A Study of Cheating Identification and Measurement of the Effect in Online Game

Emergent Research Forum Paper

Daehwan Ahn
Seoul National University
penris32@snu.ac.kr

Byungjoon Yoo
Seoul National University
byoo@snu.ac.kr

Abstract

As online games become one of the most popular Internet services, cheating has also emerged as a noticeable and important phenomenon in recent gameplay on the Internet. In regards to this potential threat, game companies focus on their efforts to protect their game from unauthorized acts throughout hacking tools under the assumption that the use of hacking tool could be particularly harmful for success of games. However, we present the usage of hacking tool causes an 102% and 79% increase in playtime and purchase respectively right after users start to use hacking tools. Also, the effects of hacking tools are not just short term. We show it is possible that cheaters can provide an online game company with more leverage and loyalty in aspect of playtime and purchase. This is exactly a contrary result to the conclusions of numerous previous studies on piracy on digital contents.

Keywords

Online game, MMORPG, cheating, hacking, user behavior, identification, effect.

Introduction

This study is on a problem of cheaters that a Korean game company faces. The game company of this research has provided MMORPG service for game users for ten years. The game users grow their avatar and make money that is needed for game playing by doing time consuming activities such as battles with monsters. If the players purchase items in games by using their cash, it makes the users save time by making their avatar get stronger easily. Players can enjoy the game for free but the game companies earn profits by selling such cash items to the users who are willing to pay more for advantages in the game.

The company faced that revenues of the game have decreased for the last two years. They suggested that one of the reasons is the existence of cheaters but it was not sure if the hypothesis is correct or not. Cheaters mean that one user logs in several IDs at the same time, and automatically make their avatars go for battles by using hacking tools. Generally, in order to make their avatars get stronger and make money, game players need to buy cash items or do battles by spending their time. However, if they use hacking tools, they can automatically make their avatars go for hunting and make money even when the users are asleep. As a result, it might be unnecessary for the game players to buy cash items to make their avatar become stronger. Even though the game company is aware of the existence of cheaters, they cannot make a decision if they have to ban the cheaters. The reasons are 1) they cannot identify who the cheaters are, 2) they cannot know whether the existence of the cheaters actually have an influence on their revenues. Cheating repeatedly occurs in most multiplayer online games, but it is hard to measure (Pritchard, 2000). Moreover, it is difficult to conduct data analysis using game data sets because the data is not stored systematically. Most companies tend to design their database not for efficient data analysis but for their optimized operation of game. Consequently, it is important to conduct data analysis using limited data sets in the game industry.

In this study, we identified cheaters and measured impacts of them using just simple and basic information such as users’ ID, playtime, and item purchase collected from a Korean online game company. With respect to identifying cheaters, we used algorithms named DTW (Dynamic Time Warping)
and JWD (Jaro–Winkler distance). We also measured effects of hacking tool usage using pseudo experiment which is composed of DTW and DND (Difference in Difference). As a result, there are 0.25% of cheaters and they hold approximately 12% of revenue. Furthermore, the usage of hacking tool causes 102% and 79% increase in playtime and purchase respectively right after users start to use hacking tools. According to additional analysis, it is shown that the positive effects of hacking tools are not just short-term. Our granger causality test also reveals that cheating users’ activity does not affect other normal users' purchases or playtime.

**Literature Review**

There are some past studies which examine positive impacts of the cheating or abnormal behaviors in online game. Gal-Oz and Zuckerman (2015) argue that cheating behavior in gamificated fitness application could actually be beneficial for encouraging physical activity. Kirman et al. (2012) suggest that mischief could serves a positive role, vital social role, in game and other applications.

One of the literature closely related with online game cheating is piracy on digital good market. With respect to piracy on digital good market, a large number of research argue that piracy has a negative effect on legitimate sales of digital contents. Ma et al. (2011) find out that pre-release movie piracy leads to a 19.1% decrease in box office revenue in comparison with piracy that occurs post-release. Hennig-Thurau et al. (2007) discover that movie piracy gives rise to substantial cannibalization of theater visits, DVD purchases and rentals responsible for yearly revenue losses of $300 million.

According to the results of literature on piracy on digital contents such as movie and broadcasting, illegal consumption of contents has a negative effect on legitimate sales in most cases. However, piracy on digital contents and cheating in online games have obvious different characteristics among them. We are not assured that the results of previous research on digital good consumption can be applied to online game cheating cases directly.

**Data**

We collected our data from a Korean online game company such as users’ ID, daily playtime and details of cash item purchase from 2014 Feb to 2016 Jan. A total of registered users are 80,368 and 30% of the users (Approximately 24,000) are active users who actually play the game. A small number of top 10% users (8,000) hold 80% of the playtime and top 2.9% of the users take up 80% of the revenues. This implies that a few number of users have big impacts and show their loyalty to the game. This phenomenon can easily be seen in other games. Also, it means that cross-subsidization effects in game platforms are comparatively bigger than other platform business.

**Method**

*Cheating Identification*

In this study, we chose to use rule-based method over data mining-based method requires enough data quality and quantity (Phua et al., 2010). One of our goals in this study is to develop methods for game companies even if they have little data. For this reason, we chose a rule-based approach.

We found that cheaters of this game log in a number of different IDs and use various avatars at the same time by using hacking tools. The reason why the game users use different IDs is that the hacking tools have functional limitations. In MMORPG, the users must win much stronger monsters in order to gain better rewards. However, if hacking tools are used, it is much harder to beat stronger monsters because the hacking tools cannot respond diverse situations as humans do. Therefore, in order to fight with stronger monsters, the game players use a way of cooperating with many different avatars by using different IDs. This type of cheating is commonly found in other online games, too.

In this study, we made a rule in order to identify who cheaters are by using distinctive characteristics of cheaters. For example, if a cheater uses different IDs, these different IDs show extremely similar playtime patterns. Figure 1, 2 and 3 below show different playtime patterns of the game users. In these figures, x-axis means each month and y-axis is the sum of playtime in the month.
Is Cheating Always Bad?

Figure 1. Each Playtime Pattern of 20 Users Who Have the Same Playtime on Average

Figure 1 displays playtime patterns of 20 different users who have the same playtime on average. We can find that each of the users has extremely diverse playtime patterns and those playtime patterns barely overlap one another.

Figure 2. Playtime Patterns of Normal Users (left) and a Cheater (right)

The left graph in Figure 2 shows playtime patterns of two different users who have the same average playtime. Generally, users who do not use hacking tools show such different playtime patterns. The right graph in Figure 2 represents playtime patterns of one cheater who uses two different IDs at the same time. We used DTW (Dynamic Time Warping) method in order to capture the playtime patterns of cheaters. DTW is a robust distance measure to calculate the distance between two different time-series data (Berndt and Clifford, 1994). The degree of differences between two time-series graphs can be represented numerically by using DTW. For example, in Figure 2, the DTW score of two different users who have similar playtime on average is 61,899 which is relatively high. However, the DTW score of two different IDs which are owned by one cheater is only 4,197.

Figure 3. Various Playtime Patterns of Cheaters

We calculated and compared DTW scores between different IDs every month in order to identify diverse types of cheaters such those who use hacking tools for a certain period of time only and who add new IDs in the middle of game playing with existing IDs (see Figure 3).

We set another rule which can complement the DTW score. As seen in Figure 3, cheaters tend to use similar IDs because it is easy to memorize. For example, they make many IDs by changing additional numbers based on same characters such as ‘blue838’ and ‘blue83’. We used those characteristics when DTW scores could not assure whether some particular IDs are owned by one cheater or not.

We used an algorithm named JWD (Jaro–Winkler distance) to compute similarities between IDs. Jaro-Winkler distance is a measure of similarity between two strings and is designed and well-fitted for short strings such as person names (Winkler, 1990). Because IDs are composed of short strings, JWD is an appropriate way to calculate distances between IDs.
Even though JWD scores between IDs are not a perfect rule which can identify cheaters but it is quite useful in aspect of practical usage. From results of analysis that we performed, we can identify more than 70% of cheaters by calculating similarities between IDs. However, although JWD analysis is a good method for cheating identifications especially for multi-ID usage, there is a critical reason why this method cannot be used alone. JWD analysis tends to misidentify normal users as cheaters if it is used alone. For this reason, this method is best suited for complemental ways to identify cheaters. In this context, we used JWD analysis when we roughly choose and guess cheater group before precise analysis is conducted. We also conducted JWD analysis when we are not sure whether some particular IDs are owned by one cheater or not with DTW analysis only.

**Measurement of Hacking Tool Usage Effect**

The method commonly used to measure causal effects of time-series data is pseudo experiment using PSM (Propensity Score Matching) and DND (Difference in Difference) (Lechner, 2011; Dehejia et al., 2002). In this study, we measured the change amount of playtime and purchase affected by hacking tool usage by conducting the pseudo experiment. However, since there are not sufficient explanatory variables which can be used for PSM, we replaced PSM to DTW in order to select control group which have similar playtime and purchase patterns with experimental group. The method that find the best matched control groups with time series data using DTW begins to be commonly used in some domains recently such as finance, medical and marketing (Larsen, 2016; Yuan, 2014).

For more details on the pseudo experiment that we performed, as experimental group, we selected all of the 25 users who do not use hacking tools at first (pre-period) then start using them later (post-period). As control group, we selected 250 users who have similar playtime and purchase patterns in pre-period by using DTW. Ten users in control group are matched to one user in experimental group (Larsen, 2016). We tested whether each of the user groups changed their playtime and purchase patterns in post-period depending on hacking tool usage and non-usage. We also measured the change amount of playtime and purchase between control and experimental groups.

**Result**

**Cheating Identification**

We found that 207 game users (= 1,026 IDs) are cheaters. It was also shown that they tend to log in five different IDs at the same time on average. One interesting point is that although they are only 0.25% of the whole registered users of the game but they are VIP users whose playtime holds 6% and 12% of the whole revenues.

**Measurement of Hacking Tool Usage Effect**

We tested how game users’ playtime pattern and purchase pattern changed before and after they started using hacking tools for three months respectively (a total of six months). As a result of DND analysis, the users who used hacking tools increased their playtime (18,504 minutes, 102%) and purchase (23,642 won = $ 20, 79%) dramatically compared to those who did not use hacking tools.

Generally, the duration of cheaters can be shortened since they use up the game contents by using hacking tools. For this reason, we checked whether the effects of hacking tools are just short-term or not. We compared play patterns of cheating group and their control group who has similar average playtime and purchase with cheating group. Surprisingly, according to our results, cheaters tend to keep spending playtime and purchase steadily compared to non-cheaters for two years. It explains that the effects of hacking tools are not just short-term.

Additionally, we tested whether the presence of cheating users would have a negative impact on other normal users. We performed a granger causality test to determine whether cheating users’ time series is useful in forecasting other normal users’ time series. As a result, playtime and purchases of cheating users cannot provide statistically significant information about future values of normal users’ playtime and purchases. In other words, cheating users’ activity does not affect other normal users’ playtime or purchases.
Conclusions

In this study, we aim to identify cheaters with limited data and examine how hacking tools change users’ playtime and purchase behavior. Consequently, we found that there are 207 users (0.25% of registered users) who use hacking tools and they are VIP users who hold approximately 12% of the whole revenue. Moreover, it was found that the users’ playtime and purchase increased to 102% and 79% respectively right after users start to use hacking tools. We also showed that the effects of hacking tools are not just short term. We assume that these phenomena can be seen for following reasons. By using hacking tools, utility of users’ time spending on game increase since hacking tools make game playing more efficient. Also, they can boost their unauthorized advantages through hacking tools by purchasing cash items. As a result, the increased utility of hacking tool usage makes hacking users spend more time and money. In our opinion, this is why hacking behavior increases users’ playtime and purchase.

We also showed through the granger causality test that the behavior of cheating users does not affect the playtime and purchase of other users. We think of two reasons for this result. First, it seems that there are no big negative effects that cheaters can affect to normal users for now since the number of cheaters is extremely small. Second, since this game has been in service for a long time, only loyal users remain. These loyal users may be dissatisfied with illegal cheating by cheating users, but they do not seem to drop out of the game or reduce their purchases. In this context, the results of our research can be derived differently in other games that have more sensitive new users.

Our study does not mean that game companies should allow users to cheat or hack. We want game companies to build better strategies that will benefit both users and companies based on our research. For example, our study shows that users’ automatic play can increase the user’s playtime (=interest) or purchase. In this respect, the decision of some game companies to incorporate an official automatic playing system into their games is correct. They took the advantage of hacking tools while preventing cheating in a clever way.

References