2014 Ninth International Conference on Broadband and Wireless Computing, Communication and Applications

Research on the user behavior-based QoE evaluation method for HTTP mobile streaming

Yajun Huang, Wen'an Zhou
Institute of Sensing Technology and Business, BUPT
Beijing University of Posts and Telecommunications
Beijing, China

Yu Du
Beijing University of Posts and Telecommunications
Beijing, China

Abstract—With the development of mobile network technology, HTTP mobile steaming is deployed by most of the video websites. Thus a lot of research works have been done on the Quality of experience (QoE) evaluation for HTTP mobile streaming. However, in most of them, only the factors in the network layer was considered, and the important information about user dissatisfaction with the video quality may be ignored. In this paper, we have made a survey to analyze the relationship between user satisfaction and user behaviors while a HTTP video is viewed, and we find that user behaviors will affect their QoE evaluation. Base on the findings, we propose a novel QoE evaluation method for HTTP mobile streaming which is based on user behaviors. The method combines user QoE with user expectations and moods inferred by user behaviors. In addition, not only the video fluency but also video clarity is considered in our work.

Keywords—HTTP mobile streaming; user behavior; QoE evaluation

I. INTRODUCTION

With the rapid development of wireless network technology, the mobile Internet has attracted widespread attention, and the types of mobile Internet business have begun rich and developed rapidly. Quality of experience (QoE) is an end to end concept, which describes the subjective feeling of a user during the interaction with a service or an application. As a comprehensive consideration of factors in network level, business level and environmental level, QoE attracts more and more attention and it becomes the key factor which network operators and service providers take into consideration to win in the fierce competition. In addition, the openness, sharing, collaboration and innovative of mobile Internet, as well as the mobility, mediumize and diversity of mobile Internet services, bring new demands and challenges to QoE evaluation in mobile networks [1].

Nowadays, the demand of mobile streaming media services also grows explosively, and one of the core technologies in streaming media service is streaming protocol. currently, HTTP mobile steaming has been employed by most of the video websites. Different to the traditional streaming protocol, e.g. RTSP / RTP, the streaming based on HTTP protocol does not require a dedicated streaming media server, and it can be supported by just a standard HTTP server, so its system configuration is relatively simpler. What's more, its firewall transversal presents a very good performance. Therefore, HTTP mobile streaming technology will become the focus of the development of streaming technology gradually. What's more, the purpose of the development of HTTP Streaming technology is to enhance the user's satisfaction on the business, thus QoE becomes a core performance indicator to evaluate HTTP streaming.

The QoE is usually expressed using a Mean Opinion Score (MOS), which is quantified by a real number ranging from 1 to 5, and 1represents the worst QoE while 5 means the QoE is really excellent [12]. It can be obtained from subjective or objective assessment. The subjective assessment is the most direct way to assess users' QoE, but it is costly, time-consuming and complicate. The objective assessment is simple and convenient, but it ignores the subjective feeling of a user.

The remaining of this paper is organized as follows. Section II introduces and analyzes the current research works on QoE evaluation for HTTP streaming. Section III analyzes the relationship between the satisfaction and each user behavior during users watching HTTP video. Based on the analysis of user behaviors, a system for personalized QoE evaluation for HTTP mobile streaming is described in section IV. Finally, conclusions of this paper are presented in Section V.

II. RELATED WORKS

It is essential for QoE evaluation and management to research on the QoE factors. Previous works on the QoE of HTTP streaming mainly concern the network layer and the application layer [2]-[7]. The QoE factors can be concluded into three aspects: service domain, user domain and environment domain. Service domain includes parameters of application layer, network layer and service layer. User domain includes user expectations, user experience and user backgrounds. Environment domain includes natural, humanistic and social environment, service operation environment and some other environmental factors. Ricky K. P. Mok et al. [2] performed experiments to evaluate the correlation between application QoS and QoE, and that

978-1-4799-4173-5/14 $31.00 © 2014 IEEE
DOI 10.1109/BWCCA.2014.69

between network and application QoS respectively, and used radar chart to correlate the network QoS with QoE. However, the results were based on the assumptions of constant network and none interaction during the playback, which is a idealization and far away from users' real experience. In [3], Steven Latré et al. proposed an algorithm, which is deployed on an intermediary node in the access network, to predict the QoE of progressive downloads in real-time, and it is only depends on the TCP data and acknowledgement packets. It is used to predict when a play-out buffer starvation occurs and how long each play-out buffer starvation takes, but it can't evaluate the QoE value. In [9], a no-reference QoE monitoring module for adaptive HTTP streaming using TCP and the H.264 video codec was proposed, but the feedback wasn't taken into account. Similar to [9], reference [11] proposed a simple model for predicting the number of rebuffering events and their duration in progressive downloads from YouTube. Barbara Staehle et al. described a tool which constantly monitors the YouTube application comfort and anticipates an upcoming YouTube QoE degradation. Some works pointed out the user layer was essential for QoE, but it is utilized in few QoE evaluation methods. Vilas et al. modeled the user behavior of a VoD website [14], but they did not correlate the behavior with the QoE. Based on the results in [2], Ricky K. P. Mok et al. also proved that the video impairments could trigger pause and screen size switching events after two and three seconds [8], but they only analyzed the user-viewing activities in PC.

The QoE evaluation is subjective and individualized, and it is affected by many factors. During users viewing videos, they will trigger related behaviors, such as pause, switch video resolution and quit. Some of these behaviors may affect the quality of video service, and some may imply the certain psychology of users. However, the existing evaluation methods tended to idealize the scene, assuming that the network parameters are constant and the user does not interact with the video during the playback. Because QoE is a subjective feeling of users, if we evaluate it only by the objective data in network layer and application layer, it will be difficult to fit users' actual experience feelings accurately. In addition, the universal evaluation results may not apply to everyone.

For creating a method to evaluate QoE exactly for each user, we make a survey to analyze the relationship between user satisfaction and each user behavior during users watching HTTP video. Then based on the results of analysis, we design a system architecture of personalized QoE evaluation for HTTP mobile streaming.

III. ANALYSIS OF USER BEHAVIOR

User behaviors refer to the activities that a user interacts with a player interface during viewing mobile video, which includes pause, forward, backward, quit and switch video resolution. To research on a QoE evaluation method from user behaviors, we need to analyze the relationship between user behaviors and user experience at first.

We can divide user behaviors into historical behaviors and current behaviors. Historical behaviors refer to the statistical activities during user viewing certain type of videos before, and current behaviors refer to the statistical activities during user viewing video when the QoE of video is needed to be evaluated. We can infer user preferences to the fluency and the resolution of certain type of videos by analyzing historical behaviors, and infer user moods during the playback by analyzing current behaviors. We have made a survey on how users behave when the playback is smooth or jerky and how users feel when they trigger the activities. Then we further analyze every user behavior on the basis of the survey results as follows.

A. Pause

The pause behavior means that the user stops playing the video playback for a short period of time. We divide this behavior into active pause and passive pause. The active pause refers to the pause behavior when the playback is smooth, and the passive pause refers to the pause behavior when the playback is jerky. The former one is always caused by the user's own reasons, and not to do with the video quality, so it has nothing to do with QoE evaluation. The passive pause is always caused by the users' dissatisfaction on experience, so it may affect the QoE evaluation. In our survey, 68.97% of users will pause the video and wait for a short time to expect video play more fluently when video is in buffering state. If the video becomes more fluent after pause behavior is triggered, 48.28% of users' experience will not change. But 17.24% of users' experience will decline when they trigger the passive pause behavior, even though the video becomes more fluent. In addition, 72.41% of users will feel more tired, and their experience will further decline, when the video playback is still not smooth after passive pause or they are forced to trigger passive pause too much.

We can conclude that the passive pause behaviors reflect user expectations and user moods for the current video, which will affect the user experience.

B. Forward and Backward

The forward behavior means that the user will watch the content after the current video position, while the backward behavior means the user will watch the content before the current video position. These two behaviors can be analyzed in the following two cases. If the user wants to watch the content in the buffer, then the behavior will not affect the quality of video and the user's QoE. If the content the user wants to watch isn't in the buffer, the video will occur a buffering event. However, more than 50% of users express that they will accept the buffering event after their forward or backward behaviors and their mood won't be affected as long as the buffering time is short. Therefore, we can conclude that the forward and backward behaviors have just few influences on user experience and we can ignore them in the QoE evaluation.

C. Quit

The quit behavior means that the user stops the video service. Similar to the pause behavior, we divide this behavior into active quit and passive quit. The active quit refers to the quit behavior when the playback is smooth, and the passive quits refers to the quit behavior when the playback is jerky. The former one is always caused by the user's own reasons, and not to do with the video quality, so it has nothing to do with QoE evaluation. The passive quit is always caused by the reason that the user has been unable to tolerate these degrees of video impairments, so it may affect the QoE evaluation. Our survey results show that 82.76% of users will quit the video service because of the high rebuffering frequency, and 51.72% of users will quit the video service because of the long buffering duration time and the low video revolution. Therefore, the degrees of video impairments (i.e. rebuffering frequency, rebuffering duration and video revolution) every time the user triggers the passive quit behavior will reflect the user's tolerance for certain type of video, and the tolerance will affect the user's QoE evaluation.

D. Switch video resolution

The switch video resolution behavior means the user will watch the video with lower or better picture quality. It reflects the expectation for the clarity of the certain type of videos. That a user always switches some types of videos to a higher revolution means the user cares the revolution of the type of video a lot. Then if the user is forced to lower video revolution for the playback smoothness, its QoE will decline more than others. Out survey results show that 88.66% of users will reduce their QoE value when they have to lower video revolution for the playback smoothness. Thus we can infer users' video clarity expectation by their switch video resolution behaviors.

IV. QOE EVALUATION BASED ON USER BEHAVIOR

In summary above, user behaviors during video playback can reflect users' subjective experience, including user preferences, user expectations and user moods. In order to predict user satisfaction exactly, we can consider to record user behaviors during video playback and analyze them to infer user expectations and moods. Most research works on QoE of HTTP streaming media have shown that the impact of the network on QoE evaluation becomes more and more indirect and vague, and the network will be reflected in the application metrics, so we propose a method to combine information in user layer with that in application layer and obtain the personalized QoE evaluation, rather than a universal one.

A. Overview of the system architecture

Figure 1 shows the system architecture of the method. It includes the terminal domain and the server domain. In the terminal domain, there are 3 modules namely 1)User behavior monitoring module (UBMM), 2)Video quality monitoring module (VQMM), 3)Data transmission module (DTM). In the server domain, there are 4 modules namely 1)Data receiving module (DRM), 2)Initial QoE evaluation module (IEM), 3)User behavior analysis module (UBAM), 4)Final QoE evaluation module (FEM).

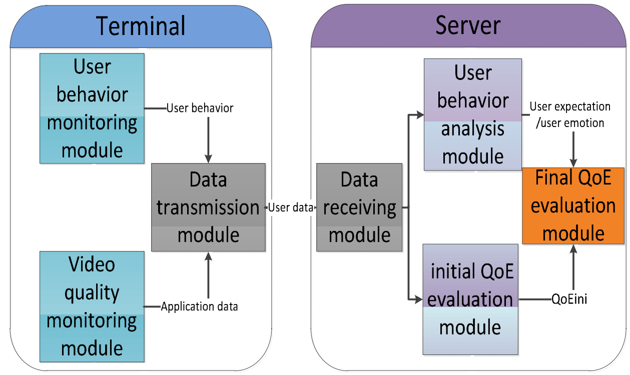

Fig. 1. System architecture

B. UBMM & BQMM

UBMM is responsible to monitor and record user behaviors during a user viewing video, which include pause, quit and switch video resolution behaviors. Once pause behavior or quit behavior is triggered during the playback, UBMM will notify BQMM to view the corresponding state of buffer. If the state of buffer is good, the behavior will be identified as active pause or active quit, otherwise, it will be identified as passive pause or passive quit and recorded. The information recorded will be send to the server via DTM.

BQMM is responsible to collect the application information, which includes the video type, initial buffering time, rebuffering duration time and rebuffering frequency. The video type and the initial buffering time will be recorded when a video is started. Then BQMM will collect the rebuffering time and the length of video playback every time the playback is jerky. The information collected by BQMM will be send to the server via DTM.

C. DTM & DRM

DTM is responsible to integrate data collected from UBMM and BQMM, and send them to the server.

DRM is responsible to receive data from the terminal and save them into a database. The fields in the database are (uID,C_i,user_act,time,T_{ini},T_{rebuf},F_{rebuf}).

- "uID" is a unique identifier of a user, which is used to identify each user by the IMEI number of a terminal, a user's ID or other ID.
- "C_i" (i.e. C_1, C_2,…, C_n) is the type of video which is playing, and it can be classified by the contents (e.g. sports, entertainment, news.) or it can be classified into "slow movement", "general walking", "rapid movement" based on the spatial and temporal feature extractions[13].
- "user_act" (i.e., "pause", "quit", "pic_in", "pic_de") is the type of user behavior, and the content included in it respectively represent "pause", "exit", "increasing the resolution" and "reduced resolution".

- "time" refers to the video time when the behavior is triggered.
- T_{ini}, T_{rebuf}, F_{rebuf} indicate the overall video impairment when the behavior is triggered, and they represent initial buffering time, mean rebuffering duration and rebuffering frequency respectively.

D. IEM

IEM is responsible for initial QoE evaluation according to the data in the application layer collected by terminal. Referring to [2], it divides initial buffering time, mean rebuffering duration and rebuffering frequency into three levels, with "1" "2" and "3" to represent "Low" "Medium" and "High" levels respectively, and then the initial QoE is calculated, as shown in (1) .

$$QoE_{inti}=4.23-0.0672L_{ti}-0.742L_{fr}-0.106L_{tr} \quad (1)\ [2]$$

Where L_{ini}, L_{fr} and L_{tr} are the respective levels of T_{ini}, and F_{rebuf} T_{rebuf}. The corresponding levels of application performance are shown in Table I.

TABLE I. THREE LEVELS OF APPLICATION PERFORMANCE[2]

Level	APMs		
	T_{ini}	F_{rebuf}	T_{rebuf}
Low (1)	0 -1 s	0-0.02	0-5s
Medium (2)	1-5 s	0.02-0.15	5-10 s
High (3)	>5 s	>0.15	>10 s

E. UBAM

As the key module of the system, UBAM is responsible to analyze user behaviors including historical behaviors and current behaviors to infer user preferences, user expectations and user moods during a user viewing video. Different to the video service based on UDP protocol, the main reason for the quality degradation of HTTP streaming video is that the arrival of packet is too late to fill the empty buffer. Thus the factor which affects the quality of user experience most is the video fluency. However, even though there is no mosaic phenomenon in HTTP video, the clarity of video affects user experience. In our survey, about 74% of users have a particular requirement for their favorite videos. For a more comprehensive consideration of users' individualized experience, we need to analyze the impact of video fluency and clarity on users themselves experience according to their own behaviors during viewing videos, and then modify their personalize QoE evaluation.

We use E_{it} to represent the user experience impact value of the video fluency and i means the type of video is C_i. In the previous analysis of behavior, we have mentioned that users trigger the passive quit to express the current degree of video impairment is unbearable. Thus we can infer users' video impairment tolerance according to the historical passive quit. In addition, users always trigger the passive pause when the video playback is jerky, so we can infer users' mood to the video fluency according to the passive pause during playback. Above all, E_{it} can be obtained in (2)-(4):

$$E_{it} = e_1I_{it} + e_2M_t \quad (2)$$
$$I_{it} = I(L_{qti},L_{qtr},L_{qfr}) \quad (3)$$
$$M_t = M(N_{pause}) \quad (4)$$

Where I_{it} and M_t denote the influence value of quit and pause on QoE respectively.

In (3), L_{qti}, L_{qtr}, and L_{qfr} respectively represent the mean initial buffering time, mean rebuffering duration and mean rebuffering frequency of all videos of Ci which is viewed and passive quit by the user before. The lower degree of video impairment every time user conducted passive quit before means the higher user expectation on the type of video. The higher user expectation will means that the user evaluate its experience more strictly, that means its QoE value will decline in the same video impairment.

In (4), N_{pause} represents times of passive pause. Users will have higher expectations when they trigger passive pauses than those without operation. With the increasing times of passive pauses, users will become more tired and their QoE value will decline.

We use E_{ip} to represent the user experience impact value of the video clarity. Different users will have different video clarity expectations on different types of video, and too many times of switching to a lower revolution will make users tired. Thus we consider E_{ip} into QoE evaluation and it includes the influence value of user expectations and user emotion during playback, as shown in (5)-(7).

$$E_{ip} = n_1I_{ip}+ n_2M_p \quad (5)$$
$$I_{ip} = I(N_{pic\_in\_ci},N_{ci}) \quad (6)$$
$$M_p = M(N_{pic\_in},N_{pic\_de}) \quad (7)$$

Where I_{ip} and M_p denote the influence value of clarity expectations and moods caused by clarity on QoE respectively.

In (6), $N_{pic\_in\_ci}$ represents the number of C_i types of videos in which the user has switched to a higher revolution before, and N_{ci} represents the total number of C_i types of videos which the user has viewed before. If the frequency of a user switching to a higher revolution is very high, we can infer that the user has a high expectation on the video clarity, and then the user experience will become lower on the same impairment. If a user has never switched video revolution, we can infer that the user has no expectation on the video clarity and we will default that users has no expectation on the clarity in the first time of QoE evaluation.

In (7), N_{pic_in} represents the number of switching to a higher resolution times during a playback, and N_{pic_de} represents the number of switching to a lower resolution times. On the same video impairment, user experience value will decline along with the number of switching to a higher resolution times.

F. FEM

FEM is responsible to combine the initial QoE evaluated by IEM with the influence value of video fluency and clarity by UBAM, and obtain the personalized QoE evaluation for each user.

$$QoE_{final} = QoE_{init} + \mu f(E_t, E_p) \qquad (8)$$

As shown in (8), the final QoE is the correction of the initial universal QoE and it is the personalized QoE for each user. The specific functions above can be obtained by regression fitting of large amounts of data or by neural network model.

V. CONCLUSIONS

In this paper, we propose a novel QoE evaluation method for HTTP mobile streaming which is based on user behaviors. At first, we introduce and analyze the current QoE evaluation methods for HTTP streaming, then we conclude that the factors in user layer are very important but few methods utilize them. Therefore, we make a survey to analyze the relationship between user satisfaction and user behaviors while a HTTP video is viewed, and prove that user behaviors will reflect user expectations and moods. What's more, we describe a system of personalized QoE evaluation for HTTP mobile streaming. In the system, the influences of video fluency and revolution are inferred by recording and analyzing user behaviors during the video playback, and they are used to obtain the personalized QoE on the basis of initial universal QoE which is evaluated by the data in application layer. The novelty of this method is that it combines user QoE with user expectations and moods by user behaviors, in addition, it considers not only the video fluency but also video clarity. Our work will provide a reference to help service providers and network operators evaluate QoE of HTTP streaming appropriately.

ACKNOWLEDGMENT

This work was supported by the National 3rd Key Program project No.2012ZX03001033 and the National Nature Science Foundation of China(61372113).

REFERENCES

[1] LIN Chuang, HU Jie, and KONG Xiangzhen. Survey on Models and Evaluation of Quality of Experience[J]. CHINESE JOURNAL OF COMPUTERS. Vol. 35 No.1,Jan. 2012:1-15

[2] Ricky K. P. Mok, Edmond W. W. Chan, and Rocky K. C. Chang . Measuring the Quality of Experience of HTTP Video Streaming[C]. Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium,Dublin, 2011:485-492

[3] Steven Latré, Nicolas Staelens,Pieter Simoens. On-line estimation of the QoE of progressive download services in multimedia access networks[C]. ICOMP 2008, Las Vegas, Nevada, USA, July 14-17, 2008

[4] Florin Dobrian,Vyas Sekar ,Ion Stoica .Understanding the Impact of Video Quality on User Engagement，SIGCOMM'11 ,ACM,2011

[5] Ozgur Oyman, Ssrabjot Singh. Quality of Experience for HTTP Adaptive Streaming Services[J].IEEE Communications Magazine.April 2012,50(4):20-27

[6] 3GPP TS 26.247 v10.1.0, "Transparent End-to-End Packet Switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming Over HTTP (3GP-DASH)," Release 10, June 2011

[7] Vlado Menkovski ,Antonio Liotta.QoE for Mobile Streaming[J].Mobile Multimedia - User and Technology Perspectives,2012

[8] Ricky K. P. Mok, Edmond W. W. Chan, and Rocky K. C. Chang . Inferring the QoE of HTTP Video Streaming from User-Viewing Activities [C]. W-MUST'11,August 19, 2011:31-36

[9] Kamal Deep Singh, Yassine Hadjadj-Aoul and Gerardo Rubino. Quality of Experience estimation for adaptive HTTP/TCP video streaming using H.264/AVC[C]. The 9th Annual IEEE Consumer Communications and Networking Conference:127-131,2012

[10] B. Staehle, M. Hirth, F. Wamser, R. Pries, and D. Staehle, YoMo: A YouTube Application Comfort Monitoring Tool, University of W¨urzburg, Tech. Rep.467, March 2010.

[11] Pablo Ameigeiras, Alba Azcona-Rivas, Jorge Navarro-Ortiz, Juan J. Ramos-Mu˜noz, and Juan M. L´opez-Soler. A Simple Model for Predicting the Number and Duration of Rebuffering Events for YouTube Flows[J], IEEE COMMUNICATIONS LETTERS, 2012,2(16):278-280

[12] ITU-T Recommendation P.800. Methods for subjective determination of transmission quality, August 1996.

[13] Asiya Khan, Lingfen Sun and Emmanuel Ifeachor. Content Clustering Based Video Quality Prediction Model for MPEG4 Video Streaming over Wireless Network [C], Proceedings of IEEE International Conference on Communications, ICC 2009, Dresden, Germany. pp.1-5. 14-18 June 2009

[14] M. Vilas, X. Paneda, R. Garcia, D. Melendi, and V. Garcia. User behavior analysis of a video-on-demand service with a wide variety of subjects and lengths. In Proc. EUROMICRO, 2005.