YouTube QoE Classification Based on Monitoring of Encrypted Traffic

Student: Irena Orsolic
University of Zagreb
Faculty of Electrical Engineering and Computing
Zagreb, Croatia
irena.orsolic@fer.hr

Advisor: Lea Skorin-Kapov
University of Zagreb
Faculty of Electrical Engineering and Computing
Zagreb, Croatia
lea.skorin-kapov@fer.hr

***Abstract*—With the move to traffic encryption adopted by many Over The Top (OTT) providers of video distribution services, Internet Service Providers (ISPs) are now facing the challenges of monitoring application performance and potential end user perceived service quality degradations. With lack of direct feedback from OTT providers, ISPs generally rely on passive traffic monitoring solutions deployed within their network for the purposes of monitoring OTT service performance. The aim of this research is to develop tools and methodologies that enable the estimation of end user QoE when watching YouTube videos using different platforms and access networks, based solely on the analysis of encrypted network traffic.**

I. Methodology and current results

Prior to the widespread use of encryption in OTT traffic, network operators could gain insight into application performance by extracting packet header information. Today, the inability to monitor service performance at an application level poses a threat to network providers, as they are potentially unable to detect problems and act accordingly. Poor performance further imposes the risk of losing customers, as customers often tend to blame network providers for poor QoE. Given the current situation, ISPs commonly rely on passive traffic monitoring solutions deployed within their network to obtain insight into user perceived degradations and their potential causes.

To address these challenges faced by ISPs, we have been studying the feasibility of estimating YouTube QoE based on monitoring of encrypted network traffic, by using machine learning (ML) techniques. To do so, we are developing a system called **YouQ**. YouQ enables application- and network-level data collection, data processing and building ML models that can subsequently be used to estimate QoE of new YouTube streaming sessions based solely on network traffic features. To develop such a system, there is a need to understand how application Key Performance Indicators (KPIs) calculated from application-level data (such as stalling duration, initial delay, etc.) affect end users' QoE (QoE modelling problem)[1]. Another challenge lies in extracting the network traffic features that can be correlated to application-level degradations [2][3].

This research is conducted in the scope of projects "Survey and analysis of monitoring solutions for YouTube network traffic and application layer KPIs" and "QoMoVid: QoE Monitoring Solutions for Mobile OTT Video Streaming", both funded by Ericsson Nikola Tesla, Croatia. This work has also been supported in part by the Croatian Science Foundation under the project UIP-2014-09-5605 (Q-MANIC). The authors would like to thank Illona Popic, Petra Rebernjak and Ivan Bartolec for their help in running experiments.

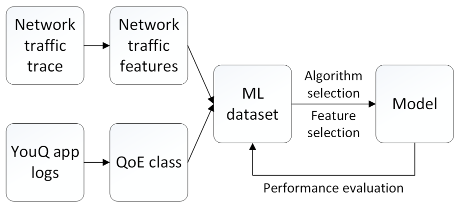

Fig. 1. Approach for QoE classification based on network traffic features

Our developed YouQ system consists of an Android application that plays a requested number of YouTube videos and logs events at an application level (e.g., video buffering, quality switch), amount of video in the buffer, and URLs from all HTTP requests. This data is collected using the YouTube IFrame API. In parallel to logging of application-layer KPIs, network traffic is also captured. After all the videos are played, both application-level logs and network traffic are uploaded to a YouQ server and processed. Processing includes calculating traffic features (e.g., average throughput, average interarrival time) for each of the videos in the experiment, calculating application-level KPIs from the collected logs (e.g., percentage of time spent on each quality level, stalling duration), and assigning a *QoE class* ("high", "medium", "low" QoE) to each video according to calculated KPIs and QoE models defined in [4]. The output of this phase is a dataset for training ML models. For each viewed video, we extract 33 traffic features (listed in [4]) based on the analysis of encrypted traffic, and classify the video into one of the three aforementioned QoE classes. The approach is depicted in Figure 1.

Our approach was tested in a laboratory environment shown in Figure 2. YouTube traffic between an Android smartphone (Samsung S6 with Android version 5.1.1) and YouTube servers is transmitted over an IEEE 802.11n wireless network and then routed through a PC running IMUNES (www.imunes.net), a general purpose IP network emulation/simulation tool enabling a test administrator to set up different bandwidth limitations and schedule bandwidth changes. Traffic is further sent through Albedo's Net.Shark device (a portable network tap used for aggregating and mirroring network traffic) where it

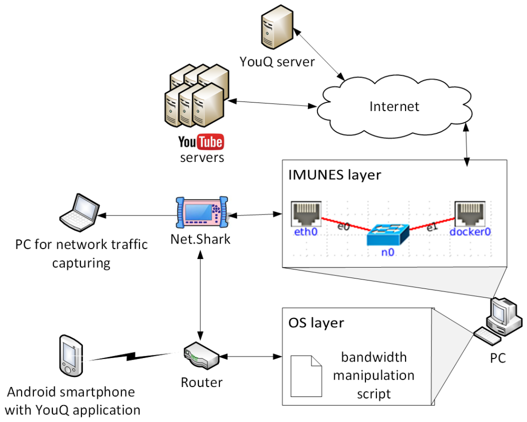

Fig. 2. YouQ lab setup

is replicated and sent to a PC designated for network traffic capturing. The PC running IMUNES also has an OS layer, accessed by the YouQ client application to run a bandwidth scheduling script according to defined experiments. The script resets the bandwidth envelope for each video in the experiment, which enables all videos to be played under exactly the same network conditions.

We collected a dataset corresponding to 1060 videos played under 39 differently defined bandwidth conditions, and trained models by using various ML algorithms. The models proved to be up to 84% accurate when 3 QoE classes were defined ("high", "medium", "low"). We also classified videos into 2 QoE classes ("high" and "low") and repeated the procedure. These models achieved an accuracy of 91%. The exact measurement procedure, definition of QoE classes, statistics of the collected dataset, along with a more detailed view of the results were published in [5], while [4] gives an even more detailed view of the test methodology used to train ML models for QoE classification, and provides a more detailed interpretation of classification results.

II. Ongoing Activities and Challenges

Our current activities aim to further develop the YouQ system so as to enable data collection, processing, and the training of QoE classification models for different usage scenarios. By this we refer to scenarios in which YouTube is delivered over different access networks (WiFi and mobile), using different clients (browser-based vs. YouTube App), using different transport protocols (TCP, QUIC), and with different types of user behaviour observed. As a standard for estimating QoE of adaptive streamed media has recently been published in ITU-T recommendation P.1203 [6], we plan to incorporate the defined model within the YouQ system instead of the simplified multidimensional model we are currently using.

A. Different client applications

Besides the YouTube IFrame API, that we used to develop the initial YouQ client application, YouTube also offers a native Android API. We have developed a version of the YouQ Android application based on this API, which enables us to observe YouTube KPIs and traffic behaviour in the case when a user access YouTube via the dedicated YouTube App. We noticed that in this case QUIC protocol [7] is used, as opposed to TCP, that was used in the IFrame case. Further comparison of the two applications also showed the differences in YouTube's adaptation algorithm. Considering these differences, models built using the data from the IFrame case are not applicable in this one, and new models should be built for this scenario. Another activity we plan, regarding the client applications, is building the YouQ application for iOS.

B. Different transport protocols

Using as a client device Samsung S6, we found that in all access network cases (WiFi, 3G, 4G) and using both the Chrome browser (version 55.0.2883.91) and the YouTube app (version 12.01.55), QUIC was used as the default protocol (Feb. 2017). As QUIC was formally proposed to the IETF as an Internet standard [7], it seems highly likely that QUIC will in the future be the base for YouTube functionality. Therefore, we plan to do an in depth study of the QUIC protocol and its impact on performance of the YouTube service. We already defined a set of QUIC traffic features that can be correlated to QoE, incorporated the calculation of these features into YouQ, and plan on running experiments to collect data and build classification models.

C. User behaviour

Finally, we are investigating different aspects of user behaviour and their effects on the service. We plan on running experiments and collecting training data involving different user interactions (e.g., using a playlist, autoplay, manually browsing through videos, seeking forward/backward, etc.), to determine the implications with respect to developing ML-based QoE classification models.

References

[1] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hoßfeld, and P. Tran-Gia, "A Survey on Quality of Experience of HTTP Adaptive Streaming," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 469–492, 2015.

[2] G. Dimopoulos, I. Leontiadis, P. Barlet-Ros, and K. Papagiannaki, "Measuring Video QoE from Encrypted Traffic," in *Proceedings of the 2016 ACM on Internet Measurement Conference*. ACM, 2016, pp. 513–526.

[3] V. Aggarwal, E. Halepovic, J. Pang, S. Venkataraman, and H. Yan, "Prometheus: Toward Quality of Experience Estimation for Mobile Apps from Passive Network Measurements," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*. ACM, 2014, p. 18.

[4] I. Orsolic, D. Pevec, M. Suznjevic, and L. Skorin-Kapov, "A Machine Learning Approach to Classifying YouTube QoE Based on Encrypted Network Traffic," *Multimedia tools and applications*, 2017, doi:10.1007/s11042-017-4728-4.

[5] I. Orsolic, D. Pevec, M. Suznjevic, and L. Skorin-Kapov, "YouTube QoE Estimation Based on the Analysis of Encrypted Network Traffic Using Machine Learning," in *Globecom Workshops (GC Wkshps), 2016 IEEE*. IEEE, 2016, pp. 1–6.

[6] T. standarization sector of ITU, "Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport," International Telecommunication Union, Tech. Rep. P.1203, nov 2016.

[7] R. Hamilton, J. Iyengar, I. Swett, and A. Wilk, "QUIC: A UDP-based secure and reliable transport for HTTP/2," *IETF, draft-tsvwg-quic-protocol-02*, 2016.