대한산업공학회/한국경영과학회 2002 춘계공동학술대회
한국과학기술원(KAIST) 2002년 5월 3일-4일

A Personalized Customer Retention Procedure For Internet Game Site Based on the Self-Organizing Map and Association Rule Mining.

Hee Seok Song [1*], Jae Kyeong Kim[2], Soung Hie Kim [3], and Kyung Hee Chae [4]
1,3 Graduate School of Management, Korea Advanced Institute of Science and Technology207-43 Cheongryangri-Dong, Dongdaemun, Seoul, 130-012, South Korea

2,4 School of Business Administration, KyungHee University#1, Hoeki-Dong, Dongdaemoon, Seoul, 130-701, South Korea

Abstract

This paper propose a personalized defection detection and prevention procedure based on the observation that potential defectors have tendency to take a couple of months or weeks. For this purpose, possible states of customer behavior are determined from past behavior data using SOM (Self-Organizing Map). For the evaluation of the proposed procedure, a case study has been conducted for a Korean online game site. The result demonstrates that the proposed procedure can assist defection prevention effectively and detect potential defectors without deterioration of prediction accuracy comparison to prediction by MLP. Our procedure can be applied to various service industries that can capture fluent customer behavior data such as telecommunications, internet access services, and content services, too

1. Introduction

Customer retention is an increasingly pressing issue in today's competitive commercial arena (Ng and Liu, 2000). The longer a customer stays with a company the more profit the customer generates.

The aim of this paper is developing a personalized procedure for defection detection and prevention using past and current customer behavior by integrating data mining techniques.

we developed a representation of the state of a customer behavior from past behavior data and used this representation in monitoring the current state of the customer behavior. For this purpose, the SOM (Self-Organizing Map) is adapted because it facilitates understanding of complex behavior dynamics (Alhoniemi et al., 1999 ; Simula et al., 1999).

2. Existing works in defection detection

In this section, we review the brief introduction of defection and summarize different approaches for defection detection. Defection which is also called to churn or customer attrition is defined as "the annual turnover of the market base".

Many researches (Berson et al., 2000; Datta et al., 2000; Eiben et al., 1998; Mozer et al., 2000; Ng and Liu, 2000; Raghavan et al., 2000; Smith et al., 2000; Yeo et al., 2001) have been conducted to detect potential defectors using data mining techniques.

These researches use static model to predict the likelihood of defection in spite of continuous and dynamic property of individual access behavior.

When we build dynamic model to predict defection using multi-period behavior data, it can be possible to develop personal procedure for defection prevention. Although potential defectors have been identified, without guidance to control undesirable behavior which leads to defection for each individuals, they will finally defect soon. However, most existing works in customer retention have not handled defection prevention procedure yet.

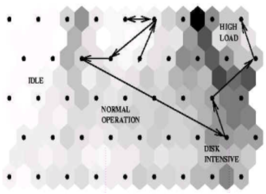
3. Background

3.1 SOM

SOM was developed in its present form by Kohonen (Kohonen, 1990 ; Kohonen, 1995 ; Kohonen et al., 1996 ; Simula et al., 1999). 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.

The SOM model is made up of input layer and output layer. Each neuron in the input layer is connected to each neuron in the output layer. Each one of these connections has a synaptic weight associated with it. In SOM, each neuron is represented by an n-dimensional weight vector, $\mathbf{m} = [m_1, m_2, \ldots, m_n]$, where n is equal to the dimension of the input vectors.

The SOM has proven to be a valuable tool in data mining and KDD with applications in full-text and financial data analysis. It has also been successfully applied in various engineering applications. Especially in application of dynamic process monitoring, the SOM is used to form a display of the operational states of the process as explained in Figure 1. The operation point (i.e., the current process state) and its history in time can be visualized as a trajectory on the map which makes it possible to track the process dynamics in a new way.



<Figure 1> Visualization of dynamic trajectory of operational states using SOM (Source : Alhoniemi et al., 1999)

3.2 Association rule mining

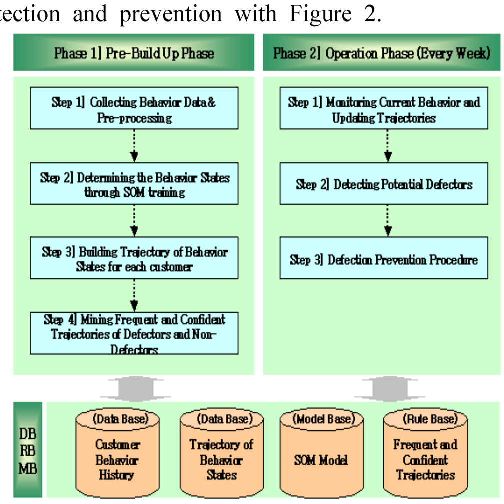
A typical association rule has an implication of the form $A \Rightarrow B$ where A is an itemset and B is an itemset that contains only a single atomic condition.

Association rule mining finds all collections of items in a database whose confidence and support meet or exceed prespecified threshold value. Apriori algorithm is one of the prevalent techniques used to find association rules (Agrawal et al., 1993; Agrawal and Srikant, 1994). Apriori operates in two phases. In the first phase, all large itemsets are generated. The second phase of the algorithm generates rules from the set of all large itemsets.

4. Proposed procedures

4.1 Overall procedure

In this section, we present the overall procedure of our personalized methodology for defection detection and prevention with Figure 2.



<Figure 2> Overall procedure for personalized defection detection and prevention

Pre-build up phase is performed once to build reliable model for defection prediction and operation phase is performed periodically (i.e. weekly sliding manner) to detect potential defectors and guide to desirable direction in their behavior pattern of those potential defectors.

The final outcomes of pre-build up phase are rules for frequent and confident trajectories of defectors and non-defectors which are developed by using past customer behavior history for actual defectors and non-defectors.

In the operation phase, potential defectors are identified by comparing new recent trajectories of customer behaviors to rules in FCT RB which have been built in pre-build up phase. For this purpose, new monitored current behavior is converted to behavior state using SOM model at the first step. At the second step, potential defectors are identified by a scoring measure (Liu et al., 2000; Ma, et al., 2000) which reflects the likelihood of defection.

4.2 Pre-build up phase

4.2.1 Step1 : Collecting behavior data and pre-processin

First of all, features which can affect defection are selected based on the prior knowledge of domain expert. Table 2 illustrates the sample records of CBH DB for an individual user.

Customer ID	Periods	Average Session Time	No. of Session per Day	No. of Congestion	...	No. of BBS Writing
10001	(t-p)th	30	10	2	...	4
...	...	...	...	...	...	...
10001	(t-2)th	40	13	2	...	7
10001	(t-1)th	34	10	2	...	5
10001	tth	40	19	5	...	7

<Table 2> Sample behavior profiles for an individual in CBH DB

In this process, data preparation is needed to build behavioral profiles. Especially in web log data, it requires huge effort to extract usage data because of difficulty in session identification and user identification.

If the number of measuring features to be considered is very large, it will be necessary to select a much smaller number of representative features to the inputs of SOM. Especially, using SOM in learning algorithm, superfluous features can make computational inefficiency and insignificant clusters which are not related to the defection.

4.2.2 Step 2 : Determining the behavior states through SOM training

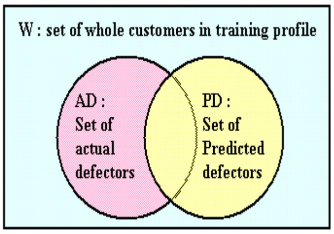
We describe how to determine the whole possible states of customer behavior which are prerequisite for building dynamic prediction model in this step. The SOM is applied to determine the states.

In determining the behavior states using SOM, the total number of states has to be given due to the nature of many clustering techniques. If we set the total number of states to be very large, then the prediction model will lose the generalization capability because of overfitting.

Decision for determining the number of states is dependent on the characteristics of domain and their goals. The optimal number of states can be given by trial and error using past multi-period behavior data and target variable (i.e. Whether they are actually defected or not ?). In this point, precision and recall which are originated from information retrieval can be applied to evaluate the alternatives. In our methodology, precision and recall is defined as follows.

Definition 1] Precision is (P) is the ratio of the actual number of defectors among predicted defectors to the predicted number of defectors.

Definition 2] Recall is (R) is the ratio of the predicted number of defectors who are actual defectors to the actual number of defectors in all the training profiles.



<Figure 3> The concept of Precision and Recall

$$P = \frac{|PD \cap AD|}{|PD|}, \quad R = \frac{|PD \cap AD|}{|AD|},$$

where $|A|$: number of element in set A

4.2.3 Step 3 : Building trajectory of behavior states for each customer

After determining the possible behavior states using SOM, we can make trajectories of customer behavior states by evaluating of which state the behavior data at each period in CBH DB belong to. This process is simply done by running SOM model which has been built at step 2. The Table 3 represents the sample trajectories of customer behavior states over time in TBS DB as a result of step 3.

Customer ID	Periods						Defect/Non-defect
	t^{th}	$(t-1)^{th}$	$(t-2)^{th}$	$(t-3)^{th}$	...	$(t-p)^{th}$	
10001	D	K	K	N	...	A	Defect
10002	A	N	P	K	...	N	Non-defect
10003	D	D	K	P	...	A	Defect
10004	D	K	A	P	...	P	Defect
10005	P	K	N	B	...	C	Non-defect

<Table 3> Sample trajectories of behavior states in TBS DB

4.2.4 Step 4 : Mining frequent and confident trajectories of defectors and non-defector

In this step, frequent and confident trajectories of defectors and non-defectors during multi-periods are discovered. For this purpose, we apply Apriori algorithm which is originally developed for association rule mining.

4.3 Operation Phas

4.3.1 Step 1 : Monitoring current behavior and updating trajectories

All the current behaviors of customers are collected and converted to behavior state using SOM model at the first step. Next, new monitored behavior state is added to prior trajectory of those customers and oldest entry is discarded.

4.3.2 Step 2 : Detecting potential defector

Potential defectors are identified by assigning the scores which represent likelihood of defection. For this purpose, we adapted scoring method based on association rules from the works of Liu et al. (2000) and Ma, et al. (2000). Scoring using association rules is challenging because when we want to score a data case, there are often many rules that can be applied.

Scoring function is developed as follows by taking into account of the above information (Liu et al., 2000; and Ma, et al., 2000).

$$S = \frac{\sum_{i \in POS} W^{i}_{positive} \times conf^{i} + \frac{1}{k} \sum_{j \in NEG} W^{j}_{negative} \times conf^{j}_{positive}}{\sum_{i \in POS} W^{i}_{positive} + \sum_{j \in NEG} W^{j}_{negative}}$$

POS : the set of positive class rules that can cover the data case

NEG : the set of negative class rules that can cover the data case

$W^{i}_{positive}$: the weight for the positive class rule i

$$W^{i}_{positive} = conf^{i} \times sup^{i}$$

$W^{j}_{negative}$: the weight for the negative class rule j

$$W^{j}_{negative} = conf^{j} \times sup^{j}$$

$conf^{i}$: the original confidence of the positive class rule

$conf^{j}_{positive}$: the confidence after converting the negative class rule j to a positive class rule,

i.e., $conf^{j}_{positive}$ = 1-' the confidence of rule j '.

The value of S is between 0 and 1 inclusively. Finally, customers who have the likelihood of defection greater than specified threshold can be targeted and defection prevention procedure is activated for those users.

4.3.3 Step 3 : Defection prevention procedure

The objective of defection prevention step is to drive their customer to desirable direction in their behavior pattern. If we can find a behavior state (i.e. a node on SOM) which will lower the likelihood of defection, then we can do campaign so as their behavior pattern to belong to that state at next period.

Driving customer closer state to current behavior state on the map means the less costs are spent to campaign for defection prevention because of topology preserving property of SOM. Therefore our key idea is driving the potential defectors to nearest neighbor node on SOM at next period that lowers the likelihood of defection. Above idea can be implemented by following algorithm as described in Figure 4. This algorithm is performed for each potential defector who has been detected at step 2.

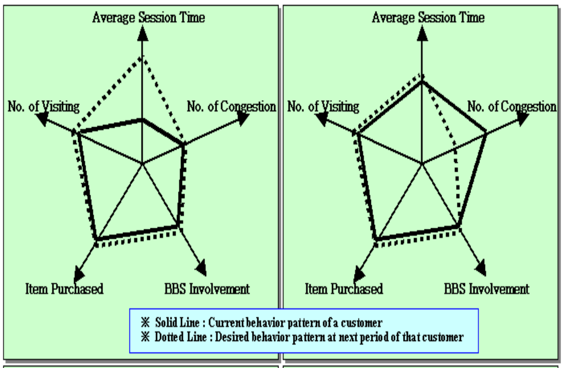
Notation	$k-NeighborNode$: set of kth nearest neighbor node ($0\text{-}NeighborNode$: only include present node)

Algorithm	$Ni = 0, MaximumGain = -1$
	$0\text{-}NeighborNode = \{n_1\}$, $1\text{-}NeighborNode = \{n_2, n_3, \ldots, n_i\}$
	for ($i = 1; i \neq \phi; i++$) do begin
	  $NewTrajectory$ = add n_i to $ExistingTrajectory$, discard oldest entry
	  scoring the likelihood of defection for $NewTrajectory$
	  if score($ExistingTrajectory$) - score($NewTrajectory$) > 0 and
	    score($ExistingTrajectory$) - score($NewTrajectory$) > $MaximumGain$
	    then $Ni = i$ and
	      $MaximumGain$ = score($ExistingTrajectory$) - score($NewTrajectory$)
	  endif
	end
	if $Ni = 0$ then "failed to find desirable node for next period"
	  else "recommend to drive to n_{Ni} node for next period"
	endif

<Figure 4> Procedure for defection prevention

In Figure 4, k-NeighborNode means the set of kth nearest neighbor nodes. In this notation, k=0 means the set of neighbor node which contains only current node, and k=1 means the set of first neighbor nodes which contain 4 or less neighbor nodes on rectangular structureSOM. In this procedure, we assume that it is impossible to drive the customers to second or more nearest neighborhood node on SOM for next period through campaign. This assumption is based on the fact that it is very difficult to control the customer behavior in a short period. Therefore, we assume that we can guide our customers to only zero or first nearest neighborhood node on SOM.

Figure 5 explains how the campaign plan can be built using the results of defection prevention procedure.



✧ Campaign for extending session time	✧ Campaign for reducing congestions
✓ Providing free of charge usage time ✓ Providing gifts according to length of usage time	✓ Recommending use of non-peak time ✓ Sending a manual for congestion

<Figure 5> Design of the personalized campaign

5. Experiments

5.1 Domain

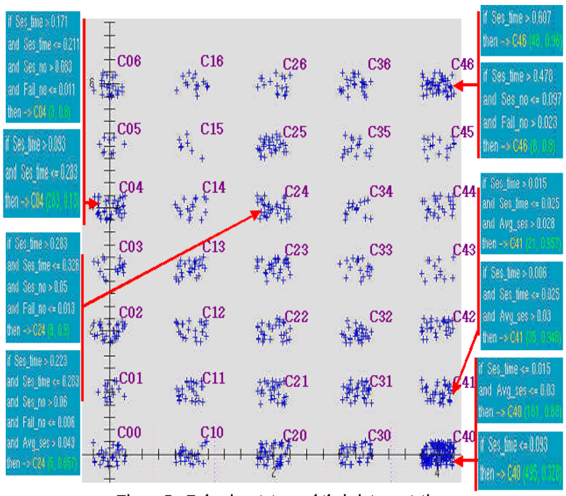
The case study has been conducted to evaluate how well the procedure performs its intended task of defection detection and prevention. The dataset is prepared from a Korean online game company who provides online multi-user game service. We sampled 255 customers and collected totally 114,736 transactions for those customers from various sources such as web log data and congestion DB. We balanced percentage of defectors and non-defectors as 35 % and 65 %. The collected input data for defection prediction contains total session length for a week, total number of sessions for a week, total number of congestions that the customer met, and averagesession length per session. Final customer behavior history (CBH) dataset maintains above 4 input data during recent 4 weeks for each customer with actual defection indicator. We selected week as a regular time interval for aggregating the behavior data because it seemed a reasonable trade-offbetween accuracy of prediction and lead-time for reaction.

5.2 Result

Figure 6 provides sample data for CBH. To interpret each state on map, decision tree analysis is additionally conducted with each state indicator as a target class. Figure 7 illustrates the possible behavior states and their interpretations.

Customer ID	Weeks	Total session time	Total number of session	Total number of congestion	Average session time	Defect/ Non-defect
ADOON6309	1	10	1	0	10.00	def
	2	0	0	0	0.00	def
	3	0	0	0	0.00	def
	4	5	1	3	5.00	def
AVAL	1	19	3	0	6.33	def
	2	39	3	0	13.00	def
	3	147	5	1	29.40	def
	4	1343	23	9	58.39	def
DANNYKIM	1	433	18	13	24.06	def
	2	2496	15	24	166.40	def
	3	8217	29	32	283.34	def
	4	3404	25	24	136.16	def
MI0935	1	6457	38	10	169.92	ndef
	2	5948	41	50	145.07	ndef
	3	2462	25	21	98.48	ndef
	4	5979	54	36	110.72	ndef
JYOA	1	3748	26	15	144.15	ndef
	2	5318	36	19	147.72	ndef
	3	3933	22	20	178.77	ndef
	4	2575	28	9	91.96	ndef

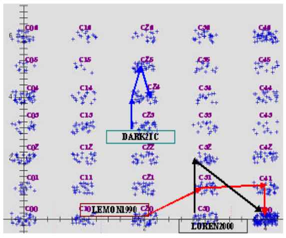
<Figure 6> The sample data for Customer Behavior History (CBH)



<Figure 7> Behavior states and their interpretations

For example, state C40 contains only the short usage behaviors because it consists of behaviors of which session length is less than or equal to 0.093. Therefore, we can interpret the state C40 as a light usage state. Figure 8 shows sample trajectories of behavior states (TBS) and their visualization on map.

Customer-ID	T-4 week	T-3 week	T-2 week	T-1 week	Defect/Non-defect
ABILITY3	C00	C14	C34	C32	ndef
BABYBEAR	C36	C25	C35	C45	ndef
BIZ212	C06	C06	C06	C16	ndef
CNLTK	C11	C00	C40	C41	ndef
DARK21C	C23	C24	C25	C24	ndef
HYUNJEE7	C40	C40	C40	C40	def
JAL8011	C31	C41	C31	C13	def
JJH0416	C04	C05	C04	C04	def
JUTMAN	C20	C02	C00	C21	def
K1KMH	C42	C42	C42	C40	def
K도사	C00	C40	C40	C40	def
KIWOOM1992	C00	C00	C40	C40	def
LEE3462	C41	C40	C41	C41	def
LEMON1990	C20	C31	C41	C40	def
LOREN2000	C30	C32	C40	C40	def



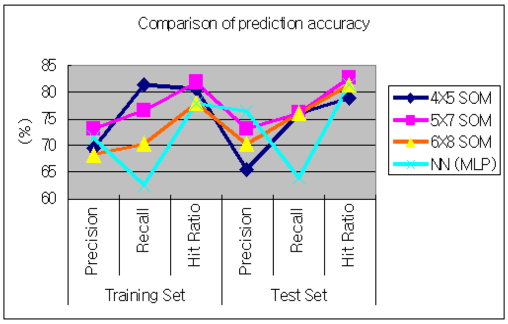
<Figure 8> The sample Trajectories of Behavior States (TBS) and their graphical representation

Next, association rule mining is performed to find frequent and confident trajectories for defectors and non-defectors. We give 3 % minimum support and 60 % minimum confidence as an input for association rule mining and get 30 negative class rules and 25 positive class rules. Using those rules, all the customers in training set and test set are scored as illustrated in Figure 9.

Customer_ID	T-4 week	T-3 week	T-2 week	T-1 week	Actual	Score	Predicted(※)
또리붕	C43	C31	C41	C40	def	0.7772791	def
렛데리아	C23	C22	C20	C40	def	0.6402468	def
르네스	C42	C40	C40	C04	def	0.7446328	def
마천사대벌	C20	C40	C40	C40	def	0.7914637	def
만두동생	C11	C40	C40	C40	def	0.7914637	def
맹꽁이지수	C40	C40	C00	C00	def	0.6778843	def
멋쟁이남자덜	C40	C40	C40	C40	def	0.7964321	def
미니제니	C13	C20	C21	C31	def	0.1103073	ndef
만루홈런	C13	C04	C22	C23	ndef	0.0568182	ndef
멋진상준	C13	C03	C03	C14	ndef	0.1250000	ndef
멋진원상	C06	C40	C01	C06	ndef	0.4835134	ndef
메탈가르몬	C30	C40	C40	C03	ndef	0.7446328	def
민들레씨앗	C46	C46	C46	C46	ndef	0.0226117	ndef
※ Defection threshold = 0.5							

<Figure 9> The sample customers who are assigned the defection score

And finally, the accuracy of prediction is measured using precision, recall and hit ratio. For the setting of optimal number of states, we repeated above experiment process many times. Figure 10 shows the comparative results of prediction accuracy among 3 alternatives (i.e. 20 states, 35 states, 48 states) and MLP neural network with 4 nodes in hidden layer. 5X7 SOM model is the most appropriate for this domain because of the high hit ratio, precision and recall. We also tried the prediction of defection using MLP neural network model based on the same input data. The proposed procedure which adapt 5X7 SOM model resulted slightly higher accuracy than MLP neural network model.



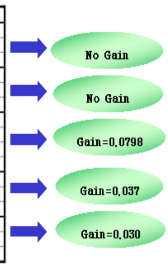
(Unit : %)

Techniques	Training Set Precision	Training Set Recall	Training Set Hit Ratio	Test Set Precision	Test Set Recall	Test Set Hit Ratio
4X5 SOM	69.3	81.3	80.6	65.5	76	78.7
5X7 SOM	73.1	76.6	81.7	73.1	76	82.7
6X8 SOM	68.2	70.3	77.8	70.3	76	81.3
NN (MLP)	71.4	62.5	77.8	76.2	64	81.3

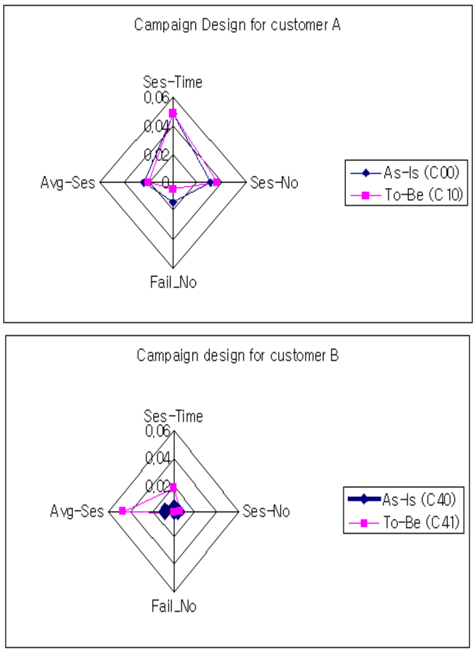
<Figure 10> Comparative results of prediction accuracy

To demonstrate how well the defection prevention procedure performs, we randomly selected 5 potential defectors and compute score gains under assumption that it can be driven to nearest neighbor node with low cost at next period. Figure 11 illustrates their new scores when they are guided to nearest neighbor node at next week and their score gains. For the two customers, we failed to find 0 or 1st nearest neighbor node which lower the defection score at next week. But with regards to the three customers, we can lower their likelihood of defection as much as their score gains. Customers who have highest score gain are targeted first to build campaign. Figure 12 provides the real case for campaign design.

Customer_ID	T-3 week	T-2 week	T-1 week	T week (To Be)	defect	New Score
또리붕	C31	C41	C40	C40	def	0.8521120
				C41	def	0.8673612
				C30	def	0.8838123
렛데리아	C22	C20	C40	C40	def	0.7466900
				C41	def	0.6470568
				C30	def	0.6936889
르네스	C40	C40	C04	C04	def	0.6648686
				C03	def	0.6648686
				C05	def	0.6648686
				C14	def	0.6648686
마천사대벌	C40	C40	C40	C40	def	0.7964321
				C41	def	0.7536564
				C30	def	0.7699426
맹꽁이지수	C40	C00	C00	C00	def	0.6832297
				C01	def	0.6470000
				C10	def	0.6470000



<Figure 11> Computing the score gains



<Figure 12> Personalized campaign design

6. Conclusion

We proposed a personalized defection detection and prevention procedure based on the observation that potential defectors have tendency to take a couple of months or weeks to gradually change their behavior (i.e. trim-out their usage volumes) before their eventual withdrawal. For this purpose, possible states of customer behavior are determined from past behavior data using SOM (Self-Organizing Map). Based on this state representation, potential defectors are detected by comparing their monitored trajectories of behavior states with frequent and confident trajectories of past defectors. Also, the proposed procedure is extended to defection prevention for potential defectors which assist building personalized campaign plan by recommending the desirable behavior state for next period so as to lower the likelihood of defection.

As a further research area, we have plan for setup the defection detection and prevention system in real business based on our suggested procedure. And it will be also interesting to check the effectiveness of the system. Also we plan to apply to various service industries that can capture fluent customer behavior data such as telecommunications, internet access services, and content services, too.

References

Agrawal, R., Imielinski, T., & Swami, A.(1993). Mining association rules between sets of items in large databases. Proceedings of the ACM SIGMOD Conference on Management of Data.
Agrawal, R., & Srikant, R. (1994). Fast algorithm for mining association rules. Proceedings of the International Conference on Very Large Databases (VLDB-94) (pp. 487-499).
Alhoniemi, E., Hollmen, J., Simula, O., Vesanto, J., Process monitoring and modeling using the Self-Organizing Map, Integrated CAE, 1999
Berson, A., Smith, S., Thearling, K., Building data mining applications for CRM, 2000, McGraw-Hill, 277-298
Eiben, A. E., Koudijs, A. E., and Slisser F., Genetic modelling of customer retention,In W. Banzhaf, R. Poli, M. Schoenauer, and T.C. Fogarty, editors, Proceedings of First European Workshop on Genetic Programming, number 1391 in LNCS, pages 178-186. Springer, Berlin, 1998.
Hall, M. A., and Smith, L. A., Practical feature subset selection for machine learning, Proceedings of the 21st Australian Computer Science Conference, Springer, 1998, 181-191
Kasslin, M., Kangas, J., Simula, O, Process state monitoring using Self-organizing maps, In I. Aleksander and J. Taylor (Eds.), Artificial Neural Networks, 2, Volume II, North-Holland, 1992, 1531-1534
Kohonen, T., The Self-Organizing Map, Proceedings of the IEEE 78(9), 1990, 1464-1480
Kohonen, T., Self-Organizing and Associative Memory. 5th ed., Springer-Verlag, Berlin, 1995
Kohonen, T., Oja, E., Simula, O., Visa, A., and Liu, B., Ma, Y., Wong, C. K., Yu, P., Target selection via scoring using association rules, IBM Research Report 21697, March 2000, NY
Ma, Y., Liu, B., Wong, C. K., Yu, P. S., and Lee, S. M., Targeting the right students using data mining, Proceedings of 6th international conference on KDD 2000, 2000, 457-464
Ng, K., and Liu, H., Customer retention via data mining, Artificial Intelligence Review 14(6), 2000, 569-590
Simula, O., Vasara, P., Vesanto, J., and Helminen, R., The Self-Organizing Map in industry analysis, CRC Press, 1999
Simula, O., Vesanto, J., Alhoniemi, E., and Hollmen, J., Neuro-Fuzzy Techniques for Intelligent Information Systems, Springer Verlag, 1999, chapter 4. Analysis and modeling of complex systems using the Self-Organizing Map
Smith, K. A., Willis, R. J., Brooks, M., An analysis of customer retention and insurance claim patterns using data mining : a case study, Journal of the operational research society 51, 2000, 532-541
Song, H. S., Kim, J. K., Kim, S. H., Mining the change of customer behavior in an internet shopping mall, Expert systems with applications 21(3), 2001, 157-168
Trubik, E., and Smith, M., Developing a model of customer defection in the Australian banking industry, Managerial Auditing Journal 15(5), 2000, 199-208
Tryba, V., and Goser, K., Self-organizing feature maps for process control in chemistry, In T. Kohonen, K. Makisara, O. Simula, and J. Kangas (Eds.), Artificial Neural Networks, North-Holland, 1991, 847-852
Yeo, A. C., Smith, K. A., Willis, R. J., and Brooks, M., Modelling the effect of premium changes on motorinsurance customer retention rates using neural networks, Computational science, Lecture notes in computer science, vol. 2074, Springer-Verlag, Berlin, 2001, 390-399