Multimedia QoE Optimized Management Using Prediction and Statistical Learning

Muslim Elkotob, Daniel Grandlund, Karl Andersson, and Christer Ahlund
Luleå University of Technology, Division of Mobile Networking and Computing, SE-931 87 Skellefteå, Sweden
{muslim.elkotob,daniel.granlund,karl.andersson,christer.ahlund}@ltu.se

Abstract—We present a scheme for flow management with heterogeneous access technologies available indoors and in a campus network such as GPRS, 3G and Wi-Fi. Statistical learning is used as a key for optimizing a target variable namely video quality of experience (QoE). First we analyze the data using passive measurements to determine relationships between parameters and their impact on the main performance indicator, video Quality of Experience (QoE). The derived weights are used for performing prediction in every discrete time interval of our designed autonomic control loop to know approximately the QoE in the next time interval and perform a switch to another access technology if it yields a better QoE level. This user-perspective performance optimization is in line with operator and service provider goals. QoE performance models for slow vehicular and pedestrian speeds for Wi-Fi and 3G are derived and compared.

Keywords- Statistical learning, QoS, QoE, multimedia traffic

I. INTRODUCTION

Autonomic control of network resources and the autonomous management of their processes is a recurring hot topic. Autonomic communication is mostly about self-awareness and self configuration of networks. As mobile nodes roam the network, resource settings change, and there are various adaptation mechanisms available to tune resources and adjust them in the best suitable way to maximize the overall goal. Example goals are: fairness, overall throughput, and profit. Proactive behavior involves some degree of foreseeing or prediction, which is also the case in our paper. Predictive QoE profiling refers to observing network parameters via active monitoring, followed by computing, using a regression best fit, the QoE for the next discrete time interval. The final step of the adaptive cycle is to select the network that achieves the best QoS level that in turn suits the so far built QoE profile including the value of the interval to come (predicted QoE for time slot n + 1 when the current time slot is n). We use statistical learning techniques and perform prediction of the target network parameter, namely video QoE for a subsequent discrete time interval.

In previous work [1], [2] we explored the issue of autonomic resource management for media streaming via a technique called parameter injection. This was limited to the open access network (OAN) [3] paradigm and operated on media streams where trans-coding, codec selection, format change and stream width were adjusted based on a predictive technique and self-configuring behavior on the terminal that performs joint signaling with the network where media streams originate. In current work, we focus more on the autonomic control loop and especially on efficient design, consistency, and accuracy for achieving best possible behavior, i.e. optimizing resource usage with respect to operator goals. We used a Wi-Fi coverage range with GPRS and 3G as alternative access technologies. The results produced in the paper provide a solution model which distinctively highlights the significant performance parameters and sorts out the insignificant ones that do not contribute to QoE depending on the scenario. The three scenarios tested and analyzed using our solution are 3G, vehicular-speed Wi-Fi, and pedestrian-speed Wi-Fi.

II. SYSTEM HIGH-LEVEL DESIGN

When conducting our experiments, we used a topology with overlapping coverage of GPRS, 3G, and Wi-Fi and in some cases even Ethernet (sockets) for testing a wired-wireless scenario. Having one user experiencing stable conditions (corresponding user) whereas the other user being mobile within a Wi-Fi network that has overlapping coverage with 3G and GPRS allowed us to test in real-time and in real indoor settings the behavior of a video (multimedia) application. We locally interfaced the video conferencing application and extracted the basic performance metrics such as round trip time, jitter, used data rate, frame rate, and packet loss rate, and bandwidth.

Our mobile node captures properties of audio streams, video streams, and the connection as an aggregate (Fig. 1) in real-time. The control loop governing the node behavior operates using discrete time intervals with a tunable timer. Every interval involves two types of action: measuring basic metrics for the current interval and predicting the QoE level for the next interval.

Figure 1. Mobile Node Control Loop and Metrics
When a mobile node passes via a network topology for the first time, it measures basic metrics (Fig. 1) for voice, video, and the aggregate flow and performs regression analysis. Determining the weights and significance impact factor of each basic variable on the overall response (being video QoE in our case) and then using those weights (coefficients) to scale the variables currently measured allows predicting the response value for the next interval. Fig. 2 presents the used architecture.

III. STATISTICAL LEARNING

By statistical learning we mean collecting network parameters and performance data, carry out statistical analysis, and to take decisions (namely handover and network selection decisions) that improve performance and optimize a certain goal. In this paper, learning is based on linear regression to derive weights for different parameters. Those weights are then used to predict the upcoming value of the target variable, namely video quality of experience (QoE). Such a setup is useful because operators try to optimize their resource usage in such a way to maximize SLA (service level agreement) fulfillment with the minimum amount of resources possible. Furthermore, operators and service providers try to keep as graceful profile as possible for any service in use. Therefore, when using video, if there is any chance to switch to a network that would help avoiding a sharp fluctuation in the QoE profile, this should be done. This can be solved using statistical analysis and learning with the help of a control loop. Algorithm 1 presents a step-wise overview of the statistical learning mechanism we used.

**Phase 1: learning, acquiring knowledge**

**Phase 2: applying knowledge (prediction + optimization)**

**Algorithm 1. Network Selection for Optimal Quality of Experience Using Statistical Learning Pseudo Code**

- **Phase 1:** learning, acquiring knowledge
  - Measure(InputValues, OutputMOS)
  - Regression(Inputs, Outputs) \( \Rightarrow \) Derive parameter weights
  - Obtain \( QoE \) Model(time range, location range)
- **Phase 2:** applying knowledge (prediction + optimization)
  - While(1):
    - For Timeslot(t):
      - PredictQoE(t+1, InputWeights(t))
      - SelectedAccessNetwork = best(PredictedQoE)
      - For Timeslot(t+1):
        - ComputePredictionError()
        - Adjust/tune model to reduce error

TABLE I. PERFORMANCE METRICS MEASURABLE IN OUR SCENARIOS

<table>
<thead>
<tr>
<th>Audio Channel</th>
<th>Video Channel</th>
<th>Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter</td>
<td>Frame rate</td>
<td>Data volume in</td>
</tr>
<tr>
<td>RTT</td>
<td>Frame resolution</td>
<td>Data volume out</td>
</tr>
<tr>
<td>Audio Codec</td>
<td>Video Codec</td>
<td></td>
</tr>
<tr>
<td>PLR</td>
<td>Data Rate</td>
<td></td>
</tr>
<tr>
<td>QoE Level</td>
<td>QoE Level</td>
<td></td>
</tr>
</tbody>
</table>

where \( A = 4.0317 \), \( B = -44.9873 \), and \( C = -0.5752 \).

EQ set 1. Mean Opinion Score QoE Output Model for Video Conferencing under Slow Vehicular and Pedestrian Speed

Since Wi-Fi is the access technology with most vastly varying QoE profile, it sometimes falls below the 3G or even the GPRS level and sometimes soars above. Therefore, prediction based on regression is a good way to determine what network to select so that it delivers the best QoE. For instance, if currently the QoE level delivered by Wi-Fi is 3.3, it is a good choice to stay there; then if the predicted QoE level for the next interval is e.g. 2.98, which is lower than what 3G can deliver, the mobile node will switch to 3G. This way, the highest aggregate QoE level of a service can be achieved.

B. Regression, Prediction, and Learning

Regression is the process of determining the relationships between different variables. Our input variables are: packet loss rate, frame rate, round trip time, bandwidth, and jitter. The output (response) variable is the perceived quality of service (QoE) represented by the universal mean opinion score (MOS). In the first run of measurements, data collection for all five...
input metrics as well as the response variable is done. Then statistical regression analysis is done to determine the weights (impacts) of different parameters on the response variable.

The regression equations obtained via Minitab software [6] for the three different scenarios we tested under the same conditions, “Coefficient” stands for the weight of a parameter for the QoE model in each scenario, “Std Err Coef” is the coefficient standard error, and “P-Value” is the strongest well known test statistic used to determine whether a statement about a parameter’s weight is accurate.

### TABLE II. 3G REGRESSION RESULTS

<table>
<thead>
<tr>
<th>Predictor/ Metric</th>
<th>Coefficient</th>
<th>Std Err Coef</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>3.69947</td>
<td>0.00415</td>
<td>0.000</td>
</tr>
<tr>
<td>Frame rate</td>
<td>0.0029732</td>
<td>0.0004128</td>
<td>0.000</td>
</tr>
<tr>
<td>PLR</td>
<td>-0.012898</td>
<td>0.008609</td>
<td>0.135</td>
</tr>
<tr>
<td>RTT</td>
<td>0.00000566</td>
<td>0.00000171</td>
<td>0.001</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.00031756</td>
<td>0.00000706</td>
<td>0.000</td>
</tr>
<tr>
<td>Jitter</td>
<td>-0.00006948</td>
<td>0.00000720</td>
<td>0.000</td>
</tr>
</tbody>
</table>

In the first cycle round (control loop in Fig. 1), target variable QoE values and input network performance parameters such as frame rate (FR), packet loss rate (PLR), round trip time (RTT), bandwidth (BW), and jitter are measured and fed into the regression mechanism on the mobile device. Jitter here refers to the one after the play out buffer which means that it reduces delay effects for video sessions. Based on the results obtained in for the scenarios low-speed vehicular Wi-Fi, pedestrian Wi-Fi, and 3G we determine the coefficients (weights) and summarize the results in Table III.

### TABLE III. PREDICTION MODEL SUMMARY

<table>
<thead>
<tr>
<th>Predictor/ Metric</th>
<th>BW</th>
<th>Jitter</th>
<th>RTT</th>
<th>FR</th>
<th>PLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicular Wi-Fi</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Pedestrian Wi-Fi</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>3G</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

The key message is to model QoE based on the settings (moving speed and available access technology) yielding:

- **For vehicular Wi-Fi**: QoE for video conferencing is best modeled as a function of bandwidth and jitter;
- **For pedestrian Wi-Fi**: QoE for video conferencing is best represented as a linear combination of a constant plus a weighted sum of bandwidth, jitter, and RTT;
- **For 3G**: QoE is represented as a linear combination of bandwidth, RTT, jitter, and video frame rate.

Optimizing and fine-tuning the coefficients with a learning loop is part of future work. For the current version the following stepwise approach was used for linking the system model to the matrix algebra of linear regression:

- Create a sufficiently large data set of derived metrics; in our case real measurements. The basic input metrics and response variable are captured. This data we assign to the m by n matrix X. m rows are for the available samples. n columns correspond to the number of metrics incremented by one;

  - We compute and verify the output values for QoE via Eq Set 1. Using customized sent RTP audio and video traffic (video conferencing application and scripts that log data from it in discrete time intervals), we measure perceived QoE on a scale of 1-5. The QoE output vector is denoted by Y;

  - Perform linear regression using: \( Y = X.\lambda + \varepsilon \); Y is our output QoE matrix in the test runs. X is the set of captured metrics (RTT, jitter, PLR, FR, and bandwidth). The variable \( \varepsilon \) corresponds to the correction factor in the formula; it has zero-mean and a normal distribution. The column vector corresponds to the coefficients (or weights);

    - To find the vector of coefficients \( \lambda \), having Y and X (measured and evaluated in the matrices), we do the following:
      
      \[
      E[Y] = E[XX^T + \varepsilon] \Rightarrow \hat{Y} = \hat{X}\lambda; \quad \text{since } \varepsilon \text{ is zero-mean}
      \]

      \[
      \Rightarrow \hat{X}^T\hat{Y} = \hat{X}^T\hat{X}\lambda.
      \]

      \[
      \Rightarrow (\hat{X}^T\hat{X})^{-1}\hat{X}^T\hat{Y} = (\hat{X}^T\hat{X})^{-1}(\hat{X}^T\hat{X})\lambda.
      \]

      \[
      \Rightarrow \lambda = (\hat{X}^T\hat{X})^{-1}\hat{X}^T\hat{Y};
      \]

    - Having all values (\( \hat{X} \) and \( \hat{Y} \)), we compute the column vector matrix \( \lambda \) entry values. For the two test runs we made, with a mobile laptop inside a slow-moving car and also carried by a pedestrian with 3G and Wi-Fi coverage. This is an extension and elaboration of the passive measurements-based elastic traffic performance prediction approach defined in [5]. We have a more complex behavior pattern and more real-time metrics rather than passive measurements.

The benefit of such a method is its simplicity, and the learning process is coupled with geographic and time awareness. In other words, traversing a particular location and measuring input and output variables allows determining the coefficients (weights) and impacts of each parameter on the target performance variable (video QoE). Any subsequent presence in the same geographical area or time interval within some vicinity would enable the mobile node to use the computed regression coefficients to predict the upcoming QoE and thus decide whether or not to switch to another access technology from Wi-Fi. The goal is optimizing the QoE yield as service providers try to achieve for subscribers. In the current implementation, geographic awareness is manually handled (i.e. signaling to the system the region index so that it uses coefficient weights from statistical learning for prediction in that region). In future work, we will systematically build geographic location awareness into the system so that the information obtained from learning at a particular location will be applied for QoE prediction during the 2nd pass for the same location.

### IV. PERFORMANCE ANALYSIS

As Fig. 3 shows, we plot the predicted QoE values our scheme produces versus the actual values measured in the subsequent discrete time interval (i.e. in interval \( I_{m} \) predict \( QoE_{m+1} \) and measure \( QoE_{m} \) then move on to \( I_{m+1} \) and so forth). Within the good performance range of Wi-Fi, corresponding to
a QoE index between 3.0 and 4.0, prediction works well, and delivers a predicted value slightly lower than the actual value.

Fig. 4 shows that on the mean opinion score (MOS) scale, the perceived video quality achievable with 3G reaches an asymptotic value between 3.5 and 4.0 even when bandwidth increases linearly. This is because of the other performance metrics within 3G pose a limitation on how high the achieved QoE can be. The point here is not to invest more network resources than needed if that does not improve QoE further.

V. RELATED WORK

A significant amount of work has been done in the areas of QoS-QoE analysis, prediction, stochastic analysis combined with prediction and optimization of resource control in wireless networks. We outline some of this work in this section and point out its relevance or delta from our own approach.

Kitawaki in [7] designs a scheme for computing perceived quality of service (called QoE in our paper) for multimedia (VoIP). Opinion models are used to process measured input values and obtain a QoE value. Prediction is an important aspect to foresee the degree of QoE to be expected. The G.107 E-model which has an additive property is used to build the prediction function. In our paper, we predict QoE using statistical learning on the already existing data sets acquired as the mobile node roams within the Wi-Fi-3G network.

Ito et al. [8, 9] study the relationship between network, node, application, and user-level QoS (which we refer to QoE in this paper) that the authors consider perceptual. Finding out the relationship between the calibrations done on QoS levels and the received improvement or change in QoE is a problem that can be optimized and is in the interest of operators. The authors provide a correlation matrix whose coefficients reflect the impact of each QoS variable in the QoE making it quantifiable and transparent for designers to see and use the relationship between QoS and QoE. We focus more on learning and applying regression to the data set in for prediction in the next interval based on the learned information.

VI. CONCLUSION

We presented a scheme for prediction and measurement combined with statistical analysis and learning in order to allow a mobile node to be proactively aware of the best access network for the next interval. The practical use of such a scheme would be interesting for operators and service providers who need to maintain graceful QoE profiles and optimize their resource usage. There are several challenges in the current existing state of the art, and as a response to them, a new approach was needed. The key challenges faced today with QoE management using QoS metrics and statistical learning are: high complexity, low accuracy (wrong decisions), and oscillations. Advantages of the scheme proposed for QoE management using prediction and statistical learning include: high accuracy, parallel threading within the control loop, simplicity, and low computational load.

REFERENCES