Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2018) 26: 2858 – 2870

doi:10.3906/elk-1712-155

An adaptive machine learning-based QoE approach in SDN context for video-streaming services

Asma BEN LETAIFA*
MEDIATRON Laboratory, SUPCOM, University of Carthage, Tunis, Tunisia

Received: 11.12.2017 • **Accepted/Published Online:** 11.06.2018 • **Final Version:** 29.11.2018

Abstract: In data service applications over the Internet, user perception and satisfaction can be assessed by quality of experience (QoE) metrics. QoE depends both on the users' perception and the used service, which together form end-to-end metrics. While network optimization has traditionally focused on optimizing network properties such as QoS, we focus in this work on optimizing end-to-end QoE metrics with the aim to deliver to the client a good QoE that can be monitored in real time. We argue that end-user QoE is a relevant measurement for network operators and service providers. In this paper, we present a machine learning approach combined with adaptive video delivery service in order to provide a better QoE for video streaming services. This solution will be established using SDN architecture. The first part of the paper deals with a brief introduction of SDN networks, QoE requirement, and ML algorithms. Secondly, we expose the rating of the web application that we developed. This will help in conducting a subjective study to collect MOS on real-time as well as objective parameters SSIM, VQM, and PSNR. At the end, we expose our QoE-aware monitoring approach and explain what it is based on.

Key words: Quality of experience, SDN, machine learning, mean opinion score, video streaming services

1. Introduction

In today's world, video streaming has risen above all other types of traffic. In fact, providing this service with high quality presents the most challenging task among all the advancements in networking technologies. Thus, researchers are trying to help ensure a high degree of video quality of experience (QoE) since the traditional network quality of service (QoS) parameters (e.g., bandwidth, jitter, packet loss) are no longer sufficient to provide satisfaction for end-users. In this work, a QoE-aware monitoring approach for video streams is described. Our system monitors the video parameters in real time. Moreover, it dynamically makes adaptive decisions based on predictions of end-user perception for the video quality, which is quantified by mean opinion score (MOS) and estimated thanks to a machine learning process. The experimental results show that our proposed approach leads to a good QoE that meets the end-users' expectations as well as video stream quality thanks to measurement done on objective metrics such as SSIM, PSNR, and VQM. This paper deals with QoE in the SDN environment and especially for video streaming services, which need real-time monitoring. We describe our developed tool to enhance, monitor, and offer adaptive QoE in the SDN context. The first part of this paper will describe the SDN environment and QoS/QoE requirements. We then present subjective and objective methods mentioned in the related work and give an overview of parameters that are useful to follow, calculate, and apply with real-time modifications in order to correctly serve the end user. Thus, this part will describe

*Correspondence: asma.benletaifa@supcom.tn

our approach. We begin by describing the Web application rating to perform Ration and collect MOS. The next step is to perform a machine-learning algorithm to predict Estimated_MOS. The following sections give a description of a test bed and Mininet environment to emulate SDN networks with VLC in order to perform a video streaming service. The last part of the article describes curves and tables obtained in the study with the aim of concluding about the impact of resolution, buffering, bitrate, and number of frames per second on MOS. At the end, we highlight the future of our work.

2. Problem and background

There has been a huge shift to software defined network architecture [1, 2] by companies since the traditional network architecture is no longer appropriate to handle the significant load on networks. To cope with this dramatic change, the SDN architecture presents a new salient solution that consists of separating the control plane from the data plane, which were typically coupled together within one piece of equipment. Software defined networks (SDNs) offer methods to make dynamic topologies. Video streaming services require good video quality for their clients. In order to achieve this condition, operators and services providers are using a huge variety of codecs. Operations such as acquisition, processing, compression, storage, transmission, reproduction, and any variety of distortion can be the source of degradation in the visual quality. How can we measure video quality? In order to correctly manage QoE for clients, operators have to measure network parameters in real time. Our proposed tool performs measurements in a cyclical manner for parameters such as RTT, delay, bandwidth, and jitter. In the next subsection, we present the main parameters used in our proposed tool.

3. Related work

Quality of service is defined by the ITU as the "totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs". It defines also other metrics helping to provide guarantees in terms of delay, jitter, and packet loss that are useful for user satisfaction. For example, ITU-T Recommendation G.1010 states that for interactive voice communication the delay should be below 150 ms and the packet loss rate should be below 3%. On the other hand, QoE is defined as "the degree of delight or annoyance of the user of an application or service". It results from the fulfillment of the user's expectations with respect to the utility or enjoyment of the application or service in light of the user's personality and current state. QoE aims to capture the user's perception when using a service. To address this challenge, we review in this paper the most common quality metrics/estimators especially for video streaming services. These metrics estimate a set of influence factors impacting QoE. For an extended overview of available metrics, we refer to [3] for speech quality, to [4, 5] for image quality, and to [6, 8] for video quality. QoE metrics can also be classified into three categories by the required amount of reference information. Full-reference (FR) metrics estimate the QoE score based on the original signal and received perturbed signal. No-reference (NR) metrics account for this challenge and estimate QoE_score based purely on the received signal. Reduced-reference (RR) metrics account for this inaccuracy and estimate it based on a subset of features.

In FR cases, several metrics can be found, such as MSU, PSNR, VQM, and SSIM [9, 11]. The challenge of QoE is encountered since the user's perception is subjective. The main objective of QoE is to quantify user perceptions of applications starting from service generation, including transport entities, until the end of the device's screen or audio unit. User satisfaction and service quality are strongly correlated with the end-to-end serving entities.

The quality of the source video, the interference introduced during transmissions, the decoding, and display on client terminals are the major factors that affect QoE [12, 14]. A lot of work was done on video quality evaluation with transport disturbances interest [15]. In a mobile system, the video streaming quality perceived by mobile clients is a function of the specific mobile environment and the terminals. The work described in [16] paid attention to transport data bit errors but not to interruptions such as initial buffering and rebuffering, which was really observed in [17]. The work in [18] dealt with a subjective test of streaming quality-based rebuffering length and rebuffering frequency as influential factors. In the last decade, there has been an increasing interest in developing objective quality metrics for evaluation of different types of digital video distortion, especially in a heterogeneous communication environment. Several objective metrics have been recently developed showing a good correspondence with the subjective mean opinion score (MOS) [19, 20]. In the literature, two different approaches are used in objective quality metrics [21] to extract important features in the original and distorted signal as well as evaluate differences between them.

ITU-R Rec. BT.500-11 presents a subjective test as video quality perceived by the human eye. It defines standards to regulate the measurement, which consist of strict viewing conditions, a different type of display used, a source signal selected, a minimum of 15 or more nonexpert observers, and a group of test sessions that should not last more than 30 min, where the length of each clip is recommended to be 5 or 10 s. Subjective approaches standardized by ITU-R BT.500 and ITU-T P.910 [22, 23] are especially cited here: DSIS: Double Stimulus Impairment Scale, DSCQS: Double Stimulus Continuous Quality Scale, SSCQE: Single Stimulus Continuous Evaluation, SDSCE: Simultaneous Double Stimulus for Continuous Evaluation, SAMVIQ: Subjective Assessment Methodology for Video Quality, ACR: Absolute Category Rating, SC: Stimulus Comparison.

Other researchers based their work on machine learning algorithms to estimate QoE for multimedia services. These algorithms give good results in real time; for example, [24] proposed the MLQoE, a modular algorithm for user-centric QoE prediction. This framework employs multiple machine learning (ML) algorithms, namely artificial neural networks, support vector regression machines, decision trees, and Gaussian naive Bayes classifiers, and tunes their hyperparameters for VoIP service. It uses also the nested cross-validation method to select the best classifier and the corresponding best hyperparameter values and predicts the performance of the final model. This model predicts accurately the QoE score. Specifically, a mean absolute error of less than 0.5 and median absolute error of less than 0.30 can be achieved. On the other hand, Mushtaq et al. [25] achieved the correlation between QoS and QoE in search of capturing the degree of user opinion, based on ML to determine the most suitable one for the task of QoS/QoE correlation in order to study the effect of the QoS metric on the QoE to deliver a better quality of service to end-users. This work evaluated six classifiers and determined the most suitable one for the task of QoS/QoE correlation. Experimental results showed that, in the case of mean absolute error rate, it is observed that DT has a good performance as compared to all other algorithms. Also, in [26], a machine learning technique was proposed using a subjective quality feedback. This technique was used to model dependencies of different QoS metrics related to network and application layer on the QoE of the network services and they summarized this an accurate QoE prediction model. Finally, the authors in [27] used ML for estimating audiovisual quality of multimedia service. They trained the models with random forests and multilayer perceptron methods; results showed that random forests-based methods outperformed multilayer perceptron methods in terms of RMSE, Pearson correlation coefficient value, and 95% confidence interval boundaries.

4. Proposed approach

Research on QoE is often based on subjective studies. In such subjective studies, users rank the perceived quality of a service or application. These studies are carried out in specialized laboratories. However, these subjective tests are tedious and costly. Moreover, this type of test is not applicable in a real-time system, which is the case for most media services. In this context, researchers have focused on new methods that approximate and estimate the quality of experience in an objective manner that can be used in real-time contexts. The main disadvantage of the existing solutions lies in the fact that they are not correlated with subjective tests and therefore cannot adequately reflect the end user's perception.

The main objective of this work is to propose an efficient method based on a machine learning algorithm to predict and estimate MOS instead of waiting for subjective MOS from clients. Our proposition helps to predict MOS under specific network conditions. In order to achieve this goal, we start by training our algorithm to learn some scenarios. Then we change some quality parameters such as buffering time, resolution, bitrate, and number of frames per second before measuring Real_MOS. Our system will be then capable to estimate the Estimated_MOS with only network parameters in order to prove, thanks to objective measurements (VQM, SSIM, PSNR), that quality conformed to that expressed by MOS. We have implemented this application. First, we have implemented a SDN network with Mininet and VLC server to deliver video streaming with high quality. We developed a rating web application to collect both Real_MOS and network parameter QoS from the simulated SDN network.

We prepared a big dataset with many scenarios involving teachers, students, and administrative staff of our school for both in-motion and steady videos with a duration of 2 min. Within the dataset and scenarios, we implement a regression ML algorithm to estimate MOS without waiting for client ratings. The algorithm is implemented in a floodlight controller, which was made to be a smart one: in each video streaming scenario, we search for network QoS (RTT, jitter, bandwidth, and delay). Then we proceed to some real-time changes in video parameters (resolution, frames per second, bitrate, etc.) to observe the impact of some of them on the video quality (SSIM, VQM, and PSNR). The proposed ML algorithm is trained to estimate the perceived quality based on network parameters. The training data were acquired from a subjective video web application called the rating web application, where participants watched and assessed short RTP video sequences (2-min duration), during which they were subjected to varying network conditions. About 100 training and 50 validation samples (25 in-motion videos and 25 steady ones) were used in order to teach our ML algorithm the relation between a degraded RTP video and expressed MOS. The used parameters are video resolution, packet loss, bandwidth, buffering time, duration, and frequency. The resulting ML algorithm has good performance when the correlation coefficient has the value of 0.8. Finally, to be sure that the quality is correct, we also implement a monitoring system that saves both the original and received video sequences and calculates objective quality parameters such as SSIM, VQM, and PSNR. Those operations aim to ensure that quality is good under the QoE system and the implemented monitoring system.

Within this work, video quality is first expressed subjectively thanks to a rating web application and scenarios. In the next step, we prove that those subjective measurements can be replaced by objective ones, thanks to parameters such as SSIM, PSNR, and VQM. Our monitoring algorithm is conducted in a way that modifies video parameters (number of frames per second, bitrate, resolution) in the function of Estimated_MOS and QoE parameters such as RTT, jitter, bandwidth, buffering, and delay. To ensure the effectiveness of our solution, a performance study based on some scenarios was realized to prove that our ML algorithm as well as our Python monitoring algorithm have good performance in terms of perceived video quality. Our algorithm and web application present a tool to estimate MOS, perform objective and subjective study of MOS, and monitor QoE in SDN networks for video streaming services. Tests and scenarios are done to prove that objective study conducted with the machine learning algorithm helps in enhancing and monitoring QoE in adaptive and real-time video services. Additionally, we realize different tests and changes in both network parameters (bandwidth, delay, jitter, etc.) and video parameters (resolution, number of frames per second, bitrate, etc.) in order to demonstrate with graphs and curves the impact of those modifications on MOS and verify objective parameters (SSIM, VQM, and PSNR). Our solution is based on entering network conditions, video parameters, and scenario characterizations to give the best quality video as an output. The result represents the quality the users are asking for and that can be performed by the network under such conditions.

Our solution is integrated in the Mininet environment. As a future work, it will be implemented in a SDN controller. In this work, we describe the impact of network parameters on MOS and we describe the ML approach to help correctly obtain Estimated_MOS. The monitoring approach will be detailed in another work.

Figure 1 details the architecture of our proposed approach. It deals with: (1) A first module that plays video on VLC over Mininet to emulate a video streaming service over a SDN network. (2) A second module that collects rates from the client: it is a rating web application that saves original video, perceived video, expressed MOS, network parameters, and objective parameters. (3) A third module that estimates MOS from a dataset based on a regression machine learning algorithm. (4) A fourth module that changes video parameters in real time thanks to estimated MOS in order to deliver the best quality for clients. This module is called the monitoring module and it is based on objective parameters.

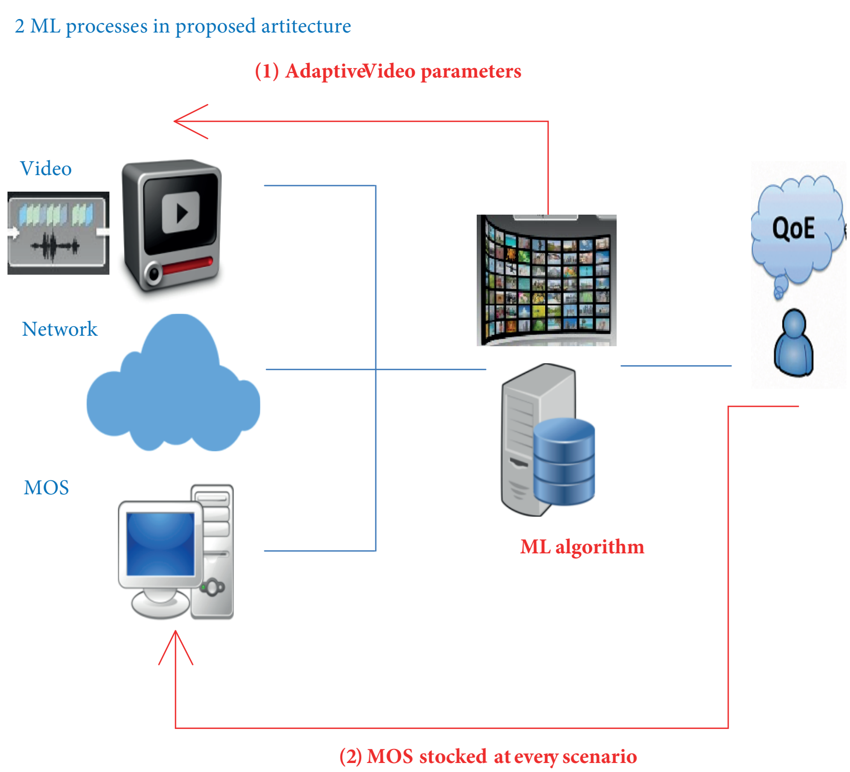


Figure 1. Proposed architecture.

5. Testbed

5.1. Problem

The aim of this work is to automatically correct multimedia settings or networks to ensure the best quality for video streaming. The work is divided into different tasks. First, we develop a rating web application that allows users to watch video streaming and rate it in order to reflect the opinion of the user regarding the quality of the video. The second goal is to automatically improve the video quality. Consequently, it is necessary to predict the user's opinion before he expresses his personal feelings towards the video. In fact, the second mission is to predict the MOS that reflects the subjective opinion of the user. Thus, with the prediction of MOS, we will know how to act and what parameters should be corrected to guarantee good video quality.

5.2. Video sequence

We consider 50 video sequences divided into two lots. The first one is composed of steady video while the second one is composed of 25 in-motion videos. We save them in a database in order to ask users to rate the quality of the Mininet environment when watching the 2-min video sequences. We choose Mininet to emulate a SDN network and VLC to emulate the video streaming server.

5.3. Rating web application and obtained dataset

We start by developing the rating application. We develop a live video streaming web application using NodeJS/AngularJS in order to make rating easier for users. They watch the stream directly in their browsers and rate the video quality instantly. Consequently, the user input will be automatically saved to the MongoDB database using RESTful API. We develop a program that stocks the original video as well as the received one after SDN network transfer and calculate the SSIM and PSNR parameters for every watched video. In order to do so, we use ffmpeg, which allows us also to calculate SSIM/PSNR. ffmpeg logs give several relevant pieces of information about the stream that is being played by the browser, like frames per second (fps).

This part of the work aims to predict the MOS parameter based on a dataset composed of network and multimedia parameters. To achieve the MOS prediction, we conducted research on the existing prediction algorithms. As a result, we selected the linear regression algorithm from the library 'Pandas' to use in our work. Below, we detail the steps used to predict the MOS. We start by presenting the different scenarios that we deployed to implement our algorithm of machine learning. We also present the results of these scenarios to realize the different measurements.

Data exploration: The dataset is an Excel table that contains different features and parameters that vary in function of the MOS feature. The dataset contains five different features. Each one of them is measured in function of the variation of the MOS. The first issue is that the data are not classified. We will find five different tables that contain some features and eventually ignore others. Consequently, we have to collect these tables to create our proper training set. In addition, another issue was noted involving the different sizes of the data. As a result, we have managed our data to be clean by making all necessary transformations and not affecting the prediction by losing the value of data at the same time. To avoid this problem, we add to each column the average of all the other values of parameters that are missing. Finally, we make a last transformation to improve data flexibility when executing our algorithm and making the measurements as well as the analysis. Data type is also modified to be a "csv" file.

Data visualization: After the cleaning process, data appear as presented in Figure 2. We can operate also by extracting some information about these data such as the minimum, maximum, and average of each feature. This will help us in deciding about the values of our parameters.

Resolution	MOS	VQM	SSIM	PSNR	Fame rate (fps)	Bit rate (kbps)
144	3	6,37412	0,76113	18,40545	34	513
240	5	7,10124	0,77893	19,36474	34	513
360	5	7,75508	0,72267	17,5815	34	513
144	4	4,24703	0,77706	17,77861	34	513
240	5	5,05528	0,77142	17,82767	34	513
360	5	6,12501	0,7729	17,76487	34	513
144	1	3,60457	0,90019	24,26753	34	513
240	1	3,52271	0,90351	24,92481	34	513
360	1	3,46201	0,90589	25,27061	34	513
248	2	1,10186	0,56327	12,4831	12	513
248	2	1,0482	0,58659	12,85736	30	513
248	2	9,49479	0,62886	13,62288	60	513
248	1	2,1568	0,96486	28,70624	12	513
248	1	2,34901	0,96559	28,71559	30	513
248	1	1,66439	0,97552	30,64262	60	513
248	3	8,7833	0,71542	14,9366	12	513

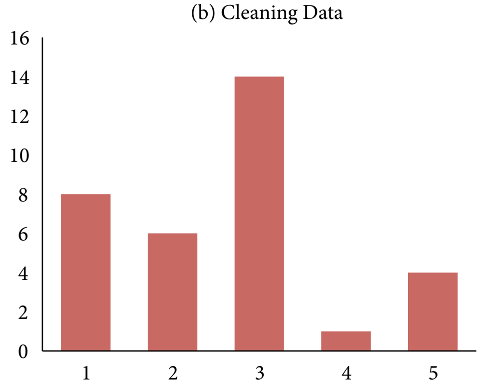
(a) Dataset

Resolution	MOS	VQM	SSIM	PSNR	Fame rate (fps)	Bit rate (kbps)
144	3	6,37412 e+00	0,76113	18,40545	34	513
240	5	7,10124 e+00	0,77893	19,36474	34	513
360	5	7,75508 e+00	0,72267	17,5815	34	513
144	4	4,247030 e+00	0,77706	17,77861	34	513
240	5	5,05528 e+06	0,77142	17,82767	34	513
360	5	6,12501 e+00	0,7729	17,76487	34	513
144	1	3,60457 e+00	0,90019	24,26753	34	513
240	1	3,52271 e+00	0,90351	24,92481	34	513
360	1	3,46201 e+00	0,90589	25,27061	34	513
248	2	1,10186 e+01	0,56327	12,4831	12	513
248	2	1,0482 e+01	0,58659	12,85736	30	513
248	2	9,49479 e+00	0,62886	13,62288	60	513
248	1	2,1568 e+00	0,96486	28,70624	12	513
248	1	2,34901 e+00	0,96559	28,71559	30	513
248	1	1,66439 e+00	0,97552	30,64262	60	513
248	3	8,7833 e+00	0,71542	14,9366	12	513
248	2	6,79065 e+00	0,78268	16,91876	30	
248	3	6,785120 e+00	0,78302	16,91377	60	
248	1	9,649900 e+00	0,7606	13,099282	12	
248	1	1,016523 e+01	0,74419	12,63443	30	

(b) Cleaning Data

	Resolution	MOS	VQM	SSIM	PSNR	Fame rate (fps)	Bit rate (kbps)
count	33	33	3,30000 e+01	33	33	33	33
mean	248	2,606061	1,531968 e+05	0,774685	18,970626	34	513;121212
std	46,861498	1,248484	8,80012 e+05	0,148265	5,310714	12,124356	193,000932
min	144	1	1,664390 e+00	0,40192	11,03869	12	128
25%	248	2	3,725150 e+00	0,72267	13,85986	34	513
50%	248	3	6,374120 e+00	0,78063	17,963	34	513
75%	248	3	9,494790 e+00	0,90262	22,97866	34	513
max	360	5	5,055286 e+06	0,97552	30,64262	60	900

(c) Result



(d) Predicted MOS

Figure 2. Data process.

Cross-validation: The goal of cross-validation is to avoid some prediction problems like overfitting. We split our data files into 3 sections: training dataset, testing dataset, and validation dataset, which is composed of a file of unknown data (or first-seen data) by the algorithm. The machine learning algorithm implemented for our scenario is the linear regression algorithm with the use of some modules of Python such as numpy, pandas, and Sklearn, which represents the main library for data science problems. The use of the linear regression model in our learned model allows us to reach a performance value of 0.61 for the algorithm.

As mentioned above, our goal is to define a linear model to predict the variable MOS as a function of many parameters. To do so, we choose in each case some parameters and observe the results of our algorithm. The results will contain the predicted MOS and the coefficients for the prediction's equation.

Case 1 ["PSNR", "frame rate (fps)", "Bit rate (kbps)"]

The algorithm's coefficients are [–0.06591784, 0.00839124, 0.00036007]

Case 2 ["resolution", "VQM", "SSIM", "PSNR" , "frame rate (fps)" , "Bit rate (kbps)"]

The coefficients are [4.65297438e-03 4.87046509e-07 –1.15165634e+00 –3.49111488e-02 8.44779387e-03 3.87424665e-04]

We will analyze the two cases by observing the results of the linear regression algorithm. For the first case, the predicted MOS is 2.8/3, which is a decent prediction. However, in this case, we just predict the feature in function of three parameters that have a significant value with a small difference in the value of each of two different parameters. This allows the algorithm to easily find a linear equation to estimate the predicted variable. For the second case, we are predicting the MOS without neglecting any parameter. We can estimate the result before running our algorithm because of the huge difference eventually between the bitrate, frame rate, and the others parameters. However, considering the average of these values, we have a predicted MOS of 2.56/3. To overcome this problem, the algorithm predicts coefficients to reduce this difference.

6. Performance analysis

6.1. Monitoring and enhancement QoE process in SDN networks

The aim of this part of our work is to conduct a subjective user study to investigate the human responses to the combined effect of video compression, initial buffering, and stalling. The goal is to understand how the network QoS affects the QoE at the end-user side. The study will be conducted to propose a model for QoE management in the case of video streaming service in the SDN context as is the case in Figure 3. At the end, we should obtain graphs presenting the effect of each parameter on MOS. QoE monitoring consists of: (1) Considering a dataset of rated videos (subjective MOS) with different parameters (PSNR, VQM, SSIM, resolution, bit rate, frame rate). (2) Calculating MOS by using a machine learning algorithm from a set of parameters. (3) If the MOS is lower than a certain threshold, we react by changing parameters (resolution, bit rate, frame rate) and recalculate the MOS. (4) Testing our algorithm by deploying a set of scenarios. (5) Analyzing the effect of each parameter on the variation of MOS.

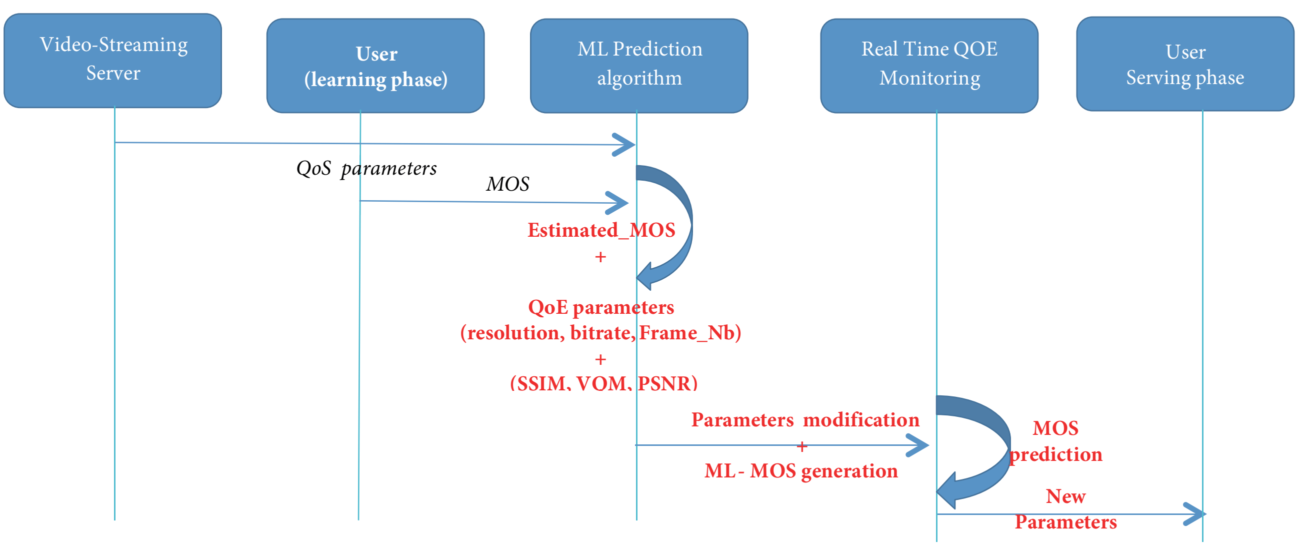


Figure 3. QoE monitoring process.

For our monitoring algorithm, the inputs are (1) a dataset of rated videos, (2) a linear regression algorithm that calculates the predicted MOS, (3) video parameters (resolution, VQM, SSIM, PSNR, frame rate, bit rate). We prepared more than one scenario, but we describe below only one.

6.2. Impact of buffering on quality of experience

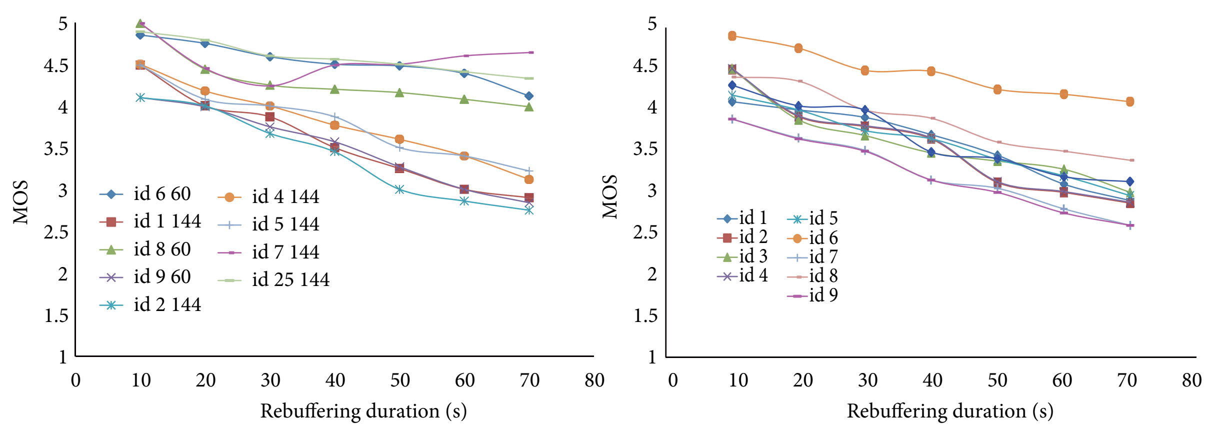
This part of the document addresses the influence of network buffers on network performance by controlling the frequency, the duration, and the time of buffering in both steady and in-motion videos. During the test, we will be able to visualize the quality of streaming while varying those three factors related to interruption, which are the duration, the number, and the position of interruptions in the video. We show below the buffering profile used for experience: Content: Steady-News; In-Motion-Trailer Film; Interruptions initial buffering duration Tinit: 5, 10 s; Rebuffering number of events Nbuff: 1, 2, 3; Rebuffering duration of events Trb: 10, 20, 30, 40, 50, 60, 70 s.

After presenting the buffering profile that we are going to use in these buffering tests, we design only 20 test videos (10 steady and 10 in-motion) for subjective tests. Each of them is approximately 2 min long. In each video, we only vary one factor and estimate the MOS attributed by each user. To realize the buffering tests, we use the VLC integrated in Mininet and we control the video streaming from the server side. This method allows us to control the interruption time. From the other side, the users will notify the server of their perceived QoE that corresponds to their subjective opinions.

6.3. Impact of buffering interruption on streaming quality

Rebuffering duration: In this part, we are going to modify the interruption duration in three different scenarios explained depending on the type of video: steady or in-motion. The first scenario is based only on one interruption located in the middle of the video. The duration of interruption will be varying from 10 s to 70 s of delay.

In motion: Figure 4 shows the impact of rebuffering duration on the video quality perceived by the viewer. This quality decreases as the rebuffering duration increases. It can be noticed that even the video length has negatively affected the rating. This is explained by the fact that viewers cannot tolerate rebuffering for short video durations. To reinforce our point of view, we observe in the graph that the quality of Id 9 is more deteriorated than others to reach 2.6 as MOS. We can then conclude that the MOS decreases while increasing the interruption duration in the middle of video.



(a) In motion videos (b) Steady videos

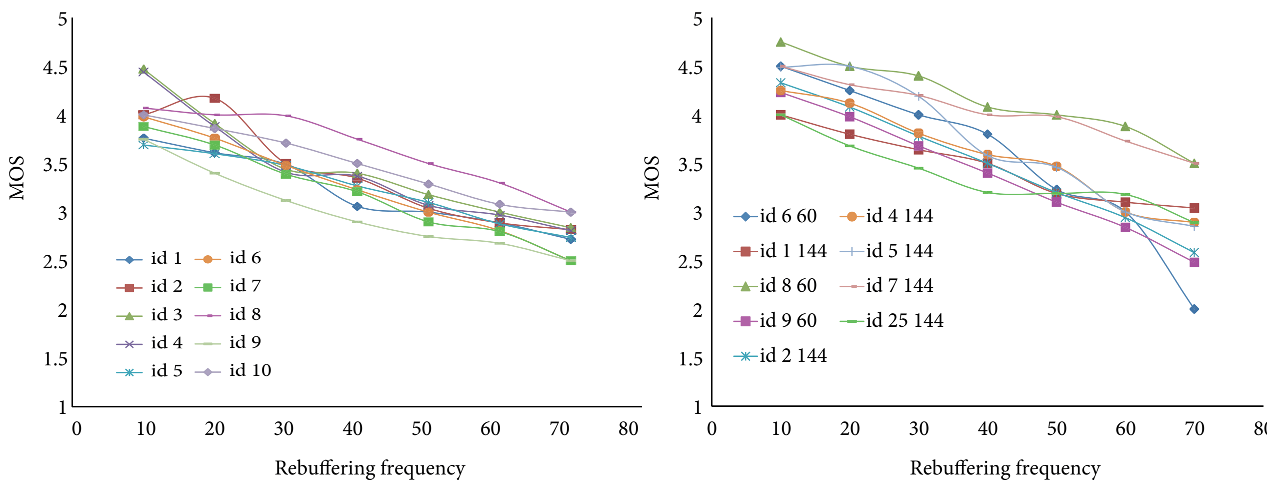
Figure 4. Impact of rebuffering duration (scenario 1).

Steady video: The evaluation of 10 steady videos is given in Figure 4. We notice that the quality of steady videos decreased, as was the case for the in-motion videos, but the difference appeared in the value of

MOS. The viewer can tolerate the interruption thanks to the specification of steadiness (there are not a lot of movements during the streaming).

In scenario 2, we will evaluate the MOS while interrupting the video streaming two times. The first interruption will be 10 s after starting the video and the second will be 20 s before the end of the video. However, the sum of both interruptions will be similar to the interruption in scenario 1, which means 10 s to 70 s of delay.

In-motion video: During the 10 in-motion video streaming tests, we succeed to obtain the average MOS of viewer ratings. As a result, we obtain Figure 5, which represents the percentage of viewer satisfaction. We can deduct from Figure 5 that the MOS value is less than for scenario 1. The viewer, in each case, prefers only one interruption even if the duration is greater than two interruptions localized in different locations.



(a) In motion videos (b) In Steady videos

Figure 5. Impact of rebuffering duration (scenario 2).

Steady video: The evaluation of 10 steady videos' results is presented in Figure 5. We notice that the MOS attributed to the quality of steady videos decreased as a function of interruption duration. Even in this type of video, the MOS degraded because of interruptions' increasing number, which affects the video and causes a bad rating by users.

Rebuffering frequency: In this part, we are going to modify the frequency of interruption. We collected the evaluations of 6 viewers while interrupting the video for one, two, and then 3 times following the three scenarios of Section 2.1. We also fixed the rebuffering duration to 20 s and we perform the tests on in-motion and steady videos.

In motion: The total rebuffering duration is constant and we vary its frequency. We evaluate 10 in-motion videos to obtain MOS. Figure 6 illustrates the fluctuation of the MOS as a function of the frequency of interruption. We found that the quality decreases more in the case of three rebuffering events. As we can observe, for example, for video id 3, MOS was rated 4 for one interruption and was reduced to 3.2 for 2 rebuffering events until reaching 2.9 for 3 rebuffering attempts. As expected, when the number of rebufferings increases, the total quality will be reduced and converge to 2. The appearance of the first interruption lowers the quality experience considerably. Consecutive interruptions decrease the quality further by a different amount. Even here, we notice that the duration of the video influences the satisfaction of viewers. Steady: We can deduct that the quality decreases as is the case for in-motion videos. However, the value of MOS is higher. For one interruption, the MOS is more than 3.5. It means that users are more tolerant with steady video. Even if the frequency goes to 3, the quality converges to 2.5 compared to 2 for in-motion videos. We clearly notice in Figure 6 the decrease of satisfaction while increasing the number of interruptions in one steady video.

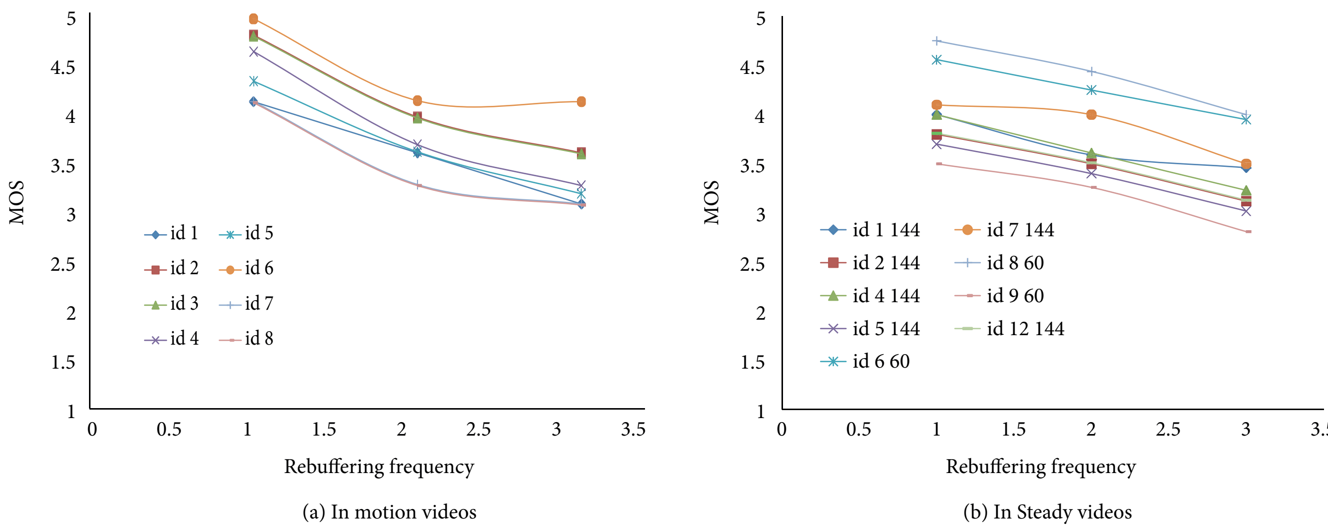


Figure 6. Impact of rebuffering duration (scenario 3).

Rebuffering duration and frequency: Graphs below illustrate the impact of duration and frequency buffering at the same time on in-motion video. Figure 7 allows us to confirm that rebuffering has a great impact on quality. The most influential factors are the rebuffering length and rebuffering frequency. It is better to have a single rebuffering independently than repeating the buffering events. Interruptions for steady videos cause less irritation than those for in-motions videos.

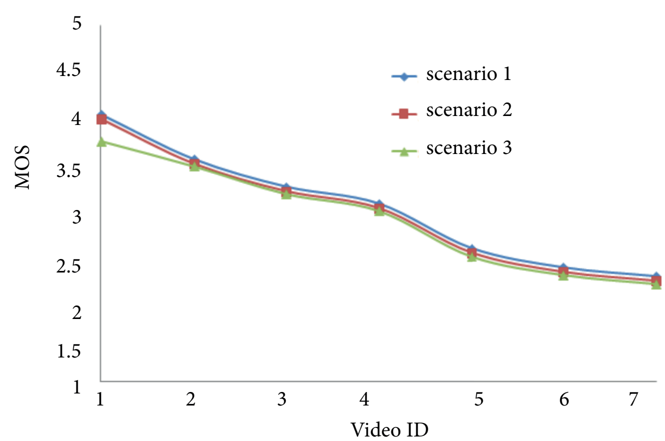


Figure 7. Impact of rebuffering frequency and duration on the quality of in motion.

7. Conclusion, perspectives, and future work

In this paper, we introduced a QoE monitoring approach based on real-time estimation of the QoE observed by end-users. The current work tries to use the ML approach to predict user QoE over SDN. It presents a study that collects MOS scores from users under varying network parameters as well as objective parameters such as SSIM, VQM, and PSNR. The MOS scores are collected by replaying videos to actual users in the SDN environment. The work presents the design of an architecture that could use the measured MOS values under varying network conditions to predict the expected MOS based on machine learning algorithms. The experimental results show that our proposed approach leads to good quality and better user satisfaction in terms of objective measurements. For the future, we plan to integrate load balancing in Mininet. In addition to video quality real-time adaptation, we suggest applying a load-balancing approach. Our ongoing work will focus on the multipath scenario of streaming video.

References

[1] Bentaleb A, Begen A, Zimmermann R, Harous S. SDNHAS. An SDN-enabled architecture to optimize QoE in HTTP adaptive streaming. IEEE T Multimedia 2017; 19: 2136-2151.

[2] Eckert M, Knoll T. QoE management framework for Internet services in SDN enabled mobile networks. In: 19th Open European Summer School; 2013; Chemnitz, Germany. pp. 112-123.

[3] Moller S, Chan W, Cote N, Falk H, Raake A, Waltermann M. Speech quality estimation: models and trends. IEEE Signal Process Mag 2011; 28: 18-28.

[4] Avcibas I, Sankur B, Sayood K. Statistical evaluation of image quality measures. J Electron Imaging 2002; 11:206-223.

[5] Laghari K, Crespi N, Connelly K. Toward total quality of experience: a QoE model in a communication ecosystem. IEEE Commun Mag 2012; 50: 58-65.

[6] Chikkerur S, Sundaram V, Reisslein M, Karam L. Objective video quality assessment methods: a classification, review, and performance comparison. IEEE T Broadcast 2011; 57: 165-182.

[7] Chang G, Ning W, Gerry F, Wilson M. Toward QoE-assured 4K video-on-demand delivery through mobile edge virtualization with adaptive prefetching. IEEE T Multimedia 2017; 19: 2222-2237.

[8] Al-Zubaidy H, Fodor V, Gyrgy D, Flierl M. Reliable video streaming with strict playout deadline in multihop wireless networks. IEEE T Multimedia 2017; 19: 2238-2251.

[9] Pinson M, Wolf S. A new standardized method for objectively measuring video quality. IEEE T Broadcast 2004; 50: 312-322.

[10] Lambrecht C. Quality assessment of motion rendition in video coding. IEEE T Circuits Syst Video Technol 1999; 9: 766-782.

[11] Ong E, Lin W, Zhongkang L, Yao S, Etoh M. Visual distortion assessment with emphasis on spatially transitional regions. IEEE T Circuits Syst Video Technol 2004; 14: 559-566.

[12] Chen-Chi W, Kuan-Ta C, Yu-Chun C, Chin-Laung L. Crowdsourcing multimedia QoE evaluation: a trusted framework. IEEE T Multimedia 2013; 15: 1121-1137.

[13] Staelens N, Moens S, Van den Broeck W, Marien I, Vermeulen B. Assessing quality of experience of IPTV and video on demand services in real-life environments. IEEE T Broadcast 2010; 56: 458-466.

[14] Brooks P, Hestnes B. User measures of quality of experience: why being objective and quantitative is important. IEEE Netw 2010; 24: 8-13.

[15] Moor D, Ketyko I, Joseph W, Deryckere T, Marez De L, Martens L, Verleye G. Proposed framework for evaluating quality of experience in a mobile, testbed-oriented living lab setting. Mobile Netw Appl 2010; 15: 378-391.

[16] Myers M. Predicting and measuring quality of service for mobile multimedia. In: IEEE 2000 International Symposium on Personal Indoor and Mobile Radio Communications; 18–21 September 2000; London, UK. New York, NY, USA: IEEE. pp. 1032-1036.

[17]] Wang Z. Study streaming video quality: from an application point of view. In: ACM 2003 International Conference on Multimedia; 4–6 November 2003; Berkeley, CA, USA. pp. 327-330.

[18] Winkler S, Dufaux F. Video quality evaluation for mobile application. In: SPIE/IS &T, Visual Communication and Imaging Processing; 2003; Lugano, Switzerland. pp. 593-603.

[19] Wang Z, Lu L, Bovik A. Video quality assessment based on structural distortion measurement. Signal Process Image Comm 2004; 19: 121-132.

[20] Masry M, Hemami S, Sermadevi Y. A scalable wavelet-based video distortion metric and applications. IEEE T Circuits Syst Video Technol 2006; 16: 260-273.

[21] De Pessemier T, De Moor K, Joseph W, De Marez L, Martens L. Quantifying the influence of rebuffering interruptions on the user's quality of experience during mobile video watching. IEEE T Broadcast 2013; 59: 47-61.

[22] Piamrat K, Viho C, Bonnin M, Ksentini A. Quality of experience measurements for video streaming over wireless networks. In: Sixth International Conference on Information Technology: New Generations; 2009. New York, NY, USA: IEEE. pp. 1184-1189.

[23] Varga P, Kun G, Sey G, Moldovan I, Gelencser P. Correlating user perception and measurable network properties: experimenting with QoE. Lect Notes Comput Sc 2006; 4268: 218-221.

[24] Charonyktakis P, Plakia M, Tsamardinos I, Papadopouli M. On user-centric modular QoE prediction for VoIP based on machine-learning algorithms, IEEE T Mobile Comput 2016; 15: 1443-1456.

[25] Mushtaq M, Augustin B, Mellouk A. Empirical study based on machine learning approach to assess the QoS/QoE correlation. In: 17th European Conference on Networks and Optical Communications; 20–22 June 2012; Vilanova ila Geltru, Spain. New York, NY, USA: IEEE. pp. 1-7.

[26] Menkovski V, Exarchakos G, Liotta A. Machine learning approach for quality of experience aware networks. In: IEEE 2010 International Conference on Intelligent Networking and Collaborative Systems; 24–26 November 2010; Thessaloniki, Greece. New York, NY, USA: IEEE. pp. 461-466.

[27] Demirbilek E, Gregoire JC. Towards reduced reference parametric models for estimating audiovisual quality in multimedia services. In: IEEE 2016 International Conference on Communications; 22–27 May 2016; Kuala Lumpur, Malaysia. New York, NY, USA: IEEE. pp. 1-5.