QoE Evaluation for Emerging Media Applications: Network-Level Analysis and Traffic Modeling

Md Tariqul Islam (Candidate), Christian Esteve Rothenberg (Advisor)
Universidade Estadual de Campinas, Brazil

***Abstract*—Technological advancements are driving the growing popularity of Emerging Media Applications (EMA), ranging from mobile applications to immersive services in eXtended Reality (XR). However, delivering satisfactory Quality of Experience (QoE) for EMA over 5G and beyond networks remains a relevant yet challenging topic due to high computational and network resource requirements and the need for network-level optimization. This PhD research addresses QoE evaluation for EMA in several directions through network-level measurement, traffic modeling, and optimization. It outlines the goals, methodologies, and potential research plans for each area.**

I. Introduction

Advancements in networking, multimedia technology, and electronic gadgets have led to the rise of Emerging Media Applications (EMA). These applications range from everyday mobile applications to fully immersive experiences through eXtended Reality (XR). Mobile applications, particularly those designed for smartphones, cover various aspects of daily life, including social networking, entertainment, productivity, and e-commerce. In contrast, immersive applications provide immense opportunities for interactive experiences in entertainment, training, education, healthcare, and more by seamlessly integrating the physical and digital realms [1].

This PhD research primarily focuses on immersive applications like omnidirectional 360-degree and volumetric/holographic videos designed for viewing in XR headsets, which include Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) technology. VR strives to simulate a 3D digital environment or replicate reality in virtual settings through headsets, while AR overlays virtual information on real-world objects. Combined, they create MR, which provides a fully immersive experience by mixing virtual content into the real world. Earlier forecasts projected a significant increase in immersive traffic and headset usage [2] [3]. These trends indicate that immersive experiences will be a significant part of the future of the internet.

In 360-degree VR video, omnidirectional cameras capture content, allowing headsets that track the user's viewport to enable three degrees of freedom (3DoF). In contrast, volumetric video presents 3D objects and scenes and offers full immersion by enabling six degrees of freedom (6DoF). These applications demand massive bandwidth (up to 1 Tbps) and ultra-low latency (in a few ms) [1] for smooth playback and a seamless experience. Failure to meet these requirements can result in cyber/motion-sickness. 6DoF volumetric video requires advanced approaches to cope with stringent interactivity and latency requirements. While transmission protocols such as QUIC, DASH/LL-DASH, and WebRTC are commonly used for streaming traditional 2D content, due to the complex nature of immersive video, particularly volumetric content, requires further investigation and redesign of these protocols to stream immersive experiences effectively. Moreover, decoding and rendering volumetric content in XR headsets is computationally intensive. 5G and beyond network is evolving with technologies like Edge Computing, Network Slicing, NFV, and SDN to meet these challenges.

Like emerging XR applications and gadgets, we have also observed the rapid growth of mobile devices and traffic, particularly in smartphone-based daily life applications [2]. Understanding and modeling both mobile and XR-associated application traffic has become crucial. The unique characteristics of these applications require tailored models, as existing research often focuses on web traffic and needs more precise models for mobile and XR applications. Developing these traffic models will help researchers better understand this traffic's behavior and design improved networking solutions.

Providing a pleasant Quality of Experience (QoE) for EMA remains a challenging topic in modern QoE-driven network management. Thus, this PhD research focuses on a comprehensive QoE evaluation for EMA using network-level analysis and modeling. We discuss goals and plans to enhance the QoE of EMA, including evaluating network performance's effect on immersive video streaming QoE by identifying immersive traffic and assessing QoE metrics derived from network-level measurement. Besides, we propose a traffic generation model that accurately replicates the behavior of mobile and XR applications based on real application traces. Furthermore, we explore the transmission protocols to optimize and improve volumetric streaming performance.

II. State of the Art & Research Questions

To optimize the network proactively or reactively for immersive XR experiences, specifically 360-VR streaming due to its significant growth [2] [3], with QoE inference network operator or management plane can only access Key Performance Indicators (KPIs) related to VR network-level traffic. This task is challenging due to widespread E2E encryption and the complex pipeline line of XR services. Prior research on

This work was supported by Ericsson Telecomunicações Ltda., and by the Sao Paulo Research Foundation FAPESP, grant 2021/00199-8, CPE SMARTNESS. This study was partially funded by *CAPES*, Brazil - Finance Code 001.

360-VR focused on subjective and PSNR/SSIM-based QoE assessments, which are impractical for operators. Although several studies [4] [5] used Machine Learning (ML) in traditional 2D video streaming, the immersive nature of 360-VR poses a gap in ML-based QoE prediction. Unlike the ML instance in [6], which used out-of-band features, we propose using in-band network KPIs from E2E encrypted traffic for a more accurate and robust QoE prediction method from the operator's perspective.

Beyond 360-VR, XR applications span domains such as gaming, social interaction, and education, generating distinct traffic patterns. Also, the rise of mobile applications has drastically altered traffic dynamics. Researchers often need to conduct controlled experiments to comprehend these applications' traffic patterns and design optimized network solutions. However, directly running these applications can be resource-intensive, inflexible, and difficult to repeat. Therefore, a tailored traffic model and generator for mobile and XR applications are essential. Existing literature on traffic modeling and generators often focuses on web application traffic [7]. A recent study [8] provided emulation-based web application traffic generators using stochastic modeling for user sessions. Another recent work [9] proposed a stochastic model for XR application traffic and presents a simulation framework. However, simulations may not effectively reflect real-world dynamics, which can be better achieved through emulation. Mobile and XR application traffic exhibit randomness and burstiness, which stochastic models alone cannot fully capture. Thus, to effectively represent these behaviors, we aim to develop a traffic model and emulation-based generator tailored to mobile and XR applications, focusing on user sessions.

In contrast, the complex nature of volumetric content streaming presents significant challenges, including content capturing, processing, transmission, and rendering. This PhD research focuses on video-on-demand (VoD) volumetric content transmission to align with 360-VR work and leverage relevant insights. Thus, live or real-time video scenarios, which are more delay-sensitive and involve protocols like LL-DASH, MoQ, and WebRTC, are beyond the scope of this work. In [1], authors showed that traditional DASH-based transmission remains still effective for streaming VoD volumetric content, typically implemented over TCP or QUIC. We aim to leverage DASH over QUIC for its advantages over TCP. A recent study [10] explored DASH-based volumetric streaming using partially reliable QUIC, and we propose a distinct approach that utilizes QUIC's stream prioritization for enhanced flexibility and adaptability in content delivery, specifically prioritizing user viewports for improved streaming performance, as motivated by prior work [11].

This PhD research endeavors to bridge existing knowledge gaps by conducting a comprehensive experiment to evaluate the QoE for emerging applications. Therefore, this research will focus on addressing the following research questions:

Research Question #1: Which network feature extraction technique do network operators leverage from encrypted traffic? How can 360-VR associated traffic be identified and QoE estimated based on available network-level KPIs? In this regard, we aim to develop a robust prediction engine using a supervised ML model that utilizes Quality of Service (QoS) KPIs as input attributes from a non-invasive (without decryption) and lightweight (packet header-based) feature extraction of E2E encrypted traffic for 360-VR video, intending to identify 360-VR traffic and infer objective QoE metrics. We previously [12] noted that this prediction engine can be integrated into the Network Data Analytics Function (NWDAF) in service-based 5G architecture. NWDAF can collect data from third-party Application Functions (AF) and network KPIs for ML training and deployment. Insights from NWDAF can then be used by other functions, such as the Policy Control Function (PCF) and User Plane Function (UPF), for optimization (e.g., traffic shaping and QoS management).

Research Question #2: How can we develop a flexible and realistic traffic generator that accurately mimics application traffic patterns for mobile and VR devices? In this matter, we aim to design and implement a hybrid traffic generator that replicates the traffic patterns of mobile and VR applications. The primary objective is to create an emulation framework that combines stochastic modeling to capture randomness with data-driven approaches to reflect the burstiness of real-world application traffic. Our approach will integrate various user interaction modes and QoE model, adapting to different network conditions and the characteristics of mobile and VR applications (hereafter referred to as apps). This will enable a more comprehensive evaluation of network performance and QoE for these emerging apps.

Research Question #3: How can the transmission protocol DASH over QUIC be adapted and optimized to support the efficient streaming of volumetric content, considering aspects like viewport prioritization? In this study, we pursue designing and implementing a client-server framework that enables DASH over QUIC streaming of Point Cloud (PC) based volumetric content with viewport-aware stream prioritization. The primary objective is to enhance users' overall QoE, specifically in scenarios with limited or fluctuating network bandwidth. Our approach will first focus on identifying the necessary modifications to make the client aware of the critical viewport data and leverage QUIC's stream prioritization to optimize transmission.

III. Challenges, Methodology, & Current Status

A. 360-VR traffic identification and QoE inference

Challenges: Network operators encounter challenges in mitigating QoE degradation issues in video services such as 360-VR due to limited control over factors affecting QoE, including human, system, and contextual factors. To evaluate QoE, operators need to identify 360-VR traffic and monitor performance metrics. However, using E2E encryption protocols such as HTTPS/QUIC makes it difficult to assess the quality of video from network traffic payloads directly. Although operators can retain IP header-level packet information for network traffic, tile-based 360-VR streaming poses challenges for correlating network QoS metrics with QoE outcomes, as spatial segment

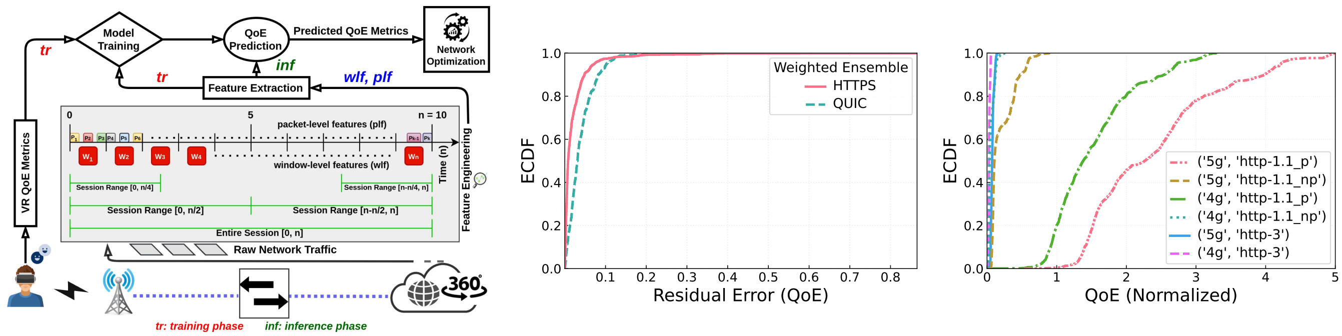

Fig. 1. Proposed Approach (*left*), ML Model Residual Error (*middle*), and QoE Distribution (*right*)

requests for each tile generate massive traffic for a small fraction of the video, unlike traditional streaming. Consequently, the question emerged of whether ML can effectively analyze these complex patterns and accurately predict QoE based on available network-level data.

Methodology: We will develop an experimental testbed to generate VR and non-VR traffic in either an emulated or real-world environment. The testbed will capture E2E encrypted traffic from a vantage network interface and record VR playout performance metrics. A non-invasive and lightweight feature extraction heuristic will be developed to extract relevant QoS features from the raw traffic IP header data. At the same time, VR playout metrics will be processed to create ground truth for objective QoE metrics. Extracted features and ground truth will then be merged to create datasets for two ML tasks: (a) VR classification concerning non-VR traffic, and (b) VR QoE metrics prediction. During data analysis, supervised ML models, including traditional and deep learning approaches, will be employed to train a prediction engine that classifies VR traffic and predicts QoE metrics using the extracted features. The model's accuracy will be assessed by comparing predicted values with actual values through statistical analysis methods such as the confusion matrix and residual error. Potential tools and libraries, including Mininet for network settings, a VR client (headless/real-headset), a network traffic sniffer, and a server for hosting prerecorded 360-VR video, will be utilized alongside Python and Scikit-learn for data processing, analysis, and ML implementation.

Current Status: We developed a testbed to emulate tile-based 360-VR streaming with a headless VR client, utilizing 4G and 5G drive test traces under various network conditions. We collected HTTPS and QUIC traffic, assuming operators can identify VR video traffic and sessions. Our non-invasive feature extraction heuristic (Fig.1- *left*) near the client extracted four basic QoS features: throughput, packet count, inter-arrival time, and packet size. These features with various statistics were used in an AutoML-based supervised learning model to predict VR streaming QoE, achieving up to 99% accuracy overall (Fig.1- *middle*). We found that current VR client was less effective with *curl*-based HTTP/3 over QUIC than native HTTP/1.1 over TCP, indicating the need for further improvements (Fig.1- *right*). More in-depth details about our approach and findings are available here [12]. Ongoing efforts include exploring HTTP/2, integrating optimized implementation of HTTP/3 over QUIC, enhancing traffic classification, and ML models generalizing aspects through transfer learning. We plan to use explainable AI to interpret models and identify key features. This research primarily focuses on offline ML model development and analysis, while real-time applications and network optimization based on QoE outputs are beyond our current scope. Finally, we aim to validate our findings with actual VR traffic using actual VR headsets and 5G setups, incorporating IP-level and radio metrics as our lab permits.

B. Hybrid Model-Based Emerging App Traffic Generator

Challenges: Developing a realistic traffic generator based on traffic modeling poses several challenges. First, it requires capturing the complex characteristics of apps, including user behavior, interaction modes, and app types. Accurately modeling each app demands real app user data, which is not easy to collect, mainly due to E2E encryption that covers detailed user behavior patterns. The traffic generator also needs to replicate the request and response patterns observed in real apps while maintaining flexibility to adapt to varying network conditions, ensuring it can generalize across different scenarios.

Methodology: To develop apps traffic model, we will start by collecting real app data across various categories, such as social media, entertainment, business, and news, under different network conditions. This data will capture user session requests, response patterns, packet sizes, and more. Since mobile apps often rely on HTTP request-response patterns (e.g., GET, POST), our modeling approach will remain generic, with an extension to VR-based apps that exhibit similar behavior. We will develop a heuristic to sequence user interactions within sessions, enabling separate modeling of each interaction. The data will be segmented into burst and non-burst traffic using a threshold-based heuristic, allowing us to capture randomness and burstiness. We will apply stochastic modeling for non-burst traffic (e.g., packet size, request number) by fitting probability distributions based on the data's characteristics and goodness-of-fit tests. Besides, a Markov model will be introduced to user interaction dynamics. For burst traffic, we will use data-driven modeling by calculating and using burst metrics such as burst probability, size, and duration from real user app session data. The traffic generator programs/applications will be developed in Python/C for both the client and server. This program will allow the client to incorporate stochastic and data-driven models, including the Markov model. It will be capable of making HTTP(S) requests and responses using either the HTTP/2 or HTTP/3 protocol, allowing for HTTP GET and POST. Moreover, we aim to implement a feature that can dynamically adjust network conditions for traffic generation. We will leverage the emulation framework, previously used for 360-VR streaming, to test the traffic generator in controlled, real-world-like virtual networks. The generated output will be evaluated for statistical similarity to the original traffic. The final tool will be released as open-source, contributing to research reproducibility.

Current Status: We have made significant progress on our mobile app traffic generator. Using a rooted Android device, Frida tool, and mitm-proxy, we collected decrypted app data in a controlled environment. The negative binomial and generalized poisson distributions proved most effective for stochastic modeling for packet sizes and request numbers, while burst metrics were derived for data-driven modeling. The client-side code uses Python's *httpx* for asynchronous HTTP/2 requests, with *hypercorn* handling the server side. Our generator shows statistical similarity to the original traffic, validated through hypothesis testing (KS test), burst metrics comparison, and analysis of overall data rates. We have also proposed a simplified QoE model based on each GET and POST method's data rates and response times. Current efforts focus on adapting models for various network conditions, exploring capturing VR app session data using the OVRseen tool to extend our approach to VR apps, and proposing a practical use case to evaluate the utility of the traffic generator.

C. Viewport-aware QUIC Prioritization for PC Content

Challenges: Volumetric content such as PC video inherent high bandwidth demands make maintaining smooth playback challenging even under good network conditions. It becomes even more problematic on mobile networks prone to fluctuations, where instability can lead to a degraded QoE. Such disruptions can have physical consequences for users, including cyber/motion sickness. Furthermore, the complexity of volumetric content, coupled with real-time user interactions such as viewport changes, calls for adequate data prioritization and adaptation. These mechanisms must effectively cope with fluctuating network conditions to deliver high QoE.

Methodology: We will extend our existing emulation-based client-server framework (used for previous research threads) to support DASH over QUIC for volumetric content streaming. The client will request pre-captured, compressed, and encoded point cloud content from a server. Using publicly available volumetric content datasets, we will compress and encode videos at various quality levels with tools like V-PCC, serving them via a simple server implementation. The client will be designed as a headless player (similar to the existing headless 360-VR player) and will handle rate adaptation and viewport prediction. Publicly available 6DoF movement data will be integrated to simulate real-time user interactions. We also aim to develop this volumetric headless player open for the research community, similar to existing tools like 2D goDASH and 360-VR video players. For the transport layer, we will use the *quiche*-based QUIC implementation, which supports stream prioritization. QUIC operates in user space, allowing for necessary customizations. The client will be developed in C/C++ using *quiche*'s API, and *nginx* will serve as the server. To evaluate performance under realistic conditions, we will simulate fluctuating network environments using open-source 5G trace datasets. Metrics such as initial delay, buffering, bitrate, and overall QoE will be monitored to assess the impact of viewport-aware stream prioritization.

Current Status: This objective is still in the proposal stage and has yet to be implemented. We are currently studying state-of-the-art volumetric/holographic content transmission, focusing on optimized streaming. Our initial plan is to adopt a gradual implementation approach, reusing existing tools and implementation logic to ensure the work remains on track for PhD completion within the expected timeframe. As we continue to gather knowledge in this area, the approach may be refined, potentially QoS to QoE estimation similar to 360-video and leveraging edge-assisted rendering to further optimize volumetric streaming if feasible within the PhD time.

IV. Concluding Remarks

With recent tech breakthroughs, the rapid growth of emerging media applications is gaining significant research attention regarding QoE in 5G and beyond networks. Thus, this research evaluates and analyzes the QoE of these applications from multiple perspectives, including 360-VR traffic QoE assessment, hybrid app traffic modeling, and optimizing network transmission for smooth volumetric delivery. The research strategy is expected to evolve as the work progresses.

References

[1] J. van der Hooft *et al.*, "A Tutorial on Immersive Video Delivery: From Omnidirectional Video to Holography," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1336–1375, 2023.

[2] V. Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2017–2022," *White paper*, vol. 1, no. 1, 2018.

[3] T. Alsop, "VR Headset Unit Sales Worldwide 2019-2024," 2022.

[4] I. Orsolic *et al.*, "A Framework for In-Network QoE Monitoring of Encrypted Video Streaming," *IEEE Access*, vol. 8, pp. 74691–74706, 2020.

[5] F. Bronzino *et al.*, "Inferring Streaming Video Quality from Encrypted Traffic: Practical Models and Deployment Experience," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 3, no. 3, pp. 1–25, 2019.

[6] R. I. T. da Costa Filho *et al.*, "Predicting the Performance of Virtual Reality Video Streaming in Mobile Network," in *Proceedings of the 9th ACM Multimedia Systems Conference*, pp. 270–283, 2018.

[7] O. A. Adeleke *et al.*, "Network Traffic Generation: A Survey and Methodology," *ACM Computing Surveys (CSUR)*, vol. 55, no. 2, pp. 1–23, 2022.

[8] M. Luglio *et al.*, "A Flexible Web Traffic Generator for the Dimensioning of a 5G Backhaul in NPN," *Computer Networks*, vol. 221, p. 109531, 2023.

[9] M. Lecci *et al.*, "An Open Framework for Analyzing and Modeling XR Network Traffic," *IEEE Access*, vol. 9, pp. 129782–129795, 2021.

[10] H. K. Ravuri *et al.*, "Partially Reliable Transport Layer for Quicker Interactive Immersive Media Delivery," in *Proceedings of the 1st Workshop on Interactive eXtended Reality*, pp. 41–49, 2022.

[11] S. Chellappa *et al.*, "Context-Aware HTTP Adaptive Video Streaming Utilizing QUIC's Stream Priority," in *Proceedings of the 2nd Mile-High Video Conference*, pp. 144–145, 2023.

[12] M. T. Islam *et al.*, "Predicting XR Services QoE with ML: Insights from In-Band Encrypted QoS Features in 360-VR," in *2023 IEEE 9th International Conference on Network Softwarization (NetSoft)*, pp. 80–88, IEEE, 2023.