Quality of Experience and Access Network Traffic Management of HTTP Adaptive Video Streaming

Michael Seufert, Phuoc Tran-Gia (supervisor)
Institute of Computer Science, University of Würzburg, Würzburg, Germany
{seufert | trangia}@informatik.uni-wuerzburg.de

Abstract — The thesis focuses on Quality of Experience (QoE) of HTTP adaptive video streaming (HAS) and traffic management in access networks to improve the QoE of HAS. First, the QoE impact of adaptation parameters and time on layer was investigated with subjective crowdsourcing studies. The results were used to compute a QoE-optimal adaptation strategy for given video and network conditions. This allows video service providers to develop and benchmark improved adaptation logics for HAS. Furthermore, the thesis investigated concepts to monitor video QoE on application and network layer, which can be used by network providers in the QoE-aware traffic management cycle. Moreover, an analytic and simulative performance evaluation of QoE-aware traffic management on a bottleneck link was conducted. Finally, the thesis investigated socially-aware traffic management for HAS via Wi-Fi offloading of mobile HAS flows. A model for the distribution of public Wi-Fi hotspots and a platform for socially-aware traffic management on private home routers was presented. A simulative performance evaluation investigated the impact of Wi-Fi offloading on the QoE and energy consumption of mobile HAS.

Index Terms — HTTP Adaptive Video Streaming; Quality of Experience; Monitoring; Traffic Management; Performance Evaluation; Wi-Fi Offloading

I. CONTRIBUTION OF THE THESIS

Video streaming is one of the most popular and demanding applications of today’s Internet. The high data volumes, bandwidth requirements, and delay constraints of video traffic pose a lot of challenges to ISPs, which want to deliver the traffic as efficiently as possible, while maintaining a high subjectively perceived service quality. To measure the satisfaction of end users with a networked service, the concept of Quality of Experience (QoE) has been established.

This thesis [84] focuses on understanding and improving the QoE of the current HTTP Adaptive Video Streaming (HAS) technology, which allows to align the streaming demands to the network conditions by adapting the video bit rate. After identifying the most important QoE factors, possibilities for traffic management are considered to improve the subjectively perceived quality of video streaming in access networks. Therefore, an accurate QoE monitoring is required to notice quality degradations of the customers and appropriate traffic management actions have to be applied. As this vast research field cannot be covered completely in a single monograph, only selected topics are studied. This thesis investigates different QoE monitoring approaches on network and application layer, and also evaluates the performance of selected QoE-aware and socially-aware traffic management approaches for HAS in access networks.

Figure 1 visualizes the research activities, identifies the selected topics of this thesis, and embeds them into the research fields. Thereby, the cartography is split vertically into different layers, which loosely resembles the OSI model of telecommunication systems. However, the layers up to transport layer have been combined into the network layer, and, typical for the Internet model (TCP/IP model), the higher layers have been combined into the application layer. Today’s highly interactive Internet services require the addition of a user layer, and the ubiquitous social networking and interactions among different users are reflected in the social layer. Horizontally, the cartography is divided into the three main research goals, which are monitoring, traffic management, and performance evaluation. Note that some of the publications might fit to more than one category. In this case, the publications are listed in the category, which yields the largest overlap. The works have been conducted within the EU FP7 project SmartenIT, the EU H2020 project INPUT, the sub-project Mobi-QoE of the EU H2020 project MONROE, the DFG projects OekoNet, QoE-DZ, and Crowdsourcing, the COST action Qualinet, the project P-SERQU of the NGMN Alliance, and several industry funded projects.

Fig. 1. Cartography of the conducted research. The notation [x]y indicates that the scientific publication [x] is discussed in Chapter y of this thesis.

The research activities can be roughly integrated into six fields. First, the QoE of Internet applications was investigated to find the most important factors that affect the subjectively perceived quality of such services. A special attention was
given to video streaming services, which are very popular and have high network demands. In Figure 1, these studies are highlighted in blue, and can be attributed to performance evaluation on user level. In this thesis, the results of the QoE research of [13], [14], [15], [39], [52] are presented. Based on the findings of previous QoE studies, the impact of adaptation-related parameters on the QoE of HAS is investigated. The impact of quality and time on each layer is identified and quantified, which allows to formulate and solve an optimization problem for HAS adaptation. It can be used to benchmark HAS adaptation logics with respect to QoE-optimal adaptation for given network conditions and video characteristics. A frequently used methodology for QoE research is crowdtesting, i.e., conducting subjective QoE studies via crowdsourcing. To obtain reliable results from a large, diverse, remote, unsupervised crowd of micro job workers is not trivial. Thus, appropriate methodology for crowdtesting had to be developed. The field of crowdtesting is categorized as monitoring on the user level and the corresponding publications are highlighted in violet, but crowdtesting methodology will not be detailed in this thesis.

After the QoE factors of Internet services, such as HAS, are known, traffic management can be applied to improve the perceived service quality. Thereby, the concepts of application-aware networks and network-aware applications were investigated and exploited. This field of QoE-aware traffic management is highlighted in green, and covers all research goals, i.e., monitoring, traffic management, and performance evaluation. This thesis studies QoE-aware traffic management for HAS in access networks, which does not require to operate on traffic aggregates, but allows to identify and manage single video flows. In this context, QoE monitoring approaches on application and network layer are presented following the publications [61], [80], [48], [45], [34]. Moreover, an analytical and simulative performance evaluation is conducted, which compares the QoE gain of different traffic management algorithms for a scenario, in which video streaming and web browsing flows share a bottleneck link [12], e.g., on a the backhaul link of a shared Wi-Fi home network or a mobile base station. Some other studies applied analytical methods to obtain accurate and scalable performance evaluation results on abstract system models, and are highlighted in red but are not included in this thesis. Dedicated performance evaluation studies were also conducted on edge computing, which is a new paradigm that extends cloud computing by additionally utilizing computing resources at the network edge, e.g., servers at mobile base stations or even devices within the homes of end users. Personalized services can be instantiated or migrated at the network edge close to end users to support user mobility and achieve a high QoE. As edge computing goes beyond access network traffic management, these results are not included in this thesis.

QoE-aware traffic management was extended by additionally considering the end user, his shared resources, as well as social information, e.g., about his preferences, his interests, or his interactions with other users. This constitutes the new research field of socially-aware traffic management (SATM), which is highlighted in orange. Similar to QoE-aware traffic management, SATM not only considers traffic management decisions and actions but also the monitoring of social information. The basic concept of SATM is introduced based on [1]. Afterwards, the focus is on Wi-Fi offloading as a SATM mechanism, which is not only widely considered by ISPs to reduce the load on stressed mobile access networks, but is also well suited in the home environment due to the prevalence of residential Wi-Fi networks. This thesis presents a hotspot location model for public Wi-Fi offloading [10], a SATM system for private Wi-Fi offloading in home environments [56], [42], and a performance evaluation on the QoE end energy efficiency of public and private Wi-Fi offloading of HAS flows [29], [47], [44].

In summary, this thesis is focused on the following research questions:

- How to model the impact of adaptation on the QoE of HAS?
- How to monitor the QoE of HAS on network and application layer?
- How to improve the QoE of HAS by resource allocation on bottleneck links?
- How to consider the end user, his shared resources, and social information for traffic management?
- How to utilize existing Wi-Fi infrastructure for improving the QoE of HAS?

II. QUALITY OF EXPERIENCE OF HTTP ADAPTIVE VIDEO STREAMING

In this chapter the impact of adaptation-related parameters on the Quality of Experience (QoE) of HTTP Adaptive Video Streaming (HAS) was investigated. This allowed to identify the most important QoE factors and quantify their impact on subjectively perceived quality of HAS services. These insights allow to improve HAS adaptation logics to mitigate the impact of adaptation on the QoE.

The investigation of the QoE of HAS started with a review of findings from previous studies [14]. These works had concluded that initial delay, stalling, and quality adaptation had the biggest impact on the QoE. Nevertheless, the detailed results on the impact of adaptation on QoE were partly contradicting. Therefore, two crowdsourcing studies were conducted to collect subjective ratings for different adaptation patterns in order to investigate the impact of adaptation-related parameters.

The results of the study on adaptation-related parameters [52] showed that only time on high layer and switching amplitude were main influence factors. In contrast, last quality level, recency time, and switching frequency had no significant impact. The results of related works on switching frequency might not be contradictory, when quality switches are correlated to time on each layer. The quality switches only call the user’s attention to a perceivable degradation/improvement, for which again the duration matters. Similar findings were
indicated in [85], which found that long low quality segments after high quality segments were perceived strongly negatively. Moreover, multiple quality switches were preferred over fewer switches, if the subject could watch a higher quality for a longer time. Thus, the impact of adaptation frequency is inferior to the impact of time on each layer. However, if the switching frequency is too high, such that the user perceives only flickering, the switching itself is the worst degradation, and the QoE might be even worse than for low video quality [86].

The results of the study on time on intermediate layer [39] supported the findings of the study on adaptation-related parameters [52]. It could be observed in both studies that the Mean Opinion Score (MOS) values of the adaptation sequences were bounded by the MOS values of the constant high and low quality sequences. Moreover, it became evident that the quality of each layer (cf. switch amplitude in [52]) and the time on each layer were the only adaptation-related parameters, which had a significant impact on the QoE of HAS. Based on these findings, QoE-based traffic management decisions could be taken in the network by monitoring the requested quality layers (e.g., by deep packet inspection or machine learning on statistical packet features) and pooling the obtained metrics over time into QoE estimators. However, in order reach a high accuracy of estimation and to create a holistic QoE model for HAS systems, the results have to be extended and consolidated for different HAS systems (e.g., more quality layers, streaming on mobile devices) and different video contents (e.g., high motion sequences, longer sequence duration).

As it directly influences the adaptation of the played out video, the implementation of the HAS adaptation logic has the most impact on the QoE in a HAS system. The results showed that it should aim to maximize the played out video quality first before reducing the number of quality switches, for which no significant QoE degradation was found. In a step towards improved HAS adaptation logics on the client side, these QoE findings were incorporated in a linear program, which can be used to compute the QoE-optimal adaptation for given network conditions and video characteristics [13], [15]. This allows for a more comprehensive assessment and benchmarking of HAS adaptation logics with respect to QoE. The applicability of the benchmark was demonstrated in an exemplary performance evaluation of four adaptation strategies. Thus, the linear program is a valuable tool for the performance evaluation of existing and design of future HAS adaptation logics, which use an improved segment download strategy to maximize the played out video quality and reach a higher QoE for HAS.

III. QoE-AWARE TRAFFIC MANAGEMENT FOR HTTP ADAPTIVE VIDEO STREAMING

This chapter investigated QoE-aware traffic management for HAS in access networks. The requirements are QoE monitoring, traffic management decisions, and traffic management actions. Two approaches were presented to monitor QoE within the network and within the application. Moreover, a direct estimation of QoE from monitored network parameters was studied. Eventually, bandwidth allocation strategies for video flows on a shared bottleneck link, such as residential broadband connections in home environments or backhaul links of mobile base stations, were evaluated and the benefits of different levels of monitoring information were discussed.

To monitor the QoE of HAS, two approaches were presented. The first approach complemented the work in [16] and was an example of network-based monitoring [61], [80]. By deep packet inspection (DPI), objective video quality metrics were extracted from the network traffic. These included the requested video quality level of the HAS stream and the video bit rate, which are objective metrics of each segment, and the SSIM metric on a per-frame base. As temporal pooling is well aligned with the QoE findings of Chapter 2, it was applied to combine these objective metrics into a single measure for the whole streaming session. The decent performance of the QoE estimators confirms that temporal pooling is a suitable approach to combine periodical objective quality metrics when monitoring HAS. Thus, its application is also promising for the increasingly used end-to-end encrypted HAS, in which objective metrics cannot be extracted by DPI, but have to be estimated from the network traffic.

The second approach monitored the QoE unobtrusively within an Android YouTube video streaming application (YoMoApp), and could thus provide accurate information on the QoE factors just as they are perceived by the end user [48], [45]. Therefore, JavaScript-based monitoring functions were injected into the streaming website, which monitored the state of the video player, the buffer, and the video quality level. The information about initial delay, stalling, and adaptation, which are the most important QoE factors of HAS, could be computed and might be signaled to the network operator as meaningful QoE feedback for QoE-aware traffic management. The concept was implemented in an Android app, which could be used for researching QoE models for HAS. Moreover, it was used in a field study on QoE estimation from monitored network parameters [34]. The results showed that application-layer information gave better insights into the QoE of HAS than simple network parameters.

Finally, four QoE-aware traffic management strategies based on bandwidth allocation were presented and their performance was evaluated in a shared bottleneck link scenario [12]. Two applications, video streaming and web browsing, were considered and different ratios between the application flows were evaluated. The strategies had different complexity, dynamics, and used different levels of information ranging from pure flow type information to real-time application-layer information about the video buffer. Analytical models based on processor sharing queues were used to obtain average download times for some of the strategies. Additionally, all algorithms were assessed with a Java discrete event simulation. Thereby, also the stalling of the video streaming sessions could be evaluated. The results showed that a joint QoE improvement was possible for both applications compared to the best effort scenario.
without traffic management. However, operators have to select the right strategy and the right parameter settings depending on the traffic mix in their networks and their objectives. Moreover, dynamic traffic management based on application-layer QoE factors proved to be feasible to react to imminent QoE degradations. Thereby, stalling, i.e., the most important QoE degradation of HAS, could be significantly reduced.

IV. SOCIALLY-AWARE TRAFFIC MANAGEMENT FOR HAS BASED ON WI-FI OFFLOADING

This chapter investigated socially-aware traffic management (SATM) [1] for improving the QoE of HAS. SATM utilizes ubiquitous social signals for improved traffic management of Internet services. It is a highly collaborative and cross-layer approach involving many stakeholders, and thus, has a huge potential for applications. As not all applications can be covered comprehensively, the focus of this chapter was on Wi-Fi offloading. Wi-Fi offloading provides a complementary Internet access over a fixed network to reduce the load on mobile networks. The applicability of Wi-Fi offloading can benefit from social information, such as information about location and mobility of users to select appropriate public hotspots, or information about trust between users to share private Wi-Fi networks. A simple hotspot location model for the performance evaluation of public Wi-Fi offloading was designed and its applicability was confirmed. A socially-aware home router platform for private Wi-Fi offloading and video content distribution was presented. Finally, the impact of public and private Wi-Fi offloading on the QoE and energy consumption of HAS sessions was investigated in a simulation.

The performance of Wi-Fi offloading in cities depends on the locations and coverage of Wi-Fi hotspots. To find a general model for such hotspot distributions, the characteristics of public Wi-Fi infrastructure in cities were analyzed [10], [83]. After a transformation to polar coordinates with respect to the city center, a uniform distribution of the angle and an exponential distribution of the distance was fitted. A simple model for generating hotspot distributions for arbitrary cities could be derived. Although the generated hotspot distributions cannot accurately recreate the spatial patterns of real hotspot locations, they could replicate the offloading potential, coverage, or signal strength in a city. Thus, the simple model can be used to generate hotspot distributions for the performance evaluation of Wi-Fi offloading in scenarios, for which real hotspot locations are not available. A performance evaluation study for public Wi-Fi offloading with a generated hotspot distribution showed similar results to the study with the original hotspot distribution, and thus, confirmed the applicability of the simple hotspot location model.

Apart from to public hotspots, also private Wi-Fi infrastructure can be utilized for SATM. However, the sharing of private Wi-Fi networks requires a trust relationship to the owner to avoid abuse by malicious users. To easily enable private Wi-Fi offloading among trusted users, the Home Router Sharing based on Trust (HORST) [56], [42] mechanism was developed, which additionally supports caching, prefetching, and content delivery on the home router. Therefore, HORST consists of a firmware for home routers, an Online Social Network (OSN) app, and a mobile device app. The firmware hosts SATM mechanisms on the private home router, which leverage social information from the OSN app and the mobile app to improve the QoE of HAS. A HORST prototype was implemented, evaluated, and demonstrated, which allowed users to request access at a friend’s HORST home router, receive the credentials for the private Wi-Fi, connect to it, and access the Internet. This proved the feasibility of HORST to increase the Wi-Fi offloading potential. The utilization of other SATM mechanisms, i.e., caching, prefetching, and content delivery, for improved QoE of HAS provided additional incentives to offload with HORST.

The performance evaluation of public Wi-Fi offloading for HAS in the city of Darmstadt, Germany, showed that QoE and smartphone energy consumption of offloaded HAS sessions deteriorated [29], [47], [44]. The reason is the lower throughput of public Wi-Fi hotspots compared to 3G and 4G access. The analysis of offloading to residential Wi-Fi hotspots, which had a fixed broadband connection and a throughput distribution similar to 3G, indicated it is possible to improve the energy consumption compared to public Wi-Fi offloading, and also to 3G access. Thereby, the results came closer to 4G access, which allowed for the best HAS performance in terms of QoE and energy consumption. The results showed that Wi-Fi offloading of HAS sessions is only beneficial for end users if the received bandwidth in the Wi-Fi network is not lower than in the mobile network. This means, the throughput and coverage of public Wi-Fi infrastructure has to keep up with the increasing 4G coverage if improved QoE and energy consumption shall be incentives to encourage offloading. This has to be taken into account by operators, which offer public Wi-Fi as an alternative Internet access. For the moment, private Wi-Fi offloading, e.g., using HORST, is a more promising approach to improve the QoE and energy consumption of HAS, and to reduce the load on cellular networks, especially taking the increasing speeds of residential fixed broadband connections into account [87].

V. SUMMARY AND CONTRIBUTIONS

The current HTTP Adaptive Video Streaming (HAS) technology allows to align the streaming demands to the network conditions by adapting the video bit rate. However, the impact of quality adaptation on the QoE has not been comprehensively investigated yet. Therefore, this thesis reviewed the previous QoE results and investigated the impact of adaptation on the QoE of HAS in subjective crowdsourced QoE studies. It was found that only the time on each layer and its respective image quality impacted the subjectively perceived quality, while the other adaptation-related parameters, such as number of quality changes, did not show a significant effect. Based on these results better adaptation logics can be designed, which reach a higher QoE by maximizing the time on high quality layers first, instead of focusing on low switching frequency or conservative up-switching behavior. These results were also
used to formulate an optimization problem as a linear program, which can be solved to compute the QoE-optimal adaptation strategy for given throughput conditions. This framework allows to benchmark and compare HAS adaptation logics in terms of distance to the optimally achievable QoE. Video streaming service providers can use this benchmark to find the best adaptation logics for realistic network conditions, in which their service is typically consumed.

The identification of the most important QoE factors of HAS allowed to develop monitoring approaches for these factors on network and adaptation layer. Temporal pooling of objective per-frame or per-segment metrics, which were collected by deep packet inspection, reached high correlations to subjective MOS values. ISPs can deploy such monitoring in their networks to keep track of the QoE of their customers in real-time. As only little information about the underlying content is used, the concept of temporal pooling might also be applied in the context of end-to-end encrypted video traffic. Irrespective of end-to-end encryption, application-layer monitoring can be implemented as it monitors the QoE factors directly at the client application. An Android app was developed to monitor the video streaming by injecting JavaScript monitoring functions into the website, which contains the video player element. Additionally, the app allows to log network and device statistics, which can also be used to estimate the QoE. Finally, the app cannot only be used for QoE monitoring, but it can also collect subjective quality ratings of users. Such an unobtrusive monitoring app can be provided by video services or network operators to their customers to gain a full understanding of all QoE factors perceived by the end user. Additionally, subjective QoE studies can be conducted with the app, which makes it a valuable tool for researching the QoE of HAS.

The monitored information is beneficial as input to QoE-aware traffic management solutions. This thesis investigated different resource allocation strategies for video streaming and web browsing flows on a shared bottleneck link. This scenario relates to access networks where multiple users and applications typically share a single backhaul link. Analytical and simulative performance evaluations were conducted and identified the benefits of QoE-aware traffic management and the trade-offs between the different resource allocation strategies. As a joint QoE improvement was possible for both applications compared to the best effort scenario without traffic management, ISPs have to invest into identifying the right traffic management strategy for the specific application mix in their networks. Moreover, it was found that dynamically considering application-layer information for traffic management has advantages over pure network-based strategies. This further confirms the potential of collaborative and cross-layer approaches for QoE-aware traffic management, such as application-aware networks and network-aware applications. Given these findings, network operators and service providers have to agree on standard interfaces to exchange application and network information, which allows for an improved QoE-aware traffic management.

Considering the end user, his shared resources, and ubiquitous social information about users, their interests, or their interactions with other users can further enhance QoE-aware traffic management. The concept of socially-aware traffic management (SATM) was introduced and the most important stakeholders were identified. These stakeholders can collaborate to deliver a personalized service to the end user to reach a high satisfaction with the service. Therefore, all members of the video service delivery chain should identify and bring together the relevant stakeholders for implementing SATM. In particular, they have to investigate what social signals they can harvest and how they can benefit from social information. Then again, standard interfaces have to be defined to exchange social information among the different stakeholders for improved services and traffic management.

Due to increasing user mobility, which puts huge loads on cellular networks, growing public Wi-Fi infrastructure, and the prevalence of residential Wi-Fi networks in home environments, Wi-Fi offloading was considered as a SATM mechanism to improve the QoE of HAS. To evaluate the potential for public Wi-Fi offloading, the distribution of Wi-Fi hotspot locations was modeled. The developed simple model reached a high accuracy in terms of offloading potential, coverage, or signal strength. It could be successfully applied in a real performance evaluation study, in which the generated hotspot distribution gave similar results to the original hotspot distribution. Based on this simple model, Wi-Fi hotspot location distributions can be generated in arbitrary cities, for which real hotspot locations are not available. This allows to design and evaluate the performance of new SATM mechanisms based on Wi-Fi offloading. Nevertheless, the simple model had limited applicability in case the spatial collocation of hotspots has to be reconstructed. This thesis also presented a trust-based SATM framework for home routers, which allows to improve the QoE of HAS by offering Wi-Fi offloading, caching, prefetching, and content delivery. As home routers are typically provided by ISPs to their customers, this SATM platform can be easily deployed to share private Wi-Fi networks, improve traffic management, and reach a higher QoE. Finally, the performance of Wi-Fi offloading of mobile HAS sessions was investigated in terms of QoE and energy efficiency. It was found that the utilization of Wi-Fi offloading deteriorated the QoE factors and energy consumption of end users due to lower throughput at public Wi-Fi hotspots. However, if a comparable throughput distribution is available at the Wi-Fi hotspots, e.g., in residential environments with fixed broadband connections, both subjectively perceived metrics can be improved. Based on these findings, in order to encourage end users to offload in terms of QoE and energy consumption, the capacity of the available public Wi-Fi infrastructure has to be increased to provide comparable throughput like current mobile networks. Moreover, the sharing of private Wi-Fi infrastructure has to be facilitated and incentivized, e.g., by utilizing the proposed SATM framework for home routers.
REFERENCES


