Pippi Longstocking Calculus for Temporal Stimuli Pattern on YouTube QoE: $1+1=3$ and $1\cdot4\neq4\cdot1$

Tobias Hossfeld
University of Würzburg
Institute of Computer Science
Würzburg, Germany
hossfeld@informatik.uni-wuerzburg.de

Dominik Strohmeier, Alexander Raake
AIPA, T-Labs
Tech. Univ. Berlin
Berlin, Germany
name.surname@telekom.de

Raimund Schatz
Telecommunications Research Center Vienna
FTW
Vienna, Austria
schatz@ftw.at

ABSTRACT

Video streaming over wireless broadband is currently responsible for half of the world-wide mobile data traffic and its share is expected to even further increase. Therefore, mobile operators need to understand how Quality of Experience (QoE) of mobile video streaming is impacted by through the network delivery. Due to HTTP video streaming used by video portals like YouTube or NetFlix, problems or insufficient resources in the networks manifest as stalling events during service consumption on an end user device. In this paper, we discuss the challenges for QoE monitoring for HTTP-based mobile video regarding three aspects. First, the impact of different patterns in terms of stalling event frequency and length on QoE is quantified on behalf of a study on YouTube. Second, we evaluate predictive power of aggregate measures like total stalling time. Third, we discuss the resulting consequences and challenges for mobile video QoE modeling, monitoring and provisioning.

Categories and Subject Descriptors

C.2.0 [**Computer Communication Networks**]: General

General Terms

Measurement, Performance

Keywords

HTTP streaming, YouTube, Quality of Experience, Stalling

1. INTRODUCTION

More than 50 % of the entire mobile data traffic in the internet are caused by mobile video. By 2016, the traffic share is expected to even increase up to 70 % [2], which will results in traffic of 7.6 exabytes. Hence, there is a strong need for mobile network operators to understand Quality of Experience (QoE) for video streaming and to satisfy their customers especially for mobile video services. On the one hand, customers find themselves in a strong position as they are able to choose between different competing providers on behalf of the offered prices as well as the service quality provided. On the other hand, operators need to effectively deliver video traffic while reducing costs in order to make profit from video traffic.

Hence, reliable QoE metrics for mobile video are required by operators in particular for video-on-demand services as provided by platforms like YouTube or Netflix. However, related work on QoE models for video applications does not consider HTTP streaming appropriately from a mobile operator perspective as they emphasize the impact of visual artifacts, but do not provide tools for assessing stalling patterns in terms of duration and frequency of rebuffering events. Especially for HTTP video streaming, insufficient resources or problems in the network make stalling a prominent quality impact factor. However, the question arises whether the aggregated information on the total stalling time is sufficient to quantify QoE. This would simplify the QoE model as well as QoE monitoring significantly. Or does reliable QoE monitoring for HTTP video streaming need more sophisticated solutions describing also interdependencies between single stalling events [8, 12]?

In this paper, we analyze subjective video quality assessments obtained via crowdsourcing for YouTube video-on-demand streaming. In particular, we consider different stalling patterns leading to the same total stalling time and answer the following questions:

1. *Can YouTube QoE be modeled on behalf of total stalling time only?*
2. *How do users perceive total stalling times caused by different stalling patterns?*
3. *Which consequences do result for QoE-based mobile video monitoring and provisioning?*

The remainder of this paper is structured as follows. Section 2 provides a background on QoE and temporal stimuli for HTTP video streaming. In Section 3, we present the results of our study on the impact of length and occuring patterns of stalling events during YouTube service consumptions. The implications of the results are then discussed for a mobile network operator in Section 4. Finally, Section 5 concludes this work and gives an outlook on current and future work.

2. BACKGROUND AND RELATED WORK

The focus of this paper is set on the Quality of Experience for HTTP video streaming and the evaluation of temporal degradations as key influencing factor [5]. In general, Quality of Experience (QoE) is defined as "the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state." It is impacted by a set of QoE Influencing Factors which are defined as "any characteristic of a user, system, service, application, or context whose actual state or setting may have influence on the Quality of Experience for the user." [16] While a lot of work is focusing on spatial degradations in the domain of video QoE, temporal effects are becoming prominent in time-sensitive media, or when service consumption or user interaction is interrupted by the service.

These interruptive effects can be categorized into three different types of temporal degradations, namely (i) initial delay, (ii) stalling, and (iii) freezing [10]. In addition, temporal effects of human perception, especially the recency effect, can become crucial in these services. The recency effect refers to the location of a "bad quality" event within a temporal service encounter. It describes that the overall quality, or the Mean Opinion Score (MOS), will be stronger influenced if a single drop to "bad quality" happens close to the end of service consumption compared to the bad quality event occuring toward the beginning. Although the recency effect cannot be expected for initial delays, since initial waiting durations are not clearly perceivable impairments during the service consumption, it is expected to become important with respect to stalling events and their occurence patterns during service consumption.

For YouTube video streaming, stalling events are caused by rebuffering of the video due to an insufficient content delivery rate over the network. If the video buffer is emptied due to slow data transmission, then the video freezes until the video buffer is filled again and the video restarts playing. To be more precise, as soon as the video buffer falls below a certain threshold Θ_0, the video stalls. As soon as the threshold Θ_1 is exceeded, the video playout starts. These parameters are derived in [20].

In [18], Pitrey at al. investigate the impact of quality impairment patterns on QoE. Although their target application is Scalable Video Coding (SVC) and not HTTP streaming as in this paper, their results provide some important findings for our research approach. In their study, they investigated the impact of impairment lengths and patterns on QoE for videos of a length of 10 s. First, the results show that there is a significant impact of the length of singular impairments on the Quality of Experience. Pitrey at al find a linear dependency between the decrease of QoE ratings and a logarithmic increase of impairment length. Second, the results also show that the number of impairments has impact on Quality of Experience. However, no significant difference is eventually found for different intervals between multiple impairments

Extending these findings into HTTP streaming, we were then able to show that stalling events are extremely critical

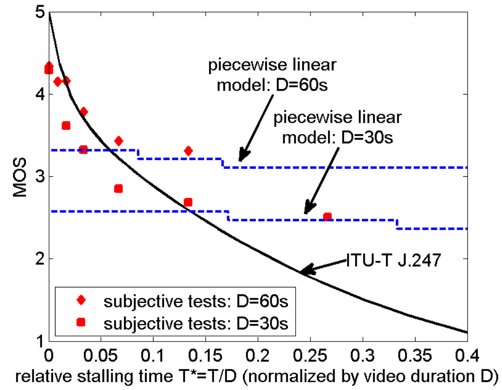

Figure 1: Comparison of subjective test model from own measurements [10], with HTTP streaming model [17] and Freezing Model [15].

for the subjective quality of such services [10]. Comparing the impact of initial delays vs. stalling events during service consumption, our results (Figure 1) demonstrate that users are extremely sensitive to interruptions and that services should be designed accordingly e.g. by increasing initial delay for prebuffering to overcome lack of resources.

Figure 1 revisits models found in literature and compares them to subjective user ratings from our experiments in Section 3. However, we consider here only a single stalling event and vary the total stalling length T. Two different video durations of D=30 s and D=60 s are considered. To be more precise, Figure 1 shows the MOS depending on the stalling time $T^*={}^T/_D$ relative to the video duration. The user ratings are given by markers. The curves of the piecewise linear model for HTTP streaming [17] and temporal model specified for freezing with skipping in ITU-T J.247 [15] show however that both models do not fully reflect YouTube QoE.

3. STALLING EVENTS AND YOUTUBE QUALITY OF EXPERIENCE

3.1 Research Method

To assess the impact of stalling events with respect to duration and occurence patterns, we conducted a subjective test based on a crowdsourcing approach. Due to the limitations of YouTube's application-layer traffic management [1], the study is conducted for fixed access conditions and results will be later discussed under the light of mobile access. For these tests, realistic test scenarios were defined that consider stalling patterns. We used D=30 s long videos which resemble typical YouTube videos of various content classes like news, sports, music clips, or cartoons. In case of a bottleneck with a fixed network data rate, periodic stalling patterns occur for YouTube [14], i.e., every Δt seconds a stalling event of almost fixed length L takes place. For N stalling events and a video duration D, it is $\Delta t={}^D/_{N+1}$. For assessment, we created a simulation environment for YouTube. An instance of the YouTube Chromeless Player was embedded into dynamically generated web pages. Via JavaScript commands, we used the video stream "pause" feature tosimulate stalling events. During the test, we varied *the number N of stalling*

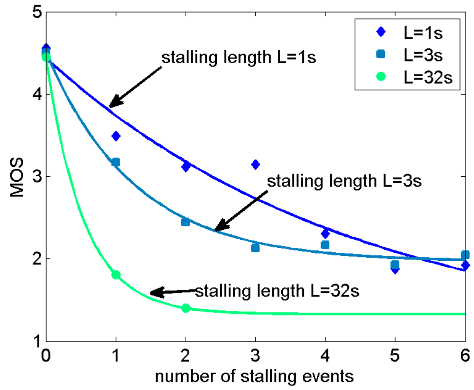

Figure 2: Longer length L of individual stalling events lead to larger total stalling times $T{=}N{\cdot}L$ and lower MOS values for varying number N of stallings.

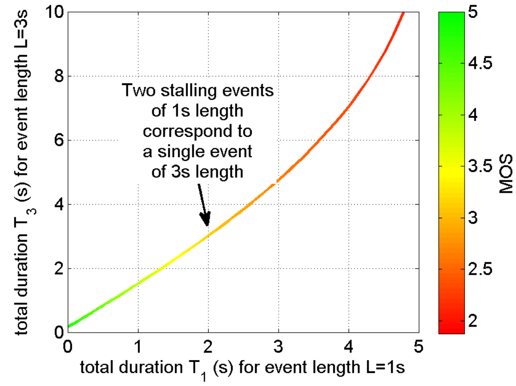

Figure 3: The inverse curve fitting functions return the total waiting times $T_L{=}L{\cdot}f_L^{-1}(y)$ for a given MOS value y. Stalling events of length $L{=}1\,$s and $L{=}3\,$s are considered. We observe $1\,$s$+1\,$s corresponds to $3\,$s.

events as well as *the length L of a single stalling event.*

For conducting the subjective user studies, we used crowdsourcing and in particular the Microworkers.com platform offering access to a large number of geographically widespread users. To identify unreliable user ratings, we applied the general test methodology developed in [11]. After verifying as real users, each test participant sequentially watched different YouTube video clips with our predefined stalling pattern according to the Absolute Category Rating (ACR). At the end of each test item, he or she was asked answered the question "Did you experience these stops as annoying?" on a 5-point scale labeled with (5) "imperceptible", (4) "perceptible", (3) "slightly annoying", (2) "annoying", (1) "very annoying". A total of 2,035 users from more than 60 countries participated in the YouTube QoE tests and rated the quality of 8,163 video transmissions. In the following, we present the key findings of our study.

3.2 QoE(1+1)=QoE(3): Interaction between length and occurence patterns

Which total stalling times lead to similiar QoE for different patterns? Our results show that the number of stalling events as well as the duration of stalling events decrease YouTube QoE significantly. Figure 2 shows the MOS y of the user ratings as well as exponential curve fittings $f_L(N){=}\alpha_L e^{-\beta_L N}{+}\gamma_L{=}y$ according to the IQX hypothesis [13, 6] which postulates the **I**nterdependency between **Q**oE and quality impairments (like stalling) to follow an e**X**ponential function. The results clearly show that users tend to be highly dissatisfied with two or more stalling events per clip, i.e. when the total stalling duration exceeds a few seconds only.

Further, it can be seen from Figure 2 that a MOS of 3 is retrieved either for two stalling events of length $1\,$s or one stalling event of length $3\,$s. For a long stalling event of $L{=}32\,$s, this value is reached only theoretically for $N{=}0.3343$. This observation already provides first evidence that total stalling times are not sufficient for modeling YouTube QoE properly.

Formalizing the interaction of total stalling durations and QoE, the function $\widetilde{f_L}(T)$ maps the total stalling time T to the MOS value based on an exponential fitting function $f_L(N)$, i.e. $\widetilde{f_L}(T){=}f_L(T/L)$. Assuming that for two different length $L{\neq}K$, the fitting functions converge towards the same value $\gamma_L{=}\lim_{N\to\infty} f_L(N){=}\lim_{N\to\infty} f_K(N){=}\gamma_K$, then the equation $\widetilde{f_L}(T){=}\widetilde{f_K}(T)$ has only the solution $T{=}\frac{KL\log\left(\frac{\alpha_K}{\alpha_L}\right)}{L\beta_K-K\beta_L}{\approx}0$ as $\alpha_L{\approx}\alpha_K$. Hence only without any stalling, we obtain a similiar MOS value for different stalling pattern $L{\neq}K$, i.e. only when $T{=}0\,$s.

Which stalling patterns do users perceive similarly? The inverse function $f_L^{-1}(y)$ returns the number of expected stalling events for a given MOS value y and can be easily derived as follows, $f_L^{-1}(y){=}-\frac{1}{\beta_L}\log\frac{y-\gamma_L}{\alpha_L}{=}N$. Then, the total waiting time T_L for a given MOS value y and single stalling length L is $T_L{=}L{\cdot}f_L^{-1}(y)$. This allows us to compare now total stalling times T_L and T_K for different stalling patterns $L{\neq}K$ which lead to the same MOS y.

Figure 3 shows the results for $L{=}1\,$s and $K{=}3\,$s. The corresponding MOS value is encoded by the color of the curve. We clearly see now that a single stalling event of length $L{=}3\,$s (i.e. $T_3{=}3\,$s) leads to the same MOS as two stalling events of length $L{=}1\,$s (i.e. $T_1{=}2\,$s). It can be seen that for small total stalling times ($T_1{\leq}3\,$s and $T_3{\leq}4.5\,$s) a linear approximation is appropriate with $T_3{\approx}1.5T_1$. For larger total stalling times, however, the shape of the curves gets a larger slope.

In general, we only observe a linear relationship between T_L and T_K resulting in same MOS value if we have the same lower limit of MOS scores, i.e. $\gamma_L{=}\gamma_K$. Then, we arrive at $T_L{=}-\frac{L}{\beta_L}\left(-\frac{\beta_K}{K}T_K{+}\log\frac{\alpha_K}{\alpha_L}\right)$. For any $\gamma_L{\neq}\gamma_K$, it holds $T_L{=}\frac{L}{\beta_L}\left(\mathrm{Log}\left[\alpha_L\right]{-}\mathrm{Log}\left[e^{-\frac{T_K\beta_K}{K}}\alpha_K{+}\gamma_K{-}\gamma_L\right]\right)$, as observed in Figure 3. Hence, there is no easy answer which stalling patterns are perceived similiarily by users.

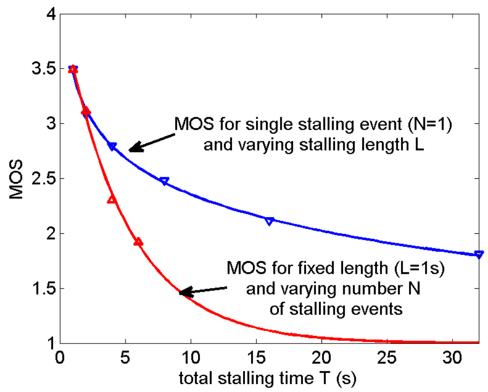

Figure 4: In scenarios leading to the same total stalling time $T{=}N{\cdot}L$, the minimum and maximum MOS values over these scenarios are considered which are obtained for maximum ($L{=}1$) and minimum ($N{=}1$) number of stalling events, respectively.

3.3 $\mathrm{QoE}(1{\cdot}4){\neq}\mathrm{QoE}(4{\cdot}1)$: Interaction of single stalling events and total stalling time

How large can the differences in MOS scores be for the same total stalling time? The total stalling time results from the number of stalling events · length of stalling events. Therefore, we can also analyze the minimum and maximum MOS values for the same $T{=}N{\cdot}L$. As a result, the maximum MOS value is obtained when there is only a single stalling event ($N{=}1$) of length $L{=}T$. Further, the minimum MOS value is observed for the maximum number of stalling events of length $L{=}1\,\mathrm{s}$, i.e. $N{=}T/1\,\mathrm{s}$. Thus, we conclude that the number of interruptions N is more dominating than the duration of the interruption itself.

Figure 4 shows that a single, longer stalling event is rated better than the corresponding overall stalling time resulting from several short stalling events for $T{\geq}4\,\mathrm{s}$. Beyond 4 s of total stalling time, the differences between minimum and maximum MOS are about 1 MOS. These findings again confirm that overall stalling time is not enough to monitor and predict YouTube Qualiy of Experience.

4. DISCUSSION: MOBILE YOUTUBE

Although the results of our study are constrained by the use of fixed access for generating stalling events, several implications arise for mobile network providers which clearly outline how YouTube QoE monitoring should be done. Firstly, stalling patterns and only total stalling times (or percentages of stalling time) have to be analyzed. Then, these stalling patterns need to be compared to corresponding patterns from the bottleneck scenario in order to check the applicability of the existing YouTube QoE mapping functions [14, 10]. Secondly, to improve the users' experienced quality despite growing mobile network usage, mobile operators have to adopt sophisticated QoE management [12]. QoE management enables to observe and react quickly to quality problems, at best before customers perceive them and ultimately decide to churn. From an economic perspective, an optimal QoE has to be achieved while constraining the application to behave as resource-efficiently as possible in order to minimize operational costs [7]. However, QoE monitoring (as prerequisite for QoE management) for YouTube requires some efforts [20] in order to accurately reconstruct the stalling events that arrive at the application layer, using network packet traces only. Based on deep packet inspection, stalling frequency and stalling duration both need to be measured, while the non-linearity of human perception demands for high QoS measurement accuracy, particularly in those cases where stalling frequency is low. Therefore, we give some practical guidelines for service provisioning to avoid QoE degradation at all in Section 4.2.

4.1 Preliminary Result: Stalling Pattern in 3G

A model of realistic stalling patterns occurring in mobile 3G environments is required for two different purposes. First, the subjective user tests presented in this work were based on a bottleneck scenario featuring static, fixed network capacity. If real-world 3G stalling patterns are too different, user studies have to be re-conducted according to the 3G stalling pattern model. Thereby 'too different' has to be defined appropriately. Second, previous studies [4, 19] have shown that users stop video playback when stalling occurs. [19] states that "34 % of the videos have a downloaded duration of 3 minutes or more, while in the case of bad reception quality their share drops to only 15 %." This can be interpreted that stalling strongly dissatifies the users fairly quickly (as shown in Figure 2) and almost half of the users react by stopping the video stream. This means that such observable user behavior patterns indicate whether QoE management or service provisioning is required for mobile video.

We conducted preliminary measurements of YouTube stalling patterns using the network of a public German 3G mobile network operator. In particular, we watched several YouTube videos and used tethering to share the 3G Internet connection of the smartphone with a laptop. The smartphone itself was connected with the laptop through USB. In order to monitor the stalling events on the application layer, the YouTube videos were embedded in a custom web page. The YouTubi API allows for embedding and playing the YouTube video. The video player status ("playing", "buffering", "ended") and the corresponding timestamp were monitored within the generated web page using Javascript.

Figure 5 shows the results from the measurement of five randomly selected video clips, focusing on stalling event length L as well as the rebuffering-free time period Δt between two consecutive stalling events. In the previously described fixed line bottleneck scenario, periodic stalling patterns have been observed resulting in constant stalling length and fixed time intervals between two consecutive stalling events. Hence, the coefficient of variation is zero for both measures. However, the 3G measurements reveal another picture. In fact, the coefficient of variation is larger than zero for all five videos. On average, the coefficient of variation is about 0.82. From the measurements, we conclude that non-periodic stalling patterns are the regular case in 3G networks.

As regards related in the wireless domain, [3] presents measurements of YouTube stalling patterns within a WiFi 802.11g environment where the bandwidth of the WiFi connection is limited to typical 3G bandwidths. The average of the

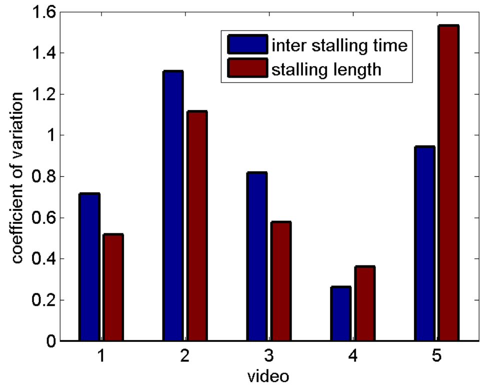

Figure 5: Smartphone YouTube stalling pattern via 3G. In contrast to bottleneck scenario, coefficient of variation is significantly larger than zero indicating non-periodic stalling patterns.

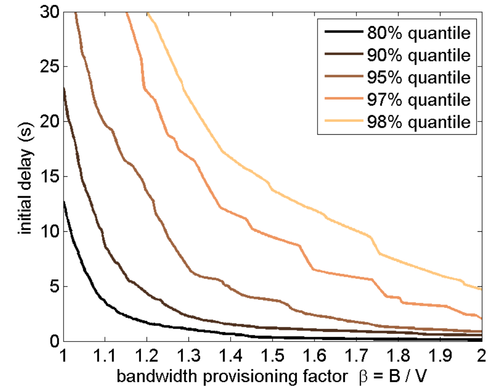

Figure 6: Bandwidth provisioning for YouTube videos and the p-th percentile of the resulting initial delay such that no stalling occurs. If the ISP provisions the video at a factor of β and the initial delay for buffering the video is y, then no stalling occurs with proability p.

measured throughput is 564 kbps. The mobile videos have an average audio bit rate of 62 kbps and an average video bit rate 109 kbps. As a result in [3], a median of 3 stalling events with a standard deviation of 2.8 is measured over 57 subjects who watched 3 videos in those conditions, i.e. 171 videos in total. For all 171 videos, the median of the length L of an individual stalling event is 0.9 s while the maximum is about 6.4 s. We approximate L by an exponential distribution with mean ${}^{\log 2}/_{0.9}$ to meet the median exactly. Then, 99.28 % of the stalling events have a length shorter than 6.4 s. For an exponential distribution, the coefficient of variation is equal to one. Hence, similar findings of non-periodic stalling patterns are observed also in this emulated 3G environment.

Our future work will address an extensive measurement campaign (similiar to [14]) within public 3G mobile networks to develop realistic stalling models for watching YouTube via smartphones.

4.2 Bandwidth Provisioning to Avoid Stalling

The subjective results in Section 3 have shown that even very short stalling events of a few seconds already decrease user perceived quality significantly. Therefore, a mobile network operator has to avoid stalling to satisfy his customers. One possibility to overcome resource limitations in the network is prebuffering the video at the cost of increasing initial delay (i.e. the startup time until playback). This QoE optimization approach takes only place before playing out the video. In this respect, [10] quantifies the impact of initial delays for filling the video buffer on QoE. The results show that initial delays up to 20 s are perceptible, but not really annoying. Nevertheless, from the user's perspective no stalling should occur and the initial delays should be minimal. In this context, the smallest initial delay for a particular video and a given network data rate which avoids stalling is referred to as *optimal initial delay*. [9] shows however that the computation of the optimial initial delay requires complex models taking into account diverse information of the network and the video contents, reflecting e.g. auto-correlation of frame sizes or scene changes within the video. To realize the required information exchange between network and application and to compute optimal initial delays before video playout, several solution approaches exist [8]. For example, the video frame structure is sent as metadata before the transmission of the video data to the application. Another option is that the application (or some other network entity) signals the video server the currently available network capacity, such that the video server (having the entire video structure information) computes the optimal initial delay and sends it back to the client.

In order to derive practical guidelines for mobile network operators, we investigate the optimal initial delay depending on the bandwidth provisioning factor of YouTube videos based on previous measurements. We conducted an extensive measurement series taking place from July to August, 2011 during which more than 37,000 YouTube videos were requested and about 35 GByte of data traffic was captured. The measurement details can be found in [14]. Figure 6 shows now the results for deriving practical guidelines. On the x-axis, the bandwidth provisioning factor β is plotted which is the ratio between the network data rate B and the video bit rate V, i.e. $\beta={}^{B}/_{V}$. On the y-axis, the optimal initial delay is depicted. The different curves in the figure show the p-th quantile over the set of video from the measurements. Hence, the p-th quantile denotes that with a probability of p there will be no stalling for the corresponding initial delay and bandwidth provisioning factor β for any video. As a rough guideline, providing a network data rate of 120 % of the video bitrate requires an optimal initial delay of 5 s in order to ensure that 95 % of the users encounter no stalling.

5. CONCLUSIONS

In this paper we have investigated Quality of Experience (QoE) quantification and modeling for HTTP-streamed mobile video, i.e. video accessed over wireless networks. Our subjective evaluation results regarding the impact of different stalling patterns on YouTube QoE reveal differences of up to 1 MOS for equivalent total stalling times. Thus, Pippi Longstocking calculus like 1+1=3 and 1·4≠4·1 applies to mobile video and consequently, aggregated metrics like total stalling time (or equivalents like stalling time percentage) are *not* sufficient for estimating QoE. The stalling pattern itself matters and therefore has to be taken into account by QoE models ensure prediction accuracy, despite the increase in complexity this implies. For the same reason, stalling patterns occuring in the wireless 3G domain have to studied in greater detail as we found them to be highly irregular (in terms of stalling intervals and durations occurding within one clip view session) and thus fundamentally different to those that have been studied so far in the context of streaming video QoE.

Our results also show that QoE management for mobile video traffic has strong potential when it comes to improving QoE while increasing resource utilization efficiency at the same time. In this respect, the probability of stalling can be decreased by technical measures such as caching and intelligent bandwidth provisioning (taking into account the native video bitrate), but also by shifting the initial delay to the optimal point. For these approaches we provided first guidelines in this paper.

As concerns future work, we see the necessity of extensive measurement campaigns in public wireless broadband networks to achieve a reliable characterization of stalling patterns. The resulting pattern models will support the design of QoE studies on irregular stalling patterns which ultimately will allow for the development of a complete mobile video QoE model covering arbitrary stalling patterns. This model (which should not necessarily be limited to YouTube only) will not only enable accurate QoE monitoring but guide QoE-based resource management in tomorrow's wireless networks.

6. ACKNOWLEDGMENTS

This work was partly funded by Deutsche Forschungsgemeinschaft (DFG) under grants HO 4770/1-1 and TR257/31-1, in the framework of the EU ICT Project SmartenIt (FP7-2012-ICT-317846), the project ACE 2.0 funded by the Austrian competence center program COMET, and the COST QUALINET Action IC1003. The authors alone are responsible for the content.

7. REFERENCES

[1] Alcock, S., and Nelson, R. Application flow control in youtube video streams. *SIGCOMM Comput. Commun. Rev. 41*, 2 (Apr. 2011), 24–30.

[2] Cisco. Cisco visual networking index: Global mobile data traffic forecast update, 2011-2016, February 2012.

[3] De Pessemier, T., De Moor, K., Joseph, W., De Marez, L., and Martens, L. Quantifying the influence of rebuffering interruptions on the user's quality of experience during mobile video watching. *IEEE Transactions on Broadcasting PP*, 99 (Mar. 2013).

[4] Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D., Ganjam, A., Zhan, J., and Zhang, H. Understanding the impact of video quality on user engagement. In *Proceedings of the ACM SIGCOMM 2011 conference* (New York, NY, USA, 2011), SIGCOMM '11, ACM, pp. 362–373.

[5] Egger, S., Hossfeld, T., Schatz, R., and Fiedler, M. Waiting Times in Quality of Experience for Web Based Services. In *QoMEX 2012* (Yarra Valley, Australia, July 2012).

[6] Fiedler, M., Hossfeld, T., and Tran-Gia, P. A generic quantitative relationship between quality of experience and quality of service. *Netwrk. Mag. of Global Internetwkg. 24* (March 2010), 36–41.

[7] Hossfeld, T., Fiedler, M., and Zinner, T. The QoE Provisioning-Delivery-Hysteresis and Its Importance for Service Provisioning in the Future Internet . In *Proceedings of the 7th Conference on Next Generation Internet Networks (NGI)* (Kaiserslautern, Germany, June 2011).

[8] Hossfeld, T., Liers, F., Schatz, R., Staehle, B., Staehle, D., Volkert, T., and Wamser, F. Quality of Experience Management for YouTube: Clouds, FoG and the AquareYoum. *PIK - Praxis der Informationverarbeitung und -kommunikation (PIK)* (Aug. 2012).

[9] Hossfeld, T., Liers, F., Volkert, T., and Schatz, R. FoG and Clouds: Optimizing QoE for YouTube. In *KuVS 5thGI/ITG KuVS Fachgespräch NG Service Delivery Platforms* (Munich, Germany, Oct. 2011).

[10] Hossfeld, T., Schatz, R., Egger, S., Fiedler, M., Masuch, K., and Lorentzen, C. Initial delay vs. interruptions: between the devil and the deep blue sea. In *under submission* (2012).

[11] Hossfeld, T., Schatz, R., Seufert, M., Hirth, M., Zinner, T., and Tran-Gia, P. Quantification of YouTube QoE via Crowdsourcing. In *IEEE International Workshop on Multimedia Quality of Experience - Modeling, Evaluation, and Directions (MQoE 2011)* (Dana Point, CA, USA, Dec. 2011).

[12] Hossfeld, T., Schatz, R., Varela, M., and Timmerer, C. Challenges of QoE Management for Cloud Applications. *IEEE Communications Magazine* (Apr. 2012).

[13] Hossfeld, T., Tran-Gia, P., and Fiedler, M. Quantification of quality of experience for edge-based applications. In *20th International Teletraffic Congress (ITC20)* (Ottawa, Canada, jun 2007).

[14] Hossfeld, T., Zinner, T., Schatz, R., Seufert, M., and Tran-Gia, P. Transport Protocol Influences on YouTube QoE . Tech. Rep. 482, University of Würzburg, July 2011.

[15] Huynh-Thu, Q., and Ghanbari, M. No-reference temporal quality metric for video impaired by frame freezing artefacts. In *Proceedings of the 16th IEEE international conference on Image processing* (Piscataway, NJ, USA, 2009), ICIP'09, IEEE Press, pp. 2197–2200.

[16] Le Callet, P., Möller, S., and Perkis, A., Eds. *Qualinet White Paper on Definitions of Quality of Experience (2012).* Version 1.1. European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Lausanne, Switzerland, June 2012. http://www.qualinet.eu/.

[17] Mok, R. K. P., Chan, E. W. W., and Chang, R. K. C. Measuring the quality of experience of http video streaming. In *IEEE/IFIP IM (Pre-conf Session)* (Dubland, Ireland, May 2011).

[18] Pitrey, Y., Engelke, U., Barkowsky, M., Pepion, R., and Le Callet, P. Subjective quality of svc-coded videos with different error-patterns concealed using spatial scalability. In *Visual Information Processing (EUVIP), 2011 3rd European Workshop on* (july 2011), pp. 180 –185.

[19] Plissonneau, L., and Biersack, E. A longitudinal view of http video streaming performance. In *Proceedings of the 3rd Multimedia Systems Conference* (New York, NY, USA, 2012), MMSys '12, ACM, pp. 203–214.

[20] Schatz, R., Hossfeld, T., and Casas, P. Passive YouTube QoE Monitoring for ISPs. In *Workshop on Future Internet and Next Generation Networks (FINGNet-2012)* (Palermo, Italy, July 2012).