Hindawi Publishing Corporation
Journal of Computer Networks and Communications
Volume 2013, Article ID 505910, 15 pages
http://dx.doi.org/10.1155/2013/505910

Research Article

Towards a QoE-Driven Resource Control in LTE and LTE-A Networks

Gerardo Gómez,[1] Javier Lorca,[2] Raquel García,[2] and Quiliano Pérez[2]

[1] *Department of Communications Engineering, University of Malaga, 29071 Malaga, Spain*
[2] *Telefonica I+D, 28006 Madrid, Spain*

Correspondence should be addressed to Gerardo Gómez; ggomez@ic.uma.es

Received 12 September 2012; Revised 9 December 2012; Accepted 19 December 2012

Academic Editor: Lea Skorin-Kapov

We propose a novel architecture for providing quality of experience (QoE) awareness to mobile operator networks. In particular, we describe a possible architecture for QoE-driven resource control for long-term evolution (LTE) and LTE-advanced networks, including a selection of KPIs to be monitored in different network elements. We also provide a description and numerical results of the QoE evaluation process for different data services as well as potential use cases that would benefit from the rollout of the proposed framework.

1. Introduction

The convergence of wireless networks and multimedia communications, linked to the swift development of services and the increasing competition, has caused user expectations of network quality to rise. Network quality has become one of the main targets for the network optimization and maintenance departments.

Traditionally, network measurements such as accessibility, maintainability, and quality were enough to evaluate the user experience of voice services [1]. However, for data services, the correlation between network measurements and user benefits is not as straightforward. Firstly, the data system, due to the use of packet switching, is affected by the performance of individual nodes and protocols through which information travels, and, secondly, radio resources are now shared among different applications. Under these conditions, the performance evaluation of data services is usually carried out by monitoring terminals on the real network.

The end-to-end quality experienced by an end user results from a combination of elements throughout the protocol stack and system components. Thus, the performance evaluation of the service requires a detailed performance analysis of the entire network (from the user equipment up to the application server or remote user equipment).

Quality of experience (QoE) is a subjective measurement of the quality experienced by a user when he uses a telecommunication service. The aim pursued when assessing the quality of service (QoS) may be the desire to optimize the operation of the network from a perspective purely based on objective parameters, or the more recent need of determining the quality that the user is actually achieving, as well as its satisfaction level. However, the QoE goes further and takes into account the satisfaction a user receives in terms of both content and use of applications. In this sense, the introduction of smartphones has been a quantitative leap in user QoE expectations.

Traditionally, QoE has been evaluated through subjective tests carried out on the users in order to assess their satisfaction degree with a mean opinion score (MOS) value. This type of approach is obviously quite expensive, as well as annoying to the user. Additionally, this method cannot be used for making decisions to improve the QoE on the move. That is why in recent years new methods have been proposed to estimate the QoE based on certain performance indicators associated with services. A possible solution to evaluate instantaneously the QoE is to integrate QoE analyzers in the mobile terminal itself [2]. If mobile terminals are able to report the measurements to a central server, the QoE assessment process is simplified significantly. Other solutions are focused on including new network elements (e.g., network analyzers, deep packet inspectors, etc.) that are responsible for capturing the traffic from a certain service and analyzing its performance [3]. For instance, the work presented in [4] investigates the problem of YouTube quality monitoring from an access provider's perspective, concluding that it is possible to detect application-level stalling events by using network-level passive probing only. In other work, the evaluation of video-streaming quality in mobile terminals is addressed by monitoring objective parameters like packet loss rate or jitter [5].

However, whatever solution intended to estimate the QoE from traffic measurements requires some kind of mapping towards a QoE value. A possible solution to perform this process is to apply a utility function associated to the particular data service in order to map the application level quality of service (QoS) into QoE (in terms of MOS value). Many research works are focused in that direction. For instance, a generic formula that connects QoE and QoS parameters (for different packet data services) is proposed in [6]. The work presented in [7] addresses the perception principles and discusses their applicability towards fundamental relationships between waiting times and QoE for web services. Other work quantifies the impact of initial delays on the user-perceived QoE for different application scenarios by means of subjective laboratory and crowd-sourcing studies [8]. Subjective experiments drawing on the evaluation of objective and subjective QoE aspects by a user panel for quantifying QoE during mobile video are presented in [9, 10].

Previous mentioned studies are mainly focused on QoS and/or QoE evaluation, but no action, procedure, or framework is proposed to enhance the end user quality. Only a few works tackle this issue; for instance, a QoE oriented scheduling algorithm is proposed in [11] to dynamically prioritize YouTube users against other users if a QoE degradation is imminent (based on the buffered playtime of the YouTube video player). Other research work provides a methodology for incorporating QoE into a network's radio resource management (RRM) mechanism by exploiting network utility maximization theory [12]. In [13], a specification and testbed implementation of an application-based QoE controller are presented, proposing a solution for QoE control in next-generation networks although they do not include any QoE modeling or estimation algorithm.

In this paper, we propose a novel architecture that enables LTE operators to be aware of the instantaneous QoE that their subscribers are experiencing. In particular, we propose some additions to existing LTE architecture for QoE-driven resource control purposes. We have also identified a set of key performance indicators (KPIs) at the network and application levels for different data services, as well as method to estimate the QoE for web browsing, video YouTube, and voice over IP. Finally, we describe a set of potential use cases that would benefit from the rollout of the proposed framework.

The remainder of this paper is structured as follows. Section 2 provides an overview of the LTE architecture. The proposed architecture for a QoE-driven control is described in Section 3. Section 4 presents a selection of KPIs to be monitored in different network elements. The QoE evaluation process from lower layers' KPIs is described in Section 5. Different use cases associated to the proposed framework are analyzed in Section 6. Finally, some concluding remarks are discussed in Section 7.

2. Overview of LTE Architecture

A general vision of the LTE architecture is described in this section, focusing on the QoS concepts specified in 3rd generation partnership project (3GPP) specifications.

Figure 1 shows the overall network architecture of the evolved packet system (EPS) including the network elements and the standardized interfaces. The network is comprised of the core network (EPC) and the access network (called Evolved Universal Terrestrial Radio Access Network, E-UTRAN). While the EPC consists of many logical nodes, the E-UTRAN is made up of essentially just one node, the evolved NodeB (eNodeB), which connects to the user Equipments (UEs).

The EPS provides the user with IP connectivity to a packet data network (PDN) for accessing the Internet, as well as for running services such as voice over IP (VoIP). One of the main concepts related to QoS in LTE is the EPS bearer, which is a logical connection (associated with a certain QoS level) between the terminal and the evolved packet core (EPC). Multiple bearers can be established for a user in order to provide different QoS streams or connectivity to different PDNs. For example, a user can be engaged in a voice call (via a VoIP bearer) while at the same time downloading a file (via a best-effort bearer).

The EPS includes a policy and charging control (PCC) subsystem, which provides advanced tools for service-aware QoS and charging control. It provides a way to manage the service-related connections in a consistent and controlled way. It determines how bearer resources are allocated for a given service, including how the service flows are partitioned to bearers, what QoS characteristics those bearers will have, and finally, what kind of accounting and charging will be applied.

The EPC is responsible for the overall control of the UE and establishment of the bearers. The main logical nodes of the EPC are (see Figure 1) as follows.

(i) Policy control and charging rules function (PCRF): it is the policy engine of PCC, and it is responsible for the QoS policy management as well as for controlling the flow-based charging functionalities in the policy control enforcement function (PCEF), which resides in the packet data network gateway (P-GW). The PCRF provides the QoS authorization (QoS class identifier and bit rates) that decides how a certain data flow will be treated in the PCEF and ensures that this is in accordance with the user's subscription profile.

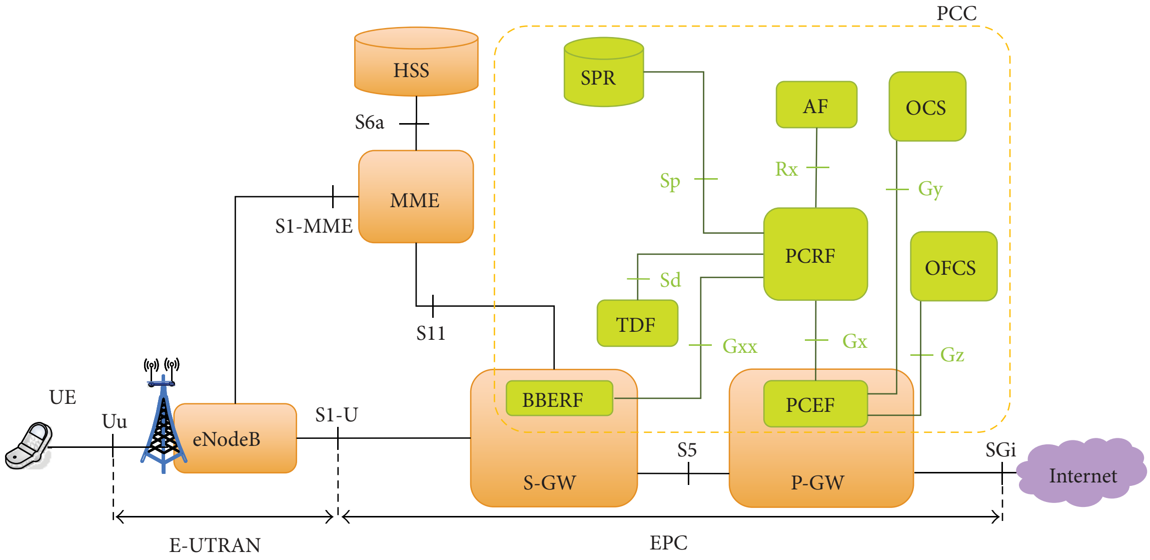


Figure 1: High level architecture for 3GPP LTE and LTE-A EPS.

(ii) Home subscriber server (HSS): it acts as a master repository of all subscriber and service-specific information. It combines the home location register (HLR) and authentication center (AuC) functionality of previous releases. The HSS contains users' subscription data such as the EPS-subscribed QoS profile and any access restrictions for roaming.

(iii) PDN gateway (P-GW): it is responsible for Internet Protocol (IP) address allocation for the UE, as well as QoS enforcement and flow-based charging according to rules from the PCRF. The P-GW is responsible for the filtering of downlink (DL) user IP packets into the different QoS bearers. This is performed based on traffic flow templates (TFTs).

(iv) Serving-GW (S-GW): all IP packets are transferred through the S-GW, which serves as local mobility anchor for data bearers when the UE moves between eNodeBs. It includes a bearer binding and event reporting function (BBERF).

(v) Mobility and management entity (MME): the MME is the control node which processes the signaling between the UE and the EPC. The main functions supported by the MME are related to bearer management (establishment, maintenance and release of the bearers), and connection management.

(vi) Application function (AF): it extracts session information from the application signalling and communicates with the PCRF to transfer this dynamic information, required for PCRF decisions.

(vii) Subscription profile repository (SPR) is the database that stores information related to network usage policies of a subscriber. For example, the SPR can indicate which final services are authorized for a user, the authorized QoS parameters per service, or the user category (e.g., business and consumer). The PCRF may use the subscription information as a basis for the policy and charging control decisions.

(viii) Traffic detection function (TDF): it has been introduced in LTE-A to help the network achieve service awareness by introducing mechanisms for service detection.

(ix) Online charging system (OCS) provides credit management and grants credit to the PCEF based on time, traffic volume, or chargeable events.

(x) Offline charging system (OFCS) receives events from the PCEF and generates charging data records (CDRs) for the billing system.

The access network of LTE, E-UTRAN, simply consists of a network of eNodeBs, which are normally inter connected with each other by means of an interface known as X2, and to the EPC by means of the S1 interface. The eNodeB plays a critical role in the end-to-end QoS. It usually performs the following QoS-related functions: admission control and preemption, rate policing (to protect the network from becoming overloaded and to ensure that the services are sending data in accordance with the specified maximum bit rates), scheduling (to distribute radio resources between the established bearers), and L1/L2 protocol configuration in accordance with the QoS characteristics associated with the bearer.

The UEs in LTE may support multiple applications at the same time, each one having different QoS requirements. This is achieved by establishing different EPS bearers for each QoS flow. EPS bearers can be classified into two categories based on the nature of the QoS they provide: guaranteed bit rate (GBR) bearers in which resources are permanently allocated and non-GBR bearers which do not guarantee any particular bit rate. In the access network, it is the eNodeB's responsibility to ensure that the necessary QoS for a bearer over the radio interface is met. Each bearer has an associated QoS

Table 1: Role of the main EPS interfaces.

Interface	Role
Gx	Used by the PCRF to convey policy enforcement to the P-GW
Gxx	Bearer binding and event reporting function
Rx	Used by application function to convey policy data to the PCRF
Sp	Retrieving per subscriber policy data
Sd	Used to identify/inform about the ongoing session details from the TDF
SGi	It is the reference point between the P-GW and the PDN (e.g., Internet)
S5	Signaling interface for establishing bearers between S-GW and P-GW
S6a	Used by the MME to retrieve subscriber data from HSS
S11	Used by the MME to control path switching and bearer establishment in S-GW
S1-MME	Signaling interface between the eNB and the MME
S1-U	User plane between eNB and S-GW
Uu	This is the air interface between UE and eNB
Gy	Online charging information
Gz	Offline charging information

class identifier (QCI)—characterized by priority, packet delay budget and admissible packet loss rate—and an allocation and retention priority (ARP) used for call admission control. IP packets mapped to the same EPS bearer receive the same bearer level packet forwarding treatment (e.g., scheduling policy, queue management policy, and rate shaping policy). Thus, the UE is not only responsible for requesting the establishment of EPS bearers for each QoS flow, but also for performing packet filtering in the uplink (UL) into different bearers based on TFTs, as P-GW does for the DL.

Table 1 summarizes the role of the main interfaces in the EPS.

3. Proposed Architecture for a QoE-Driven Control

We propose a novel architecture that enables LTE operators to be aware of the instantaneous QoE that their subscribers are experiencing. In the proposed architecture, all the information related to QoS or QoE will be managed in a centralized point that collects performance indicators from different network elements and take potential actions to improve the QoE.

Ideally, the PCRF would be the preferable candidate for this role. However, current PCRF interfaces' specifications do not provide enough flexibility to receive relevant information from any network element. This is why we propose to deploy an ad-hoc QoE-server (with a standardized interface towards the PCRF). A proper dynamic linkage between the QoE server and PCRF is recommended with the aim of achieving a dynamic control of QoS based on customer perception. As defined in the standard, PCRF may receive QoS-related information from different network elements: P-GW, S-GW, AF, and SPR. The goal of including any kind of interaction between the QoE Server and PCRF is to provide a wider vision of the quality perceived by the end users in order to take actions via policy management.

Taking into account that PCRF entity just includes standard interfaces, the communication between both entities could be fulfilled through the following alternatives (see Figure 2).

(a) *Via Gx reference point:* this option requires the QoS platform to include the capability of interchanging diameter commands with the PCRF. Note that PCRF manufacturers include Gx interface to manage policy rules between applications and policy enforcement points, such as gateways, DPIs, and so forth. We propose to reuse Gx to connect PCRF to our policy server (acting as a PCEF). It is not required that the QoE engine includes the whole PCEF functionality (like traffic filtering, monitoring, etc.) because these tasks need to be performed at the user plane. Instead, this platform just needs to include the possibility of sending/receiving certain information to/from the PCRF. This option has a higher flexibility to inform the PCRF about particular events related to the QoS.

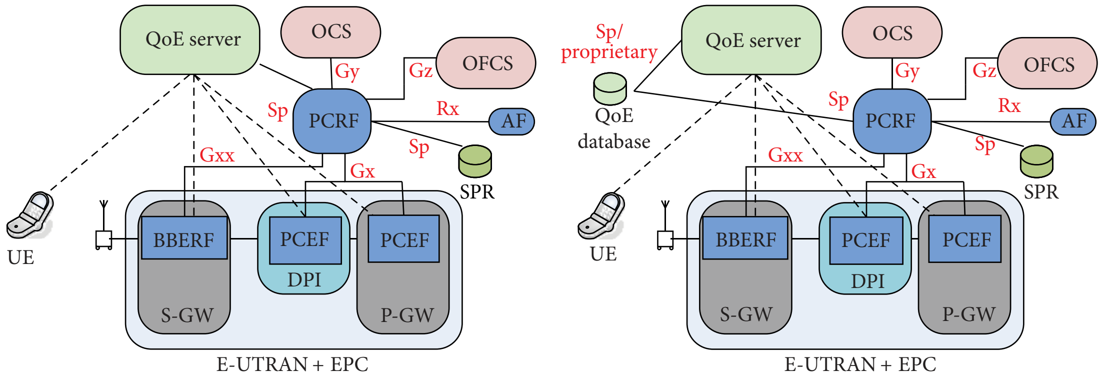
(b) *Via Sp reference point:* this option relies on storing average QoE/QoS indicators in a proprietary database, which can be accessed via Sp from the PCRF, as it is already done with the standardized SPR. This reference point is used to retrieve subscriber related information like allowed services and preemption, subscriber's usage monitoring-related information, profile configuration, priority level (used to determine the ARP), list of allowed QCIs, and so forth. A possible use of the proprietary database is to provide dynamic subscriber-related information according to their associated performance indicators collected by the QoE engine.

The QoE server will be responsible for the following tasks: (1) collecting performance indicators from different network elements; (2) estimating the QoE for specific data services from previous performance indicators; (3) triggering potential actions (depending on the use case). These tasks are further described along this paper.

3.1. Collection of Performance Indicators.

Numerous network elements may contribute to the performance monitoring process. Traditionally, performance-statistics from the operator network management subsystem (NMS) were the main source of feedback information to assess the service quality [14]. However, they are not considered very useful for the evaluation of data service quality, as NMS statistics are averaged for different services and for a long period of time (typically 1 hour).

Instead, mobile network operators usually deploy some kind of monitoring platform based on deep packet inspectors (DPIs). A DPI is a network equipment that potentially allows network providers to monitor, collect, and analyze the data



(a) QoE Server—PCRF connection via Gx

(b) QoE Server—PCRF connection via Sp

FIGURE 2: Proposed QoE architecture.

communications of millions of users simultaneously. DPIs make it possible to identify the applications being used on the network, which is very valuable information for many purposes such as QoS policy management. If a DPI is available, it will provide very valuable (real-time) information about the QoS being provided to each data flow. The location of such DPI will likely be close to the P-GW (or even within the P-GW as a hardware card). The advantage of having a DPI is that it is able to monitor above IP layers, for example, the transmission control protocol (TCP). Note that at, this location, all EPS bearers are handled by the DPI (or P-GW), each EPS being associated with a particular QoS profile. In fact, the complete QoS profile associated to each EPS bearer (i.e., QCI, ARP, GBR, maximum bitrate (MBR), etc.) is well known. That way, it may be checked whether the provided QoS is in concordance with the negotiated QoS profile; otherwise, EPS bearer renegotiation actions may be triggered via the PCRF. All statistics should be obtained per EPS bearer and calculated at a quick rate (e.g., 1 second) so that the QoE of each data service is (re)estimated every second, and real-time actions may be triggered.

In addition to the DPI monitoring process, the potential complementation of information with other sources of information from different network elements (mobile devices, gateways, etc.) is foreseen. If device-based agents are used within the network, it shall be integrated onto the QoE engine as a manner of enriching overall view of QoE for the network. Agent-based solutions are considered an interesting approach within an overall QoE monitoring strategy as representing a unique solution to access device specific issues. Nevertheless, these solutions are currently seen as a complement to network-based approach that shall help in specific issues related to QoE but are not currently seen as a replacement for network-based approaches due to the following.

(i) It is difficult to think on all-network scalability of these solutions on the medium term as per heavily dependent on handset manufacturer willingness to include them.

(ii) Agent-based solutions provide a very detailed view from end customer perspective, but have strong limitations in terms of exploring root cause for issues beyond pure device and access network-related problems.

(iii) Beyond pure technical aspects around the solutions, there are specific privacy and data protection aspects that need to be considered within the implementation of such solutions.

Based on that, it is recommended that device-based solutions are considered within the overall QoE monitoring strategy, scoping a percentage of the network and with the main function of complementing network-based solutions, specially for these aspects that are not easily seen/estimated from network perspective (device issues, network unavailability, and precise location of events).

4. KPIs Monitoring

As discussed before, our proposed architecture uses KPIs collected from different network elements, being DPIs and mobile terminals the most relevant ones. In this respect, this section describes a set of potential KPIs that might be monitored in such network elements.

4.1. KPIs to Be Monitored at a DPI. Existing DPIs are able to monitor a wide set of parameters and performance indicators at different network layers and associated to different data services. Here we list three basic network performance indicators that are key to characterize the instantaneous network status; they are useful for the estimation of the QoE associated to whatever data service.

(i) *IP level throughput:* it may be used to compare the provided throughput with the GBR and MBR values negotiated during session establishment. In the UL, this KPI provides a good performance indicator associated to the whole EPS bearer, as the statistics are taken at the output proxy of the operator network. However, the incoming DL throughput measured at the DPI may not be a proper KPI for estimating the QoE when the radio interface is the bottleneck of the network, as, in this case, the measured IP level throughput does not correspond to the IP level throughput experienced by the mobile terminal. However, this problem only occurs when user datagram protocol (UDP) is used as a transport protocol and losses may happen between the DPI and the terminal; in that case, the solution is based on obtaining the IP level throughput directly measured in the terminal, as described later on. When TCP is used, this is not a problem as the IP level throughput is regulated by the TCP congestion control mechanism, and, hence, the measured throughput will be (ideally) similar to the throughput received at the terminal.

(ii) *IP packet loss rate:* the study of packet loss rate is a challenge as packet losses may occur in any network element which data passes through. The procedure to measure the packet loss rate can be based on analyzing upper layer protocols; concretely, this procedure is applicable for services based on TCP (e.g., web browsing, HTTP YouTube progressive downloading) or real-time transport protocol (RTP), like in real time streaming protocol- (RTSP) based video streaming). In case of RTP packets, loss detection shall be based on the sequence number field included in the RTP header, checking for possible missing numbers in the incoming RTP flow. In case of TCP-based services, a simple way to detect TCP losses in the P-GW is to analyze the packet retransmissions from the server, computing the duplicated number of sequence in the DL. If selective acknowledgment (SACK) feature is used in the TCP connection, the number of retransmitted packets will be the same as the lost ones, obtaining an accurate measure of the end-to-end loss rate. On the contrary, if SACK feature is not used, when the server detects a new packet loss, all the packets with higher sequence number will be retransmitted, computing all of them as packet losses. In this case, the estimated loss rate would be higher than the actual loss rate. Note that the estimation of the loss rate has to be averaged for a large number of packets; otherwise, the result might be distorted. Other way to compute the TCP losses would be implementing a part of the TCP protocol in the DPI. Concretely, the TCP control mechanisms could be used to determine the packet losses. It would be necessary to compute the duplicated ACKs from the UL as well as the losses due to the retransmission timeout (RTO). The adjustment of the initial RTO must be set as the value of the initial RTO of the server minus the time spent from the P-GW to the server (easy computed by executing a PING command). But the RTO is a parameter calculated dynamically, so it would also be necessary to implement the corresponding Jacobson algorithm in the P-GW to dynamically estimate the RTO value according to the round trip time (RTT) and RTT variation; the results of this dynamical calculation might not be the same in P-GW and the server due to the delays experienced in the external network. This method has the advantage of detecting packet losses even before the server, but it is very costly computationally, due to the vast number of TCP connections managed by the P-GW.

(iii) *End-to-end IP RTT:* a possible method to measure the RTT is based on analyzing the TCP connection establishment of a particular data service at a particular cell. As it is well known, TCP connection establishment uses a three-way handshake where the bit SYN is active. The following steps should be followed. (1) When the DPI/P-GW receives a TCP/IP packet (from a terminal) with the bit SYN active, it must start a timer computing the RTT; (2) the contribution of the external RTT will be computed after the reception of a new TCP/IP packet with the bit SYN that is active (from the server) acknowledging the previous one. The measurement of this contribution is especially important as the load conditions in the external network are unknown to provide a theoretical estimation; (3) finally, the end-to-end RTT will be completed when a new acknowledgment from the terminal is received at the DPI/P-GW. This measurement should be performed for each TCP connection establishment procedure detected at the DPI/P-GW. The most important statistic related to the RTT is the average RTT, which have a very important impact on upper layers' performance, especially for TCP. In that sense, RTT average value should be given for each QCI, as potential actions for improving the QoE in this scenario will be taken per QCI.

4.2. KPIs to Be Monitored at the Terminals.

An important advantage of using real measurements at the terminal side is that they are highly correlated to the real QoE obtained. Thus, collecting statistical data specific for each service and terminal will allow for a better analysis of the performance of each service. Through periodic reporting of these measured values, the QoE assessment process is greatly simplified. In principle, the availability of obtaining certain performance indicators or parameters is dependent on the terminal manufacturer. The focus of this section is to list and describe the main KPIs that should be measured at the terminal side for specific data services.

Such monitoring process shall be carried out by an ad hoc application installed in mobile terminals. The software in charge of collecting terminal KPIs should be low consuming in terms of radio bitrate and processing load in order to not affect the quality of other applications. Such software will be responsible for measuring and reporting a set of KPIs to feed the theoretical model that estimates the QoE. For those mobile terminals that do not include monitoring capabilities, the theoretical model will use default values or average values from terminals located in the same cell.

Potential measurements at the terminal side are divided into the following.

(i) *Signaling KPIs:* associated to signaling delays during service establishment or attach to the EPS network, which just affect the initial service establishment and possible renegotiations of the bearers.

(ii) *Network level KPIs:* although there might be some overlapping with the network level KPIs measured in a DPI, it is always preferable to use KPIs collected by the terminals (if available) as they represent the final QoS received by the user equipment. Example of such KPIs isIP level throughput, IP packet loss rate, IP packet sizes, RTT from terminal-to-server or terminal-to-terminal (which might be periodically measured by sending PING commands from the terminal), and so forth.

(iii) *Transport level KPIs:* main KPIs at this level are TCP parameters, like the TCP advertised window (AD-WN) or the maximum segment size (MSS). The ADWN represents the receiver window size, and it is included by the receiver in every ACK segment, indicating the maximum amount of data that it is able to receive. ADWN value should be at least as large as the bandwidth-delay product (BDP); otherwise, the receiver TCP layer will limit the achievable bandwidth. For example, let us consider a LTE terminal with 10 Mbps of transmission capability in DL. Assuming a typical LTE RTT of 10 ms for a 40 bytes packet, BDP can be computed as ADWN $\geq$ BDP $=$ Bandwidth $*$RTT $=$ 10 Mbit/s $*$ 10 ms $=$ 12.5 kbytes. Additionally, during the TCP connection establishment both ends agree on the size of the largest segment that can be used within that connection, known as MSS. The value of the MSS also has an important impact on TCP performance. The larger the MSS the shorter the slow start phase will take to fill the pipe, due to the fact that during slow start the increment of transmitted bytes is in units of segments. In case of bigger segments, the bandwidth utilization during first slow start cycles is higher, and the BDP can be reached quicker. In addition, MSS also has an impact on the total packet overhead introduced by the different protocols. Another drawback of using small TCP segment sizes is the increase of the number of ACKs that are sent back to the transmitter. We recommend obtaining these two parameters values from the terminal in order to adjust the TCP model accordingly.

(iv) *Application level KPIs:* focused on obtaining some parameters that are required to estimate the QoE. One of the key issues when estimating the QoE is a proper identification of the main application performance metrics that affect the service quality, for example, number of rebufferings for streaming or end-to-end delay for VoIP. The knowledge of these application performance metrics may not be straightforward, but it may require to get lower layer QoS metrics in order to estimate them. Note that the process of mapping application QoS into user QoE may require the knowledge of some configuration parameters that cannot be estimated analytically. Such performance indicators/parameters are service-specific (web browsing, YouTube, VoIP) as described in Section 5. The way to measure some of these parameters is explained below.

(a) *Web page downloading time* (D)*:* there are two options to compute this metric. The most accurate option is based on monitoring and parsing HTTP packets. It is important to identify all HTTP transactions belonging to the same web page since the secondary objects contained in a web page might be located in external servers. The information related with the links where these objects are located is included in the main object. So, a possible way to find out all TCP connections related with the actual web page is searching all the links included in the main page HyperText Markup Language (HTML) code. Once all the segments belonging to the same transaction have been identified, the web page downloading time can be estimated as the time spent from the web page request to the last data segment received. Another simpler option is to estimate the web page downloading time from the lower layer throughput and web page size, which could be known a priori by reading the content-length HTTP header response.

(b) *End-to-end delay (d):* in order to compute the end-to-end delay, there are two options. The first option is to analyze the timestamps (if available) included in some packets. This information is included, for example, in real time control protocol (RTCP) packets, which is commonly used in standard VoIP service. In addition, this solution requires sender and receiver to be synchronized via network time protocol (NTP). Another drawback of this solution is that RTCP packet sizes may differ from the size of those packets containing the user data, leading to a difference between the measured delay (using RTCP packets) and the actual delay (corresponding to data packets). The second option is to approximate the end-to-end delay as the half of a RTT, which might only be measured at transport layer during the TCP connection establishment, or at network layer, as described before.

(c) *Loss probability at application level:* for RTP-voice services, loss detection can be based on the sequence number field included in the RTP header, checking for possible missing numbers in the incoming RTP flow.

(d) *Video buffer size:* this parameter is included in the request that the embedded player sends to the multimedia server for the download of the selected video. The "burst" field (within "videoplayback" list of parameters) indicates the

Table 2: Summary of KPIs to be measured at each protocol layer.

Layer	KPI/parameter
Signaling	Attach delay
	EPS bearer establishment delay
Transport	TCP advertised window (ADWN)
	TCP maximum segment size (MSS)
Network	Average RTT
	RTT variance
	IP packet loss rate
	IP packet sizes
Application	
Web browsing	Average throughput at application layer (r)
	Web page downloading time (d)
	DNS resolution time
VoIP (E-model)	End-to-end delay at application level (d)
	Loss probability at application level (e)
Video YouTube (HTTP/TCP)	Video buffer size (B_{full})
	Buffer length when the player paused (B_{empty})
	Video bitrate (λ)
	Video length (l)
	Average TCP goodput (β)
	Number of empty-buffer events (n_{rebuf})

buffer size in seconds, which is multiplied by the video data rate, provides the buffer size in bytes.

(e) *Video bitrate and video length:* these parameters are sent as metadata in the file downloaded from YouTube. This file is Flash Video (FLV) for the majority of non-High Definition clips and MP4 for High Definition clips.

A summary of the main KPIs to be measured at each protocol layer is listed in Table 2.

5. QoE Estimation Process

All KPIs obtained from different network elements are related to different layers below the application. For instance, KPIs measured at the gateways are mostly related to the network level, KPIs measured at a DPI may be associated to the network or transport level, whereas KPIs measured at the terminals can be associated to any level below the application. For that reason, the performance at lower layers received at the QoE server must be mapped onto application performance level, and ultimately, onto a QoE value.

We propose a methodology for estimating the QoS and QoE perceived by the user for different packet data services over wireless networks. The proposed methodology is based on network and protocol models, service-related parameters, and utility functions that map QoS objective metrics into the subjective experienced quality as perceived by the end user.

The modeling methodology follows a bottom-up approach, from the physical up to the application layer, taking into account the effects with a higher impact on the overall QoS. Therefore, layer i provides a set of performance indicators to the layer above ($i + 1$) and successively, up to the application layer. Specific equations that model each layer along the protocol stack is out of the scope of this paper although further details can be found in a previous work from one of the authors [15].

The final goal of this end-to-end model is to evaluate the application level QoS, which will be later mapped into QoE (in terms of MOS value), as shown in Figure 3. This last process is proposed to be performed by means of utility functions associated to each particular service. The goal of the utility functions is to map objective measurements (in terms of QoS) into subjective metrics (in terms of QoE perceived by the user).

Note that utility functions are very service dependent whereas MOS values will be estimated per QCI, which may aggregate different services according to the standard [16]. This may be a problem if the operator decides to aggregate a key data service with other services into the same QCI. If this is the case, it is highly recommended to use proprietary QCIs to keep separated the data services to be optimized.

This mapping process shall consider the specific characteristics of each data service. As an example, we focus on three different services:

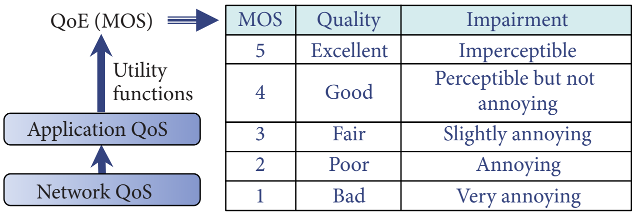
(i) *Web browsing:* the most important objective parameter to estimate the MOS in a web browsing session is the web page downloading time D. The utility function (utility functions are generally obtained through subjective tests to users, by varying the value of the application performance metrics under consideration) that estimates the MOS as a function of D (in seconds) is given by [17]:

$$\mathrm{MOS} = 5 - \frac{578}{1 + (11.77 + 22.61/D)^2}. \tag{1}$$

(ii) *Video YouTube:* among the various works devoted to estimate the MOS for video services [18–20], the analysis presented by [18] provides a utility function for hypertext transfer protocol (HTTP) video streaming as a function of three application performance metrics: initial buffering time T_{init} (time elapsed until certain buffer occupancy threshold has been reached so the playback can start, measured in seconds), mean rebuffering time T_{rebuf} (average duration of a rebuffering event, measured in seconds) and rebuffering frequency f_{rebuf} (frequency of interruption events during the playback, measured in seconds^{-1}). The final MOS expression is given by

$$\begin{aligned}\mathrm{MOS} = 4.23 &- 0.0672 T_{\mathrm{init}} \\ &- 0.742 f_{\mathrm{rebuf}} - 0.106 T_{\mathrm{rebuf}}.\end{aligned} \tag{2}$$

Note that these application layer metrics (T_{init}, T_{rebuf}, and f_{rebuf}) can be estimated (at the receiver) from



MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Figure 3: Bottom-up approach to evaluate the QoE.

performance indicators at lower layers (like the TCP throughput) as well as other configuration parameters like video coding rate or buffer size at the receiver (see [18] for further details).

(iii) *VoIP:* in this case the MOS formula just maps the result given by an intermediate model into normalized MOS values. This intermediate model, known as the E model, is specified in [21], and it provides a numerical estimation $R \in [0, 100]$ of the voice quality from a set of network impairment factors related with the signal to noise ratio (SNR) of the transmission channel, delay, distortions introduced by the coding/decoding algorithms, packet losses, and so forth. In [22], a simplification of the E-model is provided, particularizing it for VoIP communications, where the voice quality R is given by the following expression: $R = 94.2 - 0.024 \cdot d - 0.11 \cdot (d - 177.3) \cdot H(d - 177.3) - I_{e\text{-eff}} + A$ being d the end-to-end delay in milliseconds, $I_{e\text{-eff}}$ the effective equipment impairment factor, $H(x)$ the unit step function, and A the correcting factor, which takes into account the environment where the communication takes place. Besides, [22] provides a formula to translate the R value into MOS:

$$\begin{aligned} \text{MOS} = {} & 1 + 0.035 \cdot R \\ & + R \cdot (R - 60) \cdot (100 - R) \cdot 7 \cdot 10^{-6}. \end{aligned} \tag{3}$$

The impairment factors, in turn, depend on the specific codec used for the VoIP communication; the values of these factors for a number of codecs are tabulated in [22, 23].

6. Potential Use Cases

There are many use cases that would benefit the rollout of a QoE-monitoring solution as proposed in this paper. The basic utility of the proposed architecture is to monitor the QoE associated to a particular data service and user. Once the QoE server has information about the specific QoE for that service, the mobile network operator may use such information for different purposes, some of them are described next.

6.1. QoE Estimation.

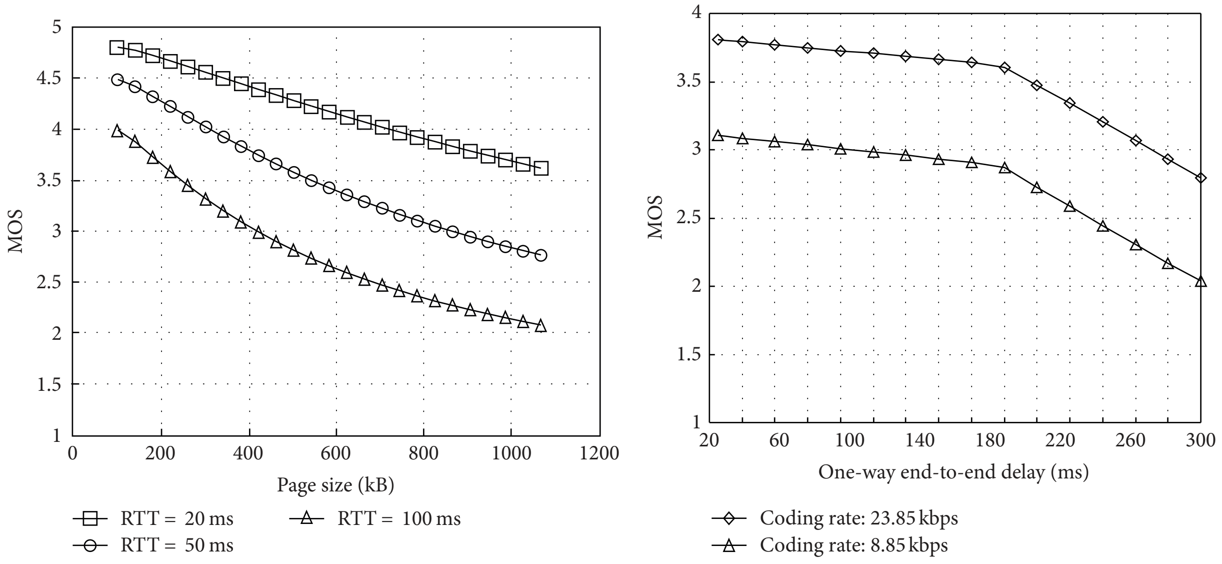
The first use case is focused on the pure QoE evaluation process, including average numerical QoE results for the three services described in Section 5. Results have been obtained from simulations assuming a LTE network whose main configuration parameters (at all protocol layers) are summarized in Table 3. A QoE module is responsible for collecting network and application performance indicators and, afterwards, for mapping QoS onto QoE in terms of a MOS value (according to the utility functions) as described in Figure 3.

Regarding the web service analysis, the exchange of information is done via HTTP/TCP, where HTTP version 1.1 has been assumed. This version includes the persistent connection feature, which makes it possible to reuse the same TCP connection for downloading subsequent objects included in the web page. The optional pipelining feature has been also assumed, thus allowing a number of object requests to be simultaneously sent without waiting for the reception of the previous object. Figure 4 on the left shows the MOS results for different network RTTs and different number of secondary objects in the web page (from 2 to 50 objects of 20 kB each). Firstly, long RTTs lead to a worse TCP performance (in terms of throughput) as a consequence of its inherent congestion control mechanisms (both during slow start and steady-state phases). Such throughput reduction has a direct impact on the web page downloading time and MOS. Secondly, a higher number of objects in the web page (assuming equal sizes) leads to longer downloading times, thus degrading the MOS.

In the case of VoIP service, it usually relies on UDP as transport layer with a configurable voice-coding rate from around 6 kbps to 40 kbps. Due to the low data rates that a VoIP flow usually needs, throughput requirements at the network side are not usually an issue over an LTE network. Instead, the network performance indicator mostly affecting the service quality is the end-to-end delay. Taking into account the characteristics of the VoIP traffic, a robust header compression (RoHC) mechanism has been considered at the packet data convergence protocol (PDCP) layer. In addition, the RLC unacknowledged mode (UM) has been selected in order to minimize the end-to-end delay, which is the application layer metric that mostly affects the MOS. Figure 3 on the right shows the MOS results as a function of the one-way end-to-end delay and the voice coding rate. In the QoE computation formulae, a correcting factor (A) value has been set according to a cellular communication inside a building whereas the impairment factor $I_{e\text{-eff}}$ has been obtained from tabulated values [24] for selected voice codecs. It can be observed that the maximum end-to-end delay that

Table 3: Configuration parameters along protocol stack.

Layer	Parameter	Value
PHY	Carrier frequency	2 GHz
	System bandwidth	20 MHz
	Antenna configuration	1-layer MIMO 2 × 2 (beamforming)
	Precoding	LTE 4-words codebook
	Channel type	Extended pedestrian A (at 4 km/h)
	Channel estimation	Zero forcing
	MIMO detection	MMSE
	Target BLER	10%
	Control channel overhead	From 1 to 3 OFDM symbols
	Modulation/coding rate	16 CQI table (4 bits)
	Channel coding scheme	Turbo codes + SOVA
	Average SNR	20 dB
MAC	HARQ model	Incremental Redundancy + Chase Combining
	Scheduling method	Proportional fair
RLC	Maximum number retransmissions	1 for web and YouTube, 0 for VoIP
PDCP	Header compression	Enabled for VoIP, disabled for web and YouTube
IP	End-to-end delay (excluding the radio interface)	Variable (from 0 to 250 ms)
TCP	Maximum segment size	1460 bytes
	initial congestion window	1 segment
	Advertised receiver window	32 kbytes
	#ACKs per segment	1
	SYN timeout	3 s
APP		
Web	Web page size	100 kB text + variable # of objects of 20 kB each
	HTTP version	1.1 (persistent TCP connection)
YouTube	Video length	250 s
	Video coding rate	512 kbps
	Client data buffer necessary to start the playback	32 s
	Buffer threshold that triggers a rebuffering event	2 s
VoIP	Voice coding rate	8.85 kbps and 23.85 kbps



(a)

(b)

Figure 4: Average MOS results for web browsing (a) and VoIP (b).

makes it possible to obtain a fair quality (i.e., MOS = 3) is around 100 ms (for a coding rate of 8.85 kbps) and 270 ms (for 23.85 kbps).

YouTube service is based on progressive download technique over HTTP/TCP; that is, the client sends an HTTP request and, as a consequence, the YouTube multimedia server delivers the requested video through an HTTP response over TCP. According to (2), the MOS for YouTube depends on three application layer metrics (T_{init}, T_{rebuf}, f_{rebuf}), which can be estimated (at the receiver) from network performance indicators at lower layers (like the TCP throughput, end-to-end RTT, or packet loss rate) as well as other configuration parameter (available at the receiver side): TCP AWND size, video coding rate, video length, play-out buffer size at the receiver, or minimum buffer threshold that triggers a rebuffering event (see [8] for further details).

Figure 5 on the left depicts the results of the three application performance metrics for YouTube as a function of the network RTT. The upper subplot represents the achievable average TCP goodput (computed from [25]). So if the average TCP goodput is higher than the video coding rate (512 kbps), then the probability of rebuffering events will be negligible. As the RTT is increased, TCP goodput is decreased until it becomes lower than the video coding rate at certain RTT value; from this RTT value and above, the parameters related to the rebuffering events (T_{rebuf} and f_{rebuf}) are higher than zero (as shown in the lower subplot). The initial buffering time (T_{init}) is also increased for higher RTTs since lower TCP goodput values lead to longer delays to reach the minimum buffer occupancy (B_{full}). The rebuffering time (T_{rebuf}) has the same behavior although it is null as long as TCP goodput is above the video coding rate (i.e., no rebufferings occur). Besides, it can be seen that $T_{\text{rebuf}} < T_{\text{init}}$ for the same RTT value due to the following reasons: (1) the amount of data needed to be filled (B_{full}) for the computation of T_{init} is greater than the amount of data ($B_{\text{full}} - B_{\text{empty}}$) required for the computation T_{rebuf} and (2) the computation of T_{init} assumes that TCP data transfer starts with a slow-start phase whereas the computation of T_{rebuf} considers the TCP steady state to be reached (being the TCP goodput higher in this second phase). Figure 5 on the right shows the MOS results for different RTTs. As mentioned above, for low RTT values (which achieve TCP goodput values higher than the video coding rate), the initial buffering time is the only metric affecting the MOS (the higher the T_{init}, the lower the MOS). When the rebuffering events start to take effect over the MOS, its value is rapidly decreased since interruptions over the playback are very annoying for the users.

For this specific data service, in which the MOS depends on many configuration parameters available at the mobile terminal, it is required to monitor most of the parameters in the own terminal, and not in the network. However, network performance indicators might be monitored either in the terminal (preferable) or in a DPI.

Although previous results correspond to average MOS values, the estimation of the MOS in a real scenario shall be performed instantaneously in order to have real-time statistics about the user's QoE so that real time actions may be taken (see Figure 6). The evaluation of additional actions from the operator side is not under the scope of this paper although a brief description of potential use cases is given in next subsections.

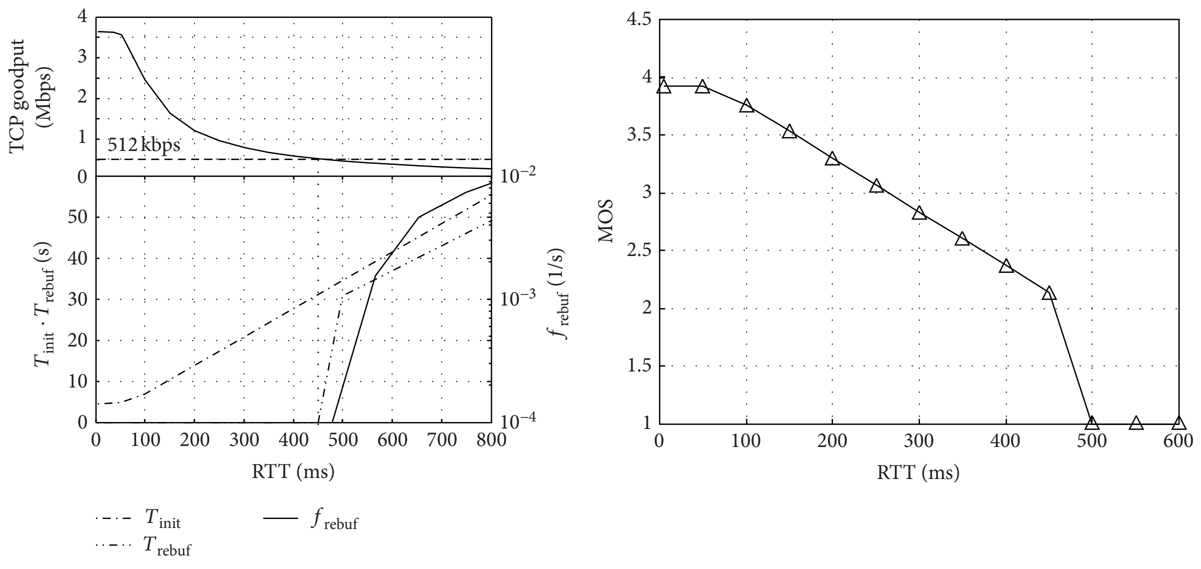
6.2. QoE/QoS Optimization. The proposed architecture can be applied to estimate the QoE perceived by the end-user for new data services over a specific wireless network. In addition, the knowledge of the instantaneous and average QoE per user may help the operator to perform other actions like for instance the following.

(a) *Modification of subscriber priority:* when a poor performance in a specific location or particular subscriber is detected, the interaction of the QoE Server with the PCRF could be considered in order to prioritize network resource usage for each cell site and/or for each individual subscriber. Such indication could be fulfilled by, for example, modifying priority levels (in the proprietary database) associated to particular subscribers (ARP and/or QCI). In case of using a different QCI, it is recommended to use proprietary QCIs that distinguish between subscriber profiles, not between QCIs associated to different services.

(b) *Flexible bandwidth limits:* it allows operators to set dynamically different bandwidth limits depending on a number of factors like: data service (e.g., streaming, gaming, downloads, and email), usage patterns, subscriber, location, time of day, and so forth. Particular QoS policies may be used to optimize the allocation of available bandwidth across subscribers, increasing fairness in network access and improving the user experience, while still taking into account real-time subscriber preferences and behavior, and network conditions.

(c) *Enforce policy rules on many different enforcement points:* the coordination between QoE server and PCRF makes it possible to enforce policy rules in many different enforcement points including access gateways, DPIs, content optimization servers, and even, subscriber devices (or any other network element with access to the QoE solution). Although PCRF only has a direct communication with P-GW, S-GW, and DPI (acting as a PCEF), it would be also possible to set policy rules on mobile devices through the QoE server (for those mobile devices with an ad hoc application). This procedure would require the QoE server to implement a PCEF entity in charge of receiving the policies from the PCRF and, afterwards, forward them to the mobile devices. This allows service providers to apply policy rules throughout the network and support a diverse set of use cases.

(d) *Send notifications to subscribers:* the QoE server might send notifications triggered from the PCRF based on real-time events, such as exceeding a usage threshold for a specific application, roaming to another network, or qualifying for a customer loyalty program.



(a)

(b)

FIGURE 5: YouTube application performance metrics (a) and MOS (b) for different RTTs.

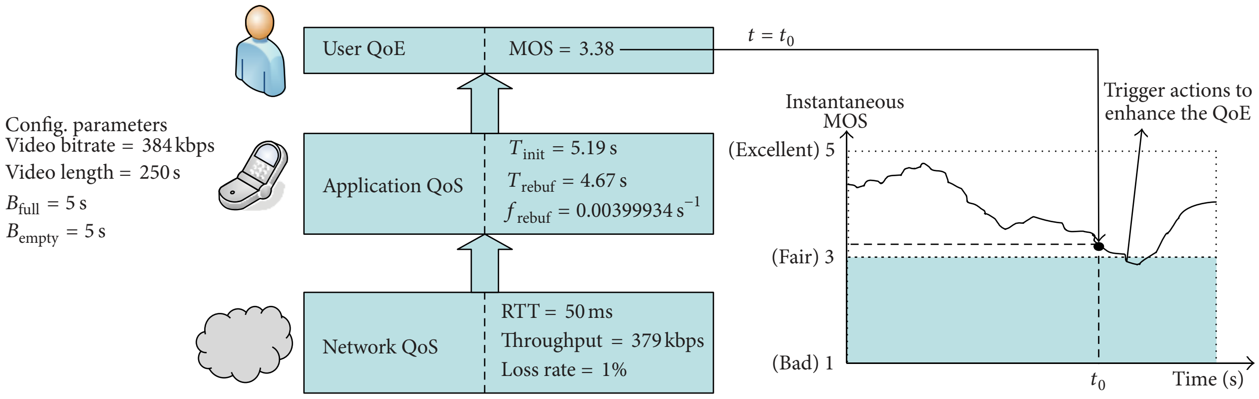


FIGURE 6: Example of MOS evaluation for YouTube service.

6.3. Network Capacity Planning. Capacity management is based on engineering limits which specify the maximum level of utilization that can be tolerated in order to provide the required quality of experience. For example, in voice legacy networks engineering limits are calculated by the use of Erlang's formula on base of maximum tolerable blocking rates. For mobile broadband networks, however, such a reliable and simple relation connecting quality with capacity does not exist.

The widely accepted processor sharing model only serves as an estimate on perceived throughput but fails in predicting quality for a heterogeneous service mix as it is observed in mobile internet traffic. Therefore, direct quality measurement has to play a more active role than just providing end-to-end control as usually employed for circuit-switched networks. Monitoring of resource utilization should therefore be complemented by monitoring of end-to-end quality, giving a complete view on QoE with regard to network topology and time. This will enable to build a reliable correlation between utilization and quality and thus serve as a basis for economically efficient capacity planning. This is particularly important for real-time data services like VoIP or video streaming.

The proposed QoE solution could help on the identification of the minimum resources required for the radio interface, the E-UTRAN and the EPC to achieve a desired QoE. With the aim of fulfilling the operator's end user quality requirements as well as minimizing CapEx and OpEx, our QoE solution could be used for both budget planning. Its end-to-end approach provides additional benefits by ensuring that all domains involved are consistently dimensioned across the whole network.

Device-based solutions are considered of interest for specific use cases (mostly precise location and device performance impact). Nevertheless, these solutions have significant limitations in terms of scalability and handset manufacturer dependencies. Based on this, device-based solutions shall be considered as a complement to a network-based solution that may address specific needs on a sample of the network.

Note that network capacity planning is not a real time process, that is, it does not require quick actions as a consequence of certain events in the network. In that sense, a quick availability of performance indicators is not an issue. Additionally, statistics from the NMS database will help in the dimensioning process as it provides both network topology and traffic load information associated to each network element. Taking into account detailed information about customer usage and traffic/usage patterns, our proposed QoE solution would be able to perform for example, the following tasks.

(i) *Identification of network bottlenecks and (re)dimension the network to ensure the targeted QoE*: QoE measurements with full network coverage can improve efficiency of bottleneck identification and extend the capability of existing load monitoring in classifying the grade of congestion according to impact on quality. This analysis makes it possible to (re)dimension those network elements and links with potential problems. This process should provide optimum network configuration for the given requirements, expected traffic mix, and QoS profiles, and it is the previous step to troubleshooting. It includes a (cell-by-cell) dimensioning process of all network interfaces including the radio (both for user plane and control plane) using real network data.

(ii) *Traffic forecasting based on actual and historical data traffic*. The proposed QoE solution could be also used to perform a traffic forecasting process based on historical data traffic stored in its database. Concretely, this solution could implement forecasting algorithms to predict traffic demand in a per-cell basis based on historical data and on the expected global traffic growth. The goal is to estimate the amount of traffic in the future by spreading forecast market data traffic to the sector level, both in terms of total amount of traffic as well as the traffic mixture.

6.4. Handset and Service Performance Benchmarking. With the growing number of mobile handsets and multimedia content launched onto the market, it is becoming increasingly important for operators to benchmark each individual terminal and measure its performance. Detail insight onto how different handsets (smartphones) do perform within the network for different services and applications, as a manner to guide handset selection and certification, and potentially feeding into device commercial negotiations.

This process enables the identification of problematic handsets and analyses of the cause for the faults. By identifying problematic handsets, operators can quickly make the required adjustments to their network to provide support for more handset models, thus improving the customer experience.

Since our QoE solution will receive performance indicators from a set of mobile devices, it will store statistical reports for quality of a wide range of handsets. These reports show the QoS experienced by the handset user over time. In addition, they show handset usage trends enabling operators to optimize the support for various types of handsets.

The main goal of this task would be to provide the following.

(i) To benchmark handset performance: how voice and data are perceived from real handsets and subscribers' perspective from any point of the network. Our QoE solution will analyze handset performance from voice and packet service statistics over time and location.

(ii) To identify, analyze, and resolve problems linked to handsets.

(iii) Handset validation process: benchmark new handsets based on specific criteria (e.g., check that handsets models used by roamers are compatible). This enables a faster handset selection and validation process and contributes to reducing the need for expensive active testing and emulating hundreds of handsets.

(iv) To deliver the best QoE for new applications: new services such as video-streaming applications are an important source of revenue for operators. In order to ensure top quality data services, handset performance monitoring helps to test applications and measure the quality of experience perceived from a handset prospective. It is important to make sure that multimedia applications are fine-tuned for the handsets that use them the most. Furthermore, using this process, marketing team can easily follow up the introduction of new services and handsets and measure their usage.

6.5. Network Troubleshooting. Current network-monitoring tools may not be the best approach for systematic network troubleshooting when issues are detected on customer experience (further than related to pure network issues without clear correlation onto customer impact). The solution shall be able to provide with customable alarming thresholds setting for different indicator functions. Automatic threshold setting, trend-tracking mechanisms, and automatic/self-learning procedures for deviation tracking availability will be positively considered.

The real challenge comes in diagnosing network problems that impact customer experience. These problems may be specific to a particular cell, device, core network element, or application. In a large network with tens of thousands of sites, each using multiple bands and carriers, and linking to hundreds of core network elements and application servers, finding the one issue underlying a problem may require analysis of Terabytes of data.

Quality of service is the most important LTE troubleshooting feature which may give one vendor (or operator) the advantage over the other. To understand the QoS issues in their networks, operators need to have more than basic analysis capabilities of the network performance in any troubleshooting system they implement. Measuring QoS in all IP networks requires an evolved solution. The engineers need appropriate KPIs (customizable built-in KPIs) to analyze service setup, service quality with dropped sessions, issues, and causes. The network's need to deliver high bandwidth data services, directly influences the capacity required of the monitoring solution and the ability needed to determine which subscribers are using the network and what services are being used.

Several nodes and interfaces are involved in transmitting the subscriber's identifiers which are used for processing the policy and charging control in LTE networks. Any failure to process this information correctly generates mistakes which come to light in the bottom line—customer satisfaction and the billing system. Therefore, it is essential that an operator is able to troubleshoot the relevant nodes preventing both service degradation and loss of income.

Our proposed QoS solution is able to receive real-time KPI alarms, which provide an at-a-glance overview of network and service performance degradations. Automatic notification of problems in the network helps to solve network failure faster, even before the subscribers are aware of such problems.

6.6. Network Monitoring and Reporting. Typically, network monitoring process is mostly based on overall network performance indicators, so that perceived experience by the end customer is not possible to be easily derived on global scale. The proposed QoE solution would allow for a combined network plus customer experience monitoring (based on QoE), being able to anticipate specific customer/service issues and reduce business impact. Detailed information about customer usage and traffic/usage patterns, which is considered of vital importance for both customer business department and for evolving towards increased level of segmentation into multiple dimensions (customer, service, etc.) is based on real trends.

The proposed QoE solution could use passive methods to infer automatically from passive measurements the user perception on the network. The goal would be to automatically derive user perception, from specific indicators being accessed purely from monitoring (eliminating the need for customer surveys) both from the network and terminal sides.

6.7. Customer Care. Currently, customer perception is evaluated mostly via periodical questionnaires and interviews with selected customers that provide views/insights onto perceived experience. The ability of linking perceived (subjective) experience with measured (objective) QoE indicators may lead to significant benefits in terms of achieving a better insight onto customer perceived quality in a much more wide approach than current one based on sampling of specific customers—evolution towards full network customer quality tracking.

QoE-monitoring solutions are linked onto Customer Care centers by means of simplified interfaces and overall status for real-time access to customer specific information, enhancing the response to customer quality and thus satisfaction. Customer care teams can rapidly diagnose problems and identify whether the root cause is linked to a badly performing network, mobile terminal, or application. This makes it possible to identify problems before they affect customers communicating more proactively thereby increasing overall customer satisfaction.

A of this use case, related to the potential active remote handling of devices (e.g., accessing remotely the PC to determine configuration issues), has been identified.

7. Conclusions

In this paper, we have proposed a novel architecture for providing QoE awareness to mobile operator networks. The proposed architecture makes it possible to link QoE engine and PCRF with the aim of achieving a dynamic control of QoS based on customer perception. Combining sophisticated metering capabilities with a highly configurable business rules engine, the PCRF can manage the QoS, optimizing high bandwidth traffic, and enforcing usage quotas. The communication between both entities could be fulfilled through the following alternatives: (a) via Gx reference point or (b) via Sp reference point (through a proprietary database). Several use cases (that take the advantage of such coordination) have been proposed, including the modification of subscriber priority for future bearer establishments, dynamic configuration of bandwidth limits, enforcement of policy rules on many different enforcement points, or the possibility to send notifications to mobile terminals.

Regarding other potential applications of the proposed QoE solution (without requiring interaction with the PCRF), other use cases that would benefit the rollout of a QoE-monitoring platform solution have been described, including network capacity planning, handset and service performance benchmarking, network troubleshooting, network monitoring and reporting based on QoE, and customer care.

Future work will be focused on a feasibility study towards a real-world practice over a LTE testbed. The final goal is to provide practical results and experience on dynamic QoE provisioning in EPS systems.

Acknowledgments

This work has been partially supported by the Junta de Andalucía (Proyecto de Excelencia TIC-06897 and TIC-03226) and by the Spanish Government (TEC2010-18451).

References

[1] G. Gomez and R. Sanchez, *End-To-End Quality of Service Over Cellular Networks: Data Services Performance Optimization in 2G/3G*, John Wiley & Sons, 2005.

[2] A. Díaz, P. Merino, and F. J. Rivas, "Customer-centric measurements on mobile phones," in *Proceedings on 12th IEEE International Symposium on Consumer Electronics (ICSE '08)*, pp. 14–16, 2008.

[3] J. K. Pathak, S. Krishnamurthy, and R. Govindarajan, "System, method, and apparatus for measuring application performance management," Patent number: WO03007115, 2003.

[4] R. Schatz, T. Hossfeld, and P. Casas, "Passive youtube QoE monitoring for ISPs," in *Proceedings of the 6th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '12)*, pp. 358–364, July 2012.

[5] I. Ketykó, K. De Moor, W. Joseph, L. Martens, and L. De Marez, "Performing QoE-measurements in an actual 3G network," in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '10)*, pp. 1–6, March 2010.

[6] 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, 2010.

[7] S. Egger, T. Hossfeld, R. Schatz, and M. Fiedler, "Waiting times in quality of experience for web based services," in *Proceedings of the 4th International Workshop on Quality of Multimedia Experience (QoMEX '12)*, pp. 86–96, July 2012.

[8] 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)*, pp. 1–6, July 2012.

[9] T. De Pessemier, K. De Moor, A. Juan, W. Joseph, L. De Marez, and L. Martens, "Quantifying QoE of mobile video consumption in a real-life setting drawing on objective and subjective parameters," in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '11)*, pp. 1–6, June 2011.

[10] T. De Pessemier, K. De Moor, W. Joseph, L. De Marez, and L. Martens, "Quantifying the influence of rebuffering interruptions on the user's quality of experience during mobile video watching," *IEEE Transactions on Broadcasting*, no. 99, pp. 1–5, 2013.

[11] F. Wamser, D. Staehle, J. Prokopec, A. Maeder, and P. Tran-Gia, "Utilizing buffered YouTube playtime for QoE-oriented scheduling in OFDMA networks," in *Proceedings of the 24th International Teletraffic Congress (ITC 24' 12)*, pp. 1–8, September 2012.

[12] G. Aristomenopoulos, T. Kastrinogiannis, V. Kaldanis, G. Karantonis, and S. Papavassiliou, "A novel framework for dynamic utility-based QoE provisioning in wireless networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '10)*, pp. 1–6, 2010.

[13] J. Sterle, M. Volk, U. Sedlar, J. Bester, and A. Kos, "Application-based NGN QoE controller," *IEEE Communications Magazine*, vol. 49, no. 1, pp. 92–101, 2011.

[14] D. Soldani, "Means and methods for collecting and analyzing QoE measurements in wireless networks," in *Proceedings of the International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM '06)*, pp. 531–535, June 2006.

[15] G. Gómez, J. Poncela González, M. C. Aguayo-Torres, and J. T. Entrambasaguas Muñoz, "QoS modeling for end-to-end performance evaluation over networks with wireless access," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, Article ID 831707, 2010.

[16] "TS 23. 203 V8. 11. 0. 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects," Policy and charging control architecture (Release 8).

[17] P. Ameigeiras, J. J. Ramos-Munoz, J. Navarro-Ortiz, P. Mogensen, and J. M. Lopez-Soler, "QoE oriented cross-layer design of a resource allocation algorithm in beyond 3G systems," *Computer Communications*, vol. 33, no. 5, pp. 571–582, 2010.

[18] R. K. P. Mok, E. W. W. Chan, and R. K. C. Chang, "Measuring the quality of experience of HTTP video streaming," in *Proceedings of the 12th IFIP/IEEE International Symposium on Integrated Network Management (IFIP/IEEE IM' 11)*, 2011.

[19] T. Porter and X. Peng, "An objective approach to measuring video playback quality in lossy networks using TCP," *IEEE Communications Letters*, vol. 15, no. 1, pp. 76–78, 2011.

[20] I. Ketykó, K. Moor, T. Pessemier et al., "QoE measurements of mobile Youtube video streaming," in *Proceedings of the 3rd Workshop on Mobile Video Delivery (MoViD' 10)*, 2010.

[21] "ITU-T recommendation G. 107," The E-model, a computational model for use in transmission planning, 2009.

[22] R. G. Cole and J. H. Rosenbluth, "Voice over IP performance monitoring," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2, pp. 9–24, 2001.

[23] "ITU-T amendment 1 for recommendation G.113," Amendment 1: revised appendix IV—provisional planning values for the wideband equipment impairment factor and the wideband packet loss robustness factor, 2009.

[24] "ITU-T recommendation G. 113," General recommendations on the transmission quality for an entire international telephone connection, 2007.

[25] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: a simple model and its empirical validation," *ACM SIGCOMM Computer Communication Review*, vol. 28, no. 4, pp. 303–314, 1998.

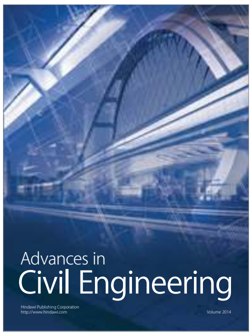
Journal of
Engineering

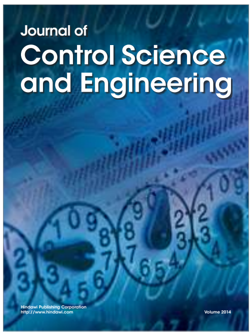
The Scientific World Journal

International Journal of
Rotating Machinery

Journal of
Sensors

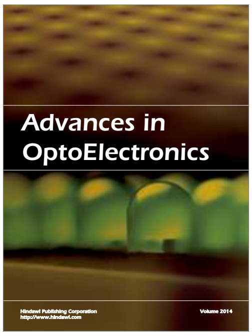
International Journal of
Distributed Sensor Networks


Advances in
Civil Engineering


Journal of
Control Science and Engineering

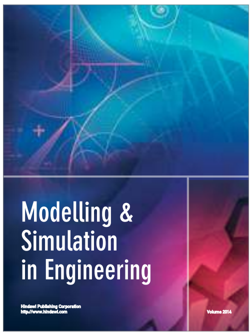
Journal of
Robotics

Journal of
Electrical and Computer Engineering

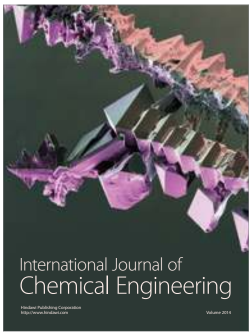

Advances in
OptoElectronics

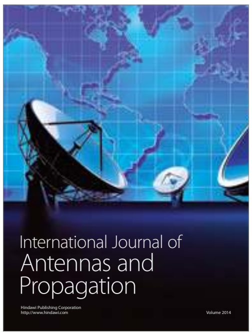
VLSI Design

International Journal of
Navigation and Observation

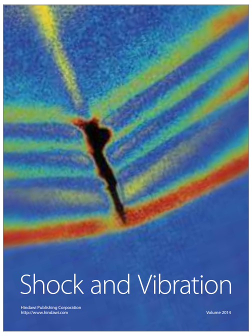

Modelling & Simulation in Engineering

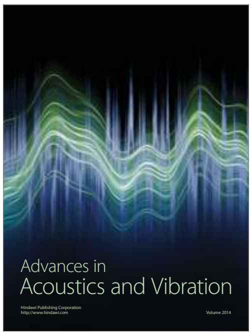
International Journal of
Aerospace Engineering


International Journal of
Chemical Engineering


International Journal of
Antennas and Propagation

Active and Passive Electronic Components


Shock and Vibration


Advances in
Acoustics and Vibration