A Fixed-Point Model for QoE-based Charging

Peter Reichl1,2
1University of Vienna
Dept. of Computer Science
Währinger Strasse 29
A-1090 Vienna, Austria
+43 1 4277 79410
{peter.reichl | patrick.maille}@telecom-bretagne.eu

Patrick Maillé2
2Université européenne de Bretagne/
Télécom Bretagne
2, rue de la Châtaigneraie
F-35576 Cesson-Sévigné, France
+33 299 127028

Patrick Zwickl3, Andreas Sackl3
3FTW Telecommunications
Research Center Vienna
Donaucitystrasse 1
A-1220 Vienna, Austria
+43 1 5052830-0
{zwickl | sackl}@ftw.at

ABSTRACT
Within the current paradigm change from Quality-of-Service (QoS) towards Quality-of-Experience (QoE), the question of how to charge for QoE is widely neglected in the research community despite of its obvious importance. In this paper, we present and analyze a fixed point model which specifically reflects the double role of prices, i.e. as regulating factor for demand size and at the same time as part of the QoE-based user context. The model is validated through comprehensive user trial results which allow interesting insights into the temporal behavior of end users who are confronted with a fine-grained scale of choices on video quality and corresponding tariffs.

Categories and Subject Descriptors
J.4 [Social and Behavioral Sciences] – Economics
K.6.0 [Management of Computing and Information Systems]: General – Economics

General Terms
Economics, Experimentation, Human Factors

Keywords
Quality of Service, Quality of Experience, Network Economics, Internet Charging and Pricing

1. INTRODUCTION AND RELATED WORK
While for almost two decades, service quality in communication networks has usually been described in terms of so-called QoS (Quality of Service) parameters like packet loss rate, delay, jitter, bandwidth etc., more recently we observe a strong trend in academic and industrial research towards introducing more user-centric quality concepts, which is considered to be a significant paradigm change [17]. In this context, the notion of QoE (Quality of Experience) [6][12] has been gaining rapidly increasing importance, as – according to the ITU-T definition – it captures the “overall acceptability of an application or service, as perceived subjectively by the end-user” [1]. As a consequence, a comprehensive body of related work has recently been devoted to developing QoE models [2][6][7][12] as well as corresponding metrics and measurement methodologies [4][8][15].

At the same time, it is interesting to note that the question of how to charge end users for their perceived quality so far has been largely neglected by the research community, despite of its obvious importance which is fully in line with R. Jain’s notorious statement that the missing integration of economical aspects into QoS research constitutes a major reason for the pertinent difficulty of introducing QoS-enabled architectures in the current Internet [5]. However, there are still not very many ongoing research projects, like for instance the EU FP7 IP ETICS (Economics and Technologies for Inter-Carrier Services) [9], which aim at developing a “techno-economic” approach to developing future Internet architectures and thus integrate both economic and technological aspects with equal importance.

For more specific related work on QoE-based charging, we would like to refer to [11][14][16], where we discuss general frameworks as well as corresponding mechanisms both from an economic and a user perspective. In contrast, the present paper follows the more fundamental way the problem has been addressed in [13] and complements our preliminary thoughts presented there with two main contributions, i.e. an analytical treatment of the fixed point model for QoE-based charging as well as its comprehensive empirical validation based on results from a large user trial which has recently been performed as part of the corresponding activities in the ETICS project (cf. also [16]).

The remainder of this paper is structured as follows: Section 2 introduces and discusses the so-called Fixed-Point Problem of QoE-based Charging. Section 3 describes the setup of our user trials and presents their results which provide empirical evidence for the analytical conclusions reached before. Section 4 concludes the paper with a brief summary and an outlook on current and future work.

1 As an example for an alternative (and much broader) QoE definition, we would like to refer especially to the one worked out and proposed by the European COST Action IC1003 QualiNet: “Quality of Experience (QoE) is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user’s personality and current state.” [8]
2. FIXED-POINT MODELS FOR QoS- AND QoE-BASED CHARGING

In order to illustrate the basic conceptual difference between traditional QoS-based charging and charging based on QoE, in [13] we have considered a simple dynamic system where a provider with limited resources offers a service to an end customer. Service demand is driven by the price (e.g. based on the concept of “price elasticity”, see [10]); on the other hand, changing demand will directly impact service quality due to resource limitations, which itself leads to price adaptation (for instance, increasing demand results in lower quality and hence in lower prices, and v.v.).

More formally, let \( p \) indicate the price, \( d \) the demand and \( q \) the QoS. Then, the cyclic dynamic system of QoS-based pricing is described by the following set of equations

\[
\begin{align*}
\text{price function} & \quad p = p(q) \quad (1) \\
\text{demand function} & \quad d = d(p) \quad (2) \\
\text{QoS function} & \quad q = q(d) \quad (3)
\end{align*}
\]

Where w.l.o.g. we assume all functions to be continuous bijective mappings of the unit interval \([0;1]\) to itself.

**Proposition 1:**
Assume \( d(p) \) to be monotonically decreasing and convex, \( q(d) \) to be monotonically increasing and concave, \( p(q) \) to be monotonically decreasing and concave, then, if all three functions are linear, the entire unit interval constitutes a set of fixed points for the system of equations (1)–(3). Otherwise, if at least one of the concavity and/or convexity conditions is strict, there are only two (trivial) fixed points: an unstable one at \((p,d,q) = (0,1,0)\) and a stable one at \((p,d,q) = (1,0,1)\).

**Proof:** see Appendix.

In contrast to the simple model considered so far, the situation is more complex if the charging mechanism is based on the QoE perceived subjectively by the end customer, mainly because prices assume a double role there: on the one hand side, the end customer still pays for the level of offered QoS (hence prices usually rise with increasing service quality), on the other hand the price to be paid forms also part of the user context [7] and thus has direct impact on the quality perception itself: the higher a price, the higher also the user’s expectations concerning the offered service quality (and the larger the probability to be disappointed).

\[
\begin{align*}
\text{QoS function} & \quad q = q(p) \\
\text{Demand function} & \quad d = d(p) \\
\text{Price function} & \quad p = p(x)
\end{align*}
\]

Figure 1. Charging for QoE

![Diagram of charging for QoE](image)

Figure 1 depicts the resulting feedback model for QoE-based charging. Observe that there is a new function involved which we term “QoE function” and which depends on both the QoS level \( q \) and the price level \( p \). Consequently, we may formulate a new system of equations where equations (2) and (3) remain identical, equation (1) becomes

\[
\begin{align*}
\text{price function} & \quad p = p(x) 
\end{align*}
\]

and we have an additional equation for the QoE function

\[
\begin{align*}
\text{QoE function} & \quad x = x(q,p)
\end{align*}
\]

In general, it is safe to assume that the QoE function in (5) should be monotonically increasing in \( q \) (better QoS leads to better QoE) and monotonically decreasing in \( p \) (higher tariffs lead to higher quality expectations which lead to lower subjective quality perception).

Following [13], we may further simplify the model by assuming \( x \) to be separable, i.e.

\[
\begin{align*}
x = x(q,p) = x_Q(q) \cdot x_E(p)
\end{align*}
\]

which we interpret as follows: the QoE depends on both the QoS offered by the provider (in terms of a monotonically increasing quality function \( x_Q \)) and the customer expectations which are triggered by the corresponding tariff (= price), which is expressed in terms of a (monotonically decreasing) expectation function \( x_E \).

Generalizing [13], however, we assume for now that \( x_Q(0) = 1 \) and \( x_Q(1) = 0 \), while \( x_E \) is not subject to further restrictions (hence it could be linear, convex, concave, sigmoidal etc. – in fact, there is little empirical evidence so far about a reasonable shape of this function at all).

**Proposition 2:**
Assume \( d(p) \), \( q(d) \) and \( p(x) \) as in Proposition 1, as well as \( x_Q : [0;1] \rightarrow [0;1] \) to be continuous and monotonically increasing and concave, and \( x_E : [0;1] \rightarrow [0;1] \) to be continuous and monotonically decreasing, moreover assume all functions to be continuously differentiable. Then, if all functions are linear, the system of equations (2)–(6) has exactly one fixed point at \((p,d,q) = (0,1,0)\) which is stable. Otherwise, if at least one of the concavity and/or convexity conditions is strict, this fixed point becomes unstable, and there is an additional stable fixed point which is non-trivial, i.e. \( p \in [0;1] \). If, in addition, \( x_E \) is concave, this non-zero fixed point is unique.

**Proof:** see Appendix.

Note that, while both propositions are original contributions of this paper, we would like to refer to [13] for an illustrating example of the corresponding convergence behavior as well as further discussions on the various functions involved.

3. EMPIRICAL VALIDATION

As a follow-up to the first ETICS [9] user trial which has taken place in autumn 2011 and whose results are described in [13][16], a second user trial series has been conducted at FTW in the autumn of 2012, in order to understand in detail the user behavior if confronted with a rather fine granular selection of quality (and tariff) choices. In this section, we describe the setup of this second user study as well as a comprehensive survey of results which serve as empirical validation of the fixed point model described in Section 2.
3.1 User Trial Setup

The technical setup of our user trial is a modified version of the one described in [16] and is depicted in Figure 2. While [16] has distinguished between four classes of Standard Definition (SD) video quality based on different packet loss levels for UDP-based transmission, our revised setup aims at being significantly closer to reality. We use TCP-based adaptive video streams in order to adapt video quality to network conditions, i.e. mainly bandwidth. The switch from SD to High Definition (HD) Blue-ray quality allows to differentiate 17 quality levels, based on logarithmically scaled bitrates, plus another three “virtual” quality classes which identically employ the best possible bitrate but still differ in terms of prices. We have included the latter classes for cross validation purposes, thus the trial subjects have been exposed to a total of 20 purchasable quality classes. Like with [3] and [16], trial users have been given a certain amount of real money – 10 EUR in our case – which they could freely spend on quality enhancements during the trial or take home afterwards. Together with the mentioned distinction into a multitude of fine granular quality classes, this setup has been chosen to allow for a clear and well observable interaction behavior on the user side.

![Figure 2. Technical Setup of the User Study](image)

Moreover, while the direct investigation of test subjects’ maximum willingness to pay has been out of scope for [16], our present study has defined three tariffs (A, B, C) to which test subjects are automatically assigned: a low tariff A ranging from 0 to 2 EUR, a middle tariff B ranging from 0 to 3 EUR, and a high tariff C ranging from 0 to 4 EUR. The choice of the (relatively high) maximum prices, i.e. 2, 3 and 4 EUR, resp., has been encouraged by observing that effective expenditures in [16] have clearly exceeded prior expectations.

Each end user watches three video sequences with a duration of 20 minutes each, which are individually chosen from a representative video library with carefully selected sequences (including e.g. highly topical blockbusters). Starting per default with the poorest quality level Q0, the end user can use a jog wheel for dynamically and interactively testing quality advancements (no charging at this stage). The user is always informed about the price tag of each selection, while the number of available quality levels is intentionally hidden for reducing unnecessary biases. Due to a highly improved setup (an own VLC client fork, precise logging mechanisms), quality changes are quickly applied (with a delay of about 1 sec only) and fine-granularly tracked in order to provide valuable clues about quality and price decisions and considerations of end users. After the free trial phase has ended after around five minutes, the latest selected quality level is assumed to be the final choice, and the corresponding price is deducted from the user’s deposit. After the completion of the experiments, the remaining deposit is paid out in cash to the subjects, as announced before the conduction of the trial.

Note that using a jog wheel as physical user interface for changing between quality classes provides significant advances, as it employs an intuitive mechanism well-known to basically all users e.g. from sound volume control, and moreover it allows to suggest the existence of an infinite number of quality levels (of course, in reality requests for lowering the quality below the lowest or above the highest level have simply be ignored). In this way, it was possible to record (with temporal granularity of 1 sec) user behavior both in terms of selecting quality classes as well as in terms of their convergence behavior, as indicated for instance by the number of “trend” changes between increasing and decreasing the quality (driving the jog wheel up and down, resp.).

Figure 3 illustrates the corresponding cumulative distribution functions for the 40 test subjects. Observe that the majority of users have adapted the video quality typically between 10 and 50 times (some individual cases going up to around 85 adaptations), in 80% of the cases the final choice has been reached while changing the direction of adaptation less than ten times which suggests a reasonable convergence behavior. Note that the present paper restricts itself to the perspective of the individual user, while additional trial results on an aggregated level will be reported in related companion papers.

![Figure 3. Distribution of price/quality changes (left) and price/quality trend changes (right) per movie](image)

3.2 User Classification

As already mentioned above, a total of 40 users has participated and completed the trial, with three movies to be analyzed for each user. Note that this number of participants is well in line with typical standard sizes for user trials (e.g. VQEG methodologies: 24 valid correlating subjects). One way to achieve a feasible classification of the results might aim at clustering the 120 runs of the trial, which however will be of limited interest. Instead, we make the key basic assumption that each individual user behaves consistently during the entire trial, i.e. follows the same quality/price selection pattern for all three movies under inspection. While this assumption cannot be claimed to be universally valid (as, for instance, certain users clearly exhibit some learning behavior, i.e. explore the entire space of quality levels during the first movie while coming to a much quicker decision for the subsequent movies), we consider it nevertheless a useful working hypothesis, while future work will address and validate this issue more in depth with the help of an appropriate consistency metric for individual user behavior. Finally note that, at least in our trial, the overall fraction of users with clearly irregular has been staying below 15% of the trial population.
Based on this assumption, we propose to classify user behavior according to two orthogonal parameters: (i) amplitude of quality/price selection, and (ii) speed of convergence. As a result, we end up with four different user classes labeled as “F” (fast convergence), “R” (regular convergence), “S” (slow convergence) and “X” (irregular convergence behavior). While Fig. 4 illustrates the three “regular” classes, the “irregular” class “X” consists of users which cannot be placed in one of the regular classes.

**Figure 4. Dimensions of User Classification**

**Type “F”: Fast Convergence – Small Amplitude**

The first general type of users is characterized by a relatively consistent convergence behavior, where the user climbs up the quality ladder until she has reached her targeted quality, and then stays there for the rest of the time without major changes (see Fig. 5). Sometimes, as a kind of learning effect, the behavior during the first movie is different and more exploratory (Fig. 5 right).

**Figure 5. Typical Examples for User Type “F”**

**Type “S”: Slow Convergence – Large Amplitude**

Contrasting the examples depicted in Fig. 3, a second user type may be characterized exactly by the opposite behavior, i.e. large amplitude and slow speed of attenuation. Two typical examples are provided in Fig. 6 (note that the change frequency is quite different between both users, while the overall temporal speed of convergence is roughly comparable).

**Figure 6. Typical Examples for User Type “S”**

**Type “R”: Regular Convergence – Medium Amplitude**

The third user type can be described as somewhat in between types F and R. It is characterized by large to medium amplitudes as well as medium speed of convergence, see Fig. 7 for two typical examples.

**Figure 7. Typical Examples for User Type “R”**

**Type “X”: Free Riding and Inconsistent User Behavior**

Finally, also some sporadic cases of free riding (Fig. 8 left) and further inconsistent user behavior (see e.g. Fig. 8 right) have been observed and sorted out, building a fourth class which is irregular. Note, however, that these cases account for only around 15% of the overall number of samples altogether.

**Figure 8. Typical Examples for User Type “X”**

### 3.3 Aggregated Convergence Behavior

In order to capture also the aggregated convergence behavior for user $i$ by an appropriate metric, we define the Root Square Deviation (RSD) $\tilde{\sigma}_i(t)$ as a function of time $t$ with respect to the convergence value $\tilde{x}_i = \lim_{t \to \infty} x_i(t) = x_i(300)$ as follows:

$$\tilde{\sigma}_i(t) = \sqrt{\left(x_i(t) - \tilde{x}_i\right)^2}$$

(7)

Moreover, define

$$\tilde{\sigma}_F(t) = 3 \cdot \exp(-0.01t)$$
$$\tilde{\sigma}_R(t) = 6 \cdot \exp(-0.007t)$$
$$\tilde{\sigma}_S(t) = 9 \cdot (1 - t/300)$$

as Reference Root Mean Square Deviations (RRMSD) of classes F, R and S, respectively. Then, for user $i$ and class $k$,

$$\Delta_i^{(k)} = \frac{1}{300} \sum_{t=1}^{300} \left(\tilde{\sigma}_i(t) - \tilde{\sigma}_i^{(k)}(t)\right)^2$$

(9)

represents the average square difference between user $i$’s RSD and the RRMSD of class $k$ over the selection period (with duration 300 sec).

Based on this metric, we specify a classification algorithm in Fig. 9. Basically, the classes are characterized by referring to equations (8) as well as to the fact that class F is more coherent (i.e. closer to the RRMSD) than class R which itself is more coherent than class S, i.e. in terms of expectation values per class:

$$E\left(\Delta_i^{(F)}\right) \leq E\left(\Delta_i^{(R)}\right) \leq E\left(\Delta_i^{(S)}\right).$$

(10)

For our purposes, we assume $\Delta_i^{(F)} \leq 1$, $\Delta_i^{(R)} \leq 2$ and $\Delta_i^{(S)} \leq 3$. 

The algorithm specified in Fig. 9 checks membership of user \(i\) with classes \(F\), \(R\), and \(S\) in hierarchical order, i.e. one after the other. If user \(i\) does not belong to any of the three “regular” classes, she is put into class \(X\). In this way, irregular behavior (which accounts for less than 15% of the cases) may easily be identified by an excessive size of the metric defined in (9).

Based on this algorithm, we were able to classify 39 out of the 40 participants of our trial correctly. The algorithm failed only in the case of one user who was converging to her final choice (maximal quality) within 32 seconds without later changes, resulting in a significant average deviation from \(\hat{c}_i(t)\). The distribution of the other 39 participants together with the corresponding mean and standard deviation values for \(\Delta_i^k\) is depicted in Table 1 (note that for class \(X\), we have averaged mean and standard deviation over the differences to all three classes).

### Table 1. Classification Results

<table>
<thead>
<tr>
<th>classes</th>
<th>(F)</th>
<th>(R)</th>
<th>(S)</th>
<th>(X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of users</td>
<td>8</td>
<td>11</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Mean ± standard deviation of (\Delta_i^k)</td>
<td>0.69 ± 0.16</td>
<td>1.58 ± 0.28</td>
<td>2.58 ± 0.36</td>
<td>8.30 ± 2.91</td>
</tr>
</tbody>
</table>

Fig. 10 depicts the three RRMSDs described in equation (8) (grey lines) together with the average value of \(\hat{c}_i(t)\) according to equation (7) per class. We conclude that the convergence metric defined above together with the classification algorithm allows a sufficient distinction between the “regular” classes \(F\), \(R\), \(S\). In total, around 85% of our user population exhibit convergence behavior towards a fixed point as suggested by our earlier analysis, while almost every second user approaches her final choice on a track which is best described as negative exponential (with different speeds of convergence between classes \(F\) and \(M\)).

### 4. SUMMARY AND OUTLOOK

In this paper, we present and analyze a fixed point model for charging users based on their Quality of Experience. The specific complexity of this model comes from the fact that with QoE-based charging, the role of prices becomes twofold: they serve as expression for the received service quality, while at the same time they constitute an essential part of the user context by forming her expectations, which has significant impact for QoE evaluation.

While in the first part of the paper we provide mathematical evidence for their existence and uniqueness, in the second half we illustrate the applicability of this model through a comprehensive user trial of QoE for video streaming. To this end, we present an advanced trial setup which allows users to choose in real time between 20 fine granular quality classes of an HD video, subject to different price plans. In order to increase closeness to reality, users are given real money which they can either spend during the trial to purchase better quality or take home with them. The physical interface consists of a jog wheel which allows intuitive and almost limitless changes between quality/price classes.

As a result, we observe that most trial participant exhibit the postulated convergence towards a fixed point choice also in practice and may be classified into three classes, according to the speed of convergence. We propose a specific convergence metric together with a simple classification algorithm based on specific properties of the three basic classes. Less than 15% of our users have shown irregular behavior which is much less than expected.

Current and future work emphasizes further psychological aspects of the problem, for instance concerning anchoring effects for market entrance pricing, and the bias of willingness-to-pay due to phenomena of cognitive dissonance, as well as more complex charging models and a model extension with respect to demand-supply optimization.

### APPENDIX

**Proof of Proposition 1**: The normalization together with the monotonicity conditions yield \(p(0) = 0, p(1) = 1, d(0) = 1, d(1) = 0, q(0) = 1\) and \(q(1) = 0\).

The existence of the two trivial fixed points can be checked in a straightforward manner. Assume now that all functions are linear. Then, for any \(p_0 \in [0;1]\), we have

\[
\begin{align*}
    d(p_0) &= 1 - p_0 \\
    q(d(p_0)) &= 1 - d(p_0) = p_0 \\
    p(q(d(p_0))) &= q(d(p_0)) = p_0
\end{align*}
\]

More generally, for non-linear functions this generalizes towards

\[
\begin{align*}
    d(p_0) &\leq 1 - p_0 \\
    q(d(p_0)) &\geq 1 - d(p_0) \geq p_0 \\
    p(q(d(p_0))) &\geq q(d(p_0)) \geq p_0
\end{align*}
\]

Hence, if any of the inequalities is strict in the open interval \([0;1[\), then \(p(q(d(p_0))) > p_0\) and hence no \(p_0 \in ]0;1[\) can be part of the coordinates of an interior fixed point.
Proof of Proposition 2:

In addition to Proposition 1, we assume also $x_Q(0) = x_E(1) = 0$ and $x_Q(1) = x_E(0) = 1$. If all functions are linear, we have

$$d(p_0) = 1 - p_0$$

$$q(d(p_0)) = 1 - d(p_0) = p_0$$

$$x(q(d(p_0)), p_0) = p_0 - p_0^2$$

$$p(x(q(d(p_0)), p_0)) = p_0 - p_0^2$$

From the last equation, we see that $(p, d, q) = (0, 1, 0)$ is the only fixed point, with $x_q(0) = 0$ and $x_E(0) = 1$, i.e. $x(0,0) = 0$. This fixed point is stable.

For non-linear functions, similarly to the proof of proposition 1, we can easily derive that

$$x_Q(q(d(p_0))) \geq p_0$$

hence $x_Q$ is a concave function of $p_0$ with $x_Q'(p_0) \geq 1$, and

$$x(q(d(p_0)), p_0) = x_Q(q(d(p_0))) \cdot x_E(p_0)$$

is a product of two continuously differentiable functions and hence continuously differentiable, with $x'(p_0) = x'(p_1) = 0$.

Hence, with the product rule we have

$$x'(p_0) = x_Q'(p_0) \cdot x_E'(p_0) + x_Q(p_0) \cdot x_E'(p_0) \geq 1$$

Eventually, in case any of the concave/convex properties is strict, the inequality in the last equation is strict, hence according to Bolzano’s theorem there exists at least one $p^* \in [0;1]$ with $x(p^*) - p^* = 0$ which is an interior (non-zero) fixed point.

Finally, if we additionally assume that $x_Q$ and $x_E$ are concave, $x$ as in (6) is concave itself, and the non-zero fixed point is unique. □

ACKNOWLEDGMENTS

Part of this work has been funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 248567 for the ETICS project. Further information on the ETICS project is available at www.ict-etics.eu. Additional support from the European Union’s Marie Curie Programme in the framework of the RBUCE WEST International Research Chair NICE and from the COST Action IC1003 (QualiNet) is gratefully acknowledged. The authors would like to thank their colleague Stefan Suette for his help with the trial.

REFERENCES


