Proceedings Paper:
Drachen, Anders orcid.org/0000-0002-1002-0414, Pastor, Jose Maria Serio, Fontaine, Dylan Jack et al. (6 more authors) (2018) To Be or Not to Be … Social : Incorporating Simple Social Features in Mobile Game Customer Lifetime Value Predictions. In: Proceedings of Australian Computer Science Week, Interactive Entertainment. ACM .

https://doi.org/10.1145/3167918.3167925

To Be or Not to Be... Social: Incorporating Simple Social Features in Mobile Game Customer Lifetime Value Predictions

Anders Drachen
University of York
York, United Kingdom
anders.drachen@york.ac.uk

Mari Pastor
Northwestern University
Evanston, IL, USA
maripastor@u.northwestern.edu

Aron Liu
Northwestern University
Evanston, IL, USA
aronliu2017@u.northwestern.edu

Dylan Jack Fontaine
Northwestern University
Evanston, IL, USA
dylanfontaine2017@u.northwestern.edu

Yuan Chang
Northwestern University
Evanston, IL, USA
yuanchang2018@u.northwestern.edu

Julian Runge
Humboldt University of Berlin
Berlin, Germany
julian.runge@gmail.com

Rafet Sifa
Fraunhofer IAIS
Sankt Augustin, Germany
rafet.sifa@iais.fraunhofer.de

Diego Klabjan
Northwestern University
Evanston, IL, USA
d-klabjan@northwestern.edu

ABSTRACT

Mobile games make up the largest segment of the games industry, in terms of revenue as well as players. Hundreds of thousands of games are available with most being free to download and play. In freemium games, revenue is predominantly generated by users making in-game purchases. As only a small fraction of users make purchases, predicting these users and their Customer Lifetime Value are key challenges in Game Analytics and currently barely explored in academic research. Furthermore, while social factors have been shown to be essential for retention in games in general, the impact on retention and monetization in mobile games is unexplored. In this paper, the problem of defining social features in freemium casual mobile games is addressed through a case study with over 200,000 players. The study evaluates the influence of specific types of social interactions typical of casual mobile games, on predictions of premium users and Customer Lifetime Value by applying classifiers and regression models respectively. Results indicate that social activity does not correlate with the tendency to become a premium user, but that social activity increases over time in a cohort.

ACSW 2018, Brisbane, QLD, Australia

ACM ISBN 978-1-4503-5436-3/18/01.
DOI: 10.1145/3167918.3167925

CCS CONCEPTS

•Computing methodologies → Supervised learning by classification; •Applied computing → Computer games;

KEYWORDS

Game Analytics, Customer Lifetime Value, Behavioral Prediction, Freemium, Computer Games

ACM Reference format:
Anders Drachen, Mari Pastor, Aron Liu, Dylan Jack Fontaine, Yuan Chang, Julian Runge, Rafet Sifa, and Diego Klabjan. 2018. To Be or Not to Be... Social: Incorporating Simple Social Features in Mobile Game Customer Lifetime Value Predictions. In *Proceedings of Australasian Computer Science Week 2018, Brisbane, QLD, Australia, January 29-February 2, 2018 (ACSW 2018),* 10 pages.
DOI: 10.1145/3167918.3167925

1 INTRODUCTION

Mobile games comprise the largest segment of the 100 billion USD games market worldwide [18], having outgrown console and PC games, both in terms of size and growth rate, with hundreds of thousands of mobile games being available. The vast majority of mobile games follow the freemium or Free-to-Play (F2P, FtP) business model, as compared to the retail model utilized by most major commercial titles across not only console and PC platforms but also non-mobile phone handheld platforms. Unlike major commercial titles ("AAA" titles) that carry the largest price tags, F2P games can be played for free. Players, however, have the option to spend money on In-App Purchases (IAPs) to purchase a wide range of incentives, such as virtual currency, temporary boosts, aesthetic augmentations to the game, and elimination of the presence of advertisements (which are a minor but commonly adopted source of revenue for F2P games) [11, 24].

In F2P games, there is typically a disparity amongst players who spend money on IAPs and those who do not [14, 22, 23, 27]. We will refer to the former as **premium players**, and the latter as **non-premium players** following the definition from [23, 26]. Non-premium players comprise the vast majority of F2P players leading to highly imbalanced datasets for prediction in mobile games [14, 23, 27, 32]. Therefore, a key challenge for mobile game developers is not only to reduce player churn and increase retention, but also to convert players from non-premium to premium players. A related goal is to increase the average Customer Lifetime Value (CLTV) [9, 21, 26, 27], since User Acquisition Costs (UAC) have notably increased in recent years for mobile applications. Very recently, UACs have been reported to be continually rising for mobile games, where costs per install regularly rise above 4.50 USD for mobile platforms using iOS, even reaching as high as 5.90 USD in April 2017 [12]. Given the heavy imbalance in player "premiumship", combined with rising UACs in the market, the capability to identify and predict players' behavioral outcomes is an integral factor for a mobile game developer's success. It can inform in-game targeting of advertising, price promotions, game difficulty, and provide critical decision support to marketing/product managers and game designers [14, 23, 26, 27, 32].

Previous work on prediction in mobile games has primarily focused on churn, whereas purchase prediction and CLTV are largely unexplored problems [26]. While social aspects of user behavior in mobile games have been shown to significantly affect retention and revenue, social indicators have not been explored in this context [1]. In most F2P casual mobile games, players have the option to log into the game using social media profiles, and send or receive requests via their contacts to compete against, collaborate with, or simply to receive benefits in the game [15, 16, 31]. Players sending game requests is a form of free advertising for the developers and should be accounted for as an added value to the player [1, 24, 30]. It is therefore of interest to investigate the impact of adding measures of the "social value" of a player to CLTV predictions.

2 CONTRIBUTION

This paper investigates two problems in mobile game analytics in an explorative manner: CLTV prediction modeling and the impact of features describing casual online social interactions in mobile titles on such models. As a case study, we adopt a cohort of players from a casual F2P mobile game published by a leading developer with the features at use being generic to such titles more broadly [23, 27]. The features derive from initial friend requests, forming a connection between players, and subsequent requests or messages between players. These simple social mechanics are common in current casual mobile titles. After a brief exploratory analysis we present binary prediction models of premium and social players. Results indicate that social and premium players rarely overlap, but that predicting both types of behaviors using typical observation windows is possible. The study then extends these results in using regression to predict the monetary value of a player, including social features in the set of predictors. Social features are shown to have minimal impact on this prediction. In summary, the present study advances previous work in the field by moving from churn and purchase prediction to an integrated prediction of CLTV, with an emphasis on social activity as predictor. Results highlight that social players, as defined in the current context of social network requests, and premium players may encompass distinctly different segments in casual mobile games.

3 RELATED WORK

While prediction modeling in freemium games is of commercial value, the vast majority of the work being done at various development and publishing houses is not publicly available [11, 23, 27]. Within published research, there are roughly two dozen publications available that directly target mobile game analytics, with the majority of these focusing on churn prediction. In this section we will thus focus on related work regarding churn, premium prediction, the definition of "social" in a gaming context, and the incorporation of social features through a network of player requests. Due to space constraints, we will focus on the key related works.

Churn Prediction

Churn and retention prediction in games is a recent topic, and even more recent outside of mobile games [28]. Hadiji et al. and Runge et al. [14, 23] formally defined the problem of churn prediction for F2P games, and proposed a range of features to integrate in models, presenting results for cross-game models. Runge et al. and Rothenbuehler et al. [22, 23] predicted the departure of high value players for casual F2P mobile games and investigated churn as a binary classification problem, comparing different classifiers and feature sets with an emphasis on Hidden Markov Models. Drachen et al. [10] introduced rapid prediction of retention using heuristics models, stressing the need to iteratively develop predictions in rapidly changing mobile games. Perianez et al. [19] describe a churn detection procedure using survival ensembles. Across the recent work on churn prediction in F2P mobile games, the importance of temporal features has been highlighted, e.g. features associated with the number of sessions per time period, the time between sessions, and average duration of sessions [14]. Features related to specific game design were generally reported to be less important for F2P titles than in other types of games, such as sandbox games [14, 23, 28, 32].

Premium and CLTV Prediction

The concept of Customer Lifetime Value has an extensive history in marketing, social media and finance research - among other domains - and is a core metric used for customer selection, segmentation, and marketing resource allocation as well as customer relationship management [21]. Prediction models for CLTV in these sectors build using a variety of models ranging from simple regression to complex machine learning [9]. These models directly inspired Game Analytics to begin adopting the same principles in the mid-2000s when *Facebook* games and other social media combined with the introduction of widespread mobile platforms provided a wealth of behavioral customer data to game development companies [1, 11, 24].

Within the context of games, Sifa et al. [26] introduced two models for predicting player purchase decisions as a three-step process, obtaining accuracies well over 80% across different observation windows. The authors formulate the process of predicting premium players (predicting that an IAP will take place) as a combination of a classification and a regression problem. Sifa et al. [26] also emphasizes the presence of rarity when analyzing premium players and provide a synthetic oversampling solution to predict rare purchase decisions, which is later combined with deep neural networks for predicting CLTV and recommending players based on their future value ranking [27]. Xie et al. [32] concentrated on predicting first purchases using standard classifiers in combination with a perspective on engagement modeling. To the best knowledge of the authors, there are no publications focused on impacts of social network features in CLTV prediction in games.

Social Network Analysis in Games

The analysis of relations between people has recently become a commonly employed tool outside of games, with online platforms such as *Facebook* and *Twitter* providing a direct vehicle for investigation. Prior work has targeted not only analysis of social networks themselves, but also their potential for product recommendation, advertising, and prediction of user behavior. Social Network Analysis (SNA) is less well-represented in games, but has applications in this context, where the value of social relationships to the players and on retention has been documented in a number of studies, e.g. Yee [33], focusing on Massively Multi-player Online Games (MMOs). In the context of eSports, Iosup et al. [17] examined networks in *DotA 2* and *StarCraft* with a specific focus on modeling social structures and Jia et al. [15] introduced networks generated from team-based match data.

Social Analytics Outside Games

There is a substantial amount of research on this topic outside games. To briefly cover a few key references from social media, Weiberg and Berger [30] introduced an extended definition of Customer Lifetime Value named *Connected Customer Lifetime Value*. It is defined as the present value of the contribution associated with purchases by that user plus the present value of the contributions associated with purchases by other users influenced by that user. This perspective on social value is seconded by Solis [29] who highlights the impact of social capital in CLTV. Wu et al. [31] working from a healthcare context, evaluated CLTV for the purpose of customer-specific marketing strategies. The authors developed a CLTV model which includes network influencing opportunities as well as churn risk assessment. The key element is regarding the market as a social network, and thus adding the social influence role of a customer into the CLTV evaluation. Jointly, the work on social value in CLTV contexts outside of games indicate that this is a perspective that bears investigation in games. Given the dearth of publicly available knowledge on the topic, we focus here on investigating and predicting social behavior itself, rather than trying to convert social behavior into a monetary figure.

4 DATASET AND METHODS

The analysis presented here was completed using tracking data of a cohort of over 219,000 players from a Free-to-Play (F2P, FtP) mobile game developed and published by a leading mobile game developer. The title of the game is omitted from this work due to the confidentiality of the data. The game is a casual puzzle game similar to Candy Crush with a Saga-based level progression, paywalls or gates every 20 levels, and offering in-app purchases. The game also includes the option to connect with a *Facebook* account, a feature that is prevelant among casual social games. The dataset was obtained in January 2017 and it consists of in-game behavior data for players that installed the game during the same month one year earlier. Three distinct types of behavior were observed: gameplay, purchase, and social. Each of these types of behavior are described by various metrics. Gameplay behavior consists of a player's playtime, the number of days that they played the game, the sessions and rounds that they completed (multiple rounds can be played within a

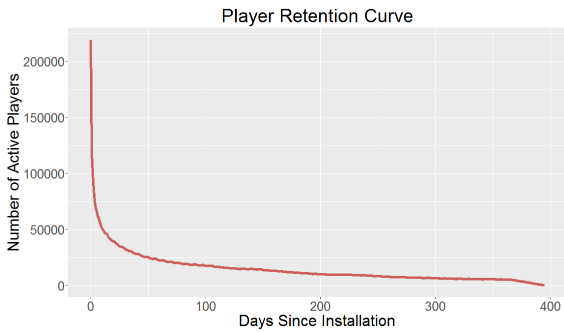

Figure 1: Number of active players over time since game installation for the 219,000 player cohort. The curve shape follows the power law commonly observed in F2P mobile games [14, 23, 32].

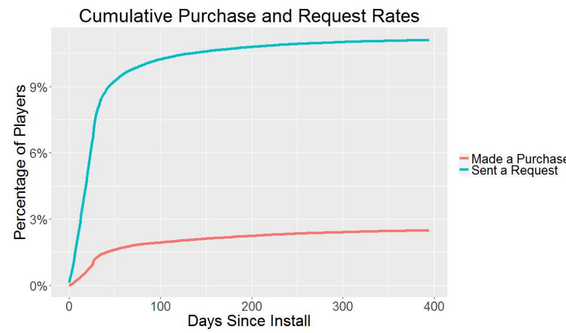

Figure 2: Percentage of players that purchase or send requests over time since game installation.

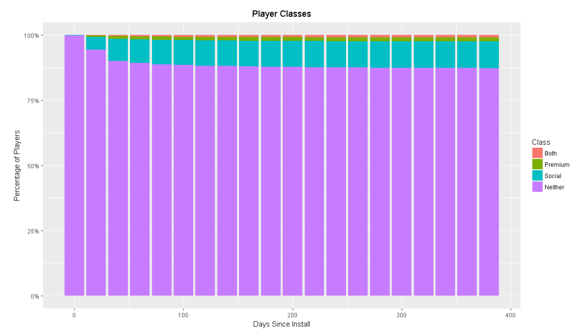

Figure 3: Distribution of the four player classes (Social, Premium, Both Social and Premium, Neither Social nor Premium) as a function of time since game install. The distribution of the population of the players stabilize around day 100. This pattern should be viewed against the rapid background drop in number of active players as seen in Fig. 1.

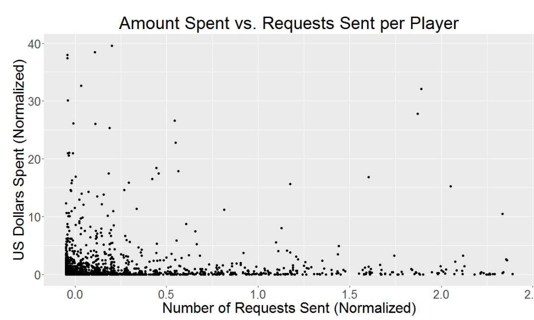

Figure 4: Spend vs. requests sent for players that did both

single session), and the player's performance in each round. Purchase behavior consists of each purchase (if any) that a player made. Social behavior consists of requests that a player sends to and receives from other players. The gameplay and purchase features just described match those used in previous work [23, 26, 32], but the social features have not been used before.

Imbalance in Dataset

It is important to understand the retention of players in our dataset. As is common with mobile games, many players "churn" (i.e. stop playing the game) only a few days after installation. However, a small portion of players play the game for many days, sometimes upwards of a year, after installation. The retention curve in Fig. 1 shows the number of active players as a function of time after game install.

It is also important to understand the imbalance present among players with respect to both social and purchase behavior. Only 11 percent of players ever sent a request, and an even smaller portion of players, 2.5 percent, ever made a purchase. Given that a player makes at least one purchase, their first purchase is likely to occur soon after installation. Only 23 percent of first-time purchases occur more than 90 days after installation. First-time requests sent are similar; only 9 percent occur more than 90 days after installation. Fig. 2 shows the cumulative purchase and social request rates across days since install.

There exists a trade-off in the number of purchases a player makes versus the number of requests they send. This is due to the structure of the game our dataset is based on. At certain points throughout the game, players are blocked from moving forward until they either make a purchase, send requests to their friends for help, or wait for a few days. Since purchases and requests can be used for the same purpose, it is rare for a player to have both made a large number of purchases and sent a large number of requests. This is visible on the scatter-plot in Fig. 4 - most players are located near the axes of the plot.

Social Network Features

In order to understand the social behavior of players on a deeper level than just the raw count of requests that they sent and received, a social network was created connecting players in the cohort. Each player was represented by a node in the network, and requests sent between players were represented by edges. The network formed was sparse, as only 11% of players sent any requests, but analysis of the network was still able to provide insight into player relationships. It was not possible to find any publicly accessible numbers on whether the sparsity of the network is comparable to other casual mobile titles. However, the kinds of social features in the game (i.e. integration with social network accounts, ability to help friends with boosts and in-game resources) are typical of casual mobile titles. Non-casual mobile games such as *Clash of Clans* and *Infinity Blade* can feature more extensive social features, including guild support and group-based activities, and it is possible that the networks are less sparse in such games.

Commonly used network metrics (closeness, degree, betweenness, page-rank, and eigencentrality) were calculated as features for each player. Also, the "number of triangles" that a player belonged to was calculated so that we could see which of the players played the game and utilized its social features with a group of friends.

Table 1 summarizes the features used in our analyses.

Table 1: Mean and Standard Deviation of All Numeric Features

	Feature	Mean	Std. Dev.
Gameplay Features	Total Playtime (hours)	12.1	34.2
	Total Sessions Played	141	354
	Total Rounds Played	305	841
	Total Days Played	27.0	51.7
	Avg. Session Duration (sec)	427	279
	Avg. Time Between Sessions (hours)	61.9	143
	Avg. Round Duration (sec)	109	44.4
	Average Moves per Round	13.23	4.5
	Max Level Reached	85.7	116
	Average Stars	104	147
	Current Absence Time (hours)	6820	2940
Social Features	Degree	7.41	42.9
	Closeness	3.03E-13	9.66E-18
	Page Rank	5.45E-07	3.51E-08
	Betweenness	1.98	60.1
	Eigen Centrality	5.20E-06	2.14E-03
	Number of Triangles	0.105	2.28
Purchase Features	Total Number of Purchases	0.236	4.54
	Total Amount Spent (USD)	0.767	20.1
Other Features	Marketing (binary)	0.329	0.470

Methods

We use a two-step process to predict the CLTV of players in our dataset. We first classify whether a player was a premium player or not, followed by predicting the monetary value that the player brings. For our classification task, algorithms such as Random Forests (RF) [4], Extreme Gradient Boosting (XGBoost) [6], which is an implementation of Gradient Boosting Machine (GBM), Adaptive Boosting (Adaboost) [13], and C5.0 Trees [20] were used. For the regression task, Random Forests and XGBoost were used. Both these tasks were done using an observation window of 7 days. 7 days is broadly used in the F2P game industry as the standard time frame to measure the behavior of a player [11, 24]. Both models were trained on the gameplay, purchase, and social features.

The Machine Learning (ML) algorithms applied are fast to implement for an iterative process, and accessible with proper documentation. Related work on churn prediction in games across industry and academia have explored similar algorithms, and these are adopted here to assist with cross-comparability of results in the domain [6, 14, 23, 26, 27, 32]. Amongst ML algorithms, tree-based methods are seen to be some of the more easily understood methods as they more closely mirror human decision-making. [14, 23]. Previous work has explored CLTV predictions in the mobile gaming space [26, 27] but does not incorporate features that measure the social behavior of players within the game and how that affects the CLTV they bring to the game. To the best of our knowledge, we are the first to take the social value into account when it comes to CLTV predictions. The insights presented here are seen as a first step in enabling companies to not only bolster revenue from players through targeted engagement, but also gain insight as to how players interact with each other and how that affects the monetary value they bring.

5 ANALYSES AND RESULTS

This section will be split into two major parts - first, we build classification models to identify players as premium or non-premium, as well as social or non-social. We build binary classification models for premium and social, as well as combining them into a four-class classification model. Then, we run regression models to predict CLTV, building further upon the customers we identified in the classification section and compare the results to other customer segmentation methods and heuristics.

Player Classification

Data Selection and Pre-processing: We begin with approaching the problem of classifying players as social or non-social, as well as premium or non-premium.

We define a social user as one who has sent at least one request to another person. This is based on the simple structure of social interactions in the game, which again is common in casual mobile titles. This is unlike the pattern in more strongly socially oriented games such as *Clash of Clans*, which can be played in single-player mode but include incentives to encourage players to form clans, which in turn gives access to further game features. Social requests, whether to form "friend" connections or subsequently to send requests along these connections, are in the current case sent through Facebook. This means that the person receiving a request may or may not be a current player of the game.

We define a premium player as one who has made at least one purchase within the game. In addition to binary classifications (social or not social, premium or not premium), we also combine the premium and social classifications into a four-class classification model: premium only, social only, both premium and social, or neither. The purpose of this four-class model is to enable evaluation of the interaction between the classes.

Given the short life cycle of mobile game players, we use a snapshot as of Day 7 within our dataset (i.e. the observation window used for model training) in order to classify premium and social players within the next week (Days 8 - 14) as well as within each player's lifetime (Days 8+). These windows are based on temporal windows are common in the publicly available work on mobile churn prediction [14, 23, 26, 32].

Because being a premium or social player is much less common than being non-premium or non-social, there is a huge imbalance in the dataset, as shown in Tables 2, 3, and 4. In order to relieve some of the issues caused by imbalance in the dataset, we use a mixture of downsampling the majority class and upsampling the minority class(es) using *Synthetic Minority Over-sampling Technique* (SMOTE) [5]. We use SMOTE since it's well-established and easy to implement.

As its name suggests, SMOTE creates synthetic examples of the minority class(es) by taking random observations within each class and perturbing it closer to a randomly chosen close neighbor. This perturbed observation becomes a new observation in the dataset, and the amount of observations in the minority class(es) grows as SMOTE continues to do this. Thus, SMOTE is generally seen as effective because it allows the decision boundary of the minority class(es) to be well-defined.

Classification Methodology: Four classification algorithms were utilized, based on consideration of related work

Table 2: Distribution of Premium vs. Non-Premium

Window	Premium	Non-Premium
Days 8 - 14	334	57,398
Days 8+	1,233	56,499

Table 3: Distribution of Social vs. Non-Social

Window	Social	Non-Social
Days 8 - 14	947	56,785
Days 8+	2,122	55,610

Table 4: Distribution of Four Class Classification

Window	Premium	Social	Both	Neither
Days 8 - 14	309	922	25	56,476
Days 8+	309	2,167	203	54,563

as described above. The models were: Random Forest [4], XGBoost [6], Adaboost [13], and C5.0 [20] algorithms. The use of multiple classifiers is common in game analytics and behavioral prediction in general for identifying the best models for a specific task [14, 23, 26]. We randomly split the dataset into training and test sets, then applied SMOTE to the training set only. Because the test set was not rebalanced, we used 50% of the data as the test set in order to capture sufficient amounts of the minority classes to test our model on. All models were subsequently ten-fold cross-validated.

We ran binary classification models for predicting premium players and predicting social players on our two time windows using all four methodologies just mentioned. For the four-class classification, we used Random Forest and XGBoost only, but still used both time windows .

Table 5: Premium Binary Classification Results

	Model Type	AUC	AUPR
Days 8 - 14	Random Forest	0.834	0.161
	Adaboost	0.799	0.134
	XGBoost	0.739	0.144
	C5.0	0.769	0.023
Days 8+	Random Forest	0.749	0.154
	Adaboost	0.740	0.113
	XGBoost	0.785	0.121
	C5.0	0.731	0.063

Classification Results: The results of the binary classification models are shown in Tables 5 and 6. We display two metrics to evaluate model performance, Area Under the Curve (AUC) and Area Under the Precision-Recall Curve (AUPR). AUC measures the probability that the model will

Table 6: Social Binary Classification Results

	Model Type	AUC	AUPR
Days 8 - 14	Random Forest	0.927	0.291
	Adaboost	0.930	0.321
	XGBoost	0.933	0.311
	C5.0	0.918	0.262
Days 8+	Random Forest	0.908	0.396
	Adaboost	0.908	0.443
	XGBoost	0.915	0.463
	C5.0	0.879	0.353

Table 7: Four-Class Classification Results

	Model Type	Accuracy
Days 8 - 14	Random Forest	0.861
	XGBoost	0.786
Days 8+	Random Forest	0.695
	XGBoost	0.673

classify a randomly chosen premium player as premium with more confidence than a randomly chosen non-premium player. However, since our test dataset is highly imbalanced, we also show AUPR which gives a more informative view on performance in such cases of high imbalance [7]. From our results, we can see that either Random Forest or XGBoost performs the best across all cases, which is why we only used those two methods for our four-class classification.

Table 8: Confusion Matrix - Four Class Random Forest, Days 8 - 14

	Premium	Social	Both	Neither
Premium	0.379	0.026	0.00	0.045
Social	0.183	0.782	0.750	0.088
Both	0.039	0.009	0.167	0.002
Neither	0.399	0.183	0.083	0.865

The results of the four-class classification are shown in Table 7. Due to the model being multi-class classification, only accuracy could be obtained as a model evaluation metric. However, Table 8 shows the confusion matrix for the Random Forest model trained on the Days 8 - 14 time window, which gives a better view into which classes the model performs better for. The confusion matrix shows for each actual class (on the columns), what proportion of those players were predicted in each class by the model (on the rows). Therefore, each column sums to 1. We can see that the model excels at predicting players that are Neither (i.e. neither premium nor social) as well as Social Only, whereas it struggles at predicting Premium Only and Both. Because Premium

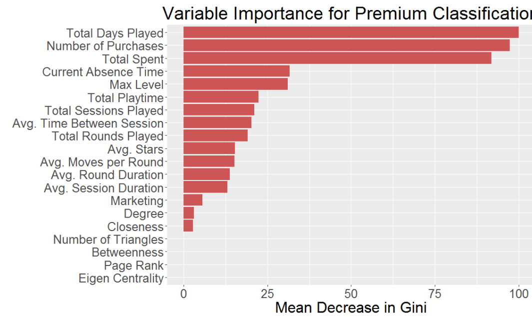

Figure 5: Feature importance for premium classification.

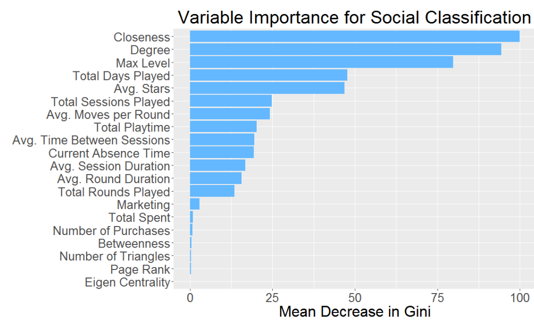

Figure 6: Feature importance for social classification.

players are a very low proportion of the total playerbase, and Premium and Social players are an even lower proportion of the playerbase, the model cannot identify the decision boundaries for these classes accurately, even with the training dataset rebalanced with SMOTE.

In Fig. 5 and Fig. 6, the variable importance from the Random Forest models (Days 8 - 14) for premium classification and social classification, respectively, are shown. In the premium classification variable importance, we see the most important features are total days played, number of purchases made, and total spent. This result agrees with previous literature that found players who have already made purchases within the game are the most likely to purchase again in the future [26]. On the other end, the social features did not seem to be predictive in the premium classification. This may be due to the sparsity in the network graph.

In the social classification variable importance, we see that the most important features are closeness and degree. However, the other network features were still unimportant in the social classification, which again indicates that the sparsity in the network graph affected those features' predictive power.

These results provide a new perspective on classification of player types within the context of mobile gaming. While binary classification on premium vs. non-premium players has been done before [26], examining social vs. non-social players and combining these classes into a four-class classification has not been explored before. These results could be useful to companies interested in finding the most influential players of its games - those who spend money in addition to having social clout.

CLTV Prediction Using Regression

Data Selection and Pre-processing: To be consistent with all previous classification models, we use same features described in the Dataset section across all regression models as well. Our response variable is each player's CLTV. In order to relieve some of the issues caused by the skewed distribution of player CLTV (i.e. an extremely long right tail), we alter the CLTV by adding 1 and applying a log transform.

We first subset the players based on their total amount spent. We build our CLTV prediction on players who have spent more than zero dollars on Days 8+ under the assumption that the classification model we developed could already distinguish between free users and paid users. This restriction reduced our sample size down to roughly 57,700 users.

We also look at other subsets of the playerbase in order to remove the assumption that we can completely distinguish between free users and paid users. We decided to use carefully selected segments that filtered out users that were not likely to spend money, as putting these users into the CLTV model would only decrease efficacy of the model. In order to be consistent with the classification models we built, we again use a snapshot as of Day 7 in the dataset and predict CLTV for Days 8+ for these alternative segments. We use four alternative segmentation methods:

- *Heuristic*: Players who have made a purchase in Days 0 - 7.
- *Predicted*: Players who we predicted to be premium on Days 8+ onwards with our best classification model.
- *Combined Users*: Combination of players who have made a purchase in Days 0 - 7 or who we predicted to be premium on Days 8+ (i.e. the union of the Heuristic and Predicted segments).
- *Play Session*: Players who had more than one play session in Days 0 - 7.

To get a better understanding of those four different segmentation methods, Table 9 indicates how many users are selected after each filtering method is applied. An immediate insight we found was since most people have more than one session from Days 0 to 7, this criterion might not help that much. While we still include it in our modeling, we focus more on the first three methods.

CLTV Prediction Methodology: We apply two popular machine learning methods, Random Forest and XGBoost.

Table 9: Summary of Four Datasets

Model Type	No. Players	No. Premium Players
Heuristic	404	215
Predicted	968	146
Combined	1,239	304
Play Session	55,582	1,458

We use these two methodologies since they were the best performing for our classification models and we are using the same features. All models were tuned to find the best parameters and ten-fold cross-validated. The same cross-validation partitions were used for each of two methodologies in order to faciliate fair comparision.

CLTV Prediction Results: The results of the models ran for the original segmentation of Days 8+ premium users is shown in Table 10. We use root-mean-square error normalized by the mean of the response variable (NRMSE) as the evaluation metric for each of the models, but show R^2 as well for informative purposes.

Table 10: Regression Results (Days 8+ Premium Users)

Model Type	NRMSE	R^2
Random Forest	0.938	0.096
XGBoost	1.062	0.097

We can tell that XGBoost outperformed Random Forest on both NRMSE and R^2. Yet, in general the predictive power of these models is not high enough, since the R^2 tells us that roughly only 9 percent of the variance can be explained by the model.

Table 11 shows the results of the four alternative segmentation methods. In addition to showing R^2 and NRMSE, we also show mean average error normalized by the mean of the response variable (NMAE).

Table 11: Regression Results (Alternative Segmentation Methods)

	Model Type	R^2	NRMSE	NMAE
Heuristic	Random Forest	0.187	1.372	1.574
	XGBoost	0.155	1.415	1.569
Predicted	Random Forest	0.301	0.791	0.578
	XGBoost	0.062	0.975	0.793
Combined	Random Forest	0.408	0.879	0.725
	XGBoost	0.365	0.922	0.760
Play Session	Random Forest	0.070	0.790	N/A
	XGBoost	0.019	0.856	N/A

Across all models, we can see that the Combined Users method has the best performance. Also, Random Forest is

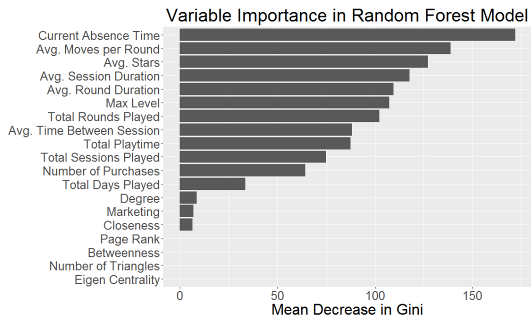

Figure 7: Feature importance for regression on days 8+ premium players.

better than XGBoost in general. Across all segmentation methods and model methodologies, the best model is Combined Users using Random Forest. These results show the importance of choosing an effective segmentation method, as the performance of any CLTV prediction is severely diminished when a segmentation method allows for a majority of the players to have a CLTV of zero.

Looking deeper at the results of the models, Fig. 7 shows the variable importance for the general regression on Days 8+ premium users using the Random Forest methodology. Current Absence Time is the strongest predictor. If we take a close look at all representing factors, playtime features and monetary features dominate the plot. This finding is largely consistent with previous research which has showed that people who play on a more consistent basis in the first week and achieve higher levels are much more likely to pay for the game [23, 26, 27]. Also, playing measures related to skill, such as Average Moves, Average Stars and Average Session Duration matter as well, since people may either play extremely well to unlock further levels of the game or make a purchase within the game. While all of these results agree with previous literature, we can see that this result seems to suggest social features are less relevant to the CLTV of users. The reason for this again seems to be the sparsity of social network features. However, these results still contribute to the research within the context of mobile gaming since CLTV models with social features as predictors have not been explored in previous literature. Games where social and purchasing features complement rather than substitute each other may find that social features are important to CLTV prediction.

6 CONCLUSION AND DISCUSSION

This paper targets two key challenges in mobile game analytics: a) Predicting the customer lifetime value of a player (CLTV); b) Evaluating the role of simple online social interactions on retention and CLTV. Results are derived from a unique dataset of 219,000 players from a F2P casual mobile title. Game-agnostic features are utilized to facilitate generalization of findings to the broaders space of casual social games [1]. Results can be summarized as follows:

1) We define a social player based on a review of selected literature and we present a generalizable operationalization in CLTV prediction for casual mobile games. Our operational definition is generalizable to other games with simple social mechanics, but cannot capture the full range of social interactions possible in titles with complex social mechanics, e.g. Massively Multi-player Online Games (MMOs) or team-based e-sports titles.

2) We build models for the classification and prediction of premium and social players. Results indicate that – at least in the title under investigation here – premium players have sparse contact with other players, and use IAPs to progress in the game. Social players, in the sense they are defined here, rarely convert to premium players, but use social connections for faster progression in the game. Both types of players are valuable, but for different reasons: Either they provide direct revenue contributions or they advertise and potentially recruit new players through social interaction and online word-of-mouth.

3) We frame CLTV prediction as a regression problem and incorporate social features in the prediction from standard observation windows [14, 26, 27]. Results indicate that players' online social interactions have no significant effect on their purchasing behaviors. The social features analyzed here do not encompass the range of potential social features in games, and analysis of games with more complex social mechanics might provide different results [8, 16].

The present work advances the state-of-the-art in mobile game analytics by adding CLTV and social perspectives on top of recent work on churn [14, 23], purchase [26] and CLTV [27] prediction. Results highlight that social and premium players can reflect distinctly different play types. It should also be noted that the multi-class classification approach – i.e. simultaneously predicting different aspects of player behavior – can be expanded to include other definitions of social behavior. Furthermore, it can be helpful in understanding other multi-category behavioral outcomes in mobile games [25].

Other studies have shown that, in freemium environments more broadly, social activity drives purchasing [3] and vice versa [2]. Hence, a key question is: Why do we not find evidence for such associations in our dataset? Many casual social games, like the game being analysed here, provide a social engagement layer that is auxiliary to the core product and game experience, they e.g. do not offer actual multiplayer gameplay [1]. Commonly, social features in casual mobile titles are limited to social media account integration and the sending of requests on the associated social network. Such a somewhat superficial integration and incentivization of social activities may not impact players' experience sufficiently to be behavior changing and habit inducing.

Summarizing, we explore the role of social behaviors in casual mobile games and their relationship to players' conversion to premium tiers and CLTV. Results are counterintuitive in showing the non-relevance of social features for purchase and CLTV prediction in casual social games. Prior studies [1] find that added social features, beyond the mere exchange of requests, can substantially drive revenue and engagement of players in casual social games. Game developers should hence strive for social interaction beyond social media account integration to drive revenue generation.

Finally, along these lines, it should be noted that social features are likely to be relevant to the prediction of CLTV in mobile games with deeper social interaction such as *Clash of Clans*, *Mobile Strike* or *Legendary – Game of Heroes*. It remains for future studies to explore CLTV prediction in such games, in further and different casual social games and online social games more broadly.

Acknowledgements

Part of this work was conducted in the Digital Creativity Labs (www.digitalcreativity.ac.uk), jointly funded by EPSRC/AHRC/InnovateUK under grant no EP/M023265/1.

REFERENCES

[1] A. Alsen, J. Runge, A. Drachen, and D. Klapper. 2016. Play With Me? Understanding and Measuring the Social Aspect of Casual Gaming. In *Proc. AIIDE*. AAAI.

[2] R. Bapna, J. Ramaprasad, and A. Umyarov. 2017. Monetizing Freemium Communities: Does Paying for Premium Increase Social Engagement?. In *Management Information Systems Quarterly*.

[3] R. Bapna and A. Umyarov. 2015. Do Your Online Friends Make You Pay? A Randomized Field Experiment On Peer Influence in Online Social Networks. In *Management Science*.

[4] L. Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[5] N. Chawla, K. Bowyer, L. O Hall, and W. Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16, 1 (2002), 321–357.

[6] T. Chen and C. Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD)* (2016).

[7] J. Davis and M. Goadrich. 2006. The relationship between Precision-Recall and ROC curves. *Proc. ICML* (2006).

[8] S. Dong-Hee and S. Youn-Joo. 2011. Why do people play social network games? *Computers in Human Behavior* 27 (2011), 852–861. Issue 2.

[9] B. Donkers, P C. Verhoef, and M. G. de Jong. 2007. Modeling CLV: A test of competing models in the insurance industry. *Quantitative Market Economy* 5 (2007), 163–190.

[10] A. Drachen, E. T. Lundquist, Y. Kung, P. Rao, R. Sifa, J. Runge, and D. Klabjan. 2016. Rapid Prediction of Player Retention in Free-to-Play Mobile Games. In *Proceedings of Artificial Intelligence in Interactive Digital Entertainment Conference*. AAAI.

[11] A. Drachen, C. Thurau, J. Togelius, G. Yannakakis, and C. Bauckhage. 2013. Game Data Mining. In *Game Analytics: Maximizing the Value of Player Data*, Drachen A. El-Nasr, M. S. and A. Canossa (Eds.). Springer.

[12] Fiksu DSP. 2017. Mobile Cost Indexes. https://fiksu.com/resources/mobile-cost-indexes/, (2017).

[13] Y. Freund and R. Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. System Sci.* 55 (1997), 119–139.

[14] F. Hadiji, R. Sifa, A. Drachen, C. Thurau, K. Kersting, and C. Bauckhage. 2014. Predicting Player Churn in the Wild. In *Proc. of IEEE CIG*.

[15] A. L. Jia, S. Shen, R. V. D. Bovenkamp, A. Iosup, F. Kuipers, and D. H. J. Epema. 2015. Socializing by Gaming: Revealing Social Relationships in Multiplayer Online Games. *ACM Trans. Knowl. Discov. Data* 10, 2 (2015), 11:1–11:29.

[16] H. Lin and S. Chuen-Tsai. 2011. Cash Trade in Free-to-Play Online Games. *Games and Culture* 6 (2011), 106–125. Issue 3.

[17] A. Iosup, R. van de Bovenkamp, S. Shen, A. Lu Jia, and F. Kuipers. 2014. Implicit Social Networks in Multiplayer Online Games. *IEEE Internet Comp.* 18, 3 (2014), 36–44.

[18] E. McDonald. 2017. The Global Games Market Will Reach $108.9 Billion in 2017. https://newzoo.com/insights/articles/the-global-games-market-will-reach-108-9-billion-in-2017-with-mobile-taking-42/, *Newszoo Insights* (2017).

[19] África Periáñez, Alain Saas, Anna Guitart, and Colin Magne. 2016. Churn Prediction in Mobile Social Games: Towards a Complete Assessment Using Survival Ensembles. In *Proc. of IEEE DSAA*.

[20] J. R. Quinlan. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.

[21] Venkatesan R. and V. Kumar. 2004. A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy. *Journal of Marketing* 68 (2004), 106–125.

[22] P. Rothenbuehler, J. Runge, F. Garcin, and B. Faltings. 2015. Hidden Markov Models for churn prediction. In *Proc. of SAI IntelliSys*.

[23] J. Runge, P. Gao, F. Garcin, and B. Faltings. 2014. Churn Prediction for High-value Players in Casual Social Games. In *Proc. of IEEE CIG*.

[24] E. Seufert. 2014. *Freemium Economics: Leveraging Analytics and User Segmentation to Drive Revenue*. Morgan Kaufmann.

[25] R. Sifa, A. Drachen, C. Bauckhage, C. Thurau, and A. Canossa. 2013. Behavior Evolution in Tomb Raider Underworld. In *Proc. IEEE CIG*.

[26] R. Sifa, F. Hadiji, J. Runge, A. Drachen, K. Kersting, and C. Bauckhage. 2015. Predicting Purchase Decisions in Mobile Free-to-Play Games. In *Proc. AAAI AIIDE*.

[27] R. Sifa, J Runge, C. Bauckhage, and D. Klapper. Customer Lifetime Value Prediction in Non-Contractual Freemium Settings: Chasing High-Value Users Using Deep Neural Networks and SMOTE.

[28] R. Sifa, S. Srikanth, A. Drachen, C. Ojeda, and C. Bauckhage. 2016. Predicting Retention in Sandbox Games with Tensor Factorization-based Representation Learning. In *Proc. IEEE CIG*.

[29] B. Solis. 2010. Social Capital: The Currency Of The Social Economy. http://www.briansolis.com/2010/03/social-capital-the-currency-of-digital-citizens/, (2010).

[30] B. Weinberg and P. Berger. 2011. Connected customer lifetime value: The impact of social media.

[31] L. Wu, L. Liu, and J. Li. 2005. Evaluating customer lifetime value for customer recommendation. In *Proc. SSSM*. IEEE.

[32] H. Xie, S. Devlin, D. Kudenko, and P. Cowling. 2015. Predicting Player Disengagement and First Purchase with Event-frequency Based Data Representation. In *Proc. IEEE CIG*.

[33] N. Yee and N. Ducheneaut. 2015. The Gamer Motivation Model in Handy Reference Chart and Slides. http://quanticfoundry.com/2015/12/15/handy-reference/. (2015).