Churn prediction in MMORPGs: A social influence based approach

Conference Paper · January 2009

DOI: 10.1109/CSE.2009.80 · Source: DBLP

Jaideep Srivastava
University of Minnesota Twin Cities

Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 EECS Building
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 09-013

Churn Prediction in MMORPGs: A Social Influence Based Approach

Jaya Kawale, Aditya Pal, and Jaideep Srivastava

May 20, 2009

Churn Prediction in MMORPGs: A Social Influence Based Approach

Jaya Kawale
Dept of Computer Science & Engg
University Of Minnesota
Minneapolis, MN 55455
Email: kawale@cs.umn.edu

Aditya Pal
Dept of Computer Science & Engg
University Of Minnesota
Minneapolis, MN 55455
Email: apal@cs.umn.edu

Jaideep Srivastava
Dept of Computer Science & Engg
University Of Minnesota
Minneapolis, MN 55455
Email: srivats@cs.umn.edu

***Abstract*—Massively Multiplayer Online Role Playing Games (MMORPGs) are computer based games in which players interact with one another in the virtual world. Worldwide revenues for MMORPGs have seen amazing growth in last few years and it is more than a 2 billion dollars industry as per current estimates. Huge amount of revenue potential has attracted several gaming companies to launch online role playing games. One of the major problems these companies suffer apart from fierce competition is erosion of their customer base. Churn is a big problem for the gaming companies as churners impact negatively in the "word-of-mouth" reports for potential and existing customers leading to further erosion of user base.**

We study the problem of player churn in the popular MMORPG EverQuest II. The problem of churn prediction has been studied extensively in the past in various domains and social network analysis has recently been applied to the problem to understand the effects of the strength of social ties and the structure and dynamics of a social network in churn. In this paper, we propose a churn prediction model based on examining social influence among players and their personal engagement in the game. We hypothesize that social influence is a vector quantity, with components negative influence and positive influence. We propose a modified diffusion model to propagate the influence vector in the player's network which represents the social influence on the player from his network. We measure a player's personal engagement based on his activity patterns and use it in the modified diffusion model and churn prediction. Our method for churn prediction which combines social influence and player engagement factors has shown to improve prediction accuracy significantly for our dataset as compared to prediction using the conventional diffusion model or the player engagement factor, thus validating our hypothesis that combination of both these factors could lead to a more accurate churn prediction.

I. Introduction

Massively Multi-player Online Role Playing Games (MMORPGs) games is a genre of computer games in which players can assume a role or a fantasy character and interact with one another in a virtual game world. One of the distinguishing factors of role playing games is a persistent world and a never ending quest for exploration. MMORPGs have achieved significant growth in the past few years and according to [1] the total number of active subscriptions in 2008 was over 16 million. The revenue generation from MMORPGs was $2 billions in 2006 and is expected to explode to a staggering $11.5 billion by 2011 [2]. There are a number of popular game providers battling for the market share and some of the more prominent ones are World of Warcraft[1], Lineage[2], Final Fantasy XI[3], Eve Online[4] and EverQuest II[5].

Churn analysis assumes importance for game publishers as it helps them understand the several factors leading to users leaving the game. It holds the key in understanding the behavior of players and the various factors that influence players to leave the game ranging from personal commitment, competing products, shifting interest to social influence. The target audience of MMORPGs is very large. World of Warcraft, which is currently the most popular MMORPG has more than 10 million paid subscription. Hence it is critical for the game publishers to identify potential churners and the reasons leading to churning, in order to sustain and grow their consumer base. By understanding such factors, the game provider can offer incentives to likely churners in order to keep them interested in the game. Additionally, acquiring new customers can be much more expensive than retaining the old ones. Therefore it is very important for service providers to identify customers who are likely to churn and plan marketing strategies for this subset of customers.

Churn Analysis has been widely studied several domains. A variety of churn analysis techniques have been developed as a solution to identifying the subset of customers who are likely to churn. In this paper we examine churn among players in the virtual world of EverQuest II which is a popular MMORPG game. MMORPGs have been studied widely in the recent past for analyzing player characteristics [3], [4], [5]. The virtual worlds of the online games provide a good indicator of understanding human behavior in the real world and this has led to increased interest from fields like social sciences, psychology and communications.

[1]http://www.worldofwarcraft.com/index.xml
[2]http://www.lineage.com/
[3]http://www.playonline.com/ff11us/index.shtml
[4]http://www.eveonline.com/
[5]http://everquest2.station.sony.com/

A. Challenges Involved

Churn prediction in the EverQuest II data posed several challenges. To study the impact of interaction among players, we needed to build a social network out of the EverQuest II data. Players in EverQuest II can play together in groups. We used the grouping information between players to build a social network. A good amount of players in EverQuest II do not group together and play solo. We have not considered them for our analysis as we do not have any indicator of social influence among these players using our algorithm. To assign weights to the edge in the graph we consider the cumulative experience points shared between players. Another aspect in the data is that churn is a rare phenomenon. The percentage of churners in one month from among a group of players who team together is less than 5%. It thus falls into a rare class classification problem. Another challenge we faced was in measuring the engagement of a player in the game. We used the session lengths of game sessions as an indicator of the engagement of the player in the game. Extracting session lengths from the raw data was another challenge. Our raw data consisted of timestamp values when a user got experience points. We stitched together these values into a session using standard techniques for session extraction in Data Mining [6].

B. Our Contribution

We propose a churn prediction model that takes into account social influence of players and their personal engagement in the game. In our model we consider that player influence is a vector quantity (instead of scalar, as used in typical diffusion models) with two components: 1) Negative influence, 2) Positive influence. In a typical diffusion model, a player would hold a scalar influence value which could be difference of positive and negative influence. In our model we consider player influence vector to have the two components which signify how strongly a player feels for the game and how strongly a player feels against the game. We argue that an influence vector helps in modeling the real world more closely.

Based on players engagement in the game and how strongly he feels for or against the game we propose a modified diffusion model (MDM) in which influence is propagated in the network. In our model, a player can convert between his negative and positive influence based on his game engagement thus increasing one type of influence in the network but conserving the total influence. The MDM takes into account the players game engagement and social influence from neighbors and their interconnections in influence propagation.

To model player engagement we use players' activity data such as session time and session length. We tried various regression and curve fitting measures to fit the user's activity data. The beta function models the churners and non-churners engagement with the least amount of error. The shape parameters, α and β of beta function are the parameters which are used to calculate the player's game engagement at any given point. The shape parameters are also used for classification of churners and non-churners for churn prediction.

II. Literature Review

Churn Prediction is an important problem studied across several areas like banking, insurance, retailing, telecommunications, etc. A wide variety of techniques have been applied to predict churn in the diverse applications.

A decision tree based approach has been most widely used in the churn prediction. Classification and decision trees have been used to determine either the class label or the churn risk. CHAMP [7] (Churn Analysis, Modeling, and Prediction) predicts churn factors for cellular phone customers using a decision tree model. Decision tree approach to predict churn using complaints data has been found to perform better in comparison with neural networks and regression [8]. Decision trees have been used to determine classification rules from which the most significant variables could be identified [9].

Another popular technique used for churn prediction is logistic regression [10], [11]. Logistic regression models are used for predicting the probability of occurrence of an event by fitting data to a logistic curve.

Latent semantic analysis has been used in predicting policy churn in insurance industries [12]. The authors conclude that prediction models that do not consider timestamp of data into consideration do not perform too well. They have converted timestamp information from the data set into features using the tf-idf notation. The paper indicates that the accuracy of the model was improved after incorporating the timestamp features.

Survival analysis is another class of statistical techniques that is used to model time to event data. It is useful to provide answers to questions such as what fraction of a population will survive after a time interval. Survival analysis techniques have been used to predict churn in the telecommunications industry [13]. It has also been used in predicting switching behavior in banking services [14].

Ordinal regression is a technique in which the response variable comes from an ordered set. [15] has proposed ordinal regression as an alternative to using survival analysis for churn prediction. Customer tenure is being modeled as an ordinal response variable to predict the time to churn. [16] proposes an approach where they have implemented an ensemble involving LR, decision trees, Random Forests, radial basis function network and SVM. A simple majority voting method is used to determine the output of majority of the classifiers and they use a recent technique called SMOTE to over sample the churn class.

Some of the recent work in this area has used Support Vector Machines [17], [18], [19] and Random forests [18] as classifiers. [18] has compared the use of SVM, Random forests and LR in predicting churn in newspaper services. It has been shown that Random Forests perform better than SVM.

Social network analysis is an important emerging area to understand the effects of the strength of ties and network characteristics on an individual behavior. Historically there have been trying to study and understand complex large scale networks. The random graph models [20], [21], the small world model[22] and the scale free model of networks[23] are some of the well studied models of large scale networks. Diffusion or the spread of a phenomenon or an idea in a network has been studied for decades in fields like sociology, biology, economics, epidemiology, marketing in understanding concepts like the spread of an infectious disease, diffusion of innovations, rumor spreading, emergence of fads, etc. Linear threshold model [24] and the independent cascade model [25], [26] are two prominent models used to model the diffusion in networks.

A new approach to predict customer churn in mobile networks using Social Network Analysis is presented in [27]. Network analysis has been used in the past to identify influentials in a network for targeting individuals in marketing campaigns [28]. [27] uses an energy propagation model to spread influence in the telecom social network. The use the frequency of calls between two people to define the strength of ties between two people. Using the underlying graph, they initiate a diffusion process where the churners are taken as a seed and they spread influence in the network. The amount of influence every neighbor of a churner gets depends upon his relative strength of the tie in the neighborhood. Once a non churner node accumulates sufficient amount of energy it is labeled as a churner and it starts propagating energy on its own. The authors show significant amount of increase in lift curves over other methods.

Understanding churn behavior in MMORPGs has been studied in [29]. The authors study the long term behavior of players in the game Eve Online so as to help game publishers provision sufficient resources to support the game. They aggregate player characteristics and model player behavior to predict workloads and analyze the effects of game updates.

III. Dataset description and analysis

A. Dataset Description

The dataset we used is from Sony Online Corporation's popular game EverQuest II. EverQuest II provides players a platform to don a fantasy character and explore the virtual world and advance through the challenges the game environment and other players provide in the form of quests and monster raids. Players can group together with other players and participate in activities like completing quests, exploring the world, killing monsters and gaining treasures and experience. In order to accomplish a task, the players can team up with others who might aid them in this exploration. In other instances, players might be on a quest to create items (e.g. a sword, a tunic, or a meat pie) commensurate with their skill level. In this case, they need to exploit resources possessed by other players or the environment.

TABLE I
Graph Characteristics

Characteristics	**Value**
Number of Nodes	6213
Number of Edges	153983
Average degree of Nodes	24.78
Average experience points shared	210897

The dataset consists of complete experience data of all the players for the month of Jan to Aug 2006. Experience data contains the time when users completed a task, quests played by users, points received by the users, etc. We also had data containing the list of churners for the month of August, September, and October 2006. We use the dataset to build the network graph and model player engagement in the game.

B. Graph Characteristics

In EQ2 several players can come together to play a quest. Since a quest might require various kinds of skills for successful completion, and since one player cannot possess all of the skills, it requires players to devise strategies and plan in advance for the quest. Players can communicate with each other using integrated voice chat, built-in mail system, global chat channels and global marketplace. We build the gaming network of players based on the games and interactions done by players with each others. When two players play a game together, we say that an edge exists between them. The weight of the edge indicates the strength of the tie between the two players, which in this case is the total number of points shared between the two players over the given period of time.

The dataset consists of details of a player's quests and the points shared by players in the quest. We extract quest information for each and every player and identify which players played together for a quest. If the quest is played by only one player we do not consider it. In general many players in EQ2 play solo. Table I describes the graph characteristics of the dataset. The degree distribution of players in the graph follows a power law and is as shown in the fig.1. Fig.2 shows the total number of active players in a given month over the period of time. As can be seen in fig. 2 that the total number of active players in month of August is much larger than 6213, but only 6213 of these players grouped together in a quest.

C. Churner Analysis

The dataset also contains the list of EQ2 players who unsubscribed from the game in the month of August,

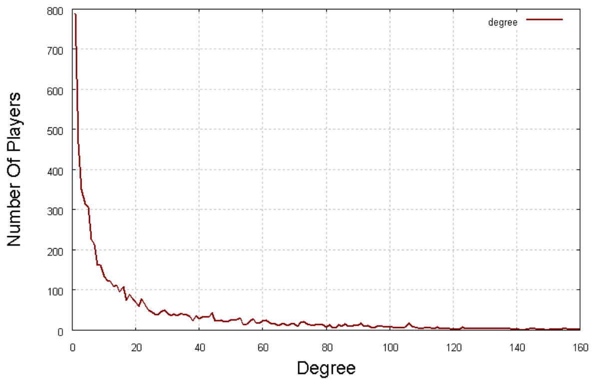

Fig. 1. Degree distribution of all players

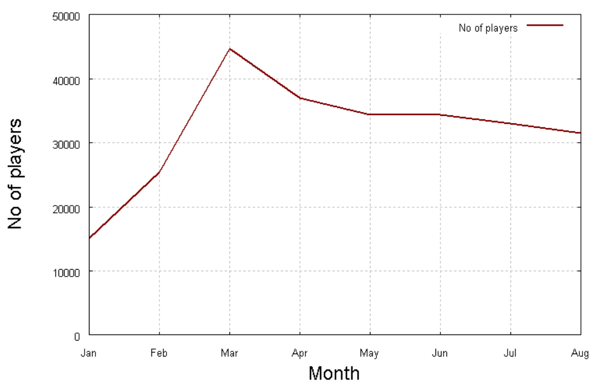

Fig. 2. Total number of active players

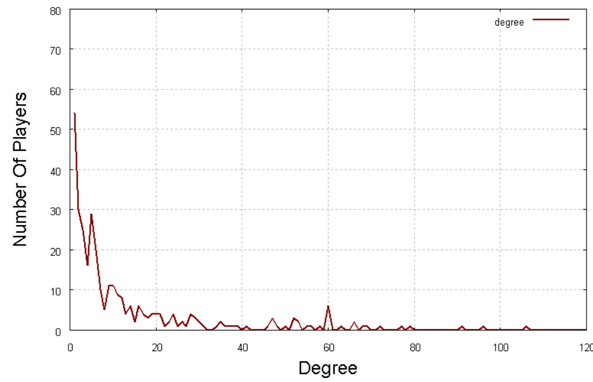

Fig. 3. Degree distribution of churners

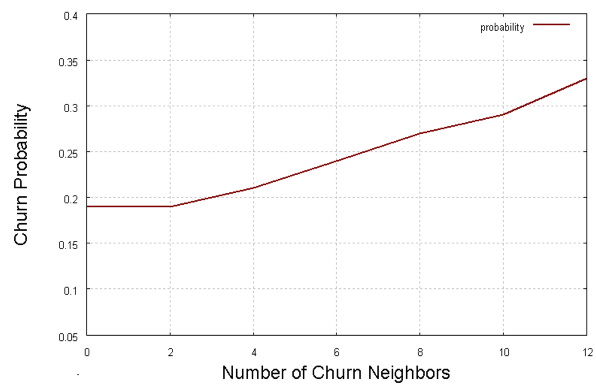

Fig. 4. Probability to Churn in subsequent months

September and October. Table II shows the number of churners for the given months. Fig.3 shows the degree distribution of churners, which follows the power law. Fig. 2 shows the total number of active players in a given month.

Network based analysis for churn prediction [27] has the basis that ties of an individual with other individuals in the network are good predictors of their interest in the games. They have successfully shown improvement in churn prediction using the energy propagation model as described in [30]. To understand the social effects of churn behavior in players, we examine the probability of nodes to churn in a network given that k of their immediate neighbors has churned. Fig.4 shows the probability of players from the August graph to churn in the subsequent months given that k of his neighbors have churned. As seen in the figure the probability of churn increases with increase in the number of churner neighbors. This explains that churn behavior has a social component.

TABLE II
NUMBER OF CHURNERS

Month	Churners
August	334
September	308
October	380

D. Player Engagement

Player engagement can be defined as the time spent by the player in the game. We measure player engagement using the session time and length. We hypothesize that the player engagement over a long span of time can be used to predict churn behavior. Fig. 5 shows the distribution of average session lengths for non churners. Fig. 6 shows the distribution of average session length of churners. It can be seen that the churners show a decreasing average session length as compared to non churners. This could partly explains their waning interest in the game.

We further studied the individual plots of the August churners over a long period of time and tried to fit Beta function to it. Fig. 7 shows a sample distribution fit for a churner. A beta function is defined by two shape parameters α and β which determine the shape of the distribution. Table III shows the distribution of alpha and beta values of the Beta function. From the table it can be seen that for nearly 2/3 rd of the Churners the distribution was positively skewed. In other words it means that for nearly 2/3rd of the Churners, the session lengths were longer in the initial periods than in the later periods of time.

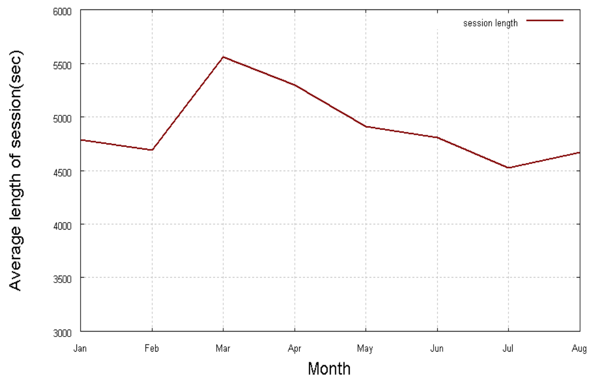

Fig. 5. Average session length of non churners

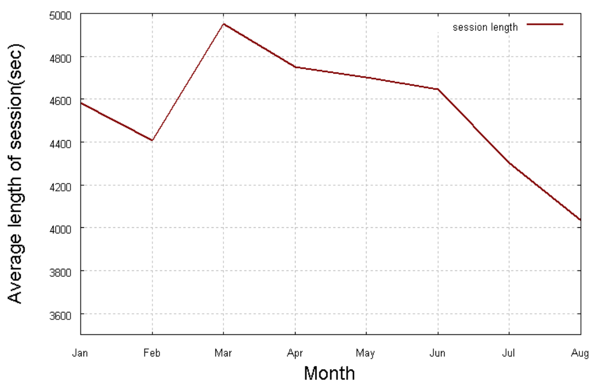

Fig. 6. Average session length of Aug churners

TABLE III
BETA FUNCTION TO FIT THE PLOT OF CHURNERS

Alpha vs Beta	Percentage	General Shape
$\alpha = \beta$	0.02%	Symmetrical
$\alpha < \beta$	75%	Positively skewed
$\alpha > \beta$	25%	Negatively skewed

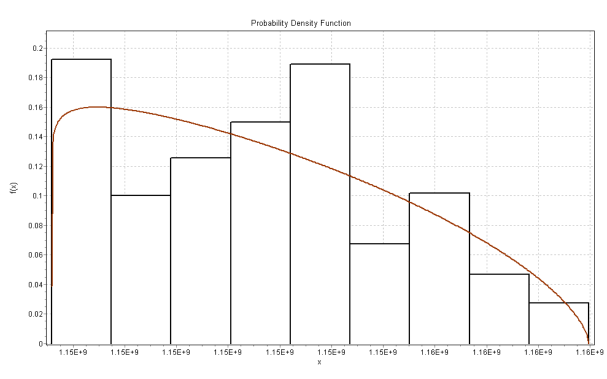

Fig. 7. Positively skewed player engagement of a churner

IV. METHODOLOGY

In the previous section we identified two main aspects that are predominant in churners in our game dataset 1) decrease in player engagement of churners over time until they finally churn, 2) increase in churn propensity of players with the increase in the number of churner neighbors. Based on these two observations we hypothesize that the social influence and player engagement combined together can be good measures for churner prediction. We define churn prediction function, P that takes two factors to predict whether a player would churn in subsequent months or not.

$$\text{Churn Prediction} = P(\text{Player engagement}, \text{Social influence}) \tag{1}$$

A. Preliminary Definitions

We represent the player network in the form of a weighted undirected graph $G(V, E)$ where $V(G)$ is the set of vertices corresponding to the players and $E(G)$ is the set of edges between the vertices, such that two vertices in $V(G)$ share an edge if the corresponding players have played a game together and the weight of the edge equals the total number of points shared between the two players. Let $N(x) = \{y \in V(G) : (x, y) \in E(G)\}$ be the set of neighbors of x. Let e_{xy} represent the weight of the edge between x and y, e_x be the sum of the weight of all edges from x.

Let Φ be the function that models player's engagement, such that $\Phi_x(t)$ represents x^{th} player's engagement in the game at time t. We can define slope $S_x(t)$ of player x at time t as follows:

$$S_x(t) = \frac{d\Phi_x(t)}{dt} \approx \frac{\Phi_x(t + \delta t) - \Phi_x(t - \delta t)}{\delta t} \tag{2}$$

We are more interested in the sign of the slope rather than the magnitude of it, hence the approximation. The main reason for not considering the magnitude of slope is that it would require us to calculate weighted average of slope and very large values of slope could skew the weighted values and the model would become biased towards some localized points. Also, to keep the model simple and computationally inexpensive, we consider just the sign of slope as an indicator of player's interest in the game.

B. Modified Diffusion Model

In our diffusion model, we consider every node to have an influence vector containing two components, namely, positive influence and negative influence. Negative influence represents how much a user is influenced against the game whereas positive influence represents how much a user is influenced in favor of the game. Two valued influence vector helps us in modeling the real world more accurately. The intuitive argument in favor of an influence vector is that when players communicate they share their good/bad experiences with each other. So a player acquires both good and bad opinion about the game from others. A player strongly interested in the game would subdue the bad things he hears and remember the good things and vice-versa. Additionally, when he tries to influence other players, he would maintain the positive influence on him and try to spread positive influence on others. In standard diffusion model, a node has to give some of his influence to others, thereby leaving less influence with a node believing strongly. In our diffusion model, node preserves his positive influence as it is (he might gain more from his neighbors) and converts his negative influence to positive and spreads on his neighbors. So in a way a user's positive influence is conserved and he is able to positively influence other nodes. We observe that the standard diffusion model[30] does not show very encouraging results on our dataset, whereas the modified diffusion model shows much better results.

The spread factor, γ is the portion of influence a user transfers to his network. A user with increasing game engagement would convert some of his negative influence to positive influence and spread γ proportion of converted influence amongst his neighbors in proportion with strength of tie with the neighbors. The total influence of the graph remains constant, whereas the positive influence or negative influence values change.

We initialize churners with negative influence $ni = 1$ and positive influence $pi = 0$ and non-churners with $ni = p$ and $pi = 1 - p$ initially. The following algorithm illustrates the propagation step for a given node x at time t:

---

if $S_x(t) < 0$ **then**
  $\{$convert positive energy to negative and propagate$\}$
  if $ni(x) < pi(x)$ **then**
    $i = ni(x)$
  else
    $i = pi(x)$
  end if
  $pi(x) = pi(x) - i * (1 - \gamma)$

  $\{$spread $i * \gamma$ to neighbors$\}$
  for $y \in N(x)$ **do**
    $ni(y) = ni(y) + i * \gamma * \frac{e_{xy}}{e_x}$
  end for
end if
$\{$if $S_x(t) > 0$, interchange $ni(.)$ with $pi(.)$ in above step$\}$
$\{$if $S_x(t) = 0$, do not consider this node$\}$

Modified diffusion model is run on the player network of August 2006. The algorithm starts with an initial time of 1st August 2006 taking discrete steps of 1 hour and finishes on 31st August 2006. At any given time, the algorithm iteratively runs on all graph nodes and applies the propagation step as described above. After the algorithm stops execution we store the influence vector, which represents the influence on this user from the player network. The influence vector is used for predicting the churners and non-churners.

C. Churn Prediction

Further, we use the shape parameters of a player engagement curve in prediction. The beta function fits the player engagement best as per our analysis and the α and β values of beta function constitute the shape parameters of the player engagement curve. We took shape parameters and the influence vector computed using the modified diffusion model and ran several clustering algorithms in order to classify the nodes as churners and non churners. The next section compares the accuracy of various clustering algorithms for our dataset.

V. Experiment and Results

The main premise of our approach is that churn prediction can be captured effectively using engagement of a player in the game and the social influence on a player by his gaming buddies. To validate our approach we ran experiments with three different models:

1) Simple Diffusion Model.
2) Classification based on Network and Player engagement.
3) Modified Diffusion Model.

A. Simple Diffusion Model

The simple diffusion model is the same as defined in [27]. The parameters of the simple diffusion model are as shown in the Table IV. The prediction accuracy for the simple diffusion model is shown in Table VI.

TABLE IV
Parameter Values in Simple Diffusion Model

Parameter	Values
Initial energy for Churners	e = 1
Initial energy for non Churners	e = 0
Spreading Factor	0.7

B. Classification based on Player Engagement and Network Statistics

To study the effect of player engagement and his raw network characteristics on churn prediction, we built another model which consisted of the feature set as defined in Table V. We used WEKA [31] for classification of churners and non-churners based on our training and test set as described in Table X. The shape parameters are the α and β parameters of the beta function that captures the player engagement the best. Fig. 8 show the plot of α and β for all the players. It indicates that churners are concentrated in a small region whereas non-churners are spread out. The summary of the results using various classifiers is shown in Table VII. The improvement in prediction accuracy over the simple diffusion model shows that engagement features and raw network characteristics are better predictor of churners over our game dataset.

TABLE VI
SIMPLE DIFFUSION MODEL

Method	Precision	Recall	Correct Predicted	Total Predicted
Simple Diffusion Model	17.9	11.2	77	430

TABLE VII
NETWORK FEATURE SET WITH DIFFERENT CLASSIFIERS

Method	Precision	Recall	Correct Predicted	Total Predicted
AdaBoostM1	42.8	14.7	101	236
ADTree	42.9	12.2	84	196
JRip	43.2	18.8	129	299
J48	47.3	11.3	78	165
NaiveBayes	46.8	12.6	87	186

TABLE VIII
MODIFIED DIFFUSION MODEL WITH DIFFERENT CLASSIFIERS

Method	Precision	Recall	Correct Predicted	Total Predicted
AdaBoostM1	50.1	29.8	205	409
ADTree	46.5	41.3	284	611
JRip	43.1	18.8	129	299
J48	38.5	21.5	148	384
NaiveBayes	49.7	23.3	160	322

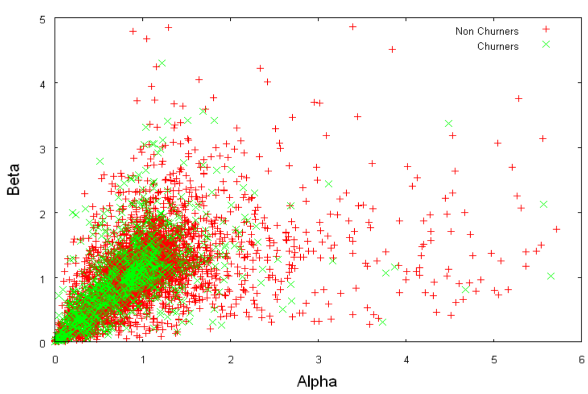

Fig. 8. Alpha vs Beta plot of all players

TABLE V
ENGAGEMENT AND RAW NETWORK CHARACTERISTICS

Feature Variable	Characteristic
Alpha	Engagement
Beta	Engagement
Number of neighbors	Network
Number of churner neighbors	Network

C. Modified Diffusion Model

Finally, we run the Modified Diffusion Model using the player graph for the month of Aug 06. We propagate the influence vector for players based on their engagement curve slope over the time period of game play in the month of August. Table IX shows the parameters value that we used in our modified diffusion model.

The shape parameters to fit the session length curve were obtained by fitting beta function on the session plots from Jan 06 to Aug 06. We ran several classifiers on the final influence vector and the shape parameters. Table X shows the description of test and training data as used by all the classifiers. Table VIII shows the precision and recall values given by the various classification techniques.

TABLE IX
PARAMETER VALUES USED IN MDM

Parameter	Values
Initial energy for Churners	pe = 0, ne = 1
Initial energy for non Churners	pe = 0.8, ne = 0.2
Spreading Factor	0.7

TABLE X
TRAINING AND TESTING DATASET

Dataset	Size	Number of Churners
Training	4026	334 (Aug)
Testing	2187	688 (Sept + Oct)

VI. CONCLUSION

The results show that MDM outperformed both Simple Diffusion Model and Network and Engagement Feature based classification in our dataset. A simple diffusion model is able to capture social influence among game players and a network and engagement feature based classification is able to capture player engagement features. An MDM scheme is able to effectively combine social influence and player engagement and provide a significant improvement in prediction accuracy which is broadly consistent with our hypothesis.

We however have not compared our approach with classification based on features that might not capture player engagement and network features. Such an approach could possibly give more prediction accuracy, however that was not the main focus of the paper.

VII. Future Work

For our future work we intend to do a deeper analysis of variables useful in capturing the engagement of a player in the game. For this paper we used the average length of sessions as an indicator of engagement in game play. We could further use other variables in the EverQuest II data like quests completed or health of a character or achievement points gained as a feature and this would provide a deeper insight into engagement of a player.

Another aspect we want to study is the impact of the current engagement of a player on the group and the impact of the engagement of the group on a player. If all the players in the group tend to move towards the mean engagement in the game, this would imply that a player's engagement in the game is also a function of the group's engagement in the game. Influence propagation can possibly accommodate this.

Classical models of Churn prediction provide a RFM (Recency Frequency and Money) analysis of Churn. In future, we will work on analyzing player engagement from the RFM perspective. The recency and frequency values in our case could possibly be defined by the recency and frequency of sessions and the money value can be defined by the amount of time a player puts in the game play. Another aspect we want to incorporate in the future is churn prediction at a time. This could possibly involve exploring the applications of queuing theory in our problem.

Acknowledgements

The research reported herein was supported by the National Science Foundation via award number IIS-0729505, and the Army Research Institute via award number W91WAW-08-C-0106. The data used for this research was provided by the SONY corporation. We gratefully acknowledge all our sponsors. The findings presented do not in any way represent, either directly or through implication, the policies of these organizations.

References

[1] "Mmog chart http://www.mmogchart.com/."

[2] "Ars technica http://arstechnica.com/gaming/news/2007/09/report-mmorpgs-revenues-to-explode-over-next-few-years.ars."

[3] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, ""alone together?": exploring the social dynamics of massively multiplayer online games," in *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, (New York, NY, USA), pp. 407–416, ACM, 2006.

[4] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, "The life and death of online gaming communities: a look at guilds in world of warcraft," in *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, (New York, NY, USA), pp. 839–848, ACM, 2007.

[5] D. Pittman and C. GauthierDickey, "A measurement study of virtual populations in massively multiplayer online games," in *NetGames '07: Proceedings of the 6th ACM SIGCOMM workshop on Network and system support for games*, (New York, NY, USA), pp. 25–30, ACM, 2007.

[6] R. Cooley, B. Mobasher, and J. Srivastava, "Data preparation for mining world wide web browsing patterns," *Knowledge and Information Systems*, vol. 1, pp. 5–32, 1999.

[7] P. Datta, B. Masand, D. R. Mani, and B. Li, "Automated cellular modeling and prediction on a large scale," *Artif. Intell. Rev.*, vol. 14, no. 6, pp. 485–502, 2000.

[8] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, "Churn prediction using complaints data," in *Proceedings Of World Academy Of Science, Engineering and Technology*.

[9] K. Ng and H. Liu, "Customer retention via data mining," *Artif. Intell. Rev.*, vol. 14, no. 6, pp. 569–590, 2000.

[10] W. Buckinx and D. V. D. Poel, "Customer base analysis: Partial defection of behaviorally-loyal clients in a non-contractual fmcg retail setting," Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium 03/178, Ghent University, Faculty of Economics and Business Administration, May 2003.

[11] D. L. . B. S. E. Jones, Michael A. ; Mothersbaugh, "Switching barriers and repurchase intentions in services," *Journal of Retailing*, 2000.

[12] K. Morik and H. Köpcke, "Analysing customer churn in insurance data: a case study," in *PKDD '04: Proceedings of the 8th European Conference on Principles and Practice of Knowledge Discovery in Databases*, (New York, NY, USA), pp. 325–336, Springer-Verlag New York, Inc., 2004.

[13] J. Lu, "Predicting customer churn in the telecommunications industry an application of survival analysis modeling using sas,"

[14] G. I. Maria Mavri, "Customer switching behaviour in greek banking services using survival analysis," *Managerial Finance*, vol. 34, pp. 186 – 197, 2008.

[15] R. K. Gopal and S. K. Meher, "Customer churn time prediction in mobile telecommunication industry using ordinal regression," in *PAKDD*, pp. 884–889, 2008.

[16] D. A. Kumar and V. Ravi, "Predicting credit card customer churn in banks using data mining," *Int. J. Data Anal. Tech. Strateg.*, vol. 1, no. 1, pp. 4–28, 2008.

[17] "Customer churn prediction using improved one-class support vector machine," vol. 3584/2005, pp. 300–306.

[18] V. Coussement K, "Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques," vol. 34(1), pp. 313–27, 2008.

[19] T.-S. Lee, C.-C. Chiu, Y.-C. Chou, and C.-J. Lu, "Mining the customer credit using classification and regression tree and multivariate adaptive regression splines," *Computational Statistics and Data Analysis*, vol. 50, pp. 1113–1130, February 2006.

[20] R. A. Erdos, Paul., "On random graphs. 1.," *Publicationes Mathematicae*, 1959.

[21] R. A. Erdos, Paul., "The evolution of random graphs," *Magyar Tud. Akad. Mat. Kutat Int. Kzl*, 1960.

[22] S. Watts, D.J.; Strogatz, "Collective dynamics of 'small-world' networks.," *Nature*, vol. 393 (6684).

[23] A.-L. B. R. Albert, H. Jeong, "Dynamics of complex systems: Scaling laws for the period of boolean networks," *Physical Review Letters*, vol. 84, p. 5660, 2000.

[24] M. Granovetter, "Threshold models of collective behavior," *American Journal of Sociology*, vol. 83.

[25] E. M. J. Goldenberg, B. Libai, "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing Letters*, vol. 12:3.

[26] E. M. J. Goldenberg, B. Libai, "Using complex systems analysis to advance marketing theory development," *Academy of Marketing Science Review*, 2001.

[27] K. Dasgupta, R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjea, A. A. Nanavati, and A. Joshi, "Social ties and their relevance to churn in mobile telecom networks," in *EDBT '08: Proceedings of the 11th international conference on Extending database technology*, (New York, NY, USA), pp. 668–677, ACM, 2008.

[28] S. Hill, F. Provost, and C. Volinsky, "Network-based marketing: Identifying likely adopters via consumer networks," *Statistical Science*, vol. 22, no. 2, pp. 256–275, 2006.

[29] W.-c. Feng, D. Brandt, and D. Saha, "A long-term study of a popular mmorpg," in *NetGames '07: Proceedings of the 6th ACM SIGCOMM workshop on Network and system support for games*, (New York, NY, USA), pp. 19–24, ACM, 2007.

[30] C.-N. Ziegler and G. Lausen, "Spreading activation models for trust propagation," in *EEE '04: Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'04)*, (Washington, DC, USA), pp. 83–97, IEEE Computer Society, 2004.

[31] S. R. Garner, “Weka: The waikato environment for knowledge analysis,” in *In Proc. of the New Zealand Computer Science Research Students Conference*, pp. 57–64, 1995.