
Quality of Experience management framework for real-time multimedia applications

Mu Mu*

Computing Department,
Lancaster University,
Lancaster LA1 4WA, UK
E-mail: m.mu@comp.lancs.ac.uk

*Corresponding author

Eduardo Cerqueira and Fernando Boavida

Department of Informatics Engineering,
University of Coimbra,
Polo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal
E-mail: ecoelho@dei.uc.pt
E-mail: boavida@dei.uc.pt

Andreas Mauthe

Computing Department,
Lancaster University,
Lancaster LA1 4WA, UK
E-mail: a.mauthe@lancaster.ac.uk

Abstract: Real-time multimedia will be among the most important applications in next generation networks. However, efficiently managing the delivery of these applications to guarantee Quality of Experience (QoE) to end-users facing network resources limitation and heterogeneity of networks is a challenge. This paper first explores key requirements to provide QoE assurance for multimedia applications in Content Distribution Networks (CDN). A new management framework, named QoE-aware Real-time Multimedia Management (QoE2M), is then introduced to provide end-to-end quality control on real-time multimedia applications over heterogeneous networks based on a combined control of video assessment, Quality of Service (QoS) and QoE-based mapping and adaptation procedures. Validation of the QoE2M framework is also discussed.

Keywords: QoE; quality of experience; video quality assessment; management; adaptation.

Reference to this paper should be made as follows: Mu, M., Cerqueira, E., Boavida, F. and Mauthe, A. (2009) 'Quality of Experience management framework for real-time multimedia applications', *Int. J. Internet Protocol Technology*, Vol.

Biographical notes: Mu Mu received his Bachelors Degree of Engineering from Nanjing University of Aeronautics and Astronautics, China in 2002. His Masters Degree of Science was awarded by Darmstadt University of Technology, Germany in 2006. He has been employed as a network engineer in DJT-software, China and as a research student in Technology Centre, Deutsche Telekom, Germany working on the topic of Seamless mobility and QoS of next generation networks. He is currently a PhD candidate in the computing department of Lancaster University, UK. His research interests include multimedia networking, objective video quality assessment and management.

Eduardo Cerqueira received the BS Degree in Data Processing Technology from the University of Amazon in 2000, MS in Computer Science from the Federal University of Santa Catarina in 2003, Brazil and PhD in Informatics Engineering from the University of Coimbra (UC) in 2008. He is an Invited Auxiliary Professor at the Department of Informatics Engineering, UC. His research interests include quality of service, quality of experience, mobility and multimedia. His publications include one book, four patents and over than 40 papers in national and international refereed journals/conferences. He participated in European Research Projects such as Q3M, SAPRA and CONTENT.

Fernando Boavida received his PhD in Informatics Engineering in 1990, and he is a full Professor and President of the Department of Informatics Engineering of the University of Coimbra. His main research interests are scalable networking and mobile and ambient networking. He was the Chairman of the Program Committee of QofIS'2001, IDMS-PROMS'2002, Networking'2006, WWIC'2007 and FMN'2008 international conferences/workshops. His publications include six book, five book chapters, and around 35 papers in national conferences and 104 papers in international refereed journals and conferences proceedings. He participated in European initiatives such as RARE, EWOS, COST263, E-NET, EuQoS, Weird, OpenNet and CONTENT.

Andreas Mauthe is Senior Lecturer at the Computing Department, Lancaster University. He has been working in the area of distributed and multimedia systems for more than 15 years. His particularly interests are in the area of content management systems and content networks, large scale distributed systems, peer-to-peer systems, and self-organisation aspects. Prior to joining Lancaster University, Andreas headed a research group at the Multimedia Communications Lab (KOM), at the Technical University of Darmstadt. After completing his PhD in Lancaster in 1997, Andreas worked for more than four years in different positions in industry.

1 Introduction

For traditional **internet** applications such as web browsing, e-mail and File Transfer Protocol (FTP), the quality of network delivery is not critical because these applications are elastic and can tolerate certain amount of network impairments. With the increasing bandwidth, more high-throughput, always-on, inelastic real-time applications, such as IPTV and MobileTV, have been introduced into packet-based Content Distribution Networks (CDN). These applications usually require prompt network respond time and low packet loss rate along their end-to-end paths to maintain their quality level. However, due to the nature of packet-based networks, various types of network impairments exist during the content delivery. The impact of the network impairments on the applications can be either trivial when it is not perceived by end-users, or vital when the deterioration on application quality is high that the application is recognised as unavailable to end-users. The design of efficient management and distribution services which guarantee the multimedia applications' Quality of Experience (*QoE*) assurance over heterogeneous (and shared-resources) networks is still a challenging research goal.

Traditional techniques which aim to maximise the quality level of multimedia applications in networking systems are focused only on network performance or QoS control operations. Network performance is measured in terms of parameters which are meaningful to the network and are used for the purpose of system design, configuration, operation and maintenance (ITU, 1993). The concept of QoS was introduced to model the application performance which determines the degree of satisfaction of users (ITU, 1994). Although considering 'user's satisfaction', QoS is mainly a metric on how applications/services are been delivered to subscribers. QoS-based schemes define a set of network-level (and packet-level) control operations to guarantee the packet differentiation of applications in CDN. Existing QoS metrics, such as packet loss rate and packet delay,

are typically used to indicate the impact of network performance on the delivery of applications. The receiving (by user devices), presenting (by displaying units) and perceiving (by end-users) of the applications are not considered. Consequently, network QoS parameters fail in capturing subjective aspects associated with human perception in network control and optimisation operations.

The study of QoE in CDN is driven by the correlation between delivery impairments and perceptual deterioration as a result of these impairments. This correlation is explored based on two aspects: QoE assessment and QoE management. The QoE assessment measures or estimates the user's experience by examining the nature and the delivery of targeting multimedia applications. Although several quality assessment metrics have been proposed (Feghali et al., 2007; Winkler, 2007; Wang et al., 2004; Brun et al., 2004), quality assessment is still a challenge in practice. Additionally, due to the heterogeneity requirement of CDN (from network and application-layer components to terminal characteristics), a modular and flexible framework must be built to support multimedia applications with QoE awareness, while taking into consideration different distribution technologies, delivery resources and other user/application/network level requirements.

The QoE management, on the other hand, utilises the correlation between delivery impairments, QoE deterioration in networks and content management to keep applications with the highest possible perceived quality facing various network impairments and to optimise network resources. Although management is not simply the reverse operation of assessment, the evaluation of the assessment approach is a critical part of any management operation.

In order to overcome the limitations of current QoE assessment schemes, our previous work has presented a framework for integrated video quality assessment in CDN (Romaniak et al., 2008). However, a QoE management framework is still required to control the quality level of applications, while optimising the usage of network

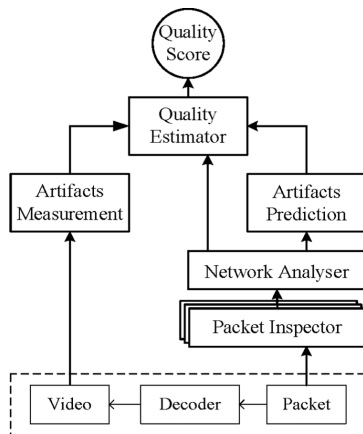
resources. With this goal in mind, the QoE-aware Real-time Multimedia Management (QoE2M) framework is proposed to integrate fore-mentioned QoE assessment framework into the management infrastructure as well as to control multimedia applications based on user's perspectives and available network resources. QoE2M is designed based on the modular integration of application, session and network-layer modules. Hence, operators can define different control mechanisms/functions to be used along the end-to-end application path according to their needs, resources and infrastructure.

The remainder of this article is organised as follows. Section 2 introduces our previous design of the QoE assessment framework. Section 3 introduces the QoE2M. Section 4 presents the validation of the QoE2M framework. Existing works on QoS architectures, QoE architecture as well as mapping and adaptation mechanisms are described in Section 5. Conclusions and future work are summarised in Section 6.

2 QoE-aware video quality assessment framework

In the face of network or terminal resource restrictions, video delivery through IP networks leads to unavoidable quality degradation and a solution to assess how well video services meet the user's expectation is a key requirement for service providers. In ITU (1994), we have presented a QoE-aware framework for integrated video quality assessment in CDN. The framework combines two evaluation approaches: Artifacts Measurement (AM) and Quality of Delivery (QoD) as presented in Figure 1. By combining these two approaches, a flexibly video quality assessment solution with high scalability and content/application independency is achieved.

Figure 1 Block diagram of the QoE assessment framework



Source: Romaniak et al. (2008)

The AM approach estimates the video quality by analysing and quantifying visual distortions (artifacts) in the video frames. The QoD approach predicts the visible artifacts from packet header and packet payload information without fully decoding the video content. Packet Inspector (PI),

Network Analyser (NA) and Artifact Prediction (AP) are three key functions of the QoD approach.

The PI studies all the packets being delivered to the application from delivery network, by accessing the packet beyond layer 2 information (as usually done by traditional shallow packet inspection). Based on the PI information, the NA verifies the transmission status of applications (e.g., type of content, transport protocol, packet loss and jitter). The AP function gives an estimation of how certain packet loss will be displayed on screens. The estimation can be rough if only general QoS parameters like packet loss rate is available or can increase its performance when advanced information is available.

The Quality Estimator (QE) aggregates four quantitative inputs (i.e., playout delay, network QoS, and the results from artifacts prediction and measurement blocks) and gives a 'quality score', which is a prediction of the end-user's experienced video quality. By selectively enabling function blocks, QE can dynamically adjust assessment methods to balance feasibility and performance for specific scenarios.

In order to realise the function blocks, such as NA and AP, we have designed the user testbed (Mu et al., 2009a) and proposed dedicated quality assessment models (e.g., utility function model and artifact-factor model) in Mu et al. (2008, 2009). These enhanced models are also utilised to select optimal management operations in order to protect user's experience during network congestions and also improve the efficiency of network resources controllers.

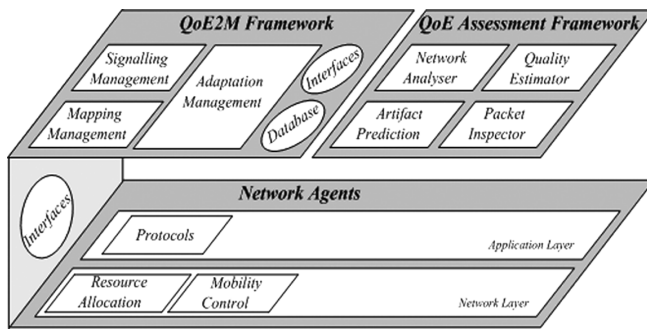
3 QoE-aware real-time multimedia management framework

The proposed QoE2M manages the end-to-end quality level of real-time multimedia group communication applications based on a combined control of video assessment, QoS and QoE-based mapping and adaptation procedures. QoE2M has been developed following a self-adaptation principle and basing on a modular integration of control components over application, session and network layers. The self-adaptation principle allows QoE2M to adjust its control functions to different underlying infrastructure, changing in network conditions, handovers and application/user level meta-information such as content characteristics. The modular design allows the inclusion (or change) of policies, technologies and services.

As illustrated in Figure 2, the QoE2M framework comprises open interfaces, one per-session database and three main components, named signalling management, mapping management and adaptation management. The database keeps QoS and QoE requirements, the quality level score (quantitative/qualitative) and the IP address of the upstream agent associated to the application. The signalling management is responsible to receive join/leave requests from static/mobile users, acquire information about application QoE/QoS requirements and coordinate QoE2M agents. The signalling management is also responsible to collect feedback control information

from users. The mapping management provides on-demand mapping control, by selecting the most suitable service class to be used for an application taking into account the user's perception and multimedia-awareness. The adaptation management adjusts the application quality level to the current network conditions, by performing application and/or network adaptation control procedures.

Figure 2 Location of QoE2M components, QoE assessment components and external application/network-based schemes



The QoE2M does not define any specific technologies or codecs to be used in the framework. Hence, operators can configure QoE2M agents according to their agreements, business models and infrastructures. This is assisted by the open interface between QoE2M control plane and components in underlay networking infrastructure such as transport protocols and resource allocation functions (Figure 2). For instance, the interface among QoE2M and resource allocation controllers allows QoE2M agents to select the most suitable service class to accommodate an application or to inform the traffic conditioner which and when packets must be dropped. Depending of the network size or network infra-structure, QoE2M agents can be implemented in centralised or decentralised manners. Centralised agents control enforcement points in edge-network agents, while decentralised agents are placed directly in edge agents to increase the system scalability.

The proposed QoE2M framework achieves the goal of QoE-aware multimedia management by two distinctive designs. Firstly, a common platform is established with well-defined interfaces and management functions. Secondly, the QoE assessment framework is integrated into QoE2M for real-time quality estimation and QoE-aware management consultant.

3.1 QoE2M components

This section introduces QoE2M components and their main functionalities.

3.1.1 Signalling management

The signalling management controls user's join/leave requests, collects information about application QoE/QoS requirements and coordinates QoE2M agents.

The QoE2M Protocol (QoEP) is designed to coordinate upstream/downstream agents over heterogeneous networks.

QoEP extends the Next Steps Signalling (NSIS) protocol (Cerqueira et al., 2006) with QoE support and operates edge-to-edge following a receiver-driven, source-initiation and soft-state fashion.

Additionally to QoEP, this module controls application setup/release operations, by interacting with Session Initiation Protocol (SIP), Real-time Streaming Protocol (RTSP) or Hypertext Transfer Protocol (HTTP), as well as, collects feedback control information, by using an interface with Real-Time Control Protocol (RTCP). In group communication environments, end-to-end signalling exchanges to setup an application are only required for a first user in an access-agent or access-network, and not for each and any application request as happened with SIP, RTSP or HTTP. Therefore, after installing an application for a first receiver in an access-point, only local operations are done by QoE2M agent to resume a second request and not end-to-end.

3.1.2 Mapping management

The mapping management module maps application requirements and user's perception into service classes. The mapping process is performed based on information about the available service classes inside or between networks (in multiple paths when possible), application QoS/QoE requirements, information about the application quality score (informed by the QoE assessment framework) and mapping policies. The later decides which, what and when mapping methods must be used by one or a set of agents.

The interaction between mapping and underlying resource allocation mechanisms allows the former to query, and receive information about the available class of services and network resources towards a receiver (or another QoE2M agent). After the selection of the most suitable service class(es) for an application, the resource allocation controller is triggered to accommodate the application (or components of an application) into a service class.

3.1.3 Adaptation management

The QoE2M adaptation management function adjusts the quality level of applications when the user's device does not support the application codec or due to unavailability of network resources – (e.g., 'downgrade' in congestion periods). The downgrade adaptation process is reversible when the conditions of fully achieving certain quality level are achieved again. A set of application and network adaptation procedures can be performed by QoE2M agents to control the quality level of new or current applications.

Congestion periods are detected by interacting with resource allocation controllers as well as video quality oscillations are notified by the QoE assessment framework. By parsing the Real-Time Transport Protocol (RTP) header, this module has information about important video-related characteristics to be used during the adaptation process, such as frame type, dependence, complexity, motion, Region of Interest (RIO) and location. The QoE2M adaptation module

can be configured with different Quality Level Adjustment (QLAdjust) values. The QLAdjust values indicate which applications must be adjusted or the amount/percentage of new or current applications to be adapted. Additionally, other QLAdjust parameters, such as population size of each application-group, cost/price, user's location, static/mobile-applications and high-rate-applications can be used for the decision process. For instance, elastic traffic must be dropped first or the system can be configured to drop only 30% of all multimedia traffic during congestions.

3.1.4 QoE2M interfaces

As presented in Figure 2, interfaces are implemented to assure the compatibility (and easy deployment) between QoE2M components, standards and external control mechanisms.

The QoE2M-QoE Assessment interface is used to allow the interconnection between their elements in order to provide multimedia QoE-aware management operations in CDN.

An interface between the signalling management module and application-related protocols, such as SIP, RTSP and HTTP, allows users to join and leave group-based multimedia applications independently of application control protocols. This interface is also used to acquire information about the application QoS and QoE, by interacting with Session Description Protocol (SDP) or other description schemes. It is assumed that an application-related proxy, such as SIP, HTTP or RTSP proxies, receives messages and forwards them to the QoE2M agent associated with the user. After QoE2M QoS/QoE operations, a correspondent message is replied to the user to resume the setup/release process.

Another signalling management interface allows QoE2M to acquire feedback control information about the application quality level in the path from the user to its access-point, by interacting with RTCP or its extensions RTCP Extended Reports (RTCP RC), Media Delivery Index (MDI) or other user feedback schemes. With QoE2M, feedback message exchanges from multiple receivers to their source are avoided, since no source-driven adaptation operations are performed. After detecting a downgrade in the quality level of applications, QoEP can be triggered to coordinate adjustment procedures in upstream agents (if need) or local adaptation operations can be accomplished to keep applications with an acceptable quality level.

Mobility controllers can be located in network and/or link-layers and are responsible to provide hard or seamless mobility to applications. Mobile IP (MIP), Hierarchical MIP (HMIP) and even SIP with handover support are examples of mobility control schemes. An interface between mobility controllers and QoE2M signalling management component allows the latter to be notified during a handover and to coordinate content delivery with QoE and multimedia support on new paths.

Resource allocation controllers are centralised or distributed mechanisms essential to manage network resources, QoS service classes, routes and inter-network

agreements in wired and wireless systems. Both QoE2M mapping and adaptation management modules implement interfaces with resource allocation controllers in order to guarantee network resources to applications and optimise the usage of network resources. Examples of resource allocation controllers for large-scale IPTV networks, wired and wireless networks that could be configured with QoE2M are proposed in Agrisani and Narduzzi (2008), Neto et al. (2007) and Alani and Mehmood (2008), respectively.

The mapping management queries information about the available service classes (in a quantitative/qualitative manner), including loss, delay, jitter parameters and available bandwidth, to the resource allocation controller. After the mapping process, the resource allocation controller is triggered to reserve resources and perform admission control procedures for multimedia applications in the selected classes inside/between wired/wireless networks. More than one class can be used to accommodate different components of a multimedia application, such as audio and video flows.

4 Validation of QoE2M framework

Integration of QoE assessment with QoE2M, realisation of signalling, mapping and adaptation management and a conceptual evaluation are explored in this section to validate the QoE2M framework.

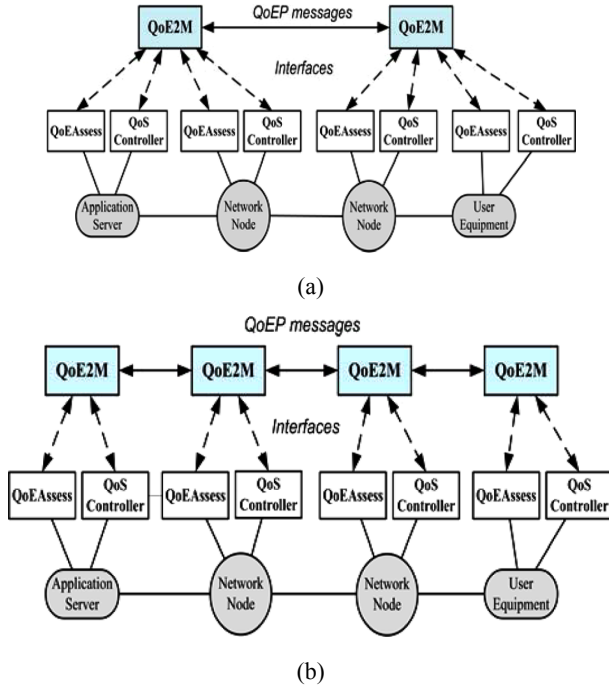
4.1 Integration of QoE2M and QoE assessment frameworks

The QoE2M framework and the QoE assessment framework are integrated to provide a QoE-aware measurement and management services and to support self-adaptive multimedia control in CDN. As depicted in Figure 3, centralised and decentralised modes are two typical deployment approaches. In the centralised mode, QoE2M connects itself with all the QoEAssess (QoE assessment functions) in its network to collect the flow information and quality evaluation results. QoE2M is aware of how packets of multimedia flows are being delivered and so a system level management can be performed within the local network. If an inter-network is established, QoE2M exchanges mapping and adaptation decisions with inter-network QoE2M agents for a global QoE-aware management (by using QoEP). Due to the nature of centralised approaches, flow quality information can be stored and updated in a global/network database. On the other hand, the implementation of QoE2M agents at network edges of networks requires decentralised databases, but permits a higher scalability.

Any QoEAssess agents on the delivery path can be informed with the delivery status and the estimated QoE deterioration at the checkpoint. The QoEAssess notifies QoE2M so management mechanisms can be performed in a pro-active manner. When the media flow departs from QoE2M agent, QoEAssess agent updates the flow quality

indication in the packet header with the append records. With the full list of delivery records along the delivery path, QoE2M agents are aware in which network segments the quality degradation is committed and the degree of its impact on perceived quality. The collaboration between QoE assessment and management frameworks allows the use of management rescue-plans to improve the quality level of multimedia applications.

Figure 3 Integration of QoE2M and QoE assessment frameworks in both (a) centralised and (b) decentralised manners (see online version for colours)



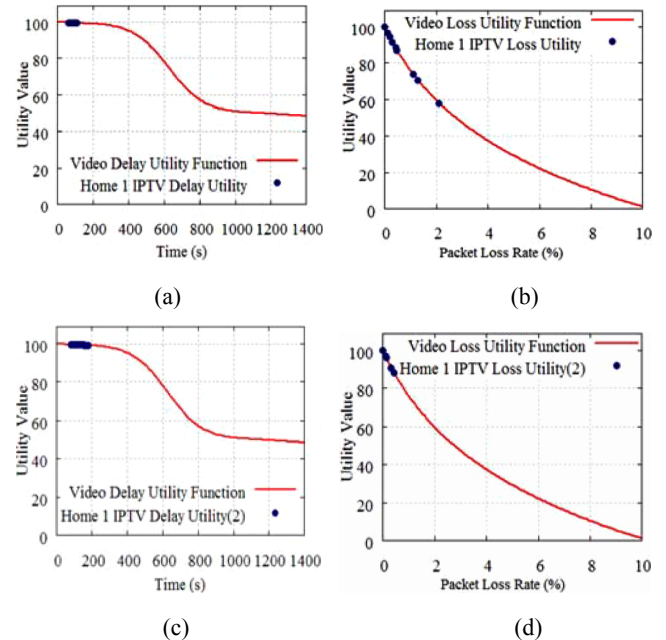
In order to improve the application quality assessment, the QoEAssess agents were extended with a utility function model. With impairment utility functions and application utility functions, the utility function model offers a user-layer extension to existing QoS models to better assess the requirements of multimedia applications. Impairment utility functions model the impact from each network QoS dimension (e.g., delay and packet loss) on the perceived quality. Utility values which are generated from all impairment functions are then aggregated as the application utility which quantifies the user's experience on target application. The utility value represents the impact of application requirements and network resources on the user's perception. The utility model is integrated into mapping management and adaptation management of the QoE2M framework in order to improve intra-flow and inter-flow QoE management operations.

As it is presented in our previous work (Mu et al., 2008), Figure 4 illustrated the impact of packet delay and loss (percentage) of two video flows on the user. The loss utility functions model the perceptual sensitivity of these two applications facing packet loss. For instance, the mapping management maps applications to different service classes

according to the loss utility function. The application which is more sensitive to packet loss can be mapped to the service class having low priority for packet dropping. For another example, Figure 4(a) and (b) show the delay and loss utility function and distribution of delay and loss of a video streaming application. It can be concluded from the figures that packets of this application are received with a delay below 200 ms and some impulse packet losses with 1% and 2% of packet loss. This information can be used to enhance intra/inter-flows adaptation procedures.

Due to the loss utility curve, a packet loss rate over 1% is annoying to end-users. QoE2M agents on the delivery path uses intra-flow traffic management operations to reduce the packet loss rate of this application to below 1% to increase the application quality. What can also be concluded from the utility curves is that this application can tolerate up to 300 ms of packet delay. Further, packet loss rate below 0.5% is unperceivable to end-users. To optimise the utilisation of network resources and potentially improve other applications sharing the same links, QoE2M provides inter-flow quality level adaptation by optimising/changing certain QoS parameters. Figure 4(c) and (d) show the utility value after a adaptation procedure due to the selection of new routes for the video streaming application. By selecting paths with extra network delay and less packet loss, the user's perceived video quality is improved. If computing resources are sufficient, QoE2M interacts with resource allocation controllers to inform the optimal paths for certain applications in real-time using network impairment utility functions and application utility functions.

Figure 4 Utility function values for two video flows (see online version for colours)



Source: Romaniak et al. (2008)

Summing-up, with the analysis of packet information with specific utility functions, the QoE assessment framework generates on-the-fly utility values of individual application

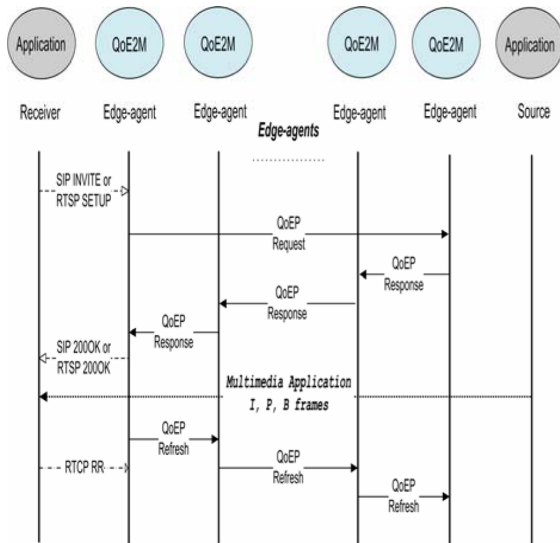
and the combination utility value of the applications which may contain several multimedia applications. These utility values together with other packets information are used for QoE2M refined application quality level control and network resources decisions.

4.2 QoE2M signalling messages

Figure 5 shows some signalling messages associated with QoE2M for a downstream application request control. SIP, RTSP and RTCP messages are exchanged between receivers and their correspondent QoE2M agent (via proxy), and QoEP messages are exchanged between QoE2M agents.

Receivers use SIP-INVITE/RTSP-SETUP messages to subscribe an application. These messages carry the *Session* SDP and are sent to a proxy, which redirects the request to the QoE2M agent controlling the access-point used by the receiver. In this agent, the signalling management collects the application QoS/QoE parameters and user's preference described in the SDP. Afterwards, the QoE2M coordinates with other agents, by using QoEP messages, the quality level to be assured to the application on the path from its source or closet branch point. After finishing its procedures, correspondent SIP200OK/RTSP200OK messages are sent by the QoE2M linked to receivers to resume the operations.

Figure 5 Signalling messages associated with QoE2M (see online version for colours)



QoEP operates in receiver-driven and source-initiated approaches. Following a receiver-driven scheme, QoEP *Request* and *Refresh* messages are exchanged to request mapping and/or adaptation operations for upstream agents, and to refresh/update their state, respectively. *Request* messages are sent towards the source and can be stopped by any upstream agent that has a branch point for the requested application. State refreshment procedures are performed between a pair of adjacent agents, rather than in an end-to-end fashion along the complete signalling path as recommended in the NSIS framework. If the state of an application is not refreshed/updated in a certain period of

time by using *Refresh* messages, its state is removed by soft-state and the resources are shared between current applications. The default value to send *Refresh* messages is 5s and was chosen based on RTCP reports interval, but it could be adjusted according to application or operator needs. To minimise signalling overhead, *Refresh* messages follow an aggregation-application basis (messages with all applications associated with the same edge agents) and not a per-application basis.

Upon receiving a *Request* message, source-initiated operations are performed to install an application, by triggering the mapping (or adaptation) management component. Upon finishing QoE2M procedures in an agent, a *Response* message is used to setup (or modify) an application with QoE assurance along the downstream communication path. The *Response* message is finished when it reaches the QoE2M agent associated with the receiver. Source-initiated signalling procedures can also be started by receiving a *Refresh* message with information about QoE score changes or congestions in downstream agents. In this case, the adaptation mechanism is triggered to adjust applications to the current downstream network conditions.

4.3 Mapping and adaptation management

Figure 6 depicts the mapping management and its interactions with other components. Seven main steps are performed to map applications and user's requirements into available service classes, independently of the underlying QoS models as follows:

Step 1: Mapping request. The QoE2M mapping starts when a request from the signalling management (alternatively, a request can also be sent by the adaptation mechanism/ QoE assessment in congestion periods) is received with information about application QoS/QoE and user's requirements.

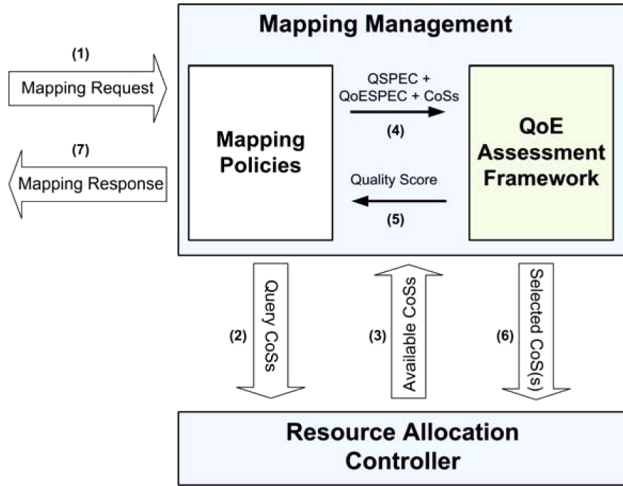
Step 2 and 3: QoS classes and network resources information. The interaction between mapping and resource allocation controllers allows the former to query and receive information about the available Class of Services (CoSs) towards the receiver (or another agent).

Step 4 and 5: Mapping of applications and user's requirements into service classes. Based on information about application and user's requirements as well as available network resources/classes, the mapping QoE assessment is triggered to generate a quality score. The quality score is a classification/weight of each CoS taking user's perspective and application type into account.

Step 6 and 7: Resource reservation and mapping response. After receiving the quality score, a class (or a set of classes) is selected and the resource allocation controller is triggered to accommodate the application or components of an application into a service class. Finally, the requester is notified about the status of the request.

The mapping policies block defines two main mapping methods to select the best class for an application (or its flows/components), named *Full-Matching* and *Partial-Matching*.

Figure 6 Mapping components and interactions (see online version for colours)



A full-matching mapping is accomplished when the quality score of an application in a class exceed the excellence level. If more than one class result in the same quality score, the policy scheme considers only the class which has more available network resources. When most suitable service class cannot assure a full-matching (due to congestions or service class with different configurations in terms of loss, delay and jitter support), the adaptation mechanism is triggered to seek potential adaptation of application to fit the current network conditions. This adaptation can be done by intra-application adjustment procedures or by requesting the re-mapping process with partial-matching mapping rules.

Depending on the business model, multimedia content and dynamism different partial-mapping approaches can be applied as follows:

- *Downgrade class mapping.* This approach chooses a less important class to accommodate the application that assures a good/acceptable quality level (excellence level > utility score \geq minimal quality level).
- *Scalable coding mapping.* This approach takes the importance of each scalable flow of an application into account during the mapping process. It maps high priority flows of applications into the best class and less priority flows into a less significant class.
- *Hierarchical component mapping.* This approach selects service classes according to the priority of different multimedia components. For example, when voice has higher priority than visual content in an application, the packets of audio flow are mapped to the best class and the packets of video flows to a less priority class.

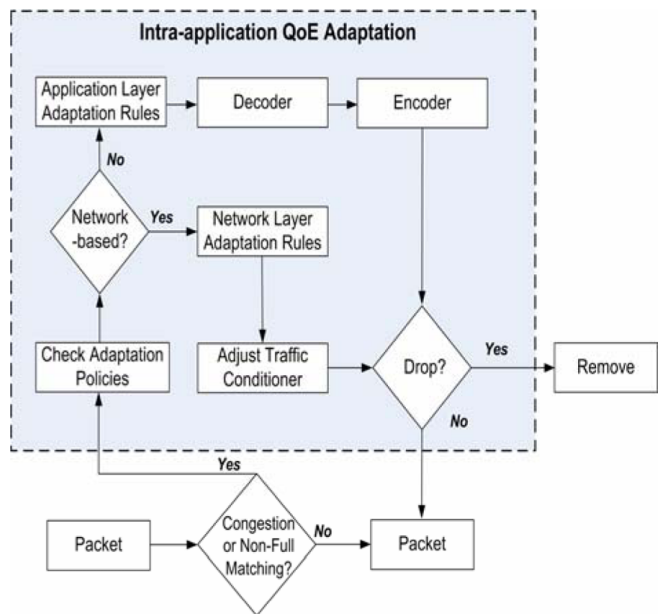
- *Hierarchical frame mapping.* This approach maps multimedia frames based on importance of each frame type. For instance, packets of *I* frames may be allocated into the best class, while packets of *P* and *B* frames are accommodated into a less important class.

The usage of one or a combined set of QoE2M mapping approaches can be determined according to pre-defined or self-adjustment policies. The former executes pre-configured mapping procedures and is configured manually by operators or on-demand by signalling messages. The latter performs mapping functions based on the system behaviour, application characteristics, heuristic schemes, traffic pattern and/or historical data.

Despite the fact that partial-mapping method keeps applications in an acceptable quality level while optimising network resources, this method has disadvantages in practice. For instance, the downgrade class mapping results in engaging resources from inferior classes. This can increase the blocking probability of new applications best suited for inferior classes. Some other partial-mapping approaches assure the application full quality only when the packet re-ordering is not crucial. The scalable coding and hierarchical component mapping schemes can be suitable for scheduled video and audio applications, where it is more important to ensure an intelligible audio component than a perfect video.

When a full-matching mapping is not possible or due to network or terminal resource restrictions, the QoE2M adaptation is triggered to adjust the quality level of multimedia applications to the current network conditions as presented in Figure 7 for an intra-application scheme. A set of application and network adaptation procedures can be performed by QoE2M agents to control the quality level of new or current real-time applications.

Figure 7 Intra-application adaptation diagram (see online version for colours)



Adaptation policies are responsible to determine which application(s) will be adapted as well as if the framework is configured to perform application or network-layer adaptation schemes. Application-layer adaptation solutions are used to reduce/change the quality level of applications, by modifying their content/objects. Networks adaptations control the drop of packets or frames associated each type application.

Application-layer adaptations are performed by using the following approaches:

- *Frame-adaptation.* Adaptations are done by varying spatial resolution, frame rate and bit rate. It is worth notice that varying one of these parameters doesn't necessarily results in changes of the other due to the adaptation scenarios. For instance, if frame resolution and frame rate is critical, the bit rate can be reduced solely by either tuning the compression parameters (compromising the frame quality) or changing frame types (decreasing the robustness).
- *Region-of-Interest (ROI) adaptation.* Instead of performing adaptation on the entire video sequence, this approach takes into consideration the spatial and temporal priority of different parts within target videos. For example, different compression parameters can be applied on a specific area of video frames (which draws more attention of end-users) than the rest of the frames.
- *Scalable coding application adaptation.* A scalable multimedia application, that covers a given bandwidth range, is mapped into one or more scalable streams covering different bandwidth and resolution ranges.
- *Codec conversion adaptation.* Due to the limited number of codecs supported by applications and due to system's heterogeneity, heterogeneous transcoding conversions are supported to convert different multimedia formats.

Although it is usually more efficient, application-layer adaptations share drawbacks such as high CPU consumption and re-encoding errors. Network-layers adaptation approach is an alternative which requires less stress to the system. Traditional QoS traffic conditioners provides a coarse/black-box QoE support for multimedia applications, where packets are dropped during the congestion based on a rate manner and capacity of service classes. In order to perform network control adaptation in a fine/glass-box QoE way, QoE2M is enabled with the capability of adjusting traffic conditioner elements (classifier, marker and dropper) to discard packets according to different rules:

- *Frame dropping adaptation.* This approach drops packets according to the visual importance of each frame. In general, *I* frames are marked with low priority and *B* frames with high priority on dropping probability. Advanced frame dropping adaptations can also be taken by considering video motion, complexity and frame dependency.

- *Scalable coding network adaptation.* This approach adjusts the quality level of applications by, dropping or adding low important flows of scalable multimedia applications.
- *Hierarchical component adaptation.* Media flows within an application can be marked with different priorities. Audio packets are marked with low priority and video packets with high priority dropping probability if voice content is more critical for the multimedia application.
- *ROI network adaptation.* This approach marks in-ROI packet with low priority and out-ROI packets with high priority dropping probability.

With QoE2M, providers evaluate the perceptual quality of multimedia applications with various genres and further guarantee the end-user's experience by controlling network resources and assigning traffic conditioner and dropper schemes according to video genre and frame type. Providers can also choose different encoding and delivery methods for different genres to reduce the impact of network impairments on the visual quality degradation.

5 Related work

This section presents existing QoS architectures and QoE architecture as well as studies on mapping and adaption mechanisms, and highlights the importance of a new QoE-based management framework.

Several QoS architectures have been designed for multimedia content delivery management. QoS Architecture for Multi-user Mobile Multimedia (Q3M) (Cerqueira et al., 2007) and Designing Advanced network Interfaces for the Delivery and Administration of Location independent Optimised personal Services (DAIDALOS) (Miloucheva et al., 2006) implemented a set of application and network controllers to distribute multimedia content with QoS guarantees over heterogeneous networks. Standard ITU-T architectures for Next Generation Networks (ITU-T TD 32 (IPTV-GSI), 2008) and IPTV services (ITU-T Recommendation Y.2001, 2004) also define a set of application and network control functions and requirements for multimedia delivery in CDN. However, these frameworks/architectures lack in QoE-based mapping and adaptation support for multimedia applications, where they do not take the user's perception and video characteristics into account in their control operations. Thus, the requirements of providing QoE assurance for multimedia applications cannot be fulfilled.

Regarding existing QoE architectures, most of the approaches require the user's interaction (expert users) to manage application quality level and do not consider QoE mapping and adaptation operations during the content distribution (Patrick et al., 2004). Other solutions only have mentioned the importance and the need

of adaptation mechanisms in networking environments, but did not present any contributions in this field (Zapter and Bressan, 2007).

In addition to current QoS or QoE architectures, existing proposals are focused on specific mapping and adaptation mechanisms. Regarding mapping solutions, several multimedia QoE-unaware static or dynamic approaches have been proposed to map the application requirements into different service classes (Cerqueira et al., 2008). However, the mapping process is performed based only on network QoS parameters and does not take the importance of each media frame in conjunction with the user's perception into account. Regarding multimedia-aware mapping, a combined DiffServ-MPEG approach maps MPEG-4 frames into DiffServ classes according to the importance of each frame (Fan et al., 2006). This solution protects most important frames in congestion periods, but fails in heterogeneity and flexibility support, where it operates only in MPEG-4 and DiffServ scenarios.

Another QoE mapping proposal maps the application requirements only into a set of pre-defined ITU-T QoS classes and requires the interaction of users during the mapping process (to improve or reduce the application quality level) (Siller and Woods, 2006). Besides requiring the use of extra modules in end-users, a one-to-one based scheme is not suitable for real-time group-based applications, because the quality level of an application cannot be decided only by a single user.

Besides mapping, multimedia adaptation control is essential in networking systems. Source-based schemes perform poorly for group-based systems, because a single transmission rate (Zhu et al., 2007) or resolution (Kung et al., 2007) cannot be used to satisfy the requirements of heterogeneous receivers and networks. Moreover, receiver-based solutions are only suitable to operate with scalable video coding applications and in end-to-end IP multicast systems (Johansen et al., 2007). Existing network-based approaches control the application quality level in congestion periods based on a 'black box' manner, by controlling the number of flows of scalable video coding applications or re-mapping their flows into another service class (Cerqueira et al., 2008).

Regarding network-based 'glass-box' approaches, several proposals use the importance of each frame or flow to drop packets during congestion periods. A MPEG scalable video coding solution for DiffServ systems marks packets to be dropped according to the importance of each layer (Zhao et al., 2005). This approach is only suitable for scalable MPEG-codec and DiffServ networks, which reduces the system flexibility. In addition, non-scalable video adaptation proposals control the application quality level, by dropping less important MPEG packets first (*B* frames) and protecting the most important ones (*I* frames) (Ke and Chilamkurti, 2008). However, this approach does not take the dependency of each frame into account and also requires changes on IP header packets and extra module in sources to accomplish source-based marking packets.

It is concluded from the related work that existing QoS architectures lack in terms of user's experience-awareness and media content-awareness. Moreover, current QoE architectures do not provide mapping and adaptation support for real-time group communication applications. Additionally, most mapping and adaptation approaches were developed to be used in systems with specific QoS models and/or codecs. Some of the studies require the interaction of users during their control procedures or need the implementation of proprietary modules in terminals. Further, current proposals do not consider the key factors which have great impact on QoE-aware evaluating and managing multimedia delivery. Some of these factors such as codec, multimedia characteristics and spatial/temporal location of the artifacts have been presented and proved to be essential in our previous work (Mu et al., 2009a). To overcome the identified limitations, a QoE-aware management framework (QoE2M) was proposed to manage the quality level of real-time group-based multimedia applications and optimise network resources in CDN.

6 Conclusion and future work

This article introduces QoE2M as a solution to perform network and application management in order to manage the user's perceived quality of multimedia applications in CDN. The QoE2M contributes to the system-level management of real-time multimedia applications with two aspects. First, a flexible integrated platform is created for collaboration between signalling management, mapping management and adaptation management. With QoE2M, mapping and adaptation solutions are dynamically selected due to best fit different management scenarios and requirements. Further, future traffic control techniques can also be easily integrated into the framework with the signalling between other functions conforming to the QoEP guideline. The second contribution is the integration with the QoE assessment framework to improve the multimedia quality control support. The responsibility of network management has then migrated from "delivering vast majority of packets to the destination with best effort" to "recognising, evaluating and distributing of content to guarantee end-user's experience and the efficiency of network resource utilisation".

As future work, a heuristic approach to combine all mapping and adaptation approaches according to historic data, multimedia characteristics and traffic patterns will be investigated. Moreover, QoE2M will be evaluated based on simulation and experimental experiments.

Acknowledgement

The work presented in this paper is supported by the European Commission, under the Grant No. FP6-0384239 (Network of Excellence CONTENT) and Agilent Laboratories, UK.

References

- Agrisani, L. and Narduzzi, C. (2008) 'Measurements for Networking: an overview', *Proc. IEEE International Instrumental and Measurement Technology Conference*, May, Victoria, Canada, pp.1328–1333.
- Alani, O. and Mehmood, R. (2008) 'Multi-component teleservice model for multimedia applications', *Proc. IEEE International Workshop on Future Multimedia Networking*, September, Cardiff, UK, Vol. 2, pp.1145–1148.
- Brun, P., Hauske, G. and Stockhammer, T. (2004) 'Subjective assessment of H. 264-AVC video for low-bitrate multimedia messaging services', *Proc. International Conference on Image Processing*, October, Singapore.
- Cerqueira, E., Neto, A., Veloso, L., Curado M., Mendes, P. and Monteiro, E. (2006) 'Multi-user session control in the next generation wireless system', *Proc. ACM International Workshop on Mobility Management and Wireless Access*, October, Torremolinos, Spain, pp.148–157.
- Cerqueira, E., Veloso, L., Neto, A., Curado, M., Mendes, P. and Monteiro, E. (2007) 'Q3M: QoS architecture for multi-user mobile multimedia sessions in 4G systems', *Proc. IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services*, October, San José, USA, pp.38–49.
- Cerqueira, E., Veloso, L., Curado, M., Mendes, P. and Monteiro, E. (2008) 'Quality level control for multi-user sessions in future generation networks', *Proc. IEEE Global Telecommunications Conference*, December, New Orleans, USA, pp.1–6.
- Fan, Y., Su, F., Li, Y. and Xu, H. (2006) 'Network-aware adaptive QoS architecture for video delivery over differentiated service network', *Proc. IEEE International Conference on Telecommunications*, May, Funchal, Portugal, pp.1330–1333.
- Feghali, F., Speranza, F., Wang, D. and Vincent, A. (2007) 'Video quality metric for bit rate control via joint adjustment of quantisation and frame rate', *IEEE Transaction on Broadcasting*, Vol. 53, pp.441–446.
- ITU (1993) *General Aspects of Quality of Service and Network Performance in Digital Networks, Including ISDNs*, ITU Recommendation I.350, ITU Telecomm. Standardization Sector (ITU-T), March.
- ITU (1994) *Quality of Service and Dependability Vocabulary*, ITU Recommendation E.800, ITU Telecomm. Standardization Sector (ITU-T).
- ITU-T Recommendation Y.2001 (2004) *General Overview of NGN*, Technical Report, December.
- ITU-T TD 32 (IPTV-GSI) (2008) *Draft Recommendation: IPTV Architecture*, Technical-Report, January.
- Johansen, S., Kim, A. and Perkis, A. (2007) 'Quality incentive assisted congestion control for receiver-driven multicast', *Proc. IEEE International Conference on Communications*, Glasgow, UK, June, pp.1642–1647.
- Ke, C. and Chilamkurti, N. (2008) 'A new framework for MPEG video delivery over heterogeneous networks', *Computer Communications*, Vol. 31, No. 11, July, pp.2656–2668.
- Kung, H., Hua, J., Chang, Y. and Lin, C. (2007) 'Seamless QoS adaptation control for embedded multimedia communications', *IEEE Transactions on Consumer Electronics*, Vol. 52, No. 1, February, pp.240–249.
- Miloucheva, I., Einsiedler, H., Gomes, D. and Jonas, K. (2006) 'QoS based multicast architecture for heterogeneous mobile IPv6', *Proc. Environment, Proc. International Conference on Telecommunication*, May, Madeira, Portugal.
- Mu, M., Mauthe, A., Casson, J., Tyson, G. and Garcia, F. (2009a) 'LA1 TestBed: evaluation testbed to assess the impact of network impairments on video quality', *Submitted to International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities*, April, Washington, USA.
- Mu, M., Gostner, R., Mauthe, A., Tyson, G. and Garcia, F. (2009b) 'Visibility of individual packet loss on H.264 encoded video stream – a user study on the impact of packet loss on perceived video quality', To appear in *Annual Multimedia Computing and Networking*, San Jose, USA.
- Mu, M., Mauthe, A. and Garcia, F. (2008) 'A utility-based QoS model for emerging multimedia applications', *Proc. IEEE International Workshop on Future Multimedia Networking*, September, Cardiff, UK, pp.521–528.
- Neto, A., Cerqueira, E., Rissato, A., Mendes, P. and Monteiro, E. (2007) 'Quality level control for multi-user sessions in future generation networks', *Proc. IEEE Symposium on Computers and Communications*, Aveiro, Portugal, July, pp.707–714.
- Patrick, A., Singer, J., Corrie, B., Noel, S., El Khatib, K., Emond, B., Zimmerman, T. and Marsh, S. (2004) 'A QoE sensitive architecture for advanced collaborative environments', *Proc. IEEE International Conference on Quality Service in Heterogeneous Wired/Wireless Networks*, Dallas, USA, October, pp.319–322.
- Romaniak, P., Mu, M., Mauthe, A., D'Antonio, S. and Leszczuk, M. (2008) Leszczuk, 'Framework for the integrated video quality assessment', *Proc. ITC Specialist Seminar on Quality of Experience*, Karlskrona, Sweden, May, pp.81–89.
- Siller, M. and Woods, J. (2006) 'Using an agent based platform to map quality of service to experience in conventional and active networks', *IEE Communications*, Vol. 153, No. 6, pp.828–840.
- Wang, Z., Lu, L. and Bovik, A.C. (2004) 'Video quality assessment based on structural distortion measurement', *Signal Processing: Image Communication*, Vol. 19, pp.121–132.
- Winkler, S. (2007) 'Video quality and beyond', *Proc. European Signal Processing Conference*, September, Poznań, Poland, pp.150–154.
- Zapter, M. and Bressan, G. (2007) 'A proposed approach for quality of experience assurance for IPTV', *Proc. IEEE International Conference on the Digital Society*, January, Guadeloupe, French-Caribbean, pp.25–30.
- Zhao, H., Ansari, N. and Shi, Y. (2005) 'Layered MPEG video transmission over IP DiffServ', *Proc. IEEE International Conference on Information Technology: Coding and Computing*, Vol. 1, April, Las Vegas, USA, pp.63–67.
- Zhu, P., Zeng, W. and Li, C. (2007) 'Joint design of source rate control and QoS-aware congestion control for video streaming over the internet', *IEEE Transactions on Multimedia*, Vol. 9, No. 2, pp.366–376.