Advanced Materials Research, vols. 931-932, pp. 1370-1374, 2014.

Departure Prediction of Online Game Players

Kittipat Savetratanakaree[1,a], Kingkarn Sookhanaphibarn[2,b], Sarun Intakosum[1,c], Ruck Thawonmas[4,d],Kuan-Ta Chen[5,e]

[1]Department of Computer Science, Faculty of Science, King Mongkut's Institute of Technology, Ladkrabang, Thailand.

[2]School of Science and Technology, Bangkok University, Thailand.

[3]Department of Human and Computer Intelligence, Ritsumeikan University, Kusatsu, Shiga, Japan

[4]Institute of Information Science, Academia Sinica, Nankang District, Taipei, Taiwan

[a]kittipatsavet@gmail.com, [b]kingkarns@gmail.com, [c]intakosumsarun@gmail.com, [d]ruck@ci.ritsumei.ac.jp, [e]ktchen@iis.sinica.edu.tw

Keywords: data mining, user behaviors, game revisitations, player behavior, massively multiplayer online role-playing game(MMORPG), Shen Zhou online

Abstract. Most business models of online game companies usually depend on sale of virtual items and the monthly subscription fees. The prediction of player departure could increase revenues by giving special promotions out to the players who are expected to unsubscribe or quit playing the game. This paper proposes a departure prediction approach by using a new feature called "*SLKdays*" and a *game revisitation*. The feature "*SLKdays*" is defined as "*Staytime*", which is the time each player spending in an online game, of the last k days, and a *game revisitation* is the playing frequency in the last month to predict the next month subscription. We explore our new feature "*SLKdays*" to determine the optimal number of k for the departure prediction. With our proposed feature, the accuracy of the departure prediction is high, 91.92%, and the precision and recall rate are 98.22% and 84.51%.

Introduction

Analyses in games and players [1, 2, 3, 4, 5, 6, 7, 8] are being enthusiastically conducted by a game research community to improve the game design and performance, and the business strategy. This paper is the extending study in the analysis of revisitations in online games [2]. Thawonmas et al. [2] analyzed game revisitations focusing on how long online-game players revisiting the game after their last logout. In this paper, we hypothesize that online-game player's information of last month (last 30-days interval) can predict that a player will continue or absent from the game in the next month (next 30-days interval).

In order to validate our hypothesis, we conduct four experiments with our new feature *SLKdays*, where *SLKdays* is defined as the stay time of the last k days of the last month, to predict online-game players'behaviors in the next 30-days interval. Four experiments are designed for a real-world use, i.e., the training set will use the information of the previous year as called one year-ahead prediction. We also tested our approach for the two-years ahead prediction. To build the prediction model, we use *SLKdays* combined with the game revisitation as the extracted feature. Our prediction approach will be explained in the rest of this paper.

Background and Related Works

Online game player's information has been shown effectively in providing useful information in a variety of research domains. There seems to be an increased interest in online-game player's information in game-oriented AI research, adaptive game research, player experience modeling, and game user research. We review various analyses related to online game player's information as follows.

Table 1: Four experiments with different year (y) for training set and test set.

Experiments	Test set (year)	Training set	
		$y-2$	$y-1$
1.1	2005	-	2004
1.2	2006	-	2005
2	2006	2004	-
3	2006	2004	2005

Based on a set of EverQuest II, Shim et al. [4] used the game's player activity data in game to construct profiles of online game players'behaviors for the pattern recognition of normal and abnormal behaviors. Analyzing the history of large behavioral data in a game provide valuable insight into a range of character types (i.e. archetype, classes, sub-classes, race) which exist in the game. The authors examined the player efficiency in terms of total experience point (XP point) and find that there is an overall trend with respect to how fast particular sub-class advance throughout the game and changes in task types as players advance throughout the game.

Whang et al. [5] explored different players' behaviors from players' lifestyles, focusing on how a game player adopted the virtual game world as part of their life. The lifestyles in the online game, Lineage, were classified into three categories as off-real world gamer, community-oriented player, single-oriented player. The study model is to understand how players with different real life backgrounds will play to the various features of a game and how they can adopt their new social identities in the virtual game world.

Chen et al.[6] focused on how network quality and network loss affect a player's decision to unsubscribe an online game prematurely. Trace collection and traffice measurement was conducted by monitoring the traffic of game servers. The results indicate that both network loss and network delay significantly affect a player's decision to quit playing a game too early.

Lou et al. [7] proposed a forecast model for online game addictiveness according to players' emotional responses. The model using additional hardware, electro-myographic measures (EMG), attached players'two facial muscles which the device can indicate brain signal in order to know humans' positive and negative emotions. The work can ensure that a game's design is on the right track in the early of game development phase.

All papers above focus on different aspects from our paper.

Proposed Methodology

Data description

We obtained the traces of 355,706 Shen Zhou Online (SZO) accounts from January 1^{st}, 2004 to April 1^{st},2007 as courtesy of User Joy Technology Co.Ltd. SZO is a massively multiplayer online role-playing game (MMORPG) in Taiwan. There are 119,082,865 sessions logged but not all of the sessions are focused. Data pre-processing are required to obtain a suitable data set for our analysis. We filtered the following kinds of records out from the data set.

1. Error records: They are those whose logout time is before the login time.

2. Duplicate records: player id, login-time, stay time, and logout-time are the same. This happens when the player's login time is during the highest network connection rate.

3. Short-term records: They are those whose players logged in less than a month. These players are absent from the game before the end of 30-day free trial period. We focus on players who login and stay in the game for more than a month, or the free trial period of time.

Table 2: Classification of 13 predefined bins of *game revisitations*.

Bin (j^{th}).	1	2	3	4	5	6	7	8	9	10	11	12	13
Time Intervals (m)	32	64	98	136	212	424	848	1696	3392	6784	13568	27136	$>$ 27136
Time Intervals (approx.)	30m	1h	1.5h	2h	3h	6h	12h	1d	2d	4d	8d	16d	$> 16d$

Note that m = minute or minutes, h = hour or hours, d = day or days

4. In-2007 records: The players information of interest is the records of 12 months in each year. Since there are the data log of only 3 months in the year 2007, the players' information in year 2007 is excluded from our consideration.

In this study, we conducted four experiments, each with a different pair of training and test data set as shown in Table 1.

Departure prediction approach

First, we consider the playing frequency on *game revisitations* [2], which it is defined as the number of times each player revisits to play the same game within the predefined time intervals (bins) of N days. We base our prediction method on the daily play time on *game revisitations* within 30-day intervals classified into 13 predefined bins as shown in Table 2. Each 30-day intervals of daily play time is 5-days overlap.

Proposed feature *SLKdays*

Second, we calculate the stay time per day (*Staytime*) for each player as follows :

$$Staytime \quad = \quad Logout\_time – Login\_time \tag{1}$$

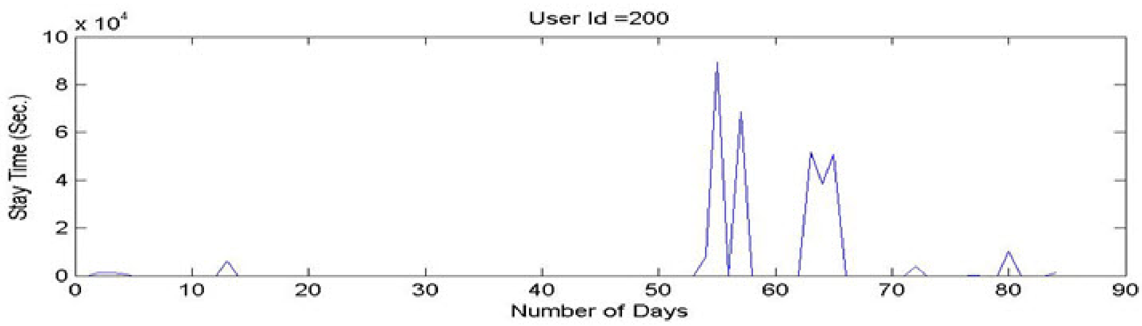

Fig. 1: Daily *Staytime* of user ID 200.

Fig. 1 shows that some players did not play the online game regularly. We smooth the daily play by averaging it in an interval such as every 5 days. From our study, we discovered that the stay time of last k days is significantly related to their continuing or absent in the next month. We also applied the logarithm to transform the linear data. In summary, we calculate the logarithm of average *Staytime* of last k days (*SLKdays*) in each 30-days interval written as follows :

$$SLKdays \quad = \quad \log\left(\frac{\sum_{i=2}^{k} S_i}{k}\right) \tag{2}$$

where S_i is daily *SLKdays* of last day $i = 2, 3, \ldots, k$ days.

Third, the normalization of *game revisitations* in each bin and that of *SLKdays* are required and its equation is shown below.

$$\frac{Bin[j]\_of\_player[i]}{average\_of\_Bin[j]\_among\_all\_users} \tag{3}$$

where $bin[j]$ of player $[i]$ is the number of times $player[i]$ revisiting SZO within $bin[j]$'s interval . Note that the normalization of *SLKdays* of each player[i] is the last bin ($j = 14$) also using Eq.3.

Lastly, we use the support vector machine (SVM) as the classifier for the departure prediction.

Experimental and Results

In our experiments, we analyzed a variety of k days for finding the best departure prediction. In every experiment, we explored 13 bins of game revisitations as shown in Table 2. We tested totally 10 cases: nine cases of k numbers, where $k = 2, 3, \ldots, 10$, one case excluding *SLKdays*,and the last case, where $k = 0$, i.e., we use only 13 bins of game revisitations without our new feature *SLKdays*. In summary, there are 10 cases analyzed in each of the four experiments as shown in Table 1. In each experiment, we show the comparison of accuracy for *SLKdays* in k days and the last case $(k = 0)$.

After we applied SVM in Matlab with the training and test data sets as experiments shown in Table 1, the accuracy percentage of all experiments is shown in Fig. 2. We found that accuracy trend

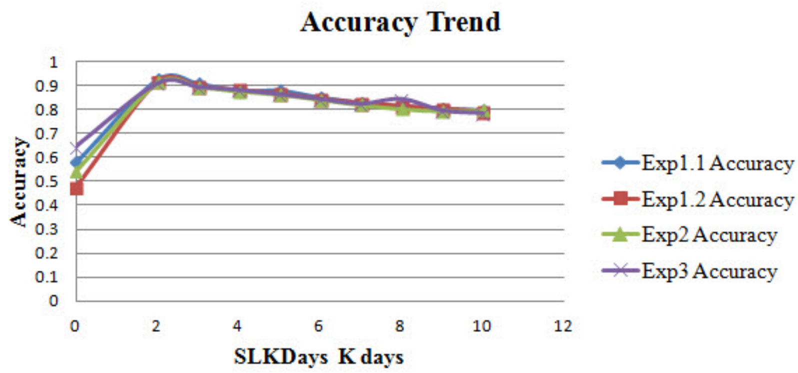

Fig. 2: Accuracy trend of *SLKdays* on different k where $k = 0$ indicates the case that uses only the *game revisitation* without our new feature *SLKdays*.

without our new feature are dropped to 47.66% - 64.67%. To compare them with our proposed feature, we included *SLKdays* from $k = 2, 3, \ldots, 10$ days and found that the highest accuracy is *SLKdays* on $k = 2$. These results can imply that the most recent information of *staytime* yields the best departure prediction. We calculate the accuracy, precision and recall of the 10 cases, whose results are shown in Table 3.

The accuracy, recall and precision of *SLKdays* when $k = 2$ is the best of all 10 cases of experiment 1.2, i.e., accuracy (91.92%) , recall (84.51%) and precision (98.22%). From the trend shown in Fig. 2, the decreasing k, the highest percentage of accuracy of the prediction is found.

Conclusion and Future Work

In this paper, we proposed departure prediction with new feature *SLKdays* when $k = 2$ days as in the $Eq.2$ with accuracy 91.92% ,recall 84.51% and precision 98.22%, appears to be the best departure prediction of all other different days studied . The high percentage of recall means that our new feature

Table 3: The accuracy, recall and precision of 10 cases of experiment 1.2.

SLKdays *k days*	Accuracy %	Recall %	Precision %
No new feature	47.66	99.80	47.54
$k = 2$	91.92	84.51	98.22
$k = 3$	89.93	79.43	97.40
$k = 4$	88.43	75.43	96.78
$k = 5$	86.98	72.24	95.39
$k = 6$	84.77	68.01	93.48
$k = 7$	82.96	66.26	90.88
$k = 8$	81.61	65.98	88.40
$k = 9$	80.59	65.70	85.93
$k = 10$	79.50	68.60	81.64

can predict the right players (up to 84.51%) who are going to quit the game. The ability to predict online game players'departure is important to both game operators and game developers in terms of current game improvement and the new design of future games. With accuracy of departure prediction approach, game operators might give special promotions to the right target players who plan to quit the game in order to keep their subscriptions and online game players'retention rate. The percentage of precision with new feature *SLKdays* is up to 98.22%. This result means that promotions given to the predicted players are worth because the predicted players (98.22%) are the right players who plan to unsubscribe the game. In future work, we will further our study using this approach combine with new features of online game players'behavioral states [3] to get better model performance in departure prediction of online game players.

References

[1] P.Y. Tarng, K.T. Chen, and P. Huang, "On Prophesying Online Gamer Departure", Network and Systems Support for Games(NetGames), 2009 8th Annual Workshop on , vol.1, no.2, pp.23-24 Nov.2009 doi: 10.1109/NETGAMES.2009. 5446225.

[2] R. Thawonmas,K. Yoshida, J. Lou,K. Chen, "Analysis of revisitations in online games", Entertainment Computing 2.4 (2011):pp 215-221.

[3] K. Savetratanakaree, K. Sookhanaphibarn, S. Intakosum,"Online Game Player's Behavior by using Hidden Markov Model", ICEAST 2013, International Conference on Engineering, Applied Sciences, and Technology 2013., PID:0144,August 21-24,2013, pp.103.

[4] K.J. Shim, J. Srivastava, "Behavioral Profiles of Character Types in EverQuest II",IEEE Conference on Computational Intelligence and Game (CIG'10), 2010, pp.186-194.

[5] L.S. Whang and G.Y. Chang, "Lifestyles of Virtual World Residents,Living in the on-line game,Lineage", Proceedings of the 2003 International Conference on Cyberworlds (CW'03), 2003, pp.18-25.

[6] K.T. Chen, P. Huang, and C.L. Lei, "Effect of Network Quality on Player Departure Behavior in Online Games", IEEE Transactions on Parallel and Distributed Systems, 2009, pp.593-606.

[7] J.K. Lou,K.T.Chen, H.J. Hsu, and C.L. Lei, "Forecasting Online Game Addictiveness", 11th Annual Workshop on Network and Systems Support for Games (NetGames), 2012, pp. 1-6.

[8] Z.H. Borbora and J. Srivastava, "User Behavior Modelling Approach for Churn Prediction in Online Games", 2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust. pp.51-60.