Mobile Quality of Experience: Recent Advances and Challenges

Vasilios A. Siris
Department of Informatics
Athens University of Economics and Business

Konstantinos Balampekos and Mahesh K. Marina
School of Informatics
The University of Edinburgh

Abstract—Quality of Experience (QoE) is important from both a user perspective, since it assesses the quality a user actually experiences, and a network perspective, since it is important for a provider to dimension its network to support the necessary QoE. This paper presents some recent advances on the modeling and measurement of QoE with an emphasis on mobile networks. It also identifies key challenges for mobile QoE.

I. INTRODUCTION

Quality of Experience (QoE) can be broadly defined as how a user perceives the usability or degree of satisfaction of a service. More specifically, ITU-T defines QoE as “the overall acceptability of an application or service, as perceived subjectively by the end-user” [1]. This definition is accompanied by two notes: 1) Quality of experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.) and 2) Overall acceptability may be influenced by user expectations and context. According to the Qualinet white paper [2], 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”.

Unlike QoE, Quality of Service (QoS) is the ability of the network to provide a service with an assured service level. QoS is measured in terms of throughput, packet loss, delay, and jitter. Comparing the two definitions, QoS is a network-centric characterization of the service quality, whereas QoE is a user-centric characterization of service quality.

Based on the above QoE definitions, it is clear that QoE depends on both human and technical factors which include the following:

- Application factors: Application type, e.g. VoIP, streaming video, Web access, but also the application sub-type, such as live video streaming, VoD with short clips, and VoD with full movies; application QoS such as video pauses (stalls), transaction delay, encoding changes, and application type; additionally, it can include characteristics of transport characteristics such as UDP, HTTP/TCP, etc.

- Content, such as type of video or videoconference.

- End-system (device) features such as screen resolution, interface, battery, and power consumption.

- Network-level QoS such as bitrate, packet loss, delay, and jitter.

QoE can be estimated using a subjective or objective approach. Subjective QoE estimation requires user involvement and quantifies the QoE in terms of a Mean Opinion Score (MOS), where the quality is assessed using a 5-point scale score: 5-Excellent, 4-Good, 3-Fair, 2-Poor, 1-Bad. On the other hand, objective QoE assessment estimates the QoE using a parametric model, without requiring the involvement of users. The parametric model can depend on the application, context, etc, and is a function of the network-level QoS which is typically estimated from measurements.

Key components of a QoE framework include the following:

- QoE model, which defines how QoE is quantified along with which factors influence the QoE and how.

- QoE measurement, which involves how QoE/QoS is measured or predicted.

- QoE-aware management and control.

In this paper we focus on outlining the state of the art for the first two components, namely on QoE modeling and measurement. We discuss the mobile QoE and associated challenges in greater depth, driven by two observations: (1) increasing reliance on mobile networks and the exponentially rising demand for mobile data services has made optimising mobile QoE a prime concern; (2) QoE assessment in mobile networks is relatively more challenging given their unique characteristics that include high levels of dynamism, resource constraints and diversity in terms of device characteristics, context, etc.

II. QOE MODELING AND MEASUREMENT

We begin by reviewing the state of the art on QoE modeling and measurement in fixed/wired networks.
A. Quantifying QoE

QoE can be quantified using the MOS value; this also allows correlation between subjective and objective QoE estimation methods. Another alternative is to quantify the QoE through some metric that is influenced by the user QoE; hence, rather than quantifying directly the user QoE, what is quantified is a user’s behavior or reaction to experiencing a particular level of QoE [3]. Examples of such metrics are video pausing and the reduction of the screen size [4] or the percentage of a video that is viewed, number of videos viewed, number of visits [5]; such metrics are referred to as user engagement [5].

B. Mapping of Network-Level QoS to QoE

Different types of relations between the QoS and QoE include the following [6]:

- Linear: Such a dependence suggests that an additive change of the QoS has a linear influence on the QoE.
- Logarithmic: Such a dependence suggests that a multiplicative change of the QoS has a linear influence on the QoE.
- Exponential: Such a dependence suggests that an additive change of the QoS has a multiplicative influence on the QoE.
- Power: Such a dependence suggests that a multiplicative change of the QoS has an exponential influence on the QoE.

We now mention a few examples of the above stated relation types between network-level QoS metrics and QoE for different applications. The QoE of VoIP and video streaming quality has a logarithmic dependence on the bitrate. A logarithmic dependence also exists between the QoE for video streaming and the initial delay for the video to start. The logarithmic dependence can be explained by the Weber-Fechner law, according to which the perception is proportional to the relative change of the stimulus [7]. The dependence of the QoE for VoIP on the loss follows an exponential dependence. Intuitively, this means that a disturbance (e.g. due to loss or delay) of the QoS will have a higher impact when the QoE is high, compared to when it is low [8], [9].

C. QoE Models for Specific Applications

In this section we discuss QoE models for some specific applications/services, such as web browsing, voice/audio, video, online multiplayer gaming, and telepresence.

1) Web browsing: For web browsing, ITU-T Recommendation G.1030 specifies a model where the QoE decreases linearly with the logarithm of the session time, which includes the time to search and download a web page. The specific function depends on the expected maximum and minimum session time, which are determined by the network access context (e.g., fast, medium, and slow network access speed). The authors of [9] show that both a decreasing logarithmic, as specified in G.1030, and an exponential dependence show a good fit to experimental data. Another important result for web browsing is that temporal correlations (or memory effects) significantly influence the QoE [10], hence QoE models should be time-dependent.

2) Audio applications: For audio, the PESQ (Perceptual Evaluation of Speech Quality, specified in ITU-T Recommendation P.862) defines a test methodology for assessing the speech quality by comparing the degraded speech signal with the original speech signal, hence is a full-reference method. PESQ does not take into account delay and echo, which influence voice quality when it is transferred across a network. Unlike PESQ, the E-model (specified in ITU-T Recommendation G.107) considers network metrics such as delay, echo, losses, and signal-to-noise ratio in order to estimate the call quality. A drawback of the E-model is that it requires knowledge of an “impairment factor”, which depends on the specific codec used. For VoIP traffic and in particular for the Internet low bit rate codec (iLBC) used by Skype and the G.711 codec, both the packet loss ratio and the reordered ratio influence the QoE in an exponential manner [9]. Moreover, experiments have shown that the QoE degradation caused by increased packet loss is significantly higher than the QoE degradation from reduced bandwidth through controlled adaptation, such as the selection of a codec with lower bit rate [11]. Indeed, this result also holds for both video conferencing and web browsing [11].

3) Video applications: For video, the Peak Signal-to-Noise-Ratio (PSNR) and Mean Squared Error (MSE) are full-reference techniques for measuring video quality. On the other hand, the Video Quality Model (VQM, specified in ITU-T Recommendation G.1070), estimates the video quality based on the type of video/audio codecs, packet delay and loss, video frame and bit rate. Similar to VoIP, the QoE of video conferencing also exhibits a higher influence from packet loss than compared to the encoding bit rate — the QoE decreases exponentially with an increasing packet loss, whereas it increases logarithmically with the encoding bit rate.

In recent years, multimedia content delivery is increasingly based on HTTP. HTTP-based video streaming systems include Apple’s HLS, Microsoft’s Smooth Streaming and Adobe’s Flash Video, and YouTube. Unlike RTSP (Real Time Streaming Protocol)UDP-based streaming used by systems such as Windows Media and Real Media, TCP used in HTTP streaming guarantees packet delivery, thus video quality is not degraded (directly) due to packet loss and reordering. On the other hand, network congestion can result in the video player not receiving data in time for playback, which leads to frame stalls (also referred to as frame pauses or re-buffering events). An approach to reduce frame stalls is to initially buffer some amount of the video stream before starting playback. Due to the above, the factors that directly influence the QoE of HTTP video streaming are the initial buffering delay, the duration of stall events, and the frequency of stall events. Among these three factors, the frequency of stall events appears to be the dominant factor influencing the QoE, while the initial delay influences the QoE least [12], [13]. Indeed, the QoE exhibits an exponential dependence on the frequency of stalls and it is not enough to consider the total stall time, since both the number and the duration of stall events influence the QoE [14]. Moreover, for YouTube video streaming an initial buffering delay up to 16 seconds has no or a very small impact on the QoE [14]. Finally, while user demographics were found not to influence the QoE ratings, the QoE depends on the video duration and stalling pattern [15].

In this section we discuss QoE models for some specific applications/services, such as web browsing, voice/audio, video, online multiplayer gaming, and telepresence.

1) Web browsing: For web browsing, ITU-T Recommendation G.1030 specifies a model where the QoE decreases linearly with the logarithm of the session time, which includes the time to search and download a web page. The specific function depends on the expected maximum and minimum session time, which are determined by the network access context (e.g., fast, medium, and slow network access speed). The authors of [9] show that both a decreasing logarithmic, as specified in G.1030, and an exponential dependence show a good fit to experimental data. Another important result for web browsing is that temporal correlations (or memory effects) significantly influence the QoE [10], hence QoE models should be time-dependent.

2) Audio applications: For audio, the PESQ (Perceptual Evaluation of Speech Quality, specified in ITU-T Recommendation P.862) defines a test methodology for assessing the speech quality by comparing the degraded speech signal with the original speech signal, hence is a full-reference method. PESQ does not take into account delay and echo, which influence voice quality when it is transferred across a network. Unlike PESQ, the E-model (specified in ITU-T Recommendation G.107) considers network metrics such as delay, echo, losses, and signal-to-noise ratio in order to estimate the call quality. A drawback of the E-model is that it requires knowledge of an “impairment factor”, which depends on the specific codec used. For VoIP traffic and in particular for the Internet low bit rate codec (iLBC) used by Skype and the G.711 codec, both the packet loss ratio and the reordered ratio influence the QoE in an exponential manner [9]. Moreover, experiments have shown that the QoE degradation caused by increased packet loss is significantly higher than the QoE degradation from reduced bandwidth through controlled adaptation, such as the selection of a codec with lower bit rate [11]. Indeed, this result also holds for both video conferencing and web browsing [11].

3) Video applications: For video, the Peak Signal-to-Noise-Ratio (PSNR) and Mean Squared Error (MSE) are full-reference techniques for measuring video quality. On the other hand, the Video Quality Model (VQM, specified in ITU-T Recommendation G.1070), estimates the video quality based on the type of video/audio codecs, packet delay and loss, video frame and bit rate. Similar to VoIP, the QoE of video conferencing also exhibits a higher influence from packet loss than compared to the encoding bit rate — the QoE decreases exponentially with an increasing packet loss, whereas it increases logarithmically with the encoding bit rate.

In recent years, multimedia content delivery is increasingly based on HTTP. HTTP-based video streaming systems include Apple’s HLS, Microsoft’s Smooth Streaming and Adobe’s Flash Video, and YouTube. Unlike RTSP (Real Time Streaming Protocol)UDP-based streaming used by systems such as Windows Media and Real Media, TCP used in HTTP streaming guarantees packet delivery, thus video quality is not degraded (directly) due to packet loss and reordering. On the other hand, network congestion can result in the video player not receiving data in time for playback, which leads to frame stalls (also referred to as frame pauses or re-buffering events). An approach to reduce frame stalls is to initially buffer some amount of the video stream before starting playback. Due to the above, the factors that directly influence the QoE of HTTP video streaming are the initial buffering delay, the duration of stall events, and the frequency of stall events. Among these three factors, the frequency of stall events appears to be the dominant factor influencing the QoE, while the initial delay influences the QoE the least [12], [13]. Indeed, the QoE exhibits an exponential dependence on the frequency of stalls and it is not enough to consider the total stall time, since both the number and the duration of stall events influence the QoE [14]. Moreover, for YouTube video streaming an initial buffering delay up to 16 seconds has no or a very small impact on the QoE [14]. Finally, while user demographics were found not to influence the QoE ratings, the QoE depends on the video duration and stalling pattern [15].
For Scalable Video Coding streaming, such as streaming using the MPEG-DASH (Dynamic Adaptive Streaming over HTTP) standard, the rate of transitions influences the QoE, and users prefer gradual rather than abrupt quality changes and lower initial quality to avoid drops in the quality [16]. Other studies have shown that both content and the type of streaming, i.e. live or video-on-demand, influences the QoE expressed through user engagement. Specifically, live content is influenced more by the duration of stalling and by the average bit rate [17].

4) Online multiplayer gaming: Two key properties of online multiplayer gaming is that they can involve multiple types of live multimedia, such as video, audio, and voice, and multiple players in different locations. These two properties yield the following two application quality factors: interactivity and consistency [18], [19], [20]. Interactivity involves the responsiveness to actions from specific players. Specific QoS metrics that influence interactivity include end-to-end network delay and jitter, and request/response latency and jitter; indeed, for First Person Shooter type games, experiments show that jitter can influence the QoE more than delay [21]. Consistency includes temporal and spatial consistency. Temporal consistency involves the time synchronization of different multimedia streams, whereas spatial consistency involves maintaining the same state in different (player) locations. In addition to the interactivity and responsiveness factors, other quality factors that pertain to the quality of individual multimedia types (such as video quality) also influence the QoE of online gaming.

5) Telepresence: Similar to multiuser online gaming, telepresence involves multiple types of live media and multiple users. Hence, the QoE of telepresence includes the interactivity and consistency quality factors. Additionally, because telepresence involves communication of data from multiple sensors, an additional quality factor is vividness [22], [23], which refers to the richness of the mediated environment, and can include both the number of sensory channels (breadth) and the resolution of each channel (depth). As expected, how each quality factor influences the QoE depends on the specific telepresence application; single-value QoE scores are not appropriate since they do not account for the smooth variation of individual quality factors, which are important since quality flickers can significantly impact a user’s overall QoE [24].

D. QoE Measurement

The QoE model can have variables that reflect quality indicators (or factors), such as network and application QoS factors, and parameters that depend on user, system, content, and context factors. To apply the model, the values of the parameters need to be first determined. This can be a complex task, due to their large number. Moreover, the quality indicators and the other factors influencing the QoE may have complex and sometimes counter-intuitive relationships with the QoE, such as the non-monotonic dependence of engagement QoE metrics on the bitrate, hence encoding level; and different quality factors can have complex interdependencies [17]. Additionally, requirements for on-line and real-time estimation of QoE levels place further constraints on the QoE measurement process. On-line and real-time QoE estimation is necessary when QoE levels are used for managing network resources and application tuning (e.g., codec selection). To address the aforementioned complex relationships and interdependencies, simple procedures such as linear regression and naive Bayes algorithms, are not sufficient. Instead, machine learning procedures such as multiple/non-linear regression and decision trees [5] can be more appropriate.

Another issue that influences QoE estimation is where and how the quality indicators are measured; this depends on whether the indicators refer to network QoS or application QoS metrics. Network QoS metrics can be estimated at nodes inside the network. On the other hand, application QoS metrics need to be measured at the application level, hence at the client side, and may require modifications to the application. Examples of client-side software tools for measuring the QoE of YouTube streaming is YoMo [25] and Pytomo [26]. Related work proposing alternate approaches that simulate or emulate the client/application is presented in [27], [28], [29]. Such approaches allow the corresponding measurement procedures to be implemented inside network nodes, hence allow QoE assessment using application QoS estimation to be performed by network operators.

III. QoE IN MOBILE NETWORKS

In this section we present some recent results on the QoE for mobile networks.

A. Importance and Benefits

The penetration of smartphones in the mobile market as well as the continuous evolution of mobile networks has led to an explosion of mobile data traffic in recent years. Current mobile networks offer their users the opportunity to stream multimedia, use social networks, shop online, etc. Furthermore, with the ongoing rollout of 4G LTE networks in many countries, it is estimated that there will be a 13-fold increase of global mobile data traffic during the five-year period of 2012-2017 [30], making QoE measurement and QoE-driven adaptation even more crucial.

Mobile network operators typically use theoretical models and in-field measurements during network planning activities to optimize network coverage and performance. However, due to the intrinsic nature of mobile networks, which is dominated by the air interface, their performance is location and time dependent. As a result, there may exist non-trivial deviations between “expected” and user-perceived performance. Traditionally, the provided performance of mobile services is expressed using QoS-related metrics such as throughput, latency and jitter. However, the application-oriented use of mobile networks necessitates a more user-centric assessment of service delivery. For instance, it is challenging to assess the quality of video content delivery through usual QoS measurements, as it would also depend on other factors such as video frame rate and codec. QoE bridges that gap.

The main benefits of QoE measurement in mobile networks are the following:

- It captures mobile experience as perceived by the users. Real user experience in mobile networks can be largely different than the “expected” performance
and the performance indicated by network-level QoS metrics.

- It provides insight on the factors that influence customer satisfaction. Such factors (e.g., application interface, device features, etc.) do not necessarily depend on the underlying network.
- It takes into account the broad range of network usage profiles. For some users making and receiving voice calls is the most influential factor, while for others data services are equally or more important.

B. Requirements

QoE is attracting increasing amount of attention from the various stakeholders of the mobile networks ecosystem (i.e., consumers, operators and regulators). For instance, Ofcom, the UK’s regulator of communications, published a Call For Input (CFI) seeking suggestions from relevant stakeholders on the metrics that should be measured and the way that these should be measured for QoE assessment [31]. The CFI received responses from a wide range of stakeholders [32]. Recent research has also tried to address the lack of a standardized QoE measurement framework [33], [34], [35].

All of the above mentioned sources agree on the fact that collecting only subjective values (e.g., MOS) in controlled environments is neither adequate nor efficient. Instead, a QoE measurement framework should consider a combination of qualitative and quantitative metrics:

- **QoS metrics.** This set of metrics includes network performance metrics (e.g., bit rate, latency, etc.), radio interface metrics (e.g., received signal strength, bit error rate, etc.) as well as device information (e.g., OS, device features, etc.).

- **Context information.** This group includes location information, which apart from the exact user location, it refers to more refined information like indoor/outdoor location, device orientation, etc. Measurements from different built-in device sensors aid in inferring the context.

- **User behavior information.** Different users have different network/device usage patterns as well as different requirements from the mobile network. Therefore, this set of information includes metrics like the proportion of time that a user experiences 3G connectivity, the number of times a user launched a particular application, etc. With the help of this information, it is possible to create user groups which can prove useful for QoE comparison.

- **Subjective experience information.** Certainly, user-originated information is helpful for a reliable and complete QoE assessment. For example, a trigger could ask the users to rate their experience, ideally immediately after a service usage.

C. QoE Measurement Approaches

There are various ways by which QoE of mobile networks can be measured and assessed:

- **Network-side passive monitoring** is also used by operators for performance assessment and network management by monitoring traffic flows. A large amount of network metrics are collected in real time, which allows network management for optimal user satisfaction. Although this is a QoS-oriented measurement method, recent research proposes a more user-centric approach. For example, an operator can take into account the percentage of user aborts due to long web response times to assess the acceptability of a specific response time [36]. On the other hand, this method cannot take into account the user’s context nor can it collect data for areas with no coverage (not-spots). Also note that this approach estimates user experience via indirect measurements.

- **Drive testing** is a method commonly used by operators for in-field network measurement campaigns in an effort to emulate user experience. Main advantage of this method are the wide range of performance metrics and the more user-centric approach compared to network-side passive monitoring. However, drive testing is not cost-efficient or real time. Additionally, it does not capture user behavior neither does it take into account users context, as typically drive testing measurements are collected only outdoors and that too in certain select locations.

- **Crowd-sourcing** is an emerging measurement method which relies on user participation in the measurement process (e.g., [37]). In the crowd sourcing approach, distributed measurement agents run on end-user devices and perform active, passive or both types of measurements in real time. It is a scalable, cost-effective approach that captures real end-user experience. Furthermore, current smartphones are equipped with enough sensors to get fine grained context information. Collecting measurement data from the end-user side also allows usage profiling for user behavior monitoring. A drawback of this method is that the measurement agents are restricted by the OS APIs. Also, incentivizing the users for contribution can prove challenging.

D. Current State-of-the-Art

QoE assessment and measurement in mobile networks has attracted significant attention in recent research. One focus area concerns the factors that influence QoE. In [38], the authors consider web response time as a major factor that influences the user perceived QoE. They define response time as the time it takes to download a web page. During a series of experiments, objective network metrics (e.g., throughput) were recorded and then mapped to QoE MOS values using a logarithmic relationship. In [33], the authors present the factors that influence the QoE of popular applications in mobile networks. By performing a 4-week, 29-user study, they collected a combination of network measurements as well as user-reported MOS values in an effort to improve the understanding of QoE in a real world environment within different contexts. The study showed that users make use of the applications while being in fixed indoor position and alone. The various factors that influence QoE were grouped
Another focus area of recent research is the proposal of QoE measurement frameworks from the end-user side, employing the concept of crowd-sourcing. In [35], the authors present a conceptual framework for the evaluation of QoE in a real-world environment. Key ideas in the proposed framework are the need to move away from highly controlled test environments, the need to take into account the context in which the user makes use of a service as well as the need to incorporate both objective and subjective aspects in QoE measurements. In [34], the authors implement the proposal in an empirical study which took place in six different contexts (at home indoor/outdoor, at work indoor/outdoor, in bus and in train), proving that the context that the user is in does play a non-trivial role on the perceived QoE. There are also several crowd-sourcing based mobile network measurement systems that have smartphone apps for different mobile platforms. OpenSignal [37] is such an application available for Android and iPhone which collects a wide range of on-device data, like network data (e.g., received signal strength, network type, throughput, etc.) and user/context data (e.g., location, time of measurement, etc.). The data are aggregated and shared publicly in the form of coverage maps which aim to increase consumer awareness. Although there is not a direct mapping to QoE-related metrics (e.g., MOS values), systems like OpenSignal measure the provided service as the users experience it.

Finally, a third research focus area is the estimation and provisioning of the provided QoE from the network side. The aim is a QoE-aware network that allocates resources to maximize user experience. Recent literature [39], [40], [41] proposes a cross-layer network design where the various network layers interact in order to optimize the provided QoE. The approach is to acquire a mapping function between a network metric (such as response time [40] or data rate [39], [41]) and QoE, and to use this function as a utility function in a dynamic resource allocation algorithm. The aim of the algorithm is the maximization of the aggregate experienced quality of all users within a cell. Following the same approach, the authors in [42] propose a resource allocation algorithm in the uplink direction for live video delivery taking into account its popularity. Discriminant Analysis is also a statistical modeling technique used to identify the degree of impact of each QoS metric on QoE [43]. In [44], the authors propose a new network architecture to optimize video content delivery based not only on network metrics, but also on subscription profiles and application requirements. The proposed architecture includes a Mobile Content Delivery Network (MCDN) which is a set of CDN nodes across the mobile network, as well as content-aware mobility management functions. They also propose a cross-layer optimizer that allocates resources based on video sensitivity streams. Industry solutions to aid the operators in optimizing QoE are also available (e.g., [45]).

### IV. Challenges

While a lot of progress has been made in the past decade on QoE measurement in general, unique challenges concerning mobile QoE are still very much open. We outline some of the key challenges in this section.

- QoE measurement in mobile networks is crucially dependent on the underlying measurement approach. On one hand, the crowdsourcing approach allows the collection of vital context information. Measurements from network / operator side, on the other hand, can more easily enable QoE-centric dynamic resource management. How the positive aspects of these distinct approaches can be combined in a more effective mobile QoE framework remains a challenge.

- Context plays a much more critical role in assessment of QoE in mobile networks than in fixed networks. For example, a user who is indoor would experience a very different service quality from those who are outdoors, all else being equal. Location and time of day are other aspects of the user context that have a big influence on the QoE. Mobile devices exhibit significant diversity in their characteristics (form factor, radio chip sets, operating systems, etc.) which again impacts the QoE, and network usage patterns also differ widely among users. The “category” of the device and usage profile also significantly impacts the QoE assessment. Reliable detection of context considering varied aspects such as the ones just mentioned is another challenge that is yet to be fully addressed.

- Finally efficient measurement of various factors that determine the QoE is yet another challenge. Mobile devices are constrained by limited battery life and bandwidth, and high data access costs and more stringent usage caps. This issue is of particular relevance when active measurements are used and also when user-side measurements need to be made available to the network side (for example, to perform QoE-driven resource management). High level of spatio-temporal variations in service quality make efficient and accurate QoE measurement a difficult problem. Continual context (e.g., location) detection is part of the problem.

### V. Conclusions

In this paper, we have surveyed the QoE literature, focusing on QoE modeling and measurement. We have then taken a more in-depth look into mobile QoE, motivated by the fact that QoE measurement in mobile networks is not only rising in importance but also more challenging compared to fixed networks. Besides presenting the state-of-the-art on mobile QoE, we also outline some of the key outstanding challenges.

### References


