Quality of experience management in mobile content delivery systems

Florence Agboma · Antonio Liotta

Abstract This study contributes towards the relatively new but growing discipline of QoE management in content delivery systems. The study focuses on the development of a QoE-based management framework for the construction of QoE models for different types of multimedia contents delivered onto three typical mobile terminals—a mobile phone, PDA and a laptop. A statistical modelling technique is employed which, correlates QoS parameters with estimates of QoE perceptions. These correlations were found to be dependent on terminals and multimedia content types. The application of the framework and prediction models in QoE management strategies are demonstrated using examples. We find that significant resource savings can be achieved with our approach by contrast to conventional QoS solutions.

Keywords Quality of Experience · Quality of Experience management · Multimedia streaming · Mobile content · Prediction models · Subjective assessment

1 Introduction

With the growth in the availability of multimedia services, coupled with the technological advances in mobile access devices, Quality of Service (QoS) has become the adopted set of technologies used in managing network traffic in the content delivery systems that provide these services. QoS enables the measurement of network parameters, and the detection of changing network conditions (such as congestion or availability of bandwidth). This information is utilised in resource management by prioritizing traffic. However, QoS processes are not always adequate since they do not take into account the user's actual perception of network performance and service quality. There is now a realisation that the users' Quality of Experience (QoE) should be an important metric to be employed during the design and management of content delivery systems and other engineering processes. This is because QoE is a metric, which relates to the perception of the user about the quality of a particular service or network [21]. It is a measure of perception that determines whether users will adopt a given service or not. QoE strongly depends on the expectations the users have about the offered service.

The proliferation of different types of access terminals possessing different display characteristics and capabilities further highlights the importance of QoE research. Delivering multimedia contents or services to these terminal types without carefully thinking about the users' quality expectations or requirements for them, might lead to service over-provisioning. In multimedia streaming services, there are several parameters, which affect the perceptual visual quality of the displayed multimedia content. These parameters can be grouped under Application-level Quality of Service (AQoS)—which deals with parameters such as content resolution, frame rate and codec type; and Network-level Quality of Service (NQoS)—which deals with parameters such as bandwidth, delay, jitter and packet loss. Traditionally, content delivery systems primarily focus on managing the network parameters, but fail to consider the user's perception of service quality at the application level. For example,

F. Agboma (✉)
Department of Computing and Electronic Systems, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK
e-mail: fagbom@essex.ac.uk

A. Liotta
Department of Electrical Engineering and Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: a.liotta@tue.nl

the maximum bitrate may be fixed at a certain value by the access network. However, this does not necessarily fix the quality of a multimedia content at say, "high." This is because there are numerous ways the multimedia content could have been encoded. The research presented here, focuses on the development of a QoE-based management framework for the construction of "QoS to QoE" prediction models for a variety of content types for mobile video services.

This work extends previous findings reported in [1, 2, 5] and [3, 6], which focused mainly on the collection of QoE data for different types of services. In [5] a QoE management methodology was proposed, which shows how QoE data may be used for the benefit of network operators. The work herein significantly extends that of [5] by providing a comprehensive and well-analysed study of QoE over a broader range of scenarios and for different types of mobile devices (mobile phone, PDA and Laptop). By means of further examples, we illustrated the applicability of the QoE management strategies for each scenario.

The paper is organized as follows. Section 2 discusses related works in the fields of QoE management. Section 3 introduces our QoE framework and details the modelling technique employed to perform QoE management. The results obtained from the statistical modelling technique are discussed in Sects. 4 and 5. QoE management strategies are presented in Sect. 6. Final remarks, conclusions and future directions are given in Sects. 7 and 8.

2 Related works

Only very few studies on QoE management exist in the public domain. The studies not in the public domain have mainly been sponsored by some of the major players in the telecommunication industry, who view the results of such studies as proprietary. Hence, the studies reported in this section are mainly from research and academic institutions.

The authors in [8, 15, 19, 20, 22–24], have proposed different QoE management solutions: [19] presented a model for measuring the QoE of mobile multimedia services, taking into account measurable parameters, which are related to the technological aspects of the service, and non-measurable parameters, which are related to the users' perception of the service, their expectations, and behaviour. The aim in [19] was to produce quantifiable quality metrics for QoE evaluation.

In [20] the authors proposed a QoE framework, using NQoS metrics and a dynamic user feedback. With this, the user can change the QoS requirements at any time via a feedback control. Using an agent platform, the user feedback is mapped onto a reference performance matrix to reflect the user's requirements. The authors in [8] proposed a framework for autonomic service delivery management in the access network. The framework consisted of three planes interacting layers, which spans the whole access network: the monitor plane, the knowledge plane and the action plane. The monitor plane provided information by means of probes monitoring the status of the access network and services. It observes all the packets that passes through the intermediate points to the end device and uses the information in these packets to deduce values of packet loss, delay and jitter. The knowledge plane analyzes the information received from the monitor plane to detect QoE degradation and determine the appropriate restorative actions (*e.g.,* the application of forward error correction on a video stream). The execution of the actions is delegated to the action plane.

QoE is a subjective measurement that generally requires translation into quantitative data. Other authors [15, 22–24] approached QoE management by means of standardized statistical procedures. Whilst [23, 24] focused on the mobile web-browsing environment, the proposed QoE management framework herein focused on mobile multimedia streaming applications using application-level (AQoS) parameters as a starting point. Understanding the user requirements at the application level is essential to allocating only the needed amount of network resources. This top-down approach is an effective technique to delivering enhanced customer value when performing network engineering [10].

3 QoE management framework

Developing the QoE management process is a non-trivial exercise since capturing QoE is very subjective. QoE is subjective because it is driven by psychological as well as technological factors. The proposed QoE framework explains how users' QoE data can be obtained and employed in developing models to be used in predicting users' QoE perceptions. The predictions can then be used in implementing management strategies. The framework is described in general terms. It consists of seven steps as illustrated in Fig. 1.

Step 1: Characterize the application: At this stage, the identification of the QoE indicators for the application or service is made. For example, the QoE indicators for a streaming application will include the audio quality, video quality, and video resolution.

Step 2: Design and define the test matrix: Once the application has been characterized, the low-level QoS factors that affect the QoE indicators the most can be identified. At this stage, the predominant QoS factors and their range of significance are selected. For instance, video quality in streaming applications will be directly affected by: (a) Network Level QoS (NQoS) parameters such as network bandwidth, delay, jitter, and packet loss (b) Application Level QoS (AQoS) parameters such as encoding bit rate and frame rate. An initial
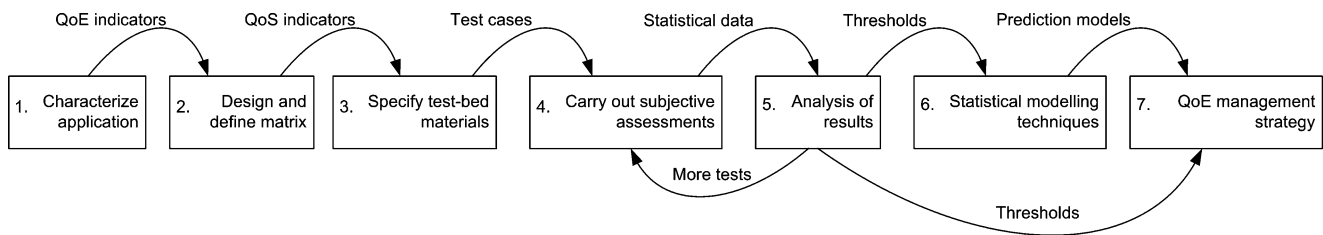
Fig. 1 QoE management framework, showing the key stages/steps as *rectangular boxes*. The results obtained from the activities of a stage (shown as *arrow label*) are used as input for the next stage

matrix of NQoS/AQoS test cases is produced using the factors. Preliminary subjective tests are conducted to reduce the initial test matrix to produce a matrix consisting of the most relevant text cases.

Step 3: Specify test-bed and materials: At this point, we specify the terminals and the materials (*e.g.*, the type of video clips) that will produce QoS test cases used in the subjective tests.

Step 4:. Carry out subjective assessments: After completing step 3, a subjective assessment is performed. There are several subjective assessment quality methods in the recommendations [11] and [12]. Subjective assessments are carried out using the largest possible sample population. Each subject is presented with the generated QoS test cases along with an automated feedback mechanism, which enables the correlation of the user's responses with the generated QoS test cases being evaluated. Care is taken to ensure that subjects are not aware of which QoS parameters they are evaluating.

Step 5: Analysis of results: At this stage, a screening process is carried out to remove unreliable results. The ITU [11] provides procedures for this screening process. Then statistical analyses are conducted to determine whether the results are statistically significant. If not, more subjective assessments are conducted until this objective is achieved. Statistical analyses are carried out to determine the minimum user acceptability thresholds, *i.e.,* the point where the users change their opinion of service quality.

Step 6: Statistical modelling technique: The ultimate aim of steps 1 through 5 is to come up with statistically reliable data in order to build a predictive model, which correlates QoS parameters (objective parameters) directly with estimates of QoE perceptions (subjective parameters). An effective way for achieving this is to employ Discriminant Analysis in determining whether the chosen predictors (*i.e.,* the QoS parameters) lead to reliable user's perception rating. The products of this step are prediction models, which correlate QoS conditions with user's QoEs. This technique is discussed further below.

Step 7: QoE management strategy: Both the analytical results, thresholds and the prediction models derived at step 5 and step 6 respectively, can be used to realize a variety of management strategies that are aimed at controlling QoS parameters to guarantee a satisfactory level of user experience. Examples of QoE management strategies are illustrated in Sect. 6.

3.1 Statistical modelling technique

A good background literature on Discriminant Analysis can be found in [14]. It is a statistical technique that can be used to predict a qualitative attribute of an object from known values of its quantitative variables [16]. Specifically, this technique can be used to predict group membership based on a set of quantitative variables. Thus, we have employed it to answer the following questions in our study: Will a chosen set of QoS parameters lead to a reliable prediction of user ratings of service quality? What is the relative importance of the QoS parameters in predicting the group membership?

To use Discriminant Analysis, a set of observations (or test cases) with known quantitative variable values, and group memberships, is needed. Then, a set of linear functions of the quantitative variables, known as discriminant functions, is constructed. These functions best separate the group memberships. The maximum number of discriminant functions for an analysis with N group memberships and q quantitative variables, is either $(N-1)$ or q whichever is the smallest. The first discriminant function maximises the differences between the group means. Subsequent functions achieve the same goal but with the added constraint that they are not correlated with the values from the previous functions.

The classification procedure provides a means used to predict the group to which an observation (a case) with unknown group most likely belongs. This process is based on either the use of the discriminant functions or the linear combination of the quantitative variables. The latter uses the theory of maximum group differences to derive the classification functions. The formula is as follows [19]:

$$h_k = b_{k0} + b_{k1}X_1 + b_{k2}X_2 + \cdots + b_{kn}X_n \qquad (1)$$

where h_k = the classification score for group k; b_{ki} = the coefficient for variable i in the equation corresponding to group k; X_i = the value of the quantitative variable for variable i, $i = 1, \ldots, n$; b_{k0} = is a constant.

In simpler terms, the classification functions can be used to determine which group each response most likely belongs. There are as many classification functions as there are vote categories. For example, a two-group membership with two quantitative variables, (1) can be expanded to

$$
\begin{aligned}
h_1 &= b_{10} + b_{11}X_1 + b_{12}X_2, \\
h_2 &= b_{20} + b_{21}X_1 + b_{22}X_2
\end{aligned}
\tag{2}
$$

A case is classified as belonging to the group for which it has the highest classification score. The Wilks's lambda and the standardized discriminant coefficients are other important statistics produced when performing discriminant analysis. These statistics help to determine how accurate a prediction is, and how much influence each quantitative variable has on predicting the group membership.

Wilks's lambda (Λ): The Wilks's lambda is used to test if the discriminant function is significant (*i.e.*, how accurate the prediction will be). The Wilks's lambda values vary between 0 and 1. Values that are near zero denote high discrimination. The significance of lambda is determined through the p parameter (or sig. for significance), similar to ANOVA's significance.

Standardized discriminant coefficients: The standardized discriminant coefficient indicates the relative importance of the quantitative variables in predicting the qualitative attribute. The relative importance is obtained by examining the magnitude of the standardized coefficients (ignoring the sign). The larger the magnitude, the greater is that variable's contribution.

4 QoE data gathering

Referring to the fundamental element of QoE management, as illustrated in Fig. 1, we have previously reported our experimental findings (subjective QoE data[1]) related to steps 1 to 5 for different applications/services in [1, 2, 5] and [3, 6]. These findings are the outcome of a 3-year research work. These experiments have been carefully designed under controlled network laboratory environments and then we collected the user's feedback using three types of terminals (mobile phone, PDA and a laptop).

The subjective QoE data used in conducting discriminant analysis was extracted from [1], and also where a detailed

description of that study is given. Thus, only a short summary is included herein. In [1] the study aimed at determining the acceptability thresholds for different content types (News, Action movie, Romance movie, Comedy, Cartoon, Music and Sports) at which video quality became unacceptable on three different terminal types. The terminals (Mobile phone—Nokia N70, PDA—HP Ipaq rx1960, and a Laptop[2]—Sony FR315B), which were functionally similar but having different characteristics in terms of display screen sizes and device capabilities, were selected to represent the common devices used for streaming services. The classical psychophysics Method of Limits [9] was used in determining the thresholds, by changing a single stimulus in successive and discrete steps either in ascending or in descending series. A series terminates when the intensity of the stimulus becomes detectable. The subjects gave a binary response of "yes" or "no" when the stimulus is perceived. In adopting this method, we gradually decreased the values of AQoS parameters (video encoding bitrate and frame rate). The idea was to gradually reduce the video quality to determine the minimum level of quality at which the subject's QoE became unacceptable. The audio quality was kept constant because audio consumes less bandwidth relative to video, and also, previous research suggests that better audio increases the overall audiovisual quality ratings [7] and [13].

When the video clips were presented in an ascending series (*i.e.,* from bad quality to good quality), users' threshold of quality acceptability was higher than the case for the decreasing series. Higher thresholds, led to resource over-provisioning. For this reason, it was concluded that studies using the decreasing series were best for implementing QoE management strategies.

5 Results of the discriminant analysis

The results obtained from applying discriminant analysis on the QoE data are presented here. The use of the classification function coefficients (1), Wilks's lambda (Λ) test for significance p and the standardized discriminant coefficients (see Sect. 3) are exemplified in the remainder of the paper. These results are a collection of QoE prediction models (a model being a set of classification functions) for a variety of content types delivered onto a mobile phone, a PDA and a laptop terminal.

5.1 Mobile phone prediction models

5.1.1 Wilks's lambda test for model significance

Table 1 shows the Wilks's lambda test of significance of the discriminant functions obtained for each content type. The

---

[1]The correlation between users' QoE perception with the generated QoS test cases.

[2]The actual image size used was 640×480 pixels.

Table 1 Discriminant analysis showing Wilks's lambda value and its significance for the mobile phone terminal

Content → types	News	Romance movie	Cartoon	Comedy	Music video	Top gear	Action movie	Music concert	Cricket	Football
Wilks's lambda (Λ)	0.760	0.760	0.747	0.780	0.678	0.742	0.516	0.468	0.481	0.578
Significance (p)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

p-values are highly significant ($p < 0.001$), indicating that for each content type, the means of the discriminant function scores are not similar between the two respective qualitative group responses. We concluded that the functions discriminated well for all of the content types.

5.1.2 The standardized discriminant function coefficients for the mobile phone terminal

Tables 2a and 2b show the standardized discriminant function coefficients for each content type. For the content types in Table 2a, the relative contribution of the video encoding bit rate was more important than that of the frame rate. Discriminant functions of content types in Table 2b were more influenced by the frame rate.

The results from Tables 2a and 2b provide invaluable insights towards network resource dimensioning strategies. The results show how AQoS parameter control mechanisms can be applied to deliver the best possible video quality experience, whilst reducing resource utilisation, based on the type of content for a mobile phone terminal. For example, for the content types in Table 2a, the frame rate can be significantly reduced to preserve the bits per frame (*i.e.*, to keep the spatial sharpness of the image).

For the content types in Table 2b, the video bit rate (bits per frame) can be reduced whilst the frame rate (temporal continuity) is kept high. An example of a resource control strategy, which uses this information, is illustrated in Sect. 6.

5.1.3 The classification function coefficients for the mobile phone terminal

Equations (3) and (4) show the derived classification functions for two content types[3]—news and action movie content respectively. The derived coefficients varied considerably depending on the type of content. For a given case (*i.e.*, a combination of video encoding bit rate and frame rate), the classification score (h) was computed for each classification function. The case was classified as belonging to the

Table 2 The standardized discriminant coefficients of the quantitative variables for the mobile phone terminal

Content types	Standardized coefficients of the quantitative variables	
	Video encoding bit rate	Frame rate
2(a)		
News	1.430	−0.449
Romance movie	1.430	−0.449
Cartoon	1.189	−0.196
Comedy	1.362	−0.376
Top gear	0.690	0.317
2(b)		
Music video	0.153	0.851
Action movie	−0.938	1.871
Music concert	−0.924	1.854
Cricket	−0.325	1.306
Football	−0.267	1.255

quality response for which it had the highest classification score.

News

$$h_{(Acceptable)} = -5.699 + (-0.080 \times \textit{Video encoding bit rate}) + (1.613 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -4.223 + (-0.104 \times \textit{Video encoding bit rate}) + (1.730 \times \textit{Frame rate}), \tag{3}$$

Action movie

$$h_{(Acceptable)} = -9.735 + (-0.111 \times \textit{Video encoding bit rate}) + (2.411 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -3.492 + (-0.087 \times \textit{Video encoding bit rate}) + (1.613 \times \textit{Frame rate}) \tag{4}$$

Taking the average across the content types, the mobile phone prediction models correctly classified 76.9% of the responses in our original set of test cases. To estimate the predictive accuracy of this model for a new set of cases with unknown QoE ratings, the leave-one-out technique[4]

---

[3]For the sake of brevity, we only show the classification functions for two content types but we have derived functions for 10 content types.

[4]Classification functions are derived based on all cases except one, and then only the omitted case is reclassified. This procedure is repeated

Table 3 Discriminant analysis showing Wilks's lambda value and its significance for the PDA terminal

Content → types	News	Cartoon	Romance movie	Comedy	Music video	Top gear	Music concert	Football	Action movie	Cricket
Wilks's lambda (Λ)	0.760	0.778	0.534	0.468	0.383	0.333	0.277	0.321	0.249	0.314
Significance (p)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

was used. We concluded that an equivalent proportion of cases would be accurately classified.

5.2 PDA prediction models

5.2.1 Wilks's lambda test for model significance

Table 3 shows the Wilks's lambda test of significance of the discriminant function obtained for each of the content types for the PDA. The means of the discriminant function scores were not similar between the two respective qualitative group responses, as indicated by the p-values ($p < 0.001$). We thus concluded that the functions discriminated well for all of the content types.

5.2.2 The standardized discriminant function coefficients for the PDA terminal

Tables 4a and 4b show the standardized discriminant function coefficients for each content type. For content types in Table 4a, the relative contribution of the video bit rate is more important than that of the frame rate, whilst the discriminant functions of content types in Table 4b were more influenced by the frame rate.

A comparison of Tables 2 and 4 shows that what quantitative variable a given content type is sensitive to, is dependent on the terminal type. For example on the mobile phone terminal, comedy and top gear contents were more sensitive to video bit rate. However, on the PDA terminal, they were more sensitive to frame rate. Information such as this one is instrumental to effective QoE management (Sect. 6).

5.2.3 The classification function coefficients for the PDA terminal

Equations (5) and (6) show the derived classification functions (prediction models), for the news and action movie content types respectively, on a PDA. Taking the average across the content types, the PDA prediction models correctly classified 86.6% of the responses in our original set of test cases. To estimate the predictive accuracy of this model

N times, and it provides as estimate of how well the classifications should predict with a new sample.

Table 4 The standardized coefficients of the quantitative variables for the PDA terminal

Content types	Standardized coefficients of the quantitative variables	
	Video encoding bit rate	Frame rate
4(a)		
News	0.833	0.171
Cartoon	1.248	−0.256
Romance movie	0.522	0.488
Cricket	0.579	0.437
4(b)		
Comedy	0.266	0.743
Music video	0.362	0.650
Top gear	0.320	0.694
Music concert	0.185	0.826
Football	0.211	0.800
Action movie	0.234	0.781

for a new set of cases with unknown QoE ratings, the leave-one-out technique was used. We concluded that an equivalent proportion of cases would be accurately classified.

News
$$h_{(Acceptable)} = -6.464 + (-0.073 \times \text{Video encoding bit rate}) + (1.747 \times \text{Frame rate}),$$
$$h_{(Unacceptable)} = -4.290 + (-0.083 \times \text{Video encoding bit rate}) + (1.706 \times \text{Frame rate}), \quad (5)$$

Action movie
$$h_{(Acceptable)} = -18.349 + (-0.065 \times \text{Video encoding bit rate}) + (2.670 \times \text{Frame rate}),$$
$$h_{(Unacceptable)} = -5.509 + (-0.076 \times \text{Video encoding bit rate}) + (1.947 \times \text{Frame rate}) \quad (6)$$

5.3 Laptop prediction models

5.3.1 Wilks's lambda test for model significance

Table 5 shows the Wilks's lambda test of significance of the discriminant functions obtained for each content type separately. The p-values were highly significant ($p < 0.001$),

Table 5 Discriminant analysis showing Wilks's lambda value and its significance

Content → types	News	Romance movie	Cartoon	Comedy	Music video	Top gear	Action movie	Music concert	Cricket	Football
Wilks's lambda (Λ)	0.631	0.631	0.631	0.469	0.469	0.432	0.413	0.441	0.165	0.214
Significance (p)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 6 The standardized coefficients of the quantitative variables for the laptop terminal

Content types	Standardized coefficients of the quantitative variables	
	Video encoding bit rate	Frame rate
News	−0.989	1.846
Romance movie	−0.989	1.846
Cartoon	−0.989	1.846
Music video	−0.780	1.657
Comedy	−0.780	1.657
Top gear	−0.758	1.634
Football	−0.749	1.624
Cricket	−0.246	1.212
Music concert	−0.703	1.526
Action movie	−0.501	1.387

indicating that for each content type, the means of the discriminant function scores were not similar between the two respective qualitative group responses.

5.3.2 *The standardized discriminant function coefficients for the laptop terminal*

Table 6 shows the standardized discriminant function coefficients for each content type. It can be seen that the contents were more sensitive to reductions in the frame rate. Due to the relatively larger image size of the laptop, a low frame rate value produced more jerkiness in the video playback as compared to the other terminals.

5.3.3 *The classification function coefficients for the laptop terminal*

Equations (7) and (8) show the derived classification functions that can be used to predict user perception based on a combination of application-level parameters for the news and action movie content types. Taking the average across the content types, the mobile phone prediction models correctly classified 83.9% of the responses in our original set of test cases. To estimate the predictive accuracy of this model

for a new set of cases with unknown QoE ratings, the leave-one-out technique was used. We concluded that an equivalent proportion of cases would be accurately classified.

News

$$h_{(Acceptable)} = -18.663 + (-0.085 \times \textit{Video encoding bit rate})$$
$$+ (2.653 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -9.156 + (-0.069 \times \textit{Video encoding bit rate})$$
$$+ (1.914 \times \textit{Frame rate}), \tag{7}$$

Action movie

$$h_{(Acceptable)} = -46.797 + (-0.107 \times \textit{Video encoding bit rate})$$
$$+ (5.021 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -19.636 + (-0.088 \times \textit{Video encoding bit rate})$$
$$+ (3.416 \times \textit{Frame rate}) \tag{8}$$

6 QoE management strategies

The following sections discuss examples that illustrate the use of the prediction models, the application of the strategies in bandwidth management, and in the provision of specific levels of service quality. We illustrate, by using selected examples, how the models enable comparisons between subjective conditions (user perceptions), and objective conditions (QoS parameter values) to be made. For instance, a mobile service provider can predict the end user perception based on the application and/or network level parameters values, and then adjust the allocation of resources such as bandwidth accordingly.

6.1 Use of prediction models

In order to demonstrate how each AQoS parameter affected the predictions, models for the following content types were chosen for illustration.

- Mobile phone—action movie and comedy
- PDA—comedy and cartoon
- Laptop—football

6.1.1 *Mobile phone models*

6.1.1.1 Action movie content model To illustrate the use of the mobile terminal's action movie prediction model (represented by the classification functions in (9)), its classification functions were evaluated for a range of AQoS cases,

Table 7 Predicted user ratings for action movie on the mobile phone. The point where the acceptability threshold (for the frame rate) is first reached is shown in bold type

Video encoding bit rate (kbps)	Frame rate (fps)	Classification scores		Predicted ratings
		$h_{(Acceptable)}$	$h_{(Unacceptable)}$	
256	25	22.124	14.587	Acceptable
224	20	13.621	9.3024	Acceptable
192	16	7.529	5.6312	Acceptable
128	**12**	**4.989**	**4.7408**	**Acceptable**
96	12	8.541	7.5217	Acceptable
64	12	12.093	10.3024	Acceptable
96	10	3.719	4.2956	Unacceptable
112	10	1.943	2.894	Unacceptable
128	10	0.167	1.5148	Unacceptable

starting with the case; 256 kbps and 25 fps for the video encoding bit rate and frame rate parameters respectively. The results of the predicted ratings for the set of cases are shown in Table 7.

$$h_{(Acceptable)} = -9.735 + (-0.111 \times \textit{Video encoding bit rate})$$
$$+ (2.411 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -3.492 + (-0.087 \times \textit{Video encoding bit rate})$$
$$+ (1.613 \times \textit{Frame rate}) \quad (9)$$

The results indicate that the values of the parameters could be reduced gracefully up to a point without affecting the user's perception of quality. It was found that frame rates below 12 fps were unlikely to be acceptable to the users, since jerkiness will be experienced. However, it was possible to reduce the video encoding bit rate further with the frame rate frozen at 12 fps. This was an unexpected result as the experimental studies reported in [1] to collate the acceptability data, did not anticipate this effect. As such, the experiments were not designed to collect subjective data regarding scenarios where all the parameter values were reduced until one of them reached an acceptability limit, whilst the other(s) could be further reduced. The result here showed that policy-based multi-parameter network optimization is possible. This is where a standard procedure for network optimization which guarantees user acceptance, is developed. An example procedure could go like this:

- For content type A, set values for the QoS parameters P1, P2 . . . PN such that the dominant parameter P1 is at its expected acceptability limit L1.
- For service type S1, additionally reduce parameter P2 to its expected acceptability limit L2
- For service type S2, additionally reduce parameters P2, P3 to their expected acceptability limits L2 and L3 respectively.

Note from the bottom part of Table 7 that the cases where the frame rate was lower than the 12 fps limit, the content

quality was still expected to be unacceptable even for increasing values of the encoding bit rate. This demonstrated that the frame rate was the major determinant of user perception.

6.1.1.2 Comedy content model The mobile terminal's comedy content prediction model is represented by the classification functions in (10). The classification functions were evaluated for a range of AQoS cases. The full results are shown in Table 8.

$$h_{(Acceptable)} = -5.732 + (-0.080 \times \textit{Encoding bitrate})$$
$$+ (1.620 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -4.229 + (-0.099 \times \textit{Encoding bitrate})$$
$$+ (1.702 \times \textit{Frame rate}) \quad (10)$$

Note from Table 8 that the cases where the video encoding bit rate was lower than the 128 kbps limit, the content quality was still expected to be unacceptable even for increasing values of frame rate. This demonstrated that the video encoding bit rate was the major determinant of user perception.

6.1.2 PDA models

6.1.2.1 Comedy content model The comedy content model for the PDA is represented by the classification functions in (11). The functions were evaluated for a range of QoS cases. The full results are shown in Table 9.

$$h_{(Acceptable)} = -9.842 + (-0.073 \times \textit{Video encoding bit rate})$$
$$+ (2.056 \times \textit{Frame rate}),$$
$$h_{(Unacceptable)} = -4.508 + (-0.079 \times \textit{Video encoding bit rate})$$
$$+ (1.741 \times \textit{Frame rate}) \quad (11)$$

From Table 9, note that the cases where the frame rate was lower than the 15 fps limit, the content quality was

Table 8 Predicted user ratings for comedy on the mobile phone. The point where the acceptability threshold (for the encoding bit rate) is first reached is shown in bold type

Video encoding bit rate (kbps)	Frame rate (fps)	Classification scores		Predicted ratings
		$h_{(Acceptable)}$	$h_{(Unacceptable)}$	
256	25	14.288	12.977	Acceptable
224	20	5.508	4.231	Acceptable
192	15	3.208	2.293	Acceptable
160	10	0.908	0.355	Acceptable
128	**10**	**0.228**	**0.119**	**Acceptable**
128	5	−7.872	−8.391	Acceptable
112	10	1.508	1.703	Unacceptable
112	12	4.748	5.107	Unacceptable
112	16	16.088	17.021	Unacceptable

Table 9 Predicted user ratings for comedy on the PDA. The point where the acceptability threshold (for the frame rate) is first reached is shown in bold type

Video encoding bit rate (kbps)	Frame rate (fps)	Classification scores		Predicted ratings
		$h_{(Acceptable)}$	$h_{(Unacceptable)}$	
128	12	5.486	6.272	Unacceptable
160	12	3.15	3.744	Unacceptable
200	12	0.084	0.426	Unacceptable
128	**15**	**11.654**	**11.495**	**Acceptable**
192	15	6.982	6.439	Acceptable
200	15	6.398	5.807	Acceptable

Table 10 Predicted user ratings for cartoon content on the PDA. The point where the acceptability threshold (for the encoding bit rate) is first reached is shown in bold type

Video encoding bit rate (kbps)	Frame rate (fps)	Classification scores		Predicted ratings
		$h_{(Acceptable)}$	$h_{(Unacceptable)}$	
256	15	0.833	−0.314	Acceptable
192	15	5.505	5.19	Acceptable
192	10	−2.985	−3.545	Acceptable
160	**10**	**−0.649**	**−0.793**	**Acceptable**
160	15	7.841	7.942	Unacceptable
128	15	10.177	10.694	Unacceptable
128	10	1.687	1.959	Unacceptable
96	10	4.023	4.711	Unacceptable
64	10	6.359	7.463	Unacceptable

still expected to be unacceptable even for increasing values of encoding bit rate. This demonstrated that the frame rate was the major determinant of user perception. This finding was contrary to that for comedy content on the mobile phone, which was more sensitive to the video bit rate.

6.1.2.2 Cartoon content model The cartoon content model for the PDA is represented by the classification functions in (12). The functions were evaluated for a range of AQoS cases. The full results are shown in Table 10.
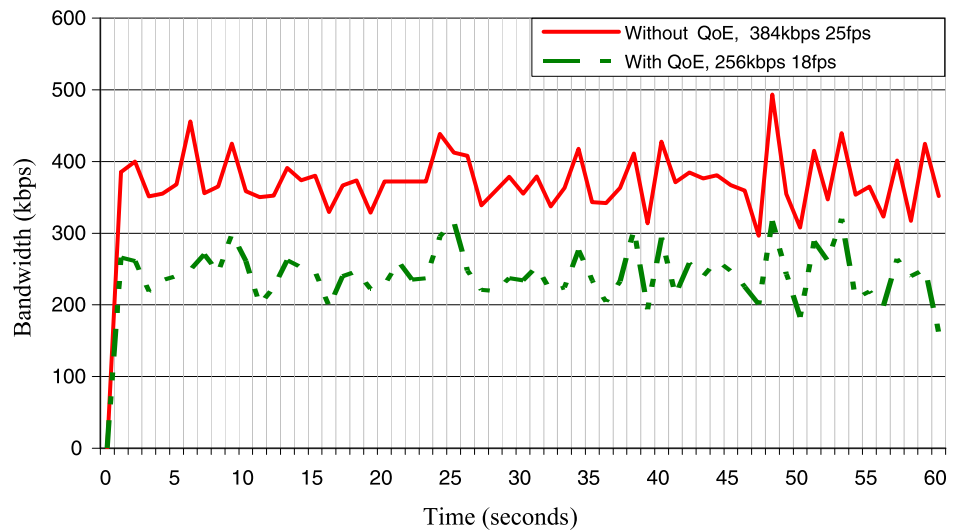
$$h_{(Acceptable)} = -14.731 + (-0.070 \times \textit{Video encoding bit rate}) + (2.420 \times \textit{Frame rate}),$$

$$h_{(Unacceptable)} = -5.209 + (-0.077 \times \textit{Video encoding bit rate}) + (1.873 \times \textit{Frame rate}) \quad (12)$$

Note from the bottom part of Table 10 that the cases where the encoding bit rate was lower than the 160 kbps

Table 11 Predicted user ratings for football content on the laptop. The point where the acceptability threshold (for the frame rate) is first reached is shown in bold type

Video encoding bit rate (kbps)	Frame rate (fps)	Classification scores		Predicted ratings
		$h_{(Acceptable)}$	$h_{(Unacceptable)}$	
256	25	33.636	27.71	Acceptable
208	20	21.351	19.464	Acceptable
192	18	16.133	15.916	Acceptable
160	18	19.173	18.412	Acceptable
144	18	20.693	19.66	Acceptable
192	**16**	**9.395**	**11.12**	**Unacceptable**
252	15	0.326	4.042	Unacceptable
224	12	−0.383	3.828	Unacceptable

Fig. 2 Comparison of bandwidth utilization by the action movie content type (with and without QoE optimization)

limit, the content quality was still expected to be unacceptable even for increasing values of frame rate. This demonstrated that the video encoding bit rate was the major determinant of user perception.

6.1.3 Laptop models

6.1.3.1 Football content model The football content model for the Laptop is represented by the classification functions in (13). The functions were evaluated for a range of QoS cases. The full results are shown in Table 11.

$$h_{(Acceptable)} = -26.269 + (-0.095 \times Video\ encoding\ bit\ rate)$$
$$+ (3.369 \times Frame\ rate),$$
$$h_{(Unacceptable)} = -12.272 + (-0.078 \times Video\ encoding\ bit\ rate)$$
$$+ (2.398 \times Frame\ rate) \quad (13)$$

From Table 11, for the cases where the frame rate was lower than the 16 fps limit, the content quality was still expected to be unacceptable even for increasing values of video encoding bit rate. This demonstrated that the frame rate was the major determinant of user perception. In fact,

all the contents types on the laptop terminal were sensitive to the frame rate parameter.

6.2 Bandwidth management using QoE on the mobile phone

An action movie clip was encoded as follows without any QoE considerations: frame rate at 25 fps, audio bit rate at 12.2 kbps and video bit rate at 384 kbps. The bandwidth consumed by this video clip over time is as shown in Fig. 2. The corresponding quantization parameter[5] (Qp) values over the same time range are shown in Fig. 3. Assuming a network LAN of 10 Mbps and a T1 connection (1.55 Mbps), about 24% of the T1 connection was utilized by the video clip.

However, if the QoE management approach were employed, the network's bandwidth utilization could be reduced whilst still providing an acceptable service quality. To minimize the use of the network's bandwidth, the values

---

[5]Low Qp values result in high quality for each encoded frame, but also in high bit rates and vice versa.

Fig. 3 Comparison of the quantization parameter values for both the action movie (with and without QoE optimization)
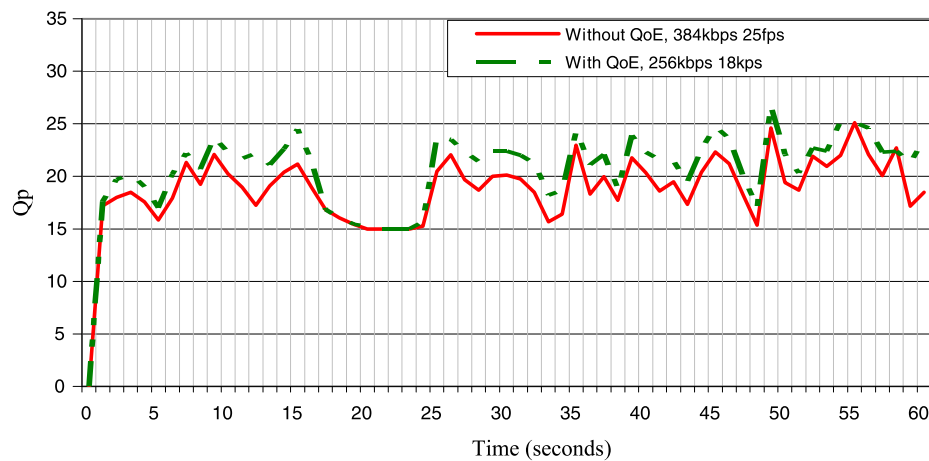
of the encoding parameters were strategically degraded using knowledge about the degree of influence each of them had on the end user perception. For example, for this action movie content played on a mobile phone terminal, we know that the typical user was more sensitive to reductions in the frame rate. Guided by this information, the video encoding bit rate for the action movie content was reduced from 384 kbps to 256 kbps, and the frame rate reduced from 25 fps to an acceptability limit of 18 fps. The bandwidth now being utilized by the action movie was reduced (see Fig. 2), whilst still maintaining the same picture quality as is shown by the similar Qp values for the "with and without" QoE scenarios in Fig. 3.

With QoE considerations employed in encoding the video clip, only 15% of the T1 connection was used, thus achieving approximately a 9% savings on network resources. When further decreases of bandwidth were necessary, the video encoding bit rate parameter was degraded first, since the users are less sensitive to it, as opposed to the frame rate. Further decrements in the frame rate were only made if the user's perception of quality was not adversely affected. With this type of control mechanism, the bandwidth usage could be decreased whilst still preserving the best possible video experience.

6.3 Provision of specific service levels using QoE on the mobile phone

QoE management can also be used in providing different levels of service quality (policy-based charging schemes). Consider a news content type that was encoded at a frame rate of 25 fps, with audio and video bit rate of 12.2 kbps and 384 kbps respectively, without any QoE considerations. Figures 4 and 5 respectively show bandwidth utilization and quantization parameter over a sixty-second time period. Assuming a network LAN of 10 Mbps, having a T1 connection (1.55 Mbps), about 24% of the T1 connection was used by this news video clip.

For the news content type on the mobile phone, the typical user was most sensitive to reductions in the video encoding bit rate. Because of this, the frame rate of the news content was significantly reduced without affecting the user's experience. It was gradually reduced to a limit of 8 fps. This limit preserved the number of bits per frame thereby preserving picture quality. The video encoding bit rate parameter was reduced to an acceptability threshold of 160 kbps. The network bandwidth was now significantly reduced as depicted in Fig. 4. It can be seen that when the frame rate was decreased, the number of bits in each frame increased thus improving the sharpness of the video clip. A comparison of the quantization parameter values for the "without QoE" and "with QoE" (Fig. 5) cases, yielded a comparable picture quality.

With QoE considerations, the news video clip used only 10% of the T1 connection, thus achieving approximately a 14% savings on network resources. When further decreases in bandwidth were required, the encoding video bit rate parameter was further degraded to a threshold of 64 kbps (see Fig. 4). In this way, the video quality was kept as high as possible within the available bandwidth whilst the user's QoE was hardly affected.

In this QoE approach, pricing could change depending on how QoE was maintained. For example, if a user were willing to accept greater degradations in QoE, the service would be cheaper thereby enabling user-centric service level agreements (SLAs). More importantly, policies and SLAs will directly target the user's experience, rather than the raw network parameters. Another advantage of the QoE approach is that service providers can selectively deliver specific levels of quality to specific market segments in such a way as to maximise their revenue.

7 Remarks

This study focused on a framework, which was developed to form the basis as an invaluable toolset that network and ser-

Fig. 4 Comparisons of bandwidth utilized by the news content type for the case (without QoE optimization, and two cases with QoE optimization)
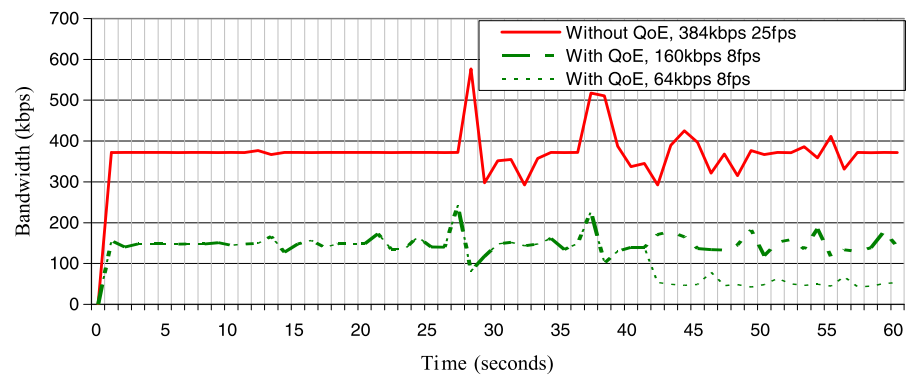
Fig. 5 Comparison of the quantization parameter values for the news content type for the case (without QoE optimization, and two cases with QoE optimization)
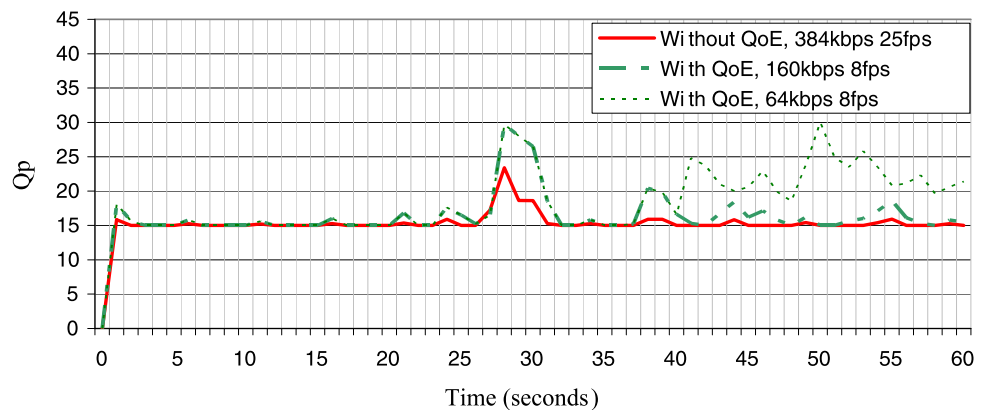
vice providers can apply to undertake network dimensioning and service provisioning strategies. The use of discriminant analysis has helped to determine the relationships between the AQoS metrics and the user's QoE. The correlations were found to be dependent on terminals and video content types. For example on the mobile phone, the quality ratings of comedy and top gear contents were mostly determined by changes in the video encoding bit rate (Sect. 5, Table 2). However, on the PDA, changes to the frame rate mostly determined the perceived video quality (Sect. 5, Table 4). On the laptop, it was found that the perceived quality for all the contents types were more sensitive to reductions in the frame rate (Sect. 5, Table 6). The strength and accuracy of the QoE prediction models laid in the fact that they were derived directly from the subjective data. The experimental results showed that the predictions correlated very well with subjective quality, which represents the user's perception of a service.

The proposed framework is sufficiently general in terms of its applicability to QoE-based management for a range of services. In order to give yet another example of how the framework developed in this study may be employed, it was applied to web browsing services [23, 24], where web browsing applications were particularly sensitive to network parameters (such as bandwidth, delay and jitter). Using discriminant analysis, prediction models were formulated to estimate user opinion of service quality based on the network parameters. Jitter appeared to have had minimal influence on the user experience in web browsing. On the other hand, the amount of delay exhibited by the network was the most influential factor that affected users' experiences, followed by the network's bandwidth. To investigate the versatility of the framework, a feasibility study was conducted where it was applied to P2P TV system (Joost) [4]. P2P systems are considered to still be in their infancy and as such, not a lot is known about their QoE characteristics. QoE prediction models were established considering network contextual conditions and the terminal types that will be used to access the services, in order to achieve context-awareness with optimal QoE management.

8 Conclusions and future works

In this paper, we proposed a QoE management methodology. Discriminant analysis, a statistical modelling technique, was used as part of the framework in implementing the methodology. The degree of influence of AQoS parameters on user perception was determined and then led for the development of prediction models for multimedia services on mobile terminals. For the mobile phone terminal, the average prediction accuracy was 76.9%. For the PDA

terminal, the average prediction accuracy was 86.6%. For both terminals, the perceived quality of some contents types was more sensitive to the video bit rate, whilst others were sensitive to the frame rate. The laptop terminal had a prediction accuracy of 83.9%; the quality of all the content types we worked with, were sensitive to the frame rate. This was probably due to its relatively larger screen size. The QoE management strategies have illustrated that, by knowing the parameters that users were most sensitive to, an application could be adapted to use minimal network bandwidth, whilst at the same time maintaining an acceptable service quality. Another advantage of these control mechanisms is that, they could aid in mitigating the effects of network congestion and data packet losses in applications such as video streaming. This is because the mechanisms reduce the output bit rate from the encoder to the network. We have shown how the prediction models will allow network operators to anticipate the user's experience based on QoS conditions, in order to allocate network resources accordingly.

This study has contributed towards the relatively new but growing discipline of QoE management in content delivery systems. It has done so by examining the degree of influence of AQoS parameters on user perception for video services on mobile terminals, and then led for the development of their QoE prediction models. However, certain areas have been identified were more research is required.

The acceptability thresholds used in this study were affected by only the artefacts generated by the video encoding process. As such, they did not account for the effects of the artefacts introduced during transmission. Further experiments need to be performed in which the network parameters are also included in the "QoS to QoE" models.

Through the process of specifying the QoE prediction methodology adopted herein we have learned some important lessons:

(1) Linear dependencies can accurately model user predictions only within limited boundaries
(2) It is important to evolve the modelling process in such a way as to reduce the amount of manual collection of subjective data.

We have now started working on building models based on machine learning techniques, which seem to be more adequate in addressing these two issues. Initial findings seem to indicate that Support Vector Machines and Decision Tree based approaches may be more accurate than linear models [17]. Furthermore the amount of subjective data required can be reduced via automated online learning tools [18]. Machine learning seems therefore to be the most promising way to further the work presented in this manuscript.

References

1. Agboma, F., & Liotta, A. (2007). Addressing user expectations in mobile content delivery. *Mobile Information Systems, 3*, 156–164.
2. Agboma, F., & Liotta, A. (2008). QoE for mobile TV services. In *Multimedia transcoding in mobile and wireless networks*, 2008.
3. Agboma, F., & Liotta, A. (2008). QoE in pervasive telecommunications systems. In *Computer communications and networks 2008*. Berlin: Springer.
4. Agboma, F., & Liotta, A. (2009). *Quality of experience management in mobile content delivery systems*. Submitted for the award of the degree of Ph.D., University of Essex.
5. Agboma, F., & Liotta, A. (2008). *QoE-aware QoS management*. Conference paper presented at the 6th international conference on advances in mobile computing and multimedia, Linz, Austria, 24–26 November 2008.
6. Agboma, F., Smy, M., & Liotta, A. (2008). QoE analysis of a peer-to-peer television system. In *Proc. of IADIS telecommunications networks and systems*, 2008.
7. Aldridge, R. (1996). *Continuous quality assessment of digitally-coded television pictures*. Ph.D. Thesis, University of Essex.
8. De Vleeschauwer, B., Van de Meerssche, W., Simoens, P., De Turck, F., Dhoedt, B., Demeester, P., Struyve, K., Van Caenegem, T., Gilon, E., Dequeker, H., & Six, E. (2008). *Autonomic QoE optimization in the access node knowledge plane*. Paper presented at the broadband Europe, 2008.
9. Fechner, G. T. (1966). In D. H. Howes, E. G. Boring (Eds.), *Elements of psychophysics*. Translated by Alder, H. E. Holt Rinehart & Winston.
10. Goodchild, J. (2005). Integrating voice, data and video. In *IP video implementation and planning guide*, 2005. United States Telecom Association.
11. ITU-R (2002). *Recommendation BT500, methodology for the subjective assessment of the quality of television pictures*.
12. ITU-T (1999). *Recommendation P.910, subjective video quality assessment methods for multimedia applications*.
13. Joly, A., Montard, N., & Buttin, M. (2001). Audio-visual quality and interactions between television audio and video. In *Sixth international symposium on signal processing and its applications*, 2001.
14. Klecka, W. R. (1980). Discriminant analysis. In *Quantitative applications in the social sciences* (Vol. 19, p. 71). Thousand Oaks: Sage
15. López, D., Gonźalez, F., Bellido, L., & Alonso, A. (2006). Adaptive multimedia streaming over IP based on customer oriented metrics. In *IEEE international symposium on computer networks*, 2006.
16. Lubin, A. (1950). Linear and non-linear discriminating functions. *British Journal of Psychology, Statistical Section, 3*, 90–103.
17. Menkovski, K., Oredope, A., Liotta, A., & Cuadra-Sanchez, A. (2009). Predicting quality of experience in multimedia streaming. In *Proc. of the 7th international conference on advances in mobile computing and multimedia*, Kuala Lumpur, Malaysia, 14–16th December 2009.
18. Menkovski, K., Oredope, A., Liotta, A., & Cuadra-Sanchez, A. (2009). Optimized online learning for QoE prediction. In *Proc. of the 21st Benelux conference on artificial intelligence*, Eindhoven, The Netherlands, 29–30 October 2009.

19. Perkis, A., Munkeby, S., & Hillestad, O. I. (2006). A model for measuring quality of experience. In *Signal processing symposium proc. of the 7th nordic*, 2006 (pp. 198–201).

20. Siller, M. (2006). *An agent-based platform to map quality of service to experience in active and conventional networks*. Ph.D. Thesis, University of Essex, 2006.

21. Soldani, D., Li, M., & Cuny, R. (2006). *QoS and QoE management in UMTS cellular systems*. New York: Wiley.

22. Watanabe, K., Yamagishi, K., Okamoto, J., & Takahashi, A. (2008). Proposal of new QoE assessment approach for quality management of IPTV services. In *IEEE international conference on image processing*, 2008.

23. Yaici, K. (2004). *User perceived Quality of Service (QoS) in 3G mobile networks (UMTS)*. MRes Thesis, University of Surrey, 2004.

24. Yaici, K., Liotta, A., Zisimopoulous, H., & Sammut, T. (2004). *User-centric Quality of Service Management in UMTS*. Poster presented at the 4th Latin American network operations and management symposium, Springer, 2004.

Florence Agboma received her Ph.D. and M.Sc. degrees from the University of Essex, United Kingdom in 2009 and 2005 respectively. She received her B.Sc. degree in Electronic Communication Systems from the University of Plymouth, United Kingdom in 2004. Her research interests include quality of experience management, multimedia applications and peer-to-peer systems. Email: fagbom@gmail.com

Antonio Liotta holds the Chair of Communication Network Protocols at the Eindhoven University of Technology (The Netherlands) where he leads the Autonomic Networks team. He was previously a Reader at the University of Essex (UK), where he led the Pervasive Services team. Antonio is a Fellow of the UK Higher Education Academy and serves the Peer Review College of the UK Engineering and Physical Sciences Research Council, and the Advisory Board of Editors of the Journal of Network and System Management (Springer) and of the International Journal of Network Management (Wiley).

He is an active member of the networking and network and service management research communities. He has co-organized and co-chaired several international conferences; has served the Technical Programme Committee of dozens of conferences; and has also contributed as keynote and tutorial speaker. Antonio has over 100 publications to his credit, in the areas of telecommunication services, distributed computing, and autonomic networks and systems. Recent articles have contributed to topical themes including: operator-mediated peer-to-peer systems; P2P streaming; quality of experience management; application-aware networks; context-aware adaptive services; and self-configurable networks.