A User Experience Study for Watching Delay Interrupted Video in the Context of Mobile Network

Hao Tan[1], Jiahao Sun[2(✉)], Bin Wang[3], Qiyong Zhao[3], Wei Li[2], and Zhengyu Tan[2]

[1] State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha, China
htan@hnu.edu.cn
[2] School of Design, Hunan University, Changsha, China
{sunjiahao2015,liwei2014}@hnu.edu.cn
[3] MBB Lab, Huawei Technologies Co. Ltd., Shanghai, China
{i.wangbin,zhaoqiyong}@huawei.com

Abstract. The development of mobile application in our daily lives has changed our way of receive various of information, the way people use mobile devices to receive information has become whenever and wherever based on the mobile network technique. End user experience in the context of Internet services is also depend on the waiting time in the beginning of service and the intervals during the service caused by the system delays. We use several factors buffer, stalling time) to investigate the relationship between the video QoE and delay. We used the results of these analyses to inform the construction of the user's satisfaction influenced by time delay. Given by the experiment results, we have conclusion some factors-based contexts which could influence the user's satisfaction and some changeable QoE-related trends based on those three factors.

Keywords: User experience · Delay · Video quality assessment · Mobile network service

1 Introduction

Nowadays, VoIP is very common in IP networks, and the important trend of video traffic has increased rapidly, meanwhile, the wide use of mobile devices with video display support and the application of wireless networks (WLANS and 3G/4G) have contributed to this scenario. In a few years, the contents related to videos transmitted through the network will be 90% (Begen et al. 2011; Cisco Systems Inc. 2013). To provide a satisfactory video service to users in real time is a certain challenge to the internet and multimedia providers.

The video streaming service is a typical internet service in which the video contents need to be encoded and stored in multimedia database, then transmitted to the customers through the core network, finally displayed on user's devices like a smartphone or a tablet PC. Based on this model of video streaming service, there are many factors such as bandwidth, delay, jitter, packet loss and so on will influence the quality of video contents which could not afford each customer. Moreover, when the customer uses

P.-L.P. Rau (Ed.): CCD 2017, LNCS 10281, pp. 724–735, 2017.
DOI: 10.1007/978-3-319-57931-3_59

mobile devices, the scenario here becomes more complicated and produces more difficulties, such as wireless signal coverage and wireless channel instability. such situation is one kind of system delays which makes users spend time on waiting for the system's response.

Also, in research field, such delays are commonly defined as system delays (Selvidge et al. 2002; Szameitat et al. 2009) or system response times (Dabrowski and Munson 2011; Schleifer and Amick 1989). System delays are caused by two parts, one is about the system itself, such as processing speed, network bandwidth or the computation on the complexity requests. Another one is about the transient factors, such as network congestion, background processes and so on (Seow 2008). From the research on Human Computer Interaction (HCI) in the past decades, system delays can evidently influence the users' experience (Ceaparu et al. 2004; Nah 2004; Thum et al. 1995).

2 Proposed Technique

2.1 QoE

The study on the quality of experience (QoE) has played an important role in the increasing popularity of various video services (Cisco Documentation 2014). Quality of Experience (QoE) is a concept commonly used to describe user's whole satisfaction reflecting the degree of delight or annoyance of a user with a (multimedia) system, service or application (Le Callet et al. 2012). In concise explanation, QoE is an assessment of the user satisfaction with the contents shown on different kinds of devices (Zepernick et al. 2011; Raake et al. 2014). Thus the network infrastructure used for the video streaming requires an evolution method to assess the video QoE. The ITU-T made a definition of QoE as an assessing system to evaluate the service quality based on users in 2007. There are many factors affect this quality measure system, including different kinds of technical and non-technical (Brooks and Hestnes 2010). These factors are related to service preparation, delivery and presentation, which makes challenge to maintaining QoE at an acceptable level. Many solutions have been introduced to tackle the challenge of video traffic quality. However, the realistic situations are required to meet the satisfaction of users and preserve the interest of service providers.

Given by the investigations above, delays lead user to wait for the system response, the waiting time here plays an important roll in user experience, although there have been many studies on how user react to delays before (Larson 1987; Clemmer and Schneider 1989; Dube et al. 1989, 1991; Hui and Tse 1996).

2.2 Delay

Given by system delays are caused by system itself and transient factors (Selvidge et al. 2002; Szameitat et al. 2009), users are influenced by these delays then lead to have bad experience or low performance (Ceaparu et al. 2004; Nah 2004; Thum et al. 1995).

Recently, many studies focus on the whether the negative effects of delays can be managed or avoided by interface design (Branaghan and Sanchez 2009; Galletta et al. 2006; Krejcar 2009). There are two elements that can effects user experience and performance during delays, the one is delays' lengths, the other one is delays' variability (Kuhmann 1989; Kuhmann et al. 1987; Schaefer 1990). It is an almost universal view that long waiting times are harmful to users' performance and experience which will have bad influence on users' overall satisfaction (Martin and Corl 1986; Schaefer 1990; Seow 2008; Simoens et al. 2011).

However, it should be noticed that only the impact of a single delay or short delay on the user have been considered in the former articles instead of multiple delay. In our study, we want to evaluate the users' experience of different kinds of delays occurred in video streaming. The main goal of this paper is to evaluate the quality of user experience in different delay context of mobile network. Therefore, delay factors are divided into initial buffer (video initial buffer time) and stalling time (video playing time of the stall) to allow users to evaluate. The experiment applies methods related to QoE in the assessment of uer satisfaction which influenced by delay factors. A typical process of video service quality evaluation is detailed and the assessment methods are divided into subjective and compared in the context of mobile network.

The study is devoted to concluded the standard of QoE which can help internet and multimedia providers plan and construct the internet, which core is video QoE, from the perspective of human factors and provide the best experience under wireless internet environment.

As the user experience is the primary productive forces, video QoE plays a very important role in improving the video service. With the trend of decreasing the traffic rate, the consumption of video traffic has become the key measure to enhance the revenue of the global internet and multimedia providers and the user experience of the video service has become the key to measure the quality of the internet and multimedia providers' network. Thus, the video QoE in our study can help them keep abreast of the user experience and adjust the network to improve quality of service. Moreover, with the consumer's subjective scores in emotional scales, a correlation model between different inquiring feedback delaying and the different emotional experience levels in online e-commerce platform was designed, which produced a trial guide of feedback delaying service design in online e-commerce platform and help to promote the online e-commerce service experience.

3 Implementation

There were lots of factors may impact to the video QoE. Baraković et al. were divided many factors into five factor categories: technological performance, usability, subjective evaluation, expectations, and context. Xue et al. reported the influence factors of the user perception also included screen size, viewing distance, lighting, and user movement when viewing videos on mobile devices. Thus, the laboratory equipment for an iPhone 6 Plus mobile phone (5.5 in., 1920 * 1080p), used to play the experimental video clips; A camera to record the process; The video subjective scoring table for experimental material used to the user to score for each experimental video clips; A video editing software to mimic the video buffer and stalling in the mobile network, allowing users to feel the real network delay. In order to test the time delay, we asked participants to perform the paradigm under one situation. We investigated the time delay during the video with the user experience. Further, we divided the time delay into the buffering time and stalling time and analyzed how these two factors influenced the video QoE.

4 Experiments

4.1 Participants

In this study, there were 38 participants (21 male/17 female) took part in. The age ranged between 18 and 55 (M = 21, SD = 10) years. They were experienced mobile device users and all had use network for 6 years on average. Prior to the study, all participants were required to accept the visual fatigue testing and fill in the individual basic information table. Finally, 30 participants (15 male/15 female) were qualified to join in this study. And 80% of the participants often use mobile device to watch the video.

4.2 Method

Thus, to assess for the experience quality of video we used a subjective modified approach. The method we use and the criterion the users relay on as a practical way to objectively measure the quality of video is MOS that is one of the subjective quality evaluate methods (ITU-T P.910 1999; ITU-R BT.500-10 2000; Jones and Atkinson 1998; Caramma et al. 1999; Watson and Sasse 1998) which is the most commonly method to assess the video quality.

We use the quality scale which included "bad", "poor", "fair", "good", and "excellent", and they are translated to the values 1, 2, 3, 4 and 5 (Fig. 1) when calculating the mean opinion score (ITU-T recommendation P-910). According the ITU-T recommendation BT 500-13, it reported that the MOS is a prevalent used metric. The observer selected a score from 1 to 5 in the quality scale to evaluate every video.

Quality scale.

5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Fig. 1. Quality scale

In this experiment, we had two tasks to research the video QoE and the time delay. In the first task, we investigated the time delay. Moreover, we compared the influence of the initial buffer time and the stalling time on the video QoE to users. In the task two, we researched that when the user watched 6 min video, other conditions in the same situation, whether playing long time shorter or increasing the stalling time have impact to the user QoE to watch video.

4.3 Tasks

Task 1
In this task, we experimented on an iphone6 plus. Please users to sit in the most comfortable position, holding the phone and adjust it to the most comfortable distance. Next, the participant was informed the experiment testing process and the purpose of the experiment.

In this experiment, we took several measures eliminate the user negative sentiment from the video clips content. Firstly, the participant selected one video that he or she was most interested in. Next, the participant selected `one type of video` about 20 s in length and clicked the play button. Next, there was a buffer of T seconds in every video. The participant then need to assess a MOS score from 5 to 1 in this experience. When the MOS score reached 1 score, all trials were finished by default (Fig. 2). The buffer time T included: 0.1 s, 0.5 s, 1 s, 2 s, 1 s, 3 s, 4 s, 5 s, 6 s, 7 s, 8 s, 9 s, 10 s (Table 1).

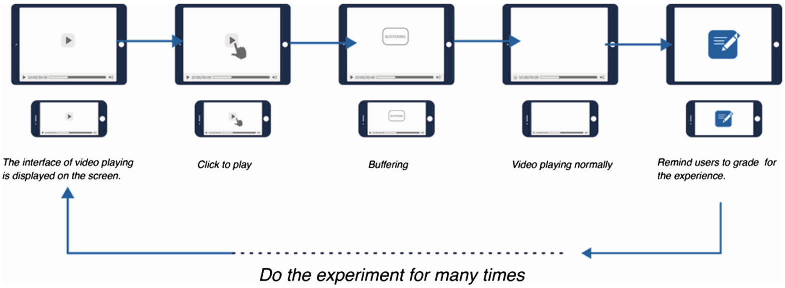

Fig. 2. The experimental process of buffer

Table 1. The video clip material about buffer.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
Buffer (s)	0.1 s	0.5 s	1 s	2 s	3 s	4 s	5 s	6 s	7 s	8 s	9 s	10 s

This step also selected one type video about 20 s in length primarily, then, the participant clicked the play button. The video clip started playing 5 s fluently, after the staling appearing T seconds. After stalling, the video clip playing continued. According this situation, the participant gave a MOS score to assess it (Fig. 3). The stalling time included (Table 2): 0.1 s, 0.5 s, 1 s, 2 s, 1 s, 3 s, 4 s, 5 s, 6 s, 7 s, 8 s, 9 s, 10 s.

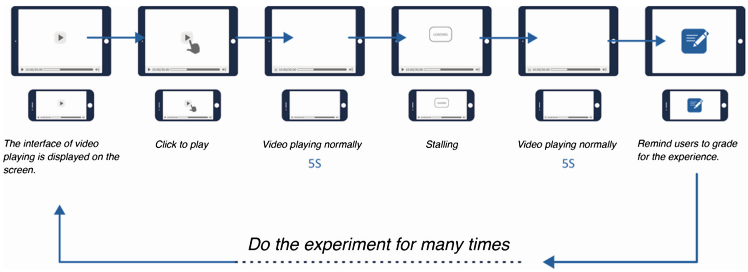

Fig. 3. The experimental process of stalling time

Table 2. The video clip material about stalling.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
Stalling (s)	0.1 s	0.5 s	1 s	2 s	3 s	4 s	5 s	6 s	7 s	8 s	9 s	10 s

Task 2

In this task, we experimented on an iphone6 plus. Ask users in the most comfortable sitting position holding the phone and adjust to the most comfortable distance. Then, the participant was informed the experiment testing process and the purpose of the experiment.

The participant selected one type video about 6 min in length, and clicked the play button to watching the video clips. In this task, we aimed to research the influence of multiple buffering and stalling during in the video viewing influence to the user QoE. After the video buffering T seconds, the video continued to play X seconds. Then, stalling was appeared during T seconds. The video continued playing X seconds until the next stalling happened. Moreover, the delay times *(T + X) was a stalling segment (e.g. the video clip had one time delay, 1 * (T + X) = 1 stalling segment) (Fig. 4). The participant finished watching the video clips, rated the MOS, besides, they started to test

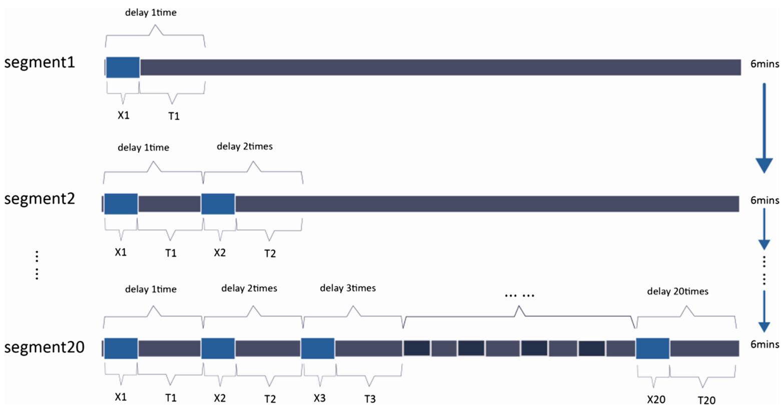

Fig. 4. The experiment material processing

the next stalling segments. We would examine each of these segments in turn. There were 20 stalling segments according to the stalling time. When the participant rated 1 score, this test was finished. The T included 0.1 s, 0.5 s, 1 s and 3 s. The X index was 2 s, 4 s, 8 s, 16 s, 32 s, 64 s and 128 s (Table 3).

Table 3. The ratio of video clip material about stalling segment

Delay time (T)	X1	X2	X3
MOS	5	4	3
Interval playing	2 s, 4 s, 8 s, 16 s, 32 s, 64 s, 128 s		
Stalling times	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20		

When the interval playing time get 32 s, 64 s and 128 s,
If the stalling times was more than 11, 6, 3 times, the next segment don not need test, respectively

Debriefing of participants revealed that the video they chose had a proper content and experimental materials were close to the actual situation. Most participants were patient enough to finish all the tasks. This fits with the fact that all participants won money in excess of the guaranteed bonus of 200RMB.

5 Results

5.1 Task 1

In task one, participants were asked to watch videos with different delay on a smart phone which is in the mobile network, and then assess the user experience using the MOS scale.

We investigated whether initial buffer time and stalling time have difference effect on user experience in the context of mobile network. For this, we average the scores of initial delays and stalling time on the phone. Results showed in Fig. 5 revealed that under the same length of waiting time, the scores of initial time were generally higher than those of stalling time, which indicated that user has a higher tolerance of initial delay than stalling time when they were watching a video in the context of mobile network. However, results showed in Fig. 5 also revealed that the overall trend of initial delay and stalling time are nearly the same. We then conducted a repeated measure one-way ANOVA including all 12 lengths of waiting time as factor levels. This analysis showed a significant main effect of these two factors on user experience (F (2, 30) ¼ 11.105; $p = 0.001$; partial Z2 ¼ .425).

The users have positive user experience and the participants' score reached the top at 5 when the delay time X1 = 0.1 s. Good user experience was suggested when the initial delay was no more than 0.5 s, with all the scores were higher than 4.5. The scores dropped to 4 when the delay time X2 = 1 s. Then, the scores dropped to approximately 3 when the delay time X3 = 3 s.

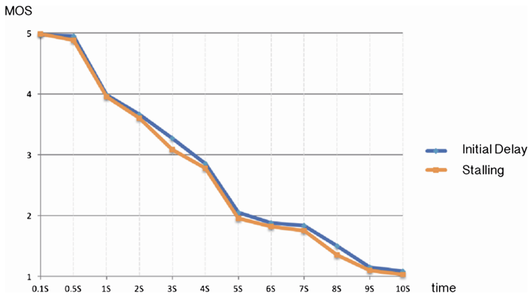

Fig. 5. The MOS's scores of initial delay and stalling time

5.2 Task 2

To test whether lengths of play duration and times of intervals exert an effect on watching experience in the context of mobile network, we conducted a repeated measures 2-way analysis of variance (ANOVA) including 7 different lengths of play duration (2 s, 4 s, 8 s, 16 s, 32 s, 64 s and 128 s) and times of intervals ranged from 1 to 15 as factor levels on three length of stalling time (X1 = 0.1 s, X2 = 1 s and X3 = 3 s) respectively. Figures 6, 7 and 8 respectively show the mean MOS scores on these three buffer levels.

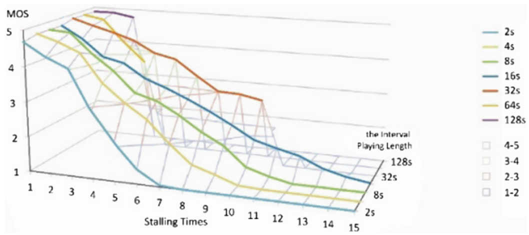

Fig. 6. Mean MOS scores with 0.1 s initial delay

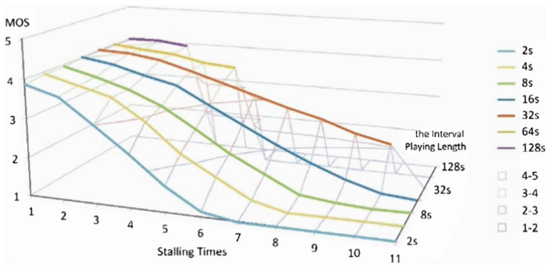

Fig. 7. Mean MOS scores with 1 s initial delay

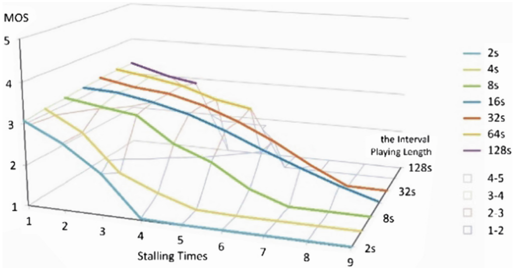

Fig. 8. Mean MOS scores with 3 s initial delay

For watching experience reflected by MOS scores of users in the context of mobile network, when the single stalling time is 0.5 s, analysis showed a significant main effect of play duration and times of intervals, as well as interaction between play duration and times of intervals with [$F(1,29) = 8.114$, $p < 0.001$, $R2 = 0.727$]. When the initial delay is 1 s, there was also a significant main effect of play duration ($p < 0.001$), times of intervals ($p < 0.001$) and interaction between play duration and times of intervals with [$F(1, 29) = 7.145$, $p = 0.000$, $R2 = 0.684$]. Same conclusion can be obtained from the analysis of 3 s-stalling-time with [$F(1, 29) = 8.024$, $p = 0.000$, $R2 = 0.519$]. Thus, the results revealed that both the decrease of single play duration time and increase of times of intervals while watching a video had negative effects on user experience.

6 Conclusion

In this paper, we firstly analyzed the most significant factor that influences the experience of watching videos in the context of mobile network: delay. It is divided into initial buffer time and stalling time. Then we calculated MOS score giving consideration of all these factors and compared the results depending on different conditions. Given by the experiment we performed, it can be figured out clearly that the MOS score is very similar in initial buffer time and stalling time. When the delay was shorter than 0.5 s, the users' MOS scores indicated excellent and good experience. However, when the delay was longer than 4 s on average, the MOS scores reflected negative emotions. All users showed a zero tolerance and an awful experience to the video if the delay was over 10 s in length. According to the emotional rating scores, after the participants finished experiment, all of them pointed out that they feel annoyed when delays occurred and they disliked delays when they were watching video. It should be noted that the effect of dislike the delays was significant when they were watching a video on a mobile device, although the participants prefer waiting for longer buffer time to endure single stalling time in the processing of video playing.

But it is worth noting that in the video playing duration ranged from 2 s to 16 s, the shorter the length of playing time, the lower the users tolerance to the times of single stalling. For example, when the video playing time is 2 s, the length of single stalling is 500 ms, the times of single stalling which the participants could tolerated is 7. However, when the video playing time was extended to 15 s, the times of single stalling which the participants could tolerated increased to 15. Meanwhile, when the video playing time was extended to 32 s, the video MOS scores of participants will not decline with the increase of the times of single stalling rapidly. For example, when the video playing time is 2 s, the times of single stalling increased from 1 to 3, the MOS score fell 0.64 points, while in 128 s' video playing duration, the score only fell 0.1 points. Moreover, at the condition of fixed playing time, the longer the single stalling time is taken, the less user want to continue watching the video. Take the 32 s playing time as an instance, when a single stalling takes only 0.1 s, user can tolerate 10 times single stalling and the MOS score is 10 points, however, when a single stalling takes 3 s, user's tolerance of stalling decrease to 8 times meanwhile MOS score decrease to 1 points. So in certain case of watching short video of 6 min or less and a fixed delay totals, we can draw two conclusions from the experiment, (1) the more delay time gathers in the initialization of the video instead of playing the better, (2) higher frequency and less single stalling time introduces worse experience than more single stalling time and lower frequency.

7 Future Work

For the future study, we want to add more factors in our investigation and compare with different video contents, and compare the QoE under different watching environments, like when the users are not single, or the video streaming service platform have some social functions in real time.

Acknowledgments. We would like to express our gratitude to Xiao Tan and Hui Yang who helped us during the experiment with his professional skill. The research was supported by Huawei Wireless Supported Projects.

References

Begen, A.C., Akgul, T., Baugher, M.: Watching video over the web: part 1: streaming protocols. Internet Comput. IEEE **15**(2), 54–63 (2011)

Cisco Systems Inc.: Cisco Visual Networking Index: Forecast and Methodology, 2012–2017, White Paper, May 2013

Selvidge, P.R., Chaparro, B.S., Bender, G.T.: The world wide wait: effects of delays on user performance. Int. J. Ind. Ergon. **29**, 15–20 (2002)

Szameitat, A.J., Rummel, J., Szameitat, D.P., Sterr, A.: Behavioral and emotional consequences of brief delays in human–computer interaction. Int. J. Hum Comput Stud. **67**, 561–570 (2009)

Dabrowski, J., Munson, E.V.: 40 years of searching for the best computer system response time. Interact. Comput. **23**, 555–564 (2011)

Schleifer, L.M., Amick III, B.C.: System response time and method of pay: stress effects in computer-based tasks. Int. J. Hum. Comput. Interact. **1**, 23–39 (1989)

Seow, S.C.: Designing and Engineering Time: The Psychology of Time Perception in Software. Addison Wesley, Upper Saddle River (2008)

Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J., Shneiderman, B.: Determining causes and everity of end-user frustration. Int. J. Hum. Comput. Interact. **17**, 333–356 (2004)

Nah, F.F.-H.: A study on tolerable waiting time: how long are Web users willing to wait? Behav. Inf. Technol. **23**, 153–163 (2004). http://dx.doi.org/10.1080/01449290410001669914

Thum, M., Boucsein, W., Kuhmann, W., Ray, W.J.: Standardized task strain and system response times in human–computer interaction. Ergonomics **38**, 1342–1351 (1995)

Le Callet, P., Möller, S., Perkis, A.: Qualinet white paper on definitions of quality of experience. In: European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003) (2012)

Zepernick, H., Engelke, U.: Quality of experience of multimedia services: past, present, and future. In: Adjunct Proceedings of the 9th European Interactive TV Conference (EuroITV 2011), pp. 115–119 (2011)

Raake, A., Egger, S.: Quality and quality of experience. In: Möller, S., Raake, A. (eds.) Quality of Experience: Advanced Concepts, Applications and Methods, pp. 11–34. Springer, New York (2014)

Brooks, P., Hestnes, B.: User measures of quality of experience: why being objective and quantitative is important. IEEE Netw. **24**(2), 8–13 (2010)

Larson, R.C.: Perspectives on queues: social justice and the psychology of queuing. Oper. Res. **35**, 895–905 (1987)

Clemmer, E.C., Schneider, B.: Toward understanding and controlling customer dissatisfaction with waiting. Report no. 89-115. The Marketing Science Institute, Cambridge, MA (1989)

Dube, L., Schmitt, B.H., Leclerc, F.: Consumers' reactions to waiting: when delays affect the perception of service quality. In: Srull, T.S. (ed.) Advances in Consumer Research, vol. 16, pp. 59–63. Association Consumer Research, Provo (1989)

Dube, L., Schmitt, B.H., Leclerc, F.: Consumers' affective response to delays at different phases of a service delivery. J. Appl. Psychol. **21**, 810–820 (1991)

Hui, M.K., Tse, D.K.: What to tell consumers in waits of different lengths: an integrative model of service evaluation. J. Mark. **60**, 81–90 (1996)

Branaghan, R.J., Sanchez, C.A.: Feedback preferences and impressions of waiting. Hum. Factors **51**, 528–538 (2009). http://dx.doi.org/10.1177/0018720809345684

Galletta, D.E., Henry, R.M., McCoy, S., Polak, P.: When the wait isn't so bad: the interacting effects of website delay, familiarity, and breadth. Inf. Syst. Res. **17**(1), 20–37 (2006). http://dx.doi.org/10.1287/isre.1050.0073

Krejcar, O.: Problem solving of low data throughput on mobile devices by artefacts prebuffering. EURASIP J. Wirel. Commun. Netw. **2009**, 802523 (2009). http://dx.doi.org/10.1155/2009/802523

Kuhmann, W.: Experimental investigation of stress-inducing properties of system response times. Ergonomics **32**, 271–280 (1989)

Kuhmann, W., Boucsein, W., Schaefer, F., Alexander, J.: Experimental investigation of psychophysiological stress-reactions induced by different system response times in human–computer interaction. Ergonomics **30**, 933–943 (1987)

Schaefer, F.: The effect of system response times on temporal predictability of work flow in human–computer interaction. Hum. Perform. **3**, 173–186 (1990)

Martin, G.L., Corl, K.G.: System response time effects on user productivity. Behav. Inf. Technol. **5**, 3–13 (1986)

Simoens, P., Vankeirsbilck, B., Deboosere, L., Ali, F.A., De Turck, F., Dhoedt, B., Demeester, P.: Upstream bandwidth optimization of thin client protocols through latency-aware adaptive user event buffering. Int. J. Commun. Syst. **24**, 666–690 (2011). http://dx.doi.org/10.1002/dac.1188

ITU-T P.910: Subjective video quality assessment methods for multimedia applications. International Telecommunication Union, September 1999

ITU-R BT.500-10: Methodology for the subjective assessment of the quality of television pictures. In: International Telecommunication Union, March 2000

Jones, C., Atkinson, D.J.: Development of opinion-based audiovisual quality models for desktop video-teleconferencing. In: International Workshop on Quality of Service (1998)

Caramma, M., Lancini, R., Marconi, M.: Subjective quality evaluation of video sequences by using motion information. In: Proceedings of International Conference on Image Processing, ICIP 1999, October 1999

Watson, A., Sasse, M.A.: Measuring perceived quality of speech and video in multimedia conferencing applications. In: Proceedings of ACM Multimedia 1998, September 1998, pp. 55–60 (1998)

Cisco Documentation: Cisco visual networking index: forecast and methodology 2013–2018, Cisco White Paper, June 2014