



What you pay is what you get?

Fundamentals of charging Internet services based on end user Quality of Experience

Peter Reichl¹ · Patrick Zwickl² · Patrick Maillé³ · Martín Varela⁴

Received: 18 April 2017 / Published online: 4 April 2018
© The Author(s) 2018

Abstract

Complementing the current paradigm change from QoS to QoE, we address fundamental QoE charging issues for Internet services from an end user perspective. Here, key issues arise from gaps of different information contexts involved, which have to be managed when introducing a QoE product. Hence, this paper analyzes the double role of prices for quality perception as well as the impact of QoE on user demand with the help of a fixed point model. Our model is consistent with real-world user behavior that we have observed during comprehensive user trials on quality perception for video on demand services. Based on these results, we propose a simple approach for convergence-based user classification, and discuss the complementarity of willingness-to-pay vs. subjective quality perception in service purchasing situations.

Keywords QoS · QoE · Network economics · Internet pricing · Service level agreement · Video on demand

Introduction and related work

For several decades, service quality in communications networks has been described more or less solely in terms of QoS (Quality of Service) parameters, like packet loss rate, delay, jitter, bandwidth etc. In the last few years, however, we observe a clear trend in academia and industry towards more user-centric concepts of service quality, eventually leading to a veritable paradigm change [24]. Especially

the notion of QoE (Quality of Experience) [12] has rapidly gained in importance as a way to capture the “overall acceptability of an application or service, as perceived subjectively by the end-user” [9]. In a much broader sense, more recently QoE has been defined as “degree of delight or annoyance of the user of an application or service”, resulting 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” [14]. Along these definitions, a comprehensive body of related work has developed QoE framework models as well as corresponding metrics and measurement methodologies [17, 13, 19, 26].

At the same time, the question of how to charge end users for their perceived quality has been largely neglected, even if appropriate charging is widely recognized as an indispensable prerequisite for putting quality differentiation into practice [10] and, in the case of QoS, has led to a plethora of related work (see for instance [3, 30, 31] for surveys and [2] for an introductory textbook). Hence we argue that, complementing the mentioned paradigm change, research on QoE-based charging deserves much more attention, while, apart from some early contributions [1, 5], research has insufficiently addressed this user-centric perspective up to now. This is especially problematic as QoE and utilities from QoE (for instance related to revenues or valuations of a service) are disparate concepts as pointed out in [34]. Of course,

✉ Peter Reichl
peter.reichl@univie.ac.at

Patrick Zwickl
patrick.zwickl@ait.ac.at

Patrick Maillé
patrick.maille@imt-atlantique.fr

Martín Varela
martin@callstats.io

¹ Research Group Cooperative Systems (COSY), University of Vienna, Währinger Strasse 29, 1090 Vienna, Austria

² AIT Austrian Institute of Technology, Center for Digital Safety & Security, Giefinggasse 4, 1210 Vienna, Austria

³ IMT Atlantique Bretagne-Pays de la Loire, Campus de Rennes, 2, rue de la Chataigneraie, 35576 Cesson-Sévigné Cedex, France

⁴ callstats.io, Helsinki, Finland

we are well aware of the fact that the effect of such pricing differentiation cannot reach the end user directly as long as operators stay with their current general attitude of turning their back on negotiating service level agreements (SLA) directly with end users. Nevertheless, understanding what the user really wants (and at which price) will be crucial both for creating future services and for successfully bringing them to the market.

Therefore, in this paper we aim at discussing several fundamental issues concerning charging for QoE, in order to lay the foundations for future research in this field. Significantly extending our earlier work [21, 27, 23], we propose and analyze a general model for QoE-based charging together with its empirical validation, which is based on the results of several comprehensive user trials described in [18, 28, 35]. As specific novel contributions, our model encompasses the joint effects of pricing on user perception and the impact of QoE on user demand, and allows to derive corresponding fixed points which characterize the equilibrium between QoS, QoE, price and demand for various types of users. The results of our user trials illuminate the convergence behavior of end users, who ponder over time the tradeoff between desired service quality and associated price, as well as a further economical and psychological aspects of QoE-based charging. Such tradeoffs mirror a conflict of interest for users around antagonistic quality and price preferences (as, rationally configured, high quality often comes with a high price and the other way round). To mitigate marketization challenges arising from this conflict, more sophisticated means for conveying such QoE products to the consumer will be needed in the future.

The remainder of this paper is structured as follows: section “[Background: context mismatches and information gaps](#)” provides a general analysis of the different information contexts involved in our problem and identifies the gaps between them. Based on this, section “[Fixed-point models for QoS- and QoE-based charging](#)” introduces several flavors of a fixed-point model for QoS- and QoE-based charging. Section “[End user convergence behaviour](#)” describes the setup of our user trials and presents some key results, which also allow deriving an efficient approach for user classification. In section “[Willingness to pay and user interaction behavior](#)”, we direct our interest towards end user willingness-to-pay (WTP), before section “[Summary and conclusions](#)” concludes the paper with a brief summary and an outlook on current and future work.

Background: context mismatches and information gaps

From a provider perspective, the transition from QoS to QoE mainly aims at better understanding the customer needs in order to improve the customer experience and/or more

efficiently use the available resources (i.e., loosely speaking, cost reduction through QoS provisioning, but only as long as customers will not be able to subjectively detect it). So far, however, the QoE concept has not been able to live up to these high promises, one of the reasons being that the economic utilization of QoE data obtained outside of concrete purchasing situations has remained obfuscated. Moreover, it is difficult to imagine that the price for a service will depend directly on the individual level of the experienced quality at the end customer side. In this sense, future “QoE products” will be restricted to offering different QoE levels for different prices, in an attempt to increase user delight. However, also in this restricted perspective it is essential to analyze the systematic influence of price (and price perception) on the quality perception of the end user (and vice versa), and to deal with the question of how to derive utilities from empirical QoE data, for instance using willingness-to-pay (WTP) as an ISP utility metric for a specific service, customer segment and context.

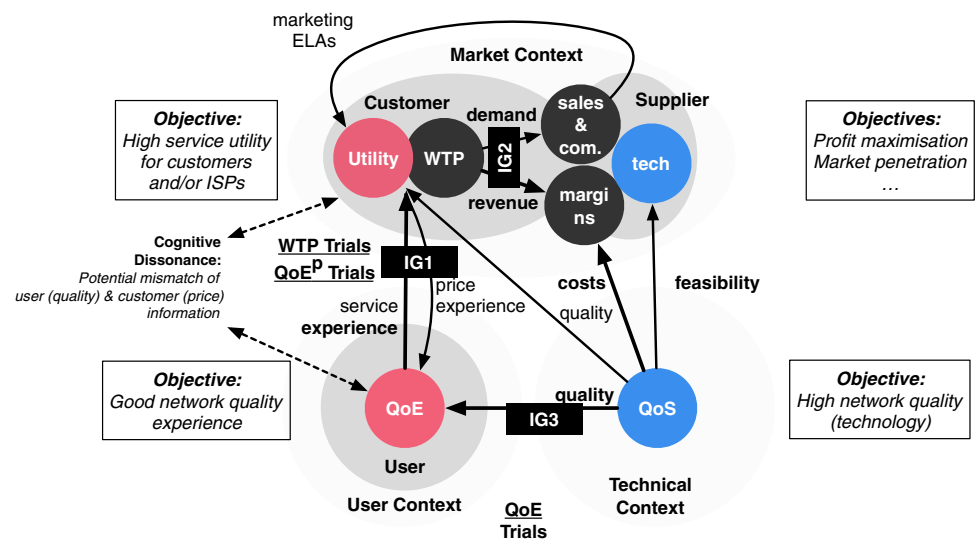
In order to illustrate this in more detail, we follow the top-down approach proposed in [33] and start with describing the big picture on information context(s) in the field of QoE, before distinguishing three corresponding information gaps that hamper the commercialization of QoE.

As depicted in Fig. 1, information (such as QoS or QoE data) is critical for the commercialization of network quality but inherently context-specific. Here, QoS is placed in a strictly technical context without involving any human role.¹ The missing human context is introduced with the creation of the QoE concept, within which a user role exists that maps technical and objective quality to a subjective quality appreciation (or “delight” according to [14]). The resulting mismatch of roles can only be resolved through empirical testing and thus has triggered numerous lab and field trials conducted in QoE research.

Hence, let us consider an abstract version of a QoE product which is offered in the context of a market with dedicated size and given characteristics (socio-economic background of population, overall size, etc.). Essentially, this QoE product refers to monetization of knowledge about quality experiences, especially by telcos and firms offering networked services (such as content providers). While direct per-service, per-user and per-usage monetization of QoE information is procedurally and legally challenging as well as highly contextual, indirect utilization of aggregate QoE information can serve the product design with much less effort. For example, QoE products may allow for QoE-aware capacity

¹ Note that, while this is largely consistent with today’s perspective on QoS versus QoE, originally QoS has been understood from a much more user-oriented perspective, for instance in ITU-T E.800 as “collective effect of service performance which determines the degree of satisfaction of a user of the service” [8]. For a more detailed account see e.g. [22].

Fig. 1 Information contexts and gaps [33]



dimensioning by telcos in order to efficiently and effectively occupy attractive market segments, or content providers may design their services for customer groups and their expected most common aggregate usage scenarios. Of course, indirectly QoE-aware products can only partially materialize on quality and price differentiation potentials for network services.²

As basic characteristic of this QoE product, we note that it is formed around the human roles of customers, users, and suppliers, which gives rise to three different types of *information gaps* (depicted as IG1–IG3 in Fig. 1). Firstly (IG1), we observe a role context mismatch between QoE assessment (= user role, quality perception metric) and commercialization (= customer role, revenue metric). Outside a purchasing situation, users may appreciate the provided quality level (i.e., high QoE), but would not have an interest in spending any money for this service or service quality (i.e., low revenue). In order to bridge this information gap, QoE could be assessed during purchasing situations where pricing p is involved. Hence, the pure quality rating (QoE) will shift to price-affected quality rating (QoE^p) where subjects can express their acceptance for a given quality level under a given price condition. Such a metric can provide solid indication for a purchase to be replicated under similar conditions. A more explicit revenue metric is given with WTP where the maximum revenue is assessed, i.e., referring to the classical customer role.

Secondly (IG2), the provider needs to optimize the QoE product to meet market objectives, typically in terms of limiting costs and maximizing revenues. While costs primarily refer to the capacity investments in order to meet

QoS requirements for the QoE product, the revenue is bound to the socio-economic factors of the market, the service expectation of the customers, and the customers' willingness to pay for the product. This optimization is non-trivial, as communicating with and/or conveying the QoE product to potential customers is critical and technically difficult. Moreover, as the QoE product can be classified as an experience good [20],³ it requires marketing and business strategy efforts especially during the market entrance phase [32].

Thirdly (IG3), QoE trials, bridging the information gap between technical parameters and experience, are inherently bound to their environmental context (scenario, tools, light, noise, stress, etc.), test scenario (e.g., video conference), and specific test conditions (e.g., parameters, bandwidth ranges or other QoS bounds). For instance, a noisy or stressful environment may alter the obtained QoE results substantially. In this sense, QoE is a *local* metric describing the subjective quality perception, and generalization may only be possible under certain, maybe quite limited circumstances, when sufficient empirical data is collected and no other information gap prevents it. Local normalization⁴ of QoE tests, e.g., through QoE training sessions as suggested by ITU-T Rec. P.10 [9], provides a standard approach to avoid this problem, however leads directly to a third type of context mismatch where global validity of QoE results is lost: measurement practices of QoE are local by definition and do not directly represent a global value, while the controllability of empirical trials (in the field as well as in the lab) requires a certain moderation of test cases and rating

² For more considerations on the procedural topics such as contractual and sales aspects, we refer to [32].

³ As customers cannot a priori assess experience goods, the set of available commercialization strategies is limited.

⁴ Normalization refers to the process of determining test ranges that define the perception and rating range.

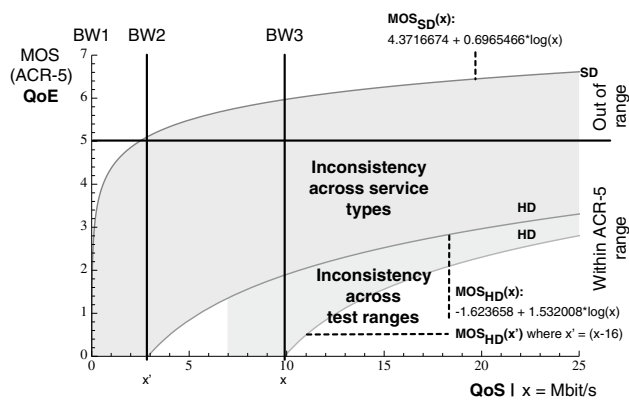


Fig. 2 Mismatch between QoE contexts: a small HD parameter range is contrasted with a global alternative and with an SD-based video QoE test [33]. MOS values (y-axis) below 1 and above 5 are considered out of range

behaviours. Hence, while the common practice to use training sessions for reducing the data noise may appear highly reasonable for reasons of testability, QoE is reduced to a perceptive sensing metric where the commercial utilization of results remains challenging.

Let us illustrate this last point with an example, and consider a QoE trial which afterwards is replicated with a different parameter range. For instance, suppose we will replicate the HD video QoE trial of [11] or alternatively the trials in [28, 29, 35], e.g., in a range between 1 kBit/s to 25 Mbit/s, and shift the tested video bitrate by 7 Mbit/s (i.e., $x' = x + 7$ Mbit/s where x is the initially tested bitrate, x' is the altered bitrate for the replicated test, and 7 Mbit/s is the bitrate offset that is exemplarily chosen). In the original and bitrate-shifted case, we may assume that the underlying laws and mechanisms remain identical [23], hence a similarly shaped QoS-to-QoE curve between the considered minimum and maximum quality should result—e.g., a logarithmic or exponential shape (cf. Fig. 2 based on the data of [4] and [11], in this case reconditioned logarithmically as recommended in [23]). Depending on the service type and scenario, a more or less sensitive reaction of QoE ratings can be expected upon QoS changes, i.e., the QoS-to-QoE mapping may vary in its steepness. However, when comparing the real and fictively replicated trial with shifted bitrates, different QoE values will be obtained for identical QoS input stimuli during the trial. Thus, the obtained QoS-QoE relations will not match across trials anymore, and we can expect that the maximum QoS of the initial trial will result in substantially higher average QoE ratings than in the rerun.⁵ Hence, the QoE results must be considered to be local to very specifically defined context, which cannot easily be transferred to a figure that

can be used for bringing enhanced network quality solutions to the market.

A similar situation may happen when a different variant of a service is tested, e.g., transitioning from high definition (HD) to standard definition (SD) videos. While the controllability of test routines aims at improving the interpretability of the data, it counteracts this effort by highly limiting the information gain from such empirical QoE trial for the commercialization of network quality.

On the other hand, monetary metrics (such as money used for market or other financial figures) are *non-local* and linear, and apart from specific contexts have a global and universal validity and value. Hence, the context-specific service experience needs to be mapped to a universal metric, which creates a knowledge gap as denormalization functions for QoE results are not yet available in general.

Altogether, these three information gaps affect the understanding and marketing of QoE products: any role context produces highly contextual information, which has an immediate consequence for QoE and WTP (or other utility-related data) that obviously are in different role contexts, and the mapping of corresponding data becomes a non-trivial task (and may require extensive user trials as a direct technique for bridging the different contexts). For instance, a clear mismatch between the appreciation for high QoE and the willingness to spend any money for this kind of service exists. Furthermore, the cognitive dissonance assessment in [28] hints towards the possibility of a direct influence of pricing on the subjectively perceived network quality, indicating a non-trivial (non-linear) mismatch between theoretical QoE tests without price considerations and QoE assessments during purchase situations. For these reasons, a better understanding of pure service valuation or preference is required for projecting QoE results to monetary means.

In the rest of this paper, we propose, analyze and discuss several approaches for resolving these information gaps. We will focus on subjects who provide QoE ratings during a purchase situation while at the same time assessing their quality appreciation, which allows to eliminate the gap between different roles, as in this case both user and customer roles are active during the assessment. The mismatch between the resulting QoE^p and QoE is characterized by an interesting fixed-problem problem which we address in detail in “Fixed-point models for QoS- and QoE-based charging”, where pricing as such creates a negative feedback that has to be incorporated into the quality experience metric. On the other hand, WTP has similar characteristics as QoE^p , but provides directly usable data for the ISP (e.g., ISP revenue data can be derived); hence, assessment methods and first results for QoE^p and WTP are presented and compared in section “Willingness to pay and user interaction behavior”.

Furthermore, our work is based on comprehensive empirical testing across services, scenarios and test ranges, which

⁵ This holds as long as the subjects have not participated in both trials.

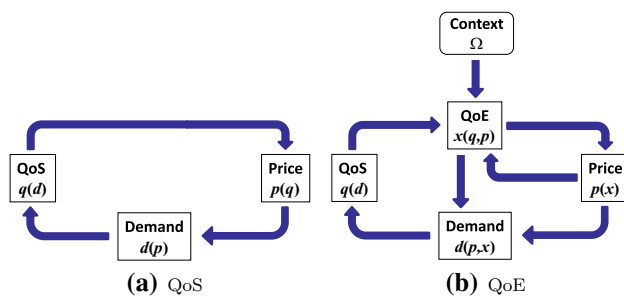


Fig. 3 Charging for network quality

altogether allows a generalization of the results, without, however, automatically getting rid of the problems associated with local normalization. This is only possible if we assume a global human perspective during purchase situation. Here, due to the user-customer role inclusion, a direct relationship to WTP appears realistic to be formalized. Thus, based on the extensive experience with highly controlled QoE trials with locally normalized output data, we are confident that the research community has by now developed the tools and obtained the experience to target QoE^p and similar output metrics.

Fixed-point models for QoS- and QoE-based charging

In order to illustrate the basic conceptual difference between traditional charging based on QoS only and QoE-based charging, in [27] we have considered a simple dynamic system where a provider with limited resources offers a service to end customers, see Fig. 3a. While QoS provisioning and pricing is up to the provider, the users are able to decide on their demand, which of course will depend on the price (e.g., according to the classical concept of “price elasticity” [16]). On the other hand, because of resource limitations in the provider’s network, the size of overall demand will influence service quality, according to some (more or less sophisticated) mechanism matching demand and supply. Finally, we assume that the QoS level delivered by the operator is reflected in a corresponding tariff structure, and thus in the price eventually charged from the customer—cf. Fig. 3a.

Note that, in this model, the user perspective is restricted to her decision to buy a certain amount of resources for the currently valid price, while on the provider side we assume a monopolistic situation such that there is only one service type offered that cannot be substituted by another service type (with maybe different price associated). In contrast, Fig. 3b depicts our structural model for QoE-based charging, where the additional QoE component has direct impact on the demand and is

influenced by the context of the user (service, environment, mood, etc.), and especially the price to be charged. In the remainder of this section, we will analyse both these models step by step.

Charging for QoS: adaptive users

For the formal analysis of the model for QoS-based charging (see Fig. 3a), where users are “non-sensitive” w.r.t. QoE at all and just adapt their demand according to price and QoS, let p indicate the price, d the demand and q the QoS. The resulting dynamic system is described by the following set of equations:

$$\text{Price function} \quad p = p(q) \quad (1)$$

$$\text{Demand function} \quad d = d(p) \quad (2)$$

$$\text{QoS function} \quad q = q(d) \quad (3)$$

where, after appropriately rescaling the boundary values of the functions, without loss of generality we assume all functions to be continuous bijective mappings of the unit interval $[0, 1]$ onto itself.

With respect to the shape of these functions, during this entire section we make the following assumptions:

- (A1): $d(p)$ is monotonically decreasing and convex. The monotonicity of $d(p)$ is straightforward, the convexity results from the asymptotic behavior, as for the case of high prices the demand will tend towards zero.
- (A2): $q(d)$ is monotonically decreasing and concave. Due to the boundedness of underlying network resources, increasing demand will result in lower network QoS. However, in general any substantial QoS degradation requires a significant reduction of the available network resources while, on the other hand, typically services (especially real-time ones) have some minimum resource requirements (e.g., basic connectivity). Together, this motivates the concavity of the function.
- (A3): $p(q)$ is monotonically increasing and concave. Services should not become cheaper if their quality is improved, while they cannot be sold at all if they become overly expensive. Both conditions together result in the assumption.

Proposition 1 Assume (A1), (A2) and (A3) to be valid. If $d(p)$, $q(d)$ and $p(q)$ are linear, the entire unit interval constitutes a set of fixed points for the system of Eqs. (1)–(3). Otherwise, if at least one of the concavity and/or convexity conditions is strict, there are only two (trivial) fixed points: one at $(p^*, d^*, q^*) = (0, 1, 0)$ and one at $(p^*, d^*, q^*) = (1, 0, 1)$.

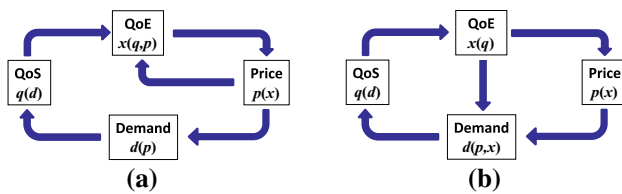


Fig. 4 Charging for QoE: price-sensitive (a) vs quality-sensitive users (b)

Proof 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 fixed point price $p^* \in [0, 1]$, we have

$$\begin{aligned} d(p^*) &= 1 - p^* \\ q(d(p^*)) &= 1 - d(p^*) = p^* \\ p^*(q(d(p^*))) &= q(d(p^*)) = p^* \end{aligned}$$

More generally, for non-linear functions these equations generalize towards

$$\begin{aligned} d(p^*) &\leq 1 - p^* \\ q(d(p^*)) &\geq 1 - d(p^*) \geq p^* \\ p^*(q(d(p^*))) &\geq q(d(p^*)) \geq p^* \end{aligned}$$

Hence, if any of the inequalities is strict in the open interval $]0, 1[$, then $p^* > p^*$, which excludes the existence of additional interior fixed points. \square

NB: Further numerical evidence suggests that $(0, 1, 0)$ is unstable and $(1, 0, 1)$ is stable.

QoE-based charging—submodel 1: price-sensitive users

In contrast to the simple model with predefined preferences considered so far, the situation is more complex when the charging mechanism integrates QoE as subjectively perceived by the end customer. Before addressing the full model for QoE-based charging (Fig. 3b), we first analyse two submodels, depending on whether the QoE perception of (“price sensitive”) users is based on their expectations due to the pricing plan, or (“quality sensitive”) users rather handle prices and QoE as independent parameters which separately influence their purchasing decision (and thus the demand), see Fig. 4.

Starting with the price sensitive submodel, note that, in our context, price plays a double role: on the one hand, like with the QoS-based model, the end customer pays for the level of offered service quality (hence prices are supposed to rise with increasing QoE), on the other hand the price to be paid forms also part of the user context [13] 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, loosely speaking, the larger the probability to be disappointed). Hence, we may postulate that higher tariffs induce the QoE evaluation to deteriorate.

Figure 4a depicts the resulting price-sensitive submodel. Observe that there is a new function involved termed *QoE function*, which reflects the mentioned janiform role of prices and hence depends on both the QoS level q and the price level p . Consequently, we may formulate a new system of equations where (2) and (3) remain identical, Eq. (1) becomes

$$\text{Price function} \quad p = p(x) \quad (4)$$

and we have an additional equation for the QoE function

$$\text{QoE function} \quad x = x(q, p). \quad (5)$$

First of all, we extend our prior set of assumptions as follows:

- (A4): $p(x)$ is monotonically increasing and concave. This is in strict analogy to (A3), just replacing q by x .
- (A5): $x(q, p)$ is monotonically increasing and concave in q , and monotonically decreasing in p . Better QoS should not deteriorate the associated QoE, while higher tariffs lead to higher quality expectations which lead to lower subjective quality perception. The concavity of the function can be argued in analogy to (A2), as it usually takes a severe degradation of QoS until the user subjectively perceives a noticeable degradation of QoE.

Indeed, for constant (e.g., flat) prices, empirical results indicate that for a broad range of scenarios, QoE depends logarithmically on the offered bandwidth (= QoS) [23], while, for constant QoS, already earlier QoE has been postulated to decrease with rising prices, see [27]. The resulting two-dimensional QoE function can of course be of a rather general form—to simplify a bit, we assume both effects to be independent from each other, and hence x to be separable, cf. [27]:

$$x = x(q, p) = x_1(q) \cdot x_2(p). \quad (6)$$

We interpret (6) as follows: the QoE depends on both the QoS offered by the provider—in terms of a monotonically increasing *quality function* $x_1(q)$ —and the customer expectations triggered by the corresponding tariff (= price), which

is expressed in terms of a (monotonically decreasing) *expectation function* $x_2(p)$. Without loss of generality we assume that $x_2(0) = 1$ and $x_2(1) = 0$, while x_2 is not subject to further restrictions (in fact, there is little empirical evidence so far about a reasonable shape of this function).

Proposition 2 Assume (A1), (A2) and (A4) to be valid, $x_1 : [0, 1] \mapsto [0, 1]$ to be monotonically increasing and concave, $x_2 : [0, 1] \mapsto [0, 1]$ to be monotonically decreasing, and all functions to be continuously differentiable. Then, if all functions are linear, the system of Eqs. (2)–(6) has exactly one fixed point at $(p^*, d^*, q^*) = (0, 1, 0)$. Otherwise, if at least one of the concavity and/or convexity conditions is strict, there is an additional stable fixed point p^{**} which is non-trivial, i.e., $p^{**} \in]0, 1[$. If, in addition, x_2 is concave, this non-trivial fixed point is unique.

Proof In addition to Proposition 1, we assume also $x_1(0) = x_2(1) = 0$ and $x_1(1) = x_2(0) = 1$. If all functions are linear, we have

$$\begin{aligned} d(p^*) &= 1 - p^* \\ q(d(p^*)) &= 1 - d(p^*) = p^* \\ x(q(d(p^*)), p^*) &= p^* - p^{*2} \\ p(x(q(d(p^*)), p^*)) &= p^* - p^{*2} \end{aligned}$$

From the last equation, we see that $(p^*, d^*, q^*) = (0, 1, 0)$ is the only fixed point, with $x_1(0) = 0$ and $x_2(0) = 1$, i.e., $x(0, 0) = 0$.

For non-linear functions, similarly to the proof of Proposition 1, we derive that

$$x_1(q(d(p))) \geq p^*,$$

hence x_1 is a concave function of p with $x'_1|_{p=0} \geq 1$, and

$$x(q(d(p)), p) = x_1(q(d(p))) \cdot x_2(p)$$

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

$$x'|_{p=0} = x'_1|_{p=0} \cdot x_2(0) + x_1|_{p=0} \cdot x'_2(0) \geq 1.$$

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-trivial) fixed point. Finally, if we additionally assume that x_1 and x_2 are concave, then x as in (6) is concave itself, and the non-trivial fixed point is unique. \square

NB: Numerical evidence suggests that in the general case p^* is unstable while p^{**} is stable.

QoE-based charging—submodel 2: quality-sensitive users

For the case of a quality sensitive user, the corresponding submodel is depicted in Fig. 4b, where now the demand function has two input parameters: price and QoE. In analogy to Submodel 1, we assume for $d(x, p)$ also here separability into a *quality-driven demand function* $d_1(x)$ and a *price-driven demand function* $d_2(p)$:

$$\text{Demand function} \quad d = d(x, p) = d_1(x) \cdot d_2(p), \quad (7)$$

while, in this submodel, QoE depends on QoS only:

$$\text{QOE function} \quad x = x(q). \quad (8)$$

While (A5) still applies for this $x(q)$ (just assuming independence from p in Eq. (6)), the underlying assumption for the new demand function in (7) is motivated as follows:

- (A6): $d(x, p)$ is monotonically increasing and concave in x and monotonically decreasing in p . The suggested monotonicity originates from the typical antagonistic demand behaviour: better quality is assumed to increase demand, higher prices decrease demand, see (A1). Concavity in x results in both competitiveness and boundedness of the market: while already a small advantage in terms of perceived quality leads to a significant increase of the market share, this effect becomes weaker for larger established market shares.

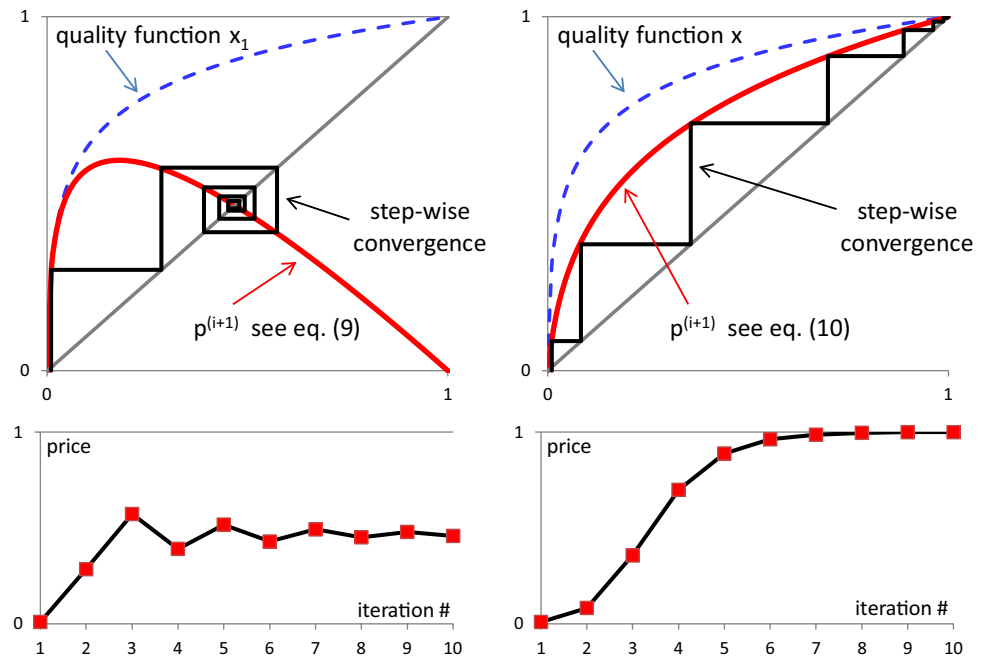
Proposition 3 Assume (A1), (A2) and (A4) to be valid, $d_1 : [0, 1] \mapsto [0, 1]$ to be continuous and monotonically increasing and concave, and $d_2 : [0, 1] \mapsto [0, 1]$ to be continuous and monotonically decreasing. Then, the system of Eqs. (3), (4), (7) and (8) has a unique fixed point at $(p^*, d^*, q^*) = (1, 0, 1)$.

Proof In addition to Proposition 1, we assume $d_1(0) = d_2(1) = 0$ and $d_1(1) = d_2(0) = 1$. If all functions are linear, we have

$$\begin{aligned} x(q(d(p^*, x))) &= q(d(p^*, x)) \\ &= 1 - (1 - p^*) \cdot x \Rightarrow x(q(d(p^*, x))) \\ &= \frac{1}{2 - p^*} p^* (x(q(d(p^*, x)))) \\ &= \frac{1}{2 - p^*} \Rightarrow (p^* - 1)^2 = 0. \end{aligned}$$

Hence $(p^*, d^*, q^*) = (1, 0, 1)$ is the unique fixed point.

Fig. 5 Step-wise convergence (top) and first 10 iterations (bottom) for strictly concave quality functions: quality-sensitive users (left column) and price-sensitive users (right column)



As in the proof of Proposition 2, for the general case we have

$$p(x(q(d^*))) \geq x(q(d^*)) \geq 1 - d^*$$

and hence, with $d_1(p^*) \leq 1 - p^* \leq d^*$ and $d_2(x) \leq 1$,

$$d^*(p, x) = d_1(p) \cdot d_2(x) \leq d^*.$$

Thus, if any of the inequalities is strict on the open interval $]0, 1[$, this leads to a contradiction, hence no additional non-trivial (interior) fixed point exists in this case. \square

Numerical examples for submodel 1 and submodel 2

For the purpose of illustrating the process of convergence in more detail, consider two simple examples, i.e., a purely price-sensitive user and a purely quality-sensitive user, whose respective quality functions $x_1(q)$, see (6), and $x(q)$, see (8), both are assumed to have a logarithmic shape [25] while all other functions are supposed to be linear. Fig. 5 depicts the resulting step-wise convergence behavior where $p^{(i)}$ refers to the i -th iteration of the price towards equilibrium and $p^{(0)}$ is assumed to be slightly larger than 0 to avoid trivial fixed points. Note that this choice corresponds to the fact that in the user trials described in section “End user convergence behaviour”, each experiment has started with the quality level corresponding to the lowest price.

Hence, for the case of the price-sensitive user (Submodel 1), Eqs. (4), (6), (3) and (2)—applied in this order—lead to

$$\begin{aligned} p^{(i+1)} &\leftarrow p(x^{(i)}) = x_1(q^{(i)}) \cdot x_2(p^{(i)}) \\ &= x_1(1 - d^{(i)}) \cdot (1 - p^{(i)}) = x_1(p^{(i)}) \cdot (1 - p^{(i)}) \end{aligned} \quad (9)$$

Similarly, for the case of the quality-sensitive user (Submodel 2), Eqs. (4), (8), (3), (7) and again (4)—applied in this order—result in

$$p^{(i+1)} \leftarrow p(x^{(i)}) = x(q^{(i)}) = x(d^{(i)}) = x(d_1(x^{(i)}) \cdot d_2(p^{(i)})) = x(p^{(i)} \cdot p^{(i)}) \quad (10)$$

Note that all functions have been normalized to the unit interval as described previously, hence all linear functions are either the identity function id or $1-id$, depending on their slope (positive or negative), while in both cases the respective quality functions are normalized logarithmic functions.

We consider these examples to reflect two very fundamental patterns that are widely observed also in practice, as will be demonstrated in the next chapter. Hence, we will come back to Fig. 5 in the course of section “Convergence-based user classification” as primary inspiration for our user classification approach presented there.

Full model: price- and quality-sensitive users

Turning now to the full model as depicted in Fig. 3b, firstly let us subsume all user-specific impact factors, e.g., due to personality or current contextual state [14], under the overarching notion of a *Context Function* Ω . In fact, as such context factors are external to the model, we can easily

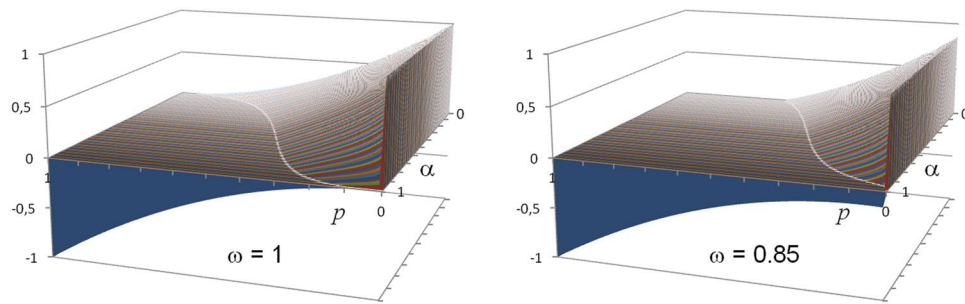


Fig. 6 Fixed point characterisation for the full model according to (13): left for $\omega = 1$, right for $\omega = 0.85$. The saddle-like surface depicts the left side of Eq. (13) depending on price p (x-axis) and α

(y-axis) and shows where this expression equals zero (curved line). Note that, for reasons of better illustration, both α - and p -axes are depicted in reverse directions

aggregate them by multiplying the *QoE Function* (5) with a scalar parameter $\omega \in (0, 1]$. This is in line, for instance, with the approach of the well-known E-model [7] for voice QoE, which derives a quality metric by subtracting certain impairment factors from an initial QoE value corresponding to the absence of any impairment. Along these lines, we hence define $\omega = 1$ as user QoE in an optimal context (i.e., without any form of distraction etc.), and describe any context-based impairment by reducing ω accordingly.

which, unfortunately, can no longer be solved in closed form.

Proposition 4 For $\alpha \in]0, 1[$, $f_\omega(\cdot, \alpha)$ is continuous, and (13) has a unique solution.

Proof First, observe that $f_\omega(0, \alpha) = \omega > 0$ and, on the other hand, $f_\omega(1, \alpha) = -1 < 0$, ensuring the existence of a solution from the intermediate value theorem. Moreover, such a solution is in the open interval $]0, 1[$, and is unique since $f_\omega(\cdot, \alpha)$ is strictly decreasing on this interval, see Eq. (14).

$$\frac{\partial f_\omega}{\partial p} = \omega \cdot \left(\underbrace{(1 + \alpha)(1 - p)^\alpha}_{<1} \underbrace{p^{1-\alpha}}_{>0} - \underbrace{(1 - \alpha)(1 - p)^{1+\alpha} p^{-\alpha}}_{>0} - \underbrace{\alpha(1 - p)^{\alpha-1}}_{>\alpha(1-p)^\alpha} \right) - 1 < \omega \cdot (1 - p)^\alpha - 1 < 0. \quad (14)$$

In order to create a unified full model which integrates both submodels as special cases, we may take strong advantage of the product form of $x(q, p) = x_1(q) \cdot x_2(p)$ in (6) and $d(x, p) = d_1(x) \cdot d_2(p)$ in (7), resp., and have to achieve that $d_1(x) \equiv 1$ for the case of Submodel 1 (price-sensitive user, cf. (2)), and $x_2(p) \equiv 1$ for the case of Submodel 2 (quality-sensitive users, cf. (8)). The easiest way to fulfill these requirements is by defining the following system of equations for a sensitivity parameter $\alpha \in [0, 1]$:

$$x = x(q, p) = \omega \cdot x_1(q) \cdot x_2(p)^\alpha \quad (11)$$

$$d = d(x, p) = d_1(x) \cdot d_2(p)^{1-\alpha} \quad (12)$$

Observe that, for $\alpha = 1$, these equations coincide with (2) and (6), resp. (i.e., price-sensitive case), whereas for $\alpha = 0$, they coincide with (7) and (8), resp. (i.e., quality-sensitive case).

For any other value of α , if all functions are assumed to be linear, the resulting fixed point equation (with $\omega \in]0, 1[$) reads

$$f_\omega(p, \alpha) = \omega \cdot (1 - (1 - p)p^{1-\alpha})(1 - p)^\alpha - p = 0, \quad (13)$$

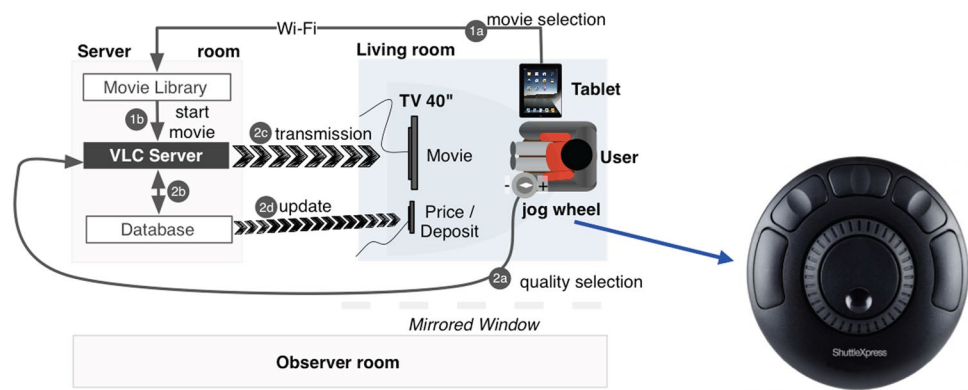
Figure 6 depicts $f_\omega(p, \alpha)$ for $(p, \alpha) \in [0, 1] \times [0, 1]$ with $\omega = 1$ and $\omega = 0.85$, illustrating existence and uniqueness of the fixed point in two typical cases. Observe that, independently of ω , $\alpha = 1$ results in the price $p = 0$ derived in Proposition 2 for price-sensitive users, whereas the high resulting price for $\alpha = 1$ is characteristic for quality-sensitive users according to Proposition 3.

NB: Also in this case, numerical evidence strongly suggests these fixed points to be stable.

Conclusions for sensitivity-based user classification

Summarizing the above fixed point (FP) analysis of our models for different user types w.r.t. charging for QoE, Proposition 1 tells that *adaptive users*, who do not care for QoE at all, go either for free low quality (unstable FP) or expensive high quality (stable FP). Proposition 2 shows that *strictly price-sensitive users* have an unstable FP at low quality plus

Fig. 7 Technical setup of the user study



a stable one who is non-trivial. In contrast, Proposition 3 states that *strictly quality-sensitive users* strongly prefer expensive high quality. Finally, according to the Full Model, *mixed users* ($0 < \alpha < 1$) end up at a stable non-trivial FP which depends on the sensitivity parameter α as depicted in Fig. 6.

A closer look to Fig. 6 reveals that the latter dependence on α is not linear, but that the FP prices stays within a relatively small interval if α is varying around 0.5, whereas the curve is rather steep for α close to 0 or 1. Hence, if we assume the sensitivity parameter for cross-sensitive users to be uniformly $U[0; 1]$ distributed, we may conclude that the FP prices for the majority of cross-sensitive users will concentrate in the middle of the unit interval, whereas both ends (low and high prices, resp.), are only lightly populated.

Altogether, if in general the overall population typically consists of some adaptive, some price- or quality-sensitive and a majority of mixed (cross-sensitive) users, we may expect the resulting distribution of FP prices to exhibit clear peaks for the lowest and the highest prices (unstable part of adaptive or price-sensitive users + quality-sensitive users, resp.), with a “bell curve” in between (cross-sensitive users + stable part of price-sensitive users). For the moment, this hypothesis is a simple prediction resulting from our rather simple models, however, in the next section we will see that it is surprisingly consistent with real-world user behavior.

End user convergence behaviour

After this extensive fixed point analysis of analytical models for QoE-based charging, we now turn towards end user convergence behavior as exhibited in practice. To this end, we refer to the results of three comprehensive user trials,⁶

⁶ Parts of these trials has been performed in the EU FP7 ETICS project [15] and the EU COST Action IC 1003 QualiNet [14].

subsequently referred to as the 2011, 2012 trial and 2015 trials. The 2011 trial was intended to assess the general readiness to pay for improved network qualities and associated services by studying purchasing behaviours of customers for packet loss-impaired UDP streams. The subsequent and more elaborate study design of 2012 and onwards, however, aimed at providing deeper insights on two aspects: (1) the separation of pricing and quality motives for the studied purchasing decisions, and (2) the identification of the absolute maximum WTP for network quality using a more realistic technological setup (i.e., video streams with various bitrates and prevailing video codecs using TCP transport). These results were compared to a retesting in Vienna (Austria) and Oulu (Finland) in 2015 [18] (i.e., the 2015 trial), where the video technologies were updated from H.264 (2012) to the new H.265 codec (2015), and the setup was reparametrized accordingly.

The subsequent analysis will primarily concentrate on our 2012 trial, but will selectively use the 2015 trials for elaborating on details or validating findings. It further builds on the outcome of the 2011 trial that a general readiness to pay exists for enhanced (video) network quality. In the following subsections, we briefly describe the trial setups, then study the convergence behaviour of end users, and present an approach for their algorithmic classification.

User trial setup and general results

The technical setup of our 2012 user trials is a modified version of the one used in 2011 and described in [28, 29], and is reused using a modernised toolset in the 2015 validation testing. The basic setup is depicted in Fig. 7. While in [28, 29] we have distinguished between four classes of Standard Definition (SD) video quality based on different packet loss levels for UDP-based transmission, the revised setup aims at being significantly closer to reality. We use the TCP-based adaptive video stream technology *HTTP Live Streaming (HLS)* in order to adapt video quality to network conditions, i.e., mainly bandwidth. Using high definition (HD) blue-ray quality allows differentiating a large

Table 1 Quality classes (in kbit/s)

Q0	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
128	181	256	362	512	724	1024	1448	2048	2896
Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19
4096	5793	8192	11585	16384	23170	32768	32768	32768	32768

Table 2 Quality classes Q_0 – Q_7 in kbit/s (2015 trial, H.265 codec)

Q_0	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7
128	256	512	1024	2048	4096	8192	16,384

number of different quality levels (17 in our case), based on logarithmically scaled bitrates (H.264 encoding). For cross-validation purposes, we have included three additional “virtual” quality classes which are identically employing the best possible bitrate but still differ in terms of prices. Hence, the trial subjects have been exposed to a total of 20 offered quality classes, see Table 1. Note that, as the virtual classes Q17–Q19 identically offer the highest available video bitrate at still growing prices, selecting a higher quality class than Q16 seems irrational, which is an essential test design addition in order to reveal dominating factors for the purchasing behavior, i.e., price-/quality-sensitivity.

In the retesting in 2015 [18], due to the newer codec (H.265 instead of H.264) lower maximal bitrates were tested, i.e., 16384 kbit/s. Instead of 20 quality classes, only 8 were used for increasing the sample sizes within each quality class, see Table 2. The proprietary HLS solution was further replaced by the similarly functioning *Dynamic Adaptive Streaming over HTTP (DASH)* standard [6]. Otherwise the technical setup closely follows the 2012 trial.

For the logarithmically increasing bitrates in from Q0 to Q16 (see Table 1) and Q0 to Q7 (see Table 2), respectively, the trial design in 2012 and 2015 used three tariffs: low-tier *tariff A*, medium-tier *tariff B* and high-tier *tariff C* with maximum prices p_{max} of EUR 2.00, EUR 3.00 and EUR 4.00, resp. Between the identical minimum price of zero for Q0 and the respective p_{max} for Q19 an increasing price curve was used which is linear with respect to the classes. Hence, for the highest bitrate 4 different prices were listed in order to test other forms of price discrimination. While using identical maximum prices of EUR 2.00, EUR 3.00 and EUR 4.00, the retests in 2015 only used 8 price steps with corresponding quality steps (i.e., the price curve was not additionally stretched for additional quality classes).

Like with [5] and [28], trial users have been given real money—10 Euro each in our case—which they could freely spend on quality enhancements during the trial or take home afterwards. Together with the mentioned fine granularity of quality classes, this setup allows for observing detailed user interaction behavior. During the trial, each test subject watches three video sequences (each 20 min long) individually chosen from a representative video library (including,

e.g., highly topical blockbusters). Starting per default with the poorest quality level Q0 (which remains free of charge during the entire trial), the subjects can use a jog wheel for dynamically and interactively testing the effect of quality adjustments during an initial period of around 5 mins free of charge. Users are always informed about the price of each selection, while the range of available quality levels and tariff designs are intentionally hidden for reducing unnecessary biases. Due to a highly improved setup (an own VLC client fork, plus precise logging mechanisms), quality changes are quickly applied (with a delay of about 1 s only) and tracked with granularity of 1 s. After the free trial phase, the latest selected quality level is taken as final choice, and the corresponding price is deducted from the user’s cash deposit. Now, users watch the remaining movie clip without any further interaction in the chosen quality. After the experiments, the remaining deposit is paid out in cash to the subjects, as announced before the trial.

Using a jog wheel as physical user interface for changing between quality classes provides significant benefits, as it employs an intuitive mechanism well-known to all users (for instance from sound volume control). Moreover, it creates the illusion of an infinite number of quality levels (of course, in reality requests for lowering the quality below Q0 or above Q19—or Q7 as in the 2015 case—have simply been ignored). In this way, it was possible to record (with temporal granularity of 1 s) 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.).

Overall, 43 test users (12 male, 31 female) have participated in our 2012 trial, 40 of whom have completed successfully, with three movies for each user. Our number of participants is well in line with typical standard sizes for user trials (e.g., VQEG methodology: 24 valid correlating subjects, ITU-T Rec. P.910: 4 to 40 test users). Most of them (26 persons) were between 18 and 30 years old, 10 persons between 31 and 45 years, and 7 older than 45 years. 38 participants had a higher school or university degree, most of them were employed (16) or students (16). 8 persons were married or living in a relationship. Only 3 users had previous

Table 3 Tariff assignment

User groups	Tariff sequence	Remarks
Group 1	$A \rightarrow B \rightarrow C$	Increasing prices
Group 2	$C \rightarrow B \rightarrow A$	Decreasing prices
Control group	$B \rightarrow B \rightarrow A/C$	Initially constant prices

experience with charged video on demand (VoD) services, with unsystematic expenditures varying from 3 EUR to 8 EUR per movie.

The validation trial in 2015 was divided in two regional cases using identical tooling and almost identical video contents: the Vienna and the Oulu campaign. In Vienna, 22 test users have participated, of whom 41% were female and 86% had graduated from a university before. Most of the subjects were between 20 and 29 years old (11 persons), 2 persons were between 10 and 19 years, 6 persons between 30 and 39, 1 subject was between 40 and 49, and 2 subjects were older. Their experiences with VoD services was still limited (32% with subscription; 55% had purchased video contents online before). In Oulu, 19 additional test users participated, of whom 21% were female and 90% had graduated from a university. Most test users were between 30 and 39 years old, 2 persons were younger, 7 persons were older. 68% had seldomly purchased video contents online before, 58% had a video service subscription, and 32% did not use any VoD service at all.

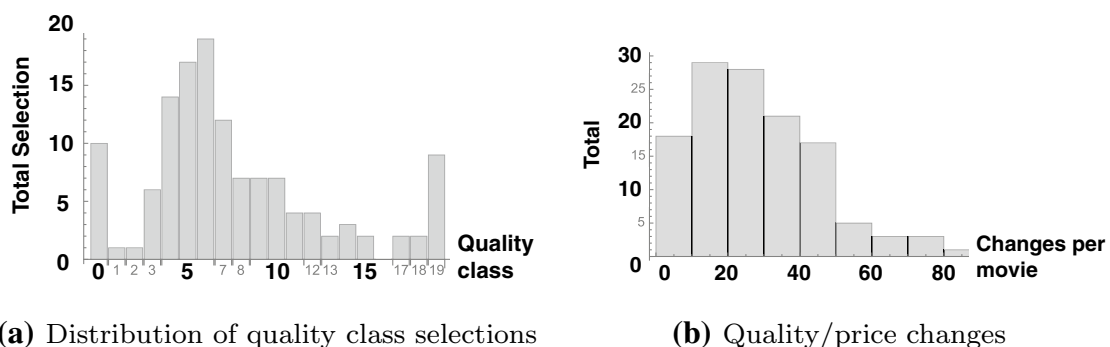
Users were randomly assigned to tariff schemes *A*, *B*, *C* as introduced above, however, due to the limited sample size, only a representative subset of the potential tariff permutations have been tested, see Table 3 for details. Note that Group 1 addresses the case of monotonically increasing prices, while Group 2 experiences decreasing prices. For the control group, prices are kept constant on medium level *B* for the first two movies, while the third movie is randomly assigned to either price plan *A* or *C*. In this way, sample sizes for the three tariffs are kept in balance, while

the tariff distribution allows further analyses, cf. [35]. To increase the sample sizes within each group, Group 2 was eliminated in the validation testing in 2015. Moreover, the reduced number of quality classes also increases the sample sizes per quality class.

Before we address an approach for user classification based on their convergence behaviour in more detail, let us have a look at some interesting general results. To start with, Fig. 8a illustrates the distribution of the quality classes eventually selected by the users at the end of the free trial phase. Most notably, the nature of this empirical distribution confirms precisely the hypothesis following from our fixed point analysis: as predicted in “[Conclusions for sensitivity-based user classification](#)”, we observe clear local maxima at both ends together with a (skewed) bell shape in between.

Complementing this result, Fig. 8b right depicts the distribution of quality/price level changes of the 2012 trial (typically between 10 and 50 per movie) during the free 5 min trial phase, while Fig. 9a shows the distribution of the corresponding number of “trend changes”, i.e., ups and downs until the user decision converges towards a final value. Here, the majority of test subjects restrict themselves to a maximum of 5 trend changes, while a few undetermined test subjects were exhibiting a total of 25 or more trend changes.

Another interesting issue concerns the points in time when quality changes happen. Figure 9b for the 2012 trial (the 2015 results are structurally similar) indicates that indeed most quality changes have been undertaken towards the end of the free trial period. This seems counterintuitive, as test users always start with *Q0* (which, by the way, introduces a certain bias we have to be aware of) and have to increase the quality until the desired level is reached. However, the last few seconds of the trial period determine the quality for the rest of the movie, hence the aggregation of those anticipated future requirements may serve as an explanation of this particular behaviour.

**Fig. 8** Histograms on the user interaction with our quality market (2012 trial)

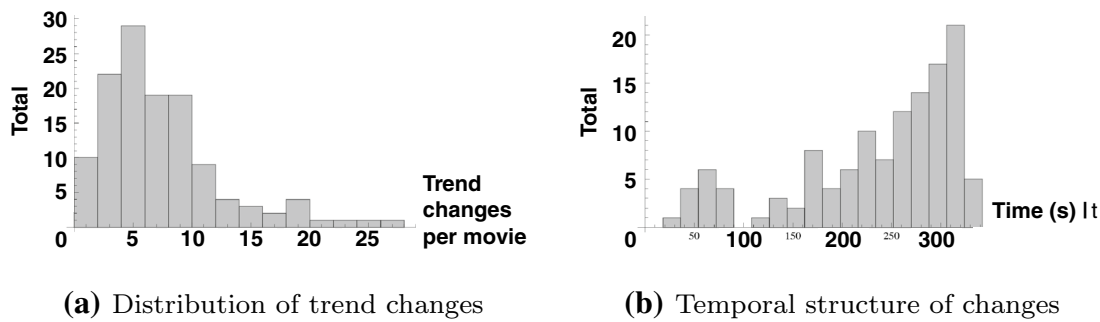


Fig. 9 Temporal histograms on user interaction with our quality market (2012 trial)

Convergence-based user classification

Having confirmed our hypothesis about a sensitivity-based user classification, we now present and discuss a quantitative approach for user classification, based on the assumption that users behave consistently during the entire trial, i.e., follow the same quality/price selection strategy for all movies. While this cannot be claimed to be universally valid (for instance, due to learning effects), for our purposes we may consider it a useful working hypothesis.

First of all, remember that the results of our model-based analysis in section “[Fixed-point models for QoS- and QoE-based charging](#)” as sketched in Fig. 5 suggest two fundamental user behavior patterns: either a damped harmonic oscillation (quality-sensitive users) or a steady increase towards the ideal price/quality level (price-sensitive users). While these two patterns strongly remind us of the two basic options (underdamped vs. overdamped case) for damped harmonic oscillators depending on the parametrization of the underlying second-order differential equations, we may simply interpret the latter pattern as a somewhat “degenerate” version of the former one, where the equilibrium price/quality is approached only from below, instead of turning an initially monotonic into an oscillatory behavior at all. Henceforth, in general we may characterize user behavior by two parameters: (1) a certain amplitude of quality/price selection, and (2) the speed of convergence towards the final choice. Within the resulting two-dimensional space, we end up distinguishing three different user classes labeled as “*F*” (fast convergence), “*R*” (regular convergence), “*S*” (slow convergence), which are illustrated in Fig. 10. Note that we have to supplement these three “regular classes” by another class “*X*” consisting of users with irregular behavior, who—for whatever reason—cannot be placed in one of the regular classes. For further illustration, we depict a couple of typical examples for the resulting classes which are taken from both Groups 1 and 2, where the different colors refer to the tariffs introduced earlier (green: *A*, blue: *B*, red: *C*). The hereinafter conducted assessment is centred around the 2012 trial data

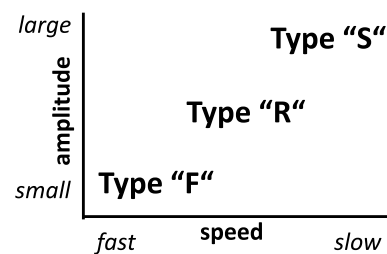


Fig. 10 Dimensions of user classification

due to its most comprehensive data due to a higher sample size and the most fine-granular tracking of user interaction behaviors.

Type “*F*”: fast convergence—small amplitude

The first general type of users is characterized by a relative consistent convergence behavior: users climb up the quality ladder until reaching the targeted quality, and stay there without major changes (see Fig. 11⁷).

Type “*S*”: slow convergence—large amplitude

Contrasting the examples depicted in Fig. 11, a second user type may be characterized exactly by the opposite behavior, i.e., large amplitude and slow speed of attenuation, see Fig. 12 (observe that the change frequency may be quite different, while the overall convergence speed seems roughly comparable).

Type “*R*”: regular convergence—medium amplitude

The third user type can be described as somewhat in between types *F* and *S*. It is characterized by large to medium

⁷ Note that in Figs. 11, 12, 13 and 14 the time on the x-axis is measured in seconds. Later on, the same applies to Eq. (16).

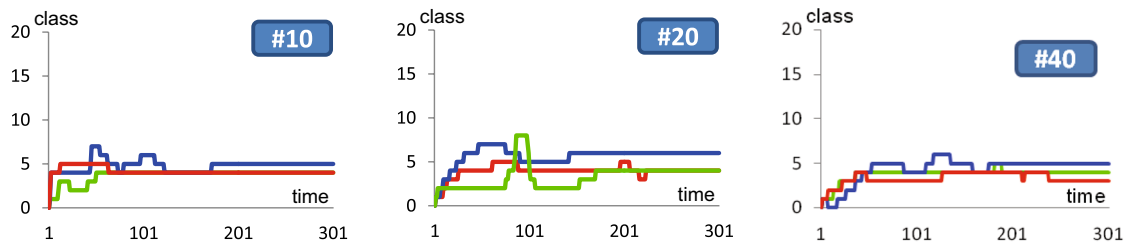


Fig. 11 Typical examples for user type “F” (test subjects #10, #20, #40)

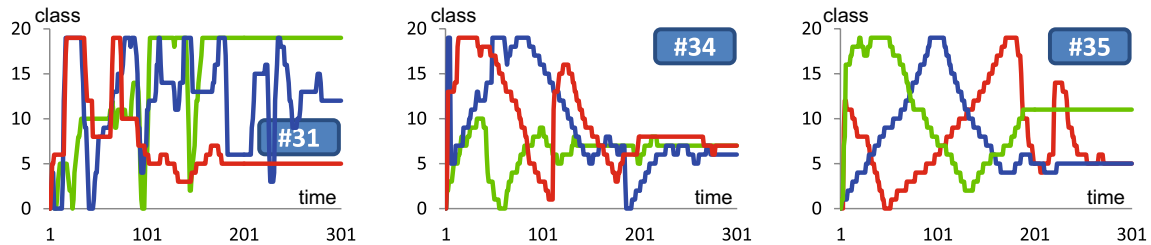


Fig. 12 Typical examples for user type “S” (test subjects #31, #34, #35)

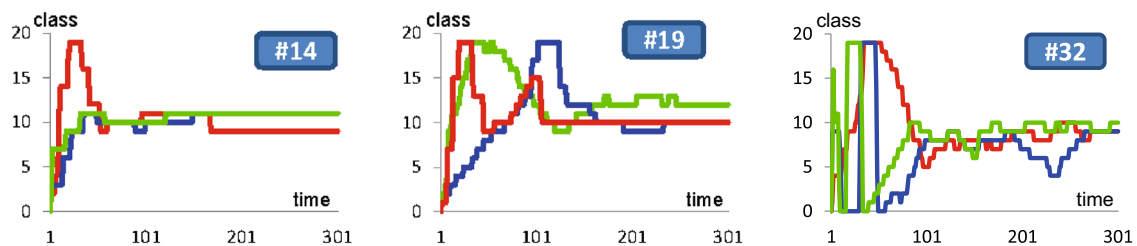


Fig. 13 Typical examples for user type “R” (test subjects #14, #19, #32)

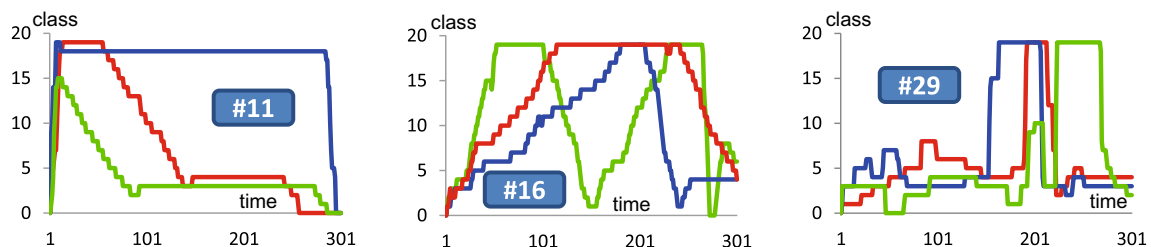


Fig. 14 Typical examples for user type “X” (test subjects #11, #16, #29)

amplitudes as well as medium speed of convergence, see Fig. 13 for typical examples. Observe how, in the first example, the behavior during the first movie is different from the others and more exploratory, suggesting a kind of learning behavior.

Type “X”: free riding and irregular user behavior

Finally, also some sporadic cases of free riding (Fig. 14 left) and further irregular user behavior (see, e.g., Fig. 14 right) have been observed (class “X”). Note, however, that

these cases account for only around 15% of the overall number of samples.

Aggregated convergence behavior

In order to capture the aggregated convergence behavior for user i , as described in [27] we define the root square deviation (RSD) $\tilde{\sigma}_i(t)$ as a function of time t with respect to the convergence value $\tilde{x}_i = \lim_{t \rightarrow \infty} x_i(t) = x_i(300)$ as follows:

$$\tilde{\sigma}_i(t) = |x_i(t) - \tilde{x}_i|. \quad (15)$$

Moreover, define

$$\begin{aligned} \tilde{\sigma}^F(t) &= 3 \cdot \exp(-0.01t) \\ \tilde{\sigma}^R(t) &= 6 \cdot \exp(-0.007t) \\ \tilde{\sigma}^S(t) &= 9 \cdot (1 - t/300) \end{aligned} \quad (16)$$

as reference root mean square deviations (RRMSD) of classes F , R and S , respectively. Then, for user i and class k ,

$$\delta_i^{(k)} = \sqrt{\frac{1}{300} \sum_{t=1}^{300} (\tilde{\sigma}_i(t) - \tilde{\sigma}^{(k)}(t))^2} \quad (17)$$

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

Note as a general remark 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:

$$\mathbb{E}[\delta_i^{(F)}] < \mathbb{E}[\delta_i^{(R)}] < \mathbb{E}[\delta_i^{(S)}]. \quad (18)$$

Hence, for our purposes we heuristically assume $\delta_i^{(F)} \leq 1$, $1 < \delta_i^{(R)} \leq 2$ and $2 < \delta_i^{(S)} \leq 3$. If user i does not belong to any of these 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 (17).

In this way we were able to classify 39 out of the 40 relevant 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 s without later

Table 4 Classification results

2012 trial				
Classes	F	R	S	X
Number of users	8	11	15	5
Mean \pm standard deviation (SD) of $\delta_i^{(k)}$	0.69 \pm 0.16	1.58 \pm 0.28	2.58 \pm 0.36	8.30 \pm 2.91

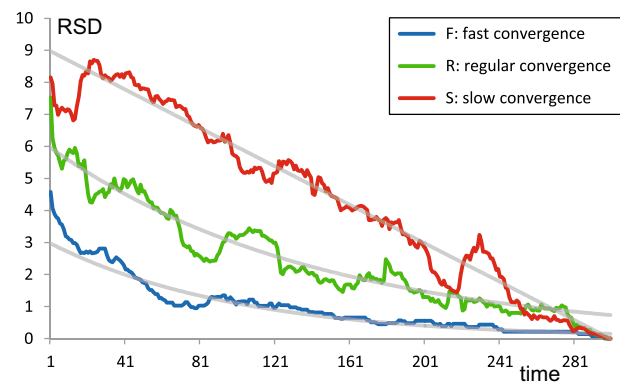


Fig. 15 User classes

Table 5 Absolute expenditures (in EUR and as % of p_{max})

	Median	Mean	Standard deviation (SD)
2012 trial	1.42 (42%)	1.52	0.95
2015 trial (Vienna) [18]	1.29 (43%)	1.48	0.87
2015 trial (Oulu) [18]	1.71 (71%)	1.90	0.86

changes, resulting in a significant average deviation from $\tilde{\sigma}^F(t)$. The distribution of the other 39 participants together with the corresponding mean and standard deviation values for $\delta_i^{(k)}$ is depicted in Table 4 (note that for class X , we have averaged mean and standard deviation over the differences to all three classes).

Figure 15 depicts the three RRMSDs described in (16) (grey lines) together with the average value of $\tilde{\sigma}_i(t)$ according to Eq. (15) 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 the mathematical analysis presented in “Fixed-point models for QoS- and QoE-based charging”, 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 S).

Willingness to pay and user interaction behavior

Having discussed user classification in some detail, we will now concentrate on monetary aspects of our trial results. First of all, observe that users’ willingness to pay unfolds

Table 6 Expenditures per round (in EUR)

	M_1	M_2	M_3
2012 trial			
Median	1.08	0.95	1.00
Standard deviation (SD)	0.79	0.75	0.94
2015 trial (Vienna) [18]			
Median	1.21	1.49	1.71
Standard deviation (SD)	0.73	0.76	1.07
2015 trial (Oulu) [18]			
Median	1.43	1.80	2.29
Standard deviation (SD)	0.68	0.69	1.05

an overall median spending over all test subjects and tariffs ranging between EUR 1.29 and EUR 1.71 for the respective trials—see details in Table 5.

While there were users spending the overall maximum (EUR 4.00) as well as others spending nothing, for further analysis we focus on two different aggregations:

Aggregation per round All movies shown at the i -th round ($i = 1, 2, 3$) are aggregated to measurement M_i , see Table 6. Note that, while overall expenditures do not exhibit significant differences, slightly lower spendings and SD values are observed for M_2 . Moreover, none of the subjects with tariff C have purchased Q19 during M_1 , while this has changed in subsequent rounds. Hence, we conclude that average purchasing behavior over time is considered rather stable, while individual behavior over time (from M_1 to M_3) may change.

Aggregation per tariff Studying each tariff independently, we observe that the median expenditure monotonically increases from EUR 0.74 to EUR 1.26 on the way from tariff A to C, indicating that absolute price increases raise absolute revenues. On the other hand, if expenditures are normalized to the unit interval (i.e., for each tariff the maximum possible expenditure is scaled to 1), resulting relative expenditures exhibit a (light) downwards slope.

Not unexpectedly, higher prices also increase the SD significantly: in the normalized data, fluctuations around 26% are observed, while a higher absolute SD suggests that higher prices trigger a broader reaction concerning possible pricing strategies. All in all, taking into account the non-linearity of absolute expenditure increase and the complex purchasing behaviors on a normalized scale, we conclude that the end user's willingness to pay does not increase linearly with respect to the tariff level, but exhibits a more complex dependency [35]. With the normalized data set, it turns out that around half of the test subjects decided to purchase quality levels priced between 20 and 40% of the respective maximum, while a significant number of users went for medium levels between 40 and 60% of the maximum. Local peaks are to be found also at 0 and 100%; in addition, we observe a non-negligible number of purchases for the virtual quality

levels (Q16–Q19), suggesting the existence of a noticeable premium segment where users seek for excellent qualities irrespectively of price. The broad majority, however, appears to be rather price-sensitive.

Complementing the convergence results presented in “End user convergence behaviour”, we may also consider aggregated user behavior over time depending on the eventually chosen quality/price level. To this end, we again distinguish *quality-sensitive users* (paying eventually more than an average user) from *price-sensitive users* (paying less than the average). Figures 16, 17 and 18 depict the aggregated convergence behavior of these two groups for the 2012 trial and the 2015 trials in Vienna and Oulu, respectively. Observe that, due to the complete reimplementations of the 2015 trial, there are some data differences to be acknowledged: the 2012 trial recorded the selected quality every second (i.e., inputs over time in seconds), while in the revised setup all inputs have been recorded (i.e., sequence of inputs).

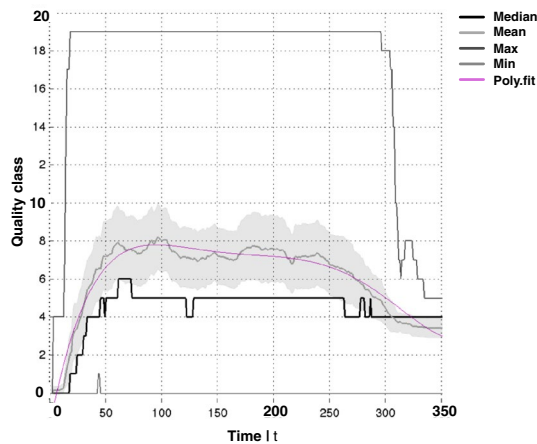
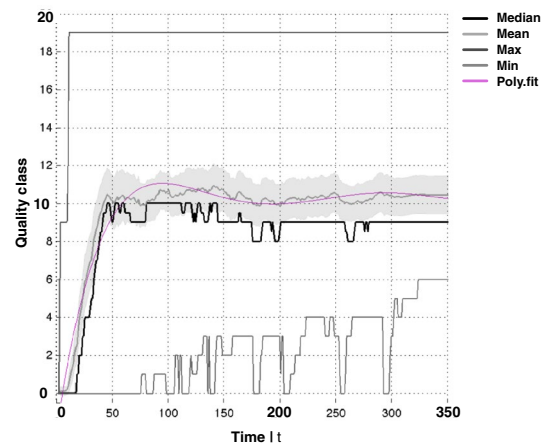
In most cases,⁸ we observe a relatively steep ascent during the first minute, where quality is the driver while users have to overcome the unacceptable initial quality condition Q0. Then, the price becomes the limiting factor and leads to an (aggregated) equilibrium state (while, of course, individual users still continue changing the quality levels, see Fig. 8b). The effect is especially pronounced for the price-sensitive users that quickly change to cheap and low-quality offers. Thus, we may conclude that end users are going for an optimal final quality level based their maximum willingness-to-pay, which can be interpreted as an active decision for an acceptable balance between price and quality, potentially involving cognitive dissonance [29]. For further details on this aggregated analysis for the case of the 2012 trial we refer to [35].

As another interesting side remark, observe that Table 7 shows how absolute expenditures have been significantly increasing along with maximum prices in all trials (i.e., from A to C). On the other hand, normalized expenditures slightly decrease in most cases (especially during the transition $A \rightarrow B$) before stabilizing ($B \rightarrow C$). This may have two reasons: either end users are not willing to accustom to lower quality levels, or effects of market entrance pricing are observed [35]. For instance, a low price in the first round may trigger lower subsequent expenditures (as users are avoiding luxury purchases), whereas high initial prices do not impact the users' purchasing behavior. However, if certain minimum quality expectations are about to be underrun, users avoid opting out (which the trial design does not allow anyhow). This illustrates different

⁸ The case of the Oulu testing is a slight exception, as the high interaction rate with the system paired with the limited sample size of such trials creates more noisy results for such assessments. However, similar trends can be observed also for this trial.

Table 7 Expenditures per tariff (in EUR)

	A: $p_{\max} = 2.00$	B: $p_{\max} = 3.00$	C: $p_{\max} = 4.00$
2012 trial			
Median	0.74	0.95	1.26
SD	0.54	0.82	0.99
Median (normalized)	0.37	0.32	0.32
SD (normalized)	0.26	0.28	0.25
2015 trial (Vienna) [18]			
Median	0.86	1.71	1.71
SD	0.43	0.80	1.16
Median (normalized)	0.43	0.57	0.43
SD (normalized)	0.22	0.27	0.29
2015 trial (Oulu) [18]			
Median	1.43	1.71	2.29
SD	0.43	0.80	1.16
Median (normalized)	0.71	0.57	0.57
SD (normalized)	0.17	0.25	0.27

**(a)** Price sensitive**(b)** Above median**Fig. 16** Aggregated temporal interaction behavior of users, 2012 trial

user mindsets with respect to price vs. quality sensitivity—see [35] for further details.

On an aggregated level, this confirms that subjects are both price and quality sensitive in video quality markets with explicitly shown price tags, which classifies a resulting representative aggregate user into the “QoE-based charging—submodel 2” (see section “[Fixed-point models for QoS- and QoE-based charging](#)”). However, the high standard deviation (e.g., see Table 6) and the previous findings in the 2011 user trial [28, 29] indicate that multiple customer segments exist, where the latter work finds quality-optimising, price-optimising and mixed quality-price-optimizing subjects, as well as strategically acting subjects seeking for bargains on the market.

Summary and conclusions

In this paper, we discuss several fundamental charging-related aspects of the current paradigm change from Quality of Service (QoS) to Quality of Experience (QoE), which significantly extends prior work on this topic. Based on a comprehensive analysis of context gaps, we have proposed and analyzed a fixed point model for QoE-based charging, whose specific complexity comes from the fact that it considers both price sensitivity and quality sensitivity of users. Here, the role of prices becomes twofold: they serve as expression for the received service quality, and at the same time they may significantly impact QoE

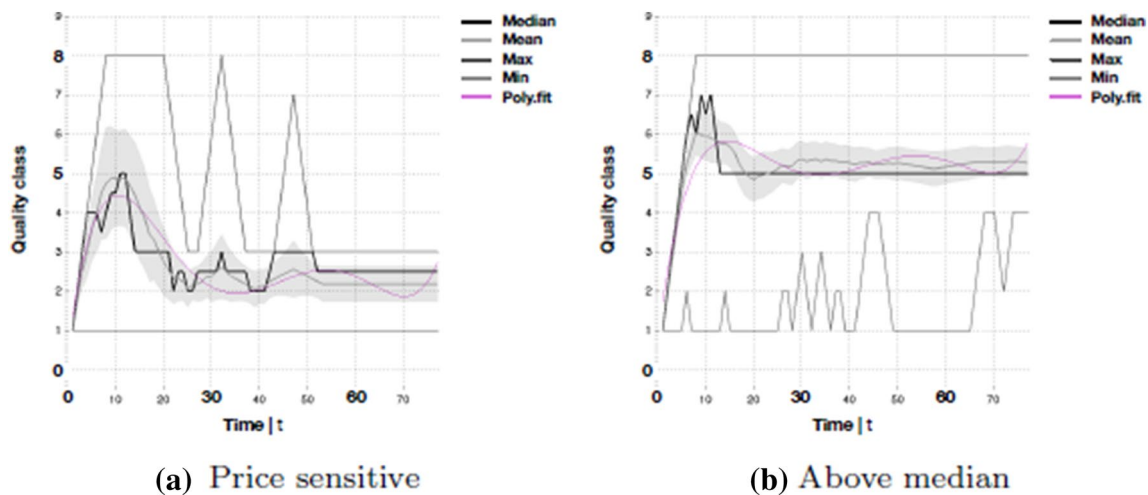


Fig. 17 Aggregated temporal interaction behavior of users, 2015 trial (Vienna)

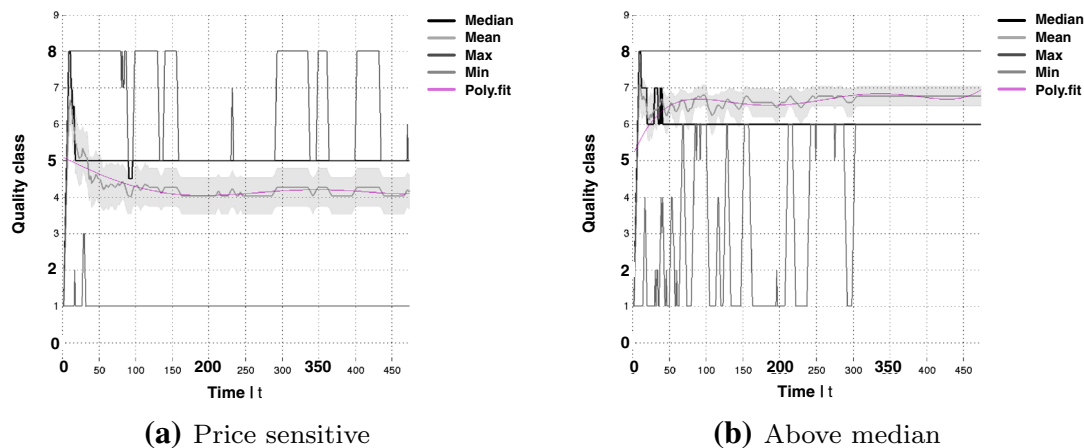


Fig. 18 Aggregated behavior of price-sensitive (left) and quality-sensitive (right) users, 2015 trial (Oulu)

evaluation. The applicability of this model is illustrated with results from a series of user trials of QoE evaluation and charging for video streaming. Our advanced trial setup allows users to choose in real time between a broad range of fine granular quality classes of an HD video subject to different price plans, and allows validating also our approach for user classification based to their convergence behaviour. Together with our quantitative investigation of user's willingness to pay (WTP), these empirical results confirm the necessity to consider price- and quality-sensitive user groups, and additionally suggest to separate charging for QoE from QoE perception itself.

Hence, the fixed-point analysis illustrates the difficulty of directly characterising the QoE market from QoE ratings that have been obtained outside of purchasing situations, while the empirical WTP data further support this perspective by

sketching a relationship for quality- and price-based network service differentiation that has no direct relationship to classical, for example, logarithmic QoE curves. This illustrates the remarkable influence of pricing on consumer decisions and their appreciation for network quality, which, however, also effects the communication to and with end customers.

This paper has aimed at laying grounds for various directions of further work. As far as the model is concerned, we are currently performing a detailed stability analysis of the resulting fixed points. Another important open issue concerns the parametrization of the various functions involved, which requires further highly focused user trials. Additional trials are also necessary to clarify the separability assumptions expressed in Eqs. (6) and (7), as well as our assumption on the consistency of user convergence behavior over time. Furthermore, the connection between charging and

perceived quality is obviously also relevant for a broad class of network-based services beyond video streaming, like, for instance, web access, VoIP, future cloud services etc. Finally, it would be very interesting to transfer our settings towards large-scale empirical studies, for instance in the framework of a dedicated field trial involving real customers and services. This would allow to overcome the methodological limitations of laboratory based experiments and to gain additional valuable insight to which extent users are indeed willing to pay for what they get in terms of quality experience.

Last, but not least, another open issue concerns the establishment of efficient frameworks for selling quality-differentiated network services to consumers. Historically, for Internet resources often the well-known concept of service level agreements (SLA) has been employed as a reference, which is however problematic for QoE products as discussed here [32], because around QoE (as the notion already suggests) only *experience products* [20] can be formed, i.e., products whose value can be appreciated by consumers only a posteriori. For this reason, the definition of new concepts may be required, for instance in the form of *Experience Level Agreements (ELA)* as initially discussed in [32], where the beneficial characteristics of SLAs meet the experience aspects required for the context of QoE products. So far, this idea has remained on a conceptual basis, and given the current lack of interest on the operator side in providing end-user SLAs, its relevance seems still somewhat limited at the moment, but further work into this direction might still become another important building block for a successful commercialization of QoE.

Acknowledgements Open access funding provided by University of Vienna. The authors would like to thank Andreas Sackl and Bruno Tuffin for helpful discussions and valuable input, Stefan Suetter and Toni Mäki for support with the trials, and the anonymous reviewers for their thoughtful comments which have helped to improve the paper significantly. On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bouch A, Sasse A (1999) It ain't what you charge, it's the way that you do it: a user perspective of network QoS and pricing. In: Proc. integrated network management (IM'99), pp 639–655. Boston, MA
- Courcoubetis C, Weber R (2003) Pricing communication networks: economics, technology and modelling. Wiley, New York
- DaSilva LA (2000) Pricing for QoS-enabled networks: a survey. IEEE Commun Surv 3(2):2–8
- Fu B, Staehle D, Kunzmann G, Steinbach E, Kellerer W (2013) QoE-aware priority marking and traffic management for H.264/SVC-based mobile video delivery. In: Proc. of eighth ACM workshop on performance monitoring and measurement of heterogeneous wireless and wired networks. ACM
- Hands D (ed) (2002) M3I user experiment results. Del. 15/2, EU FP5 Project M3I (IST 1999-11429)
- ISO/IEC: 23009-1:2014 (2014) Information technology—dynamic adaptive streaming over HTTP (DASH)—Part 1: media presentation description and segment formats
- ITU-T: Rec. G.107 (2011) The E-model: a computational model for use in transmission planning. International Telecommunications Union, Geneva
- ITU-T (1994) Definitions of terms related to quality of service. International Telecommunications Union, Geneva
- ITU-T (2008) Vocabulary for performance and quality of service, Amendment 2: new definitions for inclusion in recommendation ITU-T P.10/G.100. ITU, Geneva
- Jain R (2006) Internet 3.0: ten problems with current internet architecture and solutions for the next generation. In: IEEE/AFCEA MILCOM'06. Washington, DC, USA
- Keimel C, Redl A, Diepold K (2012) The TUM high-definition video data sets. In: Proc. QoMEX'12. IEEE
- Kilkki K (2008) Quality of experience in communication ecosystems. J Univ Comput Sci 14(5):615–624
- Laghari K, Crespi N, Connelly K (2012) Towards total quality of experience: a QoE model in communication ecosystem. IEEE Commun Mag 50(4):58–65
- Le Callet P, Möller S, Perkis A (eds) (2013) Qualinet white paper on definitions of quality of experience. European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Lausanne, Switzerland
- Le Sauze N et al (2010) ETICS: QoS-enabled interconnection for future internet services. In: Proc. future network and mobile summit. Florence, Italy
- Marshall A (1920) Principles of economics (revised edition). Macmillan, London
- De Moor K et al (2010) Proposed framework for evaluating quality of experience in a mobile, testbed-oriented living lab setting. Mob Netw Appl 15(3):378–391
- Mäki T, Zwickl P, Varela M (2016) Network quality differentiation: regional effects, market entrance, and empirical testability. In: Proc. IFIP networking
- Möller S, Raake A (eds) (2014) Quality of experience: advanced concepts, applications and methods. Springer, Berlin
- Nelson P (1970) Information and consumer behavior. J Polit Econ 78(2):311–329
- Reichl P (2010) From charging for quality-of-service to charging for quality-of-experience. Ann Telecommun 65(3):189–199
- Reichl P (2016) Quality of experience in convergent communication ecosystems. In: Lugmayr A, Dal Zotto C (eds) The media convergence handbook, vol 2. Springer, Berlin
- Reichl P, Tuffin B, Schatz R (2013) Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience. Telecommun Syst J 52(2):587–600
- Reichl P (2007) From “Quality-of-Service” and “Quality-of-Design” to “Quality-of-Experience”: a holistic view on future interactive telecommunication services. In: Proc. SoftCOM'07. Split, Croatia
- Reichl P, Egger S, Schatz R, D'Alconzo A (2010) The logarithmic nature of QoE and the role of the Weber-Fechner Law in QoE assessment. In: Proc. IEEE international conference on communications (ICC10). Cape Town, South Africa
- Reichl P, Egger-Lampl S, Möller S, Kilkki K, Fiedler M, Hossfeld T, Tsiaras C, Asrese A (2015) Towards a comprehensive

- framework for QoE and user behavior modelling. In: Proc. of QoMEX'15
27. Reichl P, Maillé P, Zwickl P, Sackl A (2013) A fixed-point model for QoE-based charging. In: Proc. ACM SIGCOMM workshop on future human-centric multimedia networking. Hong Kong, China
 28. Sackl A, Zwickl P, Egger S, Reichl P (2012) The QoE alchemy: turning quality into money. experiences with a refined methodology for the evaluation of willingness-to-pay for service quality. In: Proc QoMEX'12. Yarra Valley, Australia
 29. Sackl A, Zwickl P, Egger S, Reichl P (2012) The role of cognitive dissonance for QoE evaluation of multimedia services. In: Proc. quality of experience for multimedia communications workshop (QoEMC), IEEE Globecom, pp 1352–1356. Anaheim, CA
 30. Sen S, Joe-Wong C, Ha S, Chiang M (2013) Smart data pricing (SDP): economic solutions to network congestion. In: Haddadi H, Bonaventure O (eds) Recent advances in networking, pp 221–274. Licensed under a CC-BY-SA license
 31. Tuffin B (2003) Charging the internet without bandwidth reservation: an overview and bibliography of mathematical approaches. *J Inf Sci Eng* 19(5):765–786
 32. Varela M, Zwickl P, Reichl P, Xie M, Schulzrinne H (2015) Experience level agreements (ELA): the challenges of selling QoE to the user. In: Proc. of the ICC 2015 workshops, pp 1741–1746. IEEE. <https://doi.org/10.1109/ICCW.2015.7247432>
 33. Zwickl P (2016) End-to-end network quality: a utility-based approach. PhD Thesis, University of Vienna, Austria
 34. Zwickl P, Reichl P, Skorin-Kapov L, Dobrijevic O, Sackl A (2015) On the approximation of ISP and user utilities from quality of experience. In: Proc. QoMEX'15
 35. Zwickl P, Sackl A, Reichl P (2013) Market entrance, user interaction and willingness-to-pay: exploring fundamentals of QoE-based charging for VoD services. In: Proc. IEEE Globecom'13. Atlanta, GA, USA