QoE Models for Virtual Reality Cloud-based First Person Shooter Game over Mobile Networks

Henrique Souza Rossi*, Karan Mitra*, Christer Åhlund*, Irina Cotanis†
*Pervasive and Mobile Computing, Department of Computer Science, Electrical and Space Engineering,
Luleå University of Technology.
†Infovista AB, Sweden.
henrique.souza.rossi@ltu.se

***Abstract*—Virtual reality cloud-based gaming (VRCG) services are becoming widely available on virtual reality (VR) devices delivered over computer networks. VRCG brings users worldwide an extensive catalog of games to play anywhere and anytime. Delivering these gaming services in existing broadband mobile networks is challenging due to their stochastic nature and the user's perceived Quality of Experience (QoE)' sensitivity towards them. More research is needed regarding developing effective methods to measure the impact of network QoS factors on users' QoE in the VRCG context. Therefore, this paper proposes, develops, and validates three novel regression models trained on a real dataset collected via subjective tests (N=30); the dataset contains subjective users' QoE ratings regarding VR shooter games affected by network conditions (N=28), such as round-trip time (RTT), random jitter (RJ), and packet loss (PL). Our findings reveal that due to the nonlinear relationship of (RTT and RJ) tested together, nonlinear (mean absolute error (MAE)=0.14) and polynomial (MAE=0.15) regression models have the best performance; yet, simple linear regression model (MAE=0.19) is also suitable to predict QoE for VRCG. Further, we found that feature importance depends on the model's choice (either RTT or RJ). Finally, our models' prediction of QoE for real-world traffic measurements suggests that mobile network traffic (4G, 5G non-standalone, 5G standalone) provides a 2.5 $\leq MOS_{QoE} \leq$ 3.0 experience for VRCG, while 4.2 $\leq MOS_{QoE} \leq$ 4.4 for wired connections, suggesting the need for improvements in the current commercial 5G network deployments to deliver VRCG.**

***Index Terms*—Subjective tests, Quality of Experience, Virtual Reality, Modeling, Prediction, Games.**

I. Introduction

Metaverse goals are set to bring new applications and services to enhance the way humans entertain, socialize, and interact with the virtual world [1]. As an example, the recent move of the 18 billion US$ worth cloud gaming industry to VR[1], provides users worldwide with new forms of playing an extensive catalog of games, on the most immersive and interactive VR devices anywhere and anytime. This new type of service relies on stringent network conditions, as both cloud gaming (CG) and VR gaming are highly sensitive to latency [2], PL, and jitter. As a result, there is a need to develop new metrics to estimate users' QoE, considering factors regarding networks, VR devices, and game content. QoE is a multidimensional metric that measures users' likes and dislikes towards a particular technology, application, or service [3]. Therefore, QoE metrics can assist stakeholders in realistically understanding the customer base and improving their services and applications regarding VRCG content, accordingly.

VRCG QoE evaluation is complex. Recently, a few studies aim to address it by assessing the effect of (albeit with limited ranges of) RTT and PL on users' QoE [4]–[6]. More importantly, no models have yet been proposed and validated to predict QoE for VRCG. We assert that there is a need to develop accurate QoE models for VRCG services that account for the heterogeneous nature of mobile broadband networks such as 4G and 5G that are prone to QoS impairments such as PL, RTT, and jitter that negatively affect users' QoE [7], [8]. Therefore, to fill this research gap, this paper focuses on the network side of VRCG context and considers 28 mobile network conditions for RTT, PL, combined (RTT,PL) and combined (RTT,RJ) to assess their impact on (N=30) users QoE for the first-person shooter (FPS) game, Serious Sam VR. This paper aims to answer the following question: *"How can we model and predict users' QoE for VRCG influenced by various realistic mobile network conditions ?"*

Contributions: i. To the best of our knowledge, this is the first paper to propose, develop, and validate QoE models for VRCG based on a wide range of network conditions covering 4G and 5G. In particular, this paper proposes and develops three novel regression QoE models for VRCG; ii. we carefully assess the importance of the models' features regarding RTT, PL, and their interactions (RTT,PL), (RTT,RJ), never studied before in the context of VR gaming and VRCG; and iii. from real-world network traffic measurements, we assess whether VRCG can be supported on current networks.

II. Testbed and Dataset

To study the effect of network conditions on users' perceived QoE, we conducted subjective tests in a lab environment based on the guidelines of [9]–[12]. In total, 30 users were invited to play Serious Sam VR: The Last Hope[2]. By choosing an FPS game, a genre that entails high precision and fast response from users [13], *we assume to cover such application factors, which may or may not be available in conjunction for other immersive VR interactive content*. We used Nvidia CloudXR[3] for game streaming.

[1]https://www.theverge.com/2023/12/13/24000134/xbox-cloud-gaming-meta-quest-3-vr-headset [online: accessed June 2024]

[2]https://store.steampowered.com/app/465240 [Online: accessed June 2024]
[3]https://developer.nvidia.com/cloudxr-sdk [Online: accessed June 2024]

978-3-903176-66-9 ©2024 IFIP

TABLE I: Network conditions emulated using NetEM.

QoS Factor	N. of Conditions	Values
RTT (ms)	8	4,27,52,77,177,277,352,402 in ms
PL (%) and RTT=4ms	3	6,12,24 in %
RTT (ms) and PL (%)	8	[27ms;2,4,6], [52ms;2,4,6], [77ms;2,4,*]
RTT (ms) and RJ (std)	9	[27ms;1,3,6], [52ms;1,3,6], [77ms;1,3,6]
Total	**28****	

*PL=6% forces service crash in CloudXR. Thus not included in the tests
**Setup baseline RTT=2ms included for all conditions

In the lab, a total of 28 emulated network conditions (see Table I), using NetEM controlled by ALTRUIST [14], were applied to both up/down links (RTT values were halved for each link). The range of values for each condition were defined following the iterature [4]–[6] for RTT and PL. Regarding the combined (RTT,PL) and (RTT,RJ)[4] we considered ranges from mobile network studies [11], [15]–[17]. Subjective tests had a total duration of 1 hour and 30 minutes, since we followed a within-subjective design (each user, played all 28 conditions, randomized). After each played match, participants were asked to rate their QoE on a Likert-like scale of 1 to 5 (where 1 = "very poor" and 5 = "very good"). Each match had a duration of 90 seconds, as suggested by ITU-T Rec. P809 [18], for game tests. From the collected data, we computed the mean opinion score (MOS) as the average of the user ratings for each condition to build a training dataset. More details on data collection, analyses, and the effect of network conditions on QoE and CloudXR metrics can be found in Rossi *et al.* [19]. In the next section, we present QoE models for VRCG.

III. QoE Models for VRCG

Our modeling process considered a broad range of *linear* and *nonlinear* regression functions, ensuring the ability to model complex nonlinear interactions in various ways. As a result, we propose the following three regression models. Please refer to the table II for models' fitted coffieciensts.

Linear Models: Described in Eq. 1 named **Lin.Reg.**, is a multiple linear regression equation composed of each independent variable as single terms. All terms are statistically significant ($p < 0.05$), which emphasizes the importance of each tested network condition for VRCG. Next, after investigating the interactions among all three features, we learned that only RTT·RJ produces a statistically significant coefficient.

$$x = l_0 + l_1 \cdot RTT + l_2 \cdot PL + l_3 \cdot RJ \quad (1)$$

Moreover, we examined both individual terms and their interactions in various forms, including linear, quadratic, and cubic. The only combination deemed meaningful in terms of coefficients' p-value, the residual distribution, and the accuracy of the prediction is described in Eq. 2 refereed as **Poly.Reg.**. The coefficients associated with this model are statistically significant ($p < 0.05$) and highlight the nonlinear impact of RTT·RJ on QoE for VRCG, while RTT, PL and combined (RTT, PL) conditions were best represented by a line.

$$x = p_0 + p_1 \cdot RTT + p_2 \cdot PL + p_3 \cdot RTT^2 \cdot RJ + p_4 \cdot RTT \cdot RJ^2 + p_5 \cdot RTT \cdot RJ^3, \text{ subjective to: (RTT, RJ)} \leq 77\text{ms, 6std} \quad (2)$$

[4]Jitter values generated from a normal distribution.

NonLinear Model: Since higher order terms were necessary to model the RTT·RJ interaction, the next natural step was to explore the vast set of nonlinear functions, at the cost of higher number of coefficients, which enables better curve control [20]. Followed by ITU-T G1072 model structure [21], we propose a third model in Eq. 3, named **NonLin.Reg.**, which is composed of three impairments I_r, I_p, I_j (see Eq. 3.1), each to model a network feature impairment. The constant QoE_{Max} is the maximum MOS_{QoE} encountered in our tests (similar to [10]), reduced by each impairment function.

$$x = QoE_{Max} - I_r - I_p - I_j \quad (3)$$

$$I_r = n_1 \cdot RTT \ , \quad I_p = n_2 \cdot PL$$

$$I_j = n_4 + \frac{(n_3 - n_4)}{1 + e^{(RJ - n_5)}} \text{ , where RTT} \geq \text{2ms and RJ} \geq 0 \quad (3.1)$$

An attempt was made to apply nonlinear equations (e.g. Gompertz, power, logistic, Gaussian, etc.) for each feature separately and combined. In the features of RTT,PL and (RTT,PL), we noticed that the curve-fit converged coefficients most often make nonlinear equations resemble lines, irrespective of the equation, leading to wide or infinity coefficients' confidence intervals. Hence, we model RTT and PL as a simple line. In contrast, the (RTT,RJ) conditions were best modeled by an S-shaped logistic growth equation, which accounted for the most variance of the RJ conditions while maintaining a narrow coefficient's confidence interval.

$$MOS_{QoE} = max(1, min(x, 5)) \quad (4)$$

Since regression models' outputs can exceed our MOS range [1,5], we use the function in Eq. 4 to ensure that the outputs remain within [1,5] for untrained network conditions. Next, we present the statistical analyses of the three models.

IV. Result Analyses

This section presents the results analysis of our novel QoE models based on root mean squared error (RMSE), mean absolute error (MAE), Pearson's correlation (PLCC), coefficient of determination (R2), and adjusted R2 values. The models' fitting coefficients are listed in Table II, while the models' performance scores are listed in Table III.

TABLE II: Models' fitted coefficients.

Model	Coefficients
Lin.Reg.	l_0=4.409098, l_1=-0.006864, l_2=-0.118618, l_3=-0.255789
Poly.Reg.	p_0=4.398909, p_1=-0.006793, p_2=-0.117927 p_3=85 $\cdot 10^{-6}$, p_4=-0.006595, p_5=817 $\cdot 10^{-6}$
NonLin.Reg.	QoE_{Max}=4.4, n_1=0.007003, n_2=0.122350, n_3=-0.277909, n_4=1.456521, n_5=1.937740

A. Model Fitting and Performance Analyses:

In linear and nonlinear regression, model correctness and validity assume that residuals are normally distributed [22]. Hence, an inspection of the three model's residuals (in Fig. 1) shows that they (red dots) nearly stay on top of the normal distribution line. Therefore, the models are correct and produce a good fit.Additionally, due to inherent approximation in nonlinear regression coefficients, we verify NonLin.Reg certainty for the best fit line, by calculating the coefficients' confidence interval.. For that, the F-test method considered robust [20] was employed and the results are reported in Fig. 1d. The findings indicate that the intervals ($\alpha = 0.05$) are small and can be concluded NonLin.Reg. produced the best fit, with a reasonable degree of certainty.

TABLE III: Models' performance in various metrics.

Model	RMSE	Cross.V. (RMSE)	MAE	Cross.V. (MAE)	PLCC	R2	R2 Adj	DF
Lin.Reg.	0.26	0.22	0.19	0.22	0.95	0.89	0.88	23
Poly.Reg.	0.19	0.22	0.15	0.22	0.97	0.94	0.93	21
NonLin.Reg.	0.22	0.20	0.16	0.20	0.96	–	–	22

The prediction error for Poly.Reg. and NonLin.Reg. is similar for RMSE, MAE and PLCC, with slightly better scores for Poly.Reg. (see Table III). This indicates that the higher-order RTT·RJ terms for Poly.Reg. were comparable to NonLin.Reg. RJ function. The largest residual for both models is resid=0.47 for the condition (RTT=77ms, PL=2%) in Fig. 2c, while all the remaining conditions have the highest resid ≤ 0.4. In contrast, Lin.Reg. has the highest score errors (see Table III), and the conditions with the highest were resi=-0.52 for (RTT=52ms, RJ=3std) and resid=0.51 for (RTT=77ms, RJ=1std) in Fig. 2d. Hence, it highlights the nonlinearity of RJ conditions which were modeled by Lin.Reg. as a line. Still Lin.Reg. achieved almost 90% R2, and therefore shows *that a simple linear regression is still capable of predicting MOS for VRCG.*

To verify the models' performance on unseen data, Leave-one-out Cross-Validation (LOOCV) was utilized which excluded each network condition one at a time from the training set. LOOCV was chosen due to the small size of the dataset (N=27), where some network condition cases (e.g. PL, N=3 conditions) have insufficient number of data-points to be split in larger sets. As a consequence, applying other types of Cross-validation splits, would result in biased results due to non-representative data used in training or test sets. The results of LOOCV in Cross.V.(RMSE) and Cross.V(MAE) metrics are presented in Table III. Comparison of MAE vs. Cross.V(MAE) for all models shows that a small prediction error increases between 0.03 and 0.07 MOS[5]. Hence, we conclude *the models are robust, unbiased and they do not over-fit the dataset.*

B. Feature Importance:

From the perspective of stakeholders, better service quality involves increased investments in new hardware and software to accommodate stringent requirements. Hence, they can benefit from feature importance analyses to guide their investment toward supporting VRCG QoS. To do this, we apply the Shapley method [23], which computes the average marginal contribution of each feature (RTT, RJ, and PL) and their studied ranges per model. The results detailed in Fig. 3, indicate that feature importance is model dependent. The most important features are RJ (for Poly.Reg) and RTT for

[5]MOS varies between 1 - 5 in the dataset.

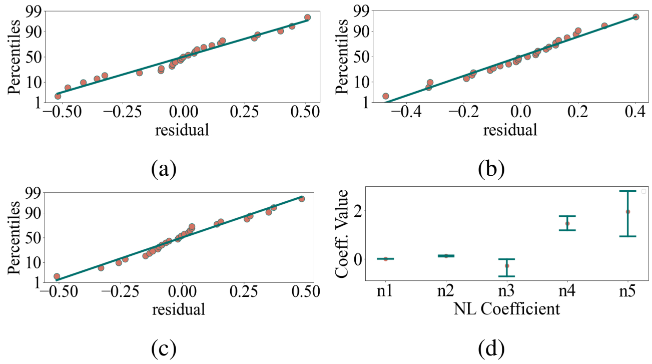

Fig. 1: Normal probability plot for (a) Lin.Reg., (b) Poly.Reg., (c) NonLin.Reg; Confidence interval for NonLin.Reg (d).

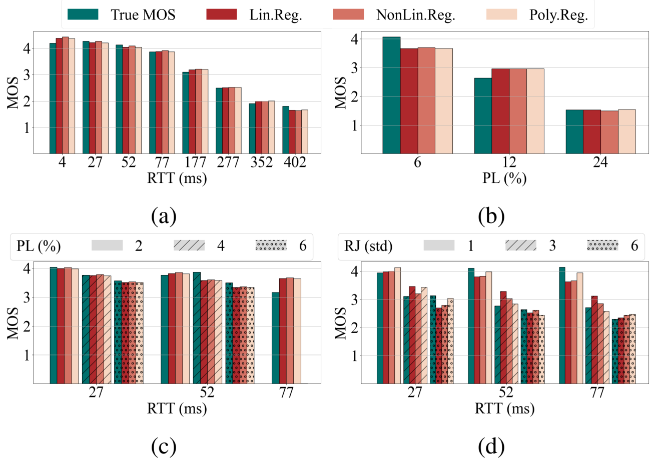

Fig. 2: Model's prediction MOS_{QoE} results as bars for the tested condition, against the True value.

NonLin.Reg and Lin.Reg. Therefore, *we recommend allocating new investments to reduce jitter and or RTT for VRCG services according to the model choice.*

C. Best Fit for VRCG:

The three models are statistically correct, they derive a good fit for the dataset, with low prediction error. However, their performance differs depending on the number of coefficients and degrees of freedom (DF). Hence, it is challenging to decide, solely on their performance, which model is the best suited for VRCG. Therefore, we follow the method suggested

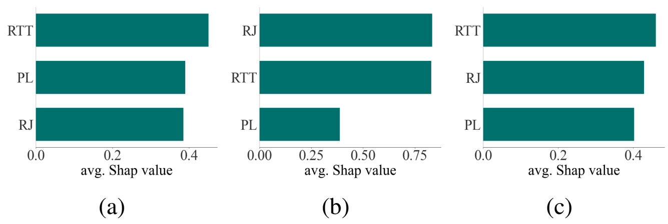

Fig. 3: Shap feature importance evaluation for (a) Lin.Reg., (b) Poly.Reg., (c) NonLin.Reg.

by [24], which entails employing a pairwise F-test, utilizing the residual sums of squares (SSR) and DF of the models, to ascertain the optimal statistical fit among them. Table IV presents the outcomes of the pairwise tests, each marked with a distinct test ID. It shows *Poly.Reg. has a statically better fit than NonLin.Reg and Lin.Reg., thereby establishing it as the most appropriate model for our VRCG dataset.*

TABLE IV: F-Test to compare model's fit performance.

Test ID	Model	SSR	F Value	P>|T|	DF	Hypothesis
0	NonLin.Reg.	1.26	–	–	22	–
0	Poly.Reg.	0.98	6.102	0.022	21	Reject NonLin.Reg.
1	Lin.Reg.	1.82	–	–	23	–
1	Poly.Reg.	0.98	9.079	0.001	21	Reject Lin.Reg.
2	Lin.Reg.	1.82	–	–	23	–
2	NonLin.Reg.	1.26	9.787	0.005	22	Reject Lin.Reg.

D. Assessment of VRCG in Real Networks:

Another important aspect related to existing commercially available broadband mobile networks is to asscess whether they can deliver sufficient QoE for VRCG. Therefore, we collected actual network traffic over seven days by sending ICMP packets every second from the city of Skellefteå in northern Sweden to five Amazon Services (AWS) data centers within Europe using three mobile network operators. This was done using four different network standards i.e., wired, 4G, 5G-NSA, 5G-SA. Subsequently, we computed the average RTT (Fig. 4a) and average Jitter (Fig. 4b) as input to the best statistically fit Poly.Reg model. The results, in Fig. 4c, reveal that the predicted QoE lies between 2.5 and 3.0 for all mobile networks tested, while the latest 5G-SA performs slightly better. The exception is for wired connections, where the MOS ranges from 4.2 to 4.4. Considering the subjective quality scale derived from the QoE question, *the quality of the VRCG performance should be between "Excellent" and "Good" (5 and 4 respectively) in wired connection, while "Average" and "Poor" (3 and 2 respectively) for mobile networks*. Therefore, it is imperative for stakeholders to allocate resources and enhance the 5G mobile network infrastructure to effectively support VRCG services as the values are not significantly better than the 4G networks.

V. Related Works

A QoE metric objectively captures users' overall quality perception affected by various context factors [3]; this is supported by studies examining differences in QoE scores due to device types and input [25] (e.g., VR vs. AR and tablet), screen sizes [26], video codecs and resolution [27], and network QoS [9], [11], [12]. Likewise, our novel models presented in this paper measure users' QoE for VRCG services affected by the VR device (MetaQuest Pro 2), the underlying network QoS, cloud-based game streaming, and the content (interactive shooter game).

To the best of our knowledge, only three studies investigated VRCG services affected by QoS factors (see Table V for a detailed comparison). They have considered either small ranges for RTT (10-90) ms and PL (0-4) % [6] or large one-way dalay (OWD) (100-500)ms. In contrast, based on the

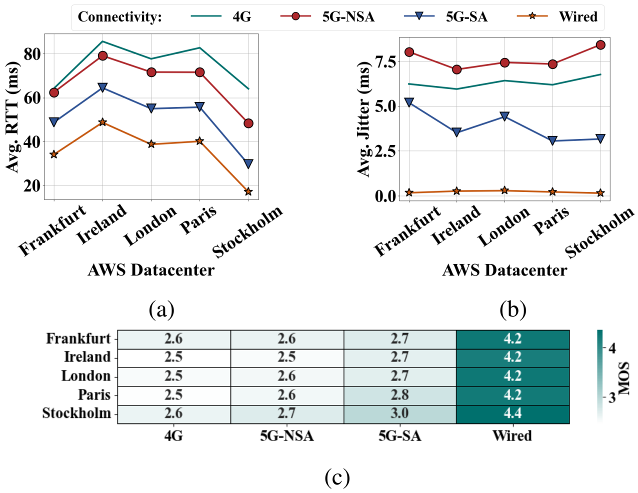

Fig. 4: Poly.Reg. MOS_{QoE} prediction (c), of real network traffic measurement, (a) RTT and (b) Jitter, to/from 5 different locations.

opinion of 30 users (the largest set), we have so far the most comprehensive QoS factors dataset (N = 28), including not only previous ranges for RTT, PL and, also, combinations of (RTT, PL) and (RTT,RJ). We cover degradations pertaining to broadband mobile networks, which occurs in both links and is very often affected by the combined effect of RTT and PL (see [7], [15]) and Jitter [7], [8] in 4G, 5G networks for stationary and mobile cases. We assess them in the most network demanding gaming content (shooter) that depends on fast and precise responses from users (similar to BeatSaber).

Foremost, no previous studies have introduced concrete and accurate VRCG QoE models. Although in [4], the authors apply models for VRCG, the models were not made available. This research thus represents the first effort to propose and validate QoE models based on QoS factors specifically for VRCG. QoE models for VR technologies and games are in high demand [28]. From the perspective of network operators and cloud providers, these models are required since QoS features are easily accessible and can be optimized regardless of the content [17]. From the research side, our proposed models can be used as a robust quality measurement guideline to further enhance studies in network traffic [29], video codecs [30], and mobile networks [31] for VRCG.

VI. Conclusion and Future Work

In this paper, from user tests (N=30) dataset, three novel VRCG regression models were proposed to predict QoE for (N=28) networks conditions RTT, PL, (RTT,RJ) and (RTT,PL) for the first time. The study reveals that simple linear regression can effectively predict VRCG QoE. Through F-tests, it has been determined that Poly.Reg has the best fit. Our model's most important features to predict VRCG QoE are RTT, and (RTT,RJ). Further, an assessment of real network traffic reveals MOS between ("Good - 4" and "Very Good - 5") and ("Average - 3 and "Poor - 2") for VRCG service in

TABLE V: Comparision of the state-of-the-art with our work

Paper	VR Assessment	Users	Games	Network Context	Network Direction	N. Network Conditions	Network Metrics	Model
[6]	VRCG (ALVR)	10	Together VR; Beat Saber; HalfLife-Alyx;	–	Up/Down Links	8	**RTT**: 0,10,30,50,70,90 in ms **PL**: 0,2,4 in %	No
[5]	VRCG (Daydream)	10	In-house (Sword Swing)	–	–	4	**OWD**:120,150,200 300 in ms	No
[4]	VRCG (ALVR)	12	AngryBird; BeatSaber; ArtPuzzle	Wired	–	4	**OWD**: 0,100,300,500 in ms	No
This Paper	**VRCG (Nvidia CloudXR)**	**30**	**Serious Sam VR**	**4G, 5G, Wired**	**Up/Down Links**	**28**	**RTT: 4, 27, 52, 77, 177, 277, 352, 402 in ms** **PL: 6,12,24 in %** **RTT and PL: (27,52,77)ms; (2,4,6) %** **RTT and Jitter: (27,52,77)ms; (1,3,6)std**	**Linear, Polynomial and Non-Linear Regression.**

wired and mobile networks (4G,5G), respectively. We aim to present our models in the ITU-T SG12 meetings in the future.

Acknowledgment: We thank David Lindero for his feedback on the models' performance analyses.

References

[1] S. Mystakidis, "Metaverse," *Encyclopedia*, vol. 2, no. 1, pp. 486–497, Mar. 2022.

[2] J.-P. Stauffert, F. Niebling, and M. E. Latoschik, "Latency and Cybersickness: Impact, Causes, and Measures. A Review," *Frontiers in Virtual Reality*, vol. 1, p. 582204, Nov. 2020.

[3] 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.

[4] K.-Y. Lee, J.-W. Fang, Y.-C. Sun, and C.-H. Hsu, "Modeling Gamer Quality-of-Experience Using a Real Cloud VR Gaming Testbed," in *Proceedings of the 15th International Workshop on Immersive Mixed and Virtual Environment Systems*. Vancouver BC Canada: ACM, Jun. 2023, pp. 12–17.

[5] T. Kämäräinen, M. Siekkinen, J. Eerikäinen, and A. Ylä-Jääski, "CloudVR: Cloud Accelerated Interactive Mobile Virtual Reality," in *Proceedings of the 26th ACM international conference on Multimedia*. Seoul Republic of Korea: ACM, Oct. 2018, pp. 1181–1189.

[6] Y. C. Li, C. H. Hsu, and C. H. Hsu, "Performance Measurements on a Cloud VR Gaming Platform," in *QoEVMA 2020 - Proceedings of the 1st Workshop on Quality of Experience (QoE) in Visual Multimedia Applications*, vol. 20. New York, NY, USA: Association for Computing Machinery, Inc, Oct. 2020, pp. 37–45.

[7] N. Ahmad, A. Wahab, J. Schormans, and A. A. Arnab, "Significance of Cross-Correlated QoS Configurations for Validating the Subjective and Objective QoE of Cloud Gaming Applications," *Future Internet*, vol. 15, no. 2, p. 64, Feb. 2023.

[8] S. Neumeier, E. A. Walelgne, V. Bajpai, J. Ott, and C. Facchi, "Measuring the Feasibility of Teleoperated Driving in Mobile Networks," in *2019 Network Traffic Measurement and Analysis Conference (TMA)*, Jun. 2019, pp. 113–120.

[9] S. Schmidt, S. Zadtootaghaj, S. S. Sabet, and S. Möller, "Modeling and Understanding the Quality of Experience of Online Mobile Gaming Services," in *QoMEX '21*, Jun. 2021, pp. 157–162, iSSN: 2472-7814.

[10] S. Zadtootaghaj, S. Schmidt, and S. Möller, "Modeling Gaming QoE: Towards the Impact of Frame Rate and Bit Rate on Cloud Gaming," in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, May 2018, pp. 1–6, iSSN: 2472-7814.

[11] H. S. Rossi, N. Ögren, K. Mitra, I. Cotanis, C. Åhlund, and P. Johansson, "Subjective Quality of Experience Assessment in Mobile Cloud Games," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Dec. 2022, pp. 1918–1923.

[12] S. Vlahovic, M. Suznjevic, and L. Skorin-Kapov, "The Impact of Network Latency on Gaming QoE for an FPS VR Game," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, Jun. 2019, pp. 1–3, iSSN: 2472-7814.

[13] M. Claypool and D. Finkel, "The effects of latency on player performance in cloud-based games," in *2014 13th Annual Workshop on Network and Systems Support for Games*, Dec. 2014, pp. 1–6, iSSN: 2156-8146.

[14] H. S. Rossi, K. Mitra, C. Åhlund, I. Cotanis, N. Ögren, and P. Johansson, "ALTRUIST: A Multi-platform Tool for Conducting QoE Subjective Tests," in *2023 15th International Conference on Quality of Multimedia Experience (QoMEX)*, Jun. 2023, pp. 99–102, iSSN: 2472-7814.

[15] D. Baltrunas, A. Elmokashfi, and A. Kvalbein, "Measuring the Reliability of Mobile Broadband Networks," in *Proceedings of the 2014 Conference on Internet Measurement Conference*. Vancouver BC Canada: ACM, Nov. 2014, pp. 45–58.

[16] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, S. Venkataraman, and J. Wang, "A first look at cellular network performance during crowded events," *ACM SIGMETRICS Performance Evaluation Review*, vol. 41, no. 1, pp. 17–28, Jun. 2013.

[17] X. Marchal, P. Graff, J. R. Ky, T. Cholez, S. Tuffin, B. Mathieu, and O. Festor, "An Analysis of Cloud Gaming Platforms Behaviour Under Synthetic Network Constraints and Real Cellular Networks Conditions," *Journal of Network and Systems Management*, vol. 31, no. 2, p. 39, Feb. 2023.

[18] I.-T. ITU-T Recommendation P809, "Subjective evaluation methods for gaming quality," 2018.

[19] 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*, Jun. 2024, pp. 4698–4703, iSSN: 1938-1883.

[20] H. Motulsky and A. Christopoulos, *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A practical guide to curve fitting*. Oxford University PressNew York, NY, May 2004.

[21] ITU-T Recommendation G.1072, "Opinion model predicting gaming quality of experience for cloud gaming services," Tech. Rep., 2020.

[22] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, fith ed. Wiley, 2013.

[23] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.

[24] H. J. Motulsky and L. A. Ransnas, "Fitting curves to data using nonlinear regression: a practical and nonmathematical review," *The FASEB Journal*, vol. 1, no. 5, pp. 365–374, 1987.

[25] 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.

[26] H.-J. Hong, C.-F. Hsu, T.-H. Tsai, C.-Y. Huang, K.-T. Chen, and C.-H. Hsu, "Enabling Adaptive Cloud Gaming in an Open-Source Cloud Gaming Platform," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 12, pp. 2078–2091, Dec. 2015.

[27] H. T. T. Tran, N. P. Ngoc, C. T. Pham, Y. J. Jung, and T. C. Thang, "A subjective study on QoE of 360 video for VR communication," in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. Luton: IEEE, Oct. 2017, pp. 1–6.

[28] T. Hossfeld, A. Seufert, F. Loh, S. Wunderer, and J. Davies, "Industrial User Experience Index vs. Quality of Experience Models," *IEEE Communications Magazine*, vol. 61, no. 1, pp. 98–104, Jan. 2023.

[29] M. Casasnovas, C. Michaelides, M. Carrascosa-Zamacois, and B. Bellalta, "Experimental Evaluation of Interactive Edge/Cloud Virtual Reality Gaming over Wi-Fi using Unity Render Streaming," Feb. 2024, arXiv:2402.00540 [cs].

[30] E. S. Korneev, M. V. Liubogoshchev, and E. M. Khorov, "Studying Cloud-Based Virtual Reality Traffic," *Journal of Communications Technology and Electronics*, vol. 67, no. 12, pp. 1500–1505, Dec. 2022.

[31] G. Maiorano, G. P. Mattia, and R. Beraldi, "Local and Remote Fog based Trade-offs for QOE in VR Applications by using CloudXR and Oculu Air Link," in *2022 International Conference on Edge Computing and Applications (ICECAA)*, Oct. 2022, pp. 95–101.