Predicting Quality of Experience in Multimedia Streaming

Vlado Menkovski
v.menkovski@tue.nl

Adetola Oredope
a.oredope@tue.nl

Antonio Liotta,
a.liotta@tue.nl

Department of Electrical Engineering
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven
The Netherlands

Antonio Cuadra Sánchez
cuadras@tid.es

Telefonica R&D
6 Emilio Vargas
28043 Madrid, Spain

ABSTRACT

Measuring and predicting the user's Quality of Experience (QoE) of a multimedia stream is the first step towards improving and optimizing the provision of mobile streaming services. This enables us to better understand how Quality of Service (QoS) parameters affect service quality, as it is actually perceived by the end user. Over the last years this goal has been pursued by means of subjective tests and through the analysis of the user's feedback. Existing statistical techniques have lead to poor accuracy (order of 70%) and inability to evolve prediction models with the system's dynamics. In this paper, we propose a novel approach for building accurate and adaptive QoE prediction models using Machine Learning classification algorithms, trained on subjective test data. These models can be used for real-time prediction of QoE and can be efficiently integrated into online learning systems that can adapt the models according to changes in the environment. Providing high accuracy of above 90%, the classification algorithms become an indispensible component of a mobile multimedia QoE management system.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations – Network Management. C.4 [**Performance of Systems**]: Modeling Techniques. I.2.6 [**Learning**]: Induction

General Terms

Algorithms, Performance, Management, Monitoring.

Keywords

QoE, QoS, QoE Management, QoS Monitoring, Machine Learning.

1. INTRODUCTION

The evolution of telecommunication environments introduces new services like streaming of multimedia content and live television. These services are proven to be exceedingly bandwidth intensive and they put particular strain on the network resources which in turn puts more significance on efficient management of network resources. In order to supply improvement in the Quality of Service (QoS) management we look at a more user-centric approach to QoS management. This approach is focused on predicting how QoS parameters actually affect the user's Quality of Experience (QoE), given the conditions in which the service is used.

The QoE approach aims at maximizing the perceived user experience while minimizing the impact on the network resources. This approach looks at a larger set of parameters in addition to the QoS attributes. These parameters affect the perceptual quality of the displayed content. It is being observed [1] that on the perceived QoE affect conditions like the type of the terminal, size of the viewing area, the type of the multimedia content, video encoding bitrate, video framerate, audio encoding framerate as well as the underlying network-related parameters, such as packet loss, delay, jitter and etc [2].

In order to effectively manage the QoS for these services it is imperative to find the correlation between all these conditions and the perceived QoE. By modeling these correlations we can build high accuracy prediction models, which then allow for online management of the QoS parameters or more precisely the efficient management of the network resources.

As a starting point we used work done in area QoE based network management, particularly subjective analysis [2]. However, our focus is on building prediction models using more advanced statistical methods like Machine Learning (ML). We are expecting to achieve high accuracy and a possibility for dynamic adaptation of these models. Looking at the initially available data from subjective tests [1] we are convinced that the ML classification techniques will provide means for successful implementation of high accuracy prediction models.

The remainder of the paper is organized in the following manner; we are exploring related work in section 2, then we continue to a discussion of the proposed ML techniques in Section 3. Next section 4 is explains the experimental design and results. We finalize with Section 5 which addresses the conclusions and future work.

2. RELATED WORK

2.1 QoE Perspectives

When we are talking about quality of experience of multimedia content most discussions agree on the general statement that the

QoE is what the end-user experiences while using the service. However, when we are talking about measuring or predicting the QoE we come to realize that many different approaches have been proposed towards this goal.

The main reason for this differentiation is the fact that QoE is a subjective measure and subjective quality assessment is the only relevant method of measurement [3]. But since subjective experiments are expensive [4] and in some applications reference content is not available many efforts are being made (including this one) in creating an objective methodology for measuring QoE.

There are efforts that are focused on the multimedia content compression, some on the transportation and others look at both compression and transport.

2.2 Objective QoE Measurements

When looking only on the fidelity of the audio and video signals researchers build models that predict the QoE by looking at the signal distortion. This data metrics remain oblivious to the content of the media as well as the human perception mechanisms.

It is shown that techniques as Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR) are lacking the understanding of the Human Visual System (HVS) due to which they deliver unsatisfactory results [4].

There are models that belong to the group of Media-layer models which observe the content and implement objective perceptual video quality measurement by modeling the HVS [5]. These models are computationally expensive because they need to execute in-depth analysis of the media content.

When focusing on the effects that the transport of the content has on the delivered quality, the direct approach is focusing alone on the QoS parameters. This yields weaker results [6]. The work in [6] proposes looking at the problem in three layers. The bottom layer being the network layer produces the QoS parameters or more particularly the Network QoS (NQoS) parameters. Above this layer they present the application layer which is concentrated on parameters like resolution, frame rate, color, codec type and so on, these are referred to as AQoS (Application QoS). The third or the top layer is the perception layer which is driven by the human perception of the multimedia content and is concentrated on spatial perception, temporal perception and acoustic bandpass [6]. The QoE which is measured on the top layer is looked at as a function of both AQoS and NQoS. $Q = f(AQoS, NQoS)$.

The proposed framework in [6] discusses that arbitrating all of the QoS parameters together are significantly more effective in maximizing the QoE than looking at each of them individually.

These efforts as well as all of the previous ones conclude that the single relevant reference that the objective methods are measured, are the subjective tests. In the following section of this paper we will discuss an approach based on initial subjective tests that correlates the QoS parameters of those tests to deliver prediction models for the QoE. These models in addition to the QoS parameters observe the content type and the terminal or console where the content is viewed to deliver accurate predictions.

2.3 Subjective Testing

The work done in [1][2] is based on subjective testing and is of significant benefit to this effort because we can be convinced that these results are accurately conveying the user satisfaction from the service. This is also the factual target of the service providers and the single relevant metric for the end-user satisfaction.

The method in [1] is known as Method of Limits [7]. It is used to detect the thresholds by changing a single stimulus in successive, discrete steps. A series terminates when the intensity of the stimulus becomes detectable. So if the quality decreases bellow the satisfactory level by the user the test series stops. The purpose is to determine the user thresholds of acceptability and the QoS parameters taking into account also the type of the content and terminal type. The following tables (Tab. 1-3) depict the different video samples that the users were subjected to. Each segment is a row in the tables and has corresponding QoS parameters given in the column values.

Segment	*Time (seconds)*	*Video bitrate (kbit/s)*	*Audio bitrate (kbit/s)*	*Frame rate*
1	1-20	384	12.2	25
2	21-40	303	12.2	25
3	41-60	243	12.2	20
4	61-80	194	12.2	15
5	81-100	128	12.2	12.5
6	101-120	96	12.2	10
7	121-140	64	12.2	6
8	141-160	32	12.2	6

Table 1. Test-bed combinations for 3G Mobile Phone descending series (174 x 144 image size)

Segment	*Time (seconds)*	*Video bitrate (kbit/s)*	*Audio bitrate (kbit/s)*	*Frame rate*
1	1-20	448	32	25
2	21-40	349	32	25
3	41-60	285	32	20
4	61-80	224	32	15
5	81-100	128	32	10
6	101-120	96	32	10
7	121-140	64	32	6
8	141-160	32	32	6

Table 2. Test-bed combinations for PDA descending series (320 x 240 image size)

Segment	*Time (seconds)*	*Video bitrate (kbit/s)*	*Audio bitrate (kbit/s)*	*Frame rate*
1	1-20	448	32	25
2	21-40	349	32	25
3	41-60	285	32	20
4	61-80	224	32	15
5	81-100	128	32	10
6	101-120	96	32	10
7	121-140	64	32	6
8	141-160	32	32	6

Table 3. Test-bed combinations for Laptop descending series (640 x 480 image size)

The test was done on a group of people and the results for each test were compiled together in Figure 1.

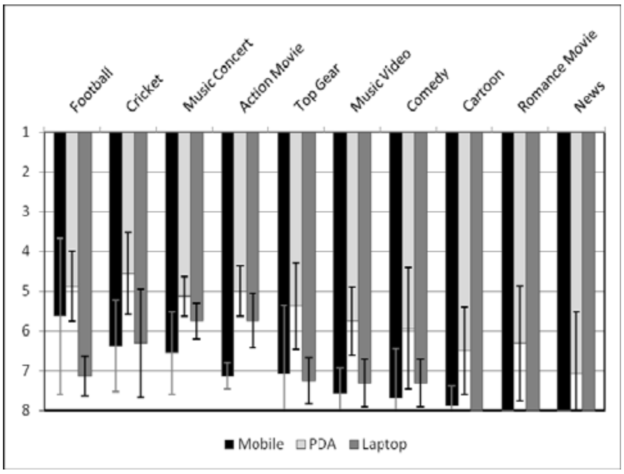

Figure 1. QoE Levels for different content types on all terminals

It is obvious from the results that the QoE changes considerably with the terminal type (screen size, position, mode of presentation) as well as the content presented. We observe that multimedia content over mobile phones requires far less bandwidth than over larger terminals as laptops, but we can also clearly observe that content like news broadcast or 'head and shoulders' shot is far less depended on high QoS in delivering satisfactory QoE on any terminal [1].

There are other correlations that are also important like in Football matches the loss of detail particularly affects the QoE even in small terminals like mobile phones in oppose to a Romance movie for example. In addition it is observed that the Audio quality is significantly more important than Video quality in most contents but particularly in News coverage [2].

It is of great importance to capture these correlations in order to derive the dependencies of the QoS and other parameters to the QoE.

2.4 Statistical Analysis on Subjective tests

Statistical analysis of the subjective tests is a very suitable technique and yields satisfactory results. In [2] the subjective test data was analyzed with the Discriminant Analysis method [8]. This method builds linear functions for each class or label that the data point can be associated with. The class of the function that returns the maximum value is associated with the class of the data point. In this analysis the data was divided in subsets for each terminal and then in smaller subsets for each content type. The discriminant functions are built based in two input parameters Video Bitrate and Video Framerate.

In Fig. 2 you can see two classification discriminant functions, one for the News content and the second for the Action move, both of which are for the mobile terminal. Now by knowing the bitrate and the framerate of the video of News broadcast we can calculate h_{accept} and $h_{unaccept}$. If h_{accept} is larger than $h_{unaccept}$ we predict that the News broadcast is with acceptable QoE and vice versa.

News

$h_{(acceptable)} = -5.699 + (-0.080 \times \mathit{VideoBitrate}) + (1.613 \times \mathit{FR})$

$h_{(unacceptable)} = -4.223 + (-0.104 \times \mathit{VideoBitrate}) + (1.613\text{x}\mathit{FR})$

Action Movie

$h_{(acceptable)} = -9.735 + (-0.111 \times \mathit{VideoBitrate}) + (2.411 \times \mathit{FR})$

$h_{(unacceptable)} = -4.223 +(-0.104 \times \mathit{VideoBitrate}) + (1.613 \times \mathit{FR})$

Figure 2. Classification functions created using the discriminant analysis method

These work generated prediction models for all of the listed content types for these three terminals. The accuracy of those models validated with the leave-out-one method for each terminal averaged are [9]:

- Mobile phones: 76.9%
- PDA: 86.6%
- Laptop: 83.9%

In the following section we will describe the motivation and implementation of a different method of statistical analysis of the data from the subjective tests based on Machine Learning techniques which generated models with superior accuracy.

3. ML TECHNIQUES USED

A major role of ML research is automated induction of models for classification based on raw data that bears statistical importance to the defined classes or labels. This approach is frequently used in decision support systems, which in this current research is also applicable. We are exploring data from the subjective QoE test that is suitable for building prediction models for classifying the QoE on yet unseen cases.

The algorithms used here belong to the group of supervised learning algorithms. The training data is classified by a human or acquired experimentally. In our case the two classes on which the data instances belong are "acceptable" and "not acceptable" referring to the expected QoE. The algorithms need to infer the correlation between the input parameters and the output and build the classification rules into the models. The models basically map the combination of input parameters to a class value. The training dataset does not posses' all possible combinations of the values of input parameters hence the ML algorithm needs to infer the decision rules based on the available data, but also in the process the model needs not to be too specific so it does not adopt errors or noise from the data. In addition the overall goal of the ML algorithm is also to build an efficient model that will classify unseen data as fast as possible with maximum precision.

The generalization problem is particularly important when the training data is introduced to noise. We want the prediction model to be general enough not to take into account specifics like introduced noise or error in the measurements. Since this is a subjectively gathered data we expect that there are certain inconsistencies between the subjective information inputted from different people, these inconsistencies can be also looked at as input noise or more precisely errors in measurements.

We will present two classification algorithms, one function based that will build a classification hyper-plane and one that is tree based that will build a decision.

3.1 Function based, Support Vector Machine

The first method is called a Support Vector Machine (SVM), is of functional type and works by first plotting the data in an n-dimensional space, n being the number of attributes. In case of nominal (or discrete) attributes the algorithm creates another axis for each value of the nominal attribute. In our case the terminal type is a nominal attribute and the algorithm will create one variable for each of the three values. These variables can take values of one or zero depending if the attribute has that particular nominal value or not (Figure 3).

After plotting the data in the n-dimensional space the SVM algorithm will build a hyperplane which separates the data in an optimal manner [10] in regards to the two classes, if there are more classes the SVM will generate one hyperplane for each combination of pair of classes. Now by substituting the values of the attributes we can see if the particular datapoint is above or below the hyperplane, meaning that it belongs to one or the other class.

The particular implementation of SVM we used is called Sequential Minimal Optimization [11].

3.2 Decision Tree based approach

The decision tree algorithm is called C4.5 [12] which is an extension of the Id3 [13] tree induction algorithm. Both algorithms work by splitting the initial set of data into smaller subset so that the subset contains as much as possible datapoints from the same class. The split is done on a single attribute by its value, the algorithm searches for the split that will bring maximum information gain. These steps are then repeated for each subset until the subset is clean enough in regards to the classes associations. At which case the node is declared a leaf and the dominant class is associated with this leaf. The procedure is recursive. When the process of building the tree has finished C4.5 executes a process of pruning the decision tree this process is focused towards achieving a more general model, which is independent of specifics in the data like measurement errors or noise.

After the decision tree is finalized the classification process begins with inputting a new unseen case. The values of the attributes for this case are tested on each node of the decision tree starting from the top one. Depending on the result of each test we follow one or another branch of the decision tree until we reach a leaf. The class association of the leaf is the classification output of the decision tree.

4. Experimental Design and Analysis

4.1 Data Adaptation

The data of the subjective tests was acquired from the questionnaires [9] grouped by terminal type and content type. In the first step of this approach we rearranged the data in a more compact format. A sample of each datset is presented in Table 4a, 4b, and 4c. Each type of multimedia content has different Spatial and Temporal Information which is given in the first two columns of Tables 4.

The datasets are randomized and ready for use as training data for classification algorithms. We used the Weka ML suite [14] for building the ML models. Weka contains implementation of the C4.5 algorithm called J48, as well as an implementation of the SMO algorithm.

Video SI	*Video TI*	*Video Bitrate (kb/s)*	*Video Frame-rate*	*QoE Accept*
70	141	64	6	no
60	153	128	12.5	yes
21	187	64	6	yes
21	187	32	6	no
54	142	32	6	no

Table 4a. Sample set of the Mobile QoE subjective test data

Video SI	*Video TI*	*Video Bitrate (kb/s)*	*Video Frame-rate*	*QoE Accept*
62	100	285	20	yes
56	87	32	6	no
62	119	349	25	yes
62	119	448	25	yes
71	125	128	10	yes

Table 4b. Sample set of the PDA QoE subjective test data

Video SI	*Video TI*	*Video Bitrate (kb/s)*	*Video Frame-rate*	*QoE Accept*
56	35	180	20	yes
70	71	128	15	no
67	70	32	10	no
23	130	363	25	yes
63	90	64	12.5	yes

Table 4c. Sample set of the Laptop QoE subjective test data

The models that are built using this data are suitable for use in an online learning system. We propose use of on online learning schema Hedge β [15] that builds an ensemble of classifiers which use weighted voting to come to a decision.

4.2 SVM Results

The results from running the SMO algorithm generated a hyper plane given in Figure 3a, 3b and 3c. The accuracy of the Support Vector Machine was validated with the 10-fold cross validation method and was measured to 88.59±2.85%, 89.38±2.77%, 91.45±2.66%; for the mobile, PDA and laptop dataset respectively. The value is given in percentage of accurately predicted instances averaged over the ten folds with the standard deviation also given as an error range for the ten values.

```
     1.4555 * (normalized) Video SI
 +   1.0459 * (normalized) Video TI
 -   5.0892 * (normalized) Video Bitrate
 -   3.7632 * (normalized) Video Framerate
 -   0.4582
```

Figure 3a. SVM hyperplane for the Mobile dataset

```
     1.4229 * (normalized) Video SI
-    0.4575 * (normalized) Video TI
-    4.2913 * (normalized) Video Bitrate
-    3.1618 * (normalized) Video Framerate
+    1.3385
```

Figure 3b. SVM hyperplane for the PDA dataset

```
     2.5405 * (normalized) Video SI
+    0.6061 * (normalized) Video TI
-    4.2157 * (normalized) Video Bitrate
-    4.3957 * (normalized) Video Framerate
-    0.7474
```

Figure 3c. SVM hyperplane for the Laptop dataset

Figure 4 shows eight projections of the SVM hyperplane built from the mobile dataset on a two dimensional plane. The two dimensions here are the Video Spatial Information and the Video Temporal information. The eight projections are for the eight different values of the Video Bitrate and Framerate for the eight tested segments. The ten tested content types are also present on the graph as diamonds point each taking its position according to its Video Spatial Information and Temporal Information. We can observe that all the points are between the seventh and the eight SVM projection. The SVM classifies all points below its boundary as QoE acceptable and all point above its boundary as QoE unacceptable. This means that this SVM will classify any of the ten types of content as QoE acceptable if the Video Bitrate and Frame rate are at level seven or above (64kbit/s and 6 fps).

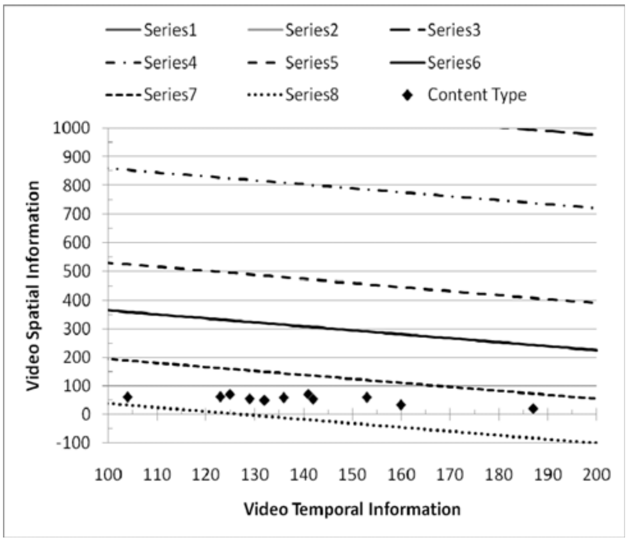

Figure 4. Projection of the SVM hyperplane from mobile data on a Video SI - Video TI plane

If in the feature for example a new type of video with higher complexity is introduced and its SI/TI point is above projection seven then the stream would need to be with video bitrate 96kbit/s and 10 frames per second for the classifier to find it QoE acceptable. For the reasons of space we only presented a projection for the mobile SVM hyperplane.

4.3 Decision Tree Result

The results from the C4.5 experiments yielded three decision trees. They all have high classification accuracy of 93.55±1.76%, 90.29±2.61% and 95.46±2.09% for the Mobile, PDA and Laptop dataset respectively. The graphical representation of the models is shown on Figure 5a, 5b and 5c.

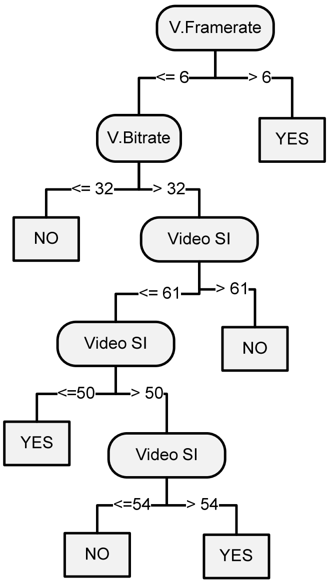

Figure 5a. Decision Tree for the Mobile dataset

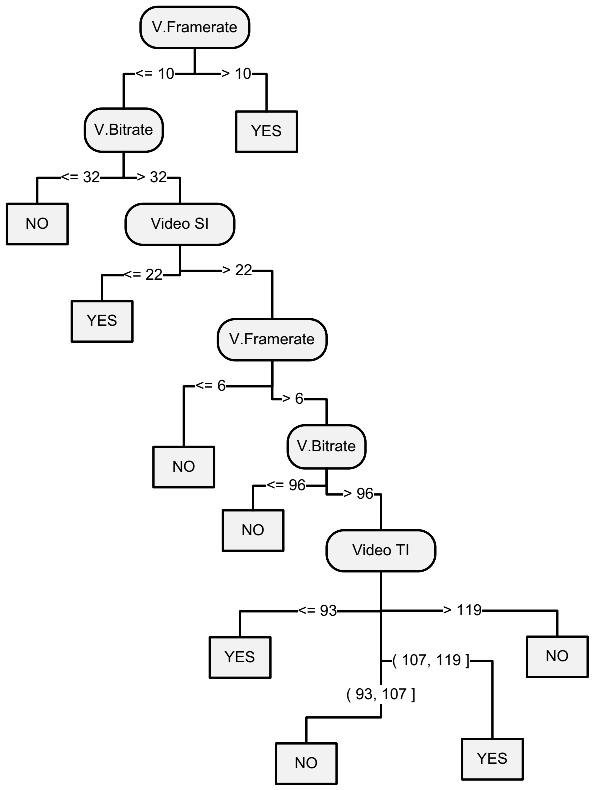

Figure 5b. Decision Tree for the PDA dataset

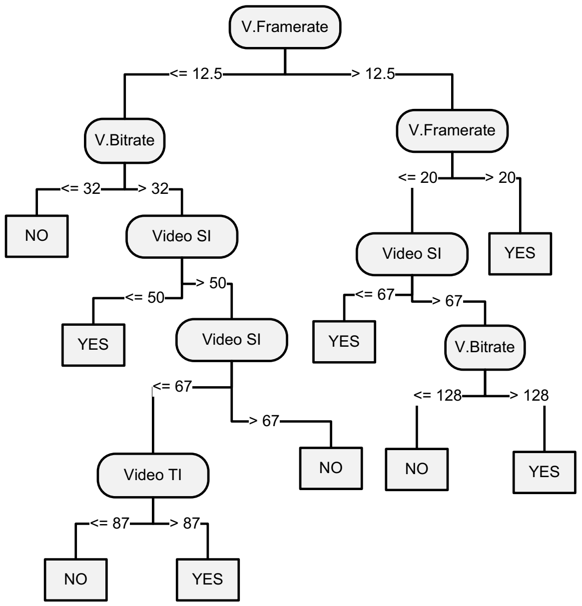

Figure 5c. Decision Tree for the Laptop dataset

4.4 Online Learning and Dataset Reduction

In order to for this prediction platform to be able to adapt to changes in user preferences we need to incorporate the prediction models into an online learning system. These models are ideally compatible for this use particularly in systems based on ensemble algorithms such as Hedge β [15]. In this algorithm a number of classifiers are trained on slightly different data and then used for classification based on weighted voting. The weights associated with each classifier are initially set during the training phase according to the accuracy of the classifier during training. After that during online operation the weights are modified based on feedback data from the users. In this way the online classifier ensemble can boost the classifier performance become resilient to noise in the data [15] and adapt to changes in the environment.

One of the first obstacles in this approach is the complexity and cost associated with QoE subjective feedback. As stated earlier subjective data is difficult to obtain so in order to address this issue we developed an approach for reducing the necessary data in a way that minimizes the impact on the classifier accuracy.

To implement a process of guided reduction that will later provide the ability for intelligent feedback selection we propose the Boundary Proximity reduction method. This method proposes that the data which is closest to the boundary of QoE acceptable and unacceptable brings most of the information about the boundary so we focus on this information.

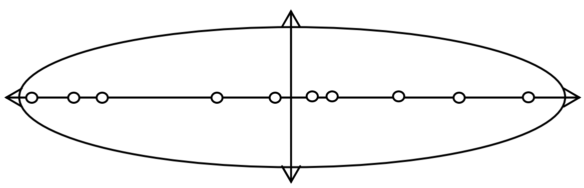

Figure 6a. Boundary Proximity Dataset Reduction method

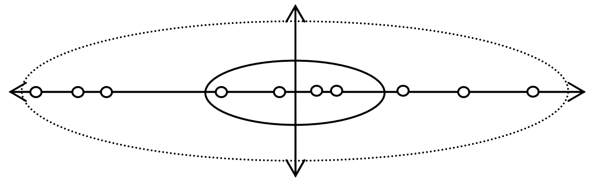

Figure 6b. Boundary Proximity Dataset Reduction method (reduced)

The first step is to build an SVM hyperplane from all of the training data. Then we use this hyperplane to calculate the distance from each datapoint to the hyperplane. As shown of Figure 6 we project each datapoint on an axis that represents the distance from the hyperplane. The implementation shown on Figure 6 is only for two class valued datasets. The reduction is done by removing the data furthest from the boundary first.

To compare our guided data reduction we executed a random data reduction so we have a basis for comparison. Figure 7 shows how the decision tree accuracy depends on the size of the dataset when we execute the stratified random dataset reduction.

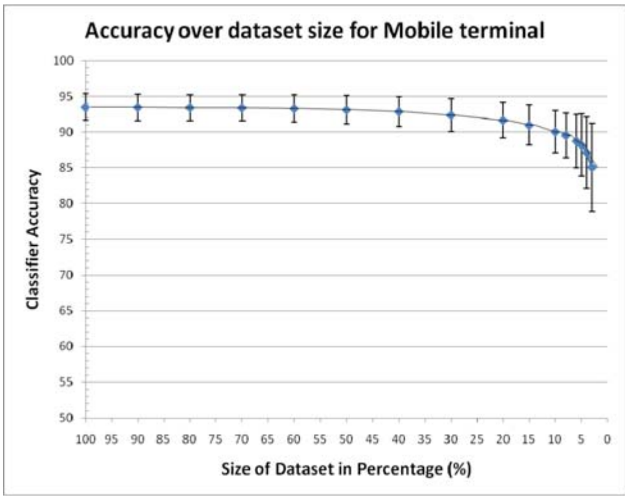

Figure 7. J48 Classifier accuracy with random data reduction

We can see that the data has a lot of redundany and decent predition models can be build even with only 20% of data. But we can also observe that the error bars towards the lower dataset size start to increase which leads to the conclustion that particular subsets bring significantly better accuracy than others. This is why we implemented a guided approach that looks for these subsets that carry the statisticaly significant data.

On Figure 8 we present an overlay of the random and guided approach. We can see that the guided approach suffers some losses initially due to the fact that some data needed for the more accurate decision tree was discared. But after discarding 50% of the data the accuracy stabilizes on 90%. The Boundary Proximity reduciton yields a stable accuracy of 90% even when we descard 98% of the dataset and are left with only 2% of the original datapoints.

This reduciton is also stratified, so we decrease the number of datapoints proportionaly for each class value, this apprach allows representative datapoints for each class value to be present until the very last datpoints are descarded. This is of course necessary for meaninful classifier training.

Now that we have shown that this method brings a stable and controlled dataset reduction we can use it to select the most valubale instances to request subjective feedback from the viewers for our QoE online prediction system.

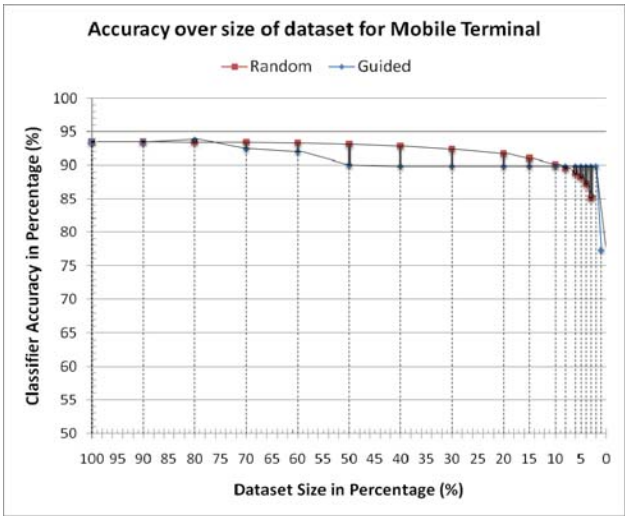

Figure 8. Overlay of random and guided data redution methods

5. CONCLUSIONS AND FUTURE WORK

The achieved accuracy for each model is tested by a k-fold cross validation. This technique splits the data set in k subsets. Then it uses k-1 of the subsets to train the model and the last one to test the model on. This is repeated k times for all combinations of subsets and the results of the testing are averaged. This is a more general technique than the leave-out-one which was used to validate the discriminate analysis models [16]. In leave-out-one only one datapoint is left to be tested on, and the procedure is repeated as much as there are datapoints. If k is equal to the number of datapoints then the k-fold cross validation is equal to the leave-out-one. In any case since both validations are of the same nature we will compare the results directly.

The accuracy of the models is:

SMO

- 88.59±2.85% (Figure 3a)
- 89.38±2.77% (Figure 3b)
- 91.45±2.66% (Figure 3c)

J48

- 93.55±1.76% (Figure 5a)
- 90.29±2.61% (Figure 5b)
- 95.46±2.09% (Figure 5c)

Compared to the discriminate analysis models:

- 76.9% (mobile phone)
- 86.6% (PDA)
- 83.9% (laptop)

We can conclude that we have achieved significantly higher accuracy of the prediction models (Figure 9). We can also state that approaches like decision tree or rule based approaches are more suitable for these datasets comparing the results from the SVM and the C4.5 models.

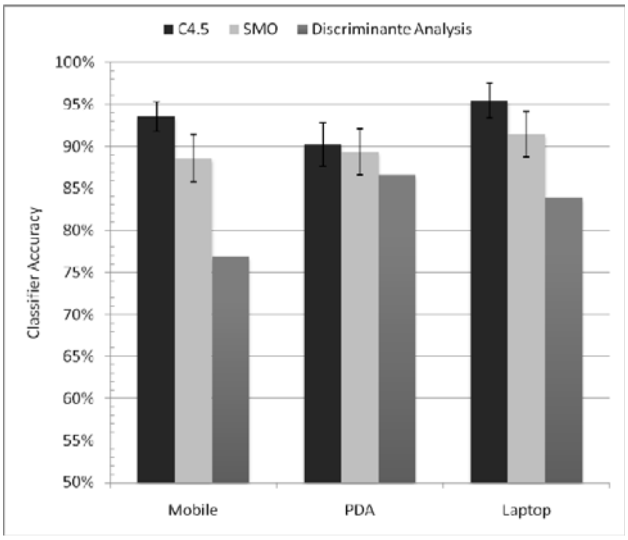

Figure 9. Comparison of models over the different datasets

We also compared the models over the distinct subsets of the datasets over the terminal type. Here we can observe that the nature and the data distribution is not same over the subsets. In the case of the mobile phone for example it shows that it is far less suitable for discriminant analysis than C4.5.

Having reached these results we can now conclude that our QoE prediction models are suitable for use in real world applications particularly as part of a control loop of a network management system. The high accuracy in determining the expected level of QoE enables us to make efficient decisions regarding the provisioning of network resources whilst keeping the customer satisfied. In addition decision trees and oblique decision trees offer very fast classification that in turn provides for applications of these models in real-time decision support systems.

The work presented herein was focused on the construction of accurate as well as practical QoE prediction models. Once we are able to directly correlate context (QoS) with user perception (QoE) it is possible to think about creating a real-time QoE monitoring system to be used for instance for user-centric SLA monitoring. Another possible development is in the area of QoE management and real-time QoE control. Hence, this work is also very relevant to those who are working on encoding and decoding techniques for real-time streaming.

6. ACKNOWLEDGMENTS

The work included in this article has been supported by Telefonica I+D (Spain). The authors thank María del Mar Cutanda, head of division at Telefonica I+D, for providing guidance and feedback. Subjective QoE data has kindly been provided by Dr. Florence Agboma, Essex University (U.K.).

7. REFERENCES

[1] F. Agboma and A. Liotta, "Addressing user expectations in mobile content delivery," *Mobile Information Systems*, vol. 3, Jan. 2007, pp. 153-164.

[2] F. Agboma and A. Liotta, "QoE-aware QoS management," *Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, Linz, Austria: ACM, 2008, pp. 111-116.

[3] A. Takahashi, D. Hands, and V. Barriac, "Standardization activities in the ITU for a QoE assessment of IPTV," *Communications Magazine, IEEE*, vol. 46, 2008, pp. 78-84.

[4] S. Winkler, "Video Quality and Beyond."

[5] S. Winkler, *Digital video quality : vision models and metrics*, Chichester West Sussex ;;Hoboken NJ: J. Wiley & Sons, 2005.

[6] M. Siller and J. Woods, "QoS arbitration for improving the QoE in multimedia transmission," *Visual Information Engineering, 2003. VIE 2003. International Conference on*, 2003, pp. 238-241.

[7] G.T. Fechner, E.G. Boring, H.E. Adler, and D.H. Howes, *Elements of psychophysics / Translated by Helmut E. Adler ; Edited by David H. Howes [and] Edwin G. Boring ; with an introd. by Edwin G. Boring*, New York :: Holt, Rinehart and Winston, 1966.

[8] W.R. Klecka, *Discriminant analysis*, SAGE, 1980.

[9] F. Agboma, "Quality of Experience Management in Mobile Content Delivery Systems," Department of Computing and Electronic Systems, University of Essex, 2009.

[10] V. Vapnik, *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (Springer Series in Statistics)*, Springer-Verlag New York, Inc., 1982.

[11] J.C. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," 1998.

[12] J.R. Quinlan, *C4.5*, Morgan Kaufmann, 2003.

[13] J.R. Quinlan, "Learning efficient classification procedures and their application to chess end games," *Machine learning: An artificial intelligence approach*, vol. 1, 1983, pp. 463-482.

[14] "Weka 3 - Data Mining with Open Source Machine Learning Software in Java."

[15] L. Kuncheva, *Combining pattern classifiers : methods and algorithms*, Hoboken NJ: J. Wiley, 2004.

[16] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *IJCAI*, 1995, pp. 1137–1145.