Delay + Buffering Delay + Auto Scale Resolution	6

Fig. 7 shows our proposed integrated QoE model and it is described by equation 18

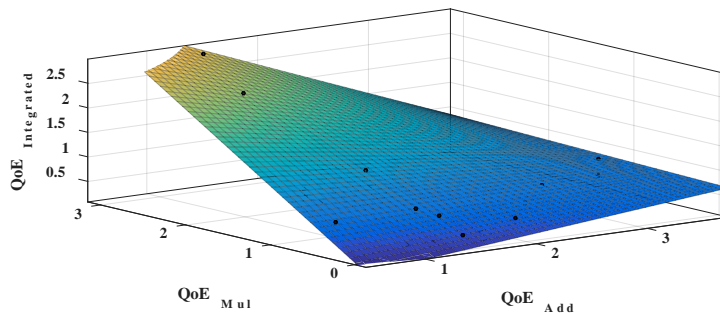- $$QoE_{Integrated} = f(QoE_{Add}, QoE_{Mul})$$

Fig. 7. Proposed QoE_{Integrated} as a function of QoE_{Add} and QoE_{Mul}

$$QoE_{Integrated} = \quad 0.18QoE_{Add} + 1.33QoE_{Mul} - 0.34(QoE_{Add} \times QoE_{Mul}) \tag{18}$$

Therefore, we are able to obtain an integrated QoE model for various network QoS factors that consider the interaction between the additive and multiplicative effects of QoE. Equation 18 suggests that QoE_{Mul} has a greater contribution towards $QoE_{Integrated}$ as compared to QoE_{Add} (coefficient values of 1.33 and 0.18 respectively). Next, we consider the accuracy of the proposed model.

5.4 Accuracy of the Proposed Model

The accuracy of our proposed model has been shown in Fig. 8. Table 19 illustrates the R^2, adjusted R^2 and PCC values for all the three stages of model building. We note that there is a gradual increase in the prediction accuracy of the model. Thus, the new integrated QoE model that we have proposed here for the various network QoS factors is feasible for an online video streaming service.

Table 19. Network Modelling Accuracy

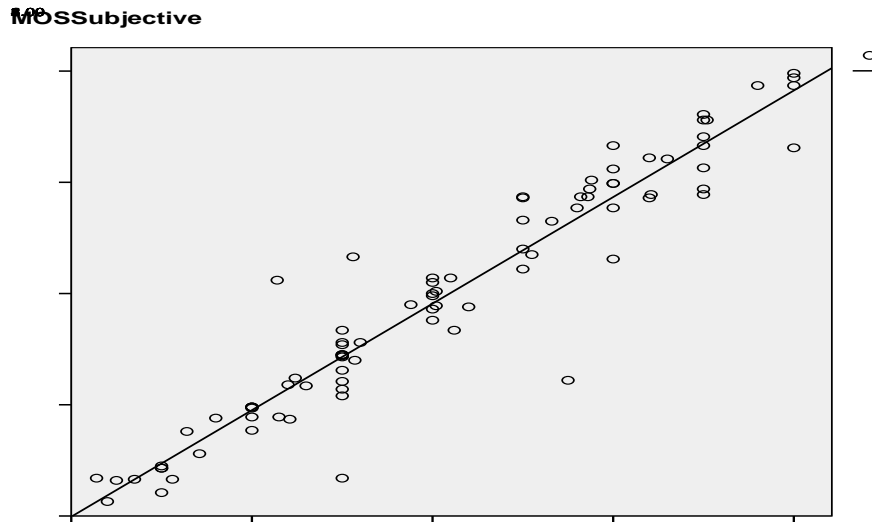| Model Stages	R^2	Adjusted R^2	PCC
Additive	0.654	0.649	0.808
Multiplicative	0.889	0.888	0.943
Regression Based	0.913	0.912	0.951

Fig. 8. Accuracy of Our Model

5. Conclusion

In this paper, we have proposed an integrated QoE evaluation model for multiple network QoS parameters in video streaming application. The experiments are carried out using the latest generation H.265 and VP9 codecs. We start the process by identifying the appropriate network QoS factors that have a potential effect on the viewing quality. In doing so, apart from the

three well-established network QoS KPI's, we introduced three more factors into consideration. A ANOVA analysis revealed the justification in including the new factors. After this, we obtained the individual QoS to QoE mappings for all the parameters considered. Optimality of the chosen model was calculated on the basis of a decision variable DV. Based upon these individual QoS to QoE mapping functions, next we calculated the overall QoE based upon an additive and multiplicative approach. AHP technique was used for the additive approach. We obtained the final integrated QoE model by using a linear regression technique that takes into consideration the effect of both the additive and multiplicative QoE's. This calculated $QoE_{Integrated}$ has a high accuracy in predicting the quality of muti-factor impaired videos that was proved from further subjective tests.

In this study, we considered the effects of network QoS only. However, other codec related parameters like bit-rate, frame rate, etc. can also affect the video quality that we did not take into account. The effect of these factors on video quality will be investigated as part of our future work.

References

[1]   Cisco Global Mobile Data Traffic Forecast Update Report, 2014-2019, Cisco White Paper (2016).

[2]   Ooyala, Q1 2013 Video Index-TV is no longer a single screen in your living room, 2013.

[3]   P. Callet, S. Möller and A. Perkis, "Qualinet White Paper on Definitions of Quality of Experience," Lausanne, Switzerland, 2013.

[4]   Quality of Experience Requirements for IPTV Services, ITU-T Recommendation G.1080, December 2008.

[5]   T. Hobfeld, R. Schatz and A.D'Alconzo, "Challenges of QoE Management for Cloud Applications," *IEEE Communications Magazine*, vol. 50, no. 4, pp. 28-36, 2012. Article (CrossRef Link)

[6]   K. Piamrat, C. Viho, J. M. Bonnin and A. Ksentin, "Quality of Experience Measurements for Video Streaming over Wireless Network," in *Proc. Of 6th Int. Conf. on Information Technology: New Generations*, pp. 1184-1189, November 2009. Article (CrossRef Link)

[7]   Subjective Video Quality Assessment Methods for Multimedia Applications, ITU-T Recommendation P.910, June 2008.

[8]   Methodology for the Subjective Assessment of the Quality of Television Pictures, ITU-T Recommendation BT.500, January 2012.

[9]   Methods for Subjective Determination of Transmission Quality, ITU-T Recommendation P.800, August 1996.

[10] Subjective Audiovisual Quality Assessment Methods for Multimedia Applications, ITU-T Recommendation P.911, December 1998.

[11] M. Vranjes, S. Rimac-Drlje and D. Zagar, "Subjective and Objective Quality Evaluation of the H.264/AVC Coded Video," in *Proc. of 15th Int. Conf. on Systems, Signal and Image Processing*, pp. 287-290, 2008. Article (CrossRef Link)

[12] Quan H. Thu and M. Ghanbari, "The Accuracy of PSNR in Predicting Video Quality for Different Video Scenes and Frame Rates," *Telecommunication Systems*, vol. 49, no. 1, pp. 35-48, January 2012. Article (CrossRef Link)

[13] Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004. Article (CrossRef Link)

[14] M. H. Pinson and S. Wolf, "A New Standardized Method for Objectively Measuring Video Quality," *IEEE Transactions on Broadcasting*, vol. 50, no. 3, pp. 312-322, September 2004. Article (CrossRef Link)

[15] Opinion Model for Video Telephony Applications, ITU-T Recommendation G.1070, July 2012.

[16] Kim. H. L and S.G. Choi, "A Study on a QoS/QoE Correlational Model for QoE Evaluation on IPTV Service," in *Proc. of 12th International Conference on Advanced Communication Technology*, pp. 1377-1382, 2010. Article (CrossRef Link)

[17] Kim. H. L, D. H. Lee, J. M. Lee, K.H. Lyu and S.G. Choi, "The QoE Evaluation Method through the QoS-QoE Correlational Model," in *Proc. of NCM'08*, pp. 719-725, 2008.
Article (CrossRef Link)

[18] A. Khan, L. Sun and E. Ifeachor, "QoE Prediction Model and its Application in Video Quality Adaptation over UMTS Networks," *IEEE Transactions on Multimedia*, vol. 4, no. 2, pp. 431-442, 2012. Article (CrossRef Link)

[19] B. Wang, X. Wen and Z. Wei, "A new approach measuring user's QoE in the IPTV," in *Proc. of Asia-Pacific Conference on Circuits, Communication and Systems*, pp. 453-456, 2009.
Article (CrossRef Link)

[20] F. Mercer Moss, K. Wang, F. Zhang, R. Baddeley and D. R. Bull, "On the Optimal Presentation Duration for Subjective Video Quality Assessment,*" IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 11, pp. 1977-1987, November 2016. Article (CrossRef Link)

[21] M. H. Pinson and S. Wolf, "Comparing Subjective Video Quality Testing Methodologies," in *Proc. of SPIE Video Communications and Image Processing Conference*, pp. 573-582, 2003.
Article (CrossRef Link)

[22] D.M. Rouse, R. Pepion R. Callet and S. Hemami, "Tradeoffs in Subjective Testing Methods for Image and Video Quality Assessment," in *Proc. of SPIE, Human Vision and Electronic Imaging Conference*, 2010. Article (CrossRef Link)

[23] P. Romaniak, M. Mu, A. Mauthe, M. Leszczuk and E. Cerqueria, "Framework for the Integrated Video Quality Assessment," Multimedia Tools and Applications, vol. 61, no. 3, pp. 787-817, December 2012. Article (CrossRef Link)

[24] S. Kanumuri, P. C. Cosman, A. R. Reibman and V. A. Vaishampayan, "Modeling Packet-loss Visibility in MPEG-2 video," *IEEE Transactions on Multimedia*, vol. 8, no. 2, pp. 341-355, April 2006. Article (CrossRef Link)

[25] J. Søgaard, S. Forchhammer and J. Korhonen, "No-Reference Video Quality Assessment Using Codec Analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 10, pp. 1637-1650, October 2015. Article (CrossRef Link)

[26] M. Sahid, A. Rossholm, B. Lövstörm and H. Zepernick, "No-reference Image and Video Quality Assessment: A Classification and Review of Recent Approaches," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1-32, August 2014. Article (CrossRef Link)

[27] R. Mok, E. Chan and R.K.C Chang, "Measuring the Quality of Experience of HTTP Video Streaming," in *Proc. of IEEE International Symposium on Integrated Network Management*, pp. 485-492, 2011. Article (CrossRef Link)

[28] S. Khirman and P. Henricksen, "Relationship between Quality of Service and Quality of Experience for Public Internet Services," in *Proc. of 3rd Workshop on Passive and Active Measurement*, pp. 23-28, March 2002.

[29] P. Romaniak, "Towards Realization of a Framework for Integrated Video Quality of Experience Assessment," in *Proc. of IEEE INFOCOM Workshops*, pp. 1-2, 2009. Article (CrossRef Link)

[30] M. Fiedler, T. Hossfeld and P. Tran-Gia, "A Generic Quantitative Relationship between Quality of Experience and Quality of Service," *IEEE Network*, vol. 24, no. 2, pp. 36-41, April 2010.
Article (CrossRef Link)

[31] P. Reichl, S. Egger, R. Schatz and A. D'Alconzo, "The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment," in *Proc. of IEEE Int. Conf. on Communications*, pp. 1-5, 2010. Article (CrossRef Link)

[32] J. A. Lozano, A. Castro, B. Fuentes, J. M. González and Á. Rodríguez, "Adaptive QoE Measurement on Videostreaming IP services," in *Proc. of 7th Int. Conf. on Network and Service Management*, pp. 1-4, 2011. Article (CrossRef Link)

[33] L. Skorin-Kapov and M. Varela, "A multi-dimensional view of QoE: the ARCU model," in *Proc. of 35th Int. Convention MIPRO*, pp. 662-222, 2012. Article (CrossRef Link)

[34] D. Yun; K. Chung, "DASH-based Multi-view Video Streaming System," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no.99, pp.1-1, April 4 2017. Article (CrossRef Link)

[35] S. Colonnese, F. Cuomo, T. Melodia and I. Rubin, "A Cross-Layer Bandwidth Allocation Scheme for HTTP-Based Video Streaming in LTE Cellular Networks," *IEEE Communications Letters*, vol. 21, no. 2, pp. 386-389, Feb. 2017. Article (CrossRef Link)

[36] K. Jia, Y. Guo, Y. Chen and Y. Zhao, "Measuring and Predicting Quality of Experience of DASH-based Video Streaming over LTE," in *Proc. of 19th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Shenzhen, China, 2016, pp. 102-107. Article (CrossRef Link)

[37] T. Mäki, M. Varela and D. Ammar, "A Layered Model for Quality Estimation of HTTP Video from QoS Measurements," in *Proc. of 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, Bangkok, 2015, pp. 591-598. Article (CrossRef Link)

[38] J. Jiang, V. Sekar and H. Zhang, "Improving Fairness, Efficiency, and Stability in HTTP-Based Adaptive Video Streaming With Festive," *IEEE/ACM Transactions on Networking*, vol. 22, no. 1, pp. 326-340, Feb. 2014. Article (CrossRef Link)

[39] Y. Chen, K. Wu and Q. Zhang, "From QoS to QoE: A Tutorial on Video Quality Assessment," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1126-1165, 2015. Article (CrossRef Link)

[40] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hoßfeld and P. Tran-Gia, "A Survey on Quality of Experience of HTTP Adaptive Streaming," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 469-492, 2015. Article (CrossRef Link)

[41] P. Juluri, V. Tamarapalli and D. Medhi, "Measurement of Quality of Experience of Video-on-Demand Services: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 401-418, 2016. Article (CrossRef Link)

[42] Sabina Barakoviï and Lea Skorin-Kapov, "Multidimensional Modeling of Quality of Experience for Mobile Web Browsing," *Computers in Human Behavior*, vol. 50, no. C, pp. 314-332, September 2015. Article (CrossRef Link)

[43] H. Kim and S. Chobi, "QoE Assessment Model for Multimedia Streaming Services using QoS Parameters," *Multimedia Tools and Applications*, vol. 72, no. 3, pp. 2163-2175, May 2013. Article (CrossRef Link)

[44] B. Belmudez and S. Möller, "Audiovisual Quality Integration for Interactive Communications," *EURASIP Journal on Audio, Speech and Music Processing*, vol.2013, no. 1, pp. 1-23, 2013. Article (CrossRef Link)

[45] W. Song, D. Tjondronegoro and M. Docherty, "Exploration and Optimization of User Experience in Viewing Videos on a Mobile Phone," *International Journal on Software Engineering Knowledge*, vol. 20, no. 8, pp. 1045-1075, 2010. Article (CrossRef Link)

[46] VQEG Standard Database maintained at http://www.its.bldrdoc.gov/vqeg/downloads.aspx

[47] J. Joskowiz and L. Ardao, "Enhancements to the Opinion-Model for Video Telephony Applications," in *Proc. of 5th Int. Latin American Networking Conference*, pp. 87-94, 2009. Article (CrossRef Link)

[48] A. Khan, L. Sun and E. Ifeachor, "Impact of Video Content on Video Quality for Video over Wireless Networks," in *Proc. of 2009 Fifth International Conference on Autonomic and Autonomous Systems*, pp. 277-282, 2009. Article (CrossRef Link)

[49] J. Nightingale, Q. Wang, C. Grecos and S. Goma, "The Impact of Network Impairment on Quality of Experience (QoE) in H.265/HEVC Video Streaming," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 2, pp. 242-250, May 2014. Article (CrossRef Link)

[50] N. Laoutaris and I. Stavrakakis, "Adaptive Playout Strategies for Packet Video Receivers with Finite Buffer Capacity," in *Proc. of IEEE Int. Conf. on Communications*, pp. 969-973, 2001. Article (CrossRef Link)

[51] M. M. Krunz and M. Hassan, "Adaptive rate control scheme for video streaming over wireless channels," in *Proc. of Conf. on Data Compression*, pp. 242-251, 2004. Article (CrossRef Link)

[52] D. Simone, F. Goldmann and T. Ebrahimi, "Towards High Efficiency Video Coding: Subjective

Evaluation of Potential Coding Technologies," *Journal of Visual Communication and Image Representation*, vol. 22, no. 8, pp. 734-748, 2011 Article (CrossRef Link)

[53] J. Korhonen, N. Burini and E. Nadernejad, "How to evaluate Objective Video Quality Metrics Reliably," in *Proc. of 4th International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 57-62, 2012. Article (CrossRef Link)

[54] Methodological Framework for specifying accuracy and Cross calibration of Video Quality Metrics, ITU-R BT.1676, 2004.

[55] Estimating end to end Performance in IP Networks for Data Applications, ITU-Recommendation G.1030, 2005.

Debajyoti Pal received the B.E degree in Electrical Engineering from Nagpur University, Maharashtra, India, in 2005 and the M.Tech. degree in Information Technology from Indian Institute of Engineering Science and Technology, Shibpur, Kolkata, India in 2007. He has recently completed the Ph.D. in Information Technology from the School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand in 2017. Presently, he is employed as a researcher with the same university. His research interests are in multimedia systems, quality evaluation of various multimedia services, Internet of Things and IT adoption just to name a few.
E-mail: debajyoti.pal@mail.kmutt.ac.th.

Vajirasak Vanijja received both the B.Sc and M.Sc. degrees in Computer Science from King Mongkut's University of Technology Thonburi, Bangkok, Thailand in 1998 and 2000 respectively. He received his PhD. in Information Science from Japan Advanced Institute of Science and Technology, Ishikawa, Japan in 2004. Currently he is an assistant professor at the School of Information Technology, KMUTT, Bangkok and the senior associate dean for business affairs. His research interests are in multimedia systems, virtual reality, image processing, computer graphics and Voice over IP (VoIP). He also has a number of patents in the area of image processing.
E-mail: vachee@sit.kmutt.ac.th