Empirical QoE Study of In-Home Streaming of Online Games

Ivan Slivar, Mirko Suznjevic, Lea Skorin-Kapov, Maja Matijasevic
University of Zagreb
Faculty of Electrical Engineering and Computing
Unska 3, Zagreb, Croatia
E-mail: {ivan.slivar, mirko.suznjevic, lea.skorin-kapov, maja.matijasevic}@fer.hr

Abstract—In-home streaming of online game content to a wide range of low-end and thin client devices via wired/wireless local area connections mitigates the effects of network limitations that pose challenges in directly delivering cloud-based games to such devices. In this paper, we study whether (and to what extent) the rendering and streaming of game content to client devices (from an in-home server) imposes a degradation in gaming QoE as compared to the case of playing using a traditional online game client. We report on the results of an empirical study whereby an in-home game streaming scenario is set up using the GamingAnywhere platform, and participants are asked to provide subjective QoE assessments while playing a Massive Multiplayer Online Role Playing game (World of Warcraft). We test the impact of different delay and loss conditions, emulated along the external Internet link from the online game server to the home. Results show that switching from an online client to an in-home streaming client consistently results in perceived game quality degradations. Further, we report on the “willingness to keep playing” under different network condition in both online and in-home game streaming cases.

I. INTRODUCTION

Game streaming architectures can be categorized into cloud gaming and in-home game streaming based on the streaming server’s location. Cloud gaming involves streaming games over a broadband Internet connection from servers located in the cloud to practically any video-enabled device, with OnLive [1] and StreamMyGame [2] as prime examples of commercial cloud gaming solutions. These platforms, however, have so far had limited success in reaching a mass market due to high operational costs and low possible user base due to high bandwidth and low network delay requirements on the network connection.

On the other hand, in-home game streaming involves streaming of (either online or local) video games to other devices in a home local network, including mobile devices and gaming consoles. This approach is applied in Sony’s Remote Play service, in which the game content is streamed from the PlayStation 4 console to PlayStation Vita handheld device, and in Valve’s Steam In-Home streaming service for the PC gaming platform Steam. In-home game streaming implies video streaming of the game content in a local area network (LAN) via a wired/wireless connection, thereby mitigating one of the major concerns related to cloud gaming – network limitations. Adaptive cloud gaming solutions continuously monitor network conditions and can adjust the video quality according to given limitations, or deny the service if the minimum requirements can not be met [3]. In LANs, network delays are very low and available bandwidth is very high, even for wireless networks (e.g., the new IEEE 802.11ac standard offers theoretical speeds up to 6933 Mbit/s). Therefore, in-home streaming approaches are designed to deliver gaming experiences which are unspoiled by poor network conditions.

This paper is motivated by the following research question: “How does using an in-home streaming platform affect the Quality of Experience (QoE) when compared to playing on a “traditional” game client for a Massively Multiplayer Online Role-Playing Game (MMORPG)?”. We conduct a user study with 35 users in which we analyse an in-home streaming scenario involving multiplayer online game play. The scenario involves participants playing an online game in which players switch from playing on a traditional game client to a thin in-home streaming client under certain network conditions. The network access in this scenario can be divided into two segments: a) local network (e.g., IEEE 802.11 network), and b) Internet connection (e.g., ADSL). In practical terms, we consider the local network as being “unlimited” with very high bandwidth, minimal network delay and loss, while the Internet connection is “limited”, i.e., has limited bandwidth, and is prone to higher values of latency and loss. Therefore, in our testing scenarios we impose degradations (i.e., packet loss and/or latency) on the Internet connection (rather than the local network) to emulate realistic conditions and to analyse how such conditions impact the perceived quality for traditional and streaming game clients.

The remainder of the paper is organized as follows: in section II we briefly discuss related work addressing gaming QoE. Section III describes the laboratory testbed and test procedure. Results are analyzed in section IV, while section V provides concluding remarks and the outlook for future work.

II. RELATED WORK

Numerous works have studied QoE for traditional online games of various genres. In particular for MMORPGS, studies have to a large extent been focused on the impact of network imperfections (delay, loss, jitter) on QoE [4], and to a lesser
extent on the impact of additional context and user-related factors [5]. In the context of cloud gaming, QoE is, in general, determined by the actual set-up of the service, and for this reason it is practically impossible to analyze and observe the effects and results of QoE evaluation in different testing environments [6]. Previous studies in the field of gaming QoE have been performed on different types of cloud gaming set-ups, ranging from commercial services, such as OnLive and StreamMyGame, to custom-built experimental cloud gaming systems. GamingAnywhere (GA) [7], the open cloud gaming system that allows researchers to perform repeatable experiments and confirm reliability of their study findings, provided a big step forward to a common cloud gaming set-up for experimental studies. Several studies have been conducted regarding the performance of GamingAnywhere [8] [7], and indicated that GamingAnywhere in some aspects outperforms commercial cloud gaming systems, specifically regarding response delay and network loads.

The impact of network parameters on cloud gaming QoE has been studied by several research groups. Lee et al. [9] investigated the impact of network latency on cloud gaming QoE while playing different game genres and observed that the frequency of game screen changes and user inputs had a high impact on QoE. Faster-paced games, such as First Person Shooters (FPS), call for a lot of game screen changes and high user interaction, thus not well suited to be played in a cloud environment, where additional cloud server delays result from video encoding/decoding and user input processing. Claypool et al. determined that packet loss higher than 12% resulted in severe drops of frame rate [10]. While high response times cause considerable QoE degradation and user dissatisfaction while playing faster-paced games, gaming QoE for such a game genre is not significantly influenced by packet loss, due to the high frequency of the game screen changes which distracts users from detecting video compression artefacts incurred due to high packet loss [11]. Quax et al. performed a similar study on the impact of the game genre on the user experience in cloud gaming and confirmed that games with high rates of user inputs are more sensitive to network delay than slower-paced games, similar to traditional online gaming. Chang et al. [12] observed the performance of thin clients while playing Pac-Man and developed a frame-based QoE model based on three network parameters: network delay, network packet loss, and network bandwidth. They found that both delay and bandwidth have a higher influence on the user’s performance than network packet loss. On the other hand, Clincy et al. [13] showed that network packet loss had a significant impact on gaming QoE while playing FPS, with packet loss smaller than 1% causing serious degradation of reported user experience. Chen et al. [14] performed an extensive traffic analysis of two commercial cloud gaming systems, but did not explicitly take into account the end user QoE while playing.

While the previously reported studies have addressed various aspects of QoS/QoE for various online games (including MMORPGs) and for cloud games, we have not found previous studies that focused on studying the impact of Internet network connection parameters on the QoE of in-home online game streaming solutions. Given the recent commercial interest in such solutions, we aim to contribute to an understanding of subjective quality perception in this domain.

III. METHODOLOGY

The subjective study was conducted in two phases. In the first phase participants filled in a pre-survey by way of an online questionnaire, and in the second they took part in a two-hour long gaming session in a laboratory environment. The participants were 35 masters level students enrolled at our university.

A. Pre-survey

The pre-survey applied the methodology introduced in our previous study [5]. Participants were asked to fill out an online questionnaire and report their previous gaming experience (with emphasis on multiplayer games and MMORPGs) several weeks before the laboratory testing. 22% reported having previous experience playing MMORPGs and only 2 of 35 participants reported having any kind of experience with cloud gaming. Participants were also asked to rate their perceived skill at gaming as “novice”, “intermediate” or “skilled” gamer. In our sample there were 31.4% novice, 51.4% intermediate, and 17.2% skilled gamers. Along with game experience, we collected data about participant’s demographics, their computer hardware and Internet connection type used while playing online games, motivation for playing games, and their opinion with regards to acceptable delays for different types of games. The results of our pre-survey, in particular as related to previous player experience, were subsequently used in forming participant test groups (as described in the following sections).

B. Laboratory set-up

The laboratory set-up is shown in Figure 1. The game we used was World of Warcraft (WoW). The WoW client was installed on PC1-PC5, five Windows 7 desktops, each with Intel 3.3 GHz i3 processor, 4GB RAM and GIGABYTE Radeon R7 250. With WoW client’s graphic settings set to high, the frame rate was around 60 fps. To emulate a cloud
gaming environment, we used the GA platform, version 0.7.5. The GA server was installed along with the WoW client on PC1-PC5, with default x264 video encoding and decoding settings, and with video bit rate set to 3Mbit/s. We were compelled to use minimal bit rate for video encoding due to limited hardware capabilities of our PCs which could not run the game and encode video stream at higher bit rates, resulting in lower video framerate and framerate drops. Therefore we acknowledge that this experimental design represents lower-end conditions under which would end-user use in-home game streaming. The GA server was running in periodic (desktop capturing) mode, implying that the entire desktop was streaming to the GA client, because we had technical problems with streaming of WoW in event-driven mode. GA clients were installed on PC6-PC10 with Windows 7 OS, Intel 3.3 GHz i3 processor, 4GB RAM, ATI Radeon HD 6450, with default x264 video decoding settings. We used IMUNES [15] network emulator/simulator for manipulating network conditions on the link from the WoW server to PCs 1-5.

As discussed in our previous work [5], two main network parameters affect the QoE of MMORPGs: delay and packet loss. High network delay postpones execution of user inputs on a server and prolongs the delivery responses to the client, whereas high packet loss leads to spikes of network delay due to use of TCP for the particular MMORPG (WoW). Therefore, in this experiment we manipulated both of these parameters. Delay was introduced in the testbed through PC 11 using the previously mentioned IMUNES tool. Three levels of one-way delay were introduced during conducted experiments (75 ms, 150 ms and 225 ms) that increased the average value of RTT by 150 ms, 300 ms, and 450 ms, respectively. We opted for these delay times based on our pre-survey results, and previous studies performed on the GA platform [16]. Note that we did not have control over the Internet connection to our testbed, resulting with the nominal RTT to WoW server between 30 and 40 ms. Likewise, three levels of packet loss were introduced on the same PC using a FreeBSD firewall: 3%, 5%, and 7%. These packet loss percentages were based on our previous study addressing gaming QoE of MMORPGs, which showed that packet loss higher than 10% leads to serious degradation of the gaming experience [5] and data regarding real wireless networks (3G) in which these values can occur [17]. Finally, we manipulated the context in which the game was played in terms of game client, with users switching from playing using a traditional online gaming client and a cloud gaming client.

C. Test procedure

Overall, 35 participants were included in the study, 21 male and 14 female. The average age of the participants was 23, with ages ranging from 22 to 28. The participants were organized into seven groups (five players in each group), based on their reported gaming skill. Each of the formed groups had at least one novice player and one skilled player. Each group had two female players and three male players.

Due to the fact that network parameters (delay, packet loss) were manipulated at three levels (and one additional condition without degradation), a total of 16 different conditions were tested and evaluated during the study. All conditions were tested by each player group. Each test scenario consisted of two phases, during which time network conditions were kept constant: in the first phase, players were requested to play WoW on standard WoW clients running on PC1-PC5 (without the cloud gaming platform), while in the second phase they switched to a PC running the GA client (PC6-PC10) and continued game play. Each phase of a given test scenario lasted 3 minutes. The fact that players knew when they would switch to in-home streaming possibly leads to bias in QoE scores, but without serious modifications of the experimental design we could not conceal this transition. The entire testing session lasted for two and half hours, with a 15-minute break allotted in the middle. Even though this was a very long period for players to hold their attention and focus on game play quality degradations, the majority of participants were highly immersed in the virtual world and engaged in playing with other players during the course of the experiment (based on user feedback). At the beginning of the session, players first played under the best (no network degradations) followed by the worst (400 ms RTT, 7% packet loss) network conditions, and were told that these were reference conditions. After playing under reference testing conditions, for the remaining test conditions players were not aware of the degradation levels of network parameters. Additionally, the sequence of test scenarios was randomly selected across different player groups to avoid a possible bias of manipulated parameters.

During the course of the experiments, there was always a test administrator present who controlled test conditions and provided players with minimal assistance in case of problems with game play (e.g., player getting lost in the virtual environment, dying, etc.). Players were instructed to fill out a questionnaire and provide subjective scores, with respect to criteria given in Table I. For a given test scenario, after the first phase players provided ratings for their overall QoE, and indicated whether or not they would continue to play the game under the current network conditions. In the second phase, players switched to playing (under the same network conditions) on the GA client, and rated the perceived degradation as compared to that in phase 1 and whether or not they would continue to play the game under the current test scenario conditions. We used a standardized 5-point degradation MOS (DMOS) scale [18] for rating the degradation of overall QoE. During the experiment each group of five players was involved in joint actions related to WoW dungeons, meaning they interacted as a group and played cooperatively to survive in the virtual world.

IV. RESULTS

A. Impact of delay and packet loss on QoE for online gaming

The mean values of perceived QoE depending on test scenario conditions are shown in Figure 2. The values for all test scenarios, even for worst scenarios, are relatively high (around 4 which can be considered “good”). In our previous work, in which we tested the influence of various system,
TABLE I
SUBJECTIVE RATINGS COLLECTED DURING THE TEST PROCEDURE

<table>
<thead>
<tr>
<th>Subjective ratings</th>
<th>Rating scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Overall QoE</td>
<td>5 pt. MOS scale (1-bad, 5-excellent)</td>
</tr>
<tr>
<td>2) Degradation of overall QoE</td>
<td>5 pt. DMOS scale (1-very annoying, 5-imperceptible)</td>
</tr>
<tr>
<td>3) Willingness to continue playing</td>
<td>yes/no</td>
</tr>
</tbody>
</table>

user, and context parameters of QoE on MMORPGs [5], we found that players’ QoE was very mildly influenced by added network delay up to 400 ms which was in contrast with findings reported in related work [4]. Our hypothesis to explain this phenomena was that delay degradations were “masked” in players’ perception by other more severe degradations (e.g., jerkiness and frame rate). In this study we have a very similar laboratory experiment, but without degradation of jerkiness and frame rate, and delay shows consistent results with our previous study and a very mild impact on QoE (i.e., added network delays of 450 ms RTT did not reduce the reported QoE below 4). On the other hand, in this study we have involved playing on a cloud gaming client which was graded very low by the players and might be what is causing the masking effect. The differences between our two studies and previous works studying WoW (such as [4]) may also be explained by improvements in the game code under test, or by the fact that a significant number of players participating in the study were inexperienced in playing the specific game.

To quantify the impact of network delay and packet loss we compute Pearson’s product moment correlation $r$, which shows a negative correlation between QoE during traditional online gaming and network delay ($r = -0.34$) and packet loss ($r = -0.29$). In addition to Pearson’s product moment correlation, we applied linear regression analysis of delay and packet loss impact on QoE. It should be noted that we consider our data as interval data and not ordinal (i.e., we consider that the intervals between points on the rating scales are equal). Also, due to the nature of the dataset we visually inspected skewness and kurtosis of data, as well as applying the Rayan-Joiner normality test (similar to Shapiro-Wilk test). The results showed that some of the test scenario results showed a higher level of skewness and kurtosis. It should be noted that while Analysis of Variance (ANOVA) is quite robust on non-normality violations [19] it should be taken into consideration when using the obtained model. ANOVA results show that both delay and loss have significant impact ($p-value < 0.01$), with delay having a stronger impact. The average perceived QoE score in this study can be modelled by the following multiple linear regression model based on only these two parameters (packet loss and network delay)(1):

$$MOS = 4.7059 - 0.00094 \times ND - 5.83444 \times PL$$

where ND represents network delay in milliseconds and PL packet loss. The accuracy of the chosen prediction model is shown in Figure 3. By observing network performances for individual users and predicting their gaming QoE, it could be possible to efficiently manage network resources, optimize game data delivery to end users and, ultimately, increase perceptive QoE affected by network state. In our future work we aim to create more complex models that capture the influence of system, context and user QoE influence factors and are based on not just one, but multiple datasets from separate MMORPG QoE experiments.

Fig. 2. Subjective ratings of overall QoE during online gaming and degradations of overall QoE during cloud gaming

Fig. 3. Accuracy of predicted MOS ratings vs subjective MOS ratings
B. QoE degradation imposed by switching to in-home game streaming

In Figure 2 we show average degradation ratings of QoE in comparison with traditional online gaming while the participants were playing on the GA client for all degradation scenarios. It should be noted that two different measures are shown (MOS for phase one of the test scenario and DMOS caused by switching to GA for phase 2). Such high DMOS values reported by the study participants, indicating severe degradations, may be attributed to default settings of GA (i.e., only 3 Mbit/s). In future studies we will take into considerations various settings.

We once again applied linear regression analysis on network parameters to test the impact on degradation of QoE in cloud gaming. This time analysis shows that packet loss has slightly more significant impact on degradation of QoE than network delay. The average perceived QoE score is modeled by the following multiple linear regression model (2):

\[
DMOS = 3.5578 - 0.00081 \times ND - 5.4473 \times PL. \tag{2}
\]

The accuracy of the chosen prediction model is shown in Figure 4.

C. Relationship between DMOS and MOS

The average QoE degradation levels depending on the overall QoE ratings across all scenarios are shown in Figure 5. In other words, the graph explains how players rate the degradation introduced by switching to GA depending on how they scored the first phase of the same scenario. There is a linear relationship between degradations ratings and overall perceived QoE ratings. This means that the greater the present degradation (i.e., greater latency and loss) in the first test phase of the scenario (inline gaming client), the switch to cloud gaming seems more severe to the player. This relationship is in line with the generic IQX hypothesis postulated in [20], showing an exponential relationship between QoS and QoE, and stating that the change in QoE with respect to QoS degradation depends on the current QoE level. For example, if the current QoE level is very high, addition of a fixed amount of service degradation will cause smaller perceived degradation compared to when the service is already degraded and the same fixed amount of service degradation is added. We illustrate our findings in the context of the IQX hypothesis in Figure 6, showing a generic relationship between QoE and introduced degradations. We note that in addition to network-related degradations due to delay and loss, a constant degradation (shown in the figure as \(x_c\)) is imposed due to the fact that a game is played on the GA client. In line with the IQX hypothesis, three areas of degradation are presented: (1) no distortion perceived, (2) user disturbed, and (3) user gives up. Within most test scenarios, players were located in the “user disturbed” area. It has been shown that that latency up to 150 ms is tolerated by players of MMORPGs [4]) (shown as \(x_{1c}\) in Figure 6), while the values of DMOS presented in the figure actually present the slope of the degradation curve.

D. Impact of delay and packet loss on willingness to play

One of the most important evaluation rating of a some online interactive application is user’s willingness to continue using it under severely degraded performances caused by system or network impairments. Willingness to play results shown in Figure 7 show that the participants have much higher willingness to keep playing under degraded conditions while playing on a WoW client than playing on the GA client, regardless of their gaming skill. This can be attributed to our experiment’s design and the participants awareness of switching to poorer gameplay conditions: after playing in the “best possible” settings on WoW client for a given test scenario, they had to physically move to other PCs and continue playing with degraded game quality (mainly graphic quality). (In future
research we intend to conceal the fact that switch to the in-home game streaming is about to happen or to reverse order of phases in the experiment. When considering the influence of player gaming skill, we observe that skilled players in greater percentage would stop playing after switching to in-home game streaming in comparison with less experienced players. Skilled player are more aware of degraded game performance and even minor system/network degradation have high impact on their eagerness to keep playing. This confirms the findings of our previous research that experienced players are more demanding of game quality.

V. CONCLUSIONS AND FUTURE WORK

In this study we observed the impact of different network delay and loss conditions on QoE during traditional online gaming and in-home game streaming. The analysis shows that while delay and packet loss have statistically significant impact on gaming QoE, the degradations of QoE caused by delay and loss, even in case of very high values, are not that severe. We also examined the degradation of QoE introduced by switching from traditional online gaming to in-home game streaming. Results show that the perceived degradation from switching to cloud gaming client changes and depends on existing degradation of quality under occurring network conditions. We also found that experienced players are more unwilling to play under the same network degradations as compared to inexperienced players. Skilled players, due their previous gaming experience, perceive even minor degradations of game quality and are unforgiving to poor game performances, culminating with their decision to not continue playing. Positive participant’s feedback during experiments showed that widespread use of in-home game streaming is possible if adequate video quality is guaranteed during streaming. In future work we aim to do additional studies on the influence of local area network characteristics (especially wireless connections) on perceived quality of in-home game streaming. Also, we aim to investigate the impact of different streaming settings on QoE. Additionally, putting to test commercial in-home game streaming solutions such as Steam In-Home streaming and comparing them with other commercial or non-commercial in-home streaming systems could be a potential topic for future research.

REFERENCES

[15] “Integrated Multiprotocol Network Emulator/Simulator,” developed at Univ. of Zagreb, Faculty of Electrical Engineering and Computing, online: http://www.imun.es.tel.fer.hr/