Churn Prediction through Matchmaking Information Analysis in PvP Mobile Game

Ui Beom Kang
Korea University
Seoul, Korea
howard1005@korea.ac.kr

Huy Kang Kim
Korea University
Seoul, Korea
cenda@korea.ac.kr

CHI'19, Extended Abstracts,, October 22–25, 2019, Barcelona, Spain
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract
We analyze the effect of the player matchmaking on user's churn in a mobile casual game. Providing proper opponent matching in Player vs. Player (PvP) genre of mobile game is one of the significant factors to retain users. If the matchmaking in a game play is not done properly, users may be dissatisfied and then leave the game due to loss of interest for the game play. In this paper, we propose a novel matchmaking method that can minimize the churn rate and propose a new churn prediction model. To this end, we adopt a machine learning method that can predict the user's churn by carefully selected features highly related to a game user's satisfaction. We test our churn prediction model to one of the famous PvP mobile game in South Korea. As a result, we can successfully find the most significant features that can affect the game user's churn. We achieve the game user's churn prediction with the high accuracy value of 82%.

Author Keywords
Matchmaking; machine learning algorithm; churn prediction model.

CCS Concepts
•Human-centered computing $\rightarrow$ **Human computer interaction (HCI);** *User models;* **•Applied computing** $\rightarrow$ **Computer games;**

Introduction

With the rapid growth of the mobile game market due to the popularity of smart phones, the mobile game industry has reached its peak. While mobile games have the advantage of being able to do it anytime and anywhere with a mobile phone simply, users of mobile games seem to play games briefly and leave immediately. User churn in a mobile game can be a critical factor that can damage game service maintenance and affect other users in the game [8]. Well-designed churn prevention technique is crucial to the game's longevity, especially for mobile games [11]. Recently, there is a tendency that PvP genre games increase in the mobile game market. As a result, the methods that match with the optimal opponent are a significant factor in mobile game and it is necessary to think about how matching with the opponent is the best match.

In PvP genre of online games, it is essential to match multiple users (N vs. N) between teams or individuals with appropriate opponents having similar play skill. In this regard, a technique for accurately measuring the difference in skill between user and opponent has been studied for a long time. There is an *Elo* rating system or Glicko, which was developed for measuring skill in a chess game for a long time ago [4, 5, 6].

More advanced ranking system such as 'True Skill' and 'True Skill2' provided the representation method of the skill score expressed in numerical value. However, the *Elo* System only shows the numerical value of an individual's skill, and matching users of a similar number with this numerical value is not optimal matching. The best matching is a matching that can maximize the satisfaction of the user who plays the game and motivation to continue the game.

In this regard, the matching system can also be used for predicting churn in online games. In particular, the matching system can be used as a useful churn factor because users tend to frequent churn in a mobile game. If the skill gap between users is too big, then the game will flow into a simple one direction that winning or losing. This mismatch will make a low-skill user feel deprived or a high-skill user get bored. If a user matches with an opponent with similar skill level, then the user can feel thrived. However, the on-going matches are always matched with the similar skill level of users, then the user can feel too tired and eventually leave the game (i.e., become a churner). Also, all the matches are like the two faces of coin. That is to say, one user in the mismatched game can feel deprived, while the other one can feel excited. Thus, the well-balanced and well-designed matchmaking scheme plays a great role of game's longevity and game user's satisfaction.

There are some difficulties in predicting user churn in mobile games. First, it is difficult to precisely define the churn because the playing period of the users is extremely intermittent due to the environment of the mobile game. Second, there is a need for an additional analysis process to identify the characteristics that can affect the churn from consideration of the genre and design factors of the game.

In this paper, we identify a churned user by establishing a reasonable criterion for churn prediction. We extract only applicable features for games using PvP genre using matching-related features, suggesting a model that does not require any additional analysis process depending on the game environment.

Related Works

We classify the existing research related to churn prediction into two parts, the churn prediction and matchmaking system design.

Prediction of user churn in online games has been studied in several ways. Borbora *et al.* [2] extracted activity logs of users in a massive multi-user online role-playing game (MMORPG) and analyzed activities by using a lifecycle-based approach. They did a comparative analysis of which behaviors occurred during the weeks between the churn and non-churn users. They did a distance-based classification model called as wClusterDist using behavior patterns of churn users and non-churn users.

Borbora *et al.* [1] investigated the churn problem in a MMORPG and developed a model that integrates layer-synchronous theory. The clustering technique was used to segment the dataset, and then the ensemble model was applied to the divided datasets to improve performance. Experimental results showed that the ensemble model was far superior to single classification model for churn prediction.

Kim *et al.* [9] proposed a model that define the user churn in existing online games. They used restrictive and straightforward criteria such as withdrawal of account or the period that the last game activity occurred. However, the existing churn prediction models could not be applied to the live mobile casual games service. They noticed the mobile casual game players tended to churn frequently and did not make periodic payments such as game fees. Based in the observation, they defined new churn criteria and suggest the comprehensive concept of the observation period (OP) and the churn prediction period (CP) with more flexibility.

Stroh-Maraun *et al.* [13] proposed the Nested Matching, an efficient matching system to increase the player's survival rate, in regards to 'Game Experience' factors such as the user's play style and skill level. They showed that the matching method with the user's characteristics feature was superior to the existing matching system only relied on the skill feature.

Sifa *et al.* [12] developed the PvP recommendation framework called the 'Destiny' based on the hybrid multi-profiling of the user's game play pattern. To generate the multi-profile of users, 'Destiny' used the K-means clustering algorithm with the user's stats and play styles. They presented the 'Destiny' system's user survey that about 80% of the framework's users said the 'Destiny' was practically helpful.

Chen *et al.* [3] proposed the EOMM (Engagement Optimized Matchmaking) framework. In order to maximize the engagement of users, they proposed unique formulas. They simulated the matchmaking formula with the game produced by Electronic Arts. Experiment results showed that the performance of EOMM was higher than other matchmaking algorithms such as WorstMM, SkillMM and RandomMM.

Herbrich *et al.* [7] proposed a Bayesian-based skill rating system called 'True skill' and 'True Skill2' [10]. In 'True Skill2', they improved the performance by adding user experience factors not considered in 'True Skill'. This system improved the existing *Elo* system by estimating a user's skill value more accurately.

Experiment data

The dataset used in this paper is a full game play records of M game, one of the most popular mobile games in South Korea. The data is large log data recorded for about two months (from November 25, 2018 to January 18, 2019) and big data of about 5.53 TB consisting of 124 tables and 1,547 fields. Each table contains information about the user's status, actions, and events. We filter four tables related to matching among a total of 124 tables as shown in Table 1. It is composed of information to be possessed if it is a game of PvP format, such as connection information, matching information, billing information and item in-

Table Name	*Fields*
Session Info	PlayerID, Time, Character
Matching Info	MatchingPoint, Level, Win, Lose, PlayTime, Status
Payment Info	Money in game or real world
Item Info	Items affecting player's skill

Table 1: Table selection related to Matchmaking.

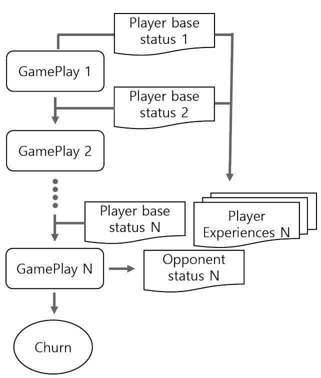

Figure 1: Feature extraction scheme in N-th game play.

formation. In the session info table, we utilize information such as a user ID, a character to be used and a connection time. In the matching info table, we utilize information such as a matching score, a level and a victory or defeat. The payment info table and the item info table use only the elements that can affect the user's skill value. In many case of PvP mobile games, billing and purchase of item information are used for the matchmaking because using high value of items are highly related to a user's skill value.

Definition of churn user

To classify churn user in the mobile casual game, the OP and the CP are set separately as like Churn Definition 1 in Fig 2. We can decide whether or not to churn if by observing the CP [9]. However, this method has a problem that the observation period is fixed and applied to to all users, not considering for each user's condition. For example, it does not work well when a user churns several times where the analysis period is set too long. To solve this problem, we set a single play as the OP. Also, we set the CP as a period between two game plays in succession. If a user does not connect for a certain period, it is defined as a churner as shown in 'Churn Definition 2' of Fig 2. Accordingly, it has a merit to estimate a user will churn or not whenever the game play finished.

Feature pre-processing

In the PvP genre online game, the game play is not a consecutive but a non-sequential and it requires matching between users. After a game, users play different games with other users. That is, each game is divided into several pieces independently. Therefore, it is necessary to process information based on one game. There are three main factors that affect the churn of a game. They are the current state of the user, the experience gained from the previous game, and the difference between the matched user and

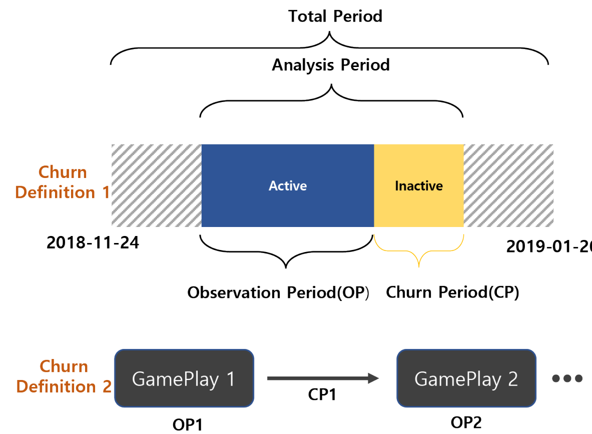

Figure 2: Churn definition based on Game play.

the reference user.

Therefore, the game information to be processed per game is divided into three major parts as shown below.

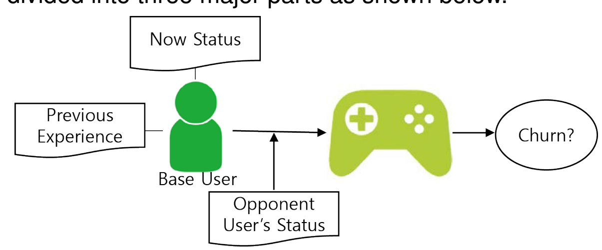

Figure 3: Three factors that can affect churn after a game.

1) Information indicating the status of the reference user
2) Knowledge of the matching user
3) The user's recent match experience information

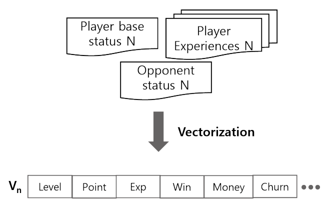

Figure 4: Feature vectorization for the N-th game play.

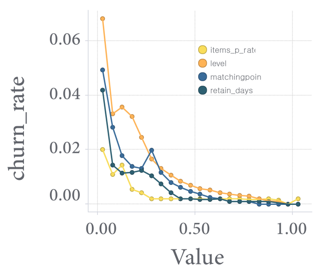

Figure 5: Line graph about Detail features.

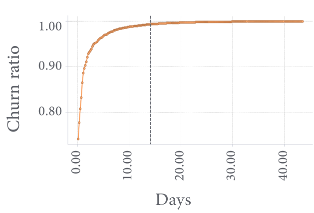

Figure 6: User Churn Ratio CDF.

Fig 3 shows three factors available in the churn prediction model of the one game standard. The proposed model is that determines whether or not churn after the game, depending on the current state of the user, the experiential elements of the previous matching, and how high ability the matching opponent has in comparing with own ability.

In this paper, the N-th game play of a user is regarded as independent data because the model predicts each user's churn based on one game. Therefore, the information extracted from the N-th game play can be regarded as the accumulated data of the game played N times and the state value before the N-th game play. Fig 1 shows that the information extracted from N games of a user is accumulated and becomes experiential data of the user.

In other words, for each game, it is necessary to process the difference between the user's existing state amount, the experience of the previous game. Then, We vectorize the extracted features for each element of the game, and the vector for the N-th play can be expressed as V_n as shown in Fig 4.

We select the features that can be used as a decisive factor for the user's churn through quantitative analysis. In total, there are 36 detailed features divided into three categories according to the existing user status, the relative user status, and the existing user's match experience. In Table 2, it shows 15 features out of 36 features.

Graphs of the detail features are shown in Fig 5 and Fig 7. We select features with discrimination between churn and non-churn users. When we draw box plots graph with Level, Matching point, retain_days, and item_p_rate as representative features, we can see that the distribution of value is obviously different. In Fig 5, the churn rate is significantly different depending on the value. Because these four features can have a significant impact on the distinction between churn and non-churn user, we adopt those features.

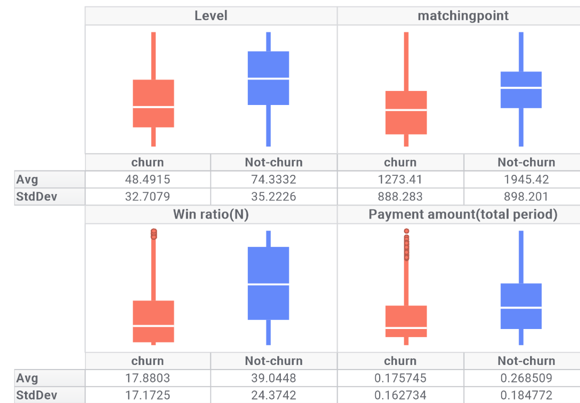

Figure 7: Box plot graph about Detail features (Level, matchingpoint, Win ratio(N), payment amount (total period)).

Learning Model

In churn prediction model, We use various machine learning models. We propose a classical learning model called Logistic Regression (LR), Multi-Layer Perceptron (MLP) using multiple neural networks, Deep Neural Network (DNN), Random Forest and XGboost. Each model is learned by dividing learning data and evaluation dataset and estimated whether or not to churn as an evaluation dataset. Finally, we compare or analysis metrics which are Accuracy, Precision, Recall, and F1-score.

Churn User Analysis

We define churn by using the length of idle time not connected since one game, so an appropriate idle time length should be set. In this paper, we set the length of the idle

Type	*Feature*
Condition (Base User)	Level, Matchingpoint, Win ratio (Total), Playtime (Day), Payment amount (Total)
Condition (Opponent User)	Level difference, Matchingpoint difference, Win ratio (Total), Payment amount (total period)
Experience (Base User)	Win ratio (N), Sequence (N), Win ratio (Day), Sequence (Day), Average Matchingpoint difference (N), Average Matchingpoint difference (Day)

Table 2: Detailed feature elements based on the existing state of base user, experience, and state of opponent. (Total: the observation period is the whole of the log, N: feature unit is number, Day: feature unit is day)

time to 2 weeks (14 days) defined by the game's service provider. We examine the distribution of the churn users. As shown in Fig 6, if the standard idle period of churn is set to 14 days, about 5% of the total users are distinguished as exit users. 5% user classified as a churn user is about 254,043 in terms of the number of games.

Experiment Result

To learn each machine learning model, we additionally pre-process features for each model. If 254,043 games classified as churn and 2,286,387 data classified as non-churn are put into INPUT as they are, there may be a problem that the weight is biased towards the non-churn. To solve this problem, we randomly extracted data as many as the number of churn from non-churn. Therefore, the number of churn and non-churn is 254,043, which is 1:1, and we divide the learning data and evaluation data into 7:3. Then, we make 6-machine learning models to learn the data and evaluate the results of models. Table 3 is the result of evaluating each machine learning model. Overall, each model achieve about 70% to 80% of accuracy and MLP show the best performance at about 82%.

Model	Accuracy	Precision	Recall	F1-score
LR	0.8192	0.8285	0.7911	0.8094
MLP	0.8246	0.8209	0.8302	0.8255
DNN	0.7362	0.7287	0.7528	0.7406
LSTM	0.7802	0.7559	0.8277	0.7902
RF	0.8124	0.8189	0.8019	0.8103
XGBoost	0.8108	0.8122	0.8076	0.8099

Table 3: Evaluation results for several machine learning models.

Conclusion

In this paper, the maximum performance is about 82%, which is similar value to other papers on churn prediction model. However, we have only used information related to matching. In this regard, this paper is meaningful in predicting churn users with only information related to matchmaking from PvP genre mobile games. As we can see in Figure 8, the prediction value of the machine learning model is linearly distinguished between the churn user and the non-churn user. These results show that this churn prediction model can classify the churn and non-churn users reasonably. Also, our model can be applied to the game matchmaking systems. With our mode, We expect game service proivders can provide better matchmaking system to game users.

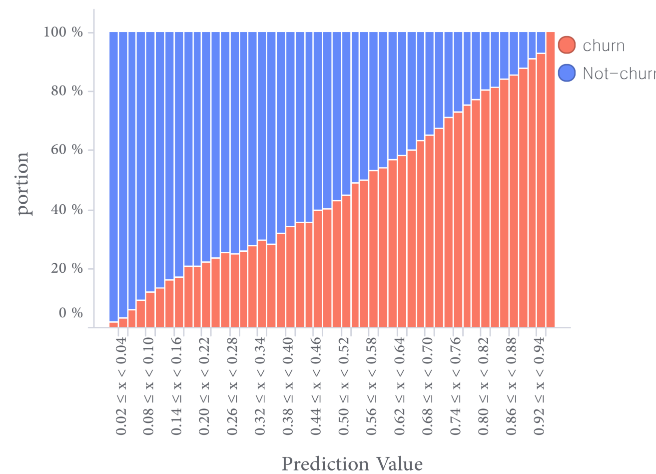

Figure 8: Ratio of churn and not-churn about evaluation result.

REFERENCES

1. Zoheb Borbora, Jaideep Srivastava, Kuo-Wei Hsu, and Dmitri Williams. 2011. Churn prediction in mmorpgs using player motivation theories and an ensemble approach. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. IEEE, 157–164.
2. Zoheb H Borbora and Jaideep Srivastava. 2012. User behavior modelling approach for churn prediction in online games. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*. IEEE, 51–60.
3. Zhengxing Chen, Su Xue, John Kolen, Navid Aghdaie, Kazi A Zaman, Yizhou Sun, and Magy Seif El-Nasr. 2017. Eomm: An engagement optimized matchmaking framework. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1143–1150.
4. Rémi Coulom. 2007. Computing elo ratings of move patterns in the game of go. *Icga Journal* 30, 4 (2007), 198–208.
5. Mark E Glickman. 1995. The glicko system. *Boston University* (1995).
6. Mark E Glickman and Albyn C Jones. 1999. Rating the chess rating system. *CHANCE-BERLIN THEN NEW YORK-* 12 (1999), 21–28.
7. Ralf Herbrich, Tom Minka, and Thore Graepel. 2007. TrueSkillâĎć: a Bayesian skill rating system. In *Advances in neural information processing systems*. 569–576.
8. Jaya Kawale, Aditya Pal, and Jaideep Srivastava. 2009. Churn prediction in MMORPGs: A social influence based approach. In *2009 International Conference on Computational Science and Engineering*, Vol. 4. IEEE, 423–428.
9. Seungwook Kim, Daeyoung Choi, Eunjung Lee, and Wonjong Rhee. 2017. Churn prediction of mobile and online casual games using play log data. *PloS one* 12, 7 (2017), e0180735.
10. Tom Minka, Ryan Cleven, and Yordan Zaykov. 2018. TrueSkill 2: An improved Bayesian skill rating system. *Tech. Rep.* (2018).
11. Kunwoo Park and Meeyoung Cha. 2017. Churn Analysis of Maximum Level Users in Online Games. *Journal of KIISE* 44, 3 (2017), 314–322.
12. Rafet Sifa, Eric Pawlakos, Kevin Zhai, Sai Haran, Rohan Jha, Diego Klabjan, and Anders Drachen. 2018. Controlling the crucible: A novel pvp recommender systems framework for destiny. In *Proceedings of the Australasian Computer Science Week Multiconference*. ACM, 39.
13. Nadja Stroh-Maraun, Daniel Kaimann, and Joe Cox. 2018. More than skills: A novel matching proposal for multiplayer video games. *Entertainment Computing* 25 (2018), 26–36.