QoE Assessment of Cloud-based Social Extended Reality Applications over Heterogeneous Access Networks

Karan Mitra, Henrique Souza Rossi, Justin Gavrell, Christer Åhlund
Pervasive and Mobile Computing, Department of Computer Science, Electrical and Space Engineering,
Luleå University of Technology, Skellefteå, Sweden
karan.mitra@ltu.se

***Abstract*—The next generation of immersive applications, such as eXtended reality (XR), will likely be cloud-based and streamed over mobile networks using myriad technologies such as WiFi and 6th-generation mobile networks. Mobile networks promise ubiquitous connectivity but are prone to stochastic network conditions that may be detrimental to end users' quality of experience (QoE). The impact of network conditions on QoE has been studied extensively by industry and academia regarding various multimedia services such as audio, video, and gaming. However, the impact of network conditions on users' QoE for XR-based social applications has yet to be thoroughly investigated. This paper presents novel results assessing the impact of network conditions (*N=20*) involving factors such as round trip time (RTT), jitter (RJ), and packet losses (PL) on users' QoE via realistic subjective tests (*N=28)* regarding social XR application. Our results show that social XR applications require stringent QoS conditions. In particular, our results show that increasing RTT values do not significantly affect users' QoE up to 77ms. Combined PL and RTT cases cause significant QoE degradation from 77ms onward with greater than 2% PL. Most importantly, results show that a very small jitter value with one standard deviation beyond 52 milliseconds can lead to significant QoE degradation. Further, jitter values beyond three standard deviations for 27ms RTT and beyond should be avoided.**

***Index Terms*—Extended reality, Subjective tests, Quality of Experience, Virtual Reality, Heterogeneous Access Networks, 6G, Metaverse**

I. Introduction

The area of Quality of Experience (QoE) aims to understand and measure the user's perception of an underlying service and/or application. It depends on the quality of service (QoS), a person's cognitive abilities, expectations, and experiences, behavior, as well as the person's surrounding environment [1]. QoE has been studied widely in terms of audio and video. Recent QoE research has focused on understanding new application classes, such as virtual reality (VR) [2] and augmented reality [3]. These applications will become the cornerstone of the eXtended Reality (XR) landscape, where these applications will combine to form cyber-physical spaces (e.g., metaverse) and enable people to interact and socialize, leading to Social XR. As in VR, Social XR applications may require stringent QoE from access networks such as Ethernet, WiFi, and 4G-6G networks to enable seamless interaction between people [4]. To the best of our knowledge and based on the state-of-the-art research [5], we assert a need for more research regarding QoE assessment of social XR applications considering the impact of network QoS factors such as jitter, delay, and packet losses.

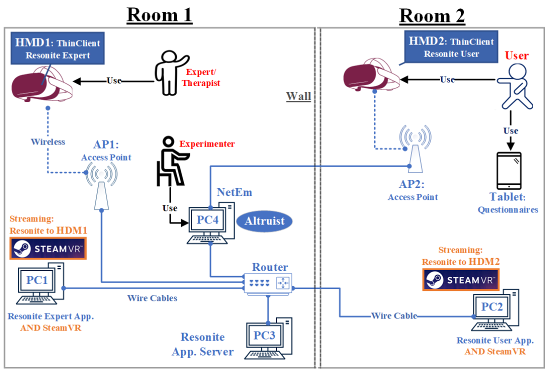


Fig. 1: Social-XR network experiment setup for user tests.

Contributions: To the best of our knowledge, this is the first paper to present results regarding QoE based on extensive subjective tests (*N*=28) in a social XR setting under realistic, diverse network conditions (*N*=20), especially, in the context of accessing the combined impact of RTT and RJ, and RTT and PL on users' QoE.

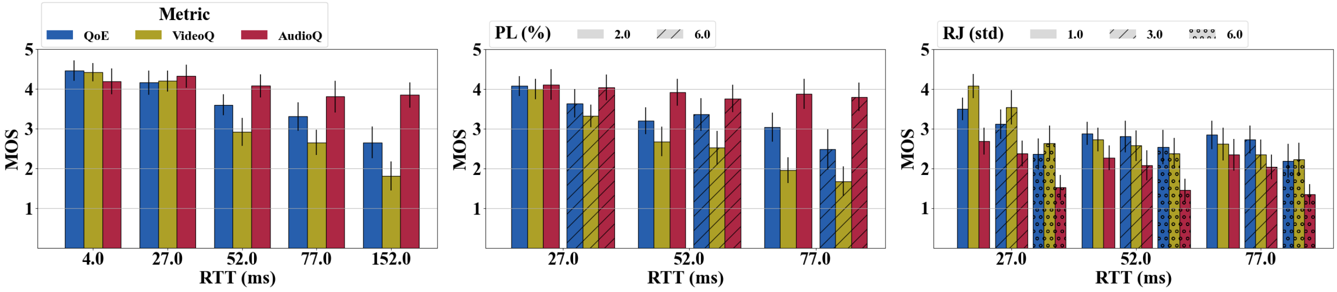
TABLE I: Network conditions evaluated during user tests. Note: Jitter values are generated from a normal distribution, and baseline RTT=2ms included for all conditions

Factors	No. of Conditions	Values
RTT (ms)	5	4, 27, 52, 77, 152
RTT (ms) and PL (%)	6	[27 ms; 2,6], [52 ms; 2,6], [77 ms; 2,6]
RTT (ms) and RJ (std-RTT in ms)	9	[27 ms; 1,3,6], [52 ms; 1,3,6], [77 ms; 1,3,6]
Total:	20	

II. Experiment Setup: Social XR Subjective Tests

Resonite, as the Social XR application that supports 6DOF was selected for collaboration, scene/avatar customization, and session hosting. We used commercial SteamVR[1] as the streaming solution to stream the PC Resonite application to VR devices in the highest quality. A laboratory environment

[1]https://store.steampowered.com/app/250820 [Online accessed: Nov. 2024]



(a) RTT (b) RTT and PL (c) RTT and RJ

Fig. 2: MOS results for three quality metrics as colors (QoE, video and audio) grouped by network conditions effect.

described in Fig. 1 was built and divided into two separate rooms; it was network-connected and capable of emulating network conditions through NetEm [6] via ALTRUIST [2]. *N=28* subjects were recruited to perform the subjective tests. We considered physiotherapy over VR as an application where Room 1 housed the expert (physiotherapist) and the experimenter; Room 2 housed the subjects who followed the exercise performed by the experts, which was displayed to them via their head-mounted displays (HMD).

Each subject followed a set of exercises performed by an expert. The 10 exercises were carefully chosen, and each was performed for nearly 90 seconds [7]. The exercises involved physically moving a subject's head, arms, shoulders, and hands, which were simultaneously tracked by the virtual avatar. During the tests, subjects and experts were free to interact with each other using voice conversations and actions performed in the VR environment. The total duration of each test was approximately 90 minutes, including filling out the pre-and post-test questionnaires. For their participation, each subject received a lunch coupon before the test. At the same time, they signed the consent form. Before the actual tests, the users were also given specific instructions regarding the tests and were invited to get accustomed to the VR device, the controllers, and the virtual environment.

III. Results

Fig. 2 demonstrates the effects of RTT, the combination of RTT and PL, and jitter on users' Overall QoE (calculated via the mean opinion score or MOS), as well as on audio quality (AudioQ) and video quality (VideoQ). Subjects rated these three metrics on the five-point Likert-like scale where '1' ="very poor"; '2' = "poor"; '3' = "average"; '4' = "good"; and '5' = "very good". Regarding RTT, see Fig. 2a, we did not notice any significant impact on the MOS values between 4 ms and 27 ms. Between the RTT of 52 ms and 77 ms, the MOS values decreased slightly, with average values hovering above 3. At 152 ms, we noticed MOS values decreased to below 3, approximately. We also noticed a similar trend in VideoQ ratings, i.e., VideoQ decreases in line with MOS values. However, the impact of RTT on AudioQ is not very apparent, with a very slight decrease in values with increased RTT values. Regarding PL and RTT, see Fig.2b, we noticed an apparent decrease in MOS values with increases in both PL(%) and RTT. A similar trend was also apparent between RTT and PL, and VideoQ. However, the combined effect of RTT and PL does not cause a significant decrease in AudioQ, showing its resiliency. Our analysis suggests that RTT greater than 52 ms with PL of more than 2% should not be exceeded for a sufficient user experience.

The impact of jitter on QoE, VideoQ, and AudioQ is very apparent, as can be observed from Fig. 2c. Even jitter of 1 std. dev. of RTT values significantly affect QoE, VideoQ, and AudioQ with reasonably low RTT values of 27ms. For higher RTT values of 52 ms and 77 ms, as well as higher jitter values, both AudioQ and VideoQ drop dramatically, leading to a drop in overall MOS. From this figure, it can also be observed that overall MOS and VideoQ values are higher than AudioQ, suggesting that the end users penalize lower AudioQ more than VideoQ; however, they can still enjoy higher overall QoE. From these results, we suggest the stakeholders aim for RTT values below 27 ms (approx.) and jitter, not more than two std. dev. of RTT.

We believe our results will assist the stakeholders in careful cloud and network resource planning for provisioning emerging immersive applications in the future.

References

[1] K. Mitra, A. Zaslavsky, and C. Åhlund, "Context-Aware QoE Modelling, Measurement, and Prediction in Mobile Computing Systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 920–936, May 2015.

[2] H. S. Rossi, K. Mitra, S. Larsson, C. Åhlund, and I. Cotanis, "Subjective QoE Assessment for Virtual Reality Cloud-based First-Person Shooter Game," in *ICC 2024 - IEEE International Conference on Communications*, Denver, CO, USA, Jun. 2024, pp. 4698–4703, iSSN: 1938-1883.

[3] C. Keighrey, R. Flynn, S. Murray, and N. Murray, "A Physiology-Based QoE Comparison of Interactive Augmented Reality, Virtual Reality and Tablet-Based Applications," *IEEE Transactions on Multimedia*, vol. 23, pp. 333–341, 2021.

[4] E. Stafidas and F. Foukalas, "A Survey on Enabling XR Services in Beyond 5G Mobile Networks," *IEEE Access*, vol. 12, pp. 59 170–59 197, 2024.

[5] C. Cortés *et al.*, "Delay Threshold for Social Interaction in Volumetric eXtended Reality Communication," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 20, no. 7, pp. 1–22, Jul. 2024.

[6] "NetEm." [Online]. Available: https://wiki.linuxfoundation.org/networking/netem

[7] ITU-T Recommendation P.809, "Subjective evaluation methods for gaming quality," 2018.