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

Roberto Irajá Tavares da Costa Filho[1], Filip De Turck[2] and Luciano Paschoal Gaspary[3]

[1]Instituto Federal de Educação, Ciência e Tecnologia (IFSUL)
[2]Ghent University (UGENT)
[3]Universidade Federal do Rio Grande do Sul (UFRGS)

Abstract. *Ranging from traditional 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. Obtained results show that the combination of the proposed methods for QoE-aware path selection outperformed state-of-the-art approaches.*

1. Introduction

The future is mobile: wireless networks will account for more than two-thirds of all IP traffic by 2020 [ITU 2017]. In this context, operators are being challenged by video traffic, which is pushing their network infrastructure to the limit [Maallawi et al. 2015]. 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 [Cisco 2018]. 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 [Cisco 2018].

VR video streaming applications are challenging due to three main reasons: (*i*) they will run mostly over mobile networks [Cisco 2018]; *ii*) mobile networks are characterized by highly variable levels of performance [da Costa Filho et al. 2016a]; and (*iii*) VR video streaming applications demand high levels of network performance to achieve a satisfactory QoE [Cisco 2018]. 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 [da Costa Filho et al. 2018b], while the bandwidth needs for the upcoming ultra high definition VR will reach 500 Mbps [Cisco 2018]. 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 [ITU 2017]. All these elements together place a lot of pressure on operators to manage their infrastructure efficiently [Maallawi et al. 2015]. 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 [Ansari et al. 2018]. 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 [Cisco 2018].

Problem statement. 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). Moreover, such an advanced infrastructure does not directly translate into improved QoE [Schlinker et al. 2017]. 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.

1.1. 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?

1.2. 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 four main research contributions, as summarized next.

2. QoE Prediction Model for Video Streaming

2.1. 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 [Katsarakis et al. 2014]. 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.

2.2. 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 a regression 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.

2.3. 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%).
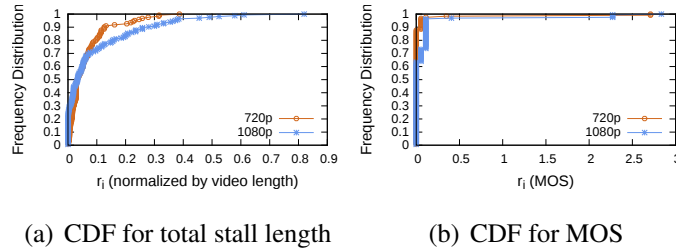

(a) CDF for total stall length   (b) CDF for MOS

Figure 1. Stall length and QoE prediction error

3. Dynamic QoE-aware Path Selection

3.1. 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 [Ahmad et al. 2016]. 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.

3.2. 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. 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.

3.3. 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 [Kuipers et al. 2002]. 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 [Tomovic et al. 2015]. 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 [Wang and Crowcroft 1996, Tomovic et al. 2015]. 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. The Dijkstra solution performed the worst, presenting a total of 3,815 seconds during 609 stalls. 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.

4. QoE Prediction for VR Videos

4.1. Challenge

QoE has shown to be a critical factor for video applications [Ahmad et al. 2016]. As such, both network operators and VR content providers are required to answer an impor-
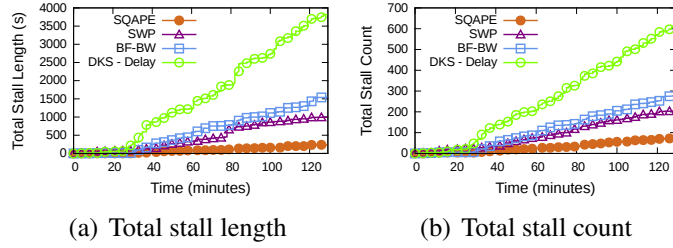
(a) Total stall length (b) Total stall count

Figure 2. Video streaming playout performance evaluation.

tant 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 constrains. 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.

4.2. 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 playout performance using network QoS indicators as predictors. Then, it uses the video playout performance metrics to model and estimate the end-user perceived quality. To cope with the VR ecosystem requirements, we introduce two novel concepts, namely, the concept of visualization zones and a VR video QoE model.

4.3. Results

The stalling time (Figure 3(a)) shows an error close to 13% for over 90% of the testing samples. In turn, Figure 3(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 3(b), the QoE estimation presents $r_i \leq 0.03922$ for over 90% of the cases.
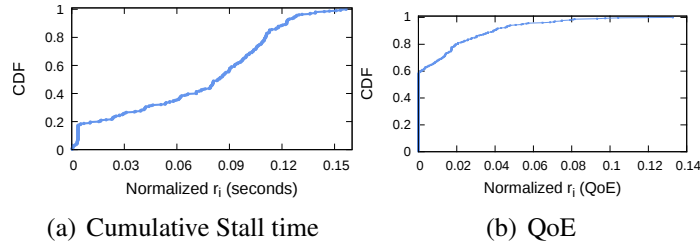

(a) Cumulative Stall time (b) QoE

Figure 3. Residual error CDFs for stall time, startup delay and QoE estimation.

5. VR Video Experimentation Platform

5.1. 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 tech-

niques 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.

5.2. 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.

5.3. 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. Figure 4 illustrate key insights provided by VR-EXP. Figure 4(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 4(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%).
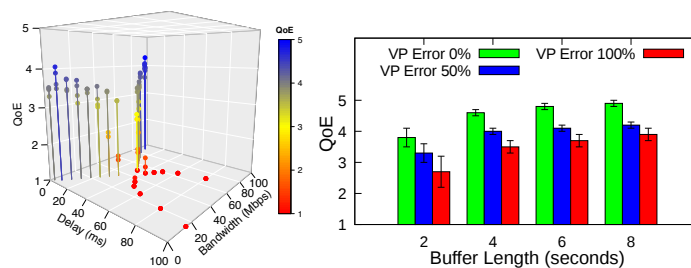
(a) Impact of VPE on QoE (b) Impact of VPE and buffer on QoE

Figure 4. The complex interplay of VPE, buffer size and QoE.

6. Research Accomplishments and Contributions

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

- Five papers in top-tier SBC, IEEE and ACM conferences/journals (first author) [da Costa Filho et al. 2016b, da Costa Filho et al. 2016a, da Costa Filho et al. 2018a, da Costa Filho et al. 2018b, da Costa Filho et al. 2020].
- Five papers in conferences/journals (co-author) [Marques et al. 2019, Kagami et al. 2019, Roesler et al. 2018, Lautenschläger et al. 2016, Lazzari et al. 2017].
- Ph.D. defense awarded with Honors Thesis High Distinction.
- Our papers provide publicly available source code and data set.
- Our VR Video prediction model [da Costa Filho et al. 2018b] 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).

7. 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.

References

Ahmad, A., Floris, A., and Atzori, L. (2016). Qoe-centric service delivery: A collaborative approach among otts and isps. *Computer Networks*, 110:168 – 179.

Ansari, R. I., Chrysostomou, C., Hassan, S. A., Guizani, M., Mumtaz, S., Rodriguez, J., and Rodrigues, J. J. P. C. (2018). 5g d2d networks: Techniques, challenges, and future prospects. *IEEE Systems Journal*, pages 1–15.

Cisco (2018). Cisco visual networking index: Forecast and trends, 2017–2022. Technical report, Cisco Systems.

da Costa Filho, R. I. T., Lautenschlager, W., Kagami, N., Roesler, V., and Gaspary, L. P. (2016a). Network fortune cookie: Using network measurements to predict video streaming performance and qoe. In *2016 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6.

da Costa Filho, R. I. T., Lautenschläger, W., Lazzari, H., Roesler, V., and Gaspary, L. P. (2016b). A rede tem a resposta: Um modelo escalavel para predicao integrada de qualidade de video e qoe em redes ip. In *Proceedings of the XXXIV Brazilian Symposium on Computer Networks and Distributed Systems (SBRC)*, Salvador, Bahia. (Honorable Mention Award - Top 3).

da Costa Filho, R. I. T., Luizelli, M., Lautenschlager, W., Kagami, N., Roesler, V., and Gaspary, L. P. (2018a). Scalable QoE-aware path selection in SDN-based mobile networks. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications (INFOCOM 2018)*, Honolulu, USA.

da Costa Filho, R. I. T., Luizelli, M., Vega, M., Hooft, J., Petrangeli, S., Wauters, T., Turck, F. D., and Gaspary, L. P. (2018b). Predicting the performance of virtual reality video streaming in mobile networks. In *Proceedings of the 8th ACM on Multimedia Systems Conference*, MMSys'18, pages 0–0, New York, NY, USA. ACM.

da Costa Filho, R. I. T., Luizelli, M. C., Petrangeli, S., Vega, M. T., van der Hooft, J., Wauters, T., Turck, F. D., and Gaspary, L. P. (2020). Dissecting the performance of VR video streaming through the VR-EXP experimentation platform. *TOMM*, 15(4):111:1–111:23.

ITU (2017). Measuring the information society.

Kagami, N., d. C. Filho, R. I. T., and Gaspary, L. P. (2019). Capest: Offloading network capacity and available bandwidth estimation to programmable data planes. *IEEE Transactions on Network and Service Management*, pages 1–1.

Katsarakis, M., Fortetsanakis, G., Charonyktakis, P., Kostopoulos, A., and Papadopouli, M. (2014). On user-centric tools for qoe-based recommendation and real-time analysis of large-scale markets. *IEEE Communications Magazine*, 52(9):37–43.

Kuipers, F., Mieghem, P. V., Korkmaz, T., and Krunz, M. (2002). An overview of constraint-based path selection algorithms for qos routing. *IEEE Communications Magazine*, 40(12):50–55.

Lautenschläger, W., Costa Filho, R., Kagami, N., Della Flora, S., de Moura, K., and Roesler, V. (2016). Netmetric: Uma plataforma para medição ativa de desempenho em redes ip de larga escala. In *Anais Estendidos do XXII Simpósio Brasileiro de Sistemas Multimídia e Web*, pages 116–119. SBC.

Lazzari, H. S., Tavares da Costa Filho, R. I., and Roesler, V. (2017). Qoe analyser: A framework to qoe knowledge base generation. In *Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web*, pages 433–436.

Maallawi, R., Agoulmine, N., Radier, B., and Meriem, T. B. (2015). A comprehensive survey on offload techniques and management in wireless access and core networks. *IEEE Communications Surveys Tutorials*, 17(3):1582–1604.

Marques, J. A., Luizelli, M. C., da Costa Filho, R. I. T., and Gaspary, L. P. (2019). An optimization-based approach for efficient network monitoring using in-band network telemetry. *J. Internet Services and Applications*, 10(1):12:1–12:20.

Roesler, V., Júnior, M. G., Husemann, R., da Costa Filho, R. I. T., and Valle, R. (2018). Mcu em software distribuído em nuvem. In *Anais Estendidos do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web*, pages 127–131. SBC.

Schlinker, B., Kim, H., Cui, T., Katz-Bassett, E., Madhyastha, H. V., Cunha, I., Quinn, J., Hasan, S., Lapukhov, P., and Zeng, H. (2017). Engineering egress with edge fabric. In *Proceedings of the 2017 ACM SIGCOMM Conference*, New York, NY, USA. ACM.

Tomovic, S., Radusinovic, I., and Prasad, N. (2015). Performance comparison of qos routing algorithms applicable to large-scale sdn networks. In *IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON)*, pages 1–6.

Wang, Z. and Crowcroft, J. (1996). Quality-of-service routing for supporting multimedia applications. *IEEE Journal on Selected Areas in Communications*, 14(7):1228–1234.