Impact of Access Line Capacity on Adaptive Video Streaming Quality - A Passive Perspective

Impact of Access Line Capacity on Adaptive Video Streaming Quality – A Passive Perspective

Martino Trevisan, Idilio Drago, Marco Mellia
Politecnico di Torino, Italy
firstname.lastname@polito.it

ABSTRACT

Adaptive streaming over HTTP is largely used to deliver live and on-demand video. It works by adjusting video quality according to network conditions. While QoE for different streaming services has been studied, it is still unclear *how access line capacity impacts QoE of broadband users in video sessions*. We make a first step to answer this question by characterizing parameters influencing QoE, such as frequency of video adaptations. We take a passive point of view, and analyze a dataset summarizing video sessions of a large population for one year. We first split customers based on their estimated access line capacity. Then, we quantify how the latter affects QoE metrics by parsing HTTP requests of Microsoft Smooth Streaming (MSS) services. For selected services, we observe that at least 3 Mbps of downstream capacity is needed to let the player select the best bitrate, while at least 6 Mbps are required to minimize delays to retrieve initial fragments. Surprisingly, customers with faster access lines obtain limited benefits, hinting to restrictions on the design of services.

CCS Concepts

•Networks $\rightarrow$ Network performance analysis; Network measurement;

Keywords

QoE-Metrics; Live Video; Access Line Capacity

1. INTRODUCTION

On-line video is among the most important Internet applications, being the number one in terms of network traffic [5]. Adaptive bitrate streaming over HTTP is a dominant technology to deliver live and on-demand video over the Internet. The key idea is to adjust video playback in real-time. The same video is produced and stored in different bitrates, and clients decide which media file to fetch based on measurements taken during video sessions. As such, video quality is adapted to changes in network conditions and surges in workloads at servers. The goal is to obtain the best possible video quality, while minimizing odds of events that would affect users' experience. Given the predominance of video in the Internet, understanding the QoE of users in adaptive video sessions as well as the factors that cause quality degradation is fundamental to improve services.

Many studies target QoE in adaptive streaming over HTTP, and readers can refer to [8, 10] for complete surveys on the topic. Whereas QoE is subjective and measuring it requires user involvement, a series of high-level metrics are identified as key factors influencing QoE [10], e.g., video bitrate and frequency of bitrate switches [6]. Works targeting measurements of such parameters are however scarce [8] and mostly rely on active methodologies. Few large-scale works have been presented, mostly relying on client instrumentation (see the works from Conviva, e.g., [4]).

Conversely, in-network passive characterizations of QoE-related factors in adaptive video mostly target a single video provider [3], or propose methodologies to reconstruct video sessions [7] and to extract QoE metrics [9]. Large-scale analyses of parameters related to QoE covering both live and on-demand video, based on in-network measurements, have been almost neglected, possibly due to the challenges for obtaining and processing relevant datasets.

Intuitively, network capacity and, in particular, access capacity offered by technologies such as ADSL or FTTH, could play a key role for video streaming quality – i.e., more capacity would favor better experience. However, we are unaware of studies that quantify the relation between access capacity and streaming QoE.

This paper is a first step toward understanding how access network capacity impacts the QoE of broadband users in video sessions. We take a passive point of view

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

This research has been funded by the Vienna Science and Technology Fund (WWTF) through project ICT15-129, "BigDAMA".

and analyze a dataset summarizing video sessions of thousands of users in Italy. We split users in groups based on estimated access line capacity, and contrast groups when watching on-line video streaming services.

We devise a passive methodology based on the parsing of Microsoft Smooth Streaming[1] requests to extract QoE figures. We focus on Sky and Rai, two popular streaming services offering live and on-demand video in Italy. We measure parameters influencing QoE, such as video bitrate, frequency of adaptations and the delay to receive initial video fragments. We show that users with slow access lines suffer more from low video quality and long delays to receive initial fragments than users with fast lines. When at least 3-6 Mbps are available, little differences are observed among users of the same content provider, suggesting access line capacity is not the bottleneck, and systems are limited by design to support only low video encoding rates.

Considering video adaptions, little differences are noticed between users with slow and fast connections contacting the same content provider. This behavior seems to be related to the fact that clients with fast connections try to obtain better qualities, resulting in more switches. Slow users instead remain with stable (low) bitrate during complete sessions. This holds for different providers, and suggests that client- and server-side factors dominate bitrate switches.

2. DATASET AND METHODOLOGY

2.1 Dataset

We rely on a dataset collected in a Point of Presence (PoP) of an Italian Internet Service Provider (ISP), where the traffic of approximately 8,000 households (i.e., broadband installations) is aggregated. ADSL lines of different capacities provide access to more than 80% of the customers. Despite customers are connected to the same PoP, downlink/uplink capacity changes based on physical constraints of the copper medium, or on selected contracts [2]. Those are limited to no more than 24/1 Mbps by the technology. 20% of households enjoy Fiber To The Home (FTTH) with either 30/10 Mbps or 100/10 Mbps. Ethernet or WiFi are used at home to connect users' devices (e.g., PCs or smartphones), which may further limit the available capacity.

We rely on Tstat to perform passive measurements. Tstat monitors each TCP connection, exposing information about more than 100 metrics, including anonymized client IP addresses, the amount of exchanged data, and the Fully Qualified Domain Name (FQDN) clients resolve via DNS queries prior to open flows [1]. The latter is instrumental to characterize general Internet usage. Tstat HTTP monitoring plugin additionally exports HTTP requests and responses. This information includes requested URLs, the `user agent` field in requests, and the Content Type and Content Length in responses. It allows us to identify user devices – e.g., PCs or mobile phones – as well as to infer details of videos services, as we will describe later.

[1] www.microsoft.com/silverlight/smoothstreaming

In total, we captured data for 12 months from 2014 to 2015. We use 10 months for our analysis, excluding holidays from the dataset. Our dataset consists of more than 17 billion flows, 29 billion records of HTTP requests or responses and around 5 TB of compressed logs. Giving the challenge of handling such amounts of data, we load and process the data using Apache Spark in a medium-sized big data cluster.

Given our focus on video streaming, we have checked which are the most popular video services regularly accessed by users. We have extracted from HTTP logs the most popular hostnames, and manually identified those that are related to streaming services. Results show that Sky, Rai and Mediaset, the most popular broadcasters in Italy, are among the most popular video services people access over the Internet too, with about 10%, 8% and 7% of users contacting them, respectively. Note that Netflix was not available in Italy at data collection time, and YouTube uses HTTPS, thus making our approach based on DPI ineffective.

In the following, we focus on Sky On-Demand, Sky Live and Rai Live. We focus on streams accessed from PCs based on MSS technology, which allows the extraction of video playback parameters directly from the HTTP requests.[2] Overall, we observe about 12 k, 9 k and 5.5 k video sessions, respectively, for a total of 76 million HTTP requests over ten months.

2.2 Grouping Households

Our goal is to examine whether and how user's access line capacity impacts quality of video streaming. Our dataset however does not include a direct indication of the access line capacity of households. Therefore, we take the following steps to create groups of households according to inferred line capacity.

First, because the studied ISP always assigns a fixed IP address per household h, we group flows per client IP address. Next, for each household, we select non-persistent HTTP flows downloading more than 5 MB. Then, we calculate the average download throughput $Tru(f,h)$ of each flow f as the ratio between the bytes downloaded from the server and the time between the last and the first server segment with payload. We assume at least some downloads are able to saturate line capacity per month. Thus, for each month of data $M \in \{M_1, \ldots, M_{10}\}$, we pick the flow with the *highest throughput*, $HighTru(h,M) = \max_{f \in M} Tru(f,h)$, obtaining 10 samples per household.

We use these 10 samples to decide whether the household will be part of the analysis and to assign it to a group. We select those households exhibiting simi-

[2] MSS is based on Silverlight, available for PCs. Smartphones and tablets contacting the services use HTTP Live Streaming standard. The same approach can be extended to other not-encrypted streaming services.

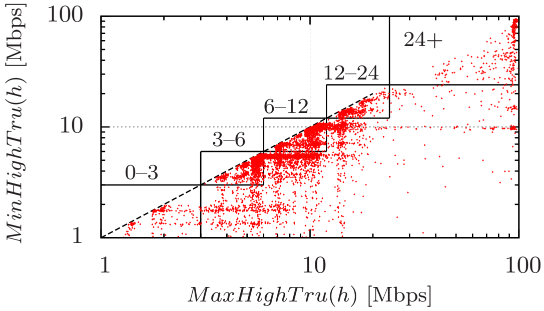


Figure 1: Groups: points represent households.

lar $HighTru(h, M)$ across the 10 months, to filter out customers that may have upgraded technology during time. For that, we evaluate the maximum and minimum highest throughput on the selected 10 samples, $MaxHighTru(h)$ and $MinHighTru(h)$. This information is depicted in Fig. 1. Points close to the diagonal represent households with consistent $HighTru(h, M)$ among the 10 months. Most points are indeed close to the diagonal, since the majority of households achieves the highest throughput in every month of collection.

The figure shows clear clusters – i.e., households connected with the same access technology and line capacity. We create 5 groups of households selecting slow [0-3) Mbps, moderate [3-6) Mbps, good [6,12) Mbps and excellent [12-24) Mbps ADSL access lines. Customers which consistently achieve [24-100] Mbps throughput are served by a FTTH access line. Classes are marked by squares in the figure. Numbers inside the figure highlight x- and y-axis ranges defining the resulting squares.

Tab. 1 summarizes the groups, showing the number and percentage of households per group, as well as the throughput range that defines each group. In total, we retain more than 5,000 households (i.e., 65% of the total), with at least 200 households per group. Finally, we have performed several analyses to characterize the behavior of users in each group and rule out macroscopic biases that could affect comparisons. We find no significant differences on video consumption across groups. Details are omitted for the sake of space.

2.3 Services & Quality Metrics

We focus on Sky and Rai, two major providers offering live and on-demand video via Internet in Italy. Both providers deliver video using MSS. Clients watching videos make HTTP requests to obtain manifest files, called *Media Presentation Description* (MPD), that contain streaming metadata. During playback, they keep downloading video and audio fragments. Every HTTP request for a fragment contains parameters that reveal requested encoding bitrate and timestamps.

By inspecting and processing HTTP requests exported by Tstat, we track how the media playback progresses and extract metrics related to video session QoE.

Table 1: Population in households' groups.

Name	Range [Mbps]	Households	Percentage
Group 1	0–3	213	2.8
Group 2	3–6	1,607	20.8
Group 3	6–12	2,509	32.5
Group 4	12–24	441	5.7
Group 5	24–100	288	3.7
Others	–	2,668	34.5
Total	–	7,726	100%

A session is defined as the consecutive group of HTTP requests issued by the same client to fetch fragments of the same video content. We sort fragments based on timestamps, since a client may open parallel TCP connections towards the same server when watching a video. We examine only sessions longer than 2 minutes, considering them expired when no more fragments are observed for 1 minute. Here, we focus on three simple metrics correlated with QoE:

• **Encoding Bitrate:** The image quality is fundamental to QoE, and can be estimated by the encoding bitrate. Since we focus on MSS, we can easily determine the requested bitrate by parsing the standardized URLs employed by the protocol. In fact, every request for a video fragment includes the `QualityLevels` parameter, which we directly leverage to obtain the instantaneous bitrate of a playback. We then calculate both the average encoding bitrate per video session, and the number of requests for fragments in each bitrate.

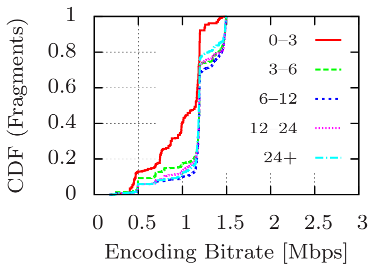
• **Initial Fragments Arrival Delay:** Playback startup delays – i.e., the time elapsed from the user action requesting a video and the playback start – can influence QoE [10]. However, measuring startup delays requires information about buffer settings on the client-side, which is not exposed in our dataset.

We thus focus on an *indirect* metric that is related to startup delay: the delay to receive initial fragments of a video. The rationale is that playback starts only after some fragments arrive at the client. We compute the time between the first HTTP request for a MPD file and the time of an *arbitrary* number of HTTP requests for video fragments. The indirect metric is then evaluated for users of the same content provider.

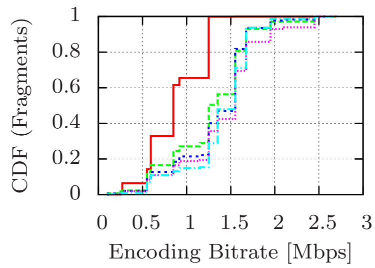
• **Frequency of Video Adaptations:** We evaluate how the video bitrate evolves in video sessions and track bitrate switches. Whereas switches are normal in adaptive video and used to prevent drastic problems such as stalls, frequent switches may affect QoE.

Bitrate switches are identified by inspecting URLs as well. We count the number of times clients experience a *reduction* on requested bitrate. We then normalize the metric obtaining the number of reductions per minute.[3]

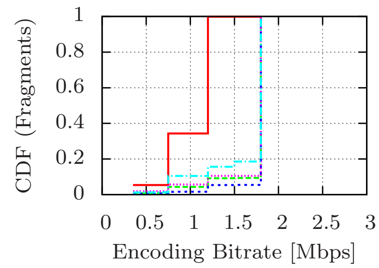
[3]The number of increases on bitrate per session is directly proportional to the number of reductions in our dataset, except for the initial transient.



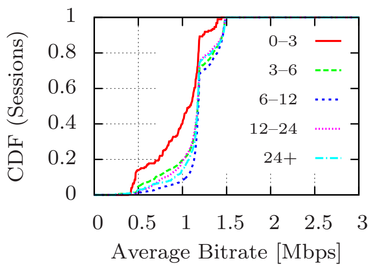
(a) Sky On-Demand



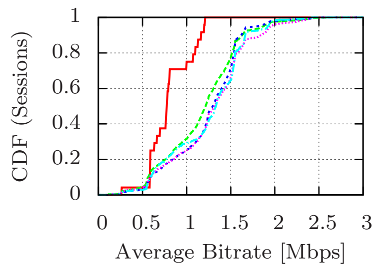
(b) Sky Live



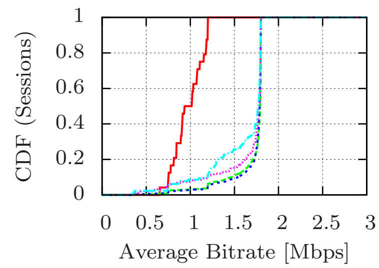
(c) Rai Live



(d) Sky On-Demand



(e) Sky Live



(f) Rai Live

Figure 2: Encoding bitrate of requested fragments (1st row) and average bitrate per video session (2nd row).

3. STREAMING PERFORMANCE

3.1 Encoding Bitrate

We first consider the requested encoding bitrate. Intuitively, this is a metric that should be closely related to the access line capacity – i.e., videos should be played at best bitrate when the network capacity allows for it.

We quantify how requested encoding bitrate varies in Fig. 2, where Cumulative Distribution Functions (CDFs) are reported. The first row of figures quantifies per-fragment CDF of encoding bitrates. The second row of figures shows per-session CDF of average encoding bitrate, i.e., we consider each video session, compute the average bitrate among all fragments, and then compute the CDF. Lines mark metrics for each group. Distinct plots report results for *On-Demand* and *Live* videos of Sky and *Live* videos of Rai from left to right.

Notice on Figs. 2a–2c how shapes of distributions are concentrated around the typical bitrate of videos offered by the providers. Each provider adopts a different configuration. For instance, Sky and Rai Live streaming platforms serve up to 2.5 Mbps and 1.8 Mbps video rates, respectively, with the latter offering a coarser set of rates. On the contrary, Sky On-Demand limits the rate to 1.5 Mbps and offers a more fine-grained set of possible encoding rates. Manually checking, we observe that some third-party videos are encoded and distributed at fixed rate – e.g., advertisements.

In all plots, we notice that the first group (0–3 Mbps – red lines) suffers from a significantly lower bitrate. In particular, focusing on Fig. 2a, note how these customers more rarely request 1.25 Mbps (or higher) rates than other users. This pattern is even more evident in Fig. 2b where bitrate for live videos is considered.

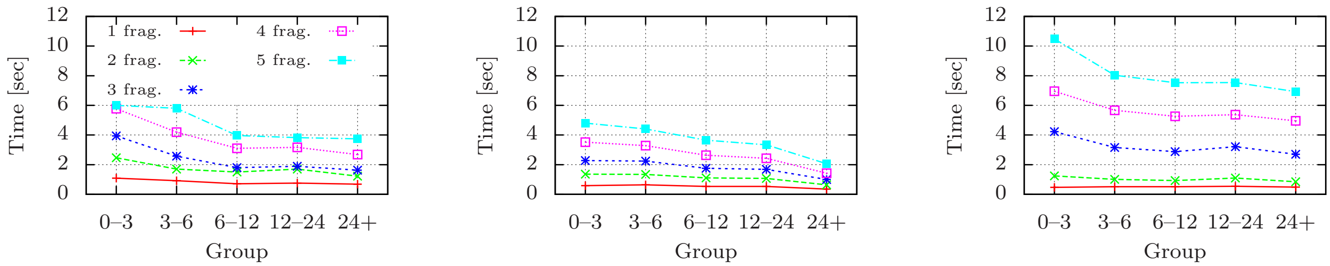
Figs. 2d–2f quantify the average encoding bitrate per session. As before, different plots per service are depicted, as well as independent lines per group. Contrast these plots to those in the first row. We see that Fig. 2d is similar to Fig. 2a, indicating that video sessions of Sky On-Demand do not experience a large number of bitrate switches. Fig. 2f instead suggests that sessions of users with slow access capacity (red line) constantly alternate between the two bitrates usually retrieved by those users when contacting Rai (see red line in Fig. 2c). As a result, video sessions have average bitrate spread in between the per-fragment rates offered by the provider. We will investigate bitrate switches in details later.

Considering customers with more than 3 Mbps, only minor differences are observed. For instance, Fig. 2b shows that households with the fastest connections (e.g., 12–24 Mbps and 24+ Mbps) seldom enjoy the best bitrate in live videos, similarly to users in 3–6 Mbps. This suggests that the content provider in question is bottlenecked by other design choices. Similar artifacts are seen for other providers.

In a nutshell, the measurements confirm and quantify the relation of line capacity and bitrate. We show that as soon as 3 Mbps is available to clients, no significant differences in encoding bitrate are noticed.

3.2 Initial Fragments Arrival Delay

We evaluate the delay clients experience to receive the initial fragments in a session, as an indirect metric



(a) Sky On-Demand (b) Sky Live (c) Rai Live

Figure 3: Average per-group delays to receive initial fragments of three video services.

of startup delay. We compute the time between the request for the initial MPD file of a session and the arrival of the first 5 video fragments. Fig. 3 reports the average arrival time of fragments for different groups.

A clear trend emerges. As the available bandwidth increases, the time to retrieve first fragments decreases significantly. Notice on Fig. 3a (On-Demand) how the time to receive the 5th fragment is on average 30% smaller for 24+ Mbps households than for those connected at 3 Mbps or less. Live video is equally impacted.

That is, access line capacity has a large impact on this metric and increases the likelihood of problems in slow lines – e.g., longer delays to fill in initial buffers. When capacity is higher than 6 Mbps, however, little differences are observed.

Differences between Sky On-Demand and Sky Live (Fig. 3a and Fig. 3b) can be explained by the size of fragments: On-demand video is shipped in fragments of 5 s, whereas live videos are split in 2 s fragments. Thus, users at 24+ Mbps obtain up to 25 s of on-demand video in 4 s (on average), while 10 s of live video arrive in 2 s.[4]

Rai instead seems to rely on smaller client-side buffers. While we cannot confirm buffer sizes based on our dataset, note how the inter-arrival time of video fragments (i.e., distance between lines in Fig. 4c) is similar across groups already after the second fragment. Thus, it is likely that the player buffer is full and the video already starts playing (i.e., reaching steady-state) after the second fragment. Note how the combination of a likely small buffer with limited line capacity results in tight margins for users in slow connections to receive fragments. In fact, only after around 10 s since the MPD file arrives, users with $\leq$ 3 Mbps receive the 5th fragment, needed to play actual 10 s of video. Yet, conclusions similar to those for Sky hold: Only users with very slow links seem to be exposed to problems.

3.3 Video Adaptations

Finally, we evaluate bitrate switches in Fig. 4. All figures report the distribution of the number of times we see a decrease of bitrate *per minute*. We use box plots to represent percentiles of the distribution: the central boxes span from the 25th to the 75th percentiles, while whiskers mark the 5th to 95th percentiles. The black bar represents the median of the distribution.

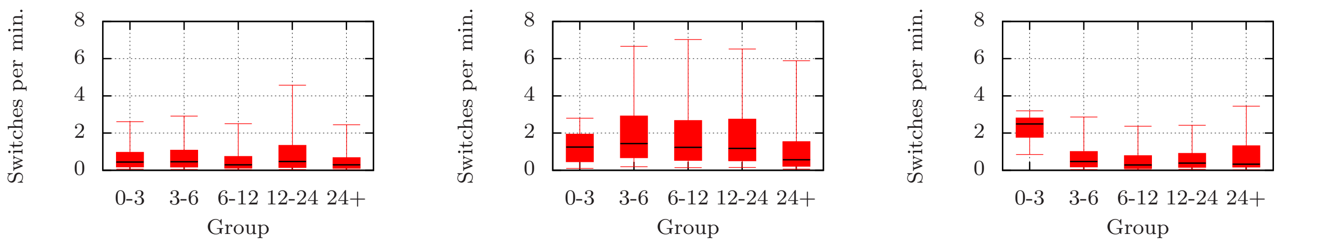
Focusing on Fig. 4a, note how the number of bitrate switches is equivalent among groups. We see that for on-demand video, 75% of playback minutes have 1 or less switches for all groups. This pattern can likely be explained by large buffers usually deployed for on-demand video. Even for users with very slow connections (i.e., $\leq$ 3Mbps), buffers on the client-side accommodate short-term fluctuations on network conditions.

Contrast now Fig. 4a to Fig. 4b. Bitrate switches are more frequent in live video than in on-demand video for the same provider. First, client-side buffers play a role, since live video is usually played with tighter buffers. Moreover, when we put these figures in perspective with those in previous sections, we see that live video of Sky is delivered in smaller pieces with higher bitrate than Sky On-Demand videos. This causes several bitrate switches, creating artifacts on the metric.

Now, compare the box for the 0-3 Mbps group in Fig. 4b with boxes for other groups. While differences in the body of distributions are small (see patterns of medians), we see that the tail of the distribution for households connected at 0-3 Mbps is shorter (see 75th and 95th percentiles). Remind Fig. 2b, where we see that users in this group typically receive a much lower bitrate. That is, by limiting the bitrate of live videos for users with slow connections, the media player obtains lower quality, but more stable playback. Households with 6-12 Mbps, for instance, experience better video quality, but 5% of them, as a consequence, see as many as 7 bitrate decreases per minute. This indicates a more aggressive setting of quality adaption policies, which results in a more oscillating behavior.

Finally, for live streaming of Rai (Fig. 4c), there is an evident trend among groups, where households with better connections obtain less switches and better bitrate. While our data does not provide a conclusive explanation for this behavior, we conjecture it is a con-

[4] The duration of fragments is estimated by the inter-arrival time of requests for fragments in steady-state.



(a) Sky On-Demand (b) Sky Live (c) Rai Live

Figure 4: Average per-group switching frequency of three popular services.

sequence of smaller client buffers deployed by Rai: slow customers are more exposed to problems in this case.

Overall, these results show that having better connectivity does not necessarily mean that less bitrate switches will be achieved. In fact, given the interplay between factors such as client buffers and network speed, it might be that faster connections experience more bitrate switches while trying to fetch video fragments of better quality.

4. CONCLUSIONS

We presented a preliminary analysis of the relation between access line capacity and streaming quality. Our goal was to quantify metrics correlated to QoE taking a completely in-network and passive point of view. We found evidences that very slow links decisively impair video playback. Users with faster links enjoy similar quality regardless of their access line speed. Interestingly, bitrate switches are similar across all users, with some services deploying aggressive policies that result in several switches per minute. Results call for a deeper look into how users perceive such variations.

4.1 Limitations & Future Work

We highlight some limitations of our measurements and possible directions for future work.

First, we relied on heuristics to determine line capacities. Whereas groups are somehow consistent in our dataset, there are many factors that hide the actual access line capacity, or artificially mix customers of different groups, e.g., occasional use of WiFi at home. Ground truth of users' connection setup at home is needed to improve group definitions, and we plan to explore this direction in collaboration with the ISP.

Second, whereas we found correlation between low capacity and QoE-metrics, the nominal capacity value imposing problems to users (e.g., ≤ 3 Mbps) is naturally dependent of the analyzed services. It is to be expected that services deploying videos with better rates will challenge these values. We plan to investigate a generic model to explain how offered bitrate, line capacity and QoE-metrics are related.

Finally, we assumed that cross-traffic equally affects results for users with different access lines. Some applications however are more aggressive than others in the network. Thus, measurements need to be put into perspective of users' habits while watching videos. We plan to study in the future how different applications running in parallel to video sessions influence QoE-metrics.

5. REFERENCES

[1] I. Bermudez et al. DNS to the Rescue: Discerning Content and Services in a Tangled Web. In *Proceedings of the IMC*, pages 413–426, 2012.

[2] F. Bota et al. Predicting ADSL Lines Data Rate Using Neural Network. In *Proceedings of the GLOBECOM*, pages 1–5, 2011.

[3] P. Casas, M. Seufert, and R. Schatz. YOUQMON: A System for On-line Monitoring of YouTube QoE in Operational 3g Networks. *SIGMETRICS Perform. Eval. Rev.*, 41(2):44–46, 2013.

[4] F. Dobrian et al. Understanding the Impact of Video Quality on User Engagement. *SIGCOMM Comput. Commun. Rev.*, 41(4):362–373, 2011.

[5] J. L. García-Dorado et al. Characterization of ISP Traffic: Trends, User Habits, and Access Technology Impact. *IEEE Trans. Netw. Service Manag.*, 9(2):142–155, 2012.

[6] T. Hossfeld, M. Seufert, C. Sieber, and T. Zinner. Assessing Effect Sizes of Influence Factors towards a QoE Model for HTTP Adaptive Streaming. In *Proceedings of the QoMEX*, pages 111–116, 2014.

[7] R. Huysegems et al. Session Reconstruction for HTTP Adaptive Streaming: Laying the Foundation for Network-based QoE Monitoring. In *Proceedings of the IWQoS*, pages 15:1–15:9, 2012.

[8] P. Juluri, V. Tamarapalli, and D. Medhi. Measurement of Quality of Experience of Video-on-Demand Services: A Survey. *Commun. Surveys Tuts.*, 18(1):401–418, 2016.

[9] R. Schatz, T. Hossfeld, and P. Casas. Passive YouTube QoE Monitoring for ISPs. In *Proceedings of the IMIS*, pages 358–364, 2012.

[10] M. Seufert et al. A Survey on Quality of Experience of HTTP Adaptive Streaming. *Commun. Surveys Tuts.*, 17(1):469–492, 2015.