Quantifying the Impact of Network Bandwidth Fluctuations and Outages on Web QoE

Andreas Sackl, Pedro Casas, Raimund Schatz, Lucjan Janowski
FTW - Telecommunications Research Center Vienna
{surname}@ftw.at

Ralf Irmer
Vodafone Group Research & Development
{name.surname}@vodafone.com

Abstract—The systematic investigation of network bandwidth fluctuations and their impact on Quality of Experience (QoE) has recently gained center stage, not only in academia but also among network providers. This development is driven by the fact that in a real world usage scenario, volatile network performance (and thus varying QoE) actually is the norm, particularly in mobile contexts. However, we are still witnessing a lack of validated fluctuation models and applicable performance indicators to represent the QoE impact of network quality variations in provider-relevant network planning & management processes. In addition, fundamental phenomena like the impact of single bandwidth outage events on Web QoE (which occur e.g. during a handover between different base stations) have not been studied yet. In this paper we present the results of an empirical Web-QoE user study which a) lay the path towards the analysis of different bandwidth fluctuation models and b) provide insights into the impact of a single outage (i.e., zero network bandwidth) events on subjective quality perception. Furthermore, our experimental setup that exerts short & repetitive testing-conditions provides methodological guidance on how to perform subjective quality testing in the context of fluctuations and Web QoE.

Keywords—Bandwidth Fluctuations; Bandwidth Outages; Quality of Experience; Quality Fluctuations; Web QoE.

I. INTRODUCTION

The scope of Quality of Experience (QoE) research is currently expanding from the investigation of constant quality conditions for various application types (e.g. Video [1], Voice [2]) towards the inclusion of studying the effects of time-varying quality on QoE. A prime example for this development is the introduction of adaptive video streaming which triggered the investigation of the QoE impact of dynamic video quality switches and related trade-offs [3]. [11]. To a large extent, this interest in time-varying quality is driven by the rapidly increasing adoption of mobile devices and the resulting proliferation of application usage over wireless networks where dynamically changing quality conditions are actually the norm (due to cell overload, user mobility, interference, occlusion, etc.). In this field, it is not only critical to assess and understand the QoE impact of network performance volatility, but also to appropriately model this kind of influence factor. However, while there exists already extensive work on time-varying quality for Voice and Video QoE (cf. [6], [9]–[12]), research addressing the impact of quality fluctuations in the context of Web QoE1 is still rare, albeit Web applications constitute the major share of mobile Internet user time (cf. [4]). Although in this field the influence of varying network throughput has been addressed in e.g. [4] to some extent, so far only repetitive fluctuation patterns have been investigated. Nevertheless, a large number real-world usage scenarios mostly feature the occurrence of single bandwidth outage events. For example, a handover between two radio cells during a train ride results in a bandwidth outage which impairs the usage experience of a passenger browsing the Web. In addition, models that capture the impact of more general bandwidth fluctuation patterns on Web QoE are still missing.

Given this gap in existing research, we study the impact of network bandwidth fluctuations on Web QoE, particularly focusing on single bandwidth outage (i.e., zero bandwidth) events. This paper presents the results and findings of a subjective Web browsing QoE lab study featuring bandwidth fluctuations as controlled variable. For the case of bandwidth outages, we quantify how annoying are such outages to users and how this relates to conscious perception of outages.

To capture the QoE-relevant impact of bandwidth fluctuations in a more general scenario, we propose to give some first steps in the modeling of such fluctuations. The main goal of such models is to define a suitable technique to measure and report fluctuations which could be used in the practice for network quality assessment. The target is to define new Key Performance Indicators (KPIs) which are actually able to capture the impact of bandwidth fluctuations on the user experience. Such new metrics must be able to describe those characteristics of bandwidth fluctuations — which are relevant to each specific application — but must not be too complex, as they would only be valuable if they could be understood intuitively. Indeed, any new metrics and KPIs must be readily interpretable in order to make them useful and applicable.

The remainder of this paper is structured as follows: Section II discusses related work in the area of time-varying QoE, focusing on bandwidth outages and bandwidth fluctuation modeling for QoE assessment. Section III presents the design and setup of the subjective Web QoE lab study, including a description of the evaluated fluctuation models and the methodology we used for their evaluation. The analysis of our experimental lab test results is presented in Section IV. Finally, we derive conclusions from our empirical findings and suggest necessary future research directions in Section V.

---

1The term Web QoE refers to “Quality of Experience of interactive services that are based on the HTTP protocol and accessed via a browser” [14].
Nevertheless, the more realistic use case of pooling methods that integrate momentary QoE values over practice, specially for ISPs, as rather than knowing how good papers generally consider the problem of how to integrate pattern shedding some light on the subject [13].

Outage patterns little has been done so far, and only some related papers with a mean age of 34.7 years (median=31y) participated in our experiment. Around half of the participants were students and almost 43% were employees, and 70% of the participants have completed university or baccalaureate studies. All tests were conducted on laptops connected to the Internet through a fixed-line network.

II. RELATED WORK

The analysis of time-varying network QoS features and their impact on QoE has been the objective of previous papers addressing voice [6] and video services [7]–[10], [12]. These papers generally consider the problem of how to integrate short-term or instantaneous quality assessments to provide an estimation of the quality experienced in a long-term and more global perspective. This problem is particularly interesting in practice, specially for ISPs, as rather than knowing how good is the experience of a user with a single snapshot of the network QoS, they are interested in the overall experience of the user when consuming a specific service. However, the directions pushed forward in these papers are still on the theoretical modeling of how to integrate multiple short-term QoE indications into a more general user experience metric.

Indeed, the approaches proposed so far apply certain temporal pooling methods that integrate momentary QoE values over time. Such temporal pooling methods are gaining increased interest for QoE modeling in more practical setups, including recent video services such as adaptive video streaming [11].

The impact of bandwidth outages on Web QoE have been recently studied in [4], where we applied various bandwidth outage patterns while the test participants browsed several Web sites. The authors show that interactive applications like Google Maps are heavily impaired by outages, whereas outages in low-interactive scenarios are tolerated by the users. Nevertheless, the more realistic use case of single bandwidth outage events are not discussed in the referenced paper. For example, in the use case of using mobile Internet during train rides, network handovers or tunnel passings may result in single outage events which impact the user experience.

When it comes to the specific problem of modeling bandwidth fluctuation patterns and their impact on QoE little has been done so far, and only some related papers shed some light on the subject [13].

III. EXPERIMENTAL SETUP & DESIGN

To investigate the influence of single bandwidth outage events on Web QoE and to verify our bandwidth fluctuation models, we conducted a user study in 2014 in our subjective testing lab at FTW. Overall, 29 users (13 female, 16 male) with a mean age of 34.7 years (median=31y) participated in our experiment. Around half of the participants were students and almost 43% were employees, and 70% of the participants have completed university or baccalaureate studies. All tests were conducted on laptops connected to the Internet through a fixed-line network.

A. Setup and Approach for Single Outage Events Analysis

For the evaluation of the single outage events on Web QoE, test participants had to browse a Photo Gallery Site (thumbnail overview plus larger single photo views) and a News Site (www.cnn.com) with 18 Mbps downlink bandwidth and 10 Mbps uplink bandwidth. Additionally, single outages with various durations were applied (0, 2, 4, 8 & 16 sec.) and occurred 10-20 seconds (randomized) after condition start. These bandwidth patterns were imposed by routing the participants’ traffic through a customized version of the very well known NetEm network emulator [17].

In Web QoE studies, the duration of a test condition (after which the participant evaluates the QoE) typically ranges from 1.5 min to 2 min. For example, in [4], we applied a condition duration of 2 minutes, similar to e.g. [14] or [16]. In our experiments we used conditions with a duration of only 30 sec. for two reasons: 1) outages occurred only once per condition and 2) because of the short condition-length, we were able to apply each condition three times: even if a participant missed an outage (e.g. a 4 sec. outage occurs while she looks at a picture), there was another chance that she recognizes the 4 sec. outage in one of the repetitions.

After each condition, participants had to evaluate the annoyance caused by the outages by means of a 5 point ACR scale, ranging from 5 (=not annoying) to 1 (=very annoying) [15]. Additionally, we asked a binary yes/no question if the participant would accept the overall quality just experienced in an everyday usage scenario (i.e. whether the participant would in real life terminate the session or not). Furthermore, we asked our participants whether they recognized an outage or not (binary yes/no).

B. Setup and Approach for Fluctuation Models Analysis

For the modeling and analysis of bandwidth fluctuations, test participants evaluated the overall quality for three different fluctuation patterns applied to the Downlink Bandwidth (DBW) of the connection, considering the usage of three different applications: browsing the aforementioned News Site, uploading a large file, and exploring different city maps using the Google Maps application (Gmaps from now on). We

Figure 1. Bandwidth fluctuation patterns evaluated in the study: (a) pattern P1 considers a progressive outage/disconnection and a subsequent recovery; (b) pattern P2 represents a fast bandwidth changing environment; (c) pattern P3 depicts a high/low bandwidth profile with fast short-scale variations. (DBW = Downlink Bandwidth; ADBW = Average Downlink Bandwidth).
focus exclusively on the downlink direction for two reasons: most applications are still generating the bulk traffic in the downlink direction, and we can not address all the potential QoS parameters in single lab tests, which are time-limited to avoid participants fatigue. Studying the impact of other QoS parameters such as Uplink Bandwidth or Access Latency is part of our future work.

For the sake of brevity we report only those results obtained for the Gmaps scenario, which represents the most interesting one in terms of network bandwidth requirements. Fig. 1 depicts the three tested fluctuation patterns, which partially mimic those patterns observed in real measurements we conducted in cellular networks on train and walking scenarios. Pattern $P_1$, depicted in Fig. 1(a), considers a progressive outage/disconnection and a subsequent recovery; pattern $P_2$, depicted in Fig. 1(b), represents a fast bandwidth changing environment; finally, pattern $P_3$, depicted in Fig. 1(c), represents a high/low bandwidth profile with fast short-scale variations. Pattern $P_1$ is most likely to occur in situations where network coverage gaps arise, for example when going through tunnels or low signal-strength regions. Patterns $P_2$ and $P_3$ are characteristic of low-motion scenarios (e.g., walking) in regions where cells are heavy loaded with multiple devices and/or with multiple interference sources (e.g., trees, buildings, etc.).

For each pattern, five different fluctuations’ models were proposed and tested. Given that the potential number of features representing a fluctuation pattern is a-priori large and unknown (i.e., the number and duration of outages, the drop of bandwidth below certain thresholds and its duration, the share of time with a bandwidth above certain threshold, etc.), the empirical evaluation of the goodness of any of these models is not a simple task. We therefore devised a practical evaluation approach, which tries to define a simple KPI able to capture the impact of a bandwidth fluctuation pattern on the QoE declared by the test participants.

The idea is as follows: in a traditional constant-bandwidth–QoE modeling approach, participants rate the QoE of some application at each constant bandwidth condition (w.l.o.g., let us consider the downlink bandwidth from now on), and a simple model, which maps bandwidth to QoE is constructed, e.g., following a simple curve-fitting approach. Let us refer to this model as $\text{MOS} = \mathcal{F}(\text{DBW}_\text{const})$. To use this model in the practice, one would measure the Average Downlink Throughput (ADT) of the corresponding application under certain fluctuation pattern $P_j$ and map it to a QoE value $\text{MOS}_{\text{ADT},i} = \mathcal{F}(\text{ADT}_j)$. What we propose is to construct an Effective ADT (EADT) KPI using the proposed models, which shall better reflect the impact of a certain bandwidth fluctuation pattern on QoE through the mapping function $\mathcal{F}(\cdot)$. Given the subjective QoE evaluation $\text{MOS}_{P_j}$ of an application under a certain fluctuation pattern $P_j$, we say that a certain model $i$ is a good model if the estimated QoE $\text{MOS}_{\text{EADT},i,j} = \mathcal{F}(\text{EADT}_{i,j})$ is close to $\text{MOS}_{P_j}$.

Fig. 2 describes this approach as used in the lab study, consisting of three steps: (1) in the first step, we construct (for each of the tested applications) the MOS = $\mathcal{F}(\text{DBW}_\text{const})$ mapping function, by asking participants to rate each DBW(condition; (2) in the second step, participants rate the QoE MOS$_{P_j}$ for each of the three tested patterns $P_j$ depicted in Fig. 1; (3) in the third step, the EADT$_{i,j}$ KPI is computed from each model $i$ used over pattern $P_j$, and the estimated QoE MOS$_{\text{EADT},i,j}$ is compared to MOS$_{P_j}$. For the sake of completeness, we additionally compare the QoE values to the estimation obtained by a direct ADT mapping for each pattern, i.e., MOS$_{\text{ADT},i}$.

Before describing the five tested models, we briefly explain how to construct the EADT KPI. As we said before, one of the main goals of such KPI is to keep a practical simplicity, fostering as such its clear understanding and potential adoption. Therefore, given a certain pattern $P_j$, the ADT$_{i,j}$ is constructed by weighting the ADT$_j$ by a Correcting Factor $CF_i$, derived from the fluctuations model $i$. As such, we simply compute $\text{EADT}_{i,j} = CF_i \times \text{ADT}_j$.

Let us now describe the five very first models we tested in our work to model bandwidth fluctuations and their effects on QoE. Since building such models is a very challenging and complex problem, the aim of this work is to discuss some first steps towards better capturing the impact of bandwidth fluctuations on QoE. Fig. 3 depicts the proposed models. Their target is to capture some of the salient features of fluctuations which have a direct impact on QoE. In particular, the total share of time in which the bandwidth drops below a certain per-application bandwidth threshold, and the “deepness” (i.e., how much below the bandwidth threshold) of such bandwidth drops are considered by these models.

1) LTD Model: Low-Throughput Duration: Fig. 3(a) describes the ideas behind the Low Throughput Duration (LTD) model. Given a certain throughput measure time span $t_0$, the LTD $CF$ accounts for the fraction of time that the throughput is below a certain threshold $\text{DBW}_T$. In the example depicted in Fig. 3(a), the LTD $CF$ is defined as $LTD = 1 - \frac{t_1}{t_0}$. The LTD is the simplest $CF$ to produce out of the bandwidth fluctuation pattern. It accounts for the fraction of time that the instantaneous throughput is below a single, constant bandwidth threshold, but it does not reflect the tolerance of some applications to short outages or bandwidth drops.

2) SLTD Model: Selective Low-Throughput Duration: The Selective LTD (SLTD) model is very similar to the LTD one, but it assumes that short time bandwidth drops are not perceived by the user. Fig. 3(b) explains the idea. Here $t_{ref}$ is a maximum bandwidth-drop duration tolerated by a specific application or application type. Therefore, the calculation of the SLTD $CF$ includes $t_1$, but excludes the other time bins: $\text{SLTD} = 1 - \frac{t_1}{t_0}$. If throughput varies at a high frequency, it is possible
A. Analysis of Single Outages

In this section we present the results of the subjective tests regarding the analysis of single bandwidth outages for Web QoE and the main findings regarding the evaluation of the proposed fluctuation models (for the specific case of Gmaps).

1) LTD Model: Limit Time Duration: The LTD model is similar to the LTD one, but considers two different bandwidth thresholds $DBW_{Th}$ and $DBW_{Th2}$. The threshold $DBW_{Th2}$ permits to account for critically large throughput drops, such as outages.

2) SLTD Model: Sliding LTD Model: The SLTD model is based on the LTD principle but uses a sliding time window $t_{ref}$ to adjust the threshold $DBW_{Th}$. In the example of Fig. 3(c), the $MADBW_{Th}$ threshold is computed through a Sliding Window Length (SLW) of 5 seconds long.

3) TJ Model: Throughput Jitter: The JT model follows the same principle of the LTD one, but instead of considering a fixed bandwidth threshold, it uses a moving average-based threshold $MADBW_{Th}$. In the example of Fig. 3(c), the $MADBW_{Th}$ threshold is computed through a Sliding Window Length (SLW) of 5 seconds long.

4) AREA Model: Area-based Model: The AREA model does not only consider the time below $DBW_{Th}$, but also partially accounts for how deep is the corresponding throughput gap between $DBW_{Th}$ and the instantaneous throughput. In Fig. 3(d), the AREA $CF$ is computed as $AREA = 1 - \frac{A_{total}}{A_{total}}$, where $A_{total} = t_{0} \times DBW_{Th}$.

5) DOUBLE Model: Double Threshold: The DOUBLE model is also similar to the LTD one, but considers two different bandwidth thresholds $DBW_{Th1} > DBW_{Th2}$. The threshold $DBW_{Th2}$ permits to account for critically large throughput drops, such as outages.

IV. EXPERIMENTAL RESULTS

In this section we present the results of the subjective tests regarding the analysis of single bandwidth outages for Web QoE and the main findings regarding the evaluation of the proposed fluctuation models (for the specific case of Gmaps).

A. Analysis of Single Outages

Fig. 4 presents the users’ annoyance in relation to the duration of the outages while the specific Web site was used. For browsing a Photo Gallery Site, outages of 2 or more

Table I. RELATIONSHIP BETWEEN OUTAGE DURATION AND RATINGS.

<table>
<thead>
<tr>
<th>Expectation Type</th>
<th>Single Outage Annoyance Model</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photo Gallery Site</td>
<td>$MOS_{Annoy} = 4.81344 \times e^{-0.02733 Dur_{Out}}$</td>
<td>0.99</td>
</tr>
<tr>
<td>News Site</td>
<td>$MOS_{Annoy} = 4.56955 \times e^{-0.02733 Dur_{Out}}$</td>
<td>0.97</td>
</tr>
</tbody>
</table>

In Fig. 5, the acceptance rate (“Would you accept this quality at home?”, yes/no) and the outage detection rate (“Have you recognized any outage?” – yes/no) is displayed for the task browsing a News Site and a Photo Gallery Site. The acceptance rate (black line) is close to 100% for both tasks for outage durations less or equal 4 seconds. For longer outages (8 & 16 seconds), the acceptance rate drops but it is still relatively high even for 16-seconds outages (approx. 50%). The acceptance rate progress is similar for both tasks, whereas the outage detection rate differ (grey line in Fig. 5): outages are more easily detected while the users browse a Photo Gallery Site in contrast to browsing a News Site e.g., outages with a duration of 4 seconds are detected by 25% of the users during browsing a News Site, whereas 55% of the users detect an outage while browsing a Photo Gallery Site. Due its visual complexity and resulting loading behaviour, 10% of the users who browse the News Site false-detects a non-existent outage (Fig. 5, grey circle on x-axis position ‘none’).

Fig. 6 depicts the influence of outage detection on quality ratings. MOS values indicated by yellow bars include only ratings from users who claimed not perceiving an outage. Obviously, these MOS values are close to 5 (=not annoying). In contrast to this, light green bars represent MOS values from users perceiving an outage. The influence of light green values (=outage detected) depends on the outage detection rate e.g., if an outage is not applied, the difference between the yellow and the dark green values is small because the
light green values only represent roughly 10% of the ratings (see Fig. 5). The significant differences between the light green bars (outages detected by users) and the dark green bars (outage detection independent) demonstrate that focusing on pure outage annoyance ratings is not sufficient; outages are not always perceived (probability increases with outage duration length, see Fig. 5), but users who detected them are definitely impaired. In contrast to previous studies, each outage duration was applied three times per user instead of single testing, see Section III. In Fig. 4, the presented MOS values include all three iterations, whereas Fig. 7 depicts iteration-dependent values. There is no significant difference between the iterations, i.e., there is no primacy/recency-effect regarding annoyance of single outages. For browsing a Photo Gallery Site, there is a little trend that outages get more and more annoying i.e. the ratings of the third iteration are little bit lower compared to the ratings of the first iteration, but the difference is rather small and within the confidence intervals. Hence, from a methodological point of view the repeated evaluation of single outage events in Web scenarios lead to consistent results. In [5], authors examined the influence of initial delay and stallings on video and music streaming QoE. In contrast to our Web QoE results, the occurrence of a single stalling event during music/video-streaming has a severe QoE impact: a single outage event during video streaming leads to a MOS ≈ 2.5 and for music streaming it leads to a MOS ≈ 2.

### Table II. Model Parametrization.

<table>
<thead>
<tr>
<th>Param.</th>
<th>LTD</th>
<th>SLTD</th>
<th>JT</th>
<th>AREA</th>
<th>DOUBLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBW_{T_h,1}</td>
<td>ADT (\frac{5}{2})</td>
<td>ADT (\frac{5}{3})</td>
<td>ADT (\frac{3}{4})</td>
<td>ADT (\frac{3}{5})</td>
<td></td>
</tr>
<tr>
<td>DBW_{T_h,2}</td>
<td>-</td>
<td>-</td>
<td>5 secs.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SWL</td>
<td>-</td>
<td>-</td>
<td>3 secs.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>\text{i}_{ref}</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### B. Fluctuation Models Evaluation

Fig. 8 depicts the overall quality reported by participants in the Gmaps tests for (a) different constant DBW values (2, 4, 8, and 16 Mbps) and (b) the three tested fluctuation patterns P1, P2 and P3 depicted in Fig. 1. Surprisingly enough, the QoE values obtained in all cases are rather low, resulting for example in a MOS ≈ 3.6 for the best tested condition, corresponding to a constant DBW = 16 Mbps. A deeper analysis shows that the desktop/laptop version of Gmaps has very high DBW requirements when using the full satellite view as it was the case in the tests.

Following the evaluation approach described in Fig. 2, the first step consists of building the mapping function $\text{MOS} = \mathcal{F}(\text{DBW})$ from the results obtained in the DBW constant tests (cf. Fig. 8(a)). Not surprisingly, the resulting fitting curve has a logarithmic shape of the form $\text{MOS} = 0.45 + \log(\text{DBW}) + 2.48$.

In the second step, the ADT, is computed for each pattern $P_j$: the resulting values are 5.9 Mbps, 5.1 Mbps and 4.3 Mbps for ADT, ADT, and ADT, respectively. Note that, for the sake of simplification, we are not directly measuring the average downlink throughput experienced by the Gmaps traffic flows under each fluctuation pattern, but rather the average downlink bandwidth of each pattern. The QoE subjective evaluations of each pattern MOS, are directly the ones depicted in Fig. 1(b). The resulting MOS scores are 2.9, 3.4 and 3.1 respectively for MOS, MOS, and MOS.

A first interesting observation from the subjective results is that the impact of short outages (pattern $P_1$) is relevant as compared to low bandwidth drops (pattern $P_3$), as MOS, even if $\text{ADT}_1 = 5.9$ Mbps and $\text{ADT}_3 = 4.3$ Mbps. Also interesting is the fact that fast DBW variations (pattern $P_2$) are only slightly perceived by users, as MOS, MOS, and MOS, = 3.4 and MOS, = $\mathcal{F}(5.2)$ = 3.2, which is almost the same.

To compute the correcting factor $CF_i$ for each of the tested models, we first need to specify the value of the different parameters used by each model. Table II reports the parametrization values selected for each model. Note in particular that we tested two different $\text{DBW}_{T_h}$ threshold for the LTD, SLTD and AREA models, corresponding to $\text{ADT}_j$ and $\text{ADT}_j/2$ respectively. This parameters’ selection is based both on the specific shape of the tested patterns (e.g., choosing $\text{DBW}_{T_h} = \text{ADT}$ or $\text{ADT}$ results in the only two possible EADT values for each pattern) as well as on the fact that they are self-contained/self-computable (e.g., no fix thresholds, but directly dependent on the ADT).

Finally, we can compute the Effective ADT values $\text{EADT}_{i,j}$ for each model $i$ and each pattern $P_j$. These are then mapped of a MOS score, using the mapping function $\mathcal{F}(\cdot)$. The obtained results are reported in Table III.

The differences among users’ QoE feedback and the estimations done through the $\mathcal{F}(\text{ADT})$ and the $\mathcal{F}(\text{EADT})$...
Table III. Model evaluation in Google Maps. MOS values marked in red correspond to the closest approximations to the participants’ ratings.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>ADT</th>
<th>MOS₀</th>
<th>MOSₐdt</th>
<th>MOSₐtdt</th>
<th>MOSſtdt</th>
<th>MOSₙt</th>
<th>MOSₐrea</th>
<th>MOSdoub</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P₁</td>
<td>5.9 Mbps</td>
<td>2.9</td>
<td>3.3</td>
<td>3 (3.2)</td>
<td>3.1 (3.2)</td>
<td>3.3</td>
<td>3.2 (3.3)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>P₂</td>
<td>5.1 Mbps</td>
<td>3.4</td>
<td>3.2</td>
<td>2.9 (2.9)</td>
<td>3.2 (3.2)</td>
<td>2.9</td>
<td>3.1 (3.2)</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>P₃</td>
<td>4.3 Mbps</td>
<td>3.1</td>
<td>3.1</td>
<td>2.5 (3)</td>
<td>2.8 (3.1)</td>
<td>2.8</td>
<td>3.0 (3.1)</td>
<td>2.3</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 8](image_url) Overall quality for constant DBW and tested fluctuation patterns.

(1) Overall quality for constant DBW. (b) Overall quality for tested fluctuation patterns.

 approaches are quite small, which makes more difficult to draw strong qualitative conclusions about the goodness of the models. Still, there are some interesting first take aways worth to discuss. From the five tested models, the LTD/SLTD and the AREA models are the ones showing the best properties for modeling bandwidth fluctuations. The (S)LTD models follow a very simple and intuitive parametrization regarding bandwidth drops, whereas the AREA model is able to account for how big is the corresponding drop. This suggests that the total share of time in which the bandwidth drops below a certain bandwidth threshold and the deepness of such bandwidth drop are good features to consider when modeling fluctuations in terms of QoE impact. A simple ADT-based QoE estimation approach tends to overestimate the QoE in the event of bandwidth outages, whereas the LTD model provides more conservative results, generally underestimating the MOS scores. The JT model is not accurate enough to capture the QoE feedback of participants, and it is outperformed by the simple ADT approach. Results obtained with the AREA model are similar with both tested thresholds DBWₖ, which suggests that the AREA model is more robust to parametrization than the other ones. Finally, the DOUBLE model provides a much better match for outages (pattern P₃) as compared to the ADT approach, but it is too drastic in the other cases, providing very low EADT values. Fast changing patterns such as P₂ and P₃ are not correctly captured by the AREA model.

V. Conclusions

In this paper we have analyzed the impact of network bandwidth fluctuations on Web QoE through a subjective lab study, particularly focusing on single bandwidth outage events. Our findings reveal that even short outages (4 seconds and shorter) significantly influence the user’s annoyance level. Nevertheless, most users (more than 80%) accept outages during Web usage up to a duration of 8 seconds. In general, the impact of outages is highly application-dependent as also confirmed by our results i.e. compared to video and music streaming, Web browsing QoE is less sensitive to outages. We have also shown that outages are not always detected by the users (e.g. the detection rate of outages with a duration of 16 sec. was only 60% in case of News Site browsing). Nevertheless, if outages are detected by users, the impact on user’s annoyance is significant. Hence, it is important to consider the detection behavior when MOS values are interpreted.

In terms of bandwidth fluctuations modeling, we introduced and evaluated some very first models which are able to correctly capture (in some cases) the impact of fluctuations on users’ experience. We introduced an evaluation approach to test the goodness of the proposed models, through the definition of a very simple and applicable KPI, referred to as the Effective ADT. Conclusions from the evaluation of the models are rather qualitative, as unfortunately, the differences among users’ QoE and the estimations done from the models and a standard ADT-based approach are rather small. Testing these models in the wild, i.e., with real traffic and bandwidth fluctuation patterns and subjective QoE feedback is our next step to verify their accuracy. We believe that the results provided in this paper have a highly applicable perspective in current and future mobile networks, specially by making the point that QoE in networking scenarios is not only about high speed and low latency connections, but also about highly stable ones (at least in terms of bandwidth).

REFERENCES