Bridging the Gap Between QoE and User Engagement in HTTP Video Streaming

Christian Moldovan
University of Duisburg-Essen
Modeling of Adaptive Systems
Essen, Germany
Email: christian.moldovan@uni-due.de

Florian Metzger
University of Duisburg-Essen
Modeling of Adaptive Systems
Essen, Germany
Email: florian.metzger@uni-due.de

Abstract—On video streaming platforms, users expect a high Quality of Experience (QoE). In contrast, service and content providers aim at high User Engagement, most notably because their revenue is usually dependent on it. In order to satisfy users, it is critical to know how QoE is related to the User Engagement. However, no model for this relation exists yet. Current approaches to managing QoE are usually based on traffic analysis. However, this will become more difficult in the future since the encryption of Internet traffic progresses.

In this paper, we present an approach to bringing QoE and User Engagement together with video streaming as the use case. We do this by fitting existing measurement data of User Engagement to obtain a model. Furthermore, we extend the existing queuing model for QoE and investigate the correlation between QoE and User Engagement in a simulation. Hereby, we model different scenarios where we quantify network bandwidth and video requirements.

Our results show that there is a strong correlation between QoE and User Engagement. Additionally, we observe that simple buffer policies, which do not rely on buffer information, can still perform well. These revelations open the way for new approaches to QoE Management in the future Internet.

I. INTRODUCTION

In the past few years, TCP-based video streaming services, like Netflix or YouTube, have garnered enormous popularity amounting to a significant share of the Internet's traffic (compare e.g. [1]). For both, service providers as well as network operators, it is in their best interest to optimize the QoE of those video streams in order to satisfy their customers.

Although service providers typically can directly derive QoE metrics from the data that flows between them and the service users, network operators usually lack this option. In order to get a faint picture of the current QoE conditions in the network, operators have to go the extra mile and intensively conduct traffic analyses and deploy middleboxes to directly investigate and influence traffic conditions. In the future, this will become even more challenging, as almost all services and protocols are slowly moving to deploy full end-to-end encryption, eliminating many traffic analysis approaches.[1]

In contrast, there is another class of metrics that could deem worthwhile, namely context factors. Using this kind of indirect measures and metadata of various origin, an astonishing amount of information can be revealed on the traffic of interest. Moreover, many of these factors could still be collected with end-to-end encryption.

Of particular interesting variant of context factors are User Engagement metrics, which give a description of the amount a user interacts with a given service. For example, in the case of video streaming, an engagement metric would be the (relative) duration of a user watching a specific video. Service providers already use such metrics, e.g. for the successful placement of advertisements or to keep users invested into their platform. Since direct QoE measures might become harder to obtain in the future, the question is, if such User Engagement metrics can be used as a replacement and, if at all possible, which trade-offs have to be taken.

This paper aims to further this exact research question of the possible correlation between User Engagement and QoE for on-demand video streaming. In addition, the correlation and the interdependence between those two metrics needs to be evaluated. Speaking of HTTP streaming, the impact of the client's buffering policy on the QoE and on the User Engagement needs to be investigated.

We conduct this investigation via a queuing model that describes the video player behavior in terms of stalling periods for arbitrary network conditions and video characteristics. The results are mapped to QoE according to our QoE model. Further, we propose a model for User Engagement that is established by fitting existing measurement results. Based on these models, we analyze the correlation between QoE and User Engagement numerically. In addition, we use measurement results from a real environment for the comparison of different buffer policies. Thus, the main contributions of this paper are as follows:

- We present an approach to align QoE and User Engagement.
- We show that our queuing model can be applied to any video with a generally independently distributed bit rate.
- We show that the choice of the buffer policy has no significant impact on either the QoE or the User Engagement.
- An important implication of our results is that User Engagement will be useful for QoE management in the future.

---

[1] https://software.dell.com/docs/2015-dell-security-annual-threat-report-white-paper-15657.pdf

The remainder of the paper is structured as follows. Section II briefly covers the importance of QoE and User Engagement in the context of future Internet developments. In Section III we provide background and related work. Section IV describes our system and user model with Section V presenting and assessing results from this model and further simulations. Finally, Section VI concludes the paper and outlines future work.

II. QOE MANAGEMENT IN THE FUTURE INTERNET

Many contradicting visions exist for a future version of the Internet, but they can be roughly divided into two schools of thought: The "clean slate" approach, that essentially wants to have a differently organized network structure created completely independent from the current structure, and the evolutionary approach, aiming to iterate on and alter current network architectures with endeavors such as Content-Aware Networking (CAN) (cf. e.g. [2] and [3]).

One of the chief goals thereof is to create the capability to select the right QoE strategy inside the network depending on the type of content. In order to achieve this, the network has to be aware of the content that the flows are transporting. This information can be attained from several angles. Either by developing new protocols and mechanisms that intentionally involve the network with the transport and thereby revealing enough metadata to it, or by using Deep Packet Inspection (DPI), which is a more likely case. The network can then start to optimize TCP-based video streaming by taking into account the client's buffering behavior.

One aspect of a future Internet is oft underrepresented in research, although this development is currently observable in today's Internet: The transition to fully end-to-end encrypted transmissions. A series of events in the recent past makes this development pretty self-evident, including:

- The Standardization of the HTTP/1.1-successor HTTP/2 in [4]. While, in contrast to earlier specifications, its final form does not mandate the use of Transport Layer Security (TLS) any more, major browser and server vendors have agreed to enforce TLS in their implementations nonetheless.[2]
- Both Google and Mozilla intent to phase out unencrypted HTTP usage completely[3][4]. Statements from major Internet organizations also strongly discourage the further use of insecure protocols for most applications, including the World Wide Web Consortium (W3C)[5], the Internet Architecture Board (IAB)[6], and the Internet Engineering Task Force (IETF) [5].
- Major service providers, including the YouTube video service, already migrated much of their infrastructure
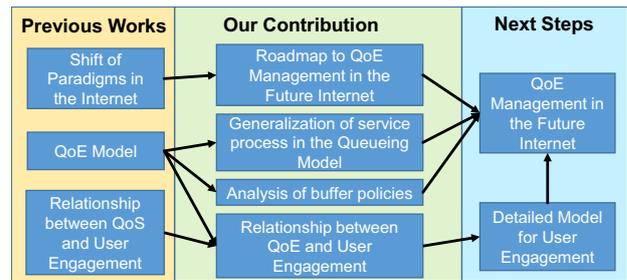

Figure 1. Contributions of this work and relationship to existing research.

to use HTTPS by default in order to reduce the attack surface of its users and to reduce the influences of middleboxes in networks. The performance cost and overhead for servers as well as to the connection is becoming more and more negligible [6] due to ongoing optimization efforts.[7]

Therefore, in the near-to-mid-term future almost all data will be transported in an encrypted manner making traffic analysis much more difficult. This is a fact, that has to be kept closely in mind for future research and also impacts the aforementioned QoE management aspect as it prevents one from successfully employing DPI.

But this might make other metrics that can still be derived from encrypted transmissions much more interesting to use. In the case of User Engagement metrics, many of them can not only be measured through monitoring but also by capturing simple flow and TCP characteristics, e.g., the length and throughput of a flow and if it is aborted prematurely. With a clear relationship between User Engagement and QoE, QoE-aware network management and dimensioning could then be easily conducted on a per-source or per-destination basis, if the User Engagement metrics detect issues for a specific source or destination.

III. BACKGROUND AND RELATED WORK

This paper touches the topics of HTTP streaming, User Engagement, as well as the monitoring and management of Quality of Service (QoS) and QoE. Each of these three will be tackled here.

A. HTTP Video Streaming Background

HTTP video streaming has been taking up a large portion of Internet traffic in recent years. Unlike past Real-time Transport Protocol (RTP)-based streaming approaches, this newer technique employs the reliable TCP as transport protocol, bringing along very distinct characteristics.

When a user opens up a video player of such a streaming platform, the player will issue an HTTP request for the video file. Video data is then progressively downloaded and put into a playback buffer, meaning that playback can be started before the whole video file is downloaded. After an initial stalling period and after the buffer contains a certain amount of video data the playback starts. Further

---

[2] http://daniel.haxx.se/blog/2015/03/06/tls-in-http2/

[3] https://blog.mozilla.org/security/2015/04/30/deprecating-non-secure-http/

[4] https://groups.google.com/a/chromium.org/forum/#!topic/blink-dev/2LXKVWYkOus/

[5] https://w3c.github.io/webappsec/specs/powerfulfeatures/

[6] https://www.iab.org/2014/11/14/iab-statement-on-internet-confidentiality/

[7] https://istlsfastyet.com/

on, if not enough data is being transmitted in time, stalling will occur. This is the key differentiator to RTP streaming, where in the same scenario, frames would have been dropped or corrupted but the playback would have continued. Stalling events lead to dissatisfaction for users which can be objectively measured as QoE.

In adaptive video streaming videos are partitioned into independently playable segments and each video will be provided in several quality levels and thus in several download volumes. This gives the player another degree of freedom and allows to reduce the video quality in order to minimize stalling events. A longer overview on HTTP streaming is given in [7] or [8]. Furthermore, e.g., [9] presents a model-based approach to evaluate the buffering characteristics of different players and strategies. This paper only takes non-adaptive video streaming into account, and leaves adaptive streaming for future work.

B. Related Work on User Engagement and QoE in non-live Video Streaming

The authors of [10] measure QoE-metrics and User Engagement from various sites, different types of content (short Video On Demand (VOD), long VOD, and live video), and also distinguish other kinds of parameters. Their results show that a high buffering ratio lowers User Engagement, with the impact being stronger for short videos. Similarly, a high bit rate has a significant impact in the live scenario while it does not in VOD. In [11] traffic during a single live event is measured and the impact of QoE metrics on User Engagement is analyzed. Their results show that the buffering ratio and the bit rate have a high impact on User Engagement. Further, they noted that the video play time may depend on various other factors such as user behavior. A correlation between QoE and User Engagement was also recognized.

A 2014 paper [12] conducted a large scale measurement study that looked at the abandonment rate — which can be another appropriate User Engagement metric — for mobile video streaming. Using data from the study, a model is proposed that can predict User Engagement in mobile video streaming with a high accuracy based on network statistics.

In two further publications, Balachandran et al. [13], [14] measured User Engagement and video session quality and run machine learning algorithms on it. Through this effort, they highlight the challenges of obtaining a robust video QoE model from such metrics. And finally, a paper from Krishnan et al. [15] puts viewer behavior in relation to video quality metrics. Of note is the observation that an increase in the initial delay of a video stream also directly leads to a higher abandonment rate.

User Engagement can be defined in many different ways. E.g., time spent on a website, abandonment rate, interactions, click rate, attention paid, number of comments. It is interesting from the perspective of the content provider and the service provider since high User Engagement leads to a higher number of ad views or sales. For video streaming services, we need an easily measurable,

objective metric that describes how much content users consume and how willing they are to view ads. Therefore, we define User Engagement as the *view time* of a video. An overview of models for User Engagement metrics for a number of online services is given in [16]. In the remainder of this paper we look at average values of the User Engagement on a per video level. In addition, users might abandon a service because of stalling, thereby reducing User Engagement. Using this definition, it seems plausible that video streaming platforms, content providers, or video service providers generate revenue based on User Engagement making it a critical metric.

C. QoE Management Related Work

So far, all these papers have looked at the significance of specific User Engagement metrics but lack in terms of mapping measured QoS values to a specific QoE and how to facilitate this information for network and QoE management aspects in light of the future Internet development. The following publications investigate this from an Internet Service Provider (ISP)'s point of view.

For an ISP it is generally more difficult to estimate the video streaming QoE in its network and may require invasive measures, such as DPI. However, this is possible with the approach suggested in [17] which was also successfully deployed in the network of a large European mobile operator [18]. Data gained from such monitoring endeavors can be further utilized, e.g., to enable flow-based traffic management for improving QoE via SDN as presented in [19].

Once such influence factors from all network layers have been collected they can be mapped to QoE according to existing QoS-QoE models and relationships. As this is not without challenges, [20] surveys current research activities on QoE management with a focus on wireless networks where QoE management has mostly just been considered in terms of resource scheduling and resource allocation decisions. Furthermore, the involved technology continuously advances and introduces new challenges. Such as the migration of services to the cloud [21], and in particular cloud gaming [22], [23]. Home gateways are also a starting point to optimize QoE at a small scale. [24] shows that even with just very basic knowledge of the users service requirements, a significant improvement in QoE can be achieved through methods such as application prioritization and traffic shaping.

In contrast, managing QoE based on User Engagement estimates the user's QoE with objective metrics. For example, [25] defines a reception ratio as the ratio between download throughput and video encoding rate. For some ISPs this may already be sufficient to determine whether stalling occurs or not and how the user reacts in response. [17] concludes that this ratio cannot be directly related to the QoE, yet it is still a good indicator if there are problems in the network. Both [11] and [26] investigate and review different engagement measures and how they are impacted by QoE metrics.

The problem with many such QoE management approaches is that for some services the models are not fully understood or there may be further, hidden influence factors which are not captured by the employed methodology, e.g., recency effects. additionally, even if the models are well established, e.g. for HTTP video streaming, it may still be difficult to measure the related parameters. Similarly, looking at research efforts involving video streaming engagement, often the reasons are unclear why a user stops watching the video. It may stem from quality issues in the streaming process, but it may very well also just be that the user lost interest in that particular content.

While many QoE metrics might not be measurable in an HTTPS environment, User Engagement can be measured more easily at large scale. Additionally, the effects of low quality will be directly visible in this User Engagement measures. Since there seems to be a lot of value in investigating the relation between QoE and engagement, the aim of this work is to bridge those two fields together, preparing the way to combine their advantages in new models. An overview of our contribution and its relationship with previous and future work is presented in Figure 1.

IV. System and User Model

In the following we discuss our system model, QoE and User Engagement in on-demand HTTP video streaming. The buffer of the video player is modeled as a queuing model with network bandwidth patterns and video characteristics (frame size, frame rate) as input and stalling patterns as output. The presented QoE model maps key performance parameters to QoE. Further, we introduce a new model for User Engagement in which a performance parameter is mapped to play time. Additionally, we point out the limitations of our model.

A. Player Model

The video player determines which frame is played out at which point in time. On application level, video frames are downloaded in order into a buffer with rate λ. Downloaded frames are replayed from the buffer with a frame rate μ. We define the ratio between the download rate and the replay rate as offered load (or reception rate) $a = \lambda/\mu$. If $a < 1$ (i.e. $\mu > \lambda$), frames are replayed faster than they are downloaded, which will lead to an empty buffer. In this case, the player pauses the replaying process, which is called stalling. In order to resume the replaying process, a condition has to be met that is determined by the players policy. E.g. playback resumes when the number of frames in the buffer surpass the predefined buffer size d. The normalized buffer size d^* relates the buffer size d to the video frame rate, i.e. $d^* = d/\mu$. Three basic player policies are discussed in Section IV-D.

We modeled the video player as a queuing model with a single service unit as follows. The arrival process of the frames follows a random distribution that can be assumed to be Markovian or generally distributed. The service process follows a Markovian random distribution.
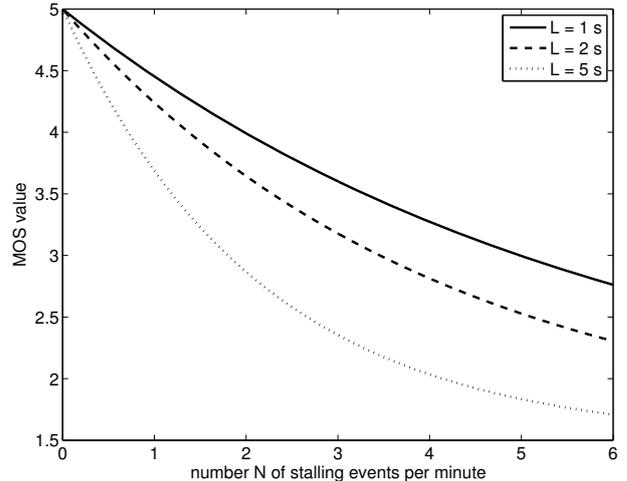

Figure 2. Impact of the number N of stalling events and the length L of stalling events on the MOS according to the QoE model for $\alpha = 0.15$ and $\beta = 0.15$.

QoE-relevant metrics were deducted in [27] as follows. The player is assumed to always be in one of two states: playing or stalling. The average length of a stalling period is given as L. The average length of a playing period is given as the busy period B. The relative amount of time spent in stalling compared to the relative amount spent replaying the video is given as the ratio of buffering events R. The number of stalling events normalized by the video duration is given as the normalized buffering ratio N^*. An overview of the notation is given in Table I.

B. QoE Model

The QoE model is provided in [17] and is based on subjective experiments in which test subjects assessed QoE for short videos with varying stalling patterns. In order to objectively assess QoE, we quantify it in terms of Mean Opinion Score (MOS). The MOS is defined as a value between 1.5 and 5 with 1.5 corresponding to the lowest QoE value and 5 to the highest. For the QoE Q, the subjective results of [17] show $Q(L, N^*) = 3.5 \cdot e^{-(\alpha L + \beta)N^*} + 1.5$ with α and β being the parameters for the user's sensitivity to stalling. This relation is depicted in Figure 2.

C. User Engagement

User Engagements describes the activity or attention of users in a system. As described in Section III-B, for video streaming we restrict the definition of User Engagement to the average amount of time (in minutes) users watch a video. In [10], several data sets were collected and analyzed. In Figure 3 we take a closer look at one of their data sets: LvodA which contains long VoD clips with a length of about $35\,\mathrm{min}$ to $60\,\mathrm{min}$. In each data point users with the same ratio of buffering events are related to an average play time. We fitted a nonlinear curve to these data points in least-squares sense using MATLAB, which provides us with a fitting function

$$U(R) = 4.2712 \cdot e^{-0.5435 \cdot R} + 25.9000 \cdot e^{-0.0339 \cdot R} \quad (1)$$

|	
video and network input parameters	
λ	network bandwidth (f/s)
μ	video frame rate, with $\mu > \lambda$ (f/s)
a	offered load $a = \lambda/\mu$ captures the relative network bandwidth (normalized by video bitrate)
player parameters and status variables	
d	buffer size (f)
d^*	buffer size normalized by video bit rate $d^* = d/\mu$ (s)
performance parameters	
B	average length of the busy period (s)
B_X	average length of the busy period (s) for policy X
T	video duration (s)
N^*	number of stalling events normalized by video duration ($1/s$)
N_X^*	number of stalling events normalized by video duration ($1/s$) for policy X
R	ratio of stalling events
L	average length of a stalling event (s)
L_X	average length of a stalling event (s) for policy X
Q	QoE value
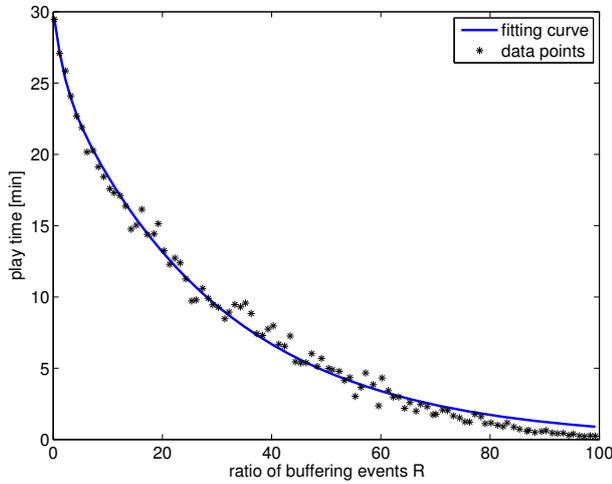
Figure 3. Fitting based on the buffering ratio R. Data taken from [10].

This function maps the ratio of buffering events R to the average play time in minutes. For the fitting, we chose a double exponential decay since this is commonly used for describing spontaneous human behavior (e.g. in [28]). The Pearson correlation coefficient for this fit is 0.996 (Spearman 0.997). The RMSE is $0.659 \, \text{min}$ (normalized RMSE $0.092 \, \text{min}$). This indicates that the fit is very accurate.

D. Impact of Player Policies on QoS and QoE

In this section we present an analytical approach to calculating key QoS parameters for the player model described in IV-A. We do this by extending an existing $M/M/1$ model from [27] to an $M/G/1$ model for three buffer policies which are presented in [29]: the D-policy, the n-policy, and the T-policy.

A detailed mean value analysis of the steady state was derived in [27] for an $M/M/1$ model. Similarly, we derive the key performance metrics from Figure 4 as follows. The video download starts at t_0 with bytes being downloaded
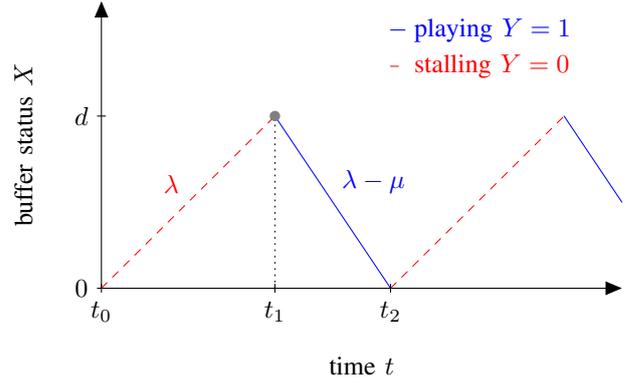

Figure 4. Video buffer status evolving over time with constant video bitrate and network bandwidth.

at rate λ. If the downloaded bytes amount to d at t_1, the video is being replayed from the buffer with video bit rate μ while the downloading continues. This means the data in the buffer is being removed at a rate of $\mu - \lambda$ until the data is the buffer is reduced to 0 at t_2. Then the replaying stalls and the process repeats. The average length of stalling events is $L = t_1 - t_0 = \frac{d^*}{a}$. Further, L is identical to the idle period. The ratio of buffering events is $R = \frac{t_2 - t_1}{t_3 - t_1} = 1 - a$. For the average length of the busy period (or playing period) we yield $B = t_2 - t_1 = \frac{d^*}{1-a}$. If we look at the frequency of busy periods during the video replaying, we get the normalized buffering ratio $N^* = \frac{1}{B} = \frac{1-a}{d^*}$. For the QoE, this leads to

$$Q(L, N^*) = 3.5 \cdot e^{-(\alpha \frac{1-a}{a} + \beta \frac{1-a}{d^*})}.$$

Next, we extend the $M/M/1$-model from above for the n-policy, the D-policy and the T-policy. In [29], the authors derive the distributions and the means of the busy and idle periods of queuing models for these three policies. In the following, we adapt these results for HTTP video streaming for the case of a generally distributed service process. For the sake of clarity, when variables for specific policies are discussed, they have the policy name as their index, e.g. for the D-policy, we use the notation L_D, B_D, N_D^* instead of L, B, N^*.

1) $M/G/1$ with D-policy: With a D-policy, the idle period ends if the sum of the service times of the units in the queue amounts to D. For the specific case of video streaming, this policy means that stalling ends after the data in the buffer amounts to a certain play time D. This policy guarantees that the length of the busy period is at least D. For the D-policy, it is

$$E[L_D] = \frac{1}{\lambda}(M(D) + 1)$$

with $M(D)$ being the renewal process of D,

$$E[B_D] = \frac{M(D) + 1}{\mu - \lambda},$$

$$E[N_D^*] = \frac{1}{B_D} = \frac{\mu - \lambda}{M(D) + 1}.$$

It follows

$$Q(L_D, N_D^*) = 3.5 \cdot e^{-(\lambda-\mu)(\frac{\alpha}{\lambda} + \frac{\beta}{M(D)+1})} + 1.5.$$

If we assume the buffer size d does not change during a video session, then $M(D) = d - 1$ is constant as well. Thus $L = L_D$ and $N^* = N_D^*$ are equal for $M/M/1$ and $M/G/1$ with D-policy. Therefore, $Q(L, N^*) = Q(L_D, N_D^*)$ is equal for both models.

2) $M/G/1$ with n-policy: With the n-policy, the idle period ends if $n = d^* \cdot \mu$ bytes are in the queue. For the n-policy, it is

$$E[L_n] = \frac{n}{\lambda} = \frac{d^*}{a},$$

$$E[B_n] = n \cdot \frac{1}{\mu(1-a)} = \frac{d^*}{1-a},$$

$$E[N_n^*] = \frac{1}{B_n} = \frac{1-a}{d^*}.$$

Since $E[L_n]$ and $E[N_n]$ are equal for M/M/1 and M/G/1 with n-policy, $Q(L, N^*) = Q(L_n, N_n^*)$ is also equal for both models.

3) $M/G/1$ with T-policy: With the T-policy, when an idle period starts, a timer is started. If the timer reaches T, the systems verifies whether a unit arrived during the idle period. If it did arrive, the busy period is started. Otherwise, the timer is restarted. The probability that no unit arrives during the idle period T is $e^{-\lambda T}$. If we can assume $e^{-\lambda T} = 0$, this policy guarantees that the length of each stalling period is exactly T. For the T-policy, it is

$$E[L_T] = \frac{T}{1 - e^{-\lambda T}},$$

$$E[B_T] = \frac{\lambda T}{(1 - e^{-\lambda T})(\mu - \lambda)},$$

$$E[N_T^*] = \frac{1}{B_T} = \frac{(1 - e^{-\lambda T})(\mu - \lambda)}{\lambda T}.$$

It follows

$$Q(L_T, N_T^*) = 3.5 \cdot e^{(1 - \frac{1}{\alpha})(\alpha + \beta \frac{1 - e^{-\lambda T}}{T})} + 1.5.$$

If we can assume $e^{-\lambda T} = 0$ and choose $T = \frac{d^*}{a}$ this leads to $Q(L_T, N_T^*) = Q(L, N^*)$. In Figure 5 we see that the impact of the policies is small for $T = d^*$. Consequently, the service process has no impact on the QoE under the assumption of a Markovian arrival process. This means that only the mean and not the variance of the video bit rate matters for the QoE, assuming an $M/G/1$-model. This result was also observed in simulation results that will be presented in Section V.

E. Model Limitations

The QoE model that we use is based on subjective experiments of short video clips ($T = 30$ s). Nevertheless, we use this model for steady state analysis in Section IV-D and in our simulation in Section V-C for a movie with a length of 12 min. However, subjective experiments for long videos are currently missing in literature and the model needs to be validated against subjective experiments
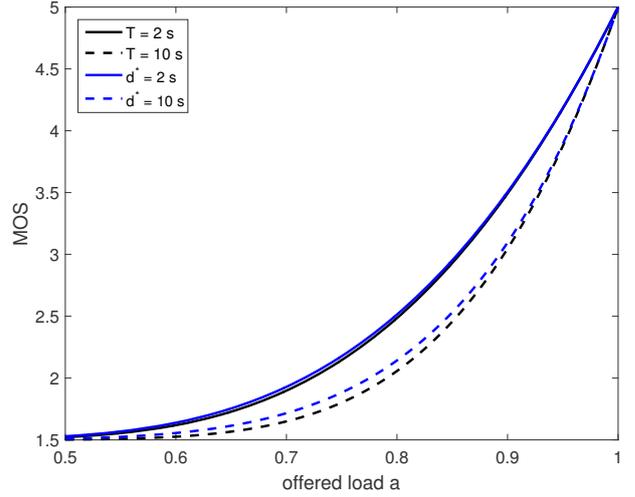

Figure 5. Difference between T-policy and D-policy if $T = d^*$ is chosen. The n-policy is omitted for the sake of clarity.

with longer videos. New objective and subjective user tests are necessary in order to provide a general QoE function that takes the duration of a video into account. This issue is a current research topic and is discussed in greater detail in [30] and [31].

While we map the ratio of stalling events to the User Engagement, there are also other factors (such as startup delay [15] or the video bit rate [11]) which influence engagement. Measurement studies for engagement need to consider those factors too, in order to derive a complete model. Future work should include such factors in order to refine the proposed model. Nevertheless, this work is an important first step in identifying the relationship between QoE and User Engagement.

V. RESULTS

This section takes a closer look at the relation between QoE and User Engagement by discussing analytic results for the queueing model described in IV-A. In addition, we look at the simulation of the download of a real video in a real network and compare it with the analytic results.

A. Analytic Results

First, we focus on the D-policy as it reflects current video player implementations of HTTP streaming, and investigate the impact of reception rate (i.e. offered load) and different buffer sizes on QoE and User Engagement. Later, in Section V-C, we compare the different policies in terms of QoE and User Engagement.

Figure 6 shows how the offered load (or ratio between network bit rate and video bitrate) a is related to the MOS value and the play time for different buffer sizes d^* (e.g. a value $a = 0.5$ means that the bandwidth is half of the video bit rate). We notice that increasing the offered load a leads to an increasing average MOS and User Engagement. It should be noted that MOS values lower than 2.5 are not considered acceptable by most users [32]. In addition, we see that an increasing buffer size
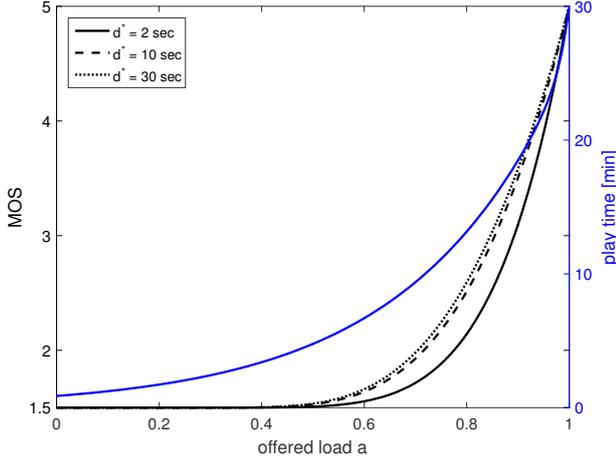
Figure 6. QoE value in an $M/G/1$ system. The offered load quantifies the ratio between average network bandwidth and video bit rate.
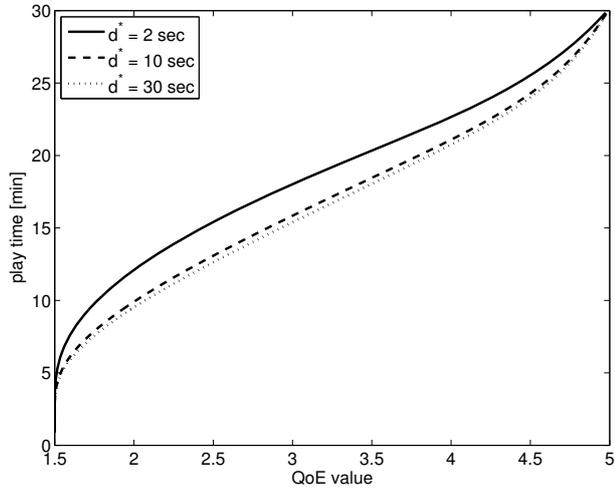

Figure 7. User Engagement in relation to the QoE in an $M/G/1$ system for $d = 2, 10, 30$ according to Equation 1 in Section IV-C.

leads to higher MOS with the optimum being reached at $Q_+ = \lim_{d^* \to \infty} Q(L, N^*) = e^{-\alpha \frac{1-a}{a}}$ as shown in [27]. In contrast, the buffer size does not have an impact on the User Engagement. A large difference between QoE and User Engagement is that for $a < 0.4$ the MOS is 1.5 and does not change while increase the play time is noticeable. This is because the QoE model that we use is based on short video clips while the user engagement model is based on long videos. User Engagement has been observed to be lower for shorter videos [15].

Next, we investigate how the QoE value is related to the User Engagement. In Figure 7 we calculated the User Engagement and the QoE for various offered loads. We observe that an increase in QoE always leads to an increase in User Engagement. Since our model for User Engagement does not take the buffer into account, more research is necessary in order to identify its impact. Furthermore, we notice that for very low QoE values, it is difficult to estimate the User Engagement as users may
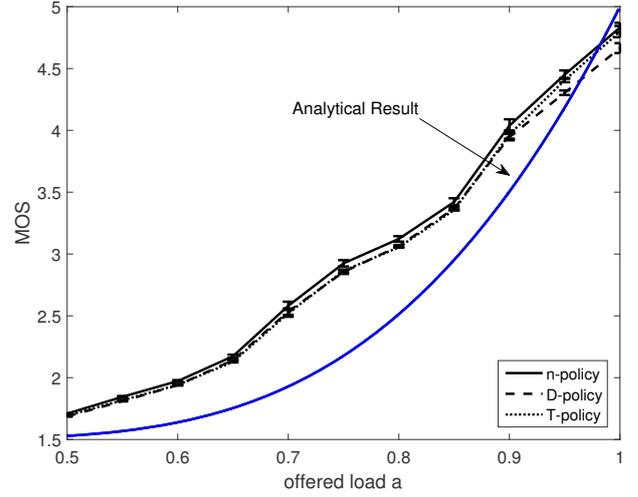

Figure 8. Relation between MOS and offered load a for three buffer policies. Simulation results for real video and real traffic pattern (black) and analytical results for $M/G/1$-model (blue). Policies have little impact on MOS.

react differently in such scenarios.

The Pearson correlation coefficient is 0.981 (Spearman 0.994). A larger buffer leads to a higher mean play time for the same QoE. This can be explained by the fact that an increase in buffer size leads to an increase in QoE, but not to an increase in mean play time. This means that users will abort videos much earlier if the QoE is low. Therefore, it is critical to ensure a high QoE if User Engagement is to be maximized.

B. Simulation Environment

In order to compare our analytic results with measurement results, we simulated replaying a real video using a real network trace that was recorded in [33]. For this simulation we chose the video "Tears of Steel" in a low spatial resolution (320×180). It is a $12\,\mathrm{min}$ short movie with a variable bit rate. The network trace was recorded by downloading a large file via HTTP using a UMTS stick while driving on a highway. The resulting trace has a strongly fluctuating bit rate. In total, we used 30 different traffic patterns that were created by adding a temporal shift to the original traffic pattern in [34]. We simulated different network capacities by adjusting the video bit rate, resulting in various offered loads a.

In our simulation, video frames are downloaded with a rate that is based on the effective network capacity and the size of the video frame in a best effort manner. Video frames are replayed at a constant rate of 24 frames per second until the video ends. If a stalling event occurs, it is resolved according to the given buffer policy. The simulator is implemented in MATLAB and is available online[8].

C. Simulation Results

In the following, we compare the simulation results to the analytic results. Figure 8 shows the impact of the

---

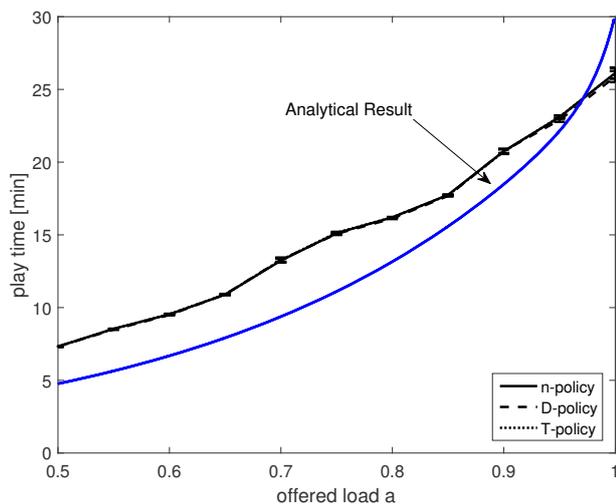[8]https://github.com/ChristianMoldovan/HAS-Simulator

Figure 9. Relation between play time and offered load a for the three buffer policies. Simulation results for real video and real traffic pattern (black) and analytical results for $M/G/1$-model (blue). Policies have no impact on play time.

offered load a on the MOS for the n-policy, the D-policy and the T-policy. It is clearly visible that the policy does not impact the QoE value significantly. In addition, the real traces lead to a higher QoE than the $M/G/1$ model. This is mainly due to video specific attributes, i.e. the distribution of frame sizes. Nevertheless, we consider the G-distribution a reasonably good approximation for the distribution of frame sizes in videos. While more advanced models may lead to more realistic results, they can not be solved analytically.

In Figure 9 we investigate how the buffer policies impact the User Engagement that was calculated based on the rate of buffering events R according to Equation 1 in Section IV-C. The main observation is that the policies have almost no impact on the mean play time. This means that since the T-policy does not require any information from the player, it provides a solid alternative to the other policies. This is particularly the case if hiding such information becomes a common practice in the future.

VI. CONCLUSION

In this paper, we studied QoE and User Engagement for an $M/G/1$ queuing model and for real measurement data. We achieved this by first creating a fit that relates video quality metrics to User Engagement. We showed analytically that $M/G/1$ with n-policy, D-policy and T-policy is equal to the $M/M/1$ queuing model in terms of QoE and User Engagement. An interesting observation was that the T-policy does not perform much different than other policies, while in contrast it does not rely on buffer information. This may open new approaches to QoE management and should be investigated in future work.

Furthermore, we noticed a strong correlation between QoE and User Engagement which indicates that User Engagement monitoring is important for QoE management. Therefore, collaboration between the community of User Engagement researchers and the community of

QoE researchers will be necessary in the future. A more precise relationship between User Engagement and QoE may be established through future subjective and objective experiments.

REFERENCES

[1] Cisco, *Cisco visual networking index: global mobile data traffic forecast update 2014–2019*, Feb. 2015.

[2] M. Grafl, C. Timmerer, H. Hellwagner, *et al.*, "Scalable media coding enabling content-aware networking," *MultiMedia, IEEE*, vol. 20, no. 2, pp. 30–41, 2013.

[3] J. Pan, S. Paul, and R. Jain, "A survey of the research on future internet architectures," *Communications Magazine, IEEE*, vol. 49, no. 7, pp. 26–36, 2011.

[4] M. Belshe, R. Peon, and M. Thomson, *Hypertext transfer protocol version 2 (HTTP/2)*, RFC 7540 (Proposed Standard), Internet Engineering Task Force, May 2015.

[5] S. Farrell and H. Tschofenig, *Pervasive monitoring is an attack*, RFC 7258 (Best Current Practice), Internet Engineering Task Force, May 2014.

[6] D. Naylor, A. Finamore, I. Leontiadis, *et al.*, "The cost of the S in HTTPS," in *Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies*, ACM, 2014, pp. 133–140.

[7] K. Ma, R. Bartos, S. Bhatia, *et al.*, "Mobile video delivery with HTTP," *Communications Magazine, IEEE*, vol. 49, no. 4, pp. 166–175, 2011, ISSN: 0163-6804.

[8] A. Begen, T. Akgul, and M. Baugher, "Watching video over the web: part 1: streaming protocols," *IEEE Internet Computing*, vol. 15, no. 2, pp. 54–63, Mar. 2011, ISSN: 1089-7801.

[9] F. Metzger, A. Rafetseder, and K. Tutschku, "A performance evaluation framework for video streaming," in *Packet Video Workshop (PV), 2012 19th International*, 2012, pp. 19–24.

[10] F. Dobrian, V. Sekar, A. Awan, *et al.*, "Understanding the impact of video quality on user engagement," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 362–373, 2011.

[11] M. T. Diallo, F. Fieau, and J.-B. Hennequin, "Impacts of video quality of experience on user engagement in a live event," in *Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on*, IEEE, 2014, pp. 1–7.

[12] M. Z. Shafiq, J. Erman, L. Ji, *et al.*, "Understanding the impact of network dynamics on mobile video user engagement," in *The 2014 ACM international conference on Measurement and modeling of computer systems*, ACM, 2014, pp. 367–379.

[13] A. Balachandran, V. Sekar, A. Akella, *et al.*, "A quest for an internet video quality-of-experience metric," in *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, ser. HotNets-XI, Redmond, Washington: ACM, 2012, pp. 97–102, ISBN: 978-1-4503-1776-4.

[14] ——, "Developing a predictive model of quality of experience for internet video," in *ACM SIGCOMM Computer Communication Review*, ACM, vol. 43, 2013, pp. 339–350.

[15] S. S. Krishnan and R. K. Sitaraman, "Video stream quality impacts viewer behavior: Inferring causality using quasi-experimental designs," *Networking, IEEE/ACM Transactions on*, vol. 21, no. 6, pp. 2001–2014, 2013.

[16] J. Lehmann, M. Lalmas, E. Yom-Tov, *et al.*, "Models of user engagement," in *User Modeling, Adaptation, and Personalization*, Springer, 2012, pp. 164–175.

[17] T. Hoßfeld, R. Schatz, E. Biersack, *et al.*, "Internet video delivery in YouTube: From traffic measurements to quality of experience," in *Data Traffic Monitoring and Analysis*, Springer, 2013, pp. 264–301.

[18] P. Casas, R. Schatz, and T. Hoßfeld, "Monitoring YouTube QoE: is your mobile network delivering the right experience to your customers?" In *2013 IEEE Wireless Communications and Networking Conference (WCNC 2013)*, Shanghai, China, Apr. 2013.

[19] M. Jarschel, F. Wamser, T. Hohn, *et al.*, "SDN-based application-aware networking on the example of YouTube video streaming," in *Software Defined Networks (EWSDN), 2013 Second European Workshop on*, IEEE, 2013, pp. 87–92.

[20] S. Baraković and L. Skorin-Kapov, "Survey and challenges of QoE management issues in wireless networks," *Journal of Computer Networks and Communications*, vol. 2013, 2013.

[21] T. Hobfeld, R. Schatz, M. Varela, *et al.*, "Challenges of QoE management for cloud applications," *Communications Magazine, IEEE*, vol. 50, no. 4, pp. 28–36, 2012.

[22] S. Moller, S. Schmidt, and J. Beyer, "Gaming taxonomy: An overview of concepts and evaluation methods for computer gaming QoE," in *Quality of Multimedia Experience (QoMEX), 2013 Fifth International Workshop on*, IEEE, 2013, pp. 236–241.

[23] M. Jarschel, D. Schlosser, S. Scheuring, *et al.*, "Gaming in the clouds: QoE and the users' perspective," *Mathematical and Computer Modelling*, vol. 57, no. 11, pp. 2883–2894, 2013.

[24] B. J. Villa and P. E. Heegaard, "Towards knowledge-driven QoE optimization in home gateways," in *ICNS 2011, The Seventh International Conference on Networking and Services*, 2011, pp. 252–256.

[25] L. Plissonneau, E. Biersack, and P. Juluri, "Analyzing the impact of youtube delivery policies on user experience," in *Proceedings of the 24th International Teletraffic Congress*, International Teletraffic Congress, 2012, p. 28.

[26] K. De Moor, M. R. Quintero, D. Strohmeier, *et al.*, "Evaluating QoE by means of traditional and alternative subjective measures: An exploratory'living room lab'study on IPTV," *Vienna, Austria*, 2013.

[27] T. Hoßfeld, C. Moldovan, and C. Schwartz, "To each according to his needs: Dimensioning video buffer for specific user profiles and behavior," in *Third IFIP/IEEE International Workshop on Quality of Experience Centric Management (QCMan 2015)*, IEEE, Ottawa, Canada, 2015.

[28] B. A. Swett, J. L. Contreras-Vidal, R. Birn, *et al.*, "Neural substrates of graphomotor sequence learning: A combined fMRI and kinematic study," *Journal of neurophysiology*, vol. 103, no. 6, pp. 3366–3377, 2010.

[29] K. Gakis, H.-K. Rhee, and B. Sivazlian, "Distributions and first moments of the busy and idle periods in controllable M/G/1 queueing models with simple and dyadic policies," *Stochastic Analysis and Applications*, vol. 13, no. 1, pp. 47–81, 1995.

[30] P. Frohlich, S. Egger, R. Schatz, *et al.*, "Qoe in 10 seconds: Are short video clip lengths sufficient for quality of experience assessment?" In *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, IEEE, 2012, pp. 242–247.

[31] A. Borowiak, U. Reiter, and U. P. Svensson, "Quality evaluation of long duration audiovisual content," in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, IEEE, 2012, pp. 337–341.

[32] R. Schatz, T. Hoßfeld, L. Janowski, *et al.*, "From packets to people: Quality of experience as a new measurement challenge," in *Data traffic monitoring and analysis*, Springer, 2013, pp. 219–263.

[33] C. Müller, S. Lederer, and C. Timmerer, "An evaluation of dynamic adaptive streaming over http in vehicular environments," in *Proceedings of the 4th Workshop on Mobile Video*, ACM, 2012, pp. 37–42.

[34] M. Seufert, C. Sieber, T. Zinner, *et al.*, "Identifying QoE optimal adaptation of HTTP adaptive streaming based on subjective studies," *Computer Networks*, 2015.