ICEB 2004 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-5-2004

A Defection Detection Procedure Using SOM and Markov Chain : A Case of On-Line Game Provider

Kyunghee Chae

Heeseok Song

Juchoel Choi

A Defection Detection Procedure Using SOM and Markov Chain : A Case of On-Line Game Provider

Kyunghee Chae[1], Heeseok Song[2], Juchoel Choi[1]

[1] School of Bsiness Administration, KyungHee University, Seoul,130-701, Korea
[2] Department of Management Information Systems, Hannam University, Taejon, 306-791, Korea
khchae@gmail.com, songorak@hanmir.com, hssong@hannam.ac.kr

ABSTRACT

Customer retention is very important issue. In this reasons, many studies have been conducted for customer retention to demonstrate the potential of data mining through experiments and case studies. But current customer retention procedures only focus on detection of potential defectors based on likelihood of defection overlooking cost. In this paper, we propose an economic procedure for profit maximizing based on the customers' LTV and behavior state transition cost. For the aim of this paper, we use past and current customer behavior by integrating data mining techniques. In this procedure, SOM is used to determine the possible states of customer behavior from past behavior data. Based on this state representation, a Markov chain is applied to generate likelihood of state transition.

Keywords: *Customer relationship management, Data mining, Life-time value, Markov chain, Self-organizing map*

1. INTRODUCTION

Customer retention and life-time value(LTV) are an important issues in today's competitive environment. It is mapping customer data to define customer behaviors so that the processes of a company are fully occupied in acquiring, selling to, and maintaining a long-term relationship to a customer. The longer a customer stays with a company the more profit the customer generates. This is the outcome of a number of factors such as the higher initial costs of attracting new customer, the increase in the number of purchases, the long-term customer's better understanding of the company, and positive word-of-mouth. In this reasons, many studies have been conducted for customer retention to demonstrate the potential of data mining through experiments and case studies. But current customer retention procedures only focus on detection of potential defectors based on likelihood of defection overlooking cost.

This paper proposes an economic procedure for profit maximizing based on the customers' LTV and behavior state transition cost. For the aim of this paper, we use past and current customer behavior by integrating data mining techniques. Here, the term 'customer behavior' means action-oriented activities like calling, visiting Websites, and making purchases. However, predicting and understanding of customers' behavior for multiple periods is not easy because of the following reasons. First, customer defection cannot be predicted by only one behavior feature. Second, these features may also be related to customer defection in a highly nonlinear way. For this reasons, SOM(Self-Organizing Map) is used. In this procedure, SOM is used to determine the possible states of customer behavior from past behavior data. Because it facilitates the understanding of complex behavior dynamics so that several variables and their interactions can be inspected simultaneously. Based on this state representation, a Markov chain is applied to generate likelihood of state transition. Also to maximize company's profit, the proposed procedure recommends the desirable behavior state for the next period based on the economic analysis of transition cost and LTV. The transition cost is based on the likelihood of state transition because that low likelihood is meant to be hard to recommend a desirable behavior state, if so we have to make more effort. Here, the term 'desirable behavior state' means that it has high LTV.

Based on the state representations, the following procedures are developed: (1) collecting behavior data and pre-processing, (2) determining the behavior states through SOM training, (3) determining a desirable state, (4) developing a marketing strategy for profit maximizing. We applied our procedures to the online game provider, because the histories of customer behavior are easily collected during multiple periods. Although our procedures are explained for a specific online game company, they can be applied to various service industries with fluent customer behavior data such as telecommunications, internet access services, and contents providing services. We begin by reviewing the previous study on defection detection in section 2 and the concept of SOM and Markov chain, which are prerequisites for our research, are summarized. In section 3, our suggested procedures are explained step by step. A case study is presented to illustrate and evaluate our procedures in section 4. Finally we summarize our contributions and outline areas for further research in section 5.

2. RELATED WORK

2.1Existing Works In Customer Retention

Many Studies [2,5,6,7] have been conducted to detect potential defectors and preserve customer using data mining techniques. Logit regression, decision trees and neural networks are the most frequently used technique to predict the likelihood of defection in these studies. However, these studies use static model to predict the likelihood of defection in spite of continuous and dynamic property of individual access behavior. Song et al. [7] and Kim et al. [5] proposed a solution that use static model to predict the likelihood of defection in spite of continuous and dynamic property of individual access behavior, but it also overlooked cost that is very important factor to make a decision. Compared to these existing works, our research builds an economic procedure for customer retention. Not only our procedure could predict potential defector, but also give the guidance to control whose undesirable behavior which leads to defection.

2.2 SOM

SOM was developed in its present form by Kohonen [4] and thus they are also known as Kohonen Maps. The SOM is able to map a structured, high-dimensional data onto a much lower-dimensional array of neurons in an orderly fashion. The mapping tends to preserve the topological relationships of the input data. Topological preserving means that the data points lying near each other in the input space will be mapped onto nearby map units. Due to this topology preserving property, the SOM is able to cluster input information and their relationships on the map. In this reasons we adopt SOM to determine every possible behavior state for a specific domain. The SOM has proven to be a valuable tool in data mining, full-text mining and financial data analysis. It has also been successfully applied in various engineering applications in pattern recognition, image analysis, process monitoring and fault diagnosis [8].

2.3 Markov Chain

A Markov chain, which is one special type of discrete-time stochastic process, can be used to predict the evolution of customer states [3].
A Markov chain has the following properties and assumptions:

? The probability distribution of the state i at time $t+1$ only depends on the state at time t.
? For all states i and j, $p(X_{t+1} = j | X_t = i)$ is independent of time t. These two assumptions can be specified by the conditional probabilities for a Markov chain:

$$p(X_{t+1} = j | X_t = j,\ X_{t-1} = i_{t-1}, \ldots,\ X_1 = i_1,\ X_0 = i_0)$$
$$= p(X_{t+1} = j | X_t = i) = p_{ij}$$
$$for\ all\ states\ i_0, i_1, \ldots, i_{t-1}, i, j$$

We adopted the probability distribution of the state i at time $t+1$, and an absorbing Markov chain in which some of the states are absorbing and the rest are transient ones. An absorbing state has a special transition probability, $p_{ij} = 1$ [9].

3. PROPOSED PROCEDURE

3.1 Overall Procedure

In this section, we present the overall procedure of our methodology for defection detection and prevention with Figure 1. As shown in Figure 1, the proposed procedure consists of four steps. In the first step, all the behavior data related to defection is collected from operational databases and Web log files. In the next step, the SOM model is developed to determine all the possible behavior states which are mutually exclusive and exhaustive. The SOM is used to represent the state of each customer behavior from collected behavior data. After determining the behavior states using the SOM, all the continuous and complex behavior data is converted into discrete states. Based on this state representation, we determine a desirable state that is low defection rate, low transition cost, and high LTV. In the final step, we make a campaign plan to focus or modifying their behavior patterns for profit maximizing.

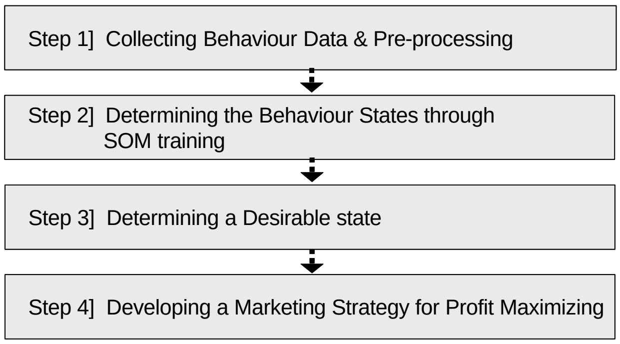

Figure 1. Overall procedure

3.2 Step 1: Collecting Behavior Data and Pre-processing

First of all, features that con affect defection are selected based on the prior knowledge of domain experts, and the value of features are collected over the multiple periods for each user from Web log files, purchasing databases, claim databases and the other sources. After collecting behavior data, data preparation is needed to build behavioral profiles. In particular the Web log data requires a great amount of effort to be transformed into usage data because of the difficulties in session identification and user identification. More detailed information for data preparation of the Web log can be found in Cooley et al.[1].

3.3 Step 2: Determining the Behavior States through SOM Training

Based on the selected input features, all the possible states of customer behavior are determined that are prerequisite for the SOM is a value array of the filtered features such as the average length of session, the number of sessions in unit period, the number of congestions per unit period, and so on. In determining the behavior states using the SOM, the total number of states has to be given. If we set the total number of states to be very large, then the prediction model will lose the generalization capability because of overfitting. On the contrary, if we specify the total number of states to be very small, then the model may not have discrimination power between defectors and non-defectors. Therefore, choosing the number of states must be done carefully. The decision for determining the number of states is usually dependent on the characteristics of the domain and the goals of decision-makers.

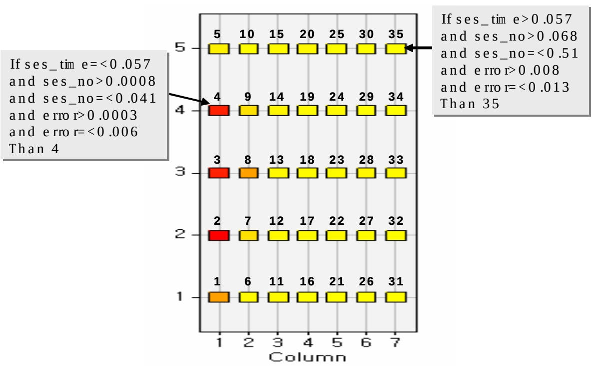

Figure 2. Behavior states and their interpretations

The satisfactory number of states may be given by trial and error using past multiple periods' behaviour data and the target variable (i.e., whether they actually defect or not). To evaluate the performance of each number of states, FalsePositive and FalseNegative are used as a measure in our procedure. FalseNegative error is more important than FalsePositive error because FalseNegative is more costly and dangerous than FalsePositive. If a non-defector is incorrectly classified as a defector (i.e., FalsePositive), the customer will receive a large concession, but if a defector is incorrectly classified as a non-defector (i.e., FalseNegative), the customer will terminate services before a company reaches the customer and thus the company will lose revenue [2]. Figure 2 illustrates the possible behavior states which are the results of SOM learning. To interpret each state on the map, a decision tree analysis(C5.0) is additionally conducted.

3.4 Step 3: Determining a Desirable State

In this step, we determine a desirable state for the aim of this paper that is the longer a customer stays. For determining a desirable state, we find a lower defection rate state at first. And then, we estimate profit rate of each state and recommend maximum profit state.

3.4.1 Average defection rate of each state

To determine a desirable state, at first, average defection rate have to be computed of each state : $Df_i = D_i / N_i$

Where, Df_i : the average defection rate of the state i

D_i : the number of defection in the state i

N_i : the total number of state i

3.4.2 Transition cost using Markov chain

The transition cost is based on the likelihood of state transition because that low likelihood is meant to be hard to recommend a desirable behavior state, if so we have to make more effort. The likelihood of state transition is generated from making an absorbing Markov chain. To make an absorbing Markov chain, we have two assumptions. First, we assume that customer behavior states have the Markovian property. It implies that given the present state of the customer, the future probabilistic behavior of the customer is independent of its past history and only depends on the current state. This is because the customer's current state information most impacts the decision of the customer's next behavior. Second, we assume that all the customers can be potential defectors because they eventually leave the company some time or other if they are not continuously satisfied with the company's service in today's greatly compressed online game industry.

Table 1. The transition probability matrix

		State at time t+1					
		1	2	3	4	---	def
State at time t	1	0	0	0.29	0.14	---	0
	2	0	0.08	0.23	0.08	---	0.08
	---	---	---	---	---	---	---
	34	0	0	0	0	---	0
	35	0	0	0	0	---	0.09
	def	0	0	0	0	---	1.00

To make a transition probability matrix, we used the states of all the customers for four periods (i.e., four weeks). We rearranged four pairs of transition data for each customer such as all of the transition data from t-4 week to t-3 week, from t-3 week to t-2 week, from t-2 week to t-1 week, and from t-1 week to t week (i.e., an absorbing state). And then we calculated transition cost, C_{ij} ,based on a transition probability from state i to another state j as shown in Table 1.

3.4.3 LTV generation using an absorbing Markov chain

In the Table 1, state 'def' means 'defector' which is an absorbing state. According to the property of an absorbing Markov chain, if the customer's current state begins in a non-absorbing state (i.e., from state 1 to state 35), then the customer will leave the non-absorbing state and end up in one of the absorbing states (i.e., state ‘def’) eventually [9]. Figure 3 describes the general transition matrix for an absorbing Markov chain. It consists of four parts which are I, O, R and Q, as illustrated in Figure 3. We can find out the elapsed time to absorbing state ‘def’ in each non-absorbing state by using the matrix $(I-Q)^{-1}$.

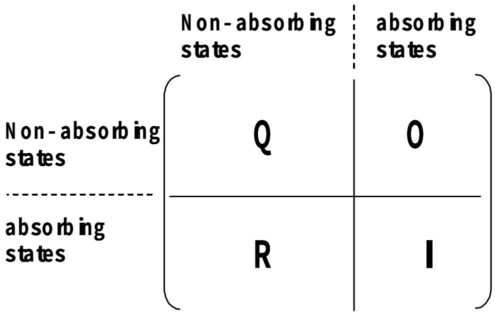

Figure 3. The transition matrix for elapsed time

Table 2 describes the outcome of the matrix $(I-Q)^{-1}$. In this Table, we can see that if a customer belongs to state 1 at time t, then the customer will defect 11.5 periods later on average. The elapsed time to absorbing state ‘def’ can be used as LTV generation. We use simple LTV that only considered period (i.e., elapsed time) and profit (i.e., fee about session time) : $LTV_i = E_i \times P_i$

Where, E_i : the elapsed time of the state i

P_i : the average profit of the state i

Table 2. Sample of the elapsed time to an absorbing state

State	Elapsed time (Week)	State	Elapsed time (Week)
1	11.5	6	14.1
2	9.8	7	12.9
3	7.5	8	9.4
4	11.3	9	10.5
5	8.2	10	12.2

3.4.4 Recommend a desirable state

In this step, we estimate gain of each state and recommend a desirable state that is maximum profit state. As it mentioned, we find a desirable state using a defection rate, and then desirable state is determined based on the profit that is transition cost and the LTV.:

$$Gain_{ij} = (LTV_j - LTV_i)/C_{ij}$$

3.2 Step 4: Developing a Marketing Strategy

For the defection prevention, we will design segmental campaigns based on the information of a To-Be state.

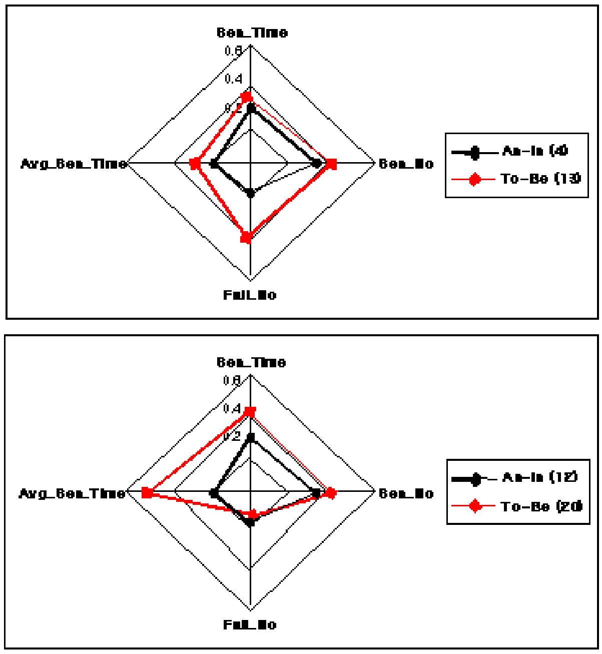

Figure 4. Segmental campaign design

Figure 4 provides samples of the campaign design. An As-Is behavior state and a To-Be behavior state for two segments are presented in this figure. From this figure, the campaign design for ‘4’ should focus on reduction error, and for ‘12’ should focus on extending session length. We will illustrate a case of campaign design in section 4.

4. A CASE STUDY

4.1 Data

The dataset for the case study is prepared from an online game provider in Korea. We sampled 255 customers and collected totally 114.736 transactions for those customers from various sources such as web log data and the transaction database. The collected input data for defection prediction contain the total session length for a week, the total number of sessions for a week, the total number of access errors caused by congestion, and the average length per session. Customer profiles were maintained for the four input features during four consecutive weeks for each customer with an actual defection indicator. We defined a defector as the customer who has no session time for one month, because there is no voluntary sign of withdrawal in this Internet game site.

4.2 Determining the Behavior States

As it mentioned, FalseNegative error is more important than FalsePositive error because if a non-defector is incorrectly classified as a defector (i.e., FalsePositive), the customer will receive a large concession, but if a defector is incorrectly classified as a non-defector (i.e., FalseNegative), the customer will terminate services before a company reaches the customer and thus the company will lose revenue. Therefore, we made an experiment on several cases and select a minimum of FalsePositive error in FalseNegative error is limited to 5%. The outcome is described in Table 3 and ‘5×7’ matrix, FalsePositive error is 17.6% and FalseNegative error is 4.6%, is appropriate to our case.

Table 3. FalsePositive and FalseNegative error matrix

Actual / Predicted	6×6		5×7		7×5	
	Def	Non	Def	Non	Def	Non
Defector	-	11.7	-	17.6	-	15.5
Non-defector	11.9	-	4.6	-	6.9	-

4.3 Determining a Desirable State

For determining a desirable state, we programmed a profit generation automatically using a Visual Basic 6.0. Table 4 provides desirable states(i.e., to-be) as a result and differences between as-is and to-be.

Table 4. Differences between as-is and to-be

As-is	To-be	Differences between as-is and to-be
1	2	Increase in session number
2	3	
3	10	Increase in average session time
4	10	
16	24	
17	18	
18	25	
9	10	Increase in session time
10	18	
11	12	
13	12	
22	25	
23	24	
12	25	Increase in session number and time
19	18	
21	25	
24	25	Decrease congestion number
27	25	Increase in session number and decrease congestion number
29	23	
32	24	

4.4 Developing a Marketing Strategy

Table 5 is a real cases and gives some idea about campaign object and feasible campaign based on Table 4. For instance, if we want to induce a customer in state ‘1’ to desirable state ‘2’, then we should focus on extending session number. And in this object, we can make a campaign plan like Table 5.

Table 5. Sample of the feasible campaign

Object	Feasible campaign
Extending session time	- Providing additional free access time - Providing gifts and awarding prizes based on the accumulation point
Extending session number	- To provide free coupon - Recommend new game information - Frequent update of the site and make an event
Reducing congestions	- Recommending the use of non-peak time by suggesting a discount rate - Sending a manual for congestion

4.5 Evaluation

In this section, we compare the prediction accuracy of the proposed procedure to that of the MLP neural network and a decision tree. To evaluate the performance of our defection detection procedure, we use FalsePositive and FalseNegative as measures. Figure 5 shows the comparative results of prediction accuracy among our suggested model and the MLP neural network(MLP NN) based on same input data in our domain(online game provider). Our proposed procedure resulted in both a slightly higher recall and hit ratio. This means that our proposed procedure can cover defection prevention as well as defection detection without deterioration of prediction accuracy compared to that of the MLP neural network.

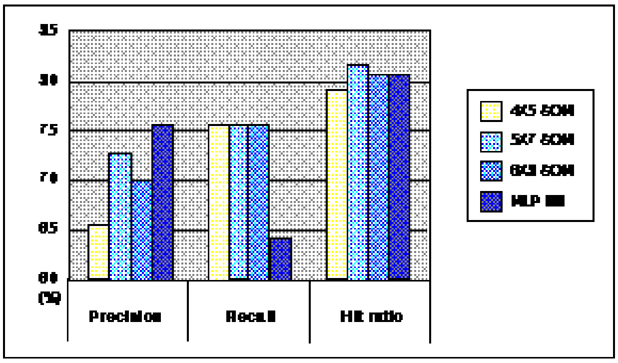

Figure 5. Comparative results of prediction accuracy

6. CONCLUSION

We proposed an economic procedure for profit maximizing based on the customers' LTV and behavior state transition cost. For this purpose, possible states of customer behavior are determined from past behavior data using SOM. Based on this state representation, the transition cost and LTV can be predicted using a Markov chain. In addition, our proposed procedure is extended to defection prevention for potential defectors and it assists in building segmental compaign plans by recommending the desirable behavior state for the next period to lower the likelihood of defection. Our proposed approach is a practical implementation procedure of eCRM because it tries to maintain a relatioonship with potential customers using an automated campaign procedure continuously

As an area for futher research, we have a plan to develop a system for defection detection and prevention based on our suggested procedure. In this study, we applied our procedure to the online game provider, but we believe that is can applied to other service industries such as telecommunications, Internet access services, and contents providing services, and check the effectiveness of our proposed procedure.

REFERENCES

[1] Cooley, R., B. Mobasher, J. Srivastava, "Data preparation for mining world wide web browsing patterns," *Knowledge and information systems*, Vol. 1, No. 1, pp5-32, 1999.
[2] Datta, P., B. Masand, D. Mani, B. Li, "Automated cellular modeling and prediction on a large scale," *Artificial Intelligence Review*, 14, pp485-502, 2000.
[3] Ha, S. H., S. M. Bea, S. C. Park, "Customer's time-variant purchase behavior and corresponding marketing strategies: An online retailer's case," *Computers&Industrial Engineering*, Vol. 43, pp801-820, 2002.
[4] Kohonen, T., "The Self-Organizing Map," *Proceedings of the IEEE*, Vol. 78, No. 9, pp1464-1480, 1990.
[5] Kim, Y. A., H. S. Song, S. H. Kim, "A dynamic procedure for defection detection and prevention based on SOM and a Markov chain," *International Conference of Korea Intelligent Information Systems Societ*, pp141-148, 2003.
[6] Ng, K., H. Liu, "Customer retention via data mining," *Artificial Intelligence Review*, Vol. 14, No. 6, pp569-590, 2000.
[7] Song, H. S., J. K. Kim, Y. H. Cho, S. H. Kim, "A personalized defection detection and prevention procedure based on the self-organizing map and association rule mining :Applied to online game sit," *Artificial Intelligence Review*, Vol. 21, No.2, pp161-18, 2004.
[8] Vesanto, J., "SOM-based data visualization methods," *Intelligent data analysis*, pp111-126, 1999.
[9] Winston, W. L., *Operations research applications and algorithms*, 3rd ed. Duxbury press, 1994.