The Impact of Network Latency on Gaming QoE for an FPS VR Game

Vlahović, Sara; Sužnjević, Mirko; Skorin-Kapov, Lea

Source / Izvornik: **2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), 2019, 1 - 3**

Conference paper / Rad u zborniku

Publication status / Verzija rada: **Accepted version / Završna verzija rukopisa prihvaćena za objavljivanje (postprint)**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:168:025039

2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)

The Impact of Network Latency on Gaming QoE for an FPS VR Game

Sara Vlahovic, Mirko Suznjevic, Lea Skorin-Kapov
University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
{sara.vlahovic, mirko.suznjevic, lea.skorin-kapov}@fer.hr

***Abstract*—In this paper we present the results of two user studies aimed at exploring the effects of network latency on user experience in a First Person Shooter Virtual Reality multiplayer game. In both studies we observed the following metrics: Quality of Experience (QoE), willingness to continue playing in given network conditions, and the final outcome of the game (survival/death). After conducting the first study, we concluded that test methodology and contextual factors (e.g., social context, level of difficulty, weapon choice) may have had a strong influence on perceived QoE, subsequently causing inconclusive results. Therefore, we aimed to mitigate those factors through different methodology choices in the second study, which led to more conclusive, statistically significant results, indicating that user experience for the chosen game begins to suffer in cases of latency greater than approx. 100 ms (round trip time between client and server). In addition to the analysis of acquired results, we discuss the significance of methodology and context when conducting studies exploring gaming QoE in multiplayer environments.**

I. Introduction

Despite the recent decrease in revenue, Virtual Reality (VR) is expected to become more popular in the following years as new and improved versions of VR headsets start to appear on the market (according to Digi-Capital[1]). As a highly immersive medium, VR exhibits potential in terms of enhancing user interaction and collaboration in shared virtual environments, such as multiplayer games. However, wide-spread acceptance of new technology is always conditioned by its ability to generate and maintain high levels of user satisfaction. Before VR is able to accomplish mainstream success, significant effort needs to be invested towards enhancing the overall Quality of Experience (QoE) in VR systems and services.

While multiplayer gaming QoE has been explored in depth throughout numerous studies, research in the field of VR multiplayer games is very limited (to the best of our knowledge). However, VR games and certain non-VR games can be compared to a degree considering they share some similarities, namely *interaction model* and *perspective*. Those two factors are noted and analyzed in the context of sensitivity to impaired network conditions in a popular article by Claypool et al. [1], which states that games using an *Avatar interaction model with a first person perspective* (most similar to typical VR games) tend to have the lowest (100 ms) latency threshold due to their demand for fast hand-eye coordination and precise movement.

[1] https://tcrn.ch/2HcZMoe

In order to begin exploring the possible latency threshold for VR games, we have conducted two studies looking into the effects of variable network latency on user experience in a cooperative First Person Shooter (FPS) VR game. The contributions of this paper are twofold. We provide and discuss the results of these studies, namely QoE scores, participants' willingness to continue playing in given network conditions, and survival rates in scenarios displaying different levels of latency. Secondly, we describe methodologies of both studies, compare their contextual differences and analyze the effect of those differences on the conclusiveness of obtained results.

II. Methodology

The VR game used in both studies was *Serious Sam VR: The Last Hope*. While initially released as a single player game, *Serious Sam VR: The Last Hope* can also be played as a cooperative (*co-op*) two-player game, which is the option we chose for both *Study 1* and *Study 2*. The networking architecture of the chosen game is based on a client-server model where one of the user computers acts as the server. Player position is the only interaction component authorized by the client. In our case, the Oculus Rift headset was used by the examined participant on the client side in both studies. HTC Vive was used on the server side (for the *active* or *passive co-player*). The gameplay mechanic is as follows - each player chooses two weapons (range weapons being the typical choice, however melee options are also available) and uses them on multiple moving targets (enemies). Enemies appear in waves and are able to harm players with their ranged and melee attacks. *Serious Sam VR: The Last Hope* does not utilize any lag compensation methods. At the beginning of both studies, participants were familiarized with tested parameters and assessment methods (i.e., questionnaire filled by the administrator), controls and rules of the chosen game, as well as the Oculus Rift system, used by the examined participant. In *Study 1*, participants were also familiarized with the HTC Vive system, used only by the *active co-player*. After completing a training session in unimpaired network conditions (initial network latency of approx. 10 ms), each participant was presented with different scenarios in randomized order, resulting in the end-to-end network delay of initial + added latency. The average duration of a scenario (ending with victory or death) was between 2 and 3 minutes for both studies, therefore exceeding the minimal recommended duration (90-120 seconds) of a short interactive test, as noted

TABLE I
DIFFERENCES BETWEEN *Study 1* AND *Study2*

	Study 1	Study 2
No. of participants	24	33
Experience level distribution	beginners: 8 intermediate: 8 advanced: 8	beginners: 17 intermediate: 14 advanced: 2
Gender distribution	female: 10 male: 14	female: 12 male: 21
Co-player activity	*active co-player*	*passive co-player*
Level of difficulty	6	1
Weapon choice	undefined	weapons displaying the full bullet path (e.g. Lasergun)
Latency simulator	Clumsy[3]	Net.Shark
Latency scenarios (RTT)	150 ms, 200 ms, 250 ms, 300 ms	50 ms, 100 ms, 150 ms, 200 ms, 250 ms, 300 ms

by ITU-T P.809 (06/2018) Recommendation [2]. A pre-test questionnaire was used to collect demographic information (age, gender, self-reported level of experience in VR), while an in-game questionnaire was used to note the overall QoE score, rated on a 5 pt. Absolute Category Rating (ACR) scale (from 5-Excellent to 1-Bad) as well as participants' willingness to continue playing in given network conditions (*yes / no*) for each scenario. Game logs were also used to collect data about the outcome of each scenario (survival / victory or death). Differences between studies are presented in Table I.[2]

In *Study 1*, participants were sorted into pairs. Latency scenarios were chosen based on results of a pilot study, which showed that even advanced players fail to notice added latency values below 150 ms in given conditions. After completing each scenario, only the participant playing on the client computer was asked about their experience. The participant playing on the server side was considered an *active co-player*. After completing all scenarios, participants would switch places and roles, the examined participant becoming a new *active co-player*, while the other participant answered questions. The testing procedure lasted approximately 40 minutes per pair.

In *Study 2* each participant played alongside a *passive co-player*, i.e., the other avatar in the co-player situation provided a static target for enemies without actively participating in the scenario. It is important to note that, while not providing a realistic co-op experience, this method of testing a multiplayer game was purposefully chosen with the goal of eliminating some of the potential noise-introducing factors and providing a more controlled context for this study, especially when compared to *Study 1*. A significantly less intense level of difficulty (1) was chosen in order to provide the participant with a more achievable scenario, due to lack of co-player participation. Since latency values used in *Study 1* were greater than the previously mentioned recommended limit for games using an *Avatar interaction model with a first person perspective* [1], and keeping in mind the possibility that the more controlled setup of *Study 2* might result in latency becoming more noticeable, participants were presented with additional latency scenarios (50 ms and 100 ms). The complete testing procedure lasted approximately 35 minutes per participant.

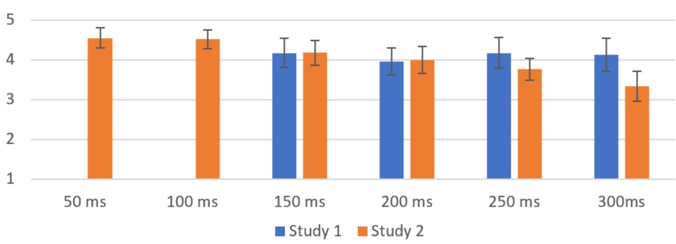

Fig. 1. Avg. QoE score per latency scenario (95% confidence intervals shown)

[2] Dataset is available at: https://muexlab.fer.hr/muexlab/research/datasets

[3] https://jagt.github.io/clumsy/download.html

III. RESULTS

When analyzing the results collected during *Study 2* (Fig. 1), it can be observed that scenarios with latency values of 50 ms and 100 ms received very similar average scores, followed by a gradual decline of scores with an increase in latency. The decrease in QoE scores after exceeding the apparent limit of 100 ms coincides with the latency threshold stated by Claypool et al. [1]. The statistical significance of these results is indicated by the single factor ANOVA test, resulting in a p-value of 3.45×10^{-8}. However, the results of *Study 1* are less conclusive, with 150 ms, 250 ms and 300 ms scenarios receiving very similar average scores. The 200 ms scenario scored slightly worse, despite not having the most severe amount of latency. Additionally, the single factor ANOVA test (p-value of 0.831) confirms there is no significant difference in QoE scores between latency scenarios in this study.

While 150 ms and 200 ms scenarios received similar average scores in both studies, a clear difference between *Study 1* and *Study 2* can be noticed when observing the average QoE score for 250 ms and 300 ms. Possible reasons for this are manifold, and noted based on participants' comments. Firstly, by not displaying the full bullet/projectile path, certain types of weapons used in the game are able to conceal the effects of latency. Defining the weapon choice in *Study 2* prevented participants from choosing latency-concealing weapons (weapons triggering the muzzle flash effect without displaying the bullet/projectile and its full trajectory), therefore indirectly making the effects of network delay more easily noticeable, in addition to providing a more controlled context compared to *Study 1*. Secondly, due to playing at a demanding and fast-paced level of difficulty, participants in *Study 1* had to be very focused on survival during gameplay, which consequently made them less perceptive of any effects of latency. It is also important to note the social component of cooperative gameplay. As mentioned in [2], analyzing data collected during a multi-player gaming session is not as straightforward as analyzing the results of a single-player experience due to additional factors influencing the results. As observed by Suznjevic et al. [3], participants in a multiplayer setting tend to give different QoE scores depending on the skill level of their co-players. Therefore, QoE scores given by the examined participant in *Study 1* could have been greatly influenced by the *active co-player*'s skill level, however the magnitude of this impact is hard to assess given that it is difficult to isolate and quantify individual contributions of each

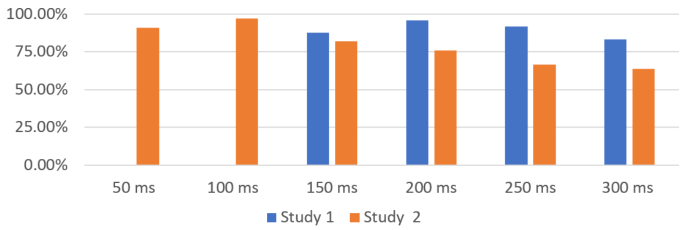

Fig. 2. Willingness to continue playing in given network conditions

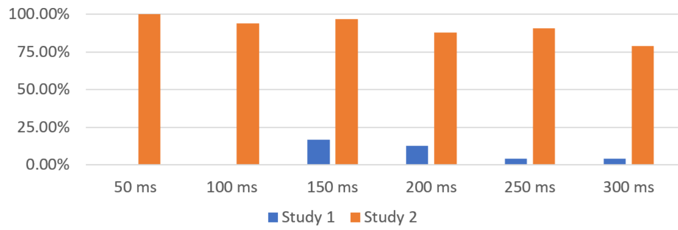

Fig. 3. Survival rate per latency scenario

co-player. Additionally, even seemingly irrelevant details, such as the *active co-player*'s gender [4], could have influenced the participants' own performance and, indirectly, their QoE assessment. The social component was omitted from *Study 2*, possibly explaining its clearer results.

As illustrated by Figure 2, the majority of users would choose to continue playing despite latency values as high as 300 ms. However, as shown by *Study 2* results, the percentage of willing participants began to gradually decrease once added latency exceeded 100 ms. In comparison to *Study 2*, participants in *Study 1* exhibited more willingness to continue playing for all tested scenarios, possibly due to contextual factors preventing them from noticing the effects of latency. Interestingly, despite being assigned the lowest average QoE score among all *Study 1* scenarios, the 200 ms scenario resulted in the highest percentage of participants willing to continue playing. On the contrary, *Study 2* results display a more apparent correlation between average QoE scores and the participants' willingness to continue playing.

Figure 3 displays a very apparent discrepancy between studies in terms of survival rates per latency scenario. The obvious reason for this are different levels of difficulty chosen for *Study 1* and *Study 2*. Nonetheless, both studies indicate that playing at a higher level of latency decreases survival rate. The decrease in survival rate in *Study 2* is not as regular as the decrease in QoE scores in the same study (e.g. the 150 ms scenario resulted in a higher survival rate compared to the 100 ms scenario), but it still portrays an apparent gradual decline, with a difference of 21% between the 50 ms scenario and the 300 ms scenario. The already low survival rate in *Study 1* has a more regular gradual decline.

When observing the analysis of subjective metrics, such as QoE and the willingness to continue playing, it is important to note that contextual factors of *Study 1* had a strong influence on participants' experience. Namely, by concealing the effects of latency (e.g., latency-concealing weapon choices), distracting the participant (e.g., social aspect of playing with an *active co-player*, playing at a fast-paced level of difficulty) and complicating the consequent analysis (e.g. the presence of two active players making it harder to separate their performances and determine their individual impacts on the outcome of the game), contextual factors had a big impact on assessment scores in *Study 1* and are therefore considered at least partially responsible for its inconclusive results. User related factors, such as gender, age and experience level, were not explored due to the unbalanced distribution of participants.

However, as indicated by declining survival rates in *Study 1*, while participants seemed unaffected by added latency in terms of their subjective experience, their performance clearly suffered. As this discrepancy between objective and subjective metrics appeared in a study which is more similar to the real gaming experience, it raises the question whether certain developer choices (e.g., weapon design), despite being seemingly unrelated to the issue of latency compensation, inadvertently mitigate negative effects of latency on subjective QoE and to what extent. If so, further research addressing this issue could create an opportunity for developers to improve user satisfaction by substituting game elements that are found to be more latency-sensitive with latency-concealing alternatives.

IV. Conclusion

We have conducted two studies with the aim of exploring the latency threshold for the chosen FPS VR multiplayer game. We discussed and compared methodologies and results of both studies. *Study 1* yielded inconclusive results in terms of subjective metrics, possibly due to contextual factors (e.g., weapon choice, social context, level of difficulty) indirectly concealing negative effects of latency. The impact of these factors was mitigated in *Study 2* through different methodology choices, leading to statistically significant results which indicate that added latency above 100 ms had a noticeable negative impact on QoE levels and participants' willingness to continue playing in given networked conditions. Additionally, the observed difference in conclusiveness between two studies confirms that contextual factors have a significant influence on subjective metrics. Therefore, these parameters should be further investigated and tested in future user studies.

Acknowledgment

This work has been supported under the project KK.01.2.1.01.0130 (Upravljanje energetskom infrastrukturom kroz kolaboraciju u proširenoj stvarnosti - ARIEN) financed by the EU from the European Regional Development Fund.

References

[1] M. Claypool and K. Claypool, "Latency and player actions in online games," *Communications of the ACM*, vol. 49, no. 11, pp. 40–45, 2006.

[2] ITU-T Recomm. P.809, "Subjective evaluation methods for gaming quality," *International Telecommunication Union-Telecommunication Standardisation Sector*, 2018.

[3] M. Suznjevic, L. Skorin-Kapov, and M. Matijasevic, "The Impact of User, System, and Context factors on Gaming QoE: a Case Study Involving MMORPGs," in *Proceedings of Annual Workshop on Network and Systems Support for Games*. IEEE Press, 2013, pp. 1–6.

[4] L. Vermeulen, E. Núñez Castellar, and J. Van Looy, "Challenging the Other: Exploring the Role of Opponent Gender in Digital Game Competition for Female Players," *Cyberpsychology, Behavior, and Social Networking*, vol. 17, no. 5, pp. 303–309, 2014.