Analysis and Monitoring of End-user Perceived QoS in Mobile Networks

Algimantas Kajackas, Antanas Vindašius

Abstract—In realistic mobile heterogeneous environment, the quality perceived by user is variable and dependent on following major factors: available bit rate variation, packet loss and latency. Even though QoS management is enabled in most modern telecommunication systems, it does not ensure the actual end-user perceived quality of service level. Therefore, to make sure that operator follows SLA provisions, QoS level has to be tracked continuously by software QoS agents, residing in user devices. Such tracking tools still do not exist. To fill the gap, in this paper we review quality management processes and analyze the possibilities to evaluate end-user perceived quality of multimedia services. Also, the principles of QoS evaluation instruments and algorithm models are presented. We intend to call attention of academic circle, telecommunication service providers and regulators to this problem.

Index Terms—mobility, perceived QoS, SLA

I. INTRODUCTION

Quality of service (QoS) in telecommunications networks is analyzed by many authors in various aspects. The QoS concept, defined by ITU-T Rec. E.800 as “satisfaction of the end user” is commonly accepted. User satisfaction is seen as one of the key components of usability and is expressed by parameters, which focus on user perceivable effects.

QoS perceived by user (PQoS) as a time function \( Q(t) \) is user’s response to received signal, represented by delivered service. When user resides in critical area of servicing base station coverage, the achieved quality degrades. In realistic mobile heterogeneous environment, the quality perceived by user is dependent on following variable major factors: available bit rate, packet loss and latency.

To define and manage relations between users and operators under variable achievable quality conditions, Service Level Agreement (SLA) concept was created. In context of SLA, the understanding of QoS is more formal and recognized as „degree of conformance of the service delivered to a user by a provider in accordance with an agreement between them” [1].

To make sure that operator follows SLA provisions, QoS level has to be tracked continuously. This may be done by software QoS agents, residing in user devices and monitoring real-time quality of online services. However, such real-time tracking tools of PQoS currently do not exist.

Some specific tools have been made for physical link characteristics and link quality monitoring in terms of technical parameters. These measures have been created purely for operator use but are neither accessible for end users, nor fully reflect the PQoS.

The main problem is that end-users have no means of establishing a quantitative PQoS value of the consumed telecommunications product and means for variable quality accounting do not exist.

In this paper we shortly review quality evaluation processes in mobile network and analyze the possibilities to evaluate and monitor end-user perceived quality of services, also present the principles of PQoS evaluation instruments and algorithm models. End user PQoS evaluation tools are required for network operators, especially for virtual network operators. These tools will help to identify network areas of degraded quality, user concentration, and will help to increase end-user trust in network services and reputation of operator, which influences service usability as well.

II. EVALUATION OBJECT AND RELATED WORKS

Perceived quality of service as a time function \( Q(t) \) depends on instantaneous quality \( q(t) \), caused by physical signal \( v(t) \), which is dependent on physical communication conditions, as well as available bit rate \( \rho \), packet loss \( \pi \) and latency \( \tau - q(t; \rho, \pi, \tau) \). \( Q(t) \) also depends on psychological properties of the user. Let’s assume that the signal \( v(t) \), instantaneous quality \( q(t) \) and perceived quality \( Q(t) \) are linked:

\[
Q(t) = \Phi[v(t); q(t; \rho, \pi, \tau); \xi],
\]

where \( \Phi[.] \) is the model of QoS perception and evaluation by end-user. The model has undefined factor \( \xi \), which reflects user’s psychological side of QoS perception. \( \Phi[.] \) is integral operator dependant on many variables (factors) and generally reflects QoS perceived by user. This operator is not clearly defined.

A. Criteria for Quality

The initial experience of evaluating perceived quality was obtained from Mean Opinion Score (MOS) expertise. MOS methodology is defined for voice [2], video [3], and multimedia [4] services. It is worth noticing, that first MOS
methodology was created in 1996 and dedicated for evaluation of voice codecs. Following this methodology, a group of volunteer experts are trained to evaluate the overall quality, which usually is being impaired by small defects. The experts form their opinion based on long duration speech segments. This way the influence of multiple small defects is integrated.

MOS based methods can be used for several different purposes including selection of algorithms, ranking of audiovisual system performance and evaluation of the quality level during audiovisual connection.

Different approach is applied for analysis of intelligibility of transmitted information or instructions [5]. When some segments are lost or packet loss rate exceeds a given threshold, received audio or video become unintelligible. The example of significance of intelligibility can be illustrated by sign language video transmission over cellular networks [6].

Previous quality evaluation experiences show, that QoS perceived by end user $Q_i(t)$ is ambiguously linked to instantaneous quality $q_i(t)$ and may differ in wide range. The answer to question what kind of methodology to choose and whether to focus on overall quality degradation, or emphasize the aspect of intelligibility, depends on the goals of the task. In this work we offer to differentiate three QoS concepts:

- General Perceived QoS – GPQoS, which is applied evaluating QoS in context of communication, when person receives information of general context (voice chat, video playback, web browsing);
- Special perceived QoS – SPQoS, which is applied if specific instructions are transmitted and decisions are made based on those instructions. It may be applied to machine-to-machine communication, also various systems, dedicated to transmission of instructions;
- Accountable Perceived QoS – APQoS, which should be used to evaluate and account perceived QoS in the period of service delivery.

**B. Thresholds and Periods of Noticeability**

When analyzing any abovementioned perceived QoS concepts, it is important to set the thresholds for parameter degradation and duration of the negative effects, which are assumed to be noticeable for the user. These effects are characterized as perceptual threshold [7], detection threshold, noticeable audio, visual and audiovisual errors. The problems of overall acceptance of quality and general research methodology, including research on user tolerable defects is described in [8].

It was discovered, that user can notice short-term quality degradation, such as individual packet loss during intelligibility tests. Our experiments show, that sometimes user can notice voice defects of extremely short duration, for example, fitting into 20 ms voice segment. Authors in [9] come to similar conclusion, stating that objective error of 30 ms was audible in all contents. In video service study [10] the minimum period of the error detection threshold was 80 ms, and 200 ms errors were visible in all contents.

However, when it comes to analysis of general quality variation in longer period, the conclusions may be different. Instantaneous quality change from “good” to “bad” at some moment in time may not be immediately noticed. The perceived quality changes more slowly than instantaneous quality, with an approximately exponential curve with a time constant of 5 seconds for the good-to-bad transition and 15 seconds for the bad-to-good transition [11]. Human ability to integrate overall influence of several defects is not evaluated in most methodologies. For example, voice quality MOS [2] evaluation uses 2 – 3 second simple sentences and ITU P.910 recommends evaluating video quality in segments of 10 s [3].

**C. On Measurements of QoS**

Objective QoS measuring methods have been created using the experience from studies of subjective evaluation of QoS. These algorithms analyze input signal (voice or video) and, like experts in MOS case, evaluate achieved level of QoS. Most of the common specialized QoS measurement tools today are PESQ for voice [12] and PEVQ for video [13].

PESQ algorithm is designed to follow human voice perception mechanism. Algorithm takes into account human abilities to notice short distortions and ability to integrate them and forms an overall score. Input signal is divided to 32 ms chunks – phonemes. Algorithm calculates spectral characteristics and deflections from reference signal (perceptual differences) for every phoneme. The overall evaluation is based on long (8 – 30 s) segment of distorted and reference signal. PEVQ works in similar way, calculating the perceptual difference comparing aligned signals: all the appropriate indicators are aggregated, forming the final result – the mean opinion score. It is worth mentioning, that attempts to use PESQ for shorter voice segments, creates uncertainty which can be eliminated by special design of input signals [14]. Basically PESQ and PEVQ algorithms estimate both short-term noticeable distortions and overall general quality of received signal. The trust level of those estimations is obtained from result comparison with subjective MOS. Both calculated and MOS evaluations strongly correlate.

In conclusion, it can be seen, that quality evaluation results can vary depending on service, measurement method and measurement tools. Furthermore, special conditions are required for abovementioned measurements. PESQ and PEVQ algorithms are invasive and may give evaluation only when original (reference) signal is available. Therefore, current QoS evaluation methods cannot be applied for service quality monitoring in real-time.

**III. QUALITY MONITOR: REQUIREMENTS AND OPPORTUNITIES**

The measures capable of evaluating perceived quality during service delivery do not yet exist. Analysis of this problem has to start with forming primary requirements. Before forming fundamental model for perceived QoS monitoring system, we have to consider the following:

i) PQoS monitoring system has to be universal and not very complex, to fit into various user devices, which have
limited processing power and memory capabilities.

ii) Currently accepted QoS evaluation and measurement methods as PESQ, PEVQ and other cannot be applied directly during service delivery and cannot be applied for real-life services quality monitoring.

iii) The system has to evaluate quality impairments or quality degradation $\Delta Q$, but not the absolute quality; the solution has to be focused to variations from defined SLA.

iv) The data about primary quality impairments during actual service delivery have to be obtained indirectly – related to parameters, used to describe network performance. Available bit rate $\rho$, packet loss $\pi$ and latency $\tau$ can be used initially for this purpose. These are well known and most influential parameters on voice and video quality.

v) The same measurable variations of parameters $\rho$, $\pi$ and $\tau$ may have different effect on perceived QoS depending on applications, information being transmitted, signal performance, coding, etc. In general, the value of short term quality variations depends on informational value of lost or corrupted packets [14].

vi) The monitor has to be defined according to the task – which of the categories (PQoS, SPQoS or APQoS) is intended to evaluate. Relations between parameter variations and perceived quality have to be defined accordingly.

Following the requirements, we state following basic rules for creating algorithm of QoS monitoring system.

Firstly, time scale is segmented:

$$t_{i+1} = t_i + T', \quad i = 0, 1, 2, \ldots, \tag{2}$$

where $T'$ is segmentation period, depending on monitor target and requirements. For GPQoS evaluation, $T'$ is chosen according to MOS requirements [2], [3]. For other cases, evaluating SPQoS, $T'$ may be equal to service quality noticeable impairment periods [9, 10]. Impairment periods for voice and video are different.

Primary quality impairments and corresponding network performance parameters $\rho$, $\pi$ and $\tau$ should be defined for every measurement period $(t_i, t_i + T')$ as well as degradation of quality

$$\Delta Q(t_i) = Q_0 - Q(t_i). \tag{3}$$

$Q_0$ in (3) shows the quality level offered by service provider or defined in SLA.

Relations between quality degradation $\Delta Q(t_i)$ and network performance parameters $\rho$, $\pi$ and $\tau$ are complex and usually cannot be clearly expressed. This problem can be solved in several ways depending on task of the monitor.

In order to find noticeable impairment periods, the quality degradation $\Delta Q(t_i)$ can be expressed as following polynomial:

$$\Delta Q(t_i) \approx \frac{\delta}{\delta \rho} \Phi[.] \cdot \Delta \rho + \frac{\delta}{\delta \pi} \Phi[.] \cdot \Delta \pi + \frac{\delta}{\delta \tau} \Phi[.] \cdot \Delta \tau. \tag{4}$$

where $\delta \Phi[.]$ is functional derivative of $x$; $\Delta \rho$, $\Delta \pi$, and $\Delta \tau$ – deviations of parameters $\rho$, $\pi$, and $\tau$ respectively.

The equation (4) expresses the degradation of quality as a sum of factors, influencing QoS. Such model, when overall QoS degradation is expressed as linear sum of separate impairments, correspond to ITU-T E-model [16].

Another way of implementing QoS degradation monitor is based on fact that quality degrading factors $\rho$, $\pi$, $\tau$ vary randomly and the accumulated effect on quality change is stochastic as well. Under these conditions, one of the possible solutions is proposed in paper [18] which allows calculating distributions of the de facto perceived quality, using traces of lost frames and using conditional rates of quality categories. Some examples of such experimentally estimated conditional rates of quality classes are presented in [18].

An overall variation of quality is evaluated by aggregating short-time noticeable quality impairments. Various aggregation mechanisms can be applied; however, simple averaging is not suitable, because average value fails to reflect actual perceived quality. Classification algorithms are more suitable for aggregation.

Two-class classification model can be successfully employed, keeping in mind, that quality and its metrics are not strict. Quality of service can be classified to two levels by setting a threshold of allowed quality degradation $\Delta Q_0$. First level represents satisfactory quality of service, where quality degradation is smaller than allowed ($\Delta Q < Q_0$). Second level represents degraded (unsatisfactory) quality ($\Delta Q > Q_0$).

Service delivery time is then divided to set of intervals $T^{\Delta Q < Q_0}$ and $T^{\Delta Q > Q_0}$ which aggregates to total durations of satisfactory ($T_h$) and degraded ($T_m$) quality:

$$T_h = \sum T^{\Delta Q < Q_0},$$
$$T_m = \sum T^{\Delta Q > Q_0}. \tag{5}$$

To provide evaluation for end-user as single value (score), it may be presented as perceived quality coefficient, derived from ratio of normal and degraded quality durations $T_h$ and $T_m$:

$$\kappa = \frac{T_h}{T_h + T_m}. \tag{6}$$

Applying this perceived quality evaluation method, users may be offered different tolerated $\Delta Q_0$.

Functional structure of proposed PQoS monitor is shown in Fig. 1. User service data flow from communications channel arrives to testing tool set, integrated into user device, where three main factor groups are measured: $\rho$, $\pi$, and $\tau$. PQoS analysis tool has the task of classification and evaluation how these factors influence PQoS of particular service. PQoS monitor processes the PQoS evaluations, stores, performs accounting functions, and also indicates PQoS level to user. It is important not only indicate the QoS impairments, but also
collect and store the measurement information for later analysis to discover the cause of impairment. For this reason system architecture should include low-level physical radio parameter and possibly GPS logging capability.

![Diagram](http://wireless.feld.cvut.cz/mesaqin2004/contributions.html)

Fig. 1. Functional structure of PQoS monitor

The form of algorithm (4) as well as the structure of PQoS monitor is universal and applicable to different services. It is obvious that the derivatives $\delta/\delta x \Phi_{i}[]$ are different for different service types and have to be defined separately: each factor may have a particular effect on voice transmission, but different effect on data services, e.g. web browsing. The structure of PQoS monitor may also be different for particular user devices – video, voice terminals. For example, voice quality on mobile circuit switched networks depends on bandwidth and packet loss; meanwhile latency variations are not common.

There are two main problems applying (4) algorithm. Firstly, it has to be defined for what period factor $\rho$, $\pi$, and $\tau$ characteristics have to be measured. Secondly, it has to be clear what influence particular factors have on quality.

Solving first problem, it is worth noticing that bit rate, packet loss and delay variations are easily obtainable by measuring these parameters at user device. This information usually is accessible to developers and vendors of mobile user devices. For example, in cellular networks packet loss is tracked with so called Bad Frame Indication (BFI), this information is used in voice decoding process, enabling lost frame substitution functions [15]. Analysis shows that BFI meets the requirements for packet loss measurement, therefore can be used to determine voice quality degradation in real time.

The available bit rate per user in mobile wireless network depends on coding scheme, which is adapted by network, therefore is always known. This algorithm can be implemented in end user devices using BFI as lost frame indication.

### IV. CONCLUSIONS AND FUTURE WORK

Main causes of perceived QoS variations in mobile wireless networks are user mobility and variable network cell load. Areas with impaired quality usually form either due to radio propagation conditions, influenced by buildings, fading, interference, or due to network planning properties and overutilization of network recourses.

In today’s mobile wireless networks, perceived QoS is not measured during service delivery. In reality, the real perceived quality of service is close to impossible to measure, but methodologies and tools to monitor and evaluate quality impairments can be created. We propose to implement perceived quality monitor, consisting of three parts: test tools for network impairments, quality analysis tool and quality monitor.

The principles of PQoS analysis are based on assumption, that instantaneous QoS $g(t)$ is stochastic process, therefore initial QoS evaluation has to be based on relatively short, but noticeable to user, time periods.

Individual monitors of perceived QoS will create the possibility for the user to obtain objective information and indication whether operator delivers sufficient level of QoS under SLA commitment. These monitoring systems may also help users to choose service provider.

### REFERENCES