

RESEARCH

Open Access



Social networks for enhanced player churn prediction in mobile free-to-play games

María Óskarsdóttir^{1*}, Kristín Eva Gísladóttir¹, Ragnar Stefánsson¹, Damian Aleman² and Carlos Sarraute²

*Correspondence:
mariaoskars@ru.is

¹ Department of Computer Science, Reykjavík University, Menntavegur 1, 102 Reykjavík, Iceland

² Independent researcher, Buenos Aires, Argentina

Abstract

Social networks have been shown to enhance player experience in online games and to be important for the players, who often build complex communities. In online and mobile games, the behavior of players is bursty as they tend to play intensively at first for a short time and then quit playing altogether. Such players are known as churners. In the literature, several attempts have been made at predicting player churn in online and mobile games using behavioral features from the games' player logs as input in supervised machine learning models. Previous research shows that information from social networks provides alternative and significant information when predicting churn, and yet the importance of networks has not been fully researched in mobile gaming. In this research, we study player churn in a mobile free-to-play game with one-versus-one matches. We build two types of networks based on how two players are matched. We train churn prediction models with features extracted from the networks to evaluate their predictive performance in terms of churn. Furthermore, we predict churn using the players' behavioral features during their first day of game playing. According to our results, the network features greatly increase the predictive performance of the models, indicating that they carry alternative information about intention to churn. In addition, the first-day features are quite predictive, which means that first day activity is sufficient to predict churn of players quite accurately, validating the bursty behavior. Our research gives an indication of which aspects of game playing are associated with churn and allow us to study influence and social factors in mobile games.

Keywords: Mobile games, Churn prediction, Social networks, Implicit networks, Explicit networks, Burstiness

Introduction

Churn prediction models are a commonly used tool in industries operating in competitive markets where customer turnover is high. From an organizational perspective, retaining an existing customer is the economical choice in contrast to attracting new customers, where, in addition, long-term customers are less costly to serve, tend to buy more and be more likely to promote the organization in their social network and thus bring in new customers through word of mouth (De Caigny et al. 2018).

In the mobile and online gaming industries, the behavior of players –i.e. the customers– is bursty, meaning that new players tend to play intensively at first and

only for a short period of time and then stop playing altogether (Karsai et al. 2011; Petersen et al. 2017). Predicting player churn in this setting is challenging, because of the high playing intensity and heavy tailed activity data which is only available for a short time (Petersen et al. 2017; Hadiji et al. 2014; Seufert 2013). As a result, the gaming providers have very limited leeway to react to the signs of their customers churning, and to make them offers to encourage them to keep playing. Clearly, it is in the interest of the gaming companies to have good churn prediction campaigns.

Social networks in online and mobile games are of great importance for the players and their experience (Wallner et al. 2019; Loria et al. 2021). They appear for example as communities and clans of players in multiplayer online and battle games and in social network games that integrate the social features of the digital social network on which they are built (Klauser et al. 2013). They have been shown to affect in-game performance, frequency and length of play and engagement, to name a few (Rattinger et al. 2016; Pirker et al. 2018).

Based on our previous research, information from social networks provides alternative and significant information when predicting churn in the telecommunication industry (Óskarsdóttir et al. 2016, 2017). Our research has established the importance of considering networks and the rich information contained in the relationships between entities, since they are often subject to influence from their neighbours and connected entities show similar behavior, which can be explained by homophily and assortativity Óskarsdóttir et al. (2017, 2016, 2021, 2019, 2018). Indeed, features extracted from the social network of customers representing the churn status of their contacts are highly predictive of churn. However, the importance of networks has not been fully studied in the mobile gaming industry, and in particular not in relation to player churn.

In this paper, we study the importance of networks on player churn in a free-to-play mobile game with one-versus-one matches. We build two types of networks based on how the players are matched, with either explicit connections, in the case of friend matches, or implicit connections, when the match is based on the players' similarity. We extract features from the networks and train churn prediction models with combinations of behavioral and network features. Our goal is to evaluate whether information from the two social networks enhances the predictive performance of player churn prediction models. According to our results, there is considerable added value of including network variables. In addition, we zoom in on the players' behavior in the first day of game playing and use such information to predict player churn. This allows us to quantify the burstyness of players in terms of the models' performance, that is, how quickly they churn. Our results show that the information from the players' networks adds information that is complementary to the simple player features, and that first day activity is sufficient to predict churn of players quite accurately.

The rest of this paper is organized as follows. In the next section, we discuss the literature related to social networks in gaming, on the one hand, and, churn prediction in online and mobile gaming, on the other hand. Then we present our methodology followed by the results of our experiments. The paper concludes with a discussion and steps for future work.

Related work

Social networks in gaming

Various types of online and mobile games exist, which offer different opportunities for social interactions and the effect they have on the players.

For example, in collaborative team-based games, social ties influence players' performance, where playing with friends is beneficial for low-skill players but not for high-skill ones (Zeng et al. 2019). Social networks in online multiplayer shooter games have been researched and social network behavior profiled (Rattinger et al. 2016). This has revealed that nodes with high influence have impact on playtime and social play of the other players, and that players with stronger social relationships have a higher in-game performance and tend to play longer and more often (Pirker et al. 2018; Canossa et al. 2019). Networks in gaming can also be studied through implicit connections (van de Bovenkamp et al. 2014).

Massive multiplayer online games have also been frequently studied, including an analysis of social groups in such a game (Schiller et al. 2018), and using a modelling approach through a graph model that captures social connections and can be used for matchmaking in the game (Jia et al. 2015). A summary of the state-of-the-art in network analysis of massive multiplayer online games pointed out that some of the biggest challenges in analysis of such networks are the network dynamics, massive datasets and real life connections, which are vital for game development (Schlauch and Zweig 2015).

Social network games are types of games that are hosted on a digital social network and integrate the social features of that network. This includes social features, such as gifting of virtual goods, sharing achievements and posting to friends' walls. Research on these games shows that adding these social features in the games increases monetization, engagement, longevity in the game and usage (Alsén et al. 2016; Drachen et al. 2018; Klauser et al. 2013). It has also been shown that, while the majority of players may not form long-term relationships, a small fraction of players connect very closely (Gu and Jia 2018).

Churn prediction in online and mobile gaming

In the literature, several attempts have been made for predicting player churn in various types of online and mobile games. The methodology typically depends on engineering representative and useful behavioral features from the games' player logs, and using them as input in supervised machine learning models (Borbora and Srivastava 2012; Kim et al. 2017; Rothmeier et al. 2020; Hadiji et al. 2014; Periañez et al. 2016, 2016; Loria and Marconi 2021). The most commonly used features describe player behavior and performance, but in addition, social interaction features (Fu et al. 2017), cognitive psychological features (Jeon et al. 2017), and influence features (Fu et al. 2017) have shown to improve the performance of the models. Unsupervised approaches have also been developed, including segmentation of players (Fu et al. 2017), difference in social neighbour influence in different user groups (Liu et al. 2019), and distinct behavioral profiles associated with churners and active players (Borbora and Srivastava 2012). The sole research we found that used social networks

to predict churn, included information from ego-nets showing that loners are much more likely to churn than the socializers, and that the propensity to churn increases with decreased socialization (Borbora et al. 2019).

More advanced methods, such as deep learning, have been studied to a lesser extent in this context. On the one hand, sentiment analysis of players' comments in games using deep learning and word embedding models showed a strong association with churn (Kilimci et al. 2020), and on the other hand, graph embeddings trained on bipartite networks linking players to games in a gaming platform were shown to capture both contextual information and relationship dynamics to predict churn of players (Liu et al. 2018, 2020).

In free-to-play games like the one we study in this paper, churn prediction is challenging because of the players' bursty behavior and the high churn rates among new players (Petersen et al. 2017). In free-to-play mobile games, retention features have been linked to player's motivation, social feature engagement, activity, different feelings and learning (Sorvari 2018; Kitola 2019; Rothmeier et al. 2020; Sifa et al. 2015). A broadly applicable churn prediction model used playtime, session length and session intervals variables, that is, variables that are not game specific, to predict churn successfully (Hadiji et al. 2014). In a survey based study on the intention to play in socially interactive games on mobile devices, results showed that both network externalities and individual gratifications significantly influenced players' intention to play (Wei and Lu 2014). Other researchers have focused on churn prediction for high value players (Runge et al. 2014), and player lifetime value prediction, where the most important features were purchase related (Vilar De Carvalho Santos 2020).

From this overview of the literature, we see that networks of players in online and mobile games in relation to churn are understudied, although it is known that social influence is an important factor. This is especially the case for free-to-play games.

Methodology

In this research, we study player churn in a mobile free-to-play game where players play against each other in a one-versus-one match. We take into account information from the players' gaming networks to assert whether it can be used to predict churn. Our methodology is divided into a few steps. First, we describe our dataset and how we define churn in this context. Then, we explain how networks emerge in our application and how network information is extracted. Finally, we describe our experimental setup.

Player data

For this research we used a dataset of players of a mobile game. We created a sample of players for our analysis as follows: Looking at 10 consecutive days in May 2020 we took a random sample of 1000 players that installed the game on each of these 10 days, resulting in a sample of 10000 players. Furthermore, we included the players' gaming behavior, i.e. all in-game activity, for one year after installing the game. Not all players start playing the game on the day of installation. Therefore, we looked in the data when the players became active in the game and considered that their first active day.

The data set we used in our analysis contains multiple features generated from the in-game activity data. These features are categorized as shown in Table 1. These features were engineered from the raw activity data, based on insights from the company's

Table 1 Description of feature categories among the benchmark features

Type	Description
Per Day	Features relating to how often per day do certain events, e.g. how many matches a player plays, wins and loses, occur on average, median, minimum and maximum for each player.
Per Match	Features relating to how often per match do certain match related events occur on average, median, minimum and maximum for each player. These events include how often a player wins or loses a point and misses a stroke.
Over all	Features relating to how often match related events occur for each player during the 14 day observation period.
Match Types	Features relating to the how many Bet-, Friend- and Tournament matches a player plays.
Playing period	Features relating to how many days players play the game during the 14 day observation period and the number of days between their first and last played days within the observation period, i.e. the length of their playing period within the observation period.
Streaks	Features relating to how many win and loses streaks a player has, the number of matches within their streak and how long a streak is. The number of matches within a winning streak is defined as the number matches won in a row and a losing streak is defined as the number matches lost in a row.
Currency	Features relating to the hard and soft currency of players.
Purchases	Features relating to the purchases made by players, with both hard and soft currency.
Friends	Features relating to matches played against friends.
Player	Features relating to the statistics of a player
Win-Loss Types	Features relating to the different types of wins/losses, e.g. a player automatically wins a match if their opponent abandons the match.

experts. The features in these categories are our *benchmark features* as they represent information used for churn prediction without any network information.

Defining churn

Churn prediction in free-to-play mobile games is challenging for two main reasons. Firstly, there is no subscription model. This means that players can stop playing the game at any time without giving notice. As a result, there is no clear or formal churn event. Instead, we use a data-driven way to determine when a player has churned. Secondly, player behavior is bursty, meaning that new players tend to play very intensively for a short period of time and then they quit –or churn. Therefore, the data distribution is very skewed.

This was clear in our data. Many players only played the game for a few days and some even played the game for a single day only. This made it apparent that the first week or two are the most crucial time when trying to predict and even prevent churn. Understanding the difference between the behavior of players labelled as churners and non-churners during those first days was therefore crucial for our purposes.

We adopted the following method to define churn. We considered the first 14 active days of a player as their observation period. We considered the subsequent 14 days as the prediction period. Furthermore, because of the bursty behavior, and to ensure we had enough activity data, we required the players to have played for at least a fixed number of days in their observation period, i.e. their first 14 days of game play, to be included in our analyses. We considered two cases for the required activity in the observation period, namely at least 3 and at least 5 days. If the player played no games in the prediction period, they were labelled as churners and otherwise, they were

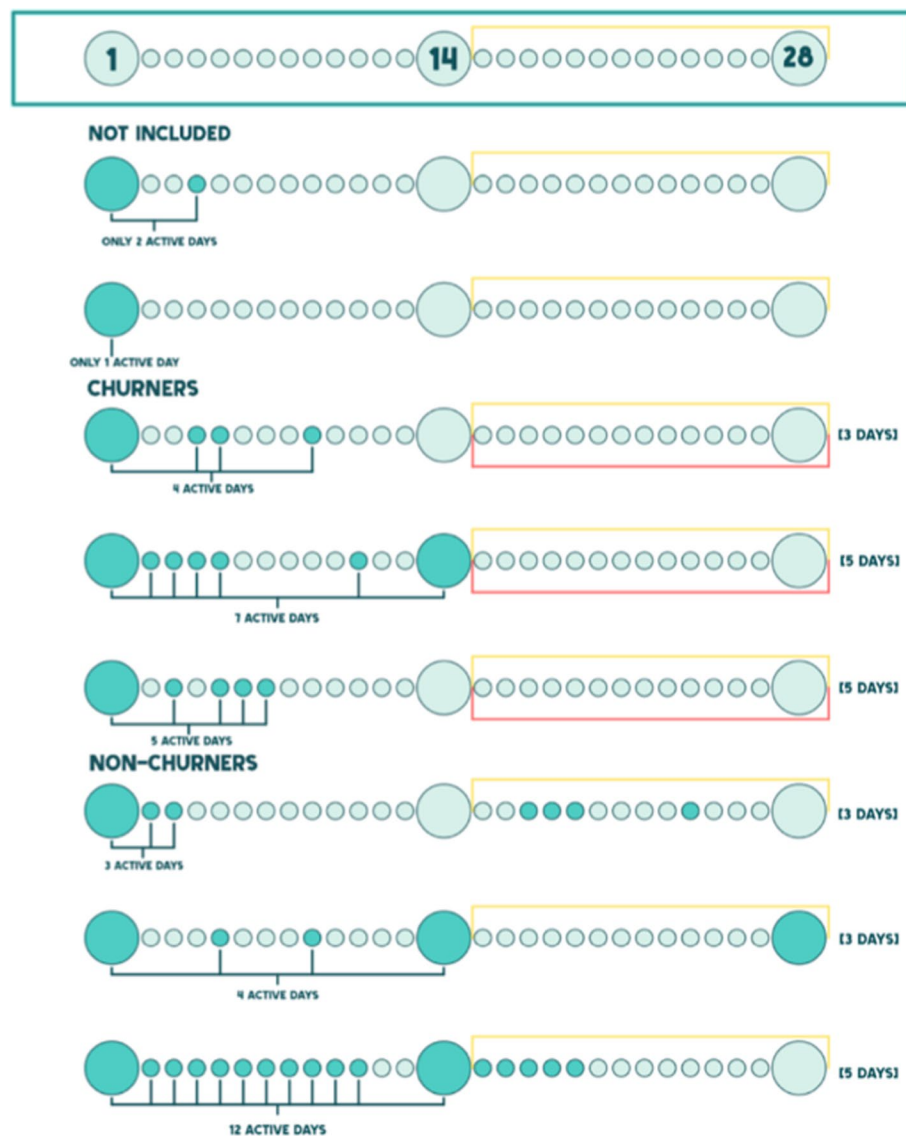


Fig. 1 Examples of excluded players, players that are churners and non-churners depending on their game playing in their observation and prediction periods

labelled as non-churners. This is shown for a few examples in Fig. 1, based on the 3 day activity requirement.

As a result of these definitions we had two datasets, namely the *3-day* dataset (based on the 3 day activity requirement) with around 3800 players and a churn rate of 50.04% and the *5-day* dataset (based on the 5 day activity requirement) with around 3200 players (in the dataset with all features), and a churn rate of 45.87%.

We note that a player that has been active for a minimum of 5 days has also been active for a minimum of 3 days, and therefore the 5-day data is a subset of the 3-day data. The minimum number of active days in the observation period was selected due to observations made during the analysis of the data. The data analysis produced after the players had been labelled with their appropriate churn label clearly showed a

correlation between a player's number of active days and their churn label. The more active days a player has in their observation period, the less likely they are to churn. Due to this, despite the 5-day data being a subset of the 3-day data, the data sets contained a different proportion of churners. The 3-day data had more players and a higher proportion of churners while the 5-day data contained fewer players, a lower ratio of churners as well as richer information of player activity. Richer information refers to the fact that the players in the 5-day data play more and therefore there exist more information about their behavior.

Networks and network features

Network features in online and mobile games can be extremely important for competitive games, as detailed above. Having friends playing the same game can make it more enjoyable but also challenges the players.

In the mobile game we study here, the players have a choice to play with a similar player or a person from their list of friends. We considered two types of matches: friend and similarity, that result in two types of networks: the friend network and the similarity network. The friend network is an explicit social network, whereas the similarity network is an implicit social network, where nodes are connected based on similar properties, that is, skill level and geographic location. They are constructed in the same way: When two players play a match, an edge is created between them. The nodes are the two players.

We used the 10 cohorts of players to create the two networks. As seen in Fig. 2, these networks include, in addition to the players we are studying, other players that did not install the game during the 10 days mentioned. In the figure, the players under consideration are shown in dark gray. Their neighbors, i.e. the players they were matched with, are shown in light gray and white. The players that fulfill our definition of being churners are gray with a red circle, those that fulfill the definition of being non-churners are gray and those which we do not know the status of are shown in white. Their label is unknown. Figure 3 shows the two networks represented as a two layer network, with dashed lines connecting the same player the two layers. The top layer shows the friend network and the bottom layers shows the similarity network.

From the networks we extracted features on the nodes' connectivity and created two types of features, friend and similarity. This enabled us to do further analysis into how friends affect game play since many players want to compete against other players which have a personal connection to them. The network features are inspired by

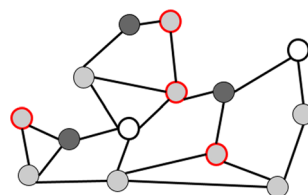


Fig. 2 One of the networks and different types of nodes. The dark nodes indicate the players being studied. The light gray nodes are other players. Known churners are circled in red. The white nodes represent players with an unknown label

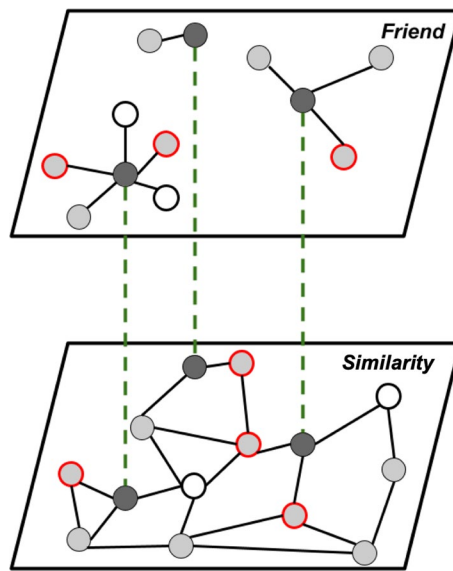


Fig. 3 The two layer network, with a similarity layer (bottom) and a friend layer (top). The dark gray nodes represent the players studied in this research. They exist in both layers

Table 2 The network features that are divided into two types: Friend and Similarity, originating from the friend and similarity networks, respectively

Type	Description
Friend	friend_churner_match_percentage friend_non_churner_match_percentage
Similarity	random_churner_match_percentage random_non_churner_match_percentage

Getoor (2005); Óskarsdóttir et al. (2017), and are so called link-based features. The two features of each type calculate the percentage of matches a player played with churners and non-churners, respectively. They are listed in Table 2. We refer to these types of feature as *friend* and *similarity* features, respectively.

Experimental setup

Now we explain the setup up of the experiments we conducted to determine the predictive power of the friend and similarity network features.

Using the two datasets, 3-day and 5-day, we trained and tested a range of machine learning algorithms to predict churn and evaluated their performance. Each algorithm was trained on both datasets, using the following combinations of feature types.

- (1) Benchmark – as listed in Table 1
- (2) Benchmark + Friends
- (3) Benchmark + Similarity
- (4) Benchmark + Friends + Similarity

The modelling process included the following algorithms: Decision Trees, Random Forests, Stochastic Gradient Descent (SGD), K-Nearest Neighbors (KNN) and Extreme Gradient Boosting (XGBoost). They were chosen because they provide a good range of interpretable and black-box models, and hence a variation in terms of supervised machine learning approaches that are commonly used in industry (Verbeke et al. 2012; Baesens 2014).

When training our models and tuning hyperparameters, we followed the extensive cross validation scheme proposed by Henckaerts et al. (2021) as depicted in Fig. 4 (Henckaerts et al. 2021). Hereby, we split the full dataset D into five disjoint subsets D_1, \dots, D_5 , and performed a five-fold cross-validation where each of the five subsets served as a test set in one fold. In each training fold, we further performed a four-fold cross validation to tune the hyperparameters. The hyper parameters that gave the highest performance were then used to train the model in the original fold, and the model validated on the test set of each respective fold. This cross-validation scheme served two purposes. First, it allowed us to tune the hyper-parameters with a four-fold cross-validation approach. Second, as the models are trained several times, we could evaluate their predictive performance on multiple datasets, instead of on a single test set. Thereby, we obtained multiple performance measures per model and thus a more accurate performance assessment. This enabled sensitivity checks to assess the stability of different algorithms. It also reduced the risk of bias in the results since we do not have to select a specific test set. Table 6 in Appendix shows the tuning grid used for the hyperparameters.

The models' performance was evaluated using precision, recall, F_1 score and the area under the receiver operating characteristic curve (AUC) (Géron 2019). Each algorithm was trained using the extensive cross-validation scheme for each feature combination and each dataset. We report the mean value from the results of the cross-validation.

Results

Impact of network features

First we reflect on the results of including network features in the churn prediction models. Tables 3 and 4 show the performance measured in precision, recall, F_1 score and AUC for the five supervised machine learning algorithms trained on the 3-day and 5-day

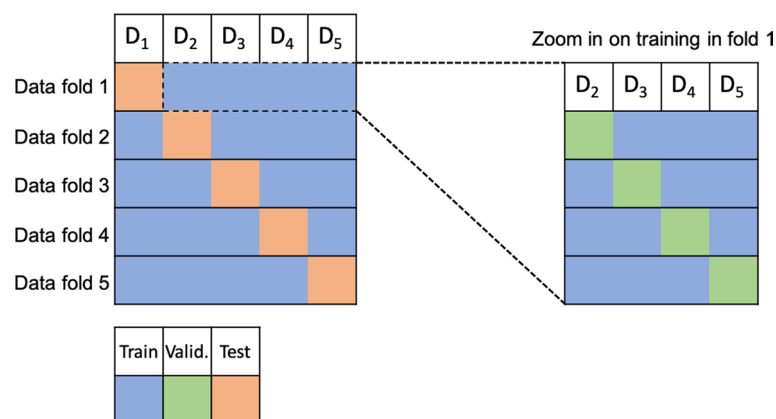


Fig. 4 The extensive cross-validation scheme. Figure adapted from Henckaerts et al. (2021)

Table 3 Performance, mean and (standard deviation), of churn prediction models for the four feature groups using the 3-day dataset

Classifier	Benchmark				Benchmark + Friends			
	Precision	Recall	F ₁	AUC	Precision	Recall	F ₁	AUC
DecisionTree	0.688 (0.017)	0.797 (0.033)	0.738 (0.018)	0.774 (0.018)	0.692 (0.019)	0.880 (0.021)	0.774 (0.008)	0.743 (0.014)
KNN	0.538 (0.036)	0.704 (0.185)	0.599 (0.073)	0.545 (0.058)	0.537 (0.036)	0.704 (0.185)	0.599 (0.073)	0.545 (0.058)
RandomFor- est	0.712 (0.015)	0.840 (0.015)	0.771 (0.014)	0.812 (0.014)	0.714 (0.015)	0.834 (0.016)	0.769 (0.012)	0.812 (0.012)
XGBoost	0.705 (0.008)	0.807 (0.022)	0.752 (0.011)	0.797 (0.016)	0.702 (0.017)	0.799 (0.055)	0.747 (0.029)	0.796 (0.019)
SGD	0.622 (0.068)	0.578 (0.288)	0.546 (0.164)	0.592 (0.042)	0.614 (0.069)	0.567 (0.332)	0.516 (0.218)	0.585 (0.047)
Classifier	Benchmark + Similarity				Benchmark + Friends + Similarity			
	Precision	Recall	F ₁	AUC	Precision	Recall	F ₁	AUC
DecisionTree	0.692 (0.019)	0.880 (0.021)	0.774 (0.008)	0.743 (0.014)	0.643 (0.042)	0.828 (0.084)	0.719 (0.017)	0.721 (0.034)
KNN	0.538 (0.036)	0.704 (0.185)	0.599 (0.073)	0.545 (0.058)	0.537 (0.036)	0.704 (0.18)	0.599 (0.073)	0.546 (0.058)
RandomFor- est	0.710 (0.013)	0.839 (0.019)	0.769 (0.013)	0.810 (0.015)	0.711 (0.016)	0.841 (0.019)	0.770 (0.016)	0.811 (0.014)
XGBoost	0.702 (0.015)	0.806 (0.038)	0.750 (0.025)	0.794 (0.025)	0.697 (0.013)	0.815 (0.046)	0.751 (0.025)	0.797 (0.019)
SGD	0.617 (0.065)	0.649 (0.219)	0.604 (0.102)	0.605 (0.043)	0.597 (0.073)	0.617 (0.342)	0.534 (0.194)	0.568 (0.040)

Table 4 Performance, mean and (standard deviation), of churn prediction models for the four feature groups using the 5-day dataset

Classifier	Benchmark				Benchmark + Friends			
	Precision	Recall	F ₁	AUC	Precision	Recall	F ₁	AUC
Decision- Tree	0.525(0.028)	0.493(0.074)	0.506(0.048)	0.715(0.030)	0.607(0.024)	0.652(0.094)	0.625(0.045)	0.750(0.018)
KNN	0.325 (0.043)	0.460 (0.194)	0.363(0.098)	0.500 (0.056)	0.339(0.049)	0.488(0.219)	0.375(0.090)	0.509 (0.061)
Random- Forest	0.620 (0.049)	0.495(0.131)	0.535 (0.082)	0.811(0.020)	0.631 (0.035)	0.546 (0.091)	0.579 (0.053)	0.810 (0.022)
XGBoost	0.570(0.023)	0.498(0.198)	0.499(0.158)	0.781(0.018)	0.584(0.054)	0.536(0.191)	0.526(0.126)	0.790(0.017)
SGD	0.455(0.144)	0.273(0.318)	0.263(0.181)	0.555(0.059)	0.571(0.219)	0.141 (0.222)	0.156 (0.199)	0.541 (0.061)
Classifier	Benchmark + Similarity				Benchmark + Friends + Similarity			
	Precision	Recall	F ₁	AUC	Precision	Recall	F ₁	AUC
Decision- Tree	0.695(0.025)	0.792 (0.071)	0.738(0.028)	0.783(0.015)	0.699(0.040)	0.716(0.152)	0.696(0.058)	0.779(0.025)
KNN	0.683(0.071)	0.573(0.082)	0.615(0.028)	0.708(0.012)	0.676 (0.067)	0.580(0.076)	0.617 (0.026)	0.704 (0.014)
Random- Forest	0.704 (0.023)	0.779 (0.020)	0.739(0.018)	0.815(0.022)	0.703 (0.029)	0.782 (0.018)	0.740 (0.023)	0.815 (0.023)
XGBoost	0.682 (0.019)	0.795(0.039)	0.733 (0.012)	0.802(0.016)	0.686(0.026)	0.786(0.017)	0.732(0.016)	0.801(0.021)
SGD	0.626 (0.086)	0.751(0.127)	0.667(0.010)	0.618(0.098)	0.602 (0.140)	0.671(0.220)	0.604(0.097)	0.579(0.128)

datasets, respectively. In the upper half is the performance of the benchmark features, which only describe the players' behavior, on the left, and on the right is the performance of models with the benchmark features and the features from the friend network.

In the lower half is the performance of the models with benchmark features and features from the similarity network on the left and on the right is the performance of the models with all features, i.e. benchmark features, the features from the friend network and the features from the similarity network.

In terms of the benchmark features, random forests and XGBoost performed the best for both datasets, followed by decision trees. For the 3-day dataset, random forest achieved 71% precision with 84% recall and XGBoost achieved 71% precision and 81% recall, meanwhile for the 5-day dataset, random forest reached 62% precision with 50% recall and XGBoost achieved 57% precision with 50% recall.

When adding the features from the friend network, random forest and decision trees performed the best for both datasets. For the 3-day, random forest reached 71% precision with 83% recall whereas decision trees achieved 69% precision and 88% recall. Meanwhile for the 5-day, random forest resulted in 63% precision with 55% recall and decision trees achieved 61% precision with 65% recall.

Similarly, when combining the features from the similarity network with the benchmark features, XGBoost, random forest and decision trees performed the best for both datasets. For the 3-day, XGBoost reached 0.70% precision and 81% recall, random forest achieved 71% precision with 84% recall and decision trees reached 69% precision and 88% recall, whereas for the 5-day, XGBoost reached 68% recall and 80% recall, random forest gave 70% precision with 78% recall and decision trees resulted in 70% precision with 79% recall.

Finally, when including all features, random forest and XGBoost performed the best for both datasets. For the 3-day, random forest reached 71% precision with 84% recall and XGBoost gave 70% precision and 82% recall, meanwhile for the 5-day, random forest resulted in 70% precision with 78% recall and XGBoost achieved 69% precision with 79% recall.

When comparing the predictive performance of the feature groups, we see that when adding the friend features there is a slight increase in the models' performance. When we add the features from the similarity network instead, we see an even greater boost in the models' performance, especially in the 5-day dataset. This indicates that the similarity network is more predictive than the friend network. Finally, when all features are included, we obtain the highest performance, especially in terms of recall. This means that the models are able to predict the churn label more accurately which indicates that the network features are adding alternative and valuable information about churners.

Interestingly, the performance on the 3-day dataset is slightly higher than for the 5-day dataset. Notably this is due to the higher recall in the 3-day dataset, where the models are better at detecting churners.

We have a closer look at the models by inspecting the feature importances of the random forest models when using all the features, see Figs. 5 and 6. For both datasets, we see that features representing wins and losses are important, as well as streaks (number of games won in a sequence) and in-game activity, i.e. strokes. In both plots, both friend and similarity network features appear among the top 30 most important features, with the friend network features appearing above the similarity network features. Interestingly, the friend network feature is the percentage of games played against churners, whereas the similarity network feature represents the percentage of games played

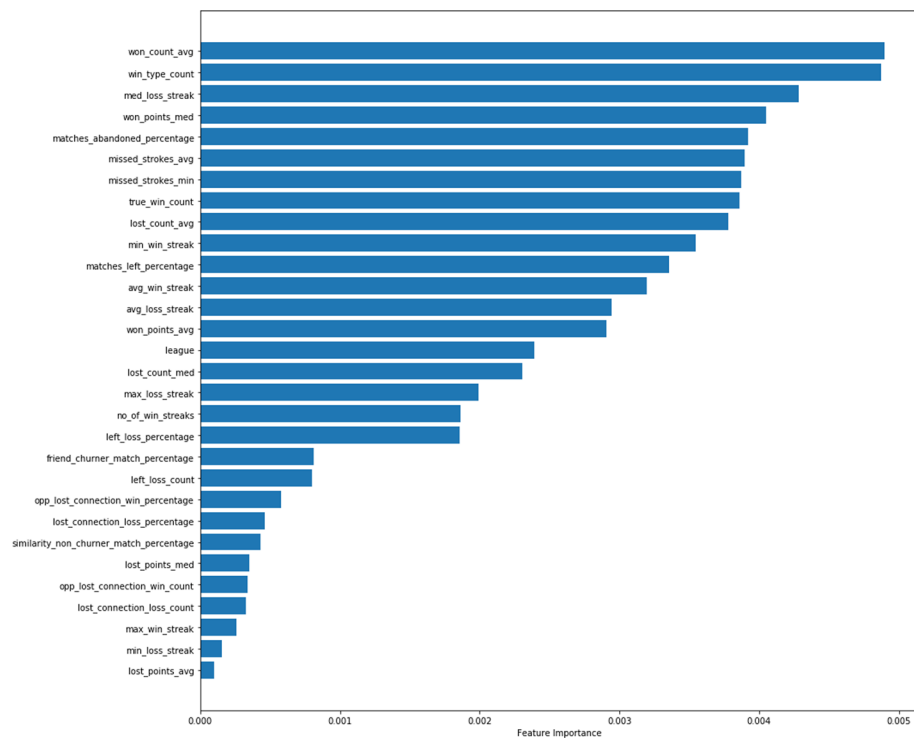


Fig. 5 Feature importance in the random forest model using the 3-day dataset and all the features

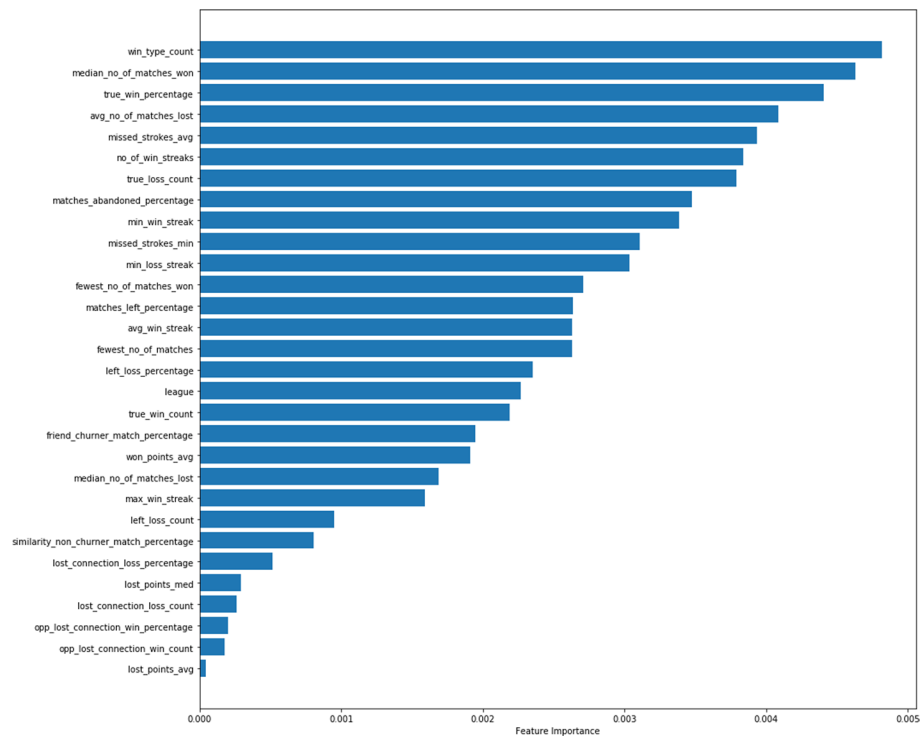


Fig. 6 Feature importance in the random forest model using the 5-day dataset and all the features

against non-churners, in both cases. This might indicate that if a player's friends churn, then the player is more likely to churn.

First day features

The data analysis and modelling indicated that the less a player played the game the more likely they were to be churners. After the modelling process was completed we were interested in examining whether it would be possible to predict a players churn label after their first day of game play. The same algorithms were applied to the features for the players first day of game play. The models showed that this is in fact possible. The features for the first day were different from the features used for modelling of the main data. The reason for why they're different is because many of the original features were e.g. based on the number of times an event occurred during the observation period. Knowing e.g. how many matches a player plays on average per day when we're only observing one day tells us nothing. Some feature were therefore represented with a Boolean value instead of a quantitative value, e.g. currency, because we decided it was more important to whether they spent currency rather than how much. Furthermore, we focus here on the 3-day dataset only as it showed better performance above.

Table 5 shows the results. Although the model performance is slightly worse than in the previous section, it is remarkably close. Especially decision trees which are now on par with the more complex models random forest and XGBoost. For example decision trees reach a recall of 78% when random forest and XGBoost achieve 75% and 78%, respectively. This means that churn can be predicted quite well using only activity of the first day of game play. This further establishes the players' bursty behavior in the game.

Discussion and conclusion

In this paper, we presented an approach for extracting network information from gaming networks of players in a free-to-play mobile game and demonstrated that such information is important for predicting churn of players, as well as that first-day behavior is highly predictive of churn. As such, our results have managerial implications, because of the novel insights they bring to customer retention in the highly competitive field of free-to-play mobile games, as well as having scientific implications, especially regarding the study of influence in networks, players' bursty behavior and a novel addition to body of literature on churn prediction with social networks.

Table 5 Performance, mean and (standard deviation), of churn prediction models for the using first day features using the 3-day dataset

Classifier	3 day			
	Precision	Recall	F ₁	AUC
DecisionTree	0.595 (0.051)	0.781 (0.130)	0.666 (0.033)	0.639 (0.049)
KNN	0.501 (0.007)	0.639 (0.117)	0.557 (0.044)	0.499 (0.036)
RandomForest	0.633 (0.015)	0.751 (0.045)	0.686 (0.025)	0.697 (0.017)
XGBoost	0.619 (0.010)	0.778 (0.075)	0.688 (0.030)	0.695 (0.019)
SGD	0.636 (0.011)	0.603 (0.096)	0.616 (0.058)	0.653 (0.030)

On the one hand, our findings indicate that churn indeed occurs quickly after people start playing the game, as shown by the high performance of the 3-day models and the models with first day features. This means that in order to run a successful retention campaign, the company must act quickly, ideally within a day after the players become active in playing the game. There are several ways to approach the retention of players. One option would be to design campaigns to increase player's engagement for the set of players that are predicted as churners. This could be achieved for example by giving the players free coins in a temporal way in order to allow them to play more matches or tournaments, decreasing the time to open chests that are subject to a time gate, or increasing the quality of cards received by those players as rewards. Another option could be to have dynamic offers for the set of players that are predicted as churners. Better offers could be made to the players predicted as churners, i.e. by decreasing the price of items during a bounded period of time. There might also be benefit in using methods that are well adapted at detecting churn early as presented in Óskarsdóttir et al. (2018).

On the other hand, our work has scientific contributions. Firstly, the literature has already demonstrated the importance of social networks in various types of online games, but free-to-play mobile games have been studied to a lesser extent. Here we take the first steps in studying the importance of social networks in mobile games. Although we are not able to study the networks in detail, the fact that players are choosing to play with their friends when given the option, is already an implication that the social networks are important in these types of games case too. Secondly, comparing the performance of models trained on the three datasets (3-day, 5-day and first-day) allows us to quantify the burstiness of the players. Indeed, the fact that the 5-day model performed worse indicates that by requiring 5 days of game play to be included in the dataset for analysis, the players that churned the earliest were already excluded and the model struggled more in distinguishing churners from non-churners. Players that played for at least 5 days were thus less likely to churn, at least in the prediction period. In contrast, when requiring less active days, there was a clearer distinction between the two groups. Then, to the best of our knowledge, we are the first to enrich churn prediction models in free-to-play games with network information. In line with our previous work on churn prediction in the telecommunication industry, we see a significant increase in the models' performance, indicating that the network features carry information for the churn prediction task that is different and complementary to the other features (Óskarsdóttir et al. 2017, 2016). This also in line with other work (Verbeke et al. 2014). Finally, it is worth noting that the performance of the simpler and interpretable models, in particular decision trees, was only slightly worse than that of the more complex, black-box models, i.e., random forests and XGBoost. This is probably due to the simplicity of the dataset, which had only a couple of thousand observations, which means that the simpler models were able to learn the patterns in the data, which were perhaps not too intricate.

The goal of this paper was to investigate the importance of social network features when predicting churn in free-to-play mobile games. For our research, we used data describing the gaming and in-app behavior of thousands of players, in addition to information about who they played against in this game with one-vs-one matches. We used this information to build two distinct networks, one with friend matches and the other

were players were matched based on similarity and extracted network features which we combined with the more traditional features in churn prediction models. In addition, we defined churn in two ways, requiring at least 3, respectively 5 days, of game play during the first 14 days of game play for players to be included in the dataset. A player was labelled as a churning player if they did not play in the 14 day period following the first 14 days. The network features greatly increased the predictive performance of the models, indicating that they carry alternative information about intention to churn. In this regard, the features from the similarity network increased the performance more than the features from the friend network. The models trained on the 3-day dataset had higher predictive performance than those trained on the 5-day dataset, and overall, random forests and XGBoost performed the best and KNN the worst. Because of the bursty game behavior of new players, we then focused on their game play in the first day and used that behavior to predict churn. The results were almost as good as when considering the first 14 days, which further strengthens the bursty gaming behavior.

Our work is not without limitations and offers many possibilities for future research. In this work, we used only two features from each network which is rather limited. Previous research used, with good results, other link-based features in addition to centrality features and propagation scores (Óskarsdóttir et al. 2017; Calzada-Infante et al. 2020; Backiel et al. 2015). In line with those works, we plan to include more features in the future. An important limitation of our work is that we chose to focus on new players and their activity in the first days of game play. As a result, social features (explicit connections) were rather sparse for the players and we had more information about the implicit connections. Therefore, we are now also studying long-term players which have played more games both with friends and similar players. Thereby, we get denser networks with higher connectivity, especially in the friend network, as more players connect the game to their social networks, so we can expect to have more social links between players. Long-term and paying customers have a higher lifetime value, making this cohort more relevant for the company. Looking at long-term customers, which have longer history and more activity data in addition to information about purchases, will give other insights on churning players than what we learned in this paper. Therefore, in addition to predict churn, we plan to also predict future purchases. In this research, we considered two distinct networks, the friend and the similarity network. But as the players can exist in both networks, an interesting approach would be to look at it as a multilayer network and study influence and behavior using multilayer network theory, an interesting and quickly advancing field of network science (Kivelä et al. 2014). We believe we could make substantial contribution in this field. Finally, we used a straight forward approach in our methodology, by extracting features from the networks and adding to the features of the players. In our future work, we would like to explore more sophisticated methods, such as graph neural networks (Wu et al. 2020). This is an ideal setting because of how dynamic the networks are and how rich they are in both node and edge features. This approach could be used not only to predict churn but also to predict the winner of a match –an example of a heterophilic network (Zhu et al. 2020)– or to match players, i.e. an application of edge prediction.

Appendix A: hyperparameters

Table 6 The grids used for the tuning of hyperparameters.

Classifier	Hyper parameters	Candidate setting
DecisionTree	Alpha	0.1, 0.01, 0.001, 0.0001
	Max tree depth	4, 8, 10, 20, 30
	Criterion	gini, entropy
KNN	Weights	uniform, distance
	# Neighbors	3, 5, 11, 14, 16, 18, 20
	Metric	euclidean, manhattan
	Algorithm	auto, ball_tree, kd_tree, brute
RandomForest	# Estimators	300, 500, 700
	Max tree depth	4, 8, 10, 20
	Criterion	gini, entropy
XGBoost	Minimum child weight	1, 5, 10
	Gamma	1, 0.01, 0.1, 0.5
	Subsample ratio of the training instances	0.6, 1.0
	Subsample ratio of columns	0.6, 1.0
	Max tree depth	3, 3, 5
SGD	Loss	hinge, log, squared_hinge, modified_huber
	Alpha	0.001, 0.01, 0.1
	Penalty	l2, l1, none

Acknowledgements

The authors are grateful to Eduardo Yada, Andreas Tobelem, and Felipe Freitas who participated in the project and gave valuable feedback.

Author contributions

MO: Designed the study, supervised the research and wrote the manuscript. KEG and RS: Performed the analysis, interpreted the results, created figures and wrote the manuscript. DA and CS: Aided in study design, contributed research expertise and edited the manuscript. All authors read and approved the final manuscript.

Funding

The authors received no specific funding for this research.

Availability of data and materials

The data that support the findings of this study are available from Wildlife Studios but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Wildlife Studios.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Received: 31 March 2022 Accepted: 28 November 2022

Published online: 15 December 2022

References

- Alsén A, Runge J, Drachen A, Klapper D. (2016) Play with me? understanding and measuring the social aspect of casual gaming. In: twelfth artificial intelligence and interactive digital entertainment conference
- Backiel A, Verbinen Y, Baesens B, Claeskens G. (2015) Combining local and social network classifiers to improve churn prediction. In: 2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp. 651–658. IEEE
- Baesens B (2014) Analytics in a big data world: the essential guide to data science and its applications. John Wiley, USA
- Borbora Z.H, Srivastava J. (2012) User behavior modelling approach for churn prediction in online games. In: 2012 international conference on privacy, security, risk and trust and 2012 international conference on social computing, pp 51–60. IEEE
- Borbora Z, Chandra A, Kumaraguru P, Srivastava J. (2019) On churn and social contagion. In: Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining, pp 833–841
- van de Bovenkamp R, Shen S, Jia AL, Kuipers F et al (2014) Analyzing implicit social networks in multiplayer online games. *IEEE Internet Comput* 18(3):36–44
- Calzada-Infante L, Óskarsdóttir M, Baesens B (2020) Evaluation of customer behavior with temporal centrality metrics for churn prediction of prepaid contracts. *Expert Syst Appl* 160:113553
- Canossa A, Azadvar A, Harteveld C, Drachen A, Deterding S (2019) Influencers in multiplayer online shooters: Evidence of social contagion in playtime and social play. In: proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–12
- De Caigny A, Coussemont K, De Bock KW (2018) A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *Eur J Oper Res* 269(2):760–772
- Drachen A, Pastor M, Liu A, Fontaine DJ, Chang Y, Runge J, Sifa R, Klabjan D (2018) To be or not to be... social: Incorporating simple social features in mobile game customer lifetime value predictions. In: proceedings of the australasian computer science week multiconference, pp. 1–10
- Fu X, Chen X, Shi Y-T, Bose I, Cai S (2017) User segmentation for retention management in online social games. *Decis Support Syst* 101:51–68
- Getoor L. (2005) Link-based classification. In: advanced methods for knowledge discovery from complex data, pp 189–207. Springer, USA
- Gu L., Jia A.L. (2018) Player activity and popularity in online social games and their implications for player retention. In: 2018 16th annual workshop on network and systems support for games (NetGames), pp 1–6. IEEE
- Géron A. (2019) Hands-on Machine Learning with scikit-learn, keras, and tensorflow: concepts, tools, and techniques to build intelligent systems. "O'Reilly Media, Inc.", ???
- Hadji F, Sifa R, Drachen A, Thureau C, Kersting K, Bauckhage C. (2014) Predicting player churn in the wild. In: 2014 IEEE conference on computational intelligence and games pp 1–8. IEEE
- Henckaerts R, Côté M-P, Antonio K, Verbelen R (2021) Boosting insights in insurance tariff plans with tree-based machine learning methods. *North Am Actuarial J* 25(2):255–285
- Jeon J, Yoon D, Yang S, Kim K. (2017) Extracting gamers' cognitive psychological features and improving performance of churn prediction from mobile games. In: 2017 IEEE conference on computational intelligence and games (CIG), pp 150–153. IEEE
- Jia AL, Shen S, Bovenkamp RVD, Iosup A, Kuipers F, Epema DH (2015) Socializing by gaming: Revealing social relationships in multiplayer online games. *ACM Trans Knowl Discov Data (TKDD)* 10(2):1–29
- Karsai M, Kivela M, Pan RK, Kaski K, Kertész J, Barabási A-L, Saramäki J (2011) Small but slow world: how network topology and burstiness slow down spreading. *Phys Rev* 83(2):025102
- Kilimci Z.H, Yörük H, Akyokus S. (2020) Sentiment analysis based churn prediction in mobile games using word embedding models and deep learning algorithms. In: 2020 international conference on innovations in intelligent systems and applications (INISTA), pp 1–7. IEEE
- Kim S, Choi D, Lee E, Rhee W (2017) Churn prediction of mobile and online casual games using play log data. *PLoS one* 12(7):0180735
- Kitola M. (2019) Impact of social features on player retention in free-to-play mobile games
- Kivela M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA (2014) Multilayer networks. *J Complex Net* 2(3):203–271
- Klauser M, Nacke LE, Prescod P. (2013) Social player analytics in a facebook health game. In: proceedings of the CHI 2013 workshop on designing and evaluating sociability in online video games pp 39–44
- Liu D-R, Liao H-Y, Chen K-Y, Chiu Y-L (2019) Churn prediction and social neighbour influences for different types of user groups in virtual worlds. *Expert Syst* 36(3):12384
- Liu X, Xie M, Wen X, Chen R, Ge Y, Duffield N, Wang N (2020) Micro-and macro-level churn analysis of large-scale mobile games. *Knowl Inf Syst* 62(4):1465–1496
- Liu X, Xie M, Wen X, Chen R, Ge Y, Duffield N, Wang N. (2018) A semi-supervised and inductive embedding model for churn prediction of large-scale mobile games. In: 2018 IEEE international conference on data mining (ICDM), pp 277–286. IEEE
- Loria E, Marconi A (2021) Exploiting limited players' behavioral data to predict churn in gamification. *Electronic Commerce Res Appl* 47:101057
- Loria E, Nacke LE, Pirker J. (2021) The quirks of being a wallflower: Towards defining lurkers and loners in games through a systematic literature review. In: extended abstracts of the 2021 CHI conference on human factors in computing systems pp 1–7
- Óskarsdóttir M, Bravo C, Sarraute C, Vanthienen J, Baesens B (2019) The value of big data for credit scoring: enhancing financial inclusion using mobile phone data and social network analytics. *Appl Soft Comput* 74:26–39
- Óskarsdóttir M, Bravo C, Verbeke W, Sarraute C, Baesens B, Vanthienen J. (2016) A comparative study of social network classifiers for predicting churn in the telecommunication industry. In: 2016 IEEE/ACM International conference on advances in social networks analysis and mining (ASONAM), pp 1151–1158. IEEE

- Óskarsdóttir M, Bravo C, Verbeke W, Sarraute C, Baesens B, Vanthienen J (2017) Social network analytics for churn prediction in telco: model building, evaluation and network architecture. *Expert Syst Appl* 85:204–220
- Óskarsdóttir M, Van Calster T, Baesens B, Lemahieu W, Vanthienen J (2018) Time series for early churn detection: using similarity based classification for dynamic networks. *Expert Syst Appl* 106:55–65
- Óskarsdóttir M, Ahmed W, Antonio K, Baesens B, Dendievel R, Donas T, Reynkens T.(2021) Social network analytics for supervised fraud detection in insurance. *Risk Anal* **n/a**(n/a) <https://doi.org/10.1111/risa.13693>. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/risa.13693>
- Óskarsdóttir M, Sarraute C, Bravo C, Baesens B, Vanthienen J.(2018) Credit scoring for good: Enhancing financial inclusion with smartphone-based microlending. In: international conference on information systems 2018, ICIS 2018, pp 1–9 . Association for information systems
- Periáñez Á, Saas A, Guitart A, Magne C. (2016) Churn prediction in mobile social games: Towards a complete assessment using survival ensembles. In: 2016 IEEE international conference on data science and advanced analytics (DSAA), pp 564–573 . IEEE
- Petersen F.W, Thomsen L.E, Mirza-Babaei P, Drachen A. (2017) Evaluating the onboarding phase of free-to-play mobile games: A mixed-method approach. In: proceedings of the annual symposium on computer-human interaction in play pp 377–388
- Pirker J, Rattinger A, Drachen A, Sifa R (2018) Analyzing player networks in destiny. *Entertain Comput* 25:71–83
- Rattinger A, Wallner G, Drachen A, Pirker J, Sifa R.(2016) Integrating and inspecting combined behavioral profiling and social network models in destiny. In: international conference on entertainment computing pp 77–89 . Springer
- Rothmeier K, Pflanzl N, Hüllmann JA, Preuss M (2020) Prediction of player churn and disengagement based on user activity data of a freemium online strategy game. *IEEE Trans Games* 13(1):78–88
- Runge J, Gao P, Garcin F, Faltings B. (2014) Churn prediction for high-value players in casual social games. In: 2014 IEEE conference on computational intelligence and games, pp 1–8. IEEE
- Schiller MH, Wallner G, Schinnerl C, Calvo AM, Pirker J, Sifa R, Drachen A (2018) Inside the group: investigating social structures in player groups and their influence on activity. *IEEE Trans Games* 11(4):416–425
- Schlauch W.E, Zweig K.A. (2015) Social network analysis and gaming: survey of the current state of art. In: joint international conference on serious games, pp 158–169. Springer
- Seufert EB (2013) Freemium economics: leveraging analytics and user segmentation to drive revenue. Elsevier, USA
- Sifa R, Hadji F, Runge J, Drachen A, Kersting K, Bauckhage C. (2015) Predicting purchase decisions in mobile free-to-play games. In: proceedings of the AAAI conference on artificial intelligence and interactive digital entertainment, vol 11, pp 79–85
- Sorvari T (2018) How the different retention and monetization features affect the user experience in free-to-play mobile games (Unpublished master's thesis). Laurea-ammattikorkeakoulu, Finland
- Verbeke W, Dejaeger K, Martens D, Hur J, Baesens B (2012) New insights into churn prediction in the telecommunication sector: a profit driven data mining approach. *Eur J Oper Res* 218(1):211–229
- Verbeke W, Martens D, Baesens B (2014) Social network analysis for customer churn prediction. *Appl Soft Comput* 14:431–446
- Vilar De Carvalho Santos G. (2020) Feature importance analysis for user lifetime value prediction in games using machine learning: an exploratory approach. PhD thesis, Politecnico di Torino
- Wallner G, Schinnerl C, Schiller MH, Calvo AM, Pirker J, Sifa R, Drachen A (2019) Beyond the individual: understanding social structures of an online player matchmaking website. *Entertain Comput* 30:100284
- Wei P-S, Lu H-P (2014) Why do people play mobile social games? an examination of network externalities and of uses and gratifications. *Internet Res* 24(3):313–331
- Wu Z, Pan S, Chen F, Long G, Zhang C, Philip S.Y. (2020) A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*
- Zeng Y, Sapienza A, Ferrara E. (2019) The influence of social ties on performance in team-based online games. *IEEE Transactions on Games*
- Zhu J, Yan Y, Zhao L, Heimann M, Akoglu L, Koutra D (2020) Beyond homophily in graph neural networks: current limitations and effective designs. *Adv Neural Inf Process Syst* 33:7793–7804

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.