No Silver Bullet: QoE Metrics, QoE Fairness, and User Diversity in the Context of QoE Management

Tobias Hoßfeld*, Poul E. Heegaard†, Lea Skorin-Kapov‡, Martín Varela§

*University of Duisburg-Essen, Modeling of Adaptive Systems, Email: tobias.hossfeld@uni-due.de

† Norwegian University of Science and Technology (NTNU), Email: poul.heegaard@item.ntnu.no

‡ University of Zagreb, Department of Electrical Engineering and Computing, Email: Lea.Skorin-Kapov@fer.hr

§ EXFO, Email: Martin.Varela@exfo.com

***Abstract*—Managing QoE is one of the most interesting direct applications of workable QoE models. Indeed, being able to predict how users perceive the quality of a service allows the service provider(s) to optimize its delivery, based on several possible criteria. It has been argued, however, that the MOS is ill-suited for this type of application, and that different measures — e.g., rating distributions or quantiles — are better suited for the task. In this paper we build on these ideas by adding the notions of QoE fairness (as opposed to QoS fairness) and user diversity, and discuss how the choice of measures used, the importance of fairness, and how the variations between users can affect the optimal QoE management choices for service providers.**

Keywords—Quality of Experience (QoE); Quality of Service (QoS); QoE metrics; QoE management; Fairness

I. Introduction

From a service provider's perspective, QoE is a means to an end: better quality leads to more satisfied users and possibly a differentiating factor against competing services, and it therefore leads to a larger user base and lower churn rates [1]. But providing higher QoE to users typically implies higher operational costs (e.g., in terms of resources used for service delivery). The goal for a service provider is, typically, to maximize their profits, and thus maximizing QoE for all users can be sub-optimal for the service provider's bottom line.

In practice, the goal will be to provide sufficient quality so that a suitably large proportion of users feel that the value they get from the service (utility) is commensurate to its cost (be it monetary or otherwise). It is also known that different users will perceive both quality and value differently [2], and that some users will be more critical than others [3]. The problem thus becomes one of finding the right trade off between the QoE provided and the costs incurred in providing it, taking into account the variations between users and a somewhat "fair" approach to provisioning quality among them, so that the service provider's goals become attainable.

Previous studies have argued that objective QoE models based on Mean Opinion Scores (MOS) may be ill-suited for QoE management purposes, as variations between users are averaged out [4], [5]. Thus considering a service provider point of view, other QoE metrics, e.g., rating distributions or quantiles, may provide more meaningful input [6]. Given a QoE optimization problem (e.g., deciding on an optimal resource allocation, service configuration, adaptive video playout strategy, etc.) considering both *quality* and *fairness* objectives, it remains unclear as to what are the implications of considering different QoE metrics (referring to the mapping between QoS and QoE) on the QoE management solution/outcome. Moreover, given the impact of *user diversity* on different QoE metrics, to what extent does user diversity impact the QoE management outcome for different chosen QoE metrics? In this paper, we address these issues and argue that there is no optimal approach to solving this problem in the general case. The right choice of QoE metrics to use will depend, not only on the service provider's constraints, but also on the diversity of the users, and the importance that the service provider assigns to being QoE-fair.

QoMEX2017 – Erfurt, Germany; 978-1-5386-4024-1/17/$31.00 ©2017 IEEE

The reminder of this paper is structured as follows. Section II gives a background of related work, while Section III discusses different QoE metrics relevant for the QoE management process. In Section IV we use a simple case study to illustrate the impact of utilizing different QoE metrics when determining an optimal resource allocation, thereby also considering fairness and user diversity. Section V highlights the main findings presented in this paper and gives directions for future work.

II. Background and Related Work

QoE management mechanisms commonly refer to some form of network or service management decisions which are driven by the aim of optimizing end user QoE [7], [1]. Examples include QoE-driven resource allocation mechanisms, service adaptation mechanisms, etc. These mechanisms are built on top of QoE models, which provide QoE estimates based on a set of underlying influence factors (system, user, and/or context factors). Focusing on system factors, such models are derived based on subjective user studies and map QoS measurements (at the network or application layer) to QoE values.

While the majority of studies to-date addressing QoE estimation models rely on MOS values derived from subjective data, other studies have shown that different metrics beyond MOS may be of potential interest to service/network providers [6]. Statistical measures providing insight into score distributions and quantiles give providers a clearer view of how quality is actually perceived by the user population, rather than estimating the quality as perceived by an "average" user. The authors in [4] argue that the arithmetic averaging inherent to the MOS may severely degrade the performance of QoE management by leading to unfairness among users, and thus propose a utility-based averaging of MOS values.

The mapping between QoS and QoE may be utilized for the purposes of benchmarking different QoE management algorithms, for monitoring user perceived quality, and finally for the development of "QoE optimal" management solutions. From a service provider's point of view, bounds imposed on the corresponding costs of optimizing QoE limit the solution space; e.g., a cost threshold will determine the maximum available bandwidth of the system (or the number of servers). Thus, given certain bounds, the objective is to distribute available resources in a QoE optimal way.

Given a system with multiple simultaneous users accessing shared resources, the aforementioned objective may be formulated in different ways, such as maximizing the average QoE, or maximizing the percentage of users rating above a certain threshold. The literature further advocates the need to consider ensuring fairness among users. Different approaches to solving this multi-objective QoE optimization problem may be considered, such as:

- A *two-step* approach whereby a solution is found maximizing average quality, followed by a second step to solve for maximum fairness while maintaining the previously determined average quality level.
- A *utility approach* where the optimization goals (such as cost minimization, average quality maximization, fairness maximization) are combined to derive a utility function.

Recent results show that a QoS fair system is not necessarily QoE fair, and have thus used various methods for calculating QoE fairness [8], [9], [10]. Such methods often rely on Jain's fairness index [11] or coefficient of variation to evaluate systems in terms of QoE fairness. A general QoE fairness index F has been defined by Hoßfeld et al. [12], which fulfills some desirable properties that are violated by Jain's index or coefficient of variation. For example, F is independent of the underlying QoE rating scale which is used to derive a mapping between QoS and QoE. A system is *absolutely QoE fair* ($F = 1$) when all users receive the same QoE value. The most unfair system leads to $F = 0$. We note that in this paper we will adhere to this definition of fairness.

The focus in the remainder of the paper is on investigating the implications of using different QoE models derived from various metrics on the QoE management decision, thereby considering the joint optimization of both quality and fairness.

III. Approach

In this section, we describe a general approach for using the results from subjective tests to estimate different QoE metrics under a specific QoS condition, as depicted in Figure 1.

A. Statistics from Subjective Tests

Under a specific QoS condition x (e.g., loss rate, loss pattern, throughput, delay) from a subjective test with k subjects we get a set of user ratings $\underline{y}_x = \{y_{x,1}, \cdots, y_{x,k}\}$.

From this we can obtain a QoS-QoE mapping function by assuming that the mean opinion score $\text{MOS}_x = \text{avg}(\underline{y}_x)$ for QoS condition x, represents the QoE metrics of interest. For a set of MOS estimates we can then do a curve fitting to a smooth QoE function, $Q_{\text{MOS},x}$. Similar mappings can be done for other metrics, e.g., QoE-QoS mapping of x to α-quantiles

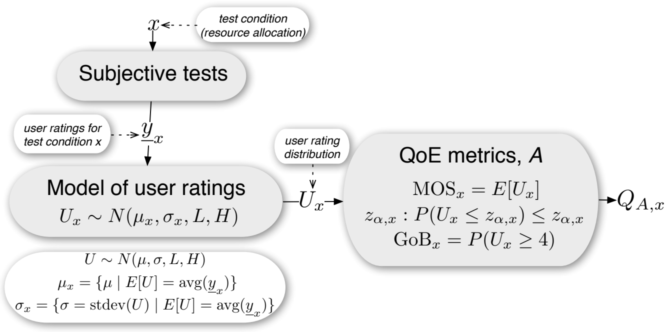


Fig. 1: QoE metrics from subjective tests and simulated user diversity

in the QoE ratings. Then the α-quantile for x, denoted $z_{\alpha,x}$, is the QoE metric of interest, and the mapping function is denoted $Q_{z_{\alpha,x},x}$

B. Modeling User Rating Distributions

The set $\underline{y}_x$ is an empirical user rating distribution for QoS test condition x over the k subjects. Since k is typically a small number, instead of using the empirical distribution, the user rating distribution can be modeled by use of a Truncated Normal distribution, $U \sim N(\mu, \sigma, L, H)$, which has been shown to be valid in [6], with probability density function:

$$f_U(u, \{\mu, \sigma, L, H\}) = \frac{\frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})}{\Phi(\frac{H-\mu}{\sigma}) - \Phi(\frac{L-\mu}{\sigma})}$$

where μ and σ are expected value and standard deviation in a Normal distribution, $N(\mu, \sigma^2)$, L and H are lower and upper limits of U, and ϕ and Φ are probability density function and cumulative probability function of a standardised Normal distribution, $N(0, 1)$, respectively. Thereby, the bounds of the normal distribution are caused by the bounds of the rating scale, e.g. 1 (lowest QoE) and 5 (highest QoE) on a typical 5-point scale.

The user rating for condition x is then a random variable $U_x \sim N(\mu_x, \sigma_x, L, H)$, where the parameters μ_x and σ_x are determined by:

$$\begin{aligned} \mu_x &= \{\mu \mid E[U] = \text{avg}(\underline{y}_x)\} \\ \sigma_x &= \{\sigma = \text{stdev}(U) \mid E[U] = \text{avg}(\underline{y}_x)\} \end{aligned}$$

This means that μ_x is the μ in the distribution of U such that the expected value of U equals the average value of the empirical distribution of user ratings. Similarly for σ.

The mean value μ is the MOS, while the σ expresses the user rating diversity. The SOS hypothesis [3] postulates thereby that the user diversity can be expressed by a single parameter a; then SOS is a function of a and the MOS. [13] derives a relation between this SOS parameter and the parameter of a truncated normal distribution. Thus, we can express the user diversity by means of σ. This allows us to simulate different user diversity behaviors by changing σ. Please note that also other models for subject rating behavior could be used instead [14].

C. QoE Metrics (from U_x)

From the modeled user rating random variable, U_x, we can obtain different QoE metrics A of x, denoted $Q_{A,x}$. Examples of A used in this paper are:

- $\text{MOS}_x = E[U_x]$ - Mean Opinion Score,
- $z_{\alpha,x} : P(U_x \le z_{\alpha,x}) \le \alpha$: this is the α%-quantile in the U_x distribution[1],
- $P(U_x \ge \theta)$ - denotes the θ-acceptability, i.e., the probability that the user rating is above a certain threshold,
- $\text{GoB}_x = P(U_x \ge 4)$ - the probability of a rating that is *Good or Better* on a 5-point rating scale[2].

Which QoE metric to use depends on what the operator considers to be most important, the average user, the most critical ones, the majority that are sensitive to changes in the delivered quality? Let us assume a user i is given resources x_i. What is that users' QoE? If we assume this is an average user, we could estimate QoE using the MOS metric. On the other hand, if we assume this is a "critical" user, we could use the 10%-quantile. We now take a second look at the QoE metrics defined above to better understand their meaning when using them in the context of QoE management:

- The average user's opinion is most important: using MOS_x quantifies QoE for QoS x for the average user.
- The most critical users' opinion are most important: e.g., 10% most critical, use $z_{0.10,x}$.
- The majority of users is most important, excluding those who are insensitive to quality changes, e.g., exclude 10% of the most happy but ignorant users, use $z_{0.90,x}$.
- The majority of users should be happy: use GoB_x or θ-acceptability. What *majority* means (probability threshold) and what *happy* means (4 or more general, θ) must be defined by the provider / operator.

D. Average QoE and QoE Fairness Among Users

For certain QoE metrics we can define the average over the users, and the fairness among them. Let $Q_{A,\underline{x}}$, be the QoE function for a specific QoE metric A under resource allocation $\underline{x} = \{x_1, \cdots, x_k\}$, where x_i is the resources allocated to user i, $(i = 1, \cdots, k)$. Then we obtain for QoE metric A:

$$\bar{Q}_{A,\underline{x}} = \frac{1}{k}\sum_{i=1}^{k} Q_{A,x_i} \quad \text{Average QoE metric over } k \text{ users}$$
$$F_{A,\underline{x}} = 1 - 2\frac{\sigma_{A,\underline{x}}}{H-L} \quad \text{QoE fairness over } k \text{ users [12]}$$

The QoE management decision aims at obtaining the optimal resource allocation $\underline{x}_o$ over the k users which maximises the average QoE and QoE fairness in $Q_{\underline{x}}$, i.e., the QoE metric A for resource allocation $\underline{x}$ over the k users. See Figure 2 for an illustration.

IV. Numerical Results: Resource Assignment for Two Video Users

We illustrate the impact of the different aforementioned QoE metrics on the user-centric evaluation of a concrete resource assignment. We assume an operator has some shared, but constrained resources (such as network bandwidth) which

[1]We use Q10 and Q90 for the 10%- and 90%-quantile.

[2]θ-acceptability with $\theta = 4$

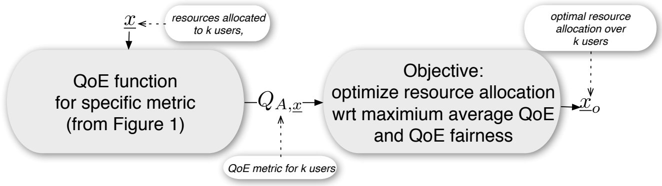


Fig. 2: QoE management maximises the resource allocation $\underline{x}$. Choosing different metrics for quality estimation ($Q_{A,\underline{x}}$) corresponds to optimizing the system for different types of users (avg. users, critical users, etc.) and thus leads to different optimal resource allocations.

are to be allocated to the users in the system (no over-provisioning).

The goal of the operator is in a first step to maximize the QoE per user and in a second step to provide QoE fairness across the users, cf. Section II. To this end, the operator relies on a QoE model which is fed with objectively measurable QoS parameters. The operator has thereby different options how to consider the users in the system, as discussed in Section III-C.

A. Two Users Scenario: Optimal QoE Resource Assignment

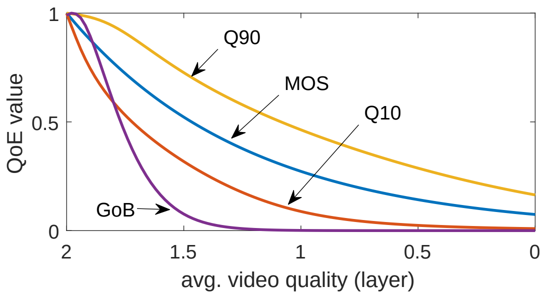
For illustration purposes, we consider the simplest scenario consisting of two users only. We argue that this scenario is, however, sufficient to demonstrate without loss of generality the impact of the QoE metrics and the operator's dilemma on how to assign resources to the users.

We consider HTTP adaptive streaming (HAS) as an example application. We use a simple QoE model from [15], derived from subjective experiments. The model maps the average HAS video quality layer (equivalent to the time on the highest layer) to QoE. Given that only MOS values are derived, we model the user rating distribution for a test condition with a truncated normal distribution so as to explore the potential impact of different user diversities, see Section III-B.

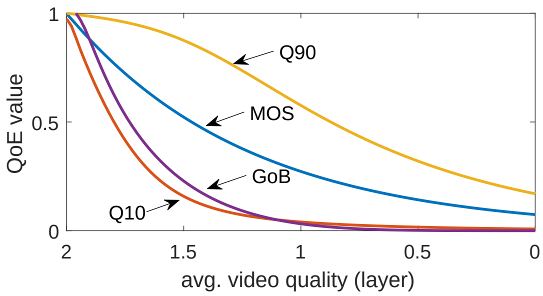
We must emphasize that the approach used here for simulating user diversities is supported by the literature [6]. It is based on real subjective data, and has been shown to lead to very good approximations. Furthermore, we highlight that the approach used to obtain different QoE metrics is irrelevant for our conclusions. Any set of realistic QoE models (e.g. based on objective metrics like PESQ or PSNR) could be substituted here to demonstrate the effects of choosing different metrics.

1) QoS-QoE Mapping Function for HAS: The simple QoS-MOS mapping function provided in [15] is an exponential function taking the relative time spent on the highest quality representation into a MOS value. To be more precise, only two video layers are considered in the subjective experiments (high quality $V = 2$ and low quality $V = 1$) and the relative time on the highest quality layer is equal to the average video quality. In the considered scenario in this paper, we also consider two layers which allows to use the experimentally derived QoS-MOS mapping from [15].

Based on the user rating model described in Section III-B, we can numerically derive additional QoE metrics such as quantiles and GoB for a certain test condition. Thereby, the test condition leads to an average video quality which we use in the result figures. We use Matlab's Piecewise Cubic Hermite



(a) User diversity $\sigma_N = 0.16$



(b) User diversity $\sigma_N = 0.40$

Fig. 3: QoS-QoE mapping — Matlab's Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) is used to fit the curves numerically.

Interpolating Polynomial (PCHIP) to fit the curves numerically. Thus, we get as QoE metrics: MOS, quantiles, and GoB as functions of the average video quality x (Figure 3). We see that more sensitive metrics such as 10%-quantile (denoted as Q10) and GoB portray a stronger decay of the QoS-QoE mapping function. We also note that GoB corresponds to the ratio of θ-satisfied customers. We keep the user diversity fixed (i.e., a certain σ_N) to the value which is derived from subjective studies in [6]. However, we can also simulate other user behavior by changing σ_N, which leads to a different set of QoE metrics, e.g., for the MOS $f_{\sigma_N}(x)$. We see the strongest impact of the user diversity on the GoB and Q10 curves (while clearly there is no impact on the MOS curve since diversity is averaged out). Please note that we use normalized QoE values with 1 indicating highest QoE and 0 indicating lowest QoE.

2) System Description: In the scenario considered, there are two users sharing network resources (a total capacity of C). The capacity is sufficient to serve one user with high quality ($V = 2$) and one user with low quality ($V = 1$) at the same time, but not sufficient enough to provide both users high quality. A simple QoE management solution would be to give both users low quality. This leads to a fair system, but the average QoE is not maximized.

We consider a better QoE management scheme which always utilizes the existing capacity completely. As a result, user 1 gets high quality with probability p. A video of length t is played in high quality for time tp for user 1. Accordingly, user 2 watches the video in high quality with probability $1 - p$. Thus, the resource management problem is reduced to *determining which value of p is appropriate according to the operator's goals.* From a QoS perspective (i.e. video quality), it does not matter how to assign p. The average video quality (over both users) is 1.5 (independent of p).

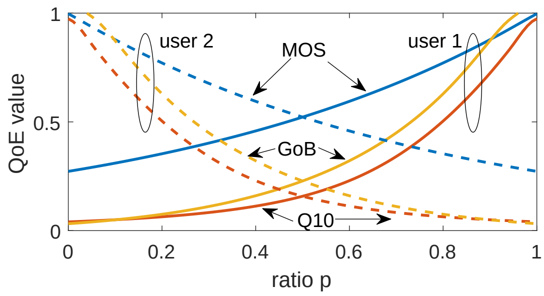


Fig. 4: Value of the selected QoE metric for $\sigma_N = 0.40$ and varying QoE management decisions (p). Solid lines represent the QoE value for user 1 for those QoE metrics; dashed lines represent user 2.

- Video quality of user 1: $V_1 = 2p + 1(1 - p)$
- Video quality of user 2: $V_2 = 2(1 - p) + 1p$
- Average video quality: $V = (V1 + V2)/2 = 1.5$

Thus, an optimal solution would be $p = 0.5$ which is QoS fair and leads to optimal QoS. However, $QoS \neq QoE$, and due to the non-linear relation between QoS and QoE, it is not clear what is a QoE optimal value of p. We discuss this with concrete numbers in the next sections.

B. Results: Average QoE per User

Figure 4 shows the value of the selected QoE metric for the two users; solid and dashed lines represent QoE for user 1 and 2, respectively. The ratio p (indicating the QoE management decision) is varied and different QoE metrics are considered. We exemplarily show the results for high user diversity ($\sigma_N = 0.40$). Figure 5 shows the average value of the selected QoE metric over both users. We see that the average QoE (in contrast to average QoS), depends on the resource assignment.

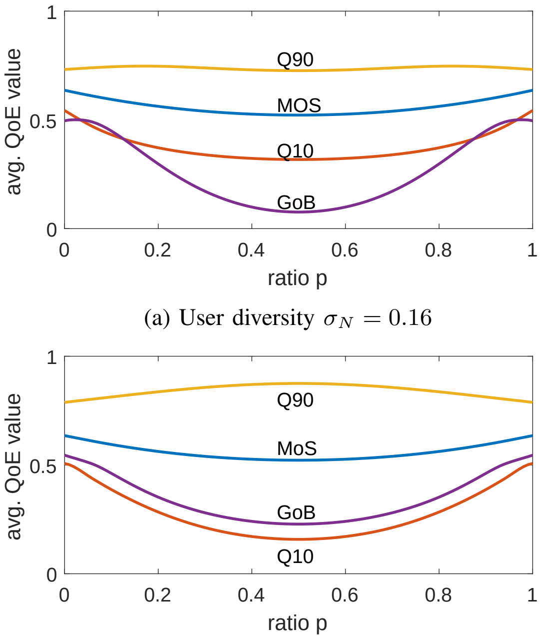
Now, let's take a closer look at the different QoE metrics. Critical users (10%-quantiles) will lead to lower QoE values than average users (MOS) or insensitive users (90%-quantiles). Thus, the selected QoE metric significantly changes how the operator sees the overall QoE in the system. The numerical results show how significant the changes can be. As expected, absolute QoE values are very different. We further see that the GoB is also a very sensitive QoE measure and allows the operator to clearly discriminate "bad" configurations.

We also note that user diversity has the strongest influence on GoB, followed by Q10. The higher the user diversity, the lower the Q10 curve (as a consequence of the mapping function). The Q90 curve is only slightly affected. Nevertheless, for Q90, the optimal resource assignment changes according to the user diversity (which is not the case for the other QoE metrics).

We observe that the optimal resource assignment p (which is determined by the maximum of the average QoE value) is achieved if we simply serve one user with the highest quality (except for Q90), i.e., $p = 0$ or $p = 1$! However, this clearly leads to unfairness.

C. Results: Fairness per User

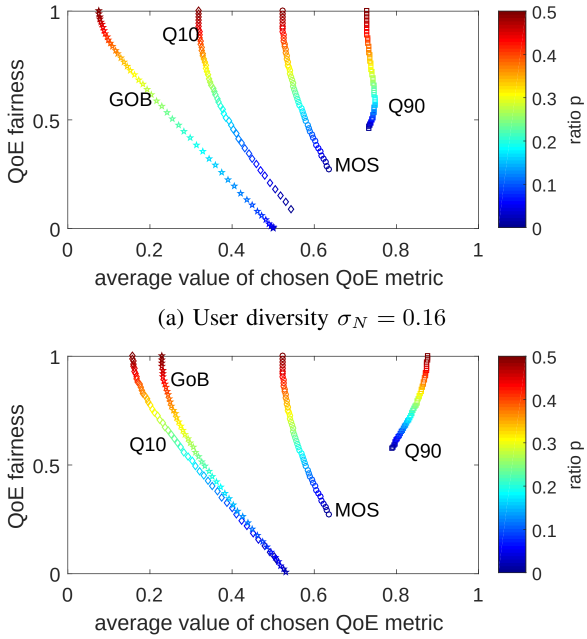
In addition to average QoE, we compute the standard deviation over the QoE values and derive the QoE fairness



(a) User diversity $\sigma_N = 0.16$

(b) User diversity $\sigma_N = 0.40$

Fig. 5: The average value of the selected QoE metric is plotted against the QoE management decision (p). It can be seen that the average MOS does not depend on user diversity.



(a) User diversity $\sigma_N = 0.16$

(b) User diversity $\sigma_N = 0.40$

Fig. 7: QoE management trade-off: average value of chosen QoE metric vs. QoE fairness depending on the resource assignment p, the selected QoE metric and the user diversity σ_N.

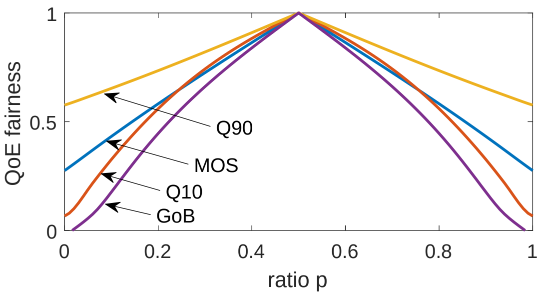


Fig. 6: QoE fairness of the system when using different QoE metrics in dependence of the QoE management decision p for $\sigma = 0.4$.

index F_x, as shown in Figure 6. Thereby, $F = 1$ means a perfectly fair system, while $F = 0$ describes the unfairest system [12]. Clearly, optimal fairness is achieved for p=0.5, where both users get the same quality – independent of the used QoE metric. However, we see strong differences between the selected QoE metrics if we assign one user 100% high quality and the other only low quality (i.e., p=0). This may lead to a completely unfair system ($F = 0$) when considering GoB or Q10, or a moderately fair system (with F around 0.5) when considering MOS or Q90. We observe also an overlap in the curves for a range of p from approx. 0.4 to 0.6. Thus, *the interpretation of the resource assignment for an operator strongly depends on the used QoE metric*. An operator has to decide the QoE metric for estimating average QoE and fairness. This reflects the operator's decision with regards to which user population to focus on: *average* users (reflected by using MOS as the QoE metric), *critical* users (using Q10), *θ-satisfied* users (using GoB).

Figure 7 depicts the trade-off between fairness and average QoE (depending on the assignment p). It combines the results from Figure 5 and Figure 6. Except for Q90, a higher average QoE decreases QoE fairness. Clearly, it remains up to an operator to determine the relative importance of fairness as compared to quality. In the case of Q90, the trade-off changes with a higher user diversity, and a clear optimum exists at p=0.5. Moreover, for low user diversity (σ=0.16), the average Q90 is almost independent of p (only fairness is affected).

D. Results: Maximum Utility

A straightforward approach to combine average QoE $\bar{Q}_A$ and fairness F_A into a single utility value U_A is as follows. A relevance factor ρ indicates the importance of fairness in the management decision. A indicates the selected QoE metric.

$$U_A = (1 - \rho)\bar{Q}_A + \rho F_A$$

The goal of the operator is then to maximize the utility (based on the decision of ρ). Figure 8 plots the optimal value of p (leading to maximum utility) depending on the fairness relevance ρ for the different QoE metrics.

The main observation is as follows. For Q90, using QoE fairness or overall QoE will give the same optimal solution. For the other three metrics there is a certain threshold, i.e., there are clear "break-even" points (optimal trade off between average QoE and QoE fairness) - these depend on the metric. Depending on ρ, the QoE metrics lead the operator to either $p = 0$ or $p = 0.5$. Different values than $p = 0, 0.5$ are caused by numerical inaccuracies (due to the numerical fitting of the QoS-QoE mapping functions). However, the critical point is at which "break-even" points the assignment changes. Those critical points change depending on the QoE metric.

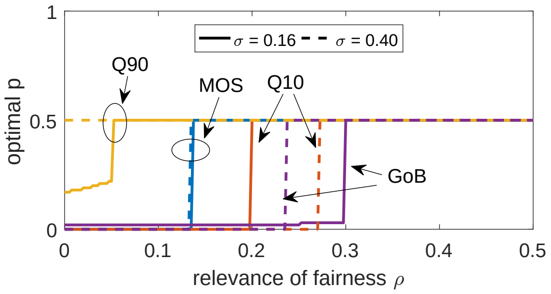


Fig. 8: The optimal resource assignment p leads to maximum utility U_A depending on the QoE metric A and the fairness relevance ρ. Solid lines indicate user diversity $\sigma_N = 0.16$; dashed lines represent $\sigma_N = 0.40$.

Q90 considers insensitive users which will always lead to higher QoE values (see average QoE for p=0, in Figure 5). Thus, the utility value is mainly determined by the fairness. Q10 and GoB focus more on maximizing average QoE (due to the stronger decay of the QoS-QoE mapping function in Figure 3) at the cost of fairness. Thus, only a higher fairness relevance ρ will change the assignment (from $p = 0$ to $p = 0.5$). MOS lies in-between Q90 and Q10, but does not allow to differentiate user diversity! We observe the same break-even point independent of the user diversity σ_N. On the other hand, we see that user diversity changes the ordering of the break-even point for Q10 and GoB.

Depending on the QoE metric, we can obtain significantly different conclusions for the operator with respect to which p value to use.

V. Conclusions and Future Work

QoE management mechanisms deployed either at the network or application layers inherently rely on underlying mapping functions between QoS and QoE when optimizing the system. Using an illustrative numerical example, we have shown that choosing different metrics for quality estimation corresponds to optimizing the system for different types of users (average users, critical users, etc.). We provide novel insights into how this in turn leads to different QoE management outcomes, such as optimal and QoE-fair resource allocations. Ultimately, the decision of which metrics to use may depend on the market situation and target user category, e.g., operators/providers focused on meeting the requirements of "innovative/early adopters" (potentially for marketing reasons) may choose a conservative approach and apply the Q10 or GoB metrics, whereas operators focused on meeting the requirements of "followers/laggers" may opt to use the Q90 metric. Deciding for a "majority" approach may lead to the decision to use the MOS metric.

A simple example is sufficient to derive the key observations and conclusions. Larger scenarios with more users, more complex user rating models, or other underlying subjective user ratings used to derive the QoE metrics will lead to the same observations.

A further important contribution of the paper is the analysis of the implications of different user score diversities on the QoE management outcome when applying different metrics. Previous studies have shown that for certain services, ratings are found to be more diverse than for others (e.g., scores are less diverse when rating impaired videos as opposed to Web page loading times [6]). While using MOS averages out user diversity, the most significant impact of diversity is reflected in GoB and Q10 curves.

In addition to understanding the meaning of using different QoE metrics for QoE management, another research question is related to how to formulate the actual QoE optimization problem with respect to multiple objectives such as quality, fairness, and costs. In future work, we plan to further investigate the impacts of different metrics and QoE optimization formulations on QoE management outcomes both at a theoretical and practical level.

References

[1] A. Ahmad, A. Floris, and L. Atzori, "QoE-aware service delivery: A joint-venture approach for content and network providers," in *Eighth Int. Conference on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal, Jun. 2016.

[2] V. A. Zeithaml, "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence," *The Journal of marketing*, vol. 52, no. 3, Jul. 1988.

[3] T. Hoßfeld, R. Schatz, and S. Egger, "SOS: The MOS is not enough!" in *Third Int. Workshop on Quality of Multimedia Experience (QoMEX)*, Mechelen, Belgium, Sep. 2011.

[4] J. Xu, L. Xing, A. Perkis, and Y. Jiang, "On the properties of mean opinion scores for quality of experience management," in *IEEE Int. Symposium on Multimedia (ISM)*, Dana Point, CA, Dec. 2011.

[5] R. C. Streijl, S. Winkler, and D. S. Hands, "Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives," *Multimedia Systems*, vol. 22, no. 2, Mar. 2016.

[6] T. Hoßfeld, P. E. Heegaard, M. Varela, and S. Möller, "QoE beyond the MOS: an in-depth look at QoE via better metrics and their relation to MOS," *Quality and User Experience*, vol. 1, no. 1, Sep. 2016.

[7] R. Schatz, M. Fiedler, and L. Skorin-Kapov, "QoE-based network and application management," in *Quality of Experience: Advanced Concepts, Applications, and Methods, T-Labs Series in Telecommunication Services*. Springer, Mar. 2014.

[8] P. Georgopoulos, Y. Elkhatib, M. Broadbent, M. Mu, and N. Race, "Towards network-wide QoE fairness using openflow-assisted adaptive video streaming," in *ACM SIGCOMM workshop on Future human-centric multimedia networking*, Hong Kong, Aug. 2013.

[9] A. Mansy, M. Fayed, and M. Ammar, "Network-layer fairness for adaptive video streams," in *IFIP Networking Conference (IFIP Networking)*, Toulouse, France, May 2015.

[10] S. Petrangeli, J. Famaey, M. Claeys, S. Latré, and F. De Turck, "Qoe-driven rate adaptation heuristic for fair adaptive video streaming," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 12, no. 2, Oct. 2015.

[11] W. R. H. Raj Jain, Dah-Ming Chiu, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," DEC Report, Tech. Rep. TR-301, Sep. 1984.

[12] T. Hoßfeld, L. Skorin-Kapov, P. Heegaard, and M. Varela, "Definition of QoE Fairness in Shared Systems," *IEEE Communications Letters*, vol. 21, no. 1, Jan. 2017.

[13] T. Hoßfeld, M. Fiedler, and J. Gustafsson, "Betas: Deriving quantiles from mos-qos relations of iqx models for qoe management," in *IFIP/IEEE International Workshop on Quality of Experience Management (QoE-Management 2017)*, Lisbon, Portugal, May 2017.

[14] L. Janowski and M. Pinson, "The accuracy of subjects in a quality experiment: A theoretical subject model," *IEEE Transactions on Multimedia*, vol. 17, no. 12, Dec. 2015.

[15] T. Hoßfeld, M. Seufert, C. Sieber, and T. Zinner, "Assessing effect sizes of influence factors towards a QoE model for HTTP adaptive streaming," in *Sixth Int. Workshop on Quality of Multimedia Experience (QoMEX)*, Singapore, Sep. 2014.