

From 2D to Next Generation VR/AR Videos: Enabling Efficient Streaming via QoE-aware Mobile Networks

Roberto Irajá Tavares da Costa Filho^{*}, Filip De Turck[†], Luciano Paschoal Gaspar[‡]

^{*}Federal Institute of Education, Science and Technology - IFSul

[†]Ghent University - imec

[‡]Federal University of Rio Grande do Sul - UFRGS

^{*}robertocosta@charqueadas.ifsul.edu.br, [†]filip.deturck@ugent.be, [‡]paschoal@inf.ufrgs.br

Abstract—Ranging from traditional video streaming to Virtual Reality (VR) videos, the demand for video applications to mobile devices is booming. In the context of mobile operators a challenging problem is how to handle the increasing video traffic while managing the interplay between infrastructure optimization and QoE. Solving this issue is remarkably difficult, and recent investigations do not consider large-scale networks. In this dissertation paper we explore the solution space of efficient video streaming over mobile networks. First, we propose a model to predict video streaming quality based on the observation of performance indicators of the underlying IP network. Second, we introduce a novel QoE-aware path deployment heuristic for large-scale SDN-based mobile networks. Third, based on the lessons learned with QoE prediction for traditional video streaming, we finally explore the VR video domain by proposing PERCEIVE and VR-EXP. PERCEIVE is a two-stage method for predicting the perceived quality of adaptive VR videos when streamed through mobile networks. In turn, VR-EXP consists of an experimentation platform that allows in-depth evaluation of state-of-the-art VR video optimization techniques. Obtained results show that the combination of the proposed methods for QoE-aware path selection outperformed state-of-the-art approaches.

Index Terms—Quality of Service, Quality of Experience, Video Streaming, Virtual Reality, Path Selection, Mobile Networks.

I. INTRODUCTION

The future is mobile: wireless networks will account for more than two-thirds of all IP traffic by 2020 [1]. In this context, operators are being challenged by video traffic, which is pushing their network infrastructure to the limit [2]. According to Cisco, mobile video accounted for 60% of total Internet traffic in 2016. And there is more to come, since mobile video is expected to increase 9-fold by 2021, reaching 78% of total data traffic [3]. According to the same source, Virtual Reality (VR) videos will significantly increase this challenge as the traffic generated by this application is expected to increase 12-fold by 2022 [3].

VR video streaming applications are challenging due to three main reasons: (i) they will run mostly over mobile networks [3]; (ii) mobile networks are characterized by highly variable levels of performance [4]; and (iii) VR video streaming applications demand high levels of network performance to achieve a satisfactory QoE [3]. To provide a notion of how

demanding these applications are, recent studies have shown that, to provide adequate levels of QoE, current VR video applications require a network delay lower than 9 ms [5], while the bandwidth needs for the upcoming ultra high definition VR will reach 500 Mbps [3]. At this level of demand, not only will network operators struggle to provide cost-effective services, but VR video content providers and developers will also be challenged by such resource-intensive applications.

To deal with the huge growth in data traffic, mobile operators have to constantly invest (*i.e.*, CAPEX and OPEX) to increase capacity, to switch technology (*e.g.*, 3G, 4G, 4G+, 5G), as well as to improve outdoor and indoor coverage. In the opposite direction, the Average Revenue Per User (ARPU) for mobile broadband has fallen from USD 23 in 2013 to USD 13 in 2015 [6]. All these elements together place a lot of pressure on operators to manage their infrastructure efficiently [2].

Aiming at increasing efficiency, mobile operators have been relying on offloading technologies such as Small Cells (Femtocell, Picocell), Wi-Fi offloading, Content Delivery Networks (CDNs), and, in the near future, 5G Device-to-Device communication (D2D) and 5G Mobile Edge Computing [7]. These technologies play a fundamental role in the network infrastructure optimization as they can take advantage of network path diversity to distribute traffic across multiple network segments (*i.e.*, edge, aggregation, core and peering). Such approaches have the ability to shorten the distance between subscriber and content while avoiding network congestion by spreading the traffic among alternative paths. As an indication of how important these offloading techniques are, only in 2016, 60% of mobile data was relocated to alternative paths, just considering Wi-Fi and Femtocell offloading [3].

The adoption of offloading techniques introduces a multitude of possible paths through which user traffic can be forwarded and, as an immediate consequence, raises the complexity of the network management (*e.g.*, path selection, configuration and troubleshooting). While very important, such an advanced infrastructure does not directly translate into improved QoE [8]. This is notably true if considering that some offloading techniques may rely on shared and third-party infrastructure, which would possibly exacerbate the unpredictability regarding the delivered QoE.

A. Research Questions

The objective of the Ph.D. thesis was to answer the following two main research questions.

Research Question 1. Considering that mobile operators already have tools in place to measure network performance, would it be possible to employ monitored network indicators to predict playout performance and QoE for both traditional 2D and VR video streaming applications?

Research Question 2. How to employ QoE prediction to dynamically select and deploy paths that maximize QoE and minimize infrastructure utilization over time?

B. Research Goals

To be able to answer the research questions above, the thesis was organized around five objectives: (i) devise a prediction model capable of estimating video streaming playout performance and QoE based on available network information; (ii) formalize the QoE-aware path selection problem; (iii) formulate a QoE-aware path selection heuristic with the ability to operate in dense networks; (iv) propose a QoE prediction model for VR videos; and (v) provide an in-depth evaluation of state-of-the-art VR video optimization techniques. These objectives unfolded into five main research contributions, as summarized next.

II. QOE PREDICTION MODEL FOR VIDEO STREAMING

A. Challenge

Both the scientific community and the industry agree that maximizing the user's Quality of Experience (QoE) regarding video streaming applications represents a relevant research challenge [9]. An essential aspect in this direction is to systematically determine the quality of the provided video services. To this end, service providers require a solution with low intrusion, scalability, and a reasonably accurate way to measure the quality of service delivered. This task becomes particularly challenging if encompassing cellular networks, in which highly intrusive measurement techniques have the potential impair quality. In our approach, we characterize intrusiveness as the volume of data injected in the network in order to measure IP performance indicators (*e.g.*, throughput, packet loss, delay). For further details concerning intrusiveness, interested readers should refer to our paper [4].

B. Our Approach

In this investigation we propose a Lightweight application QoS and QoE Predictor (LEAP). It is capable of providing a detailed view of how the network performance affects video streaming applications and, moreover, the corresponding user experience. The model is designed to receive four network performance indicators as input: (i) delay, (ii) jitter, (iii) throughput and (iv) packet loss. To capture video playback performance, the model predicts three video playout performance indicators: (i) startup time, (ii) stall count and (iii) total stall length. To estimate each application QoS indicator, the four network QoS indicators are combined and given as input to each decision tree. In a second stage, the three application QoS indicators are combined to estimate QoE. The QoE indicator is estimated in the same range of the Mean Opinion Score

(MOS), which is graded from 1 to 5, where 1 means the worst possible experience and 5 means the best one.

LEAP leverages the Regression Decision Tree technique to construct the three application QoS prediction models. It was chosen mainly due to its suitability to handle complex and non-linear relationships between attributes. Additionally, decision trees have three characteristics that are desirable in the network management realm. First, the resulting prediction structure is simple to implement and integrate with third-party network management systems (they require only decision structures). Second, the prediction scheme presents low time complexity, being suitable for large scale environments. Third, decision trees provide decent accuracy. Although a more sophisticated approach for machine learning (*e.g.*, deep learning) might present higher accuracy than decision trees, for this task, the accuracy provided by decision trees is good enough. More accurate algorithms would not justify increased complexity for implementation and integration.

C. Results

The proposed approach produces accurate estimates (average error of less than 10%) while keeping intrusiveness around twenty times lower than traditional techniques. The low intrusiveness allows the service provider to configure systematic measurements, with reduced polling interval, without an excessive usage of network resources. Figures 1(a) and 1(b) depict the variation of r_i (residual error for the sample i) in the horizontal axis, associating each value in this axis to a portion of the sampling group (in the vertical axis) in which r_i is lesser or equal to the set threshold. The normalized residuals r_i are defined by the equation $r_i = |\hat{x}_i - x_i|/N$, where x_i represents the observed values and $\hat{x}_i$ the predicted ones. We use the factor N to normalize the error values. For each of the three application indicators, the value of N is derived from the duration of the videos, which enables a generalization of the evaluation method for videos of any duration. The average of r_i in 90% of the samples for all indicators, considering both 1080p and 720p, has a value of $\bar{r}_i = 0.0982$ (9.8%). We deem this error rate to be satisfactory and in line with recent research work.

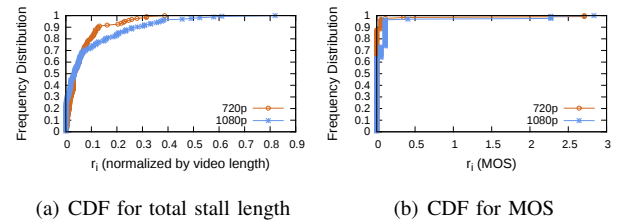


Fig. 1. Stall length and QoE prediction error

III. DYNAMIC QOE-AWARE PATH SELECTION

A. Challenge

A challenging task for mobile operators consists in how to handle the increasing end-user traffic while optimizing infrastructure utilization and managing user's QoE. In fact, from the operator's perspective, addressing this challenge

is crucial for improving competitiveness, since the effective management of the interplay between perceived QoE and infrastructure investments is the main factor for increasing the return of investment [10]. Given the context above, the main research challenge is how to take advantage of the path diversity introduced by upcoming technologies (*e.g.*, 5G D2D, Edge Computing, and Fog Computing) to dynamically select QoE-aware paths capable of delivering cost-effective video applications.

B. Our Approach

In a nutshell, the QoE-aware path selection task can be decomposed into two challenging problems. The first problem consists of timely predicting QoE for network paths. In turn, the second problem encompasses the large-scale path selection algorithm, which should be able to take constraints (*e.g.*, target QoE and available network resources) into account and select optimized paths. The first problem is complex because the information that is closely related to QoE (*e.g.*, subjective evaluations and objective measurements) are not largely available or feasible to obtain in a systematic approach for large-scale networks. Regarding the second problem, it can be easily solved considering small deployments. However, it becomes notably complex when combined with additional constraints (*e.g.*, resource utilization) and applied to ultra-dense networks.

In this investigation we introduce SQAPE, a novel scheme for network path provisioning that employs active QoS measurements to predict video streaming performance and QoE, which will be used for deployment of QoE-aware paths in an SDN-enabled mobile network. This investigation contributes significantly to the state of the art since, unlike the related work, it does not rely on third-party information. Our approach considers neither video streaming flow data nor client/server-side information. The proposal is able to provide QoE-aware paths, in networks characterized by high path diversity, with high accuracy and low network resource consumption.

In our first contribution we formalize the QoE-aware path selection problem. As for the second contribution, we address the QoE-aware path selection problem by proposing a polynomial time complexity path selection heuristic. This approach takes advantage of the previously proposed QoE prediction model (LEAP) and introduces a novel algorithm for performing efficient QoE-aware path selection, in Software-defined Networks, based on per-link QoE composition. The proposed method has the ability to perform dynamic path selection by providing a feasible balance between QoE maximization and network resource minimization over time.

MOS-based link composition. We propose an innovative approach to combine QoS measurements and MOS estimation in a link composition scheme. The proposed strategy is able to predict MOS in end-to-end paths composing multiple QoS metrics. We define functions $T_p(\tau)$, $D_p(\tau)$, and $L_p(\tau)$ which represents, respectively, TCP throughput, delay and packet loss.

Given the inputs of the model, that is, the infrastructure G with its associated metrics and a set S of video connection requests, the problem consists of finding a valid simple path

for each $(i, c) \in S$ in given time slot t so as to *maximize* observed MOS. Therefore, the output of the model is a set of paths that maximize the overall observed MOS given a set of video requests S (Equation 1).

$$\text{Max.} \sum_{\forall (k,l) \in S} \sum_{j \in P(l)} \sum_{\forall \tau \in \Psi(k,j)} \Phi(T_p(\tau), D_p(\tau), L_p(\tau)) \quad (1)$$

Large-scale path selection algorithm. Next, we present a heuristic algorithm capable of finding optimized QoE-aware paths in polynomial time. The proposed algorithm runs on top of an SDN architecture and takes advantage of the northbound interface (NBI) to coordinate related mechanisms such as QoS active measurements, QoE prediction and network rule management.

Algorithm 1 consolidates the proposed path selection approach. As one can observe, the measurement phase collects the results of the latest measurement snapshot and updates the QoS graph for each of its links (line 2). The path selection algorithm operates on top of an adaptation of the Dijkstra shortest-path algorithm in which the distance metric is a combination of MOS estimation and a link utilization contribution. The algorithm starts navigating from node k and determines the path with the best local distance metric to each node $u \in V$ (lines 7-14). For each neighbor v , we estimate MOS based on QoS measurement to path τ (line 10-11) and then combine it with an estimation of residual bandwidth (line 12). The objective of this strategy is to capture the interplay among: (i) path MOS; (ii) path length (hop count); and (iii) residual bandwidth. Our algorithm simultaneously maximizes MOS and residual bandwidth while minimizing path length. We achieve this by means of linear combination (line 12). Parameter α determines the importance of the residual bandwidth in relation to MOS. After navigating through all nodes $u \in V$, SQAPE deploys a path for the video request (k, l) such that the $\chi_{[j]}$ is minimum over all $j \in P(l)$. In other words, it takes the minimum path between k and an offloading node j in $P(l)$. Then, we deploy the appropriate SDN forwarding rules (line 15).

C. Results

To evaluate our approach SQAPE is compared to three other path selection algorithms that attempt to minimize one or more constraints of QoS indicators. The first contender (DKS - Delay) minimizes distance according to the Dijkstra algorithm, using delay as the distance metric [11]. This technique is commonly employed by routing mechanisms such as OSPF. The second contender (BF - BW) maximizes path's available bandwidth bottleneck, using the Bellman-Ford algorithm. This strategy operates in an iterative manner, finding the widest path with the least hops [12]. The third contender is known as the Shortest Widest Path (SWP), which optimally solves the problem of maximizing TCP throughput and minimizing delay in the selected path [12], [13].

The potential of SQAPE can be observed in Figures 2(a) and 2(b), which present the total stall duration and stall count for the considered cases. In such configuration, SQAPE led to 238 seconds of stall, distributed among 72 stall events. The contender solution with the closest performance (SWP) registered 204 stalls, amounting to 998 seconds of duration.

Algorithm 1 QoE-aware path selection

Input: $G = (V, E)$: network infrastructure
Input: $(k, l) \in S$: video content request
Input: $P(l) \subseteq N$: set of available network nodes (caches) with content l
Input: α : path selection factor; β : throughput estimation factor

- 1: **for** every link $(i, j) \in E$ **do**
- 2: gather measured QoS indicator for $T(i, j), D(i, j), L(i, j)$
- 3: **end for**
- 4: **for** every node $u \in V$ **do**
- 5: $\chi_{[u]} \leftarrow \infty, \varphi_{[u]} \leftarrow \text{NILL}$
- 6: **end for**
- 7: $\chi_{[k]} \leftarrow 0$
- 8: **while** $V \neq \emptyset$ **do**
- 9: $u \leftarrow \text{Extract-Min}(V, \chi)$
- 10: **for** each edge (u, v) **do**
- 11: $\{T_\tau, D_\tau, L_\tau\} \leftarrow \text{QoS metrics to path } \tau \text{ from } s \text{ to } v \text{ passing through } u$
 according to equations (1), (2), and (3)
- 12: $M \leftarrow \Phi(T_p(\tau), D_p(\tau), L_p(\tau))$
- 13: $\chi' \leftarrow M + \alpha \cdot \left(\left(\sum_{(i,j) \in \tau_{(k,u)}} \frac{1}{T(i,j)} \right) + \frac{1}{T(u,v)} \right)$
- 14: **if** $\chi_{[v]} > \chi'_{[u]}$ **then**
- 15: $\chi_{[v]} \leftarrow \chi'; \varphi_{[v]} \leftarrow u$
- 16: **end if**
- 17: **end for**
- 18: **end while**
- 19: **deploy** (k, l) based on $\{\chi, \varphi\}$ such that $\chi_{[j]} : j \in P(l)$ is minimum
- 20: **QoS update:** $\forall (i, j) \in \{\chi, \varphi\}$: estimate $T(i, j)$ according to β
- 21: **return** χ, φ

The Dijkstra solution performed the worst, presenting a total of 3,815 seconds during 609 stalls.

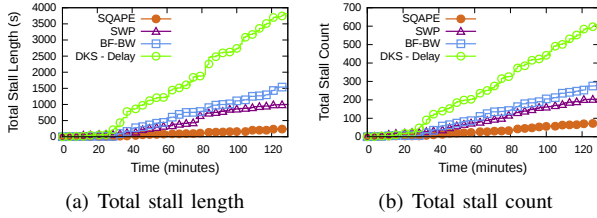


Fig. 2. Video playout performance evaluation using realistic workload and HAS.

Considering real mobile operator topology and video traffic traces, we show that the proposed algorithm outperformed state-of-the-art approaches by reducing impaired videos in aggregate MOS by at least 37% and lowering accumulated video stall length four times.

IV. QoE PREDICTION FOR VR VIDEOS

A. Challenge

QoE has shown to be a critical factor for video applications [10]. As such, both network operators and VR content providers are required to answer an important question: considering the wide range of performance levels of IP networks, to which extent are the currently observable network conditions able to provide users of VR applications with adequate QoE? Answering this question is remarkably complex due to two constraints. First, the influence of the network on VR video performance is unknown; and second, the state-of-the-art on video QoE estimation modeling does not consider the VR context.

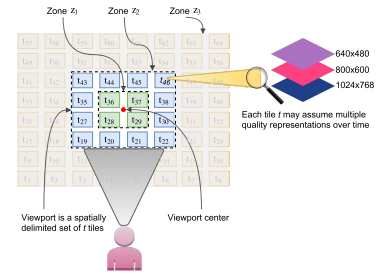


Fig. 3. Example of an adaptive tile-based VR video structure split in 3 zones.

B. Our Approach

Bringing the QoE prediction model to the virtual reality arena, as the fourth contribution we introduce PERCEIVE, a two-stage adaptive VR performance assessment model. It employs regression decision trees to predict VR video play-out performance using network QoS indicators as predictors. Then, it uses the video playout performance metrics to model and estimate the end-user perceived quality.

The employed machine learning strategy is similar to the one described in our first contribution (Section II). However, in order to apply this strategy to the VR ecosystem, we introduce two novel concepts, namely, the concept of visualization zones and a VR video QoE model.

To explain the concept of visualization zones, let us consider that a VR video V can be represented by a set of k spatially divided zones $Z = \{Z_1, \dots, Z_k\}$ such that $\bigcap_{\forall k} Z_k = \emptyset$. The same video V is temporally split into a discrete number of m segments $S = \{S_1, \dots, S_m\}$ such that $\bigcup_{\forall m} S_m = V$. Each zone Z_k is composed of a set of tiles $t \in Z_k$. A tile t is time-divided into m chunks $C = \{C_{t_1}, \dots, C_{t_m}\}$, and may assume different bitrates (qualities) $R(C_{t_m})$ over time. Finally, we refer to a segment as the set of all chunks for a given time frame such as $S_m = \bigcup_{\forall t} C_{t_m}$. In tile-based approaches, the encoding process defines how the video will be spatially divided (*i.e.*, tiling scheme), which bitrates will be available in the HAS context (*i.e.*, quality representations), and the segment length (*i.e.*, number of seconds).

An example of this type of structure is shown in Figure 3. There are three quality zones $Z = \{Z_1, Z_2, Z_3\}$, each one composed by a set of adjacent tiles. Z_1 is a set of tiles adjacent to the viewport center ($t_{28}, t_{29}, t_{36}, t_{37}$), Z_2 is the border of the viewport ($t_{43}, t_{44}, t_{45}, t_{46}, t_{38}, t_{30}, t_{22}, t_{21}, t_{20}, t_{19}, t_{27}, t_{35}$) and Z_3 is composed by all tiles outside the viewport.

The proposed QoE model is defined per zone as a function of the VR playout characteristics predicted by the machine learning procedure within a given zone. This strategy is aligned with the notion that the influence of VR playout characteristics on user perception is different depending on the zone where they are observed (*e.g.*, quality switches for tiles outside the viewport are less important than quality switches inside the viewport). Thus, the per-zone quality function ($\phi(Z_k)$) is defined as the weighted sum of the four playout characteristics (Equation 2).

$$\phi(Z_k) = \underbrace{\sum_{\forall t \in Z_k} \sum_{\forall c \in C(t)} q(R(C_{t_m}))}_{\text{Quality}} - \underbrace{\mu \cdot \sum_{\forall t \in Z_k} \sum_{\forall c \in C_{t_m}} \left(\frac{d_c(R_c)}{C_c} - B_c \right)}_{\text{Stalls}} + \underbrace{- \lambda \cdot \sum_{\forall t \in Z_k} \sum_{\forall c \in C_{t_m}} |q(R(C_{t_m+1})) - q(R(C_{t_m}))|}_{\text{Quality switches}} - \underbrace{\omega \cdot T_s}_{\text{Startup}} \quad (2)$$

In Equation 2, $R(C_{t_m})$ represents the bitrate (*i.e.*, quality) of a given chunk. Recall that a tile t is time-divided into m chunks $C = \{C_{t_1}, \dots, C_{t_m}\}$, (adapted from [14], [15]). Function q is a mapping function that translates the bitrate of chunk C_{t_m} belonging to tile t to the quality perceived by the user (*i.e.*, in terms of bitrate sensitivity). The second term of the Equation is used to track stall time. Stalls can be characterized either by tile (*i.e.*, it is possible to have stall in some tiles and video playout in other, for the same segment) or by segment (*i.e.*, the video will stall until all the tiles for a given segment have enough buffer). To keep the model as general as possible, we consider for each chunk c , that a stall event occurs when the download time $\frac{d_c(R_c)}{C_c}$ is higher than the playout buffer length (B_c) when the chunk download started. Hence, the total stall time is given by $\sum_{c=1}^C \left(\frac{d_c(R_c)}{C_c} - B_c \right)_+$. In addition, $|q(R_{ct+1}) - q(R_{ct})|$ considers the quality switches between consecutive chunks and T_s tracks the startup delay. Finally, constants μ, λ, ω are the non-negative weights used to tune the model for different user importance regarding QoE events. For example, a higher value of μ , with respect to the other weights, means that the user is more likely to be annoyed when video stalls occurs. Consequently, these events should greatly steers the QoE score.

C. Results

The stalling time (Figure 4(a)) shows an error close to 13% for over 90% of the testing samples. In turn, Figure 4(b) depicts the residual error for the QoE estimation. By applying the previously defined QoE model to each sample i , it is possible to estimate QoE for both the predicted playout values and the original ones. Then, the residual error can be calculated. Through this procedure, the QoE estimation error induced by the proposed prediction scheme can be assessed. As shown in Figure 4(b), the QoE estimation presents $r_i \leq 0.03922$ for over 90% of the cases.

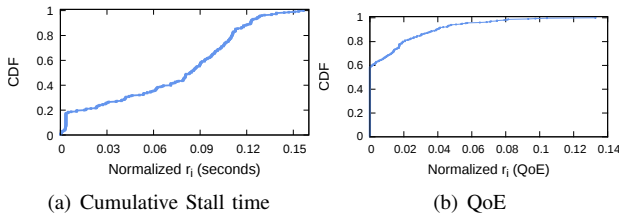


Fig. 4. Residual error CDFs for stall time, startup delay and QoE estimation.

V. VR VIDEO EXPERIMENTATION PLATFORM

A. Challenge

As most of the VR video traffic is expected to be delivered through mobile networks, a major problem arises: both the

network performance and VR video optimization techniques have the potential to influence the video playout performance and the Quality of Experience (QoE). However, the interplay between them is neither trivial nor has it been properly investigated. When considering both the multitude of approaches to optimize a VR video streaming and the highly variable mobile network performance, it becomes a difficult challenge to understand how different (combinations of) optimization techniques perform under varying infrastructure conditions. The lack of a publicly-available method and tools for systematic and reproducible evaluation exacerbate this challenge.

B. Our Approach

To fill in this gap, in this investigation, we propose VR-EXP, an adaptive VR video streaming experimentation platform. The platform is capable of systematically evaluating different combinations of VR video streaming optimization approaches. Also, VR-EXP allows pinpointing the interplay between a set of optimization techniques and variable network performance. Comprised of an evaluation method and software components, VR-EXP assumes as input tile-based VR videos, network datasets, and parameters (e.g., network performance conditions, users' head-tracking information, ABR heuristics, and tile fetching methods). Then, it emulates essential components of the VR video streaming ecosystem, measuring key VR video playout performance indicators. Finally, our platform produces, as output, detailed VR video playout performance and QoE estimation reports. Using VR-EXP, we carry out an in-depth analysis of (combinations of) state-of-the-art VR video optimization approaches under varying network conditions.

In a nutshell, the VR-EXP platform enables evaluating the interplay between a set of adaptive tile-based VR streaming optimizations and variable network performance conditions. The proposed method consists of systematically fetching VR videos through a controlled network environment. From a client perspective, the adaptive VR video client emulator coordinates the use of several VR video techniques upon requesting VR videos from a content server. During the streaming session, a Network Performance Enforcement Point enforces realistic network conditions on the network links between the VR Video Client Emulator and the Content Server. Once the VR video streaming session is finished, VR-EXP reports key VR video playout performance metrics.

C. Results

The results obtained evidence that the relationship between different optimization techniques for video VR optimization is not trivial. Mainly, because certain combinations can benefit one aspect of reproduction and impair others. For example, the increased buffer size, combined with the Full Delivery Basic approach, may lead to increased Viewport Prediction Error (VPE). In this case, the viewport bitrate will be degraded and the stall time will be reduced. By combining an objective assessment of VR video streaming playout performance and a comprehensive QoE model, VR-EXP allowed pinpointing the components of the VR video ecosystem that most affect the performance of VR video playout and, ultimately, QoE.

Figure 5 illustrate key insights provided by VR-EXP. Figure 5(a) demonstrate the impact of viewport prediction errors on QoE. One can observe that severe prediction errors may lead to a decrease of up to 2 points in the QoE score when compared to the baseline case (QoE equal to 5). However, when dealing with realistic performance levels, increasing the playout buffer size may potentially lead to a better QoE score, even considering the likely increase in the viewport prediction error. For example, in Figure 5(b) we can observe that, when using 8 seconds of playout buffer, the worst case scenario for the QoE score (*i.e.*, viewport prediction error of 100%) performs on par with the best case scenario of the 2 seconds buffer (*i.e.*, viewport prediction error of 0%). Furthermore, due to the human randomness, prediction algorithms may present low accuracy even when considering small buffers (*e.g.*, 2 sec). Therefore, using higher values for dimensioning the playout buffer (*e.g.*, 8 sec) will probably outperform smaller buffers setups in most cases.

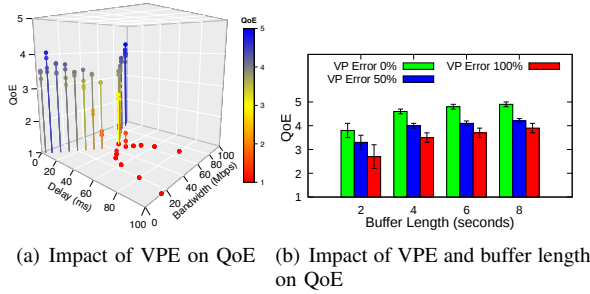


Fig. 5. The complex interplay of VPE, buffer length and QoE.

VI. RESEARCH ACCOMPLISHMENTS AND CONTRIBUTIONS

The main accomplishments and contributions of this research can be summarized as follows:

- Four papers in top-tier IEEE and ACM conferences/journals [4], [5], [16], [17].
- Ph.D. defense awarded with Honors Thesis High Distinction.
- Our papers provide publicly available source code and data set.
- Our VR Video prediction model [5] was awarded with the ACM Reproducibility Badge.
- Best-in-session Presentation Award at INFOCOM 2018.
- Doctoral Stay at Ghent University - imec with joint supervision (Prof. Filip De Turck).

This thesis is publicly available for download from [18].

VII. CONCLUSION

The work conducted throughout this thesis provided us with satisfactory answers to the research questions, namely (i) would it be possible to employ monitored network indicators to predict playout performance and QoE for both traditional 2D and VR video streaming applications? and (ii) How to employ QoE prediction to dynamically select and deploy paths that maximize QoE and minimize infrastructure utilization over time? In the scope of traditional 2D video streaming, research questions 1 and 2 were answered in Sections II and

III, respectively. Finally, in the context of VR video streaming, research question 1 was answered in Sections IV and V.

ACKNOWLEDGMENT

This research was partially funded by CAPES, CNPq, FAPERGS and IFSul.

REFERENCES

- [1] ITU, "Measuring the information society," 2017.
- [2] R. Maallawi, N. Agoulmine, B. Radier, and T. B. Meriem, "A comprehensive survey on offload techniques and management in wireless access and core networks," *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1582–1604, thirdquarter 2015.
- [3] Cisco, "Cisco visual networking index: Forecast and trends, 2017–2022," Cisco Systems, Tech. Rep., 2018.
- [4] R. I. T. da Costa Filho, W. Lautenschlager, N. Kagami, V. Roesler, and L. P. Gaspary, "Network fortune cookie: Using network measurements to predict video streaming performance and qoe," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.
- [5] R. I. T. da Costa Filho, M. Luizelli, M. Vega, J. Hooft, S. Petrangeli, T. Wauters, F. D. Turck, and L. P. Gaspary, "Predicting the performance of virtual reality video streaming in mobile networks," in *Proceedings of the 8th ACM on Multimedia Systems Conference*, ser. MMSys'18. New York, NY, USA: ACM, 2018, pp. 0–0.
- [6] ITU, "Measuring the information society," International Telecommunication Union, Tech. Rep., 2016.
- [7] R. I. Ansari, C. Chrysostomou, S. A. Hassan, M. Guizani, S. Mumtaz, J. Rodriguez, and J. J. P. C. Rodrigues, "5g d2d networks: Techniques, challenges, and future prospects," *IEEE Systems Journal*, pp. 1–15, 2018.
- [8] B. Schlinder, H. Kim, T. Cui, E. Katz-Bassett, H. V. Madhyastha, I. Cunha, J. Quinn, S. Hasan, P. Lapukhov, and H. Zeng, "Engineering egress with edge fabric," in *Proceedings of the 2017 ACM SIGCOMM Conference*. New York, NY, USA: ACM, 2017.
- [9] M. Katsarakis, G. Fortetsanakis, P. Charonyktakis, A. Kostopoulos, and M. Papadopoulou, "On user-centric tools for qoe-based recommendation and real-time analysis of large-scale markets," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 37–43, 2014.
- [10] A. Ahmad, A. Floris, and L. Atzori, "Qoe-centric service delivery: A collaborative approach among otts and isps," *Computer Networks*, vol. 110, pp. 168 – 179, 2016.
- [11] F. Kuipers, P. V. Mieghem, T. Korkmaz, and M. Krunz, "An overview of constraint-based path selection algorithms for qos routing," *IEEE Communications Magazine*, vol. 40, no. 12, pp. 50–55, Dec 2002.
- [12] S. Tomovic, I. Radusinovic, and N. Prasad, "Performance comparison of qos routing algorithms applicable to large-scale sdn networks," in *IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON)*, Sept 2015, pp. 1–6.
- [13] Z. Wang and J. Crowcroft, "Quality-of-service routing for supporting multimedia applications," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 7, pp. 1228–1234, Sep 1996.
- [14] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli, "A control-theoretic approach for dynamic adaptive video streaming over http," *SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 4, pp. 325–338, Aug. 2015.
- [15] H. Mao, R. Netravali, and M. Alizadeh, "Neural adaptive video streaming with pensieve," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, ser. SIGCOMM '17. New York, NY, USA: ACM, 2017, pp. 197–210.
- [16] R. I. T. da Costa Filho, M. Luizelli, W. Lautenschlager, N. Kagami, V. Roesler, and L. P. Gaspary, "Scalable QoE-aware path selection in SDN-based mobile networks," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications (INFOCOM 2018)*, Honolulu, USA, Apr. 2018.
- [17] R. I. T. da Costa Filho, M. Luizelli, M. Vega, J. Hooft, S. Petrangeli, T. Wauters, F. D. Turck, and L. P. Gaspary, "Dissecting the performance of vr video streaming through the vr-exp experimentation platform," *ACM Trans. Multimedia Comput. Commun. Appl.*, 2019, accepted for publication Sep/2019.
- [18] R. I. T. da Costa Filho and L. P. Gaspary, "From 2d to next generation vr/ar videos:enabling efficient streaming via qoe-aware mobile networks," UFRGS, 2019, available at: <http://tiny.cc/2n2yflz>. Accessed on: Nov. 5, 2019.