A Change Detection Model for Credit Card Usage Behavior

CHIEH-YUAN TSAI*, JING-CHUNG WANG, CHIH-JUNG CHEN
Department of Industrial Engineering and Management, Yuan-Ze University
135 Yuan-Tung Rd., Chung-Li, Taoyuan
TAIWAN, R.O.C.

Abstract: In recent year, credit card has been one of the most attractive financial products all over the world. The magnificent increase in credit card market leads card issuers put more attentions on how to understand their customers. This research proposes a detection model to identify usage behavior change patterns of credit card customers in two time periods. In the model, customer profiles and purchase transactions of two time periods are retrieved from card issuer databases. Then, a usage behavior rule set for each time period is generated using Apriori association rule algorithm. Finally, significant usage behavior change patterns are identified through rule set comparison using defined similarity and difference measures. The proposed model has been successfully implemented using real credit card data provided by a commercial bank in Taiwan. Several marketing strategies are suggested according the analysis finding. With the proposed model, card issuers can detect critical usage behavior changes and allocate their limited resource to establish suitable marketing strategy for their customers.

Key-words: Data Mining, Change Detection, Association Rule, Credit Card Usage, and Customer Relationship Management.

1 Introduction

According to the Nilson Report, purchases of goods and services billed to credit card issued in the U.S. reached $1.472 trillion in 2003. Credit card and debit card purchase volume in 2002 was $1.852 trillion, up 9.8% from 2001 [1]. Virtually, credit cards have be used everywhere from vacations, business trips, grocery stores, and restaurants. Whether you call them bank cards, gas cards, retail cards, travel and entertainment cards, or simply plastic cards, there is no doubt that the cards have revolutionized the business model and become an essential element of our daily life.

One strategy to increase the profit of card issuers is to intensify their competition in market through increasing satisfaction, retention, and loyalty of customers [2] [3] [4]. That is, card issuers emphasize on understanding what their customers purchase, when they use the card, and how often they consume. When the usage behavior information is available, the card issuers can encourage customers use their cards more frequently through offering better products and services.

Data mining is the technique to discover meaningful patterns (rules) and construct models from large databases. Much of existing data mining research has focused on devising techniques to build accurate models and to discover rules. Relatively little attention has been made to mining changes in databases collected over time [4] [5] [6].

Emerging pattern mining is the process to discover significant changes or differences from one database to another [7]. 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 [8]. Liu et al. [9] 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. [10] 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.

This research proposes a detection model to identify credit card usage behavior changes in two time periods. Three change patterns (emerging pattern, unexpected change, added/perished rule) are defined and detected using designated similarity and difference measures. The rest of this paper is organized as follows. Section 2 introduces the framework of the proposed change detection model. Section 3 provides an implementation case using a real credit card database provided by a commercial bank in Taiwan to demonstrate the benefit of the proposed model. A summary and future works are concluded in Section 4.

2 A behavior change detection model

The framework of the proposed change detection model for credit card usage behavior is divided into three stages as shown in Figure 1.

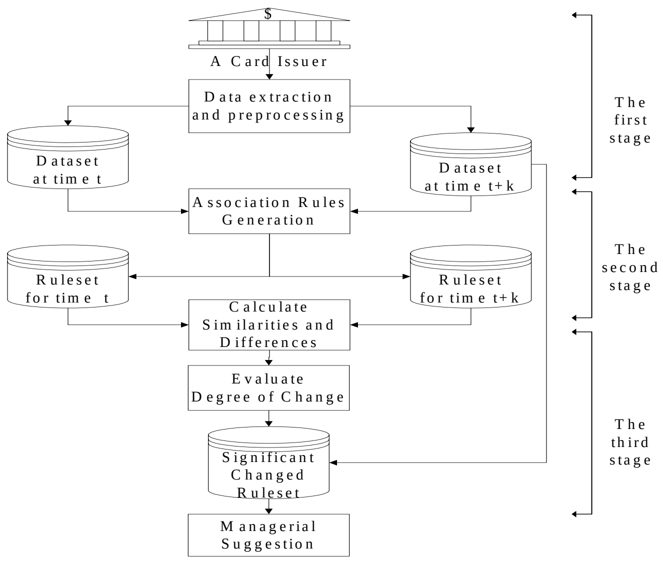

Fig. 1: The framework of the proposed system.

2.1 Data extraction and preprocessing

The first stage is to retrieve the customer profile and their transaction data from databases. The transaction data is divided into two datasets according to the transaction date we want to study. To understand the usage behavior of credit card customers, the data related to customer profile and consumer behavior need to be extracted. Practically, demographic attributes such as *Age*, *Gender*, *Marriage*, *Education*, *Occupation*, *Address*, *Credit Status* and are often used to describe a customer profile, while the transitional data attributes such as *Transaction Date*, *Store Address*, *merchant category codes* (*MCC*), and *Transaction Amount* are used to describe consumer behavior in credit card business. Notes that, continuous values in the data such as *Age*, *Income*, and *Transaction Amount* should be discretized to appropriate categorical values to facilitate the association rules generation in the next stage.

Let $I=\{i_1,i_2,\cdots,i_m\}$ be a set of items where each item is an attribute-value pair. The pair could be a customer profile or consumer behavior attribute with respective categorical value. Let the dataset $D=\{t_1,t_2,\cdots,t_n\}$ be a set of records and $t_i=\{I_{i1},I_{i2},\cdots,I_{ik}\}$ be a set of attribute-value pair items for a customer i where $I_{ij}\in I$. Since this research intends to detect the usage behavior change patterns in two time periods, the dataset D is separated to datasets D^t and D^{t+k} where D^t and D^{t+k} represent the datasets storing the records transacted in time period t and t+k respectively.

2.2 Association rules generation

The second stage is to generate rule sets from the two datasets using association rule technique. Each dataset (D^t or D^{t+k}) can be used to generate a rule set for its time period (t or t+k) using association rule technique. A rule set is a collection of association patterns describing interesting causal relationship between customer profile and consumer behavior. An association rule is an implication of the form $LHS\Rightarrow RHS$ where $LHS, RHS\subset I$ and $LHS\cap RHS=\phi$. The support for an association rule $LHS\Rightarrow RHS$ is the percentage of records in the database that contains $LHS\cup RHS$. If the support is high, the itemset $LHS\cup RHS$ is worth to put into discussion. Notes that a set of items is referred to an itemset. The confidence for an association rule $LHS\Rightarrow RHS$ is the ratio of the number of records that contain $LHS\cup RHS$ to the number of records that contain *LHS*. Therefore, the goal of the association rule problem is to find out all association rules $LHS\Rightarrow RHS$ with a minimum support ($s_{\min}$) and a minimum confidence ($\alpha_{\min}$) where the $s_{\min}$ and $\alpha_{\min}$ are specified by a user.

A well-known association rule algorithm, called Apriori algorithm [11], is utilized in this research. The algorithm uses a so-called large itemset property to reduce the computational time. The property states that any subset of a large itemset must be large. After the large itemsets have been found, it is straightforward to generate strong association rules, where strong association rules satisfy both minimum support and minimum confidence.

2.3 Behavior change detection

In the third stage, the two rule sets are compared using similarity and difference measures. After obtaining the rule set R^t from time period t and R^{t+k} from time period $t+k$, the next stage is to find out the usage behavior changes between the two time periods. In the following description, r_i^t represents the ith rule in R^t and r_j^{t+k} be the jth rule in R^{t+k}. $Sup^t(r_i)$ represents the support of r_i in time t.

In this research, three usage behavior change patterns are defined. First, r_j^{t+k} is considered as an ***emerging pattern*** respective to r_i^t when both *LHS* and *RHS* of r_i^t and r_j^{t+k} are the same but there is a significant difference in $Sup^t(r_i)$ and $Sup^{t+k}(r_j)$ [7]. Second, r_j^{t+k} is an ***unexpected change*** with respect to r_i^t, if *LHSs* of r_i^t and r_j^{t+k} are similar but their *RHSs* are quite different [8]. Third, r_j^{t+k} is an ***added rule*** if all the *LHSs* and *RHSs* are quite different from any of r_i^t in R^t, and r_i^t is a ***perished rule*** if all the *LHSs* and *RHSs* are quite different from any of r_j^{t+k} in R^{t+k} [12]. Whether rules are similar or quite different can be judged by a *Rule Matching Threshold* (*RMT*) [13]. The value is subjectively provided by a user. The relation between *RMT* and the three change patterns is illustrated in Figure 2.

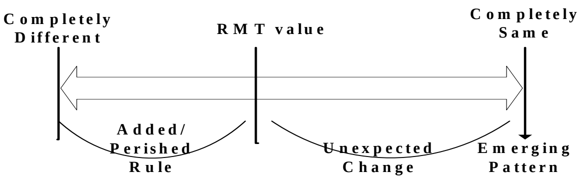

Fig. 3: The three usage behavior change patterns.

For better explanation, some notations and variables are defined:

- s_{ij} is the similarity measure between r_i^t and r_j^{t+k} where 0 s_{ij} 1.
- δ_{ij} is the difference measure between r_i^t and r_j^{t+k} where -1 δ_{ij} 1.
- ℓ_{ij} is the degree of attribute match in the *LHS* part, where $l_{ij} = |A_{ij}| / \max(|X_i^t|, |X_j^{t+k}|)$.
- c_{ij} is the degree of attribute match in the *RHS* part.
- $|A_{ij}|$ is the number of the same attributes in the *LHS* parts of r_i^t and r_j^{t+k}.
- $|X_i^t|$ is the number of attributes in the *LHS* parts of r_i^t.
- $|X_j^{t+k}|$ is the number of elements in the *LHS* parts in r_j^{t+k}.
- x_{ijk} is the degree of value match of the kth matching attribute in the *LHS* parts.
- y_{ij} is the degree of value match in the *LHS* parts.

The change pattern can be detected through the following steps [13]:

1. First, calculate the maximum similarity value for each rule in time period t and $t+k$. The maximum similarity value of r_i^t is defined as $s_i = \max(s_{i1}, s_{i2}, \ldots, s_{i|R_{t+k}|})$ and the maximum similarity value of r_j^{t+k} is defined as $s_j = \max(s_{1j}, s_{2j}, \ldots, s_{|R_t|j})$, where s_{ij} is the similarity measure between rule r_i^t and r_j^{t+k}. The description of s_{ij} is defined as

$$s_{ij} = \begin{cases} \dfrac{\ell_{ij} \times \sum_{k \in A_{ij}} x_{ijk} \times c_{ij} \times y_{ij}}{|A_{ij}|}, & if\ |A_{ij}| \neq 0 \\ 0 & ,if\ |A_{ij}| = 0 \end{cases}$$

If both *LHS* and *RHS* is the same, the value of similarity will be 1. Others will have the value between 0 and 1.

2. Second, for each rule r_i^t, calculate the difference measures δ_{ij} and modified difference measures $\delta_{ij}^{'}$ between r_i^t and r_j^{t+k}. The difference measure δ_{ij} can be defined as:

$$\delta_{ij} = \begin{cases} \dfrac{\ell_{ij} \times \sum_{k \in A_{ij}} x_{ijk}}{|A_{ij}|} - y_{ij}, if |A_{ij}| \neq 0, c_{ij} = 1 \\ -y_{ij} \qquad ,if |A_{ij}| = 0, c_{ij} = 1 \end{cases}$$

The modified difference measure can be defined as:

$$\delta_{ij}^{'} = |\delta_{ij}| - k_{ij},\ k_{ij} = \begin{cases} 1, & \text{if} \quad \max(s_i, s_j) = 1 \\ 0, & \text{otherwise} \end{cases}$$

3. Third, the usage behavior change patterns are judged using the maximum similarity values (s_i and s_j), difference measures (δ_{ij}), and modified difference measures ($\delta_{ij}^{'}$). If $\delta_{ij} = 0$ $(\sum_{k \in A_{ij}} x_{ijk} > 0\ or\ y_{ij} > 0\ or\ \ell_{ij} > 0)$, then the change type is an ***emerging pattern***. If $\delta_{ij} > 0, \delta_{ij}^{'} \geq RMT$, then the change type is an ***unexpected consequent change***. If $\delta_{ij} < 0, \delta_{ij}^{'} \geq RMT$, then the change type is an ***unexpected condition change***. If $s_j < RMT$, then

the change type is an ***added rule***. If $s_i < RMT$, then the change type is a ***perished rule***.

Notes that the number of changes could be large in most cases. Therefore, only those patterns whose degrees of change are no less than a minimum degree of change specified by users $\alpha_{\min}$, called ***significant changes***, need to be examined. The degree of change for each change pattern can be shown in Table 1.

Table 1: The change degree for different change types.

Type of changes	Measures
Emerging Patten	$(Sup^{t+k}(r_i) - Sup^t(r_i)) / Sup^t(r_i)$
Unexpected change	$Sup^{t+k}(r_{i\cap j}) / Sup^{t+k}(r_j)$
Added rule	$(1 - s_i) \times Sup^t(r_i)$
Perished rule	$(1 - s_j) \times Sup^{t+k}(r_j)$

3 A case study

The proposed framework is implemented using the data provided by a major credit card issuer in Taiwan. The first time period for this study is year 2001 and the second period is year 2002. There are totally 464,621 active and inactive credit card issued until 2002.

3.1 Data preprocessing

After a comprehensive discussion with the marketing managers of the bank, the data selection criteria are set such as “card issued for more than 6 months,” “no delay payment in the last two months,” “the total spending greater than 200,000 NT in the last 12 months” and so on. In addition, twelve critical customer profile and consumer behavior attributes are identified. They are *Customer ID*, *Gender*, *Age*, *Marriage*, *Education*, *Occupation*, *Location*, *Credit Status*, *Yearly Usage Frequency*, *Yearly Spending Amount*, *Merchant Category Code*, and *Spending Amount*. A serial of COBOL (common business oriented language) and JCL (job control language) programs are coded to retrieve customer profiles and consumer behavior data from the VSAM (virtual storage access method) files in OS/390 operation system of an IBM 9121main frame computer.

There are 10,940 customers who satisfy the selection criteria in both year 2001 and 2002. 509,066 transaction records are found in year 2001 and 467,504 transaction records in year 2002 for those customers. The discretization is processed iteratively based on the data distribution of each attribute.

3.2 Association rules for the two time periods

Depending on practical need, a user can decide an appropriate minimum support value ($s_{\min}$), minimum confidence value ($\alpha_{\min}$), and the largest itemsets length. If the minimum support is high, the number of association rules will be low. Alternatively, if the minimum support is low, the number of association rules will be many. For instance, in year 2001 there are 63,288 association rules generated when minimum support is 1% and 6,312 when minimum support is 9% (while minimum confidence is 70% and the largest itemset is 4). It is found that the number of rules decreases dramatically when the minimum support increases. In the following discussion, the minimum support is set as 1%, minimum confidence as 70%, and the largest itemset as 4. With the setting, we get 63,288 association rules in year 2001 and 56,571 association rules in year 2002. After deleting the meaningless rules, 1,986 association rules in year 2001 and 1,622 association rules in year 2002 are left for the following change detection analysis.

3.3 Change detection analysis

As described in section 2.3, the number of unexpected changes rules and added/perished rules is affected by the *RMT*. For example, 33 unexpected changes and 11,724 added/perished rules are identified if the *RMT* is set as 0.4, while 6,323 unexpected changes and 614 added/perished rule are found if the *RMT* is 0.8. Noted that the number of emerging patterns is not affected by the *RMT* value.

Due to the number of rules for the three change types is large, we need to screen out some changes that might not be significant enough to pay attention on it. That is, the degree of change greater than minimum degree of change ($\alpha_{\min}$)will be further considered. For example, if we set $\alpha_{\min}$ =0, then 8,122 significant emerging patterns, 759 significant unexpected changes, and 264 significant added/perished rules will be significant.

■ Significant emerging patterns

Significant emerging patterns are the rules that appear in year 2001 and 2002 but their support levels are quite different. There are 8,122 significant patterns found if we set $\alpha_{\min}$ =0. Three of them are summarized in Table 2. As shown in Table 2, the support levels of Pattern 1 to Pattern 3 increase from about 1 % in 2001 to 45% in 2002. It is about 40 times increase for those three patterns. With those patterns, we know that male customers who live in Taipei, Taichung, or Kaoshiung (top three largest cities in Taiwan) with college or lower education level, dramatically increase their purchase in auto service and gas station from year 2001 to year 2002. We further query the number of customers that satisfy the three patterns and find that there are 3,211, 1,508, and 1,606 persons satisfying the Pattern 1 to 3 respectively. If the bank wishes to promote "auto service and gas station" program to its customers, up to 42.1% of mailing fee can be saved by just sending the mails to the interesting group instead of all customers.

Table 2: Significant emerging patterns.

Pattern	Year 2001 (or 2002)	Year 2001 Sup	Year 2002 Sup	$\alpha_{i,j}$
1	Gender=Male, City=Taipei, Education=College → Behavior=Auto Service & Gas Station	1.0030	45.8268	44.69
2	Gender=Male, City=Taichung, Education=High School & Others → Behavior= Auto Service & Gas Station	1.0122	45.8268	44.27
3	Gender=Male, City=Kaoshiung, Education=High School & Others → Behavior= Auto Service & Gas Station	1.1227	45.8268	39.81

■ Significant unexpected changes

The second type of usage behavior patterns is the unexpected change. Three typical significant unexpected changes are illustrated in Table 3. In the table, we found that both Change 1 and Change 2 are unexpected consequent changes due to δ_{ij} = 1. That is, both rules are the same in *LHS* but different in *RHS*. For Change 1, we discovered that single women with high school education level increase their spending from 0-70,000 in year 2001 to 70,001-140,000 in year 2002. For Change 2, women who work in banking with age 31-40 have their credit card usage from entertainment in year 2001 to department store & duty free shop in year 2002. When we further study the Change 2 case, it is found that more than 1/5 of customers in database (2819/10940) conform to the *LHS* of rules. Therefore, it is strongly suggested that marketing persons should be aware of this significant unexpected change.

Table 3: Significant unexpected changes

Change	Year 2001	Year 2002	δ_{ij}	δ'_{ij}	α_{ij}
1	Education=High School & Others, Marriage=No, Gender=Female → Yearly Spending Amount=0-70,000 Sup 1.8312	Education=High School & Others, Marriage=No, Gender=Female → Yearly Spending Amount = 70,001-140,000 Sup 17.5301	1	1	29
2	Occupation=Banking, Age= 31-40, Gender=Female → Behavior=Entertainment Sup 1.3067	Occupation=Banking, Age= 31-40, Gender=Female → Behavior= Department Store & Duty Free Shop Sup 2.1923	1	1	17
3	Gender=Male, Occupation=Own Business, City=Kaoshiung → Behavior=Entertainment Sup 16.8952	Gender=Male, Occupation=Own Business, City=Taipei → Behavior=Entertainment Sup 9.8812	-1	1	21

■ Significant Added/Perished Rules

The third type of usage behavior change is added rules or perished rules. Three typical significant added rules and three significant perished rules are illustrated in Table 4 and Table 5. We found that the supports for those added or perished rules are relatively low. The Rule 1 in Table 4 revealed that customers living in Hsin-Ju with college education consume on home decoration and improvement in Year 2002. Alternatively, the Rule 1 in Table 5 is a significant perished rule indicating that customers who have own-business and lives in Pintung city consume in department store and duty free shop in Year 2001. It is suggest that marketing persons should note the reasons why those rules are added in year 2002 or no longer valid in year 2001.

Table 4: Significant added rules.

Rule	Year 2002 (Added Rule)	Sup	α_{ij}
1	Education=College, City=Hsinchu → Behavior=Home Decoration & Improvement	1.0030	25
2	Occupation=Service Industry, Age=41-50 → Behavior= Restaurant & Pub	1.0490	21
3	City=Taipei, Credit Limit=200,000-250,000, Gender=Female → Behavior=Bookstore & Musical Store	1.8220	17

Table 5: Significant perished rules.

Rule	Year 2001 (Perished Rule)	Sup	α_{ij}
1	Occupation=Own Business, City=Pingtung → Behavior= Department Store & Duty Free Shop	1.1319	23
2	Marriage=Yes, City=Changhua → Behavior= Department Store & Duty Free Shop	1.0214	19
3	City=Taipei, Age=41-50, Credit Limit=0 → Behavior=Home Decoration & Improvement	1.0030	11

4 Conclusions

In recent year, credit cards have been one of the most popular financial products. The magnificent increase in credit card markets leads card issuers put more efforts to understand their usage behavior. To fulfill this need, this research proposes a detection model to identify usage behavior change patterns of credit card customers in two time periods. First, customer profiles and transaction data are retrieved, preprocessed, and stored as two datasets. Then, association rule sets are generated for each dataset using Apriori algorithm. The rules in the two time periods are classified as emerging pattern, unexpected change, or added/perished rule using defined similarity and difference measures. With the proposed model, card issuers can detect critical usage behavior changes and allocate their limited resource to establish a more suitable marketing strategy for their customers.

The proposed model has been successfully implemented using real credit card data provided by a commercial bank in Taiwan. Several marketing strategies have been suggested in this paper according to the analysis finding. However, there are still some rooms for model improvement in the future. Currently, discretization process in the stage two is conducted manually by marketing persons of the card issuer. The decision for setting appropriated numerical interval could be very subjective. It is suggested that automatic discretization algorithms can be applied to improve the efficiency of the model. Besides, it is worthwhile to explore variant customer groups and study how different marketing strategies will affect their behavior.

References:

[1] The Nilson report, U.S. credit and debit cards projected thru 2007, October 2003, pp. 797.

[2] J. Peppard, Customer Relationship Management (CRM) in financial Services, *European Management Journal*, Vol. 18, No. 3, 2000, pp. 312-327.

[3] D. Peppers, M. Rogers, and R. Dorf, Is your company ready for one-to-one marketing, *Harvard Business Review*, Jan-Feb 1999, pp. 151-160.

[4] P. Giudici, and G. Passerone, Data mining of association structures to model consumer behavior, *Computational Statistics and Data Analysis*, Vol. 38, 2002, pp. 533-541.

[5] J. Han and M. Kamber, *Data mining: concepts and techniques*, San Francisco, Morgan Kaufmann Publishers, 2001.

[6] G. Nakhaeizadeh, C. Taylor, and C. Lanquillon, Evaluating usefulness of dynamic classification, *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*, 1998.

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

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

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

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

[11] R. Agrawal and R.Srikant, Fast algorithm for mining association rules in large databases, *Proceeding of Int'l Conf. VLDB*, 1994, pp. 487-499.

[12] C. Lanquillon, Information filtering in changing domains, *Proceedings of the International Joint Conference on Artificial Intelligence*, 1999, pp. 41-48.

[13] H. S. Song, J. K. Kim, and S. H. Kim, Mining the change of customer behavior in an internet shopping mall, *Expert Systems with Applications*, Vol. 21, 2001, pp. 157-168.