

# Mining changes in customer behavior in retail marketing

Mu-Chen Chen<sup>a,\*</sup>, Ai-Lun Chiu<sup>b</sup>, Hsu-Hwa Chang<sup>c</sup>

<sup>a</sup>*Department of Business Management, National Taipei University of Technology, No. 1, Section 3, Chung-Hsiao E. Road, Taipei 106, Taiwan, ROC*

<sup>b</sup>*SAS, Taipei, Taiwan, ROC*

<sup>c</sup>*Department of Business Administration, National Taipei College of Business, Taipei, Taiwan, ROC*

---

## Abstract

During the past decade, there have been a variety of significant developments in data mining techniques. Some of these developments are implemented in customized service to develop customer relationship. Customized service is actually crucial in retail markets. Marketing managers can develop long-term and pleasant relationships with customers if they can detect and predict changes in customer behavior. In the dynamic retail market, understanding changes in customer behavior can help managers to establish effective promotion campaigns. This study integrates customer behavioral variables, demographic variables, and transaction database to establish a method of mining changes in customer behavior. For mining change patterns, two extended measures of similarity and unexpectedness are designed to analyze the degree of resemblance between patterns at different time periods. The proposed approach for mining changes in customer behavior can assist managers in developing better marketing strategies.

© 2005 Elsevier Ltd. All rights reserved.

*Keywords:* Data mining; Association rules; Retailing; Customer behavior; Change patterns

---

## 1. Introduction

Customized service is actually crucial in retail markets. Customized service has become a key issue in developing customer relationships. Marketing managers can develop long-term and pleasant relationships with customers if they can detect and predict changes in customer behavior. In the past, researchers generally applied statistical surveys to study customer behavior. Recently, however, data mining techniques have been adopted to predict customer behavior (Giudici & Passerone, 2002; Song, Kim, & Kim, 2001).

Data mining techniques search through a database without any specific pre-determined hypothesis to obtain implicit, previously unknown, and potentially useful information including knowledge rules, constraints and regularities (Chen, Han, & Yu, 1996). Data mining is a stage in Knowledge Discovery in Databases (KDD), involving the application of specific algorithms for pattern extraction (Mitra, Pal, & Mitra, 2002). Various successful applications

have been reported in areas such as marketing, finance and banking. Applications in these domains generally involve the collection and storage of large amounts of data.

Data mining brings various techniques together to discover patterns (rules) and to construct models from databases. Currently businesses face the challenge of a constantly evolving market where customer needs are changing all the time. In such a situation, change mining can enable market analysts to better understand changes in customer needs and how those needs change. Change mining is more appropriate in dynamic business environments, and involves extensive human intervention (Song et al., 2001).

Retail market managers must not only provide high-quality products and services, but also must react appropriately to changes in customer needs. Data mining can be applied to identify useful customer behavior patterns from large amounts of customer and transaction data (Giudici & Passerone, 2002). As a result, the discovered information can be ascertained to support better decision-making in retail marketing. Data mining techniques have mostly been adopted to generate predictions and describe behaviors. Relatively little research has focused on mining changes in databases collected over time (Liu, Hsu, Han, & Xia, 2000).

---

\* Corresponding author. Tel.: +886 2 2771 2171x3417; fax: +886 2 2776 3964.

E-mail address: [bmcchen@ntut.edu.tw](mailto:bmcchen@ntut.edu.tw) (M.-C. Chen).

In the dynamic retail market, understanding changes in customer behavior can help managers to establish effective promotion campaigns (Song et al., 2001). Liu et al. (2000) devised a method of change mining in the context of decision trees for predicting changes in customer behavior. Since decision tree is a classification-based approach, it cannot detect complete sets of changes (Song et al., 2001). Association rule extraction was widely used for analyzing the correlation between product items purchased by customers, and to support sales promotion and market segmentation (Changchien & Lu, 2001; Changchien, Lee, & Hsu, 2004). Song et al. (2001) employed an approach based on association rules to identify changes in customer behavior.

Most previous customer behavior studies applied customer demographic variables to analyze customer behavior (Song et al., 2001). However, valuable customer behavioral variables, such as recency, frequency, and monetary (RFM), can be used to differentiate customer contributions to a business (Miglautsch, 2000; Stone, 1995; Suh, Noh, & Suh, 1999; Tsai & Chiu, 2004). Researchers have observed that RFM is a widely used technique for customer behavioral analysis that can effectively investigate customer values and segment markets. Customer behavioral variables, RFM, respectively, measure the recency of customer purchasing behavior, the frequency of purchasing, and the average monetary expenditure on purchasing. RFM can be transformed using customer and transaction databases.

This study attempts to integrate customer behavioral variables (RFM), demographic variables, and transaction database to establish a method of mining changes in customer behavior. Song et al. (2001) designed two measures of similarity and unexpectedness to analyze the degree of resemblance between patterns at different time periods. However, these two measures are limited to the analysis of patterns with a single attribute on the right-hand-side (consequent part) of an association rule. This study designs two modified measures of similarity and unexpectedness to overcome the above limitations. In this study, the data required for analysis are integrated and transformed from customer, product, and transaction databases. The proposed approach for mining changes in customer behavior can assist managers in developing better marketing strategies.

## 2. Mining customer behavior changes

In this study, customer behavior patterns are first identified using association rule mining. Following the association rules of customer behavior are discovered, the changes in customer behavior are identified by comparing two sets of association rules generated from two datasets of different periods. Based on previous studies, changes in customer behavior include emerging patterns, added patterns, perished patterns, and unexpected patterns (Dong & Li, 1999; Liu & Hsu, 1996; Liu, Hsu, Mun, & Lee, 1999; Padmanabhan & Tuzhilin, 1999; Song et al., 2001). The discovered change patterns can be further explained and assessed to provide a basis for formulating marketing strategies. Fig. 1 illustrates the flowchart of change mining for customer behavior. Further details of the change mining procedure are discussed below.

### 2.1. Data pre-processing

Prior to analysis, data accuracy and consistency must be ensured to obtain truthful results. Generally, some useful variables can be hidden in a large quantity of raw data, and thus can be obtained through data integration and transformation. Customer behavioral variables (RFM) are hidden in customer and transaction databases, and can be extracted from data integration and transformation. Since the data required for analyzing association rules must be discrete, continuous variables are transformed to discrete variables, and a simple 3-4-5 rule is applied (Han & Kamber, 2001).

In RFM, recency represents the interval between the most recent transaction time of individual customers and the evaluation time (Stone, 1995). Moreover, frequency represents the number of purchases by individual customers during a specific period. Additionally, monetary represents the average expenditure of a customer during a specific period. Individual customers' recency, frequency, and monetary are scored to calculate the value of the purchasing behavior of each customer. This study adopts the RFM scoring approach of Miglautsch (2000) to transform the customer behavioral variables.

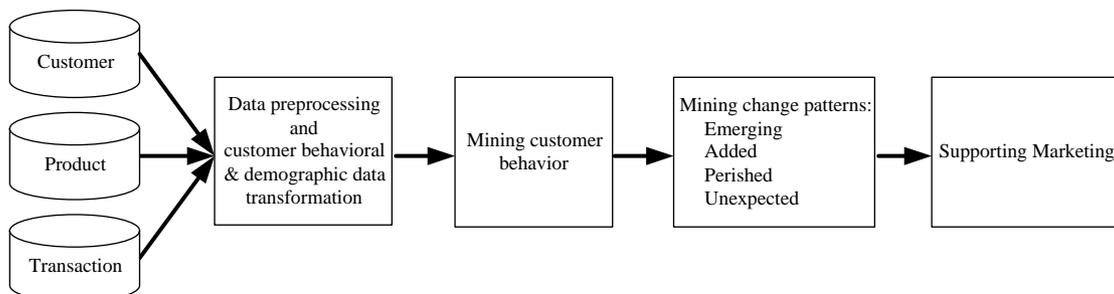


Fig. 1. Flowchart of mining changes for customer behavior.

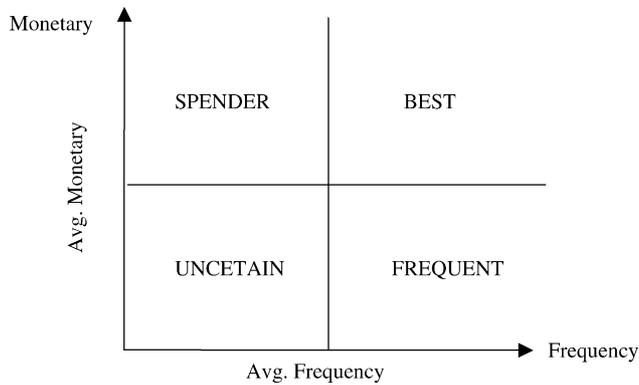


Fig. 2. BSG matrix of customer value.

## 2.2. Customer segmentation

Customers are segmented into various target markets in terms of customer value obtained by scoring RFM. Marcus (1998) conducted market segmentation based on customer values obtained from the growth matrix of Boston Consulting Group (BSG), and classified customers into four clusters in terms of average purchase expenditure (monetary) and purchase frequency. These four customer groups include best customers, frequent customers, spenders, and uncertain customers, as illustrated in Fig. 2. To develop a target market, another customer behavioral variable, recency, can be used for cross analysis to differentiate customer contributions to businesses.

## 2.3. Mining customer behavior

Association rules were initially applied to analyze the relationships of product items purchased by customers at retail stores (Agrawal, Imielinski, & Swami, 1993; Srikant, Vu, & Agrawal, 1997). In data mining, association rules are descriptive patterns of the form  $X \Rightarrow Y$ , where  $X$  and  $Y$  are statements regarding the values of attributes of an instance in a database.  $X$  is termed the left-hand-side (LHS), and is the conditional part of an association rule. Meanwhile,  $Y$  is called the right-hand-side (RHS), and is the consequent part. The most typical application of association rules is market basket analysis, in which the market basket comprises the set of items (namely itemset) purchased by a customer during a single store visit.

In customer behavior research, association rules can be applied to identify the correlations between customer profiles represented by demographic variables and purchased products by examining customer and product databases (Song et al., 2001). In Song et al. (2001), LHS of association rules chooses customer profile variables such as gender, age, yearly income, and so on; whereas RHS includes the product items bought. Customer behavioral data are generally the most effective predictive data in customer relationship management (Rud, 2001). In this study, the customer behavioral

variables (RFM) are associated with demographic variables to predict customer purchasing behavior. The association rules discovered at different periods of time are adopted for change mining to identify customer behaviors that vary over time.

Association rules can include any number of attributes on either side of the rule. Not all association rules are interesting to decision makers. Rule support and confidence are two measures of rule interestingness. An interesting rule must satisfy the minimum support and confidence determined by domain experts. In the algorithms for association rule mining, Apriori is one of the most widely used algorithms (Srikant et al., 1997). In this study, Apriori is applied to discover customer behavior patterns.

## 2.4. Change mining

In this study, measures of similarity and unexpectedness are developed for rule matching to investigate changes in customer behavior. This study first explains the changes in customer behavior, and then mathematically designs the similarity and unexpectedness measures.

### 2.4.1. Change patterns

Based on previous studies, four patterns are identified to measure changes in customer behavior (Dong & Li, 1999; Liu & Hsu, 1996; Padmanabhan & Tuzhilin, 1999; Song et al., 2001). These patterns include emerging pattern, added pattern, perished pattern, and unexpected change. These four change patterns are explained below.

**2.4.1.1. Emerging patterns.** Emerging patterns are defined as rules whose supports increase extensively between time stamped datasets (Dong & Li, 1999; Song et al., 2001). In marketing management, emerging patterns imply the same consumer behavior that exists in different periods of time with trend. The positive pattern growth rate (i.e. the support of a rule increases over time) indicates that the customer behavior becomes robust over time. Meanwhile, a pattern growth rate below zero indicates that the customer behavior is getting weak. For emerging patterns, the conditional and consequent parts are identical for  $r_i^{t_1}$  and  $r_i^{t_2}$ , but support for the two rules differs significantly between different time periods.

**2.4.1.2. Added patterns.** A rule at period  $t_2$ ,  $r_j^{t_2}$ , is identified as an added pattern if all conditional and consequent parts differ significantly from any rule,  $r_i^{t_1}$ , at period  $t_1$  (Lanquillon, 1999; Song et al., 2001). The rule matching threshold (RMT) is used to measure the degree of change.

**2.4.1.3. Perished patterns.** A rule at period  $t_1$ ,  $r_i^{t_1}$ , is identified as a perished pattern if all conditional and consequent parts differ significantly from any rule,  $r_j^{t_2}$ , at period  $t_2$  (Lanquillon, 1999; Song et al., 2001). A perished

pattern is a vanished pattern found in the past but not the present.

**2.4.1.4. Unexpected changes.** Unexpected changes can be found in previous studies on mining interesting patterns (Liu & Hsu, 1996; Silberschatz & Tuzhilin, 1996; Song et al., 2001). Liu and Hsu (1996) classified unexpected changes into unexpected conditional changes and unexpected consequent changes. If the conditional parts of  $r_i^{t_1}$  and  $r_j^{t_2}$  are similar, but their consequent parts are different, then  $r_j^{t_2}$  is an unexpected consequent change with respect to  $r_i^{t_1}$  (Liu & Hsu, 1996; Song et al., 2001). Moreover, if the consequent parts of  $r_i^{t_1}$  and  $r_j^{t_2}$  are similar, but their conditional parts are different, then  $r_j^{t_2}$  is an unexpected conditional change with respect to  $r_i^{t_1}$ . The patterns of customer behavior described in this study reflect the associations between customer attributes and their purchased products. Therefore, the conditional part of association rule (LHS) of the rule, takes the form of customer profile presented by demographic variables and customer behavioral variables (RFM). While the consequent part (i.e. RHS) denotes the product purchased by customers. In this study, the unexpected changes of customer behavior can be identified in the form of unexpected purchasing (consequent) patterns and customer shifting (conditional) patterns.

After explaining the four customer changes, this study elaborates on the measures used to detect these changes.

#### 2.4.2. Change measures

Before introducing the similarity and unexpectedness measures for mining changes in customer behavior, some notations are defined as follows.

- $R^{t_1}$  the set of association rules for  $t_1$ ;
- $R^{t_2}$  the set of association rules for  $t_2$ ;
- $r_i^{t_1}$  an association rule in  $R^{t_1}$ ,  $r_i^{t_1} \in R^{t_1}$ ;
- $r_j^{t_2}$  an association rule in  $R^{t_2}$ ,  $r_j^{t_2} \in R^{t_2}$ ;
- $A_{ij}$  the set of attributes that simultaneously appear in conditional part (LHS) for  $r_i^{t_1}$  and  $r_j^{t_2}$ ;
- $|A_{ij}|$  the number of attributes in  $A_{ij}$ ;
- $B_{ij}$  the set of attributes that simultaneously appear in consequent part (RHS) for  $r_i^{t_1}$  and  $r_j^{t_2}$ ;
- $|B_{ij}|$  the number of attributes in  $B_{ij}$ ;
- $|M_i^{t_1}|$  the number of attributes in LHS for  $r_i^{t_1}$ ;
- $|M_j^{t_2}|$  the number of attributes in LHS for  $r_j^{t_2}$ ;
- $|N_i^{t_1}|$  the number of attributes in RHS for  $r_i^{t_1}$ ;
- $|N_j^{t_2}|$  the number of attributes in RHS for  $r_j^{t_2}$ ;
- $\ell_{ij}$  the similarity of attributes in LHS for  $r_i^{t_1}$  and  $r_j^{t_2}$ ;
- $h_{ij}$  the similarity of attributes in RHS for  $r_i^{t_1}$  and  $r_j^{t_2}$ ;
- $X_{ijp}$  a binary variable, where  $X_{ijp} = 1$ , if the  $p$ th attribute in  $A_{ij}$  has the same value for  $r_i^{t_1}$  and  $r_j^{t_2}$ , and otherwise  $X_{ijp} = 0$ ,  $p = 1, 2, \dots, |A_{ij}|$ ;
- $Y_{ijq}$  a binary variable, where  $Y_{ijq} = 1$ , if the  $q$ th attribute in  $B_{ij}$  has the same value for  $r_i^{t_1}$  and  $r_j^{t_2}$ , and otherwise  $Y_{ijq} = 0$ ,  $q = 1, 2, \dots, |B_{ij}|$ ;
- $S_{ij}$  a measure of the similarity between  $r_i^{t_1}$  and  $r_j^{t_2}$ ;

- $S_i^{t_1}$  the maximum similarity for  $r_i^{t_1}$ ;
- $S_j^{t_2}$  the maximum similarity for  $r_j^{t_2}$ ;
- $\delta_{ij}$  a measure of the unexpectedness between  $r_i^{t_1}$  and  $r_j^{t_2}$ ;
- $\delta'_{ij}$  an adjusted measure of the unexpectedness between  $r_i^{t_1}$  and  $r_j^{t_2}$ ;
- $k_{ij}$  a binary variable,  $k_{ij} = 1$ , if  $\max(S_i^{t_1}, S_j^{t_2}) = 1$ ; otherwise,  $k_{ij} = 0$

To identify the degree of similarity and difference in customer behavior changes for different periods of time, this study designs two measures of similarity and unexpectedness. Similarity can be used to measure the degree of likeness between two rules, and unexpectedness can be used to identify the disparity between dissimilar rules. The measures of similarity and unexpectedness developed in the literature only analyze patterns with a single product item in RHS of a rule. This study designed two modified measures of similarity and unexpectedness to overcome the above limitation. These two modified measures are constructed as follows.

**2.4.2.1. Similarity.** Liu and Hsu (1996) developed the similarity measure to analyze the degree of resemblance between patterns at different periods of time. However, the application of Liu and Hsu's measure is limited to analyzing patterns with a single product item in RHS (i.e. RHS includes only one attribute). As a result, when the RHS of a pattern involves two or more products, the similarity measure of Liu and Hsu cannot be used for rule matching. This study extends the measure of Liu and Hsu measure by enabling two or more attributes (product items) in RHS to detect customer behavior changes. The modified similarity measure can be represented as:

$$S_{ij} = \begin{cases} \ell_{ij} \frac{\sum_p X_{ijp}}{|A_{ij}|} \times \frac{h_{ij} \sum_q Y_{ijq}}{|B_{ij}|}, & \text{if } |A_{ij}| \neq 0, \text{ and } |B_{ij}| \neq 0 \\ 0 & \text{if } |A_{ij}| = 0, \text{ or } |B_{ij}| = 0 \end{cases} \quad (4)$$

where  $\ell_{ij}$  and  $h_{ij}$  are defined as follows:

$$\ell_{ij} = \frac{|A_{ij}|}{\max(|M_i^{t_1}|, |M_j^{t_2}|)} \quad (5)$$

$$h_{ij} = \frac{|B_{ij}|}{\max(|N_i^{t_1}|, |N_j^{t_2}|)} \quad (6)$$

In Eqs. (5) and (6),  $\ell_{ij}$  and  $h_{ij}$  represent the similarity of the conditional and consequent parts, respectively. The degree of similarity,  $S_{ij}$ , is between 0 and 1, where 0 indicates that the two patterns are completely dissimilar, and 1 indicates that the two patterns are identical. Following calculating the similarity of patterns is calculated, the maximum similarity degrees of Rules  $r_i^{t_1}$  and  $r_j^{t_2}$  are determined to measure the change of patterns during periods  $t_1$  and  $t_2$ . The maximum degrees of similarity are

represented using Eqs. (7) and (8), as below.

$$S_i^{t_1} = \max(S_{i1}^{t_1}, S_{i2}^{t_1}, \dots, S_{i|R_1^{t_1}|}^{t_1}) \quad (7)$$

$$S_j^{t_2} = \max(S_{1j}^{t_2}, S_{2j}^{t_2}, \dots, S_{i|R_1^{t_1}|}^{t_2}) \quad (8)$$

In this study, the maximum similarity provides the basis for differentiating emerging patterns, added patterns, and perished patterns during various periods. If the maximum similarity of Rule  $r_i^{t_1}$ ,  $S_i^{t_1}$ , equals 1 (or  $S_j^{t_2}$  equals 1), then the rule exists in both time periods  $t_1$  and  $t_2$ , and thus represents an emerging pattern. If a rule exhibits positive growth ( $\text{Sup}_2 > \text{Sup}_1$ ), then the rule represents a pattern of customer behavior that becomes robust with time. Vice versa, a growth rate below zero indicates negative trend of customer behavior change.

If the maximum similarity of Rule  $r_i^{t_1}$ ,  $S_i^{t_1}$ , lies between 0 and 1, the two rules share a partial resemblance. The decision maker determines a rule matching threshold (RMT) to judge whether the similarity of a specific rule satisfies the criteria set by the individual user. If the maximum similarity of Rule  $r_i^{t_1}$ ,  $S_i^{t_1}$ , is smaller than RMT ( $S_i^{t_1} < \text{RMT}$ ), this rule gradually perishes in time period  $t_2$ , and is thus considered a perished pattern. Meanwhile, if the maximum similarity of Rule  $r_j^{t_2}$  is below RMT ( $S_j^{t_2} < \text{RMT}$ ),  $r_j^{t_2}$  in period  $t_2$  is quite different to the rules in period  $t_1$ , and thus it is considered an added pattern.

**2.4.2.2. Unexpectedness.** Unexpectedness was originally used as a subjective measure for interestingness of pattern. Patterns are interesting if they are ‘surprising’ to the user (Silberschatz & Tuzhilin, 1996). When the similarity between two rules equals 0, an unexpectedness measure is used to judge whether the two rules consist of unexpected changes. This study also extends the previous unexpectedness measure designed by Liu and Hsu (1996) and Song et al. (2001) by allowing for two or more attributes (product items) on RHS to detect changes in customer behavior. The modified unexpectedness measure is denoted by Eq. (9), as follows:

$$\delta_{ij} = \begin{cases} \frac{g_{ij} \sum_p x_{ijp} - h_{ij} \sum_q Y_{ijq}}{|A_{ij}|} & \text{if } |A_{ij}| \neq 0, \text{ and } |B_{ij}| \neq 0 \\ 0, & \text{if } |A_{ij}| = 0, \text{ or } |B_{ij}| = 0 \end{cases} \quad (9)$$

In this study, if  $\delta_{ij} > 1$ , then Rule  $r_j^{t_2}$  is an unexpected purchasing rule (i.e. unexpected consequent change) according to  $r_i^{t_1}$ . In this case, customers with same characteristics shift their purchasing behavior or buy different products. If  $\delta_{ij} < 0$ , then Rule  $r_j^{t_2}$  is an unexpected shifting rule (i.e. unexpected conditional changes) according to  $r_i^{t_1}$ . This change indicates that the consumer group of specific products has been changed to another group. If the unexpectedness value equals 0, the two rules are either the same or completely different.

Before unexpected behavior changes are determined, the change patterns are first investigated to ascertain whether they are emerging patterns. If the rule is already an emerging pattern (i.e. it exists in both time periods  $t_1$  and  $t_2$ ), it cannot be categorized into unexpected behavior rules. The adjusted unexpectedness measure is expressed as:

$$\delta'_{ij} = |\delta_{ij}| - k_{ij} \quad (10)$$

When the adjusted unexpectedness value  $\delta'_{ij} = 0$ , the Rule  $r_j^{t_2}$  may have been categorized as an emerging pattern, and consequently can be excluded from unexpected behavior rules. If  $\delta'_{ij} > \text{RMT}$  (RMT for unexpected change is set to 1), Rule  $r_j^{t_2}$  can be categorized as an unexpected change according to  $r_i^{t_1}$ .

The above rule matching process for mining changes in customer behavior is described as follows.

- Step 1. Input customer, product and transaction databases.
- Step 2. Pre-process data via data transformation and discretization.
- Step 3. Partition the data for periods  $t_1$  and  $t_2$ .
- Step 4. Generate association rules for customer behavior,  $R^{t_1}$  and  $R^{t_2}$ , describing the relationships between customer profile (demographic and behavioral attributes) and products purchased using the Apriori algorithm (minimum support = 20%).
- Step 5. Obtain similarity and unexpectedness measures between Rules  $r_i^{t_1}$  and  $r_j^{t_2}$ ,  $r_i^{t_1} \in R^{t_1}$  and  $r_j^{t_2} \in R^{t_2}$ .
- Step 6. Identify the type of change patterns for customer behavior according to the following criteria:
  - (a) Emerging pattern:  $S_i^{t_1} = 1$  (or  $S_j^{t_2} = 1$ ),  $\delta'_{ij} = 0$ ;
  - (b) Added pattern:  $S_j^{t_2} < \text{RMT}$ ;
  - (c) Perished pattern:  $S_i^{t_1} < \text{RMT}$ ;
  - (d) Unexpected condition pattern:  $\delta_{ij} = -1$ ,  $\delta'_{ij} = 1$ ;
  - (e) Unexpected consequent pattern:  $\delta_{ij} = 1$ ,  $\delta'_{ij} = 1$ .

### 2.5. Change patterns for marketing decision-making

When emerging patterns are applied in marketing, positive emerging patterns require concerted marketing efforts. Meanwhile, negative emerging patterns suggest that the use of marketing resources should be reduced by avoiding less profitable customers. The added patterns provide a new marketing target for exploring new consumer markets. Observing the customer behavior changes indicated by perished patterns helps management to plan the reallocation of existing marketing resources. Unexpected purchasing rules (i.e. unexpected consequent patterns) indicate that customers with the same characteristics no longer purchase the same products. Consequently, marketing decision-makers can promote new products based on the projection of unexpected purchasing rules, or can strengthen their marketing efforts to retain customers. Unexpected customer shifting rules (i.e. unexpected conditional patterns) indicate changes occurring in the target customers of products.

### 3. Implementation

A prototype system for mining for changes in customer behavior is devised based on the scheme discussed above. The data for analysis are derived from a sample database, Foodmart, in Microsoft SQL Sever 2000. This database includes the product, customer, and transaction databases of a chain store. Since this study is intended to identify the patterns of changes in customer behavior during different periods, the dataset is first segmented into two for periods of  $t_1$  and  $t_2$ .

#### 3.1. Data transformation

Customer behavioral variables (recency, frequency, and monetary, RFM) are the important variables hidden in the database. To reduce the complexity in data explanation, the recency is assessed using five-score in terms of customer value. Customer purchase frequency is calculated based on weighted average and is differentiated using five-score. The monetary quantity of customer purchase indicates average customer expenditure and divides into five behavior scores for market segmentation.

#### 3.2. Customer valuation and segmentation

According to the RFM variables mentioned in Section 3.1, the total customer value scores for individual customers are calculated to analyze market segmentation and target marketing. In retailing, recency of store visit is a more important indicator than purchase frequency and average spending per visit; hence, recency is assigned a greater weight (say 5) than frequency (say 3) and monetary amount (say 2). Consequently, the maximum score of individual customer value is 50 (i.e.  $5 \times 5 + 3 \times 5 + 2 \times 5$ ), whereas the minimum score is 10 ( $5 \times 1 + 3 \times 1 + 2 \times 1$ ).

The score of customer value provides a basis for customer clustering intended to demonstrate significant difference in customer behavior scores of individual customer clusters. Customers are divided into four clusters based on differentiated value scores. Following customers are grouped based on customer values, the growth matrix of Boston Consulting Group (BSG) is employed to differentiate the value of each customer cluster based on purchase frequency and average monetary expenditure and to segment customers into four clusters with different values based on purchase frequency and average monetary expenditure. The clusters are: best customers (most valuable), frequent buyers, spenders, and uncertain customers (least valuable).

Cluster 4 represents the most valuable customers, while Clusters 1 and 2 represent peripheral customers with lower purchase frequency and monetary expenditure. When recency is used for cross analysis of the four customer clusters, customers in Clusters 1 and 2 have not made regular purchases recently. Meanwhile, Cluster 4 consists of customers who have recently made regular purchases, and

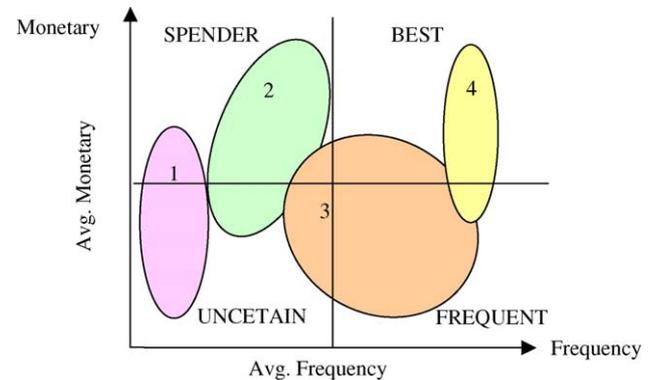


Fig. 3. Customer segmentation via value matrix.

also have higher average purchase size and purchase frequency. Therefore, Cluster 4 is concluded to be the most valuable for the business. This study applies association rules to discover patterns and changes in the behavior of customers in Cluster 4 (Fig. 3).

#### 3.3. Association rules for customer behavior

Association rules are used to analyze the patterns of customer behavior of different time periods for each customer cluster. For mining changes in customer behavior during different periods, the data in this study are grouped into two periods. The data for each observation period is divided into four market segments of uncertain, spender, frequent, and best based on the scores based on customer behavior.

To analyze the relationship between customer profile and purchased products, LHS (conditional part) of association rules consists of customer demographical and behavioral variables, and RHS (consequent part) consists of the products purchased by customers. This study uses the Apriori algorithm to mine customer behavior patterns. The minimum support and minimum confidence of the association rules are set to 20% and the frequent itemset is assumed to include up to six items.

Following the generation of association rules, the association rules of Cluster 4 for two different time periods are compared to understand the customer behavior patterns of the most valuable customers. The generation of customer behavior changes for the most valuable customers is illustrated in the following.

#### 3.4. Changes in customer behavior

The most valuable customers in Cluster 4 are used to provide an example to explain the generation of patterns of change in customer behavior. The association rules of customer behavior generated by the Apriori algorithm for the first and second periods involve 111 and 78 rules, respectively. The rules obtained from the two periods are

then entered into the system for rule matching for customer behavior. The rule matching threshold (RMT) is set to 0.4.

Owing to the large number of changes in customer behavior patterns, a few examples of change pattern are selected from each change category to provide an explanation.

3.4.1. Emerging patterns

Sex = Female → Buy = Snack Foods  
 (Sup<sup>t1</sup> = 20.16%; Sup<sup>t2</sup> = 24.47%; growth rate = 21.38%)

The above rule shows that female customers generally purchase snack food. The support of the rule is 20.16% during the first period and increases to 24.47% in the second period. The growth rate of the rule is 21.38% implying that the rule grows more robust over time.

3.4.2. Added patterns

30 days < Recency < 60 days → Buy = Snack and Vegetables  
 (Sup<sup>t2</sup> = 20.29%; S<sub>i</sub><sup>t2</sup> = 0)

The above rule is a newly added pattern, which provides a reference for developing promotion plans to stimulate consumer needs.

3.4.3. Perished patterns

Sex = Male; Recency < 30 days; Frequency > 13 times → Buy = Snack Foods  
 (Sup<sup>t1</sup> = 21.09%; S<sub>i</sub><sup>t1</sup> = 0.33)

The above decomposed pattern shows that during the first period, among male customers, those whose most recent

purchase date falls within the last 30 days and whose total purchase frequency exceeds 13 times are more likely to purchase snack food. However, this rule shows that the maximum similarity to all the rules during the next period is 0.33, meaning that during the second period, this customer behavior decomposes over time.

In marketing, when developing promotion programs for snack food, the original focus of marketing strategy on male customers should be replaced by a focus on customer groups with other characteristics. The new target customer groups can be determined based on the conclusions from unexpected purchasing behavior (consequent) patterns and unexpected customer shifting (conditional) patterns.

3.4.4. Unexpected consequent patterns

t<sub>1</sub>: Sex = Male → Buy = Snack Foods  
 t<sub>2</sub>: Sex = Male → Buy = Vegetables

The above rules show that the initial pattern of customer behavior is for male customers to purchase snack food. However, in the second period, the customer behavior pattern changes to be that male customers purchase vegetables. This unexpected consequent pattern can lead marketing decision-makers to enforce marketing efforts in promoting vegetables to male customers and to reduce promotions of snack food to male customers, thus increasing customer value.

3.4.5. Unexpected condition patterns

t<sub>1</sub>: Member Card = Bronze; Recency < 30 days → Buy = Vegetables  
 t<sub>2</sub>: No. of Children = 0; 12 < Monetary < = 18; 30 < Recency < = 60 → Buy = Vegetables

Rule ID	LHS	RHS	Confidence	Support
JT1	Marital State=No	Snack Foods	40.78%	22.46
JT10	Marital State=No; Number Children at home=0; Monetary between 12 and 18; Recency<30 days; Frequency>13 times	Snack Foods	40.14%	20.23
JT100	Recency<30 days	Snack Foods	42.26%	42.26
JT101	Recency<30 days; Frequency>13 times	Vegetables	41.84%	40.60
JT102	Recency<30 days; Frequency>13 times	Meat	21.59%	20.95
JT103	Recency<30 days; Frequency>13 times	Fruit	21.96%	21.31
JT104	Recency<30 days; Frequency>13 times	Dairy	28.12%	27.29
JT105	Recency<30 days; Frequency>13 times	Snack Foods	41.77%	40.53
JT106	Frequency>13 times	Vegetables	41.84%	40.60
JT107	Frequency>13 times	Meat	21.59%	20.95
JT108	Frequency>13 times	Fruit	21.96%	21.31
JT109	Frequency>13 times	Dairy	28.12%	27.29
JT11	Marital State=No; Number Children at home=0; Monetary between 12 and 18; Recency<30 days; Frequency>13 times	Vegetables	41.57%	20.95
JT110	Frequency>13 times	Snack Foods	41.77%	20.23
JT12	Marital State=No; Number Children at home=0; Monetary between 12 and 18; Recency>13 times	Snack Foods	40.14%	20.23

Fig. 4. Partial list of patterns for customer behavior.

Original Rule "T100": 「Recency<30 days=> Snack Foods(Support 0.4226)」

the condition parts will be changed:

RID	Unexpected Condition Patterns	Support
2T5	Member Card=Enroaze; Frequency>13 times	0.2044
2T13	Sex=Female; Frequency>13 times	0.209
2T16	House Owner=Yes; 12	0.2097
2T17	House Owner=Yes; 12	0.2097
2T18	Number of Children=0; 12	0.2112
2T19	Number of Children=0; 12	0.2112
2T20	House Owner=Yes; 3013 times	0.212
2T28	Number of Children=0; 3013 times	0.2204
2T30	House Owner=Yes; Frequency>13 times	0.228
2T31	Sex=Female; 30	0.2287
2T32	Number of Children=0; Frequency>13 times	0.2287
2T37	Marital Status=NO; 30	0.2318
2T39	Marital Status=NO; Number of Children=0; 30	0.2318

Fig. 5. Example of search for a change pattern (unexpected conditional pattern).

Unexpected customer shifting patterns show that the same products may shift their initial target between customer clusters at different times. For the future promotion of new vegetable products, marketing personnel can shift to different markets and target new customer clusters to create new markets.

### 3.5. Online system

Many patterns of change in customer behavior require the review of various related data to obtain valuable information for market analysts. In this study, four types of behavior changes obtained from rule matching are saved. An online query system is then built to facilitate the timely searching of change patterns.

All customer behavior patterns generated by association rule mining can be recorded, as illustrated in Fig. 4. Moreover, Fig. 5 shows that the user can select a customer behavior rule for a specified time period and observe whether the rule implies emerging patterns, added patterns, perished patterns, unexpected condition patterns, or unexpected consequent patterns.

By searching for change patterns, market analysts can rapidly acquire the required information) via visualization, and can devise appropriate marketing strategies by searching for possible changes in customer behavior over time.

## 4. Conclusions

The advent of data mining has enhance the customer behavior prediction accuracy. Mining changes for customer behavior is useful for satisfying customer needs in dynamic business environments. Change mining can extract further value from customer, product and transaction databases. In this study, the behavioral variables, RFM, coupled with growth matrix of customer value, are

applied to estimate the value that individual customers contribute to the business. Association rules are used to identify the association between customer profile and product items purchased. The improved measures of similarity and unexpectedness are developed for mining changes in customer behaviors at different time snapshots. Finally, an online query system provides marketing managers a tool for rapid information search, and valuable information based on prompt feedback. The developed system enables marketing managers to rapidly establish marketing strategies.

## Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC 93-2416-H-027-006.

## References

- Agrawal, R., Imielinski, T., & Swami, A. (1993). *Mining association rules between sets of items in large databases*, ACM SIGMOD Conference Washington, DC, USA , 254–259.
- Changchien, S. W., Lee, C.-F., & Hsu, Y.-J. (2004). On-line personalized sales promotion in electronic commerce. *Expert Systems with Applications*, 27(1), 35–52.
- Changchien, S. W., & Lu, T. C. (2001). Mining association rules procedures to support on-line recommendation by customers and products fragmentation. *Expert Systems with Applications*, 20(4), 325–335.
- Chen, M.-S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866–883.
- Dong, G., & Li, J. (1999). Efficient mining of emerging patterns: discovering trends and differences, KDD-99.

- Giudici, P., & Passerone, G. (2002). Data mining of association structures to model consumer behavior. *Computational Statistics and Data Analysis*, 38, 533–541.
- Han, J., & Kamber, M. (2001). *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers.
- Lanquillon, C. (1999). Information filtering in changing domains. *Proceedings of the Joint Conference on Artificial Intelligence (IJCAI99)*, 41–48.
- Liu, B., & Hsu, W. (1996). Post-analysis of learned rules. *Proceedings of the Thirteenth National Conference on Knowledge Discovery in Databases (PAKDD)*, 220–232.
- Liu, B., Hsu, W., Han, H. S., & Xia, Y. (2000). Mining changes for real-life applications. *Second International Conference on Data Warehousing and Knowledge Discovery*, 337–346.
- Liu, B., Hsu, W., Mun, L. F., & Lee, H. Y. (1999). Finding interesting patterns using user expectations. *IEEE Transactions on Knowledge and Data Engineering*, 11(6), 817–832.
- Marcus, C. (1998). A practical yet meaningful approach to customer segmentation. *Journal of Consumer Marketing*, 15(5), 494–504.
- Miglautsch, J. R. (2000). Thoughts on RFM Scoring. *International Society for Strategic Marketing, The Journal of Database Marketing*, 8(1), 67–72.
- Mitra, S., Pal, S. K., & Mitra, P. (2002). Data mining in soft computing framework: a survey. *IEEE Transactions on Neural Networks*, 13 (1), 3–14.
- Padmanabhan, B., & Tuzhilin, A. (1999). Unexpectedness as a measure of interestingness in knowledge discovery. *Decision Support Systems*, 27, 303–318.
- Rud, O. P. (2001). *Data Mining Cookbook*. New York: Wiley.
- Silberschatz, A., & Tuzhilin, A. (1996). What makes patterns interesting in knowledge discovery systems. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 970–974.
- Song, H. S., Kim, J. K., & Kim, S. H. (2001). Mining the change of customer behavior in an internet shopping mall. *Expert System with Applications*, 21(3), 157–168.
- Srikant, R., Vu, Q., & Agrawal, R. (1997). Mining association rules with item constraints. *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*, 67–73.
- Stone, B. (1995). *Successful Direct Marketing Methods*. Lincolnwood, IL: NTC Business Books.
- Suh, H., Noh, K. C., & Suh, C. K. (1999). Customer list segmentation using the combined response model. *Expert Systems with Applications*, 17(1), 89–97.
- Tsai, C.-Y., & Chiu, C. C. (2004). A purchase-based market segmentation methodology. *Expert Systems with Applications*, 27, 265–276.