

# Mining Player In-game Time Spending Regularity for Churn Prediction in Free Online Games

Wanshan Yang\*, Ting Huang†, Junlin Zeng†, Gemeng Yang†, Jintian Cai‡,  
Lijun Chen\*, Shivakant Mishra\*, Youjian (Eugene) Liu§

\*Department of Computer Science  
University of Colorado, Boulder  
Boulder, CO, USA

Email: {wanshan.yang, lijun.chen, mishras}@colorado.edu

†Department of Data Analytics  
Yoozoo Games  
Shanghai, China

Email: {thuang, zengjl, yanggm}@yoozoo.com

‡Department of Statistical Sciences  
University of Toronto  
Toronto, ON, Canada

Email: jintian.cai@mail.utoronto.ca

§Department of Electrical, Computer, & Energy Engineering  
University of Colorado, Boulder  
Boulder, CO, USA

Email: eugeneliu@ieee.org

**Abstract**—In the free online game industry, churn prediction is an important research topic. Reducing churn rate of a game significantly helps with the success of the game. Churn prediction helps a game operator identify possible churning players and keep them engaged in the game via appropriate operational strategies, marketing strategies, and/or incentives. Most churn prediction models are based on game-specific features, which limits their applicability to other games that do not share those features. In this paper, we consider developing universal features for churn predictions for long-term players. In particular, we mine player time spending regularity from data sets of two free online games. We leverage information from players' in-game time spending regularity in the form of universal features for churn prediction. Experiments show that our developed features are better at predicting churners, compared to the baseline features. The performance of our developed features is satisfactory even without game-specific features.

**Index Terms**—free-to-play games, churn prediction, data mining, supervised learning, feature engineering

## I. INTRODUCTION

Free online games allow players to access games for free. As in other freemium products and services, revenue of a free online game company depends on in-game purchases, and a larger player base indicates greater potential revenue. Retaining current players is usually much easier and less costly than recruiting new players. Therefore, these game companies strive to identify potential churners in order to retain them via proper operational strategies and incentive mechanisms.

Recent efforts on churn prediction in game industry have employed various methods and models such as binary predictions, survival ensembles and Cox model, and have utilized different features including playtime, login frequency, player in-game state and player in-game activity such as purchases; see, e.g., [1]–[9]. In [10], the churn prediction model is built based on user-app relationships in a game launcher platform.

However, almost all of the aforementioned work leverages game-specific features that are different for different games or may exist in one game but not in others. This limits their applicability. There is a lack of research in identifying universal features to identify potential churners that may work for most of the online games. In this paper, we consider the players' in-game time spending to develop such universal features for churn prediction. The players' in-game time spending is one of the most fundamental records in any game database, irrespective of any game-specific characteristics. It is also more reliable than other records such as the players' login data that sometimes depends on network conditions instead of the player behavior.

Specifically, we consider long-term players, who stay in the game for a sufficiently long period of time, and calculate the empirical distributions and related entropies of these players' in-game time spending. After observing the differences of related entropies of these players' in-game time spending between churners and non-churners, we develop universal features for churner prediction based on these distributions and entropies.

To the best of our knowledge, this paper is the first to take

advantage of data of long-term players' in-game time spending distribution, mine corresponding in-game time spending regularity for each player, and develop features from player in-game time spending regularity. The main contributions of our work are listed below.

- We model the long-term player in-game time spending regularity based on the data of players' in-game time spending distribution.
- We inspect churners' and non-churners' evolution of in-game time spending regularity across two free online games. Then we propose features based on player time spending regularity of long-term players.
- We conduct experiments to evaluate our developed features across the two free online games' data sets and show that these features could help achieve a better prediction performance than the baseline features.

The rest of the paper is organized as follows. Section II describes the game data sets we use. Section III explains how we split up the in-game time of a player into periods and how different player time spending distributions can be defined to capture the player time spending regularity of a player. Section IV illustrates how churners and non-churners evolve differently over time. Section V presents the process of feature engineering from player in-game time spending regularity and Section VI evaluates the performance of our proposed features. Section VII concludes the paper.

## II. FREE ONLINE GAME DATA SETS

This paper utilizes non-game-specific time spending data sets of two free online games *Thirty-six Stratagems* [11] and *Thirty-six Stratagems Mobile* [12].

### A. Background of *Thirty-six Stratagems* and *Thirty-six Stratagems Mobile*

*Thirty-six Stratagems* and *Thirty-six Stratagems Mobile* are two free online games published by Yoozoo Games. These two games share similar in-game topics. *Thirty-six Stratagems* runs on PC platform, while *Thirty-six Stratagems Mobile* runs on mobile platforms. They are operated separately by different game operators and have many differences in their respective gaming design, mechanisms, and events. In both *Thirty-six Stratagems* and *Thirty-six Stratagems Mobile*, millions of registered players manage their own resources, engage in land wars, and play as ancient lords. Each player develops his/her city, trains troops, and interacts with others in exchange for resources and honors [13]. The reasonable number of active players and the diverse in-game mechanisms make the data from these two games highly suitable to extract and evaluate our proposed features. Fig. 1 shows sample screenshots of the two games.

### B. Data Selection

Since a short playing time implies little information, we consider those long-term players who have played the game for at least 15 days.



(a) *Thirty-six Stratagems*



(b) *Thirty-six Stratagems Mobile*

Fig. 1. Screenshots of *Thirty-six Stratagems* and *Thirty-six Stratagems Mobile*.

To define churners and non-churners in the free online game, notice that the churners are unlikely to withdraw their accounts even if they stop playing the game for a long time. We thus define a churner as a player who does not access the game consecutively for a certain number of days. Motivated by [9], in both *Thirty-six Stratagems* and *Thirty-six Stratagems Mobile*, we define a churner as a player who stops playing for more than 3 days, because we find that in our data sets, over 95% of such players do not return to the game.

For our analysis, we randomly select a set of long-term players with the same numbers of churners and non-churners for both games. Among the players in the data set of *Thirty-six Stratagems* that were registered from April 23, 2018 to July 16, 2018, 7194 long-term players in total (3597 churners and 3597 non-churners) are randomly selected with corresponding 10865221 in-game time spending records. Among the players in the data set of *Thirty-six Stratagems Mobile* that were registered from November 1, 2018 to December 14, 2018, 3062 long-term players in total (1531 churners and 1531 non-churners) are randomly selected with corresponding 3225601 in-game time spending records. Those selected players all played in the same version of the corresponding game and had a consistent game experience.

## III. PLAYER TIME SPENDING DISTRIBUTION

The player time spending distribution describes how a player allocates his/her time spent in a given game. In this section, we consider the time spending distributions at different aggregation levels during the latest playing periods of a player. As will be seen later, these distributions will be the

basis for the proposed feature engineering and churn prediction method.

### A. Latest Playing Times and Periods

We consider the latest  $n$  days of playing times of a user. Let  $t_{u,d,r}$  be the playing time spent by user  $u$ , on day  $d$ , within hour  $r$ , where  $d = 1, 2, \dots, n$ ,  $r = 1, 2, \dots, 24$ . For example, for  $n = 15$ , if user 2 kept playing the game for 15 days continuously and his/her latest playing date is December 15th, then  $t_{2,15,1} = 0.1$  means user 2 spent 0.1 hour on December 15th between 0AM and 1AM. Similarly,  $t_{2,1,2} = 0.4$  means user 2 spent 0.4 hour on December 1st between 1AM and 2AM. A different user may have a different latest playing date. For example, if user 3's latest playing date is November 20th, then  $t_{3,15,4} = 0.7$  means user 3 spent 0.7 hour on November 20th between 3AM and 4AM. Formally,  $d$  indexes the  $n+1-d$ -th day to the latest playing date.

We partition  $n$  days into  $m$  periods of equal days, assuming  $n$  is divisible by  $m$ . For example, if  $n = 15$ ,  $m = 5$ , then each period has 3 days and the first period includes days in the set  $D_1 = \{1, 2, 3\}$  and the second period includes days in the set  $D_2 = \{4, 5, 6\}$ . Formally, the  $k$ -th period includes days in

$$D_k = \left\{ (k-1) \frac{n}{m} + i : i = 1, 2, \dots, \frac{n}{m} \right\}, \quad k = 1, 2, \dots, m.$$

Based on the latest playing times, we will calculate empirical probability distributions related to the in-game time spent by a player.

### B. Daily Time Spending Distribution

We first consider the total in-game time spent on each day and how each player distributes his in-game time over different days of a period. To this end, we define the individual (empirical) probability of the in-game time spending for player  $u$  on day  $d$  within period  $k$  as

$$p_{\text{ind}}(d|u, k) = \frac{\sum_{r=1}^{24} t_{u,d,r}}{\sum_{w \in D_k} \sum_{r=1}^{24} t_{u,w,r}}.$$

For instance, consider the latest  $n = 15$  days of playing the game for a certain player, with  $m = 5$  periods of 3 days. Assume that during the first period player  $u$  spends 5 hours, 6 hours, and 8 hours on the 1st, 2nd, and 3rd days, respectively. Then within the first period, the probabilities of the daily in-game time spending of this player over the 3 days are  $p_{\text{ind}}(1|u, 1) = \frac{5}{5+6+8}$ ,  $p_{\text{ind}}(2|u, 1) = \frac{6}{5+6+8}$ ,  $p_{\text{ind}}(3|u, 1) = \frac{8}{5+6+8}$ .

In addition, in order to capture the daily time spending distribution of the entire game community of  $N$  users, we introduce a “global” probability of the total daily time spending of all the players on the  $d$ -th day within the  $k$ -th period as

$$p_{\text{global}}(d|k) = \frac{\sum_{u=1}^N \sum_{r=1}^{24} t_{u,d,r}}{\sum_{u=1}^N \sum_{w \in D_k} \sum_{r=1}^{24} t_{u,w,r}}.$$

### C. Hourly Time Spending Distribution

We next consider the in-game time spent in each hour and how it is distributed over different days of a period. We define the empirical probability of the in-game time spending for player  $u$  in hour  $r$  on day  $d$  within period  $k$  as

$$P_{\text{ind}}(d|u, k, r) = \frac{t_{u,d,r}}{\sum_{w \in D_k} t_{u,w,r}}.$$

For instance, consider the same example as in the last subsection, and assume that during the first period player  $u$  spent 0.1 hour, 0.2 hour, and 0.3 hour in the hour 8 : 00 – 9 : 00 on the 1st, 2nd, and 3rd days, respectively. Then within the first period, the probabilities of the in-game time spending in the hour 8 : 00 – 9 : 00 of this player over the first period are  $P_{\text{ind}}(1|u, 1, 9) = \frac{0.1}{0.1+0.2+0.3}$ ,  $P_{\text{ind}}(2|u, 1, 9) = \frac{0.2}{0.1+0.2+0.3}$ , and  $P_{\text{ind}}(3|u, 1, 9) = \frac{0.3}{0.1+0.2+0.3}$ .

Similarly, in order to capture the hourly time spending distribution of the entire game community, we introduce a global probability of the time spending of all the players in hour  $r$  on the  $d$ -th day within the  $k$ -th period as

$$p_{\text{global}}(d|k, r) = \frac{\sum_{u=1}^N t_{u,d,r}}{\sum_{u=1}^N \sum_{w \in D_k} t_{u,w,r}}.$$

The afore-introduced player time spending distributions will be the basis to extract new features for the churn prediction.

## IV. PLAYER TIME SPENDING REGULARITY: CHURNERS VERSUS NON-CHURNERS

Since we aim to predict possible churners in this paper, for each game, the data set is divided into a training data set and a test data set via a 50% : 50% split. In this section, we examine player time spending patterns of churers and non-churers from the training data set at different timescales for each game, with the aim to identify possible differentiator between churers and non-churers.

### A. Entropy and In-game Time Speeding Pattern

Based on the player time spending distributions introduced in Section III and motivated by [14], we use the notion of *entropy* from information theory as the metric to characterize variance and change in the in-game time spent by a player [15]. Given a probability distribution  $p(\cdot)$ , its entropy is defined as

$$H(p) = \sum_x p(x) \log \frac{1}{p(x)}.$$

A higher entropy means a more even distribution and more regular time spending pattern, and a smaller entropy implies a less even distribution and more irregular/casual time spending pattern.

For each game and the data set described in Section II, we consider the latest  $n = 15$  days playing time and divide it into 5  $m = 5$  periods of 3 days. We calculate the corresponding time spending distributions and entropies for each player. The mean entropies and 95% confidence intervals of churers and

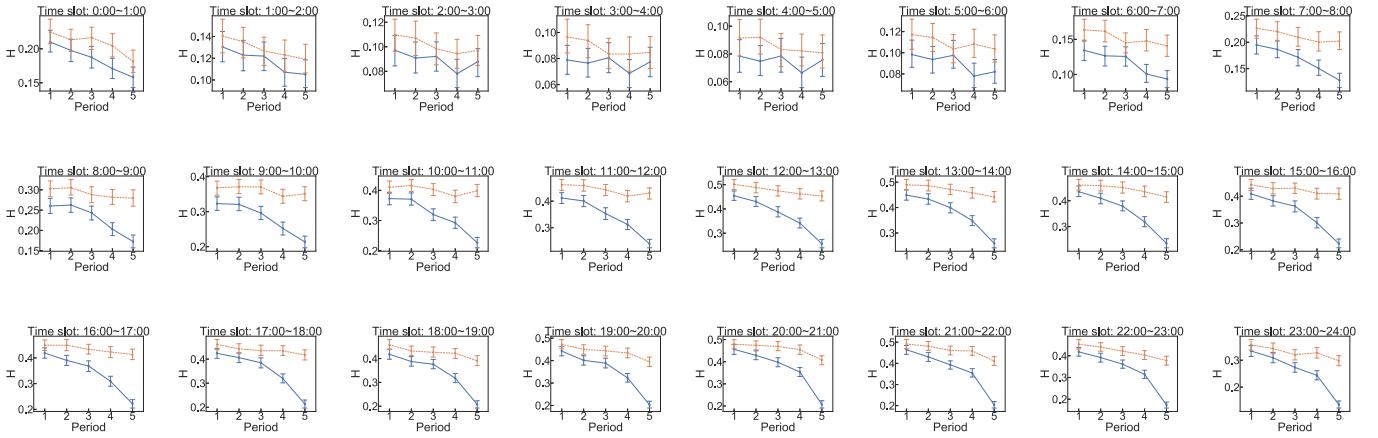


Fig. 2. The mean entropies of the distributions of hourly time spending of *Thirty-six Stratagems* players in different periods. The orange dashed line shows the non-churners, while the blue solid line shows the churners.

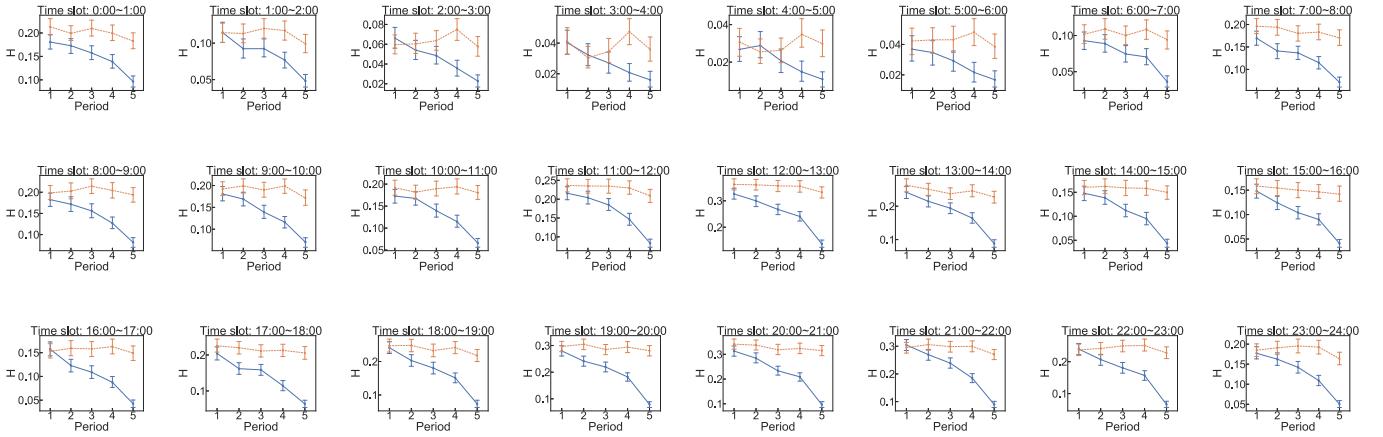


Fig. 3. The mean entropies of the distributions of hourly time spending of *Thirty-six Stratagems Mobile* players in different periods. The orange dashed line shows the non-churners, while the blue solid line shows the churners.

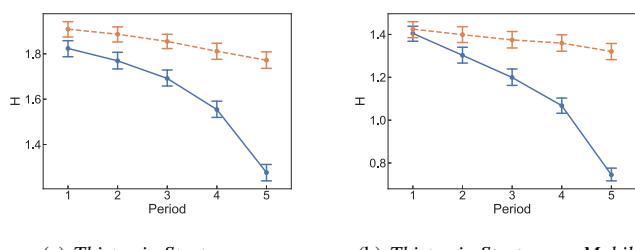


Fig. 4. The mean entropies of the distributions of daily time spending of players in different periods. The orange dashed line shows the non-churners, while the blue solid line shows the churners.

non-churners in different periods are shown in Fig. 4, Fig. 2, and Fig. 3. Fig. 4 illustrates the entropy distributions of churners and non-churners in the two games on a granularity of each day, while Fig. 2 and Fig. 3 illustrate the entropy distributions on a granularity of each hour for the two games

respectively.

We see that non-churners have significantly a higher mean value of entropy than churners, in both daily and hourly time spending distributions and in both games. This implies that non-churners have much more regular in-game time spending pattern than churners. Moreover, the entropies of non-churners exhibit a small decrease as the time goes on, while that of churners exhibit a significant decrease. This implies that non-churners have a more regular in-game time spending pattern than churners across different timescales, while churners spend their in-game time more and more casually as the time goes on.

A further look at the hourly entropies in Fig. 2 and Fig. 3 shows that the above mentioned difference between churners and non-churners is more significant in the hours from 8 : 00 to 24 : 00 and less significant from 0 : 00 to 8 : 00. This is consistent with the fact that 0 : 00 – 8 : 00 is the most common sleep time, and it is hard for the majority of players (churners or non-churners) to maintain a regular in-game time

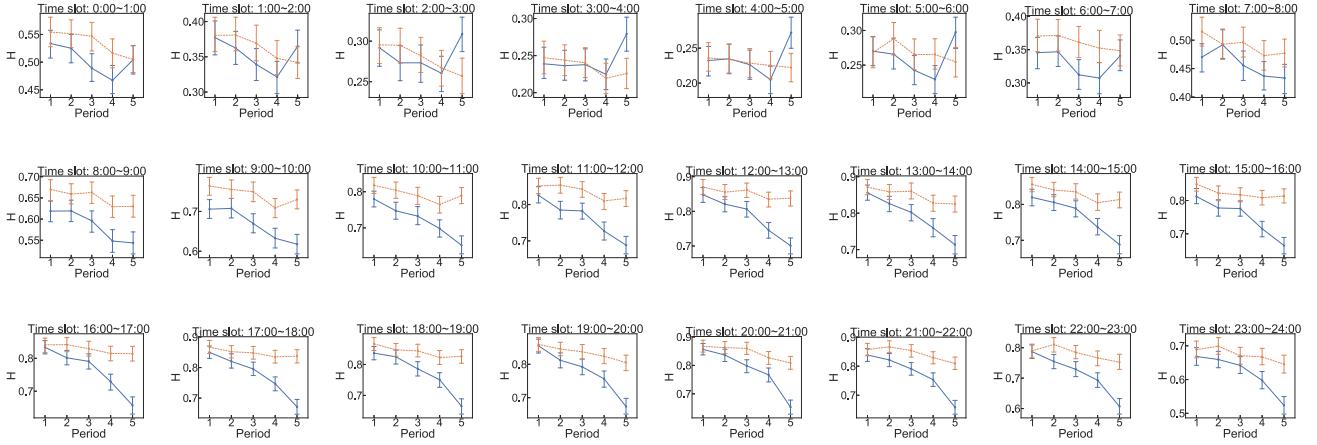


Fig. 5. The mean cross-entropies of the distributions of hourly time spending of *Thirty-six Stratagems* players in different periods. The orange dashed line shows the non-churners, while the blue solid line shows the churters.

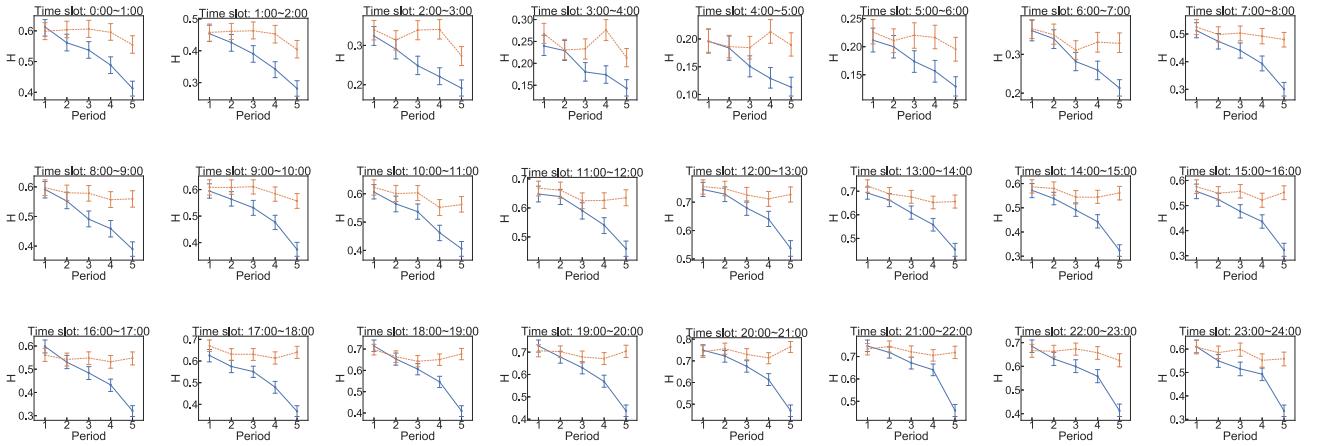


Fig. 6. The mean cross-entropies of the distributions of hourly time spending of *Thirty-six Stratagems Mobile* players in different periods. The orange dashed line shows the non-churners, while the blue solid line shows the churters.

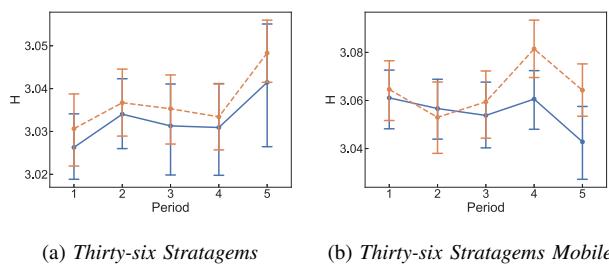


Fig. 7. The mean cross-entropies of the distributions of daily time spending of players in different periods. The orange dashed line shows the non-churners, while the blue solid line shows the churters.

spending pattern during that period of time.

#### B. Cross-entropy and Correlation with Aggregate Pattern

We now examine how each player compares to the entire game community that he/she is playing with over the same corresponding playing period using the notion of *cross-entropy*.

from information theory. Given an individual player's time distribution  $p_{\text{ind}}$  and the global time distribution  $p_{\text{global}}$ , the cross entropy between these two distributions is defined as:

$$H(p_{\text{ind}}, p_{\text{global}}) = \sum_x p_{\text{ind}}(x) \log \frac{1}{p_{\text{global}}(x)}.$$

A smaller cross-entropy means that the time spending pattern of a player is more similar to that of the game community as a whole.

For each game, we calculate the corresponding global time spending distributions and cross-entropies. The mean cross-entropies and 95% confidence intervals are shown in Fig. 7, Fig. 5, and Fig. 6. We see from Fig. 7 that there is no significant difference in the cross-entropy of the daily in-game time distribution between churters and non-churters.

On the other hand, Fig. 5 and Fig. 6 show that non-churters have a significantly higher mean value of cross-entropy in the hourly in-game time distribution than churters. Further, the difference between churter and non-churter is more significant in the hours from 8 : 00 to 24 : 00 and less

significant from 0 : 00 to 8 : 00. This characteristics of the hourly cross-entropy is similar to that of the hourly entropy presented in the last subsection.

### C. Player Time Spending Regularity for Churn Prediction

To summarize, churners and non-churners exhibit different in-game time spending regularities or patterns as captured by the entropies in the time spending distributions:

- Churners have lower entropies and larger time spending irregularity, as well as larger decrease in entropy than non-churners as the time moves on.
- Churners have increasingly lower cross-entropies in the hourly time spending distribution with the game community as the time moves on.

In the next two sections, we will exploit these differences to engineer entropy-based features for churn prediction.

## V. FEATURE ENGINEERING

The observation in Section IV shows that churners and non-churners exhibit different in-game time spending regularity that can be captured by the corresponding entropies. In this section, we propose several features based on entropies that will be used for churn prediction in the next section.

### A. Static Feature and Rate Feature

For a given time spending distribution for player  $u$  in  $k$ -th period ( $k \leq m$ ), we define a function  $f_{\text{distribution}}(u, k)$  which returns the corresponding entropy or cross-entropy. We call the calculated value of  $f_{\text{distribution}}(u, k)$  a static feature. Based on the static feature, we define a rate feature:

$$g_{\text{distribution}}(u, k) = \frac{f_{\text{distribution}}(u, k) - f_{\text{distribution}}(u, k+1)}{f_{\text{distribution}}(u, k)}$$

to capture the change in entropy as the time moves on where  $k < m$ . Recall from the last section that churners exhibit smaller entropies but with a greater entropy decrease as the time goes. The rate feature amplifies the differences between churners and non-churners.

### B. Feature Selection

As seen in Section IV-A and Section IV-B, churners and non-churners exhibit differences in entropy of the daily time spending distribution and of the hourly time spending distribution, as well as in cross-entropy of the hourly time spending distribution. We therefore select four types of features as follows:

- The combination of the static feature and rate feature of entropy of the daily time spending distributions. We call this type of features as the 1st type of developed features.
- The combination of the static feature and rate feature of entropy of the hourly time spending distributions. We call this type of features as the 2nd type of developed features.
- The combination of the static feature and rate feature of cross-entropy of the hourly time spending distributions.

We call this type of features as the 3rd type of developed features.

- The combination of 1st, 2nd, and 3rd types of features. We call this type of features as the combined type of developed features.

## VI. CHURNER PREDICTION

In this section, we evaluate the efficacy of churn prediction using the entropy features with several typical classifiers.

### A. Evaluation Strategy

1) *Baseline Features*: To evaluate the effectiveness of our proposed features, we use the following baseline features for comparison:

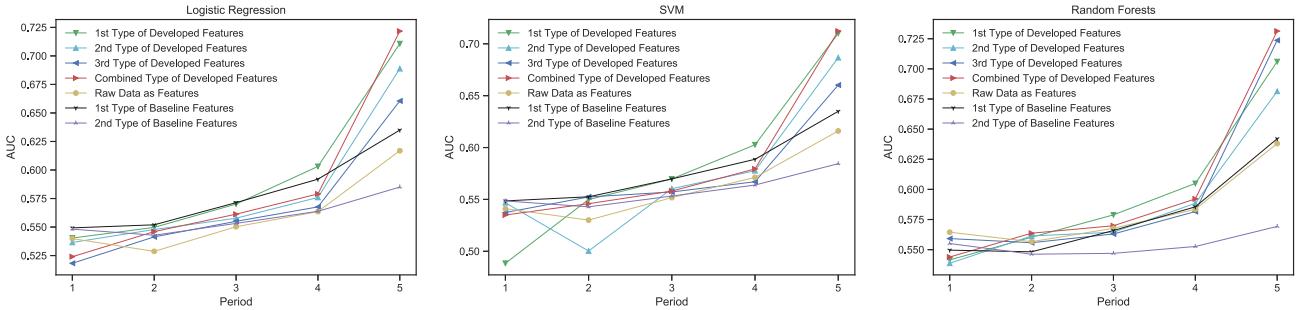
- The raw data of the in-game time distribution of players.
- The daily total time spent, which we call the 1st type of baseline features.
- The combined feature of the total time spent, the last day of login, and the number of time slots played. This baseline feature is designed based on RFM model [16]. We call this type of feature as the 2nd type of baseline features.

2) *AUC Evaluation*: Given the binary nature of the prediction/classification, we use the area under the Receiver-Operating-Characteristics (ROC) [17] curve (AUC) [18] to evaluate the overall performance of a classifier. Note that, AUC has been used as a metric to evaluate the performance of a certain churn prediction model in many previous work, such as [1], [4], [6], [19].

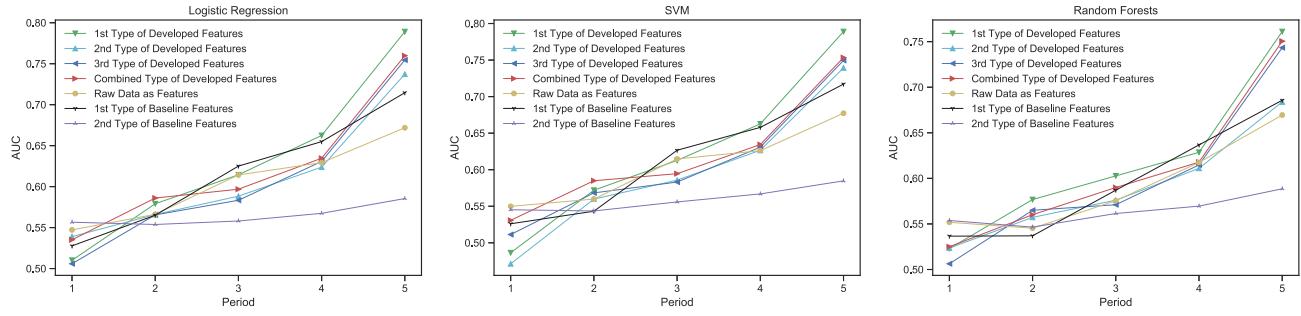
### B. Evaluation Results

We use multiple classifiers (Logistic Regression, SVM, Random Forests) and the aforementioned features for churn prediction. For each game, the data set is divided into a training data set and a test data set via a 50% : 50% split. We train the classifiers using the training data set and evaluate their performance using the test data set. Since we are also interested in evaluating how early we could detect the possible churning via our model, we evaluate the performance of the classifiers that are trained with the data of the first  $q$  latest playing periods on the test data set ( $1 \leq q \leq 5$ ). The results are shown in Fig. 8 and Table. I.

We see that, as expected, larger AUC can be achieved with training data of longer period. However, the “marginal gain” in AUC is moderate with the data of the first  $q$  periods when  $q \leq 4$ , while there is a significant jump in AUC with the data of 5th period. This implies that the most recent data is critical for performance of churn prediction and the performance of early churn detection is not good. In particular, the proposed entropy features have the best performance with the highest AUC (0.731350) for *Thirty-six Stratagems* and AUC (0.789158) for *Thirty-six Stratagems Mobile*, as opposed to the baseline features with the highest AUC (0.642020) for *Thirty-six Stratagems* and AUC (0.717088) for *Thirty-six Stratagems Mobile*.



(a) *Thirty-six Stratagems*



(b) *Thirty-six Stratagems Mobile*

Fig. 8. AUC result for the different proposed features with training data of the first  $q$  periods of the latest playing periods for players in test data set ( $1 \leq q \leq 5$ ).

TABLE I  
AUC FOR DIFFERENT PROPOSED FEATURES (CLASSIFIERS TRAINED WITH THE DATA OF ENTIRE 5 PERIODS)

		<i>Thirty-six Stratagems</i>			<i>Thirty-six Stratagems Mobile</i>		
		Logistic Regression	SVM	Random Forests	Logistic Regression	SVM	Random Forests
Developed Features	1st Type	<b>0.710750</b>	0.710065	0.706014	<b>0.789158</b>	0.788974	0.761047
	2nd Type	<b>0.688832</b>	0.686657	0.681387	0.737263	<b>0.739304</b>	0.683868
	3rd Type	0.660539	0.660212	<b>0.723693</b>	<b>0.754539</b>	0.749525	0.743388
	Combined Type	0.721714	0.712261	<b>0.731350</b>	<b>0.759767</b>	0.753405	0.750510
Baseline Features	Raw Data	0.616890	0.615981	<b>0.638047</b>	0.671846	<b>0.677245</b>	0.669476
	1st Type	0.634996	0.634947	<b>0.642020</b>	0.714608	<b>0.717088</b>	0.685689
	2nd Type	<b>0.585044</b>	0.584443	0.569255	0.585449	0.584704	<b>0.588557</b>

The evaluation shows that the proposed entropy features outperform the baseline features in churn prediction, and indeed capture the differences in the in-game time between churners and non-churners. In other words, although players' in-game time spending serves as an important feature in previous work, our proposed entropy features could exploit the information extracted from players' in-game time spending more effectively.

## VII. CONCLUSIONS

In this paper, we address the problem of predicting churners in free online games using universal features that are agnostic to the specific characteristics of the games. The goal is to develop models that can predict churners in a variety of games, each with different characteristics. We consider long-term players and understand their in-game time spending regularity among churners and non-churners. We observe that there is

a significant difference between churners and non-churners in terms of entropies exhibited in in-game time spending regularity. Based on this observation, we propose prediction models using new features extracted from mining players' in-game time spending regularity. After experiments are conducted, the corresponding result shows that our developed features are better at predicting churners, compared to the baseline features. Even without other game-specific features (e.g., login records, in-game events frequency, and payment information), we can still leverage the information from our examination of player in-game time spending regularity in the form of features and achieve satisfactory prediction performance. Thus, game companies can benefit from our algorithms without worrying about the specific characteristics of their games. Our findings can also help game developers design better in-game mechanisms to reduce churn rate.

## ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their valuable comments and suggestions.

## REFERENCES

- [1] H. Xie, S. Devlin, D. Kudenko, and P. Cowling, "Predicting player disengagement and first purchase with event-frequency based data representation," in *2015 IEEE Conference on Computational Intelligence and Games (CIG)*, Aug. 2015, pp. 230–237.
- [2] F. Hadjii, R. Sifa, A. Drachen, C. Thurau, K. Kersting, and C. Bauckhage, "Predicting player churn in the wild," in *2014 IEEE Conference on Computational Intelligence and Games*, Aug. 2014, pp. 1–8.
- [3] A. Drachen, E. T. Lundquist, Y. Kung, P. S. Rao, D. Klabjan, R. Sifa, and J. Runge, "Rapid Prediction of Player Retention in Free-to-Play Mobile Games," *arXiv:1607.03202 [cs, stat]*, Jul. 2016.
- [4] J. Runge, P. Gao, F. Garcin, and B. Faltings, "Churn prediction for high-value players in casual social games," in *2014 IEEE Conference on Computational Intelligence and Games*, Aug. 2014, pp. 1–8.
- [5] R. Sifa, S. Srikanth, A. Drachen, C. Ojeda, and C. Bauckhage, "Predicting Retention in Sandbox Games with Tensor Factorization-based Representation Learning," in *2016 IEEE Conference on Computational Intelligence and Games*, Sep. 2016, pp. 1–8.
- [6] M. Tamassia, W. Raffe, R. Sifa, A. Drachen, F. Zambetta, and M. Hitchens, "Predicting player churn in destiny: A Hidden Markov models approach to predicting player departure in a major online game," in *2016 IEEE Conference on Computational Intelligence and Games*, Sep. 2016, pp. 1–8.
- [7] P. Bertens, A. Guitart, and Á. Periéñez, "Games and big data: A scalable multi-dimensional churn prediction model," in *2017 IEEE Conference on Computational Intelligence and Games*, Aug. 2017, pp. 33–36.
- [8] J. Jeon, D. Yoon, S. Yang, and K. Kim, "Extracting gamers' cognitive psychological features and improving performance of churn prediction from mobile games," in *2017 IEEE Conference on Computational Intelligence and Games*, Aug. 2017, pp. 150–153.
- [9] E. Lee, B. Kim, S. Kang, B. Kang, Y. Jang, and H. K. Kim, "Profit Optimizing Churn Prediction for Long-term Loyal Customer in Online games," *IEEE Transactions on Games*, pp. 1–1, 2018.
- [10] X. Liu, M. Xie, X. Wen, R. Chen, Y. Ge, N. Duffield, and N. Wang, "A Semi-Supervised and Inductive Embedding Model for Churn Prediction of Large-Scale Mobile Games," *arXiv:1808.06573 [cs, stat]*, Aug. 2018.
- [11] Yoozoo Games, "Thirty-six Stratagems," [Online]. Available: <http://36.youzu.com/>, May 2017.
- [12] ——, "Thirty-six Stratagems Mobile," [Online]. Available: <http://36m.youzu.com/>, Jul. 2018.
- [13] W. Yang, G. Yang, T. Huang, L. Chen, and Y. E. Liu, "Whales, Dolphins, or Minnows? Towards the Player Clustering in Free Online Games Based on Purchasing Behavior via Data Mining Technique," in *2018 IEEE International Conference on Big Data (Big Data)*, Dec. 2018, pp. 4101–4108.
- [14] M. Rowe, "Mining User Development Signals for Online Community Churner Detection," *ACM Trans. Knowl. Discov. Data*, vol. 10, no. 3, pp. 21:1–21:28, Jan. 2016.
- [15] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (2. Ed.). Wiley, 2006.
- [16] P. S. Fader, B. G. Hardie, and K. L. Lee, "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, vol. 42, no. 4, pp. 415–430, Nov. 2005.
- [17] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [18] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.
- [19] E. G. Castro and M. S. G. Tsuzuki, "Churn Prediction in Online Games Using Players' Login Records: A Frequency Analysis Approach," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 3, pp. 255–265, Sep. 2015.