On Prophesying Online Gamer Departure

Pin-Yun Tarng[1], Kuan-Ta Chen[2], and Polly Huang[1]
[1]Department of Electrical Engineering, National Taiwan University
[2]Institute of Information Science, Academia Sinica

I. Introduction

The current business model of online game companies usually depends on the sale of virtual items or monthly subscriptions in which gamers must pay for credits to continue their adventures in the virtual world. From the perspective of game operators, being able to predict *how long people will stay in the game* is crucial, since it directly affects their revenue. The duration for which a player will remain active in a game should correlate with the extent of his involvement in the virtual world. More often than not, this can be inferred from the player's external behavior, such as how quickly his avatar advances to new levels and how much time he spends in the game every day.

This study is an extension of our previous game hour analysis [1], in which we analyzed what time players enter the game's virtual world and how long they stay in the game, and investigated whether a player's future game hours can be predicted with his observed behavior. Our goal in this study, then, is to provide a practical scheme for predicting player unsubscription that takes a player's game hours as input and determines whether or not he will renew an expiring subscription. Our rationale is that, *if we can predict the departure of a player before he actually quits a game, the game operator can take remedial measures to prevent it from happening and improve the game along the way based on the feedback provided by such a player.*

Our study is based on real-life traces collected from Shen-Zhou Online [2], a mid-scale commercial MMORPG in Taiwan sustaining at any moment thousands of players online. The traces we acquired from the operator contain the playing histories of $162,980$ accounts over a span of four years. To play on the safe side, we only use the un-curtailed traces of $20,514$ accounts. We propose a classification method to first identify the involvement pattern of the players in game along time with 90% accuracy, and then devise a prediction model that detects whether gamers are leaving the game in the near future. For "hardcore" players who subscribe to the game for more than a year, the prediction reaches 85% accuracy. Furthermore, we analyze the generalizability of our scheme by collecting $2,132$ questionnaires from real-life gamers, and find that players of different MMORPGs have similar playing patterns towards their unsubscription. We also apply our methods to World of Warcraft avatar traces, and find that we can detect with 80% accuracy whether gamers are discarding their avatars in the near future.

II. Classification of Online Gamers

An intuitive categorization of unsubscribing players would be dividing them into "fade-out" and "sudden-out" groups, the former featuring *ever-decreasing daily playtime* and the latter exhibiting no noticeable tendency in daily playtime or login frequency. We figure that if we focus on the fade-out group of gamers, the prediction of departure will attain a higher degree of success. Consequently, a scientific method is needed to separate the "predictables" from the "unpredictables."

We base our automated classification method on gamers' *average daily playtime* and *playing density*. First, we randomly choose $2,000$ gamers from our traces and classify them with the human eye. Among all the sample gamers, 613 are fade-out and $1,387$ are sudden-out. Second, we divide each gamer's history into k periods of equal length, and evaluate the average daily playtime and playing density in each period. The playing density is the occurrence of a gamer's playing days within all available days. For example, if a gamer has at least logged in the game once for 15 days in June, his playing density in June will be 0.5.

We use the support vector machine (SVM) as the classifier. The traces of the $2,000$ sample gamers, along with their k-period features and predetermined categories, serve as the training data set. To find the optimal value of k, we experiment within the range of $[2, 20]$, ten-fold cross-validating for each value (as with every accuracy estimate plotted hereafter), and conclude that 10-period features yield the best classification. Of the $18,514$ remnants, $5,503$ (29.7%) are deemed fade-out while $13,011$ (70.3%) are dubbed as sudden-out.

In real-life prediction, however, only *incomplete data* is available, and the approximate time of a gamer's final login is to be predicted. To verify the predictivity of our classification method, the traces of the $2,000$ sample gamers, their last n days cut off ($n \in [3, 60]$), are fed into the SVM model as the training data set along with their re-computed 10-period features. The result is shown in Figure 1, where lines are drawn for various subscription lengths. It can be seen that gamers with longer subscription lengths tend to be more resilient to the cutting, as 82% of the hardcore players (subscribing for more than one year) are correctly classified even with their last month truncated.

III. Model for Predicting Gamer Unsubscription

To predict whether a gamer is leaving in d days ($n \in [3, 60]$), we first construct a SVM for each d. Similar to our classification method, we use the traces of the $2,000$ sample gamers as the training data set. For each sample gamer, we
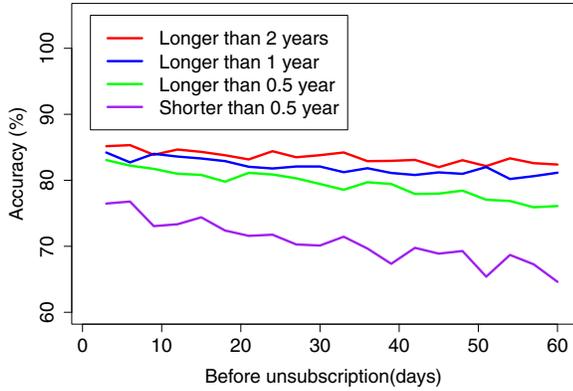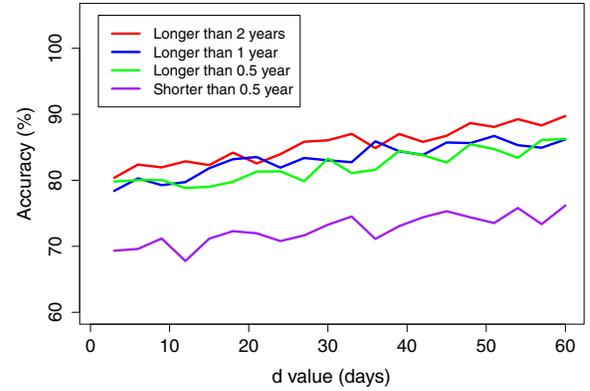
Fig. 1. Predictivity of our classification method
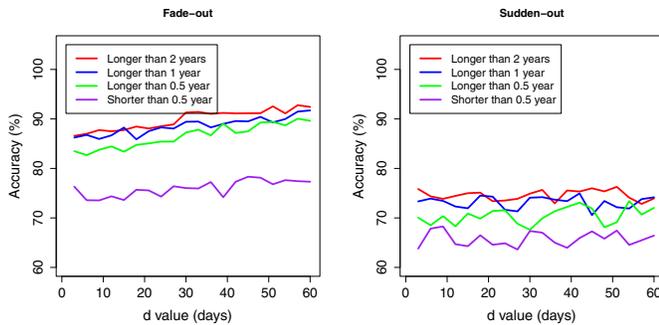

Fig. 2. Unsubscription prediction accuracy

Fig. 3. Accuracy of our final prediction scheme

assign the *prediction point* at d days before his quitting. Two random observation windows, counting from the gamer's first login day, are derived for each gamer. One of the observation window, dubbed as *leaving*, contains the prediction point, since it is implied that after the last day of the window, the gamer will unsubscribe within d days. The other window, tagged as *staying*, does not contain the prediction point, so it is implied that the gamer will stay in the game for at least d days after the last day of the window. The 10-period features are extracted from each window and fed to the SVM along with their corresponding window type.

We applied the above procedure on both fade-out and sudden-out groups to give Figure 2. It can be seen that our method is especially useful for predicting hardcore fade-out gamers, reaching an accuracy as high as 90%. On the other hand, the difficulty of predicting sudden-out gamers lies in their irregular behavior, which may or may not be due to their own social activities.

IV. THE COMPLETE SCHEME

The complete unsubscription prediction scheme is the combination of the classification method and the prediction model and takes a player's incomplete trace as input. A gamer categorized as sudden-out is difficult to predict, so we opt

to leave him be, not applying the prediction model on him. As a result, the complete scheme will give a *three-way output*:

- The player is of sudden-out pattern and just unpredictable;
- The player is staying in the game for the time being;
- The player is leaving within a specific number of days.

The ten-fold cross-validated prediction accuracy of our scheme, shown in Figure 3, can be seen as a logical aggregation of Figures 1 and 2. We can see that the accuracy for predicting whether a gamer is going to leave in a month is around 82% for gamers who subscribed for more than one year, and 85% for gamers who subscribed for more than two years. The results indicate that, for hardcore gamers, it is feasible to use our scheme to predict whether they are quitting from the game in the near future.

The probability of wrongly identifying a leaving hardcore player as staying is only about 10%. Put it the other way, we can detect 90% leaving gamers correctly with our final prediction scheme, if we do not count the false unpredictables in. Nevertheless, a certain amount of resource inefficiency is inevitable, as there is a 15% false leaving rate.

V. CONCLUSION

The ability to predict a gamer's departure is coveted by the MMORPG industry as it allows the game operators to target their resources on keeping subscribers motivated and to benefit from these loyal customers, not only financially, but also in terms of the improvement of current games and the design of future ones. To this end, we hope that our scheme will prove helpful to operators, as well as gamers who may enjoy a better gaming environment because of it.

REFERENCES

[1] P.-Y. Tarng, K.-T. Chen, and P. Huang, "An analysis of WoW players' game hours," in *NetGames '08: Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games*. ACM, 2008, pp. 47–52.

[2] UserJoy Technology Co., Ltd., "ShenZhou Online," http://www.ewsoft.com.tw/index.php.