A CREDIT CARD USAGE BEHAVIOUR ANALYSIS FRAMEWORK - A DATA MINING APPROACH

Chieh-Yuan Tsai
Department of Industrial Engineering and Management, Yuan Ze University
No. 135, Yuantung Rd., Chungli City, Taoyuan County, Taiwan

Keywords: e-Commerce, Customer Relationship Management, Data Mining, LabelSOM, Fuzzy Decision Tree.

Abstract: Credit card is one of the most popular e-payment approaches in current online e-commerce. To consolidate valuable customers, card issuers invest a lot of money to maintain good relationship with their customers. Although several efforts have been done in studying card usage motivation, few researches emphasize on credit card usage behaviour analysis when time periods change from *t* to *t*+1. To address this issue, an integrated data mining approach is proposed in this paper. First, the customer profile and their transaction data at time period *t* are retrieved from databases. Second, a LabelSOM neural network groups customers into segments and identify critical characteristics for each group. Third, a fuzzy decision tree algorithm is used to construct usage behaviour rules of interesting customer groups. Finally, these rules are used to analysis the behaviour changes between time periods *t* and *t*+1. An implementation case using a practical credit card database provided by a commercial bank in Taiwan is illustrated to show the benefits of the proposed framework.

1 INTRODUCTION

Companies across all sectors and of all sizes are now benefiting from Internet e-commerce where services, information and/or products are exchanged via the web. To finalize the online transaction, providing convenient e-payment approaches for consumers are very important. Among several e-payment mechanisms, credit card is one of the most welcome approaches for online stores. Card issuers earn their profit from the fee charged to the store that accepts the credit card, charging interest on outstanding balances, and fee charged to customers. The major fees contributed from customers are (1) payments received late (past the "grace period"); (2) charges that result in exceeding the credit limit on the card (whether done deliberately or by mistake); (3) cash advances and convenience checks; and (4) transactions in a foreign currency. However, raising these fees or interesting rates to increase the profit of card issuers could be very difficult in current competitive credit card markets.

Another strategy to increase the profit of card issuers is to forge closer and deeper relationships with customers by emphasizing on customer relationship management (CRM) (Giudici and Passerone, 2002; Tsai and Chiu, 2004). CRM focuses on customer's need and regard customer life cycle as important assets of an enterprise. Due to the advance of the Information Technology (IT), it is easy to discover the usage information of what their customers purchase, when they use the card, and how often they consume. When the usage information is available, the card issuers can encourage customers use their cards more frequently through offering suitable products and services.

Data mining is the technique to discover meaningful patterns (rules) from large databases. Much of existing data mining researches in credit card fields has focused on building accurate models for risk and scoring analysis (Lee *et al*. 2006), cross selling (Wu and Lin, 2005), and fraud detection (Kou *et al*. 2004; Chen *et al*., 2005b). Relatively little attention has been made to analyze pattern changes in databases collected over time (Donato *et al*., 1999). However, customer behaviour usually changes over time. Some frequent patterns at one time period may not be valid for another time period (Chen et al., 2005a; Tsai *et al*., 2007). For example, a group of customer has preference in shopping at department stores this year and might change their preference to doing outdoor activities and travelling

Tsai C. (2007).
A CREDIT CARD USAGE BEHAVIOUR ANALYSIS FRAMEWORK - A DATA MINING APPROACH.
In *Proceedings of the Second International Conference on e-Business*, pages 219-226
DOI: 10.5220/0002108102190226

in the following year. If issuers cannot capture the behaviour change dynamically due to the time difference, it will be hard to retain customers by tailoring appropriate products and services to satisfy their real needs.

This paper is organized as follows. Section 2 reviews the literatures related to the change analysis models. Section 3 introduces the proposed credit card usage analysis framework that adopts LabelSOM algorithm and fuzzy decision tree algorithm. Section 4 provides an implementation case using a practical credit card database provided by a commercial bank in Taiwan to demonstrate the benefit of the proposed framework. A summary and future works are concluded in Section 5.

2 LITERATURE REVIEW

Current businesses face the challenge of a constantly evolving market where customer's needs are changing rapidly. Some researches applied customer demographic variables such as recency, frequency, and monetary (RFM) to analyze customer behavior (Tsai and Chiu, 2004). Although RFM analysis can effectively investigate customer values and segment markets, it is not a suitable tool for detect the customer behavior changes. Therefore, to better understand customer behaviors, developing suitable change detection models becomes an important research topic in the financial business.

Except the studies of rule maintenance in the changed database, some researches focus on discovering emerging patterns. Emerging pattern mining can be defined as the process to discover significant changes or differences from one database to another (Dong and Li, 1999). Emerging pattern captures emerging trends in time stamped database. Another related research trend is subjective interestingness mining. Interestingness mining is to find unexpected rules with respect to the user's existing knowledge. Unexpected changes compare each newly generated rule with each existing rule to find degree of difference (Liu and Hsu, 1996). Liu et al. (Liu et al., 1999) proposed a DM- II (Data Mining-Integration and Interestingness) system which has classification and association rule mining tasks to help users perform interestingness analysis of the rules. Its analysis compares each newly generated rule with each existing rule to find degree of difference, which is useful and important for real-life data mining applications. Han et al. (1999) presented several algorithms for efficient mining of partial periodic patterns, by exploring some interesting properties related to partial periodicity. The algorithms show that mining partial periodicity needs only two scans over the time series database to make efficient in mining long periodic patterns.

3 ANALYSIS FRAMEWORK

The proposed credit card usage behaviour analysis framework consists of four major stages as shown in Figure 1. The first stage is data extraction and pre-processing. In this stage, the customer profile and their transaction data at time period t are retrieved from databases. The second stage is to conduct customer segmentation using the LabelSOM neural network. The LabelSOM adaptively cluster customers into groups and automatically identifies critical demographic features for each group. In the third stage, the usage behaviour of the customers in the interesting group is generated using fuzzy decision tree (FDT) algorithm that represents usage behaviour as a set of IF-THEN rules. After obtaining the usage patterns of interesting customer group at time period t, we can trace the behaviour changes of these customers from time period t to t+1 when retrieving their corresponding data at time period t+1. The LabelSOM algorithm in the second stage and the FDT the third stage are further introduced in the following sub-sections.

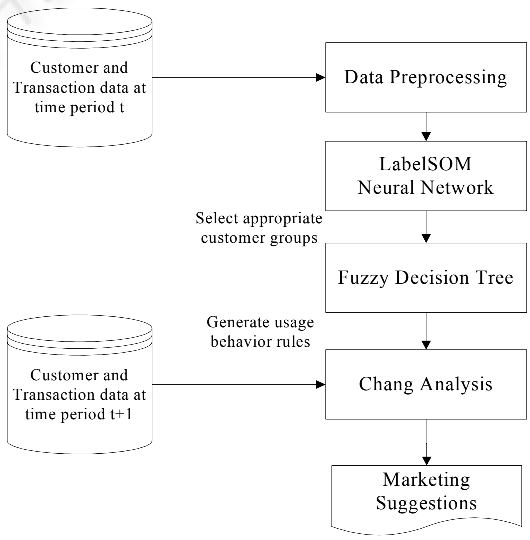

Figure 1: The proposed credit card usage behaviour analysis framework.

3.1 Label Self-Organizing Map (LabelSOM)

Self-Organizing Map (SOM) neural network, proposed by Kohonen (1990), is recognized as one of the popular clustering methods. The SOM employs a competitive unsupervised learning technique to project high dimension data into a two-dimensional grid without destroying data topology. Data points that are close each other in the input space are mapped to nearby output neurons in the SOM. An input neuron represents a data feature in N dimension and an output neuron represents the clustering result in two-dimensional space. The output neuron that has the highest similarity with an input data point is claimed as the winner (the best matching unit). The weights of the winner node and its neighbouring neurons are then adjusted automatically to force the weights closer to the input vector. After completing this learning process, each neuron represents a set of data.

Self organizing maps are an unsupervised neural network model which lends itself to the cluster analysis of high dimensional input data. However interpreting a trained map could be difficult because the features responsible for a specific cluster assignment are not evident from the resulting map representation. To solve this difficulty, the LabelSOM was developed to automatically label every node of a trained SOM (Rauber and Merkl, 1999). That is, LabelSOM algorithm can not only conduct cluster operation well but also distinguish the difference between each cluster clearly. Therefore, the LabelSOM is adopted for customer segmentation in our second stage.

The operations of the LabelSOM algorithm are summarized into the following five steps.

1. Input the primary parameters of the LabelSOM. The primary parameters include the number of input neurons, number of output neurons, number of input data, learning iteration, learning rate α, and radius η. In addition, initial weight matrix **W** is set randomly.

2. Conduct the following three sub-steps for each input vector $x_i = (x_{i1}, x_{i2}, ...x_{ik})$ sequentially where $i = 1, 2, ..., m$, m is the total number of input data points and k is the number of neurons in the input layer. Notes that k is also the number of features for an input point.

- Calculate the Euclidean distance d_j between the input vector i and output neuron j. That is, $d_j = \|x_i - w_j\|$ where $w_j = (w_{j1}, w_{j2}, ..., w_{jk})$, $j = 1, 2, ..., n$, and n is the number of output neurons.
- Find the output neuron $j*$ with minimum Euclidean distance between output neurons and the input vector. Mathematically, it is represented as $d_{\min} = \min_{j=1,2,...,n} d_j$.
- Update the weights of output neurons $j*$ using $w_{j^*}^{new} = w_{j^*}^{old} + \zeta \times \alpha \times (x_i - w_{j^*}^{old})$ where $\zeta = \exp(-R/\eta)$, α is the learning rate, η is the radius, and R is the closed distance.

3. Decrease the learning rate α and radius η and repeat Step 2 until the stopping criteria of the learning process are reached.

4. After the above learning process is completed, the quality of this network is evaluated using an average total distance measure which is defined as: $G_N = {d_j^p}/{n}$ where $d_j^p = \sqrt{\sum_p (x_p - w_j)^2}$ is the distance between data p and jth output node, and n is the number of data in cluster N.

5. Let D_j be the set of input vector x_i mapped onto node j. Summing up the distances for each vector element over all the vectors x_i ($x_i \in D_j$) yields a quantization error vector q_j for every node. This can be represented as:

$$q_{jl} = \sum_{x_i \in D_j} \sqrt{(w_{jl} - x_{il})^2}, \quad l = 1...k$$

3.2 Fuzzy Decision Tree (FDT)

Decision tree algorithms are supervised learning models that express knowledge rules using a tree structure. Among several decision tree algorithms, the ID3 is one of the most popular algorithms since it is efficient and easy to operate (Quinaln, 1986). However, data in the input vector to the ID3 should be in categorical format. All numerical data need to be discretized into proper categorical data format in advance. If the discretization process is not well performed, the classification quality of the ID3 will be erroneous. In addition, when a new input vector is fed to the ID3, only one branch of the tree is initialized and the end node of the active branch returns the class label for the new input vector.

Although this approach is straightforward, the class information of other branches, which are similar to the active branch, is not considered. This might significantly increase the classification error rate. To solve this difficulty, fuzzy set theory is integrated with ID3 for generating customer behaviour rules in the third stage, called the fuzzy decision tree (FDT). The rules generated using the FDT are easier to be interpreted by human beings. In addition, discretization process is not required since fuzzy membership functions will map the numerical data into proper membership value. Moreover, more than one branch of the fuzzy decision tree might be initiated. Therefore, class labels suggested by multiple branches will be fused by majority voting into a more trustable class label.

Let $E = \{\tilde{g}_j \mid \tilde{g}_j = (g_j^1, ..., g_j^n, C_j)\}$ be the set of our interesting customer group decided in the last stage where g_j^i represents the value of attribute i for customer j and C_j is the predefined class label for customer j. $A = \{a_1, ..., a_n\}$ is the attribute set where $a_i \in A$ with value [0,1]. $D_i = \{a_{i1}, ..., a_{ip_i}\}$ denotes the set of fuzzy linguistic terms for attribute a_i where a_{ip_i} can be described using membership function $\mu_{a_{ip_i}}(x)$. Thus, $\mu_{a_{ip_i}}(g_j^i)$ is the membership value for attribute i of vector j for fuzzy linguistic term a_{ip}. In addition, for each node N in the fuzzy decision tree, F^N denotes the set of fuzzy restrictions on the path leading to N. $N \mid a_{ip}$ denotes the particular child of node N created using a_i to split node N and following the branch a_{ip_i}.

The training process for the FDT algorithm contains five steps and is introduced as follows (Janikow, 1998).

1. Set all input vector set of the interesting customer group E in the root node of the tree. At node N to be expanded, compute the number of data included in the node which need to be subdivided as:

$$P^N = \sum_{k=1}^{|C|} P_k^N$$

where $|C|$ is the number of predefined classes, P_k^N is the number of vectors belonging to kth class and is computed as $P_k^N = \sum_{j=1}^{|E|} X_j^N \times 1_{\{C_j = k\}}$ where $1_{\{C_j=k\}} = 0$ if $C_j \neq k$ and $1_{\{C_j=k\}} = 1$ if $C_j = k$.

2. Compute the information gain at node N as $I^N = -\sum_{k=1}^{|C|} (\frac{P_k^N}{P^N} \times \log \frac{P_k^N}{P^N})$ where a_i means all clustering attributes which have not appeared at path F^N.

3. Compute the extended attached nodes from the above attributes $N \mid a_{ip}$. The information quantity of $a_{ip} \in D_i$ is $I^{N \mid a_{ip}}$.

4. Choose the attribute a_i^* with largest information gain. The formula to compute information quantity is $G_{ai} = I^N - I^{N|a_i}$ where $I^{N|a_i} = \sum_{p=1}^{|D_i|} (w_p \times I^{N|a_{ip}})$. The weight w_p is the proportion of examples belonging to node $N \mid a_{ip}$ and is shown as $w_p = P^{N|a_{ip}} / \sum_p P^{N|a_{ip}}$.

5. Subdivide the node N again by clustering attribute a_i^* and delete the attached nodes with few vectors.

4 A CASE STUDY

The proposed framework is implemented using the database provided by a major credit card issuer in Taiwan.

4.1 Data Extraction and Data Preprocessing

In the database, there are 314,339 activate card users who generated 2,153,062 transactions in year 2001 (time period t) and 2,561,202 transactions in year 2002 (time period t+1). Marketing managers want to concentrate on the customer behaviour of their VIP customers. The VIP selection criteria are based on the corporation regulations, credit assessment policies, and customer life value evaluations. "No delayed payment is made in recent nine months" and "lowest limit amount are paid in the past two months" are two typical criteria they set. A serial of COBOL (common business oriented language) and

JCL (job control language) programs are coded to retrieve customer profiles and customer behaviour data from the VSAM (virtual storage access method) files in OS/390 operation system of an IBM 9121main frame computer. As a result, 9,086 VIP customers are identified. In addition, these customers made 354,063 transactions in year 2001 and 440,010 transactions in year 2002.

4.2 Customer Segmentation Using LabelSOM

Based on the available data in the database, demographic attributes of gender, age, marital status, education, occupation, card holding period, and credit limit are used to describe a VIP customer profile. Therefore, seven input nodes are required for the LabelSOM neural network. In addition, a two-dimension rectangle topology is selected as output layer. Since the clustering quality of the LabelSOM might be affected by different parameter settings, a number of experiments are conducted based on literature suggestion (Vesanto and Alhoniemi, 2000; Zhang and Li, 1993) and our own experience. Table 1 shows the primary parameter settings in our experiments.

Table 1: The parameters of LabelSOM.

Parameter	Value
Number of Input nodes	7
Topology	Two-dimension rectangle
Number of output nodes	3~7 for each side
Learning coefficient	0.3~0.9
Neighbourhood radius	1~35
Epoch	9,086
Training number	272,580

After a systematic experimental design, the best clustering result with minimal distance 0.346 is found. The best clustering result consists of five VIP customer groups where the number of customers and the average total distance for each group are shown in Table 2. Table 3 shows the quantization error vectors, introduced in Section 3.1, of all attributes for the five groups. The smaller vector value indicates that the attribute is more important for distinguishing data among clusters.

Table 2: The grouping results for the best experiment trial.

Group No.	Number of Customers	Average Total Distance
1	1697	0.263
2	3279	0.429
3	1594	0.369
4	1379	0.379
5	1137	0.294

Table 3: The quantization error vectors for all attributes.

Group No.	Attributes						
	Gender	Martial Status	Edu.	Occp.	Age	Holding period	Credit Limit
1	0	0	0	6.376	6.929	8.385	1.380
2	0	0	18.140	8.852	9.894	13.439	2.698
3	0	0	11.354	5.714	5.558	7.639	1.141
4	0	0	11.030	5.739	4.911	7.114	1.209
5	0	0	3.945	5.464	5.314	7.915	1.307

4.3 Customer Behaviour Pattern Generation Using FDT

After discussion with marketing managers, they are interesting in the customers in VIP Group 2. Table 4 shows the comparison between the customers in Group 2 and all customers.

Table 4: The comparisons between Group 4 and all VIP customers.

	Group 2	Average for all VIP
Gender	Male: 100%	Female: 49% Male: 51%
Martial Status	Married: 100%	Married: 67% Single: 33%
Edu.	High School: 8% Undergraduate: 54% Graduate: 38%	High School: 5% Undergraduate: 47% Graduate: 48%
Occup.	Self or Intl. Business: 88% Finance or Service: 10% Others: 2%	Self or Intl. Business: 88% Finance or Service: 10% Others: 3%
Age	20~29: 0.4% 30~39: 24% 40~49: 44% 50~59: 25%	20~29: 5% 30~39: 39% 40~49: 36% 50~59: 16%
Holding Period	3~7 Year: 59%	3~7 Year: 65%
Credit Limit	Below 100K: 10% 100K~200K: 27% 200K~300K: 37% 300K~400K: 14%	Below 100K: 17% 100K~200K: 37% 200K~300K: 31% 300K~400K: 9%

For managerial reasons, each customer was classified as one of the four types (Type 1 to 4) according to their RFM scores in year 2001 (time period *t*). Marketing managers want to know whether this classification method is still valid in year 2002 (time period *t*+1). Therefore, a classification model for year 2001 needs to be constructed first.

There are two factors that might affect the inference result of the FDT algorithm. They are the number of linguist terms for each variable, and the shape of membership functions for each linguistic term. To understand the influence, the following experiments are conducted.

Assume that a trapezoid membership function, which can be described as $T(a, b, c, d, e)$, is to represent an numeric interval in this study where a is the left-bottom corner point, b is the left-top corner point, c is the middle point between b and d, d is the right-top corner point, and e is the right-bottom corner point. If we adjust top and/or bottom widths of $T(a, b, c, d, e)$, the fuzzy degree will be different. Therefore, an experiment *Shape*(x, y) denotes that a trapezoid fuzzy number $T(a, b, c, d, e)$ is modified as T ($a\times(1-y)$, $c-(c-b)\times x$, c, $c+(d-c)\times x$, $e\times(1+y)$). For example, Figure 2(a) shows a linguistic term "Medium" with a crisp membership function T(210000, 210000, 250000, 290000, 290000) for a "credit limit" attribute, while Figure 2(b) shows the fuzzy membership functions T(199500, 240000, 250000, 260000, 304500) after *Shape*(25%, 5%) is applied to T(210000, 210000, 250000, 290000, 290000).

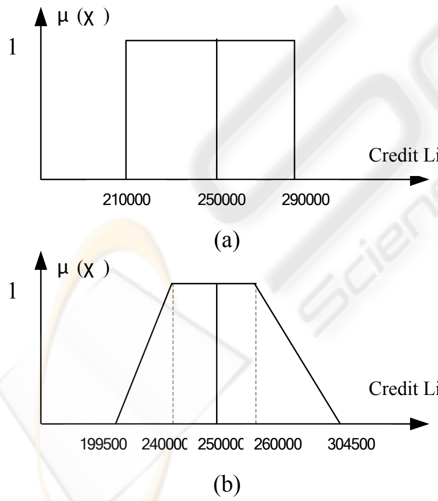

Figure 2: Crisp and fuzzy membership functions.

Table 5 shows examples of our experiment result when an attribute is represented using 3, 5, and 7 linguistic terms respectively. For each linguistic term, one crisp and four fuzzy membership functions with different bottom widths are experimented. It is clear that the classification accuracy using fuzzy membership functions are higher than the one using crisp membership function in all cases. In addition, the classification accuracy using 5 linguistic terms is higher than ones using 3 or 7. It also indicates that when the bottom with of the trapezoid fuzzy number increases, a more accurate classification result can be obtained. Table 6 shows the experiment results when the top widths of fuzzy membership functions change.

Table 5: The classification accuracy using different number of linguist terms and membership functions (I).

Number of Linguistic Terms	Membership Functions	Classification Accuracy
3	crisp	59%
	Shape(25%, 5%)	68%
	Shape(25%, 10%)	69%
	Shape(25%, 15%)	72%
	Shape(25%, 20%)	72%
5	crisp	63%
	Shape(25%, 5%)	70%
	Shape(25%, 10%)	74%
	Shape(25%, 15%)	74%
	Shape(25%, 20%)	77%
7	crisp	62%
	Shape(25%, 5%)	69%
	Shape(25%, 10%)	69%
	Shape(25%, 15%)	68%
	Shape(25%, 20%)	68%

Table 6: The classification accuracy using different number of linguist terms and membership functions (II).

Number of Linguistic Terms	Membership Functions	Classification Accuracy
3	*Shape*(20%, 10%)	68%
	Shape(25%, 10%)	69%
	Shape(30%, 10%)	68%
	Shape(35%, 10%)	68%
5	*Shape*(20%, 10%)	72%
	Shape(25%, 10%)	74%
	Shape(30%, 10%)	72%
	Shape(35%, 10%)	72%
7	*Shape*(20%, 10%)	68%
	Shape(25%, 10%)	69%
	Shape(30%, 10%)	68%
	Shape(35%, 10%)	68%

4.4 Change Analysis

Based on the experiment result of Table 5 and Table 6, managers decide to use the FDT for the following analysis where the number of linguist terms for each variable is 5 and the shape of membership functions for each linguistic term is fuzzy(25%, 20%). Under these settings, the FDT generate 109 rules and has the highest classification accuracy. Among these rules, 16 rules are to identify Customer Type 1, 30 rules are to identify Customer Type 2, 38 rules are to identify Customer Type 3, and 25 rules are to identify Customer Type 4. Figure 3 shows some of these rules.

Rule1:
Usage=Very Low (0~17) &
Interest Amount = Very Low（0~432）&
Expenditure Amount = Very Low (0~27423)
=>Customer Type 1

Rule 2:
Usage =Very Low (0~17) &
Interest Amount=High（8929~18251）&
Expenditure Amount=Very Low (0~27423) &
Credit Limit=Low (125000~250000) &
=> Customer Type 2

Rule 3:
Usage =High (37~81) &
Expenditure Amount=High (119140~393753) &
Interest Amount=Very Low (0~432) &
Credit Limit=High (255000~360000)
=> Customer Type 3

Rule 4:
Usage =Very High（63~317）&
Expenditure Amount=Very High (316215~4944133) &
Credit Limit=High (255000~360000) &
Interest Amount=Low (0~432) &
=> Customer Type 4

Figure 3: Example rules generated by the FDT algorithm.

For example, rule 4 indicates that, in year 2001 (time period *t*), if a customer has the usage behaviour such as "Usage = Very High (63~317) AND Expenditure = Very High (316215~4944133) AND Credit Limit = Very High (310000~3600000) AND Interest Amount = Very Low (0~432)," then he/she should be "Customer Type 4". When we further check the database, 143 customers in year 2001 (time period *t*) satisfy this rule. However, when this rule applies to these 143 customers in year 2002 (time period *t*+1), only 107 customers still confirm this rule. 24 customers change to Type 3, 11 customers change to Type 2, and 1 customers change to Type 1. Table 7 summarizes basic changing information. It is surprised that all changing persons are male customers, married, and own business. The company should note the changes, since customer Type 4 is most valuable for the company.

Table 7: The changing information for Type 4 customers.

	Type (4 → 1) (1 person)	**Type (4→2) (11 persons)**	**Type (4→3) (24 persons)**
Gender	Male	Male	Male
Martial Status	Married	Married	Married
Edu.	High School and Below	High School: 2 Undergraduate: 6 Graduate: 3	High School: 4 Undergraduate: 15 Graduate: 5
Occup.	Self or Intl. Business	Self or Intl. Business	Self or Intl. Business
Age	46	30~39: 3 40~49: 4 50~59: 3 60~69: 1	30~39: 3 40~49: 8 50~59: 11 60~69: 2
Holding period	8 Year	5~10 Years	4~12 Years
Credit Limit	500K	325K~650K	320K~1850K
Interest Amount	Y2001: None Y2002: None	Y 2001: None Y2002: 1	Y2001: 1 Y2002: 3
Average Spending Amount	Y2001: 362704 Y2002: 242075	Y2001: 688302 Y2002: 303967	Y2001: 499210 Y2002:: 354909

5 CONCLUSIONS

The magnificent increase in credit card markets for e-commerce leads card issuers put more efforts to understand their usage behaviour. In reality, customer behaviours usually change over time. Some frequent patterns at one time period may not be valid for another time period. To fulfil this need, this research proposes an integrated data mining approach for credit card usage behaviour analysis.

The proposed credit card usage behaviour analysis framework consists of four major stages. The first stage is data extraction and pre-processing. In this stage, the customer profile and their transaction data at time period *t* are retrieved from databases. The second stage is to conduct customer segmentation using the LabelSOM neural network. The LabelSOM adaptively cluster customers into groups and automatically identifies critical demographic features for each group. In the third stage, the usage behaviour of the customers in the interesting group is generated using fuzzy decision tree (FDT) algorithm that represents usage behaviour as a set of IF-THEN rules After obtaining the usage patterns of interesting customer group at time period *t*, we can trace the behaviour changes of these customers from time period *t* to *t*+1 when retrieving their corresponding data at time period *t*+1. The proposed model has been successfully implemented using real credit card data provided by a commercial bank in Taiwan. The provided analysis procedure should provide card issuers a systematic approach to set up marketing strategies for interesting customer groups. However, there are still some rooms for improvement in the future. Currently, only the fuzzy number with trapezoid shape is considered. It is suggested that automatic membership function fitting algorithms can be incorporated into the proposed framework. Besides, it will be worthwhile to explore variant customer groups and study what marketing strategies can affect their behaviour.

REFERENCES

Chen, M.-C., Chiu, A.-L. and Chang, H.-H., 2005a. Mining changes in customer behavior in retail marketing, Expert Systems with Applications, Vol. 28, No. 4, pp. 773-781.

Chen, R., Chen, T., Chien, Y., and Yang, Y., 2005b. Novel questionnaire-responded transaction approach with SVM for credit card fraud detection, Lecture Notes in Computer Science (LNCS), Vol. 3497, pp. 916-921.

Donato, J. M., Schryver, J. C., Hinkel, G. C., Schmoyer, R. L., Leuze, M. R., and Grandy, N. W., 1999. Mining multi-dimensional data for decision support, IEEE Future generation Computer Systems, Vol. 15, No. 3, pp. 433-441.

Dong, G., Li, J., 1999. Efficient mining of emerging patterns: discovering trends and differences, In Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining, pp. 43-52.

Giudici, P., Passerone, G., 2002. Data mining of association structures to model consumer behavior, Computational Statistics and Data Analysis, Vol. 38, No. 4, pp. 533-541.

Han, J., Dong, G. and Yin, Y., 1999. Efficient mining of partial periodic patterns in time series database. In Proceedings of the Fifteenth International Conference on Data Engineering, pp. 106-115.

Janikow, C. Z., 1998. Fuzzy decision trees: issues and Methods, IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, Vol. 28, No. 1, pp. 1-14.

Lee, T.-S., Chiu, C.-C., Chou, Y.-C. and Lu, C.-J., 2006. Mining the customer credit using classification and regression tree and multivariate adaptive regression splines, Computational Statistics & Data Analysis, Vol. 50, No. 4, pp. 1113-1130.

Kohonen, T., 1990. The self-organizing map, In Proceedings of the IEEE, Vol. 78, No. 9, pp. 1464-1480.

Kou, Y., Lu, C.-T., Sirwongwattana, S., Huang, Y.-P., 2004. Survey of fraud detection techniques, In Proceedings of IEEE International Conference on Networking, Sensing and Control, Vol. 2, pp. 749- 754.

Liu, B., Hsu, W., 1996. Post-analysis of learned rules, In Proceedings of the Thirteen National Conference on Artificial Intelligence, pp. 220-232.

Liu, B., Hsu, W., Ma, Y. and Chen, S., 1999. Mining interesting knowledge using DM-II, In Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining, pp. 430-434.

Quinaln, J. R., 1986. Induction of decision trees. Machine Learning, Vol. 1, No. 1, pp. 81-106.

Rauber, A. and Merkl, D., 1999. The SOMLib digital library system. In Proceedings of the Third European Conference, ECDL'99, pp. 323.

Tsai, C.-Y., Chiu, C.-C., 2004. A purchase-based market segmentation methodology, Expert Systems with Applications, Vol. 27, No. 2, pp. 265-276.

Tsai, C.-Y., Wang, J.-C., Chen, C.-J., 2007. Mining usage behavior change for credit card users, WSEAS Transactions on Information Science and Applications, Vol. 4, No. 3, pp. 529-536.

Vesanto, J., Alhoniemi, E., 2000. Clustering of the Self-Organization Map, IEEE Transactions on Neural Networks, Vol. 11, pp. 568-600.

Wu, J. and Lin, Z., 2005. Research on customer segmentation model by clustering. In Proceedings of the 7th international Conference on Electronic Commerce (ICEC '05), pp. 316 – 318.

Zhang, X., Li, Y., 1993. Self-organizing map as a new method for clustering and data analysis, In Proceedings of International Joint Conference on Neural Networks, pp. 2448-2451.