Opportunities and Challenges of Using Game Video Stream Data for Games Research

Julian Frommel, Jan Gugenheimer, Katja Rogers
Ulm University, Ulm, Germany
<firstname>.<lastname>@uni-ulm.de

ABSTRACT

Games research is highly dependent on the collection of data such as measures of player-game interaction or realistic player behavior. While collecting such data is frequently a long and tedious process, large video game streaming platforms such as *Twitch* constantly generate vast amounts of game play data. With this position paper, we want to highlight the potential of conducting games research with videos generated through game streaming platforms. We use data of a third-party analytics site for *Twitch* to show opportunities as well as challenges that might impede potential research. We hope this position paper sparks interest in collaborative work on potential guidelines for researchers who want to use this data.

INTRODUCTION

The collection of data is crucial to many areas of games research. Research on player experience such as *Games User Research* [5] frequently requires realistic player behavior to assess how game design affects the players. Similarly, research on artificial intelligence in games is interested in player data to predict the behavior of human players (e.g., [2]).

The collection of such data is frequently a lengthy, expensive, and tedious procedure. There are various approaches such as publishing a game, e.g., as free-to-play, and integrating data collection functionality (e.g., [6]). However, the success of a game in the public game market requires a lot of development time and resources. Being successful in a very crowded free-to-play games market generally requires more than "just" a good game. Because even good games are not necessarily successful, the effort involved just for data collection purposes often does not warrant this approach

for many institutions. They frequently rely on lab studies that provide a controllable environment for specific research questions but can suffer from limitations such as external validity, low sample sizes, and narrow populations.

Yet game play data is generated by streamers on video game stream platforms in large amounts. While this data is not generated for the purpose of research, it could be well suited for various research questions. The *Twitch* science team uses machine learning methods to categorize video clips by themes based on emotes used in the user chat [11]. Similarly, previous work examined how machine learning methods can be applied for highlight detection [10]. While this type of research is often specifically focused on streaming or a streaming platform in particular, other broader games research questions could be studied using such data as well. Academic games research in general might be similarly interested in synchronized multimodal game play data, which frequently consists of a combination of in-game video data, webcam data, and text chat. For example, previous work investigated emotional challenge in games [1]. In investigating such a research question (e.g., "What is the emotional reaction of players to emotionally challenging game situations?"), researchers might want to gather data on players' reactions to emotionally challenging game situations; video game streams contain a wide range of such reactions. Researchers could identify situations where emotional challenge might emerge and study the streamers' responses. However, whether this type of data can be easily used for such a question is unclear.

We explored descriptive data from a third-party analytics site of *Twitch* data as an example to assess the potential of data gathered from video game stream platforms. While this preliminary analysis highlights the potential size of the data base, there are a variety of challenges that potentially could interfere with conducting games research with this kind of data. With this position paper, we want to spark discussion on these challenges of using game videos for games research. We hope to work together on a set of guidelines for researchers who want to use this kind of data to navigate these challenges and avoid potential pitfalls that could limit researchers' findings.

DATA ON TWITCH

It is difficult to gather general insights on data that is available on video game stream platforms, but this information is necessary to assess its value for games research. While previous work has collected data on the content of *Twitch* [4, 7], its statistics collected in 2011–2012 and 2014 respectively do not necessarily provide an up-to-date reflection of current data. Therefore, we use *SullyGnome*[1], a service for analytics and statistics on *Twitch*, to gather first insights. The site states that it queries the *Twitch API* every 15 minutes and aggregates this data [12]. We report data collected on 2019–06–02. Overall, there were 1,347,669 channels recorded in the last seven days up to the collection date. We manually analyzed the top 100 games based on streamed amount of time. There were 9 content categories that can be considered not traditional video game play (see Table 1). The remaining 91

Category
Just Chatting
Art
Slots
Poker
Music & Performing Arts
Talk Shows & Podcasts
Science & Technology
Travel & Outdoors
VRChat

Table 1: Content type of non-traditional video game play in the top 100 streamed "games" from 2019-01-30 – 2019-02-06

[1]https://sullygnome.com

games generated a total of 6,928,550 hours of stream time; this highlights how much data is generated. While certainly not all data can be accessed later, it is likely that many videos are stored on *Twitch* or on other platforms such as *Youtube*, because streamers can access multiple revenue streams by diversifying to other platforms [8]. Overall, 12,281 different games were streamed in the last seven days up to collection date. Even when accounting for non-gameplay channels this showcases the vast range of games.

OPPORTUNITIES

These statistics show the potential benefits this data offers to games research. First, the amount of available data is very large. Time streamed seems a valid statistic for assessing the amount of data, as it is more interesting than time viewed for our purposes: a stream showing game play could be used for games research questions even if there were few or even no viewers. While there might be no interaction with a chat, it still shows game play and potentially video reactions of players. Even if only a marginal percentage of the data can be used, e.g., due to irrelevance, so much data is generated that other approaches of data generation can hardly compare. For example, 0.01% of the examined data equal 693 hours of game play data. This amounts to almost 100 hours per day, which is hardly achievable in lab studies. Second, while interestingly quite some non-game content is generated, the majority of content features videos of video game play. Overall, there was a wide variety of different games. If researchers need game play data of a specific game or genre, it is likely that there will be a stream containing suitable game play. Third, while we currently have no statistics on the prevalence of web cam streams, it appears that many streamers use those in their streams. This would allow insights into player reactions such as facial expressions, body posture, or gaze behaviour that is synchronized to game play, allowing an assessment of player experience in these situations. These aspects show that it could be very beneficial to use this type of data for games research. Finally, game play on *Twitch* is arguably more natural than in controlled lab studies. Therefore, this data could potentially be used to investigate phenomena that occur in natural gaming such as *griefing* (cf. [3]).

POTENTIAL CHALLENGES

While there are great opportunities for games research with this data base, there are some potential challenges that are inherent in the platform that could hinder games research:

- How can researchers select relevant and valid data from a huge data base of available data?
- Which research questions are applicable to this data and which should be avoided?
- How do external motivations such as social factors, monetary incentives, or effects of playing in front of an audience [9] factor into a natural interaction of player and game?

- How to assess information that is usually collected via self-report but not always available (e.g., gender or age), yet vital to allow insights into the generalizability of findings?
- What are the ethical and legal implications of using data that is openly available but not generated for the purpose of games research?

While some questions are addressable, they can certainly impact what research can be conducted to produce valid and sound data with game videos such as those generated through *Twitch* streams.

OUTLOOK

This position paper aims to highlight that there is huge potential in video game stream data for games research, but there are also challenges. Therefore, we hope that this position paper sparks discussion on how to handle these challenges in hope of working towards a set of guidelines that help researchers to use this type of data for gathering valid and generalizable findings.

REFERENCES

[1] Julia Ayumi Bopp, Klaus Opwis, and Elisa D. Mekler. 2018. "An Odd Kind of Pleasure": Differentiating Emotional Challenge in Digital Games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 41, 12 pages. https://doi.org/10.1145/3173574.3173615

[2] Elie Bursztein. 2016. I am a legend: Hacking hearthstone using statistical learning methods. In *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*. IEEE, 1–8.

[3] Thomas Chesney, Iain Coyne, Brian Logan, and Neil Madden. 2009. Griefing in virtual worlds: causes, casualties and coping strategies. *Information Systems Journal* 19, 6 (2009), 525–548.

[4] Jie Deng, Felix Cuadrado, Gareth Tyson, and Steve Uhlig. 2015. Behind the game: exploring the twitch streaming platform. In *Proceedings of the 2015 International Workshop on Network and Systems Support for Games*. IEEE Press, 8.

[5] Anders Drachen, Pejman Mirza-Babaei, and Lennart E Nacke. 2018. *Games user research*. Oxford University Press.

[6] Niels Henze, Enrico Rukzio, and Susanne Boll. 2011. 100,000,000 Taps: Analysis and Improvement of Touch Performance in the Large. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 133–142. https://doi.org/10.1145/2037373.2037395

[7] Mehdi Kaytoue, Arlei Silva, Loïc Cerf, Wagner Meira Jr, and Chedy Raïssi. 2012. Watch me Playing, I am a Professional: a First Study on Video Game Live Streaming. In *Proceedings of the 21st International Conference on World Wide Web*. ACM, 1181–1188.

[8] Mark Longhurst. 2018. A Twitch Success Guide: What Else Should You Be Thinking About? https://medium.com/the-emergence/a-twitch-success-guide-what-else-should-you-be-thinking-about-42e21471dd6e. Accessed: 2019-02-06.

[9] Anthony J Pellicone and June Ahn. 2017. The Game of Performing Play: Understanding Streaming as Cultural Production. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 4863–4874.

[10] Charles Ringer and Mihalis A Nicolaou. 2018. Deep unsupervised multi-view detection of video game stream highlights. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*. ACM, 15.

[11] Ariel Sagalovsky. 2018. Using chat emotes as signals for content themes. https://blog.twitch.tv/using-chat-emotes-as-signals-for-content-themes-13c31d4a0f74.

[12] SullyGnome.com. 2018. SullyGnome – detailed Twitch statistics and analysis: About SullyGnome.com. https://sullygnome.com/about. Accessed: 2019-02-06.