Application-layer monitoring of QoE parameters for mobile YouTube video streaming in the field

Michael Seufert, Pedro Casas, Florian Wamser, Nikolas Wehner, Raimund Schatz, Phuoc Tran-Gia

Angaben zur Veröffentlichung / Publication details:

Seufert, Michael, Pedro Casas, Florian Wamser, Nikolas Wehner, Raimund Schatz, and Phuoc Tran-Gia. 2016. "Application-layer monitoring of QoE parameters for mobile YouTube video streaming in the field." In *2016 IEEE Sixth International Conference on Communications and Electronics (ICCE), 27-29 July 2016, Ha-Long, Vietnam*, edited by Nguyen Quynh, Hoang Minh Son, Randy Giles, Yong Soo Cho, Thang Manh Hoang, and Nguyen Hoang Dung, 411–16. Piscataway, NJ: IEEE. https://doi.org/10.1109/cce.2016.7562672.

Application-Layer Monitoring of QoE Parameters for Mobile YouTube Video Streaming in the Field

Michael Seufert*, Pedro Casas†, Florian Wamser*, Nikolas Wehner*, Raimund Schatz†, Phuoc Tran-Gia*
*University of Würzburg, Institute of Computer Science, Würzburg, Germany
{seufert | florian.wamser | nikolas.wehner | trangia}@informatik.uni-wuerzburg.de
†AIT Austrian Institute of Technology, Vienna, Austria
{pedro.casas | raimund.schatz}@ait.ac.at

***Abstract*—YouTube video streaming is one of the most popular and most demanding services in cellular networks. Thus, operators are concerned about the quality of the streaming delivered by their networks and would like to monitor the Quality of Experience (QoE) of the end users. In this work, we conduct a field study of mobile YouTube video streaming, in which both network flow parameters and application-layer streaming parameters were monitored, and present the characteristics of current mobile YouTube streaming. The impact of both approaches is investigated showing that monitoring network parameters is not sufficient to directly infer the resulting QoE. In contrast, the streaming parameters, which can be obtained from application-layer monitoring, show high correlations to the subjectively experienced quality, and thus, are better suited for QoE monitoring.**

I. Introduction

In today's Internet, YouTube is one of the most popular and volume-dominant services. Every day, people watch hundreds of millions of hours on YouTube and generate billions of views. More than half of these views come from mobile devices [1]. To satisfy their customers, the performance of YouTube Mobile is essential for mobile operators, who must cope with the huge amount of traffic within the constrained cellular networks. Therefore, they need to understand how the streaming quality is perceived by the end users in order to keep the Quality of Experience (QoE) at satisfying levels.

The QoE of video streaming is mainly affected by waiting times, such as initial delay and stalling, and the video quality [2]. Due to HTTP adaptive streaming technology, the video quality can be changed according to the current network conditions in order to avoid or shorten these waiting times. In case of YouTube, the resolution of the video is the quality level, which is switched according to a client-side adaptation logic when the network conditions change.

To quantify the QoE of end users, network providers used to estimate streaming characteristics based on traffic characteristics or deep-packet inspection of video packet flows (e.g., [3]). However, the recent trends towards encrypted HTTP traffic, impede them to obtain accurate estimates. Thus, they can only rely on the monitoring of network parameters of identified video flows, such as flow size, duration, signal strength, or throughput. A different approach is the monitoring of application-layer streaming parameters, such as initial delay, stalling, and video quality, directly at the end user device (e.g., [4]). This concept includes that the monitored data need to be communicated to the network operator, however, QoE-related parameters can be obtained directly.

To inspect both approaches, a YouTube mobile field study was conducted. Participants were asked to use their own cellular ISPs and smartphones, which were equipped with monitoring applications for both network flow parameters and application-layer streaming parameters, to stream YouTube videos. After watching a video, the participants were asked to fill a web-based questionnaire on their subjective perception of the streaming. In this paper, the study and the used applications are presented in detail. The characteristics of current mobile YouTube streaming as measured in the field study are presented. Moreover, the applicability of both monitoring approaches to estimate the QoE of end users is investigated.

The remainder of the paper is structured as follows. First, related work is presented in Section II in order to give an overview of work related to HTTP video streaming and monitoring of QoE. In Section III, the used applications are described and the field study is summarized. Afterwards, the results on mobile streaming characteristics, the impact of network flow parameters, and streaming parameters are presented in Section IV. Section V concludes the paper.

II. Related Work

The problem of QoE assessment in HTTP video streaming is already well-known. Initial delays and stallings are the key parameters defining the QoE of video streaming [5]–[7]. It was shown that while most users can tolerate moderate initial delays, stalling has a huge impact, as already little stalling severely degrades the perceived quality.

Whilst adaptive streaming concepts are well-known for a long time, their broad commercial usage has only risen recently, and the topic is getting more and more attention within the research community. Authors in [8] found that quality adaptation could effectively reduce stalling by 80% when bandwidth decreased in a mobile environment, and was responsible for a better utilization of the available bandwidth when bandwidth increased. However, quality switches have an impact on perceived quality themselves, as they increase or decrease the video quality according to the switching direction [9]. Authors in [10] found that only the time on each quality layer has a significant impact on QoE, but not the number of quality switches. In [11], authors found that resolution is a key parameter for video QoE on small displays. They concluded that low resolutions contributed to enhanced eyestrain of the subjects. However, authors found in [12] that QoE for YouTube in modern smartphones is actually not much impaired by the resolution switches, as the size of the screens are small and users are already much used to watching YouTube in such devices. A more comprehensive survey of the QoE of adaptive streaming can be found in [2].

When it comes to the specific study of YouTube QoE in mobile networks and mobile devices, there are some recent papers worth mentioning. In [13], authors study the characteristics of YouTube traffic for both Android and iPhone mobile devices connected to a cellular network, showing that mobile devices have a non-negligible impact on the characteristics of the downloaded traffic (for example, in terms of video resolution and flow download control behavior). Closer to our work, authors in [14] describe a subjective QoE evaluation framework for mobile Android devices in a lab environment. Additionally, they perform some basic QoE-based study on the classical, non-adaptive YouTube streaming using very low bit rate videos. Authors in [15] study the QoE of YouTube in mobile devices through a field trial, but completely neglect the analysis and impact of adaptive streaming as we do. In [3], authors took a further step and introduced an on-line monitoring system for assessing the QoE of YouTube in cellular networks using network-layer measurements only. Newer papers have evaluated the QoE of current smartphone apps from both lab subjective tests [12] and field trials [16].

There has also been a recent surge in the development of tools and software libraries for measuring network performance on mobile devices: some examples are Mobiperf [17], Mobilyzer [18], and the Android version of Netalyzr [19]. Authors in [4], [20] presented YoMoApp, an Android app for passively monitoring QoE-relevant parameters for YouTube video streaming in smartphones. Authors in [21] introduced Prometheus, an approach to estimate QoE of mobile apps, using both passive in-network measurements and in-device measurements, applying machine learning techniques to obtain mappings between QoS and QoE. In [22], authors introduced QoE Doctor, a tool to measure and analyze mobile app QoE, based on active measurements at the network and the application layers. Additional papers in a similar direction tackle the problem of modeling QoE for Web [23] in cellular networks, and video [24].

III. Study Description

A. YoMoApp

YoMoApp (YouTube Performance Monitoring Application) [4], is an Android application, which passively, non-intrusively monitors application-level key performance indicators (KPIs) of YouTube adaptive video streaming on end-user smartphones. The monitored KPIs can be used to analyze the QoE of mobile YouTube video sessions. The tool is currently unique on the market. Comparable apps so far measure only parameters such as data volume or latency. However, according to [2], the main influence parameters of the YouTube QoE are *stallings* and *video quality*. To obtain these parameters, we monitor the buffer and the resolution of the YouTube videos.

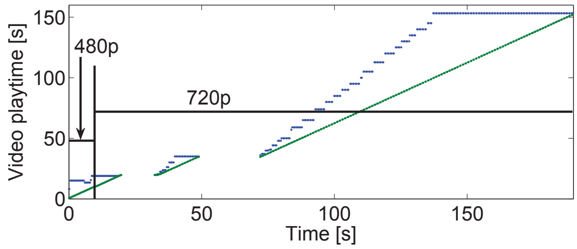


Fig. 1. Buffer and video quality of an exemplary video streaming: current video playtime (green), buffered video playtime (blue), played out video quality/resolution (horizontal black lines).

The approach is as follows. The original YouTube App is fully replicated in functionality and design. To this end, existing libraries from YouTube are used that are available for YouTube developers. An Android WebView browser element is embedded to display the YouTube mobile web site on which HTML5 video playback, including adaptive streaming, is possible. Additional functions are added, which ultimately perform the monitoring of the application parameters in the newly created app. The monitoring is done at runtime via JavaScript, which queries the HTML5 $\langle video \rangle$ object (i.e., player state/events, buffer, and video quality level). More details can be found in [4].

Fig. 1 shows the buffer and video quality data of an exemplary run in their processed form. Postprocessing of the data is required because JavaScript can sometimes introduce inconsistencies and obvious errors, e.g., missed player events, non-equidistant data queries, missing/incorrect values. However, as demonstrated in [4], YoMoApp proved to perform accurate measurements on a sufficiently small time scale ($\sim$1 s).

B. Network Flow Measurement Tool

To monitor the network usage of the field-trial participants, we used a simple Android-based passive monitoring tool, which captures several metrics for all the traffic flows on the device. Other tools available from the literature (e.g., [17]–[19], [22]) either rely on active measurements only or are too specific and could not be used.

Table I lists the different metrics passively monitored for each traffic flow by our network measurement tool. A flow is associated to the specific app generating it by using the Android Developers' APIs. The first metric is a the IMEI (International Mobile Station Equipment Identity), which is a unique number identifying a 3GPP device. Metrics 2 to 6 report results of traffic flow measurements, including the flow start time, the flow direction (uplink or downlink), the flow duration, the size of the flow, and the average flow transfer throughput, which is computed as the ratio between the flow size and the flow duration. Metric 7 identifies the app, which

TABLE I
METRICS RECORDED FOR EACH DATA FLOW, WHICH ARE EXTRACTED VIA THE ANDROID DEVELOPERS' API.

Metric	Metric Name	Units	Example
1	Device ID (IMEI)	-	352668049725157
2	Flow start time	s	1430825689
3	Flow direction (up/down)	-	downlink
4	Flow duration	s	10.24
5	Flow size	KB	4041.00
6	Avg. flow throughput	kbps	3157.03
7	App	-	de.yomoapp
8	Signal strength	dBm	-71
9	Operator (MCC.MNC)	-	295.4
10	Cell ID	-	16815
11	Cell location {lat;lon}	degree (°)	{48.194;16.348}
12	RAT	-	LTE

generated the corresponding flow, using the Android naming scheme. Metric 8 provides the strength of the signal at the smartphone. Metrics 9 to 11 report the operator providing the Internet access, the cell to which the smartphone is attached at the flow start, and the cell's position (i.e., longitude and latitude). Metric 12 indicates the Radio Access Technology (RAT) used by the smartphone (e.g., LTE, 3G, 2G, EDGE, etc.). Metrics 7 to 12 are recorded at the time at which a new flow starts. All metrics are logged locally at the smartphone, and are periodically uploaded to a server for post-processing and analysis.

C. Web-based Rating App

The users were asked to provide QoE feedback through a web-based app immediately after using YoMoApp. On the webpage, the users were required to rate the overall quality on a 5-point ACR scale ranging from bad (1) to excellent (5), and to indicate whether the session quality was acceptable. Moreover, the users were asked to indicate if to what extend they were annoyed by the initial delay on a 5-point ACR scale ranging from very annoyed (1) to not at all (5), and if they noticed any interruptions or stops during the streaming. If yes, they had to indicate whether they experienced these interruptions as annoying on the same scale as for initial delay. The ratings were stored on the web server for later analysis.

D. Field Study

The field trial consisted of 30 participants using their own smartphones and cellular ISPs to stream videos using YoMoApp in Vienna, Austria, for a total span of 2 weeks in January 2015. Field trial participants were compensated with vouchers for their participation, which proved to be sufficient for achieving correct involvement in the study.

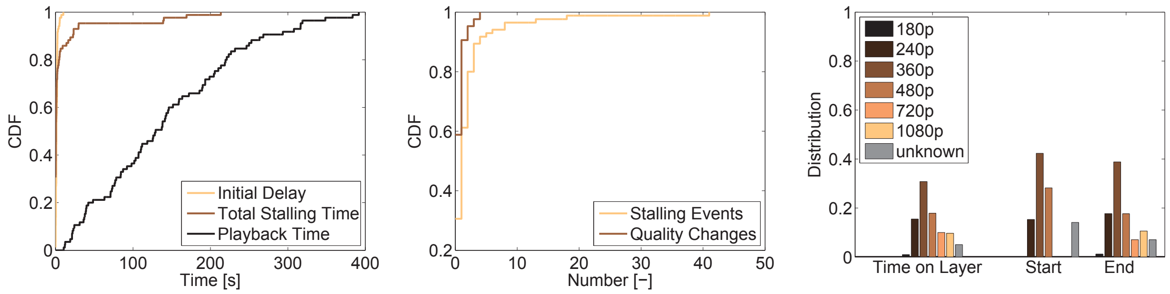
After the study, the log files from three sources were collected, namely, the YoMoApp logs, the network measurements, and the QoE ratings. In total, 85 video sessions were monitored. To map the corresponding network measurements, we identify the traffic flows that overlap with the streaming log. Although this is a straightforward approach, only 30 streaming logs could be mapped to network measurements. For the other streaming sessions, the network monitoring app was not actively running on the participants' devices or an Internet connection via WiFi was used. WiFi sessions had to be excluded because the signal strength parameter was not available, and they performed significantly better than the mobile sessions. Moreover, for all streaming logs, we matched the corresponding QoE rating based on the device identifier and the rating time. Thereby, we only accepted QoE ratings, which were submitted latest 15 minutes after the streaming. The other ratings are considered unreliable as users were asked to rate immediately after the playback. Further filtering was done based on the noticing of stalling. If the users rated the presence of interruptions contrary to the monitored stalling, the rating was not accepted. This resulted in 30 rated streaming sessions in total. Combining all three logs, only 10 sessions remained, for which both network measurements and QoE ratings are available.

IV. RESULTS

In the course of the field study, 85 streaming sessions were logged by the YoMoApp application. Figure 2 gives statistical insights in the monitored sessions. Figure 2a presents the CDF of initial delay (yellow), total stalling time (brown), and playback time (black). The playback times of a single video range up to 391.7 s, having a mean of 142.9 s. This means, the participants did not deliberately watch short video clips just to finish their task, but playback times show a reasonable involvement of the participants with the field study and suggest that liked content was selected. Over all streaming sessions, short initial delays are observed (avg.: 1.60 s, max.: 4.02 s), although stalling times are considerably higher (avg.: 10.78 s, max.: 213.43 s). Nevertheless, the waiting times are generally low, which results in around 90 % of the sessions having an initial delay of less than 2.5 s, and around 75 % of the sessions having a total stalling time of less than 2.5 s.

Figure 2b shows the CDF of the number of stalling events (yellow) and the number of quality changes (brown). A quality change means the switching between two different quality layers, i.e., in the context of YouTube, the switching between two different video resolutions. The average number of stalling events is 2.22, which corresponds to the low total stalling times, but still a maximum number of stalling events of 41 was monitored. In contrast, the number of quality switches is low (avg.: 0.58, max: 4), which indicates that the adaptation logic of YouTube is rather conservative and avoids too many quality changes.

Finally, Figure 2c presents more details on the played out quality. The bar plot presents the distribution of the playout time of the different video qualities (time on layer). Moreover, it shows distributions of the quality played out at the start and the end of a streaming session. It can be seen that all qualities were used, although mostly resolutions 240p (15.5%), 360p (30.8%), and 480p (17.8%) are streamed. Additionally, almost 20% of the time HD content (720p or 1080p) can be watched on the mobile devices. Note that for some times, the video resolutions could not be determined by YoMoApp (unknown). Looking at the start quality, again a conservative behavior of YouTube can be observed as no session downloads a HD



(a) Time-related streaming parameters (b) Event-related streaming parameters (c) Played out video quality

Fig. 2. Monitoring of streaming during the field study.

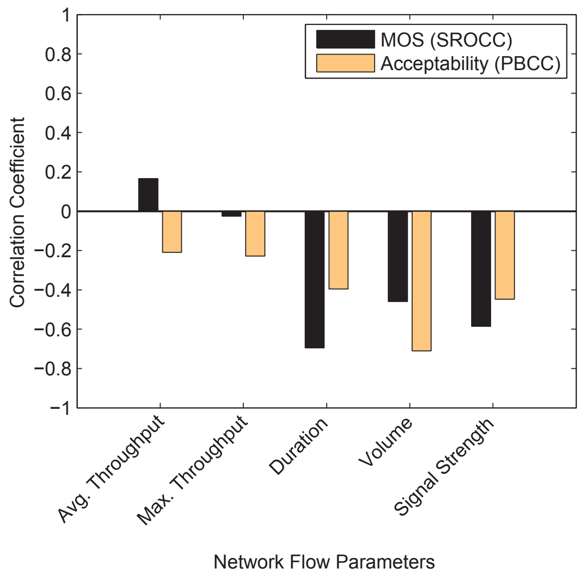


Fig. 3. Correlations between network parameters and QoE ratings.

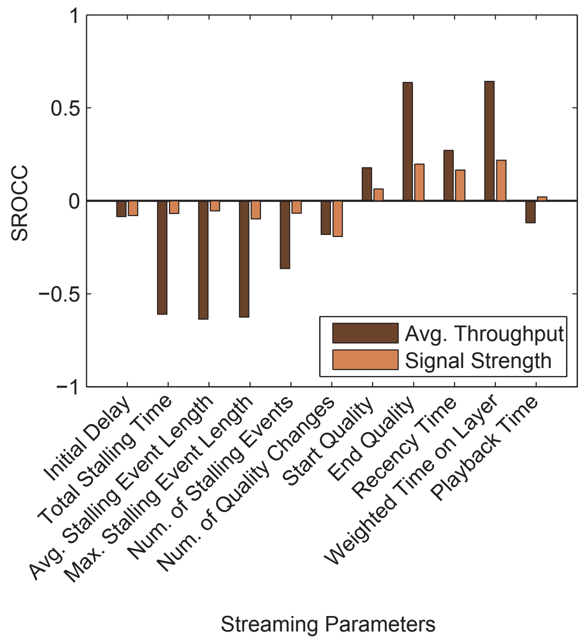


Fig. 4. Correlations between network parameters and streaming parameters.

quality from the start. The end quality distribution, on the other hand, is similar to the distribution of the time on layer, and indicates that the streaming sessions tend to improve from the low start qualities if the network conditions permit.

In the following, the impact of the different network and streaming parameters will be investigated in detail.

A. *Impact of Monitored Network Flow Parameters*

Figure 3 presents the correlations of the network measurements to the QoE ratings of the participants. The bar plot shows the different network parameters on the x-axis. The black bars indicate the Spearman rank-order correlation coefficient (SROCC) of the network parameter values and the mean opinion scores (MOS) of the users. The yellow bars indicate the correlation between the network parameters and the acceptability. As acceptability is a dichotomous variable, the point-biserial correlation coefficient (PBCC) is plotted.

Both the average flow throughput and the maximum flow throughput show little correlations to both MOS and acceptability. This means that a higher throughput does not necessarily result in a better streaming experience. For the other network parameters flow duration, flow volume, and signal strength, negative correlations can be observed. This is especially counter-intuitive for signal strength as a higher signal strength seems to reduce the streaming quality.

Note that the data set consists of only 10 complete streaming sessions (both network measurements and QoE ratings available) is too small to show generalizable results. However, it becomes clear that the network measurement is not sufficient for an accurate QoE estimation. In the following, we will investigate the impact of network parameters on the streaming parameters in more detail.

Figure 4 shows a bar plot of the SROCC between the two network parameters average flow throughput (dark brown bars) and signal strength (light brown bars), and different streaming parameters. The investigated streaming parameters on the x-axis are initial delay, total stalling time, average and maximum length of a stalling event, number of stalling events, number of quality changes, start and end quality, recency time (i.e., time after last quality change), weighted time on layer (i.e., a linearly weighted sum of the time spent on different

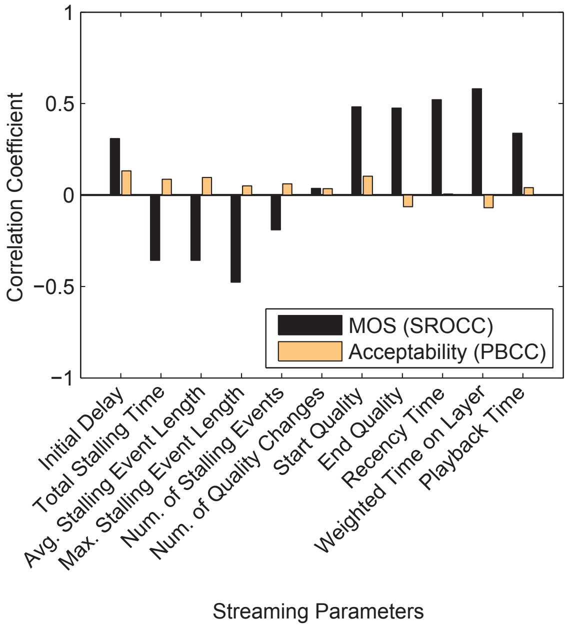


Fig. 5. Correlations between streaming parameters and QoE ratings.

quality layers), and playback time. Note that quality (layer) and quality change refers to the different resolutions used by the YouTube streaming. The data set consists of 30 records.

The average flow throughput shows a high correlation to end quality (0.64) and the weighted time on layer (0.64). This means, a high throughput seems to generally improve the quality level of the streamed video. Moreover, negative correlations from -0.64 to -0.36 for all stalling time parameters is visible. Thus, a high maximum flow throughput also reduces the stalling experienced by the end user. A similar behavior could be observed for the maximum flow throughput (similar SROCCs) and flow volume (smaller SROCCs). Flow duration, in contrast, shows an inverse behavior having positive correlations to stalling parameters and negative correlations to quality parameters. Confirming the findings from above, signal strength does also not perform well as an indicator for streaming parameters. It shows only low correlations having little SROCCs ranging from -0.19 to 0.22.

To sum up, some network parameters, especially average or maximum flow throughput, are closely linked to the performance of the streaming and the resulting experienced streaming quality. Nevertheless, as observed above, these network parameters cannot be directly used to infer the resulting QoE.

B. Impact of Monitored Streaming Parameters on QoE

The impact of streaming parameters on the subjective rating is investigated on a data set containing 30 streaming sessions. Figure 5 presents the correlations between streaming parameters and the QoE ratings. The bar plot depicts the SROCCs between the streaming parameters, which are on the x-axis, and the MOS (black), as well as the PBCCs between the streaming parameters and the acceptability ratings (yellow).

Interestingly, no negative effect is visible for initial delay, which might be due to the fact that only small initial delay times were monitored during the field study. Also the ratings of the dedicated initial delay question were high (avg.: 3.53), which indicates that initial delay was not an issue here.

It can be seen that only the stalling parameters have a clear negative correlation to the MOS, which confirms previous findings that stalling is the worst quality degradation of video streaming (cf. [2]). The correlation is highest for maximum stalling event length (-0.48), still reasonably high for total stalling time (-0.36) and average stalling event length (-0.36), but smallest for number of stalling events (-0.19). In contrast, when directly asked about the annoyance caused by stalling, the correlations are generally small, but are highest for the maximum stalling event length.

Taking a look at the video quality parameters, positive correlations to the MOS can be observed. The weighted time on quality layer shows the highest correlation (0.58), but also start (0.48) and end quality (0.48) have high SROCCs. Confirming the results from [10], the number of quality switches has no correlation to MOS (0.04). However, the recency time, which is an indicator for quality level fluctuation, has a high correlation to the MOS (0.52). This could still suggest that the users prefer a stable streaming quality. Finally, a positive correlation of playback time and MOS is visible (0.34), which confirms that users tend to watch longer when the streaming quality is better.

In contrast to the MOS ratings, the acceptability shows only low correlations to the streaming parameters having PBCCs ranging from -0.07 to 0.13. This can be explained by the fact that acceptability of a streaming session is not strongly influenced by a single streaming parameter but rather depends on a more complex combination of them. Still, a high positive PBCC of 0.54 between MOS and acceptability could be observed from the collected ratings.

All in all, the results of the field study show that it is important to monitor the parameters of the video streaming at the application layer. The streaming parameters proved to provide better insights into the subjective experience of users than the network flow parameters. This necessitates the usage of tools like YoMoApp for future QoE monitoring. Based on these monitored parameters, a holistic QoE model has to be developed, which allows for an accurate QoE estimation.

V. Conclusion

In this work, a field study of mobile YouTube video streaming was conducted with YoMoApp. This custom Android QoE monitoring application allows the usage of YouTube and passively measures streaming parameters on the application layer, e.g., initial delay, stalling, and quality changes. Participants were asked to use YoMoApp to stream and watch videos on their own smartphones over their cellular ISPs. In addition to the application-layer monitoring, network flow parameters were monitored with a special purpose application, and the participants rated their subjective perception of the streaming quality via a web-based questionnaire.

Bringing together all three logs, as initially intended, resulted in a very small data set because of the unreliability of some participants and the partly missing usage of the network monitoring application. This means that not only for future research studies, but also for practical operation, efforts should be undertaken to have all monitoring unified into one application. Nevertheless, also the results for the incompletely logged streaming sessions provided valuable insights. First, we observed that the YouTube mobile streaming starts on a rather low video quality level (i.e., resolution), which results in short initial delays. The adaptation logic is very conservative avoiding too many quality changes, but generally tends to improve the video quality level if the network conditions permit. Second, it became clear that the network measurements are not sufficient for an accurate QoE estimation and cannot be directly used to infer the resulting QoE. However, some network flow parameters, especially average or maximum flow throughput, are closely linked to the performance of the streaming. In contrast, other parameters, like signal strength, show only little correlations to the streaming parameters. Finally, the study revealed that the streaming parameters, like stalling times and times on quality layers, provide much better insights into the subjective experience of users than the network flow parameters. They show high correlations to the subjective experience rated by the participants, which confirms previous QoE studies.

The field study practically tested the usage of tools like YoMoApp for QoE monitoring. The app proved to be able to passively, non-intrusively measure valuable streaming parameters on application layer. Thus, future streaming applications could be equipped with such monitoring to gain better insights into the user's QoE. Still, a holistic QoE model has to be developed, which takes these monitored parameters into account and allows for an accurate QoE estimation.

ACKNOWLEDGMENT

This work was partly funded in the framework of the EU ICT projects mPlane (FP7-2012-ICT-318627) and INPUT (H2020-2014-ICT-644672), and was partly performed within the project ACE 3.0 at the Telecommunications Research Center Vienna (FTW). The authors alone are responsible for the content.

REFERENCES

[1] YouTube, "Statistics," Accessed: January 11, 2016. [Online]. Available: https://www.youtube.com/yt/press/statistics.html

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

[3] P. Casas, M. Seufert, and R. Schatz, "YOUQMON: A System for On-line Monitoring of YouTube QoE in Operational 3G Networks," *ACM SIGMETRICS Performance Evaluation Review*, vol. 41, 2013.

[4] F. Wamser, M. Seufert, P. Casas, R. Irmer, P. Tran-Gia, and R. Schatz, "YoMoApp: a Tool for Analyzing QoE of YouTube HTTP Adaptive Streaming in Mobile Networks," in *Proceedings of the European Conference on Networks and Communications (EuCNC)*, Paris, France, 2015.

[5] T. Hoßfeld, R. Schatz, M. Seufert, M. Hirth, T. Zinner, and P. Tran-Gia, "Quantification of YouTube QoE via Crowdsourcing," in *Proceedings of the International Workshop on Multimedia Quality of Experience - Modeling, Evaluation, and Directions (MQoE)*, Dana Point, CA, USA, 2011.

[6] R. K. P. Mok, E. W. W. Chan, X. Luo, and R. K. C. Chan, "Inferring the QoE of HTTP Video Streaming from User-Viewing Activities," in *Proceedings of the ACM SIGCOMM Workshop on Measurements Up the STack (W-MUST)*, Toronto, ON, Canada, 2011.

[7] T. Hoßfeld, S. Egger, R. Schatz, M. Fiedler, K. Masuch, and C. Lorentzen, "Initial Delay vs. Interruptions: Between the Devil and the Deep Blue Sea," in *Proceedings of the 4th International Workshop on Quality of Multimedia Experience (QoMEX)*, Yarra Valley, Australia, 2012.

[8] J. Yao, S. S. Kanhere, I. Hossain, and M. Hassan, "Empirical Evaluation of HTTP Adaptive Streaming Under Vehicular Mobility," in *Proceedings of the 10th International IFIP TC 6 Networking Conference: Networking 2011*, Valencia, Spain, 2011.

[9] B. Lewcio, B. Belmudez, A. Mehmood, M. Wältermann, and S. Möller, "Video Quality in Next Generation Mobile Networks – Perception of Time-varying Transmission," in *Proceedings of the 2011 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)*, Naples, FL, USA, 2011.

[10] T. Hoßfeld, M. Seufert, C. Sieber, and T. Zinner, "Assessing Effect Sizes of Influence Factors Towards a QoE Model for HTTP Adaptive Streaming," in *Proceedings of the 6th International Workshop on Quality of Multimedia Experience (QoMEX)*, Singapore, 2014.

[11] H. Knoche, J. D. Mccarthy, and M. A. Sasse, "How Low Can You Go? The Effect of Low Resolutions on Shot Types in Mobile TV," *Multimedia Tools and Applications*, vol. 36, no. 1-2, pp. 145–166, 2008.

[12] P. Casas, R. Schatz, F. Wamser, M. Seufert, and R. Irmer, "Exploring QoE in Cellular Networks: How Much Bandwidth do you Need for Popular Smartphone Apps?" in *Proceedings of the 5th ACM SIGCOMM Workshop on All Things Cellular: Operations, Applications and Challenges (ATC)*, London, UK, 2015.

[13] J. J. Ramos-Muñoz, J. Prados-Garzon, P. Ameigeiras, J. Navarro-Ortiz, and J. M. López-Soler, "Characteristics of Mobile YouTube Traffic," *IEEE Wireless Communications*, vol. 21, no. 1, pp. 18–25, 2014.

[14] I. Ketykó, K. De Moor, T. De Pessemier, A. J. Verdejo, K. Vanhecke, W. Joseph, L. Martens, and L. De Marez, "QoE Measurement of Mobile YouTube Video Streaming," in *Proceedings of the 3rd Workshop on Mobile Video Delivery (MoViD)*, Florence, Italy, 2010.

[15] G. Gómez, L. Hortigüela, Q. Pérez, J. Lorca, R. García, and M. C. Aguayo-Torres, "YouTube QoE Evaluation Tool for Android Wireless Terminals," *EURASIP Journal on Wireless Communications and Networking*, vol. 164, 2014.

[16] P. Casas, B. Gardlo, M. Seufert, F. Wamser, and R. Schatz, "Taming QoE in Cellular Networks: from Subjective Lab Studies to Measurements in the Field," in *Proceedings of the 11th International Conference on Network and Service Management (CNSM)*, Barcelona, Spain, 2015.

[17] "Mobiperf, Measuring Network Performance on Mobile Platforms," Accessed: January 11, 2016. [Online]. Available: https://sites.google.com/site/mobiperfdev/home

[18] A. Nikravesh, H. Yao, S. Xu, D. Choffnes, and Z. M. Mao, "Mobilyzer: An Open Platform for Controllable Mobile Network Measurements," in *Proceedings of the 13th International Conference on Mobile Systems, Applications and Services (MobiSys)*, Florence, Italy, 2015.

[19] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson, "Netalyzr: Illuminating the Edge Network," in *Proceedings of the Internet Measurement Conference (IMC)*, Melbourne, Australia, 2010.

[20] F. Wamser, M. Seufert, P. Casas, R. Irmer, P. Tran-Gia, and R. Schatz, "Understanding YouTube QoE in Cellular Networks with YoMoApp – a QoE Monitoring Tool for YouTube Mobile," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom)*, Paris, France, 2015.

[21] V. Aggarwal, E. Halepovic, J. Pang, S. Venkataraman, and H. Yan, "Prometheus: Toward Quality-of-Experience Estimation for Mobile Apps from Passive Network Measurements," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications (HotMobile)*, Santa Barbara, CA, USA, 2014.

[22] Q. A. Chen, H. Luo, S. Rosen, Z. M. Mao, K. Iyer, J. Hui, K. Sontineni, and K. Lau, "QoE Doctor: Diagnosing Mobile App QoE with Automated UI Control and Cross-layer Analysis," in *Proceedings of the Internet Measurement Conference (IMC)*, Melbourne, Australia, 2014.

[23] A. Balachandran, V. Aggarwal, E. Halepovic, J. Pang, S. Seshan, S. Venkataraman, and H. Yan, "Modeling Web Quality of Experience on Cellular Networks," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom)*, Paris, France, 2015.

[24] M. Z. Shafiq, J. Erman, L. Ji, A. X. Liu, J. Pang, and J. Wang, "Understanding the Impact of Network Dynamics on Mobile Video User Engagement," in *Proceedings of the ACM SIGMETRICS*, Austin, TX, USA, 2014.