

Quality of Experience on Smartphones: Network, Application, and Energy Perspectives

Selim Ickin¹, Markus Fiedler²,
Ericsson, Sweden¹
Blekinge Institute of Technology, Sweden²

Email: selimickin@gmail.com, markus.fiedler@bth.se

Abstract—For service and mobile operators, it is important to monitor and keep high user engagement levels. Quality of Experience (QoE) on video streaming applications is an important engagement measure for video consumer customers. In this paper, video QoE (with the focus on stalling events) is studied from network, application, and energy perspectives with various instrumentations on a smartphone. This enables the understanding of inter-relation between the perspectives and also how they influence the video QoE. Results show that packet delay variation and the maximal burst size in the network level; inter-picture times in the application layer; and also fluctuations in the energy consumptions are strong indicators for QoE. Enabled by extensive QoE experiments and energy measurements on smartphones, we obtain a set of telling QoE models capturing the impact of jitter and freezes, and quantifying the insights that energy consumption can be both reduced or increased in the case of stalling events during a video stream.

Index Terms—QoE, QoS, Energy, Mobile, Video

I. INTRODUCTION

Mobile video streaming traffic has exceeded 50% of the world's mobile data traffic in 2012, and it is expected to be three-fourths of the world's mobile data traffic by 2019 [1] thanks to the 4-and-beyond-G radio access technologies, which have raised the perceived quality of video streaming applications and services. The degree of delight or the annoyance, here of a user of a particular video streaming service, is named *Quality of Experience* [2]. There is high competition amongst the operators with the aim to enable the highest QoE levels on the used video streaming services, which is important to increase revenues. QoE is related to many influential factors including the network, application, and energy on the end user device. The influential factors of QoE from different perspectives as well as the inter-relation amongst them needs to be well understood, in order to actuate high QoE levels, which is typically done on network level by ISPs.

In this paper, we present important findings in our previous work [6] [7] [4] [9]. Some of the influential factors in smartphone-based video QoE are studied from different perspectives such as network, application, and energy. We also discuss the inter-relation in-between different layers in the Internet stack. As an analogy, one can imagine a heavy traffic jam at a highway during a peak hour such that the vehicles

consume fuel at a stand-still state. This is a situation where the cars barely move, and at the same time consume fuel as the engines are still running. The same applies to a video streaming application on a smartphone in communication networks domain; the video streams are stalled occasionally due to the Internet packet latency in the mobile network caused often by coverage issues or heavy load, e.g., a high number of users in the mobile network cell. This might eventually increase energy consumption at the end user's side. This paper focuses on the stalling, as previous research shows that it is one of the most important influential factor on video QoE [13]. In this paper, we study the energy perspectives of QoE together with the network and the application as all of these have complex inter-relationships between each other.

The paper is structured as follows. In Section II, we present some of the important influential factors belonging to aforementioned perspectives. Sections III-V present instrumentation for and modeling of QoE, for mobile video streaming, from the network, application, and the energy perspectives, respectively. The conclusive remarks from our studies relating the QoE to network, application and energy are given in Section VI.

II. INFLUENTIAL FACTORS ON MOBILE VIDEO QOE

The degradation of QoE levels in cellular-based video streaming is often, although not necessarily, caused by degradation of QoS level in the radio network level. A high packet latency caused by a heavy load in the network cell influences the delivery of video data, e.g., the received throughput on the smartphone. For example, the choice of a transmission protocol might influence the energy consumption when there is a problem in the network. If a video packet has not been received within a particular time window, then the video packets are re-transmitted from the source video server, which in turn might cause video packets to be accumulated in large queues at the radio link, e.g., in the base stations. This impacts QoE indirectly in many aspects including user's monthly data usage offered by the network operator, the presentation of the video content to the user in the video streaming application, and the energy consumption of the device due to increased duration of the cellular network module's active state. Increased mobile data usage caused by re-transmission of video data may impact QoE as it may increase the monthly data cost for customers. The presentation of the video content to the end user through

The work of Selim Ickin has been conducted at Blekinge Institute of Technology.

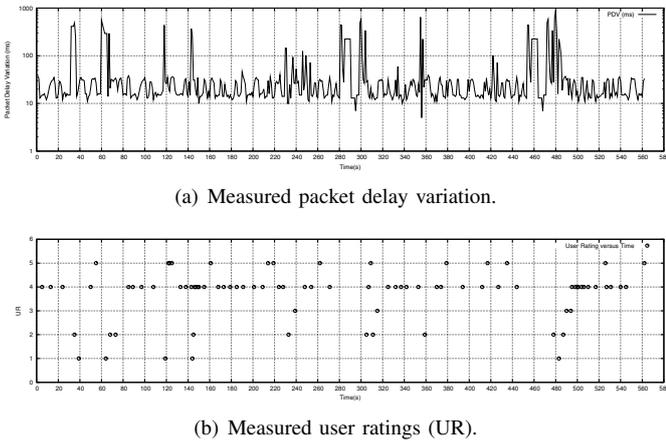


Fig. 3. Timeseries illustration of the video packet delay variation measured at the mobile video client device. [6]

against the current and the previous QoS metrics. Thus, we imitated the human perception by using EWMA on PDV values [11].

We used a user interface similar to Fig. 4(a), and asked each user to rate the video quality by using one of the five buttons located in the user interface at own will while the video is being streamed. There were in total 15 participants in the study, and they were asked to watch the video as they have being watching it in daily life settings.

Various video packet queues on the way from the streaming server to the video player might cause the PDV. Moreover, the 3G operator contributes to increases in PDV by attempting to compensate packet loss by retransmissions. In addition, TCP, while trying to recover from lost packets, might amplify the PDV. This might yield bursty packet traffic causing stallings. For this reason, we study the packet delay variation in another metric called Maximal Burst Size (MBS). This measures the amount of packets being received at a smartphone terminal in a short time interval. We assumed that this way the abnormal bursty behaviour of the video streaming traffic via 3G can better be captured.

First, we observe that the UR obtained at a given time is strongly impacted by the previously measured PDV values as expected. The power-law model fitted better (with a R^2 value increased by 0.1) than an exponential IQX-type one [5]. Furthermore, the goodness-of-the-fit value R^2 for the power model is improved by over 100 % when the EWMA technique is applied. We obtained the models for the relation between the PDV and the MOS that is strengthened via EWMA as given with Eq. 4.

$$UR = -9.10 (PDV_{EWMA}/ms)^{0.08} + 16.18, \quad (4)$$

The relation between the UR and the MBS is also studied, and observed that the peaks in the number of transmitted packets per given time interval have a negative impact on the UR. This can be explained by the fact that large MBS values indicate a *ON/OFF* behavior. In this case, when there is no available bandwidth the packets are queued, and then flushed all at once,

causing a stall in the video streaming application. Eq. 5 shows the obtained model with respect to the MBS and the MOS with a better ($R^2 = 0.78$) goodness-of-fit as compared to Eq.4 ($R^2 = 0.68$).

$$UR = 59.96 MBS_{[packets/ms]}^{-0.036} - 51.71 \quad (5)$$

In order to find out the influential factors on video QoE, we investigate at a level in the Internet stack that is closest to the user. Typically, the user interface level is where a particular service is sensed and experienced by a human user.

IV. QOE STUDY IN THE APPLICATION

In the application, we consider the user interface, the point where the user interacts with a video streaming application/service. Measurement points are deployed at relevant parts of the open-source VLC video player application code base such as the recording of timestamps when the video frame is rendered and displayed on the video screen. This modified version is called VLQoE. Then, the inter-frame (or inter-picture) time during a video stream are computed. The deviation of the inter-picture times are then matched to the perceived video QoE. We measured and quantified QoE with opinion scores that are collected at user's own will while the video is being streamed. We conducted the measurements on an Android device.

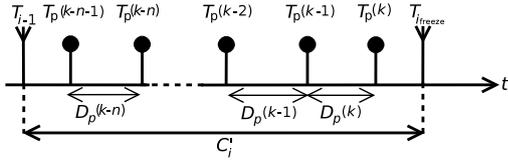
The users in this study are asked to watch a 250 seconds long video clip that consists of various scenes including racing scenes of sailing boats. The video clip contains a sequence of 6251 pictures and is encoded with a nominal frame rate of 25 fps with a bitrate of 1000 kbps. The same video is watched twice (first with RTSP, then with Hypertext Transfer Protocol (HTTP)) by each user on a smartphone (with a video screen size of 196×117 pixels, which is provided to them. Each user was encouraged to rate the temporal quality based on the five-level MOS scale, while pressing one of the five user rating buttons at her/his own will during the ployout. In addition, a 'freeze' button is horizontally placed on top of the five buttons at the user interface. The user interface used in the experiments is given in Fig. 4(a).

30 subjects performed the study at various location in Karlskrona, Sweden by using exactly the same smartphone. In total, 60 ($= 30 \text{ users} \times 2 \text{ protocols}$) user experiments were conducted. We asked the users to hold the smartphone at a comfortable distance with convenient illumination level, *i.e.*, at a familiar physical context as in daily life. This way, we imitated peoples' natural daily life settings in the subjective tests. The video was muted such that the users could focus only on the visual freezes. A visual freeze can be measured at the user interface with an inter-picture time D_p , as the time gap between two consecutive pictures displayed on the smartphone screen. This is calculated in Eq.6. $T_p(k)$ is the timestamp when the k^{th} picture is displayed on the smartphone screen. An illustration of the D_p metric during a video stream is given in Fig. 4(b).

$$D_p(k) = T_p(k) - T_p(k-1) \quad (6)$$



(a) VLQoE tool user interface.



(b) Inter-picture time measurements.

Fig. 4. (a) presents the user interface of VLQoE tool; (b) illustrates the inter-picture time measurement. [7]

The distributions of the inter-picture time values for the user ratings 1 – 5, as well as freeze indications collected from all users, are given in Fig. 5. The corresponding mean D_p values for ‘UR 5’, ‘UR 4’, ‘UR 3’, ‘UR 2’, ‘UR 1’, and ‘freeze’ are 152 ms, 282 ms, 321 ms, 768 ms, 831 ms, and 1289 ms, respectively. Van Kester *et al.* [10] state the acceptable freezing duration as 360 ms, which also confirms our results when considering an acceptable UR of 3.

By using the VLQoE tool, we assumed that a 3G video stream follows a two-state ON/OFF model. The *ON* (smooth playout) and *OFF* (freeze) states are set based on the D_p metric. We considered the state-of-the-art 100 ms as the maximum tolerance threshold for a user to feel that a system reacts instantaneously [3]. Then, D_p values less than 100 ms are assigned to an ON state; while the D_p values higher than 100 ms are assigned to an OFF state. The state durations were modeled by exponential distributions. The mean Maximum Likelihood Estimate (MLE) of ON and OFF durations for all 58 iterations were calculated as 9.7 s and 642 ms, with mean R^2 values of 0.81 and 0.93, respectively. This model will be applied for controlled local-based video streams in the user tests as described in Section IV.

V. ENERGY AS AN OBSERVABLE METRIC FOR QOE INDICATION

Amongst many other influential factors, energy (and particularly the remaining battery level of a smartphone) is one of the most important ones that influence the overall QoE of a smartphone user [4]. And energy is highly consumed in video streaming applications due to (1) the high bandwidth demand,

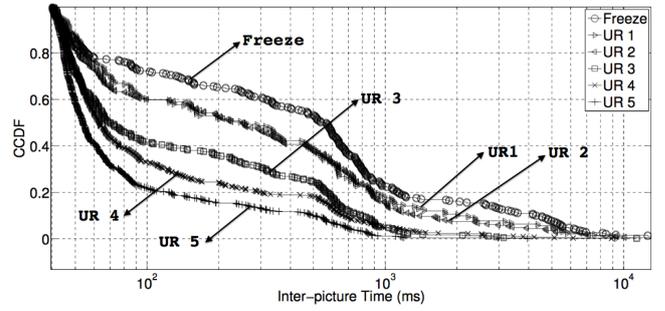


Fig. 5. The distribution of inter-picture times in between two user indications are given for all five UR scores. [7]

keeping the network module in its active state for long durations; (2) high CPU utilization for processing/rendering the multimedia packets; and (3) good-enough screen illumination for a clear presentation of video content on a battery-powered smartphone display. For these reason, it is important to understand the energy consumption patterns during a video stream and how these patterns are influenced when there are video quality issues such as stalling events.

By using the VLQoE tool, the inter-picture time during a video stream is measured. In parallel, the Monsoon energy measurement tool is used to measure the total instantaneous power consumption of a smartphone terminal. Then, the relation between the power measurements and the freezes are observed. Fig. 6 illustrates a snapshot, during a video stream, of the relation between the inter-picture time and the total power consumption of smartphone. Phase 1 is the low power state, can also be called as the initial rebuffering state where no pictures are yet displayed on the screen, of the video stream. Phase 2 is the steady playout state. It is observed that the high inter-picture OFF time causes a slight reduction in the power consumption values. The reason for this is that the video stalls, as there are no packets to present on the smartphone screen. Thus, there is a direct relationship between the inter-picture time and the duration of the power consumption staying at a rather lower level.

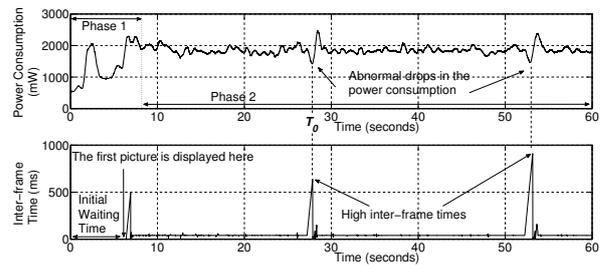


Fig. 6. Illustration of the power consumption and the inter-picture time values during video streaming. [8]

Indeed, Fig. 6 is a high-level sketch of a smartphone’s power consumption pattern when there is a stalling event during a video stream. We hypothesize that freezes actually can cause either an energy saving or an energy waste depending on the

transmission protocol. In the case of a freeze during a video stream, if there is a packet retransmission initiated by the underlying Transport Control Protocol (TCP) connection, then all the video content is shown to the user with an extended view duration, which implies an increased energy consumption. During the freeze time, energy is still being consumed by CPU, AMOLED display, and the current power state of the radio data module. Thus, the energy consumed during the stall duration must be considered as *wasted*. Therefore, the smoother the video playout, the lower the total energy consumption and the higher the QoE.

In contrast, if there is no retransmission to compensate the lost packets, such as in video streams established over UDP, then there is a picture jump, which causes some content to be skipped, but leaves the total video view duration unchanged. The energy drop during a freeze duration can be considered as *saved* energy. However, the freezes influence the QoE of a user during a video stream in an unfortunate way. Thus, in the case of a video transmission with skipped pictures, there is a tradeoff between the energy saving and the MOS. It is important to find out the maximum saved energy without impacting QoE. For the case of video streams, where there is no picture jump, the waste of energy increases with the duration of a freeze. Thus, for the latter scenario, it can be said that the better the streaming quality, the less energy is wasted. The two scenarios are illustrated in Fig. 7.

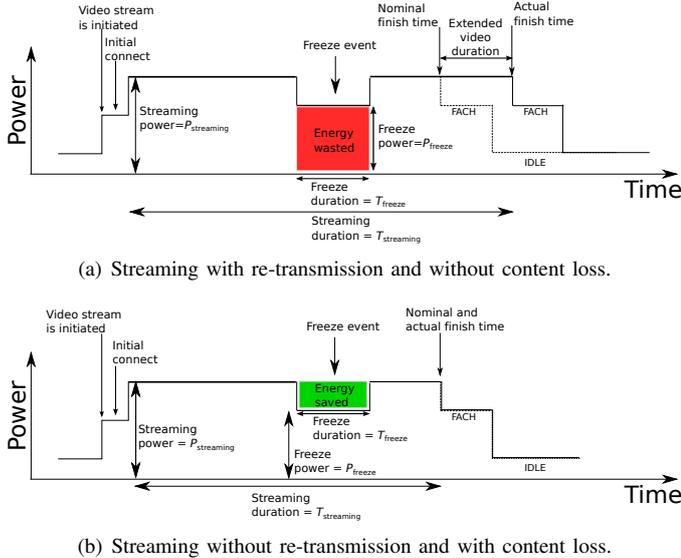


Fig. 7. Illustration of the power consumption on a smartphone with freeze scenarios. [9]

Next, the amount energy waste and the energy savings for the two scenarios are computed by considering QoE models. The procedure that is followed during the video experiments is given in Fig.8. Each user is asked to watch three versions of the same video content in a raw, and are asked to give a MOS score at the end of each video. The film clip was three minutes long, with a 6 Mbps bitrate and 25 fps, and converted into MP4 multimedia format. The video is streamed via the local drive

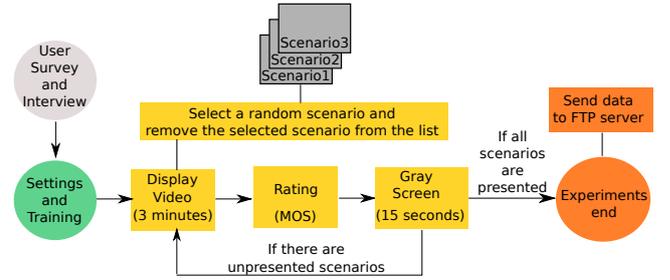


Fig. 8. Video QoE experiment procedure. [9]

of the smartphone to enable controlled experiments. We emulate the ON/OFF exponential model obtained in Section IV, which means that play and freeze durations are exponentially distributed over the video session. The mean OFF duration is set to 2 s with varying mean ON durations of 4 s, 8 s, and 16 s. The first version of the video is the original one that has no temporal distortions (scenario 1); the second version contains freezes and the video pictures are not skipped (scenario 2); and the third version also contains freezes and the corresponding video pictures are skipped (scenario 3). The three versions of the videos are shown to the user in random order, and in-between each video session, the user is asked to rate the video quality of the previous film clip. After the user registers the MOS score, she is asked to watch a gray screen for 15 seconds on the smartphone to reduce the memory effect. In total, 60 users were involved in the study. We computed the OFF probability P_{OFF} as the ratio of the mean OFF state duration to the sum of the mean ON and OFF state durations. Then, the relation in-between the P_{OFF} and the MOS scores are computed.

In total, it was observed that, with the presented settings in the experiment, there was no statistically significant difference in QoE, regardless of whether or not skips follow freezes. The important factor that influences the QoE is whether or not a freeze happens. Thus, we merge the data for scenario 2 and 3, *i.e.*, the scenarios involving freezes regardless of the fact that the pictures are skipped or not. When the P_{OFF} values are fitted to the MOS scores, an exponential relation is obtained in Eq. 7 with $R^2 = 0.73$:

$$MOS = 4.59e^{-3.44 \cdot P_{OFF}} \quad (7)$$

$$= 4.59e^{-3.44 \cdot \frac{T_{freeze}}{T_{video}}} \quad (8)$$

Eq. 8 can further be used to calculate the relation between the energy saving or energy waste, as both are highly related with the ratio of the total freeze duration to the whole video duration, cf. Fig. 7. The power saving during a freeze P_{saving} is measured to be 185 mW, and the total freeze duration T_{freeze} can be further substituted with $\frac{E_{saving}}{185 \text{ mW}}$, yielding the model expressed by Eq.9, where the total video duration T_{video} is three minutes in the experiment setup.

$$MOS = 4.59 \cdot e^{-18.59 \cdot \frac{E_{saving}/J}{T_{video}/s}} \quad (9)$$

With this tradeoff model between the MOS and the energy

saving, at most 4.25 J can be saved while keeping the MOS level at 3.

The power consumption during a freeze P_{freeze} is measured to be 728 mW. Applying similar calculations to model the relation between the MOS and the energy waste, Eq. 10 can be obtained showing that the MOS degrades further with the increase in the energy waste caused by a freeze:

$$MOS = 4.59 \cdot e^{-4.72 \frac{E_{waste}/J}{T_{video}/s}}. \quad (10)$$

We can conclude that the amount of energy saving in a three minutes long video stream is insignificant for a commercial smartphone with battery capacity of 9.88 Wh. Throughout extensive energy measurements on the smartphone, we identified potential scenarios where energy saving might be possible and we recommended approaches to increase energy saving while maintaining QoE. It has shown, for TCP-based streams, that a lower total energy consumption can be achieved with a smoother video playout causing in parallel a higher QoE. In contrast, for UDP-based streams, there is a tradeoff between the QoE and the energy saving, *i.e.*, the stalling events both decrease QoE and the energy consumption.

VI. CONCLUSION

In this paper, we have presented various studies on video Quality of Experience (QoE) from different perspectives such as network, application, and energy. Starting on the network level, we have identified the Packet Delay Variation (PDV) and the Maximal Burst Size (MBS) metrics as strong indicators of QoE. Although an exponential IQX-type of model represented the relation between QoE and QoS well, a power-law model was found to yield a slightly better QoE model for both PDV and MBS. A major improvement was reached by incorporating exponentially weighted moving averages into the calculation of the corresponding network-level parameter.

Turning to the application level, we quantified stalling via inter-picture times trespassing a given threshold (100 ms). The corresponding instrumentation was done by modifying the open-source VLC player for Android OS. Indeed, the inter-picture time of a 3G-based video stream could be modeled by a two-state exponential ON/OFF model with a mean OFF value of around 600 ms and a mean ON value of around 10 s, respectively. We obtain a QoE model as a function of the freeze or OFF probability P_{OFF} , measured at the user interface.

Through extensive energy measurements on the smartphone, we were able to relate QoE to energy consumption. We leveraged QoE models that the impact of stallings at the user interface, which then allowed to relate QoE to energy waste and savings on the smartphone. It was shown for TCP-based streams that a lower total energy consumption can be achieved with a smoother video playout, which comes along with higher QoE. In contrast, for UDP-based streams, there is a tradeoff between the QoE and the energy saving, *i.e.*, the stalling events both decrease QoE and the energy consumption. However, the achievable savings are in no proportion to the potential loss of QoE.

One of our main findings is that smoother video delivery yields both better QoE and better energy efficiency in case of retransmission-based video streams. The QoE of a video streaming application/service highly depends on the network level QoS metrics including the packet delay, throughput, etc. Thus, the most obvious way of monitoring and preventing poor video quality is via QoS management by an Internet Service Provider (ISP), *e.g.*, with smart scheduling and increasing bandwidth. However, due to the wide variety of video streaming applications with different streaming characteristics and over-the-top streaming protocols, it is hard to pinpoint a low QoE level by solely relating it to the QoS metrics collected in the network level. As future work, collaborative data analytics and machine learning techniques can be suggested to study many cross-layer metrics simultaneously to predict quality degradations in advance, and to actuate timely robust decisions to improve QoE and to save energy for smartphones at the same time.

VII. ACKNOWLEDGEMENT

This work is supported by the European Celtic-Plus project CONVINCe and was partially funded by Finland, France, Sweden and Turkey.

REFERENCES

- [1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2014–2019." Cisco, White Paper, 2014.
- [2] Qualinet White Paper on Definitions of Quality of Experience, Dagstuhl Seminar 12181, Version 1.1, Dagstuhl, June 2012.
- [3] S. C. Seow, "Designing and engineering time: The psychology of time perception in software," *Addison-Wesley Professional*, 2008.
- [4] S. Ickin, K. Wac, M. Fiedler, L. Janowski, J. H. Hong, and A. K. Dey, "Factors influencing Quality of Experience of commonly-used mobile applications," *IEEE Communications Magazine*, 50(4):48–56, 2012.
- [5] M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A generic quantitative relationship between Quality of Experience and Quality of Service," *IEEE NETWORK*, vol. 24, no. 2, pp 36–41, March 2010.
- [6] S. Ickin, K. D. Voegelers, M. Fiedler, D. Erman, "The effects of Packet Delay Variation on the perceptual quality of video," in *Proc. IEEE 35th Conference on Local Computer Networks (LCN2010)*, pp 663–668, 2010.
- [7] S. Ickin, M. Fiedler, K. Wac, P. Arlos, C. Temiz, K. Mkocha, "VLQoE: Video Quality of Experience Instrumentation on the Smartphone," *Multimedia Tools and Applications. Special Issue on Advances in Tools, Techniques and Practices for Multimedia QoE*, Springer US, 74(2):381–411, 2014.
- [8] S. Ickin, M. Fiedler, K. Wac, "Energy-based anomaly detection in Quality of Experience," in *Proc. 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pp 1–6, USA, 2013.
- [9] L.G. Ballesteros, S. Ickin, M. Fiedler, J. Markendahl, K. Tollmar, and K. Wac, "Energy Saving Approaches for Video Streaming on Smartphone based on QoE Modeling," in *Proc. 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 2016.
- [10] S. Van Kester, T. Xiao, R. E. Kooij, K. Brunnström, and O.K. Ahmed, "Estimating the Impact of Single and Multiple Freezes on Video Quality," *SPIE'11*, vol. 7865, 2011.
- [11] F. Guyard and S. Beker, "Towards real-time anomalies monitoring for QoE indicators," *Annals of Telecommunications*, 65(1–2):47–57, Jan/Feb 2010.
- [12] Z. Chen and H. Nakazato, "A model for time-varying quality of speech services," *IEEE Globecom*, pp. 240–244, 2005.
- [13] T. Hossfeld, M. Seufert, M. Hirth, T. Zimmer, P. Tran-Gia, and R. Schatz, "Quantification of Youtube QoE via crowdsourcing," in *2011 IEEE International Symposium on Proc. Multimedia (ISM)*, pp. 494–499, 2011.
- [14] S. Ickin, "Quality of Experience on Smartphones: Network, Application, and Energy Perspectives," Ph.D. Thesis, Blekinge Institute of Technology, 2015.