January 2014

QoE-based network-centric resource allocation for on-demand uplink adaptive HTTP streaming over LTE network

Farhan Pervez
COMSATS Institute of Information Technology

Muhammad Salman Raheel
University of Wollongong, msr949@uowmail.edu.au

Follow this and additional works at: https://ro.uow.edu.au/eispapers

Recommended Citation
Pervez, Farhan and Raheel, Muhammad Salman, "QoE-based network-centric resource allocation for on-demand uplink adaptive HTTP streaming over LTE network" (2014). Faculty of Engineering and Information Sciences - Papers: Part A. 4723.
https://ro.uow.edu.au/eispapers/4723

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
QoE-based network-centric resource allocation for on-demand uplink adaptive HTTP streaming over LTE network

Abstract
With the advent of community portals and video sharing websites, production and consumption of multimedia content has increased tremendously. It is because of excellent video recording capabilities of cameras available with mobile devices, and enhanced data rates offered by cellular networks, that the uploading of high quality multimedia content is more realistic. Nowadays most of the video traffic is HTTP-based, that provides rate adaptive streaming, known as Dynamic Adaptive Streaming over HTTP (DASH). In this work we generate DASH video streams that can be supported on LTE uplink and then assign resources to the video producers such that the QoE of on-demand viewers is maximized.

Keywords
allocation, demand, uplink, adaptive, qoe, http, lte, streaming, network, over, centric, resource

Publication Details

This conference paper is available at Research Online: https://ro.uow.edu.au/eispapers/4723
QoE-based Network-Centric Resource Allocation for On-demand Uplink Adaptive HTTP Streaming over LTE Network

First Author: Farhan Pervez
Department of Electrical Engineering
COMSATS Institute of Information Technology
Lahore, Pakistan
farhanpervez@ciitlahore.edu.pk

Second Author: Muhammad Salman Raheel
Department of Engineering
University of Wollongong
New South Wales, Australia
msr949@uowmail.edu.au

Abstract—With the advent of community portals and video sharing websites, production and consumption of multimedia content has increased tremendously. It is because of excellent video recording capabilities of cameras available with mobile devices, and enhanced data rates offered by cellular networks, that the uploading of high quality multimedia content is more realistic. Nowadays most of the video traffic is HTTP-based, that provides rate adaptive streaming, known as Dynamic Adaptive Streaming over HTTP (DASH). In this work we generate DASH video streams that can be supported on LTE uplink and then assign resources to the video producers such that the QoE of on-demand viewers is maximized.

Keywords—DASH (Dynamic Adaptive Streaming over HTTP), QoE (Quality of Experience), resource management, Uplink

I. INTRODUCTION

4th generation (4G) mobile networks based on 3GPP LTE [1] provide high data rates for mobile users which enable them to use high quality Internet services. According to traffic and services forecasts, mobile video traffic in 2017 will exceed 66% of all generated traffic which will increase 26-fold compared to 2012 [2]. A large amount of this video traffic will be user generated. Some of the applications that are source for user generated content (UGC) are for example a car on a road monitoring its surroundings and sharing the content through mobile radio network, a camera used in an operating room for haptics communication between a doctor and his patient, and a mobile user recording and uploading videos of a live football match. A large share of video traffic nowadays is HTTP-based [3], that provides dynamic rate adaptive streaming. There has been efforts, largely by 3rd Generation Partnership Project (3GPP), to standardize this HTTP-based streaming for mobile networks, known as Dynamic Adaptive Streaming over HTTP (3GP-DASH). In the near future it is expected that this type of dynamic adaptive streaming will also be a large share of the uplink traffic with LTE networks providing the data rates to support it. In this paper we first take into account the challenges that can be faced for the video producers and generate DASH video streams that can be supported on LTE uplink air interface. Then, we design an algorithm that allocates physical resources to the video producers such that the QoE for the on-demand viewers is maximized. Simulation results are presented for LTE uplink resource distribution that demonstrate the significance of the proposed optimization on the overall user experience. To the best of knowledge this work is one among the first in the direction of assigning optimized resources to DASH video producers over uplink.

The rest of the paper is organized as follows. In the next section, we first generate DASH video streams using a suitable video encoding scheme for uplink and then provide a discussion of related work. In Section III, we present our network-centric system. Section IV describes the LTE uplink model and the resource allocation framework for LTE uplink. Section V presents the DASH rate adaptation algorithm along with simulation results and Section VI concludes the paper.

II. VIDEO ENCODING FOR UPLINK DASH

A. Generating DASH Video Streams

According to the principle of the 3GP-DASH specification [4], the media is generated in video segments of specific duration at different bit rates. The client then HTTP request's the generated segments of different bit rates according to its network conditions. There are different ways of providing multiple bitrate representations of the media segments that need to be send from the producer to the viewer. One method is Advanced
Video Coding (H.264/AVC) [5], in which a segment is completely and independently encoded at different bitrates. This can be done using a single layer codec. Another method is by encoding a segment in multiple layers by using Scalable Video Coding (SVC) [6]. This allows to store layers of the video as different representations which are additive to each other. In respect to AVC this type of encoding is referred as SVC extension of H.264/AVC.

Largely, the segments that are available at different bit rates are generated using H.264/AVC. As H.264/AVC results in a single-layer so this means that the encoded bit stream does not provide several spatial resolutions. It neither provides several bit-rates for a specific spatio-temporal resolution. On the other hand, SVC provides subset lower quality bit streams from a high quality video bit stream. This is done by dropping packets from the larger video so that the bandwidth required to send the subset bit stream is lower. These subset bit streams are referred as layers, and generally there is a Base Layer (BL) and one or more than one Enhancement Layers (EL).

In this work as we are dealing with UGC, the computation required to generate media segments for DASH remains an issue. Due to the amount of computation power required and limited memory of the UE, it is not practical to use H.264/AVC to generate multiple single layer streams of different bit rates and to store them at the UE. In comparison SVC provides the benefit of reduced computation power and less memory requirement to store segments.

Figure 1 compares an example of SVC and H.264/AVC in terms of storage requirement. Figure 1 compares an example of SVC and H.264/AVC in terms of storage required. It shows, for AVC, a single segment of one second of video length needs a total of (100 + 200 + 300) Kbits of storage. The three layers that provide different qualities are named as base layer, enhancement layer 1 and enhancement layer 2. If the client requests for lowest quality then only base layer (100 Kbits) is transferred. Whereas, if it requests for a better quality then enhancement layer 1 (100 Kbits + 200 Kbits), i.e. 200 Kbits in total, is transferred. In this way multiple streams at different bit rates are accessible, minimizing the total computation required at the UE and its battery utilization. Therefore, the selected video coding scheme for the implementation done in this work is SVC.

We decompose a video stream into three video layers where the base layer (BL), encoded at 15 fps, provides a basic video quality and the remaining enhancement layers (EL1 and EL2), encoded at 30 fps, provide a refined video quality. Figure 2 shows the Rate-MOS utility curves generated for different test videos we use in our simulation.

B. Related Work

Research work on resource allocation for DASH video streams has largely been focused on downlink. For example in [7] Wei Pu et. al presented a wireless proxy approach for DASH, WiDASH proxy, in order to maximize the viewers QoE. WiDASH proxy locates at the edge between Internet and wireless core networks. Different from conventional DASH whose rate adaptation logics are either implemented locally in the user equipment or in the DASH server, WiDASH proxy is in charge of video adaptation, which makes it feasible to perform global optimization over multiple concurrent DASH flows. Another example of resource allocation for DASH downlink streams is proposed in [8] in which an over the top (OTT) approach is used that requires no adaptation of the media content. The advantage of the proposed scheme is obvious in terms of gain in QoE in comparison to both reactive QoE-optimized and to standard-DASH HTTP streaming.
For uplink user generated content, but that only for UDP/RTP based streaming, some work has been done by El Essaili et. al in [9] in which the proposed approach assigns more resources for popular contents while maintaining a minimum guaranteed QoE for the less popular ones. The results show that a significant gain in terms of average user satisfaction can be achieved.

III. NETWORK-CENTRIC APPROACH

SVC decomposes a video stream into three video layers where the base layer (BL) provides a basic video quality and the remaining enhancement layers (EL1 and EL2) provide a refined video quality. We consider time-shifted up streaming of scalable video from a video producer to a video portal which acts as an intermediate node for archiving of video streams for on-demand retrieval. The uplink UEs are the video producers. They have the SVC streams generated that are available for being transmitted over the radio access network. The eNodeB acts as a centralized resource allocation entity that decides on the resources to be provided to the video producers. The video content that is uploaded by the video producers is first stored on a video portal. This uploaded content is then accessed, from the video portal, by the viewers.

A schematic overview of a system for uplink distribution of user-generated content, from a set of video producers to a video portal, to be viewed on-demand is illustrated in figure 3.

![Figure 3: System Image for QoE-driven Uplink Resource Allocation](image)

IV. QoE OPTIMIZED RESOURCE ALLOCATION

A. LTE Uplink Model

The link layer model used for uplink is recommended by 3GPP LTE in Release 8 [10]. It determines the achievable throughput per PRB for a given Signal-to-Noise ratio ($\gamma$). The model approximates the throughput in the uplink, after link adaptation and hybrid automatic repeat request, by an implementation loss $\beta = 0.4$ compared to the Shannon capacity, as defined in equation 1. It further defines range of uplink parameters with $\gamma_{\text{min}} = -10\text{dB}$, $\gamma_{\text{max}} = -15\text{dB}$ and a maximum throughput ($\text{Thr}_{\text{max}}$) of 2bps/Hz.

$$\text{Thr} = \begin{cases} 0 & \text{for } \gamma < \gamma_{\text{min}} \\ \beta \log_2(1 + \gamma) & \text{for } \gamma_{\text{min}} \leq \gamma \leq \gamma_{\text{max}} \\ \text{Thr}_{\text{max}} & \text{for } \gamma > \gamma_{\text{max}} \end{cases}$$ (1)

B. Resource Allocation Framework

As we are working with DASH videos that are encoded using scalable video coding, the number of different video quality representations for the streams are limited. In our case we only have three SVC layers, that means we have less flexibility when trying to optimize the resource distribution between the uplink video producers in order to maximize the MOS of the viewers. We introduce a gradient based resource allocation algorithm that provides discrete optimization which is more suitable to limited number of data rate representations. The objective function is defined in equation 2.

$$\text{arg max}_{\alpha_1, \alpha_2, \ldots, \alpha_K} \sum_{k=1}^{K} U_k(\alpha_k)$$

subject to

$$\sum_{k=1}^{K} \alpha_k <= 1$$ (2)

The objective of the uplink resource allocation is to determine the resource share $\alpha_k$ of each user that maximizes the overall QoE. For the given resource share, the utility function of user $k$ is denoted by $U_k$. The inequality in the constraint is due to the limited number of SVC layers. As we only use three Rate-MOS representations available on the utility curve in order to achieve the objective function defined, we may end up with few physical resource blocks not assigned if they are not enough to obtain the next higher rate representation.

The gradient based resource allocation algorithm, determines the resource share of every video producer. It starts by first assigning the resources, to all video producers, that are enough to transmit their base layers. After that, in order to send the enhancement layer, the algorithm then assigns the resources to the producer whose video provides the maximum gradient on the utility curve. Here the gradient is defined as the increase in utility with respect to the amount of resources required for that increase. The algorithm runs assigning resources to the next video producer that has the maximum gradient on the utility curve. It stops if all resources are utilized or when EL2 of all video producers is transmitted or if the remaining resources are not sufficient to get the next higher layer of any of the videos.
V. SIMULATION RESULTS

A. Scenario

The simulation model that has been implemented for on-demand uplink DASH was described in section III. Here we assume that there is no delay or congestion experienced by the packets over the core network. Furthermore, we consider a single LTE cell with one eNodeB. The number of video producers uploading the video content is 8 and so as the number of viewers. All the video producers upload different videos that are available in three different qualities (base layer, enhancement layer 1, enhancement layer 2). The number of optimization cycles is 100 and the total simulation time is 100 seconds.

B. Rate Adaptation Algorithm for DASH

Here we also implement a receiver-driven rate adaptation algorithm, as defined in [11], which gives a more realistic QoE perspective of DASH viewers. This algorithm is based on the segment fetch time, and not on instantaneous TCP transmission rate, that decides on the data rate representation to be requested for the next DASH segment. The client measures the segment fetch time to determine if the bitrate of the current representation matches the available end-to-end bandwidth capacity. The segment fetch time (SFT) denotes a period of time from the time instant of sending a GET request for a media segment to the instant of receiving the last bit of the requested media segment. The client switches up the request rate if SFT is less than or equal to 1 segment, or in this case 1 sec when the segment duration is itself 1 sec, and buffered media time is larger than the predefined minimum. Whereas the switch down takes place if SFT is greater than 1 segment. There is a complementary switch down if buffered media time is lower than the predefined minimum to prevent client buffer under flow. From here we state this rate adaptation as DASH SFT.

C. On-demand Simulation Results

In this section we evaluate the QoE-driven gradient based algorithm, with a classical round robin (RR) scheduling. Figure 4 shows the comparison between the average MOS of QoE-driven algorithm and RR scheduling. The pink curve shows the theoretical QoE optimized average MOS using the gradient algorithm, whereas the black curve shows the theoretical average MOS achieved by using RR scheduling. We also analyze the SFT-based rate adaptation algorithm here which gives a more realistic QoE perspective of DASH viewers. The blue curve shows the average MOS of the viewers for QoE optimized SFT scenario, whereas the red one shows the average MOS of the viewers for RR scheduling based SFT scenario.

The results show the advantage of our QoE scheme in comparison to RR scheduling, both theoretically and in case of a more practical rate adaptation (SFT-based).

Figure 4: Average MOS for QoE gradient-based algorithm and RR scheduling

Figure 5 shows the comparison between the cumulative distribution function (CDF) for mean MOS of the QoE gradient-based algorithm and RR scheduling.

Figure 5: CDF for mean MOS of QoE gradient-based algorithm and RR scheduling

It confirms the advantage in terms of average MOS of our QoE-based resource allocation in comparison to RR scheduling. As we can see, the gain in average MOS is 0.5 for both theoretical and in case of rate adaptation (SFT-based) perspectives. In this section the number of scenarios used to
generate the CDF is five, where the video producers experience different channel qualities in every different scenario.

Figure 6a shows the comparison between a SFT-based QoE-based scenario and a SFT-based RR scheduling scenario in terms of total number of buffered video segments at the viewers’ end. It shows the buffer build-up for RR scheduling scenario although there is no gain in average MOS, hence unnecessarily increasing the uplink network load. Along with it, for the same scenario, figure 6b shows the average channel quality for all the uplink video producers. We can see that the total buffer depends on the uplink channel quality of the video producers.

VI. CONCLUSION

We focused on optimized resource allocation for transmitting multiple DASH video streams in a LTE cell in the uplink direction. First, taking into account the challenges that can be faced for the video producers, DASH video streams that can be supported on LTE uplink air interface were generated. Then, an algorithm was designed that allocates physical resources to the video producers such that the QoE for the on-demand viewers is maximized. Simulation results were presented for LTE uplink resource distribution that demonstrate the significance of the proposed optimization on the overall user experience.

In future we look forward to implement a de-centralized approach for the network-based QoE optimization that can reduce exchange of information between the video producers and the eNodeB and thus reduce the overall optimization complexity.

REFERENCES