A Post-Analysis of Decision Tree to Detect the Change of Customer Behavior on Internet Shopping Mall

Jae kyeong Kim[a*], Hee Seok Song[b], and Tae Sung Kim[c]

a, c School of Business Administration, Kyung Hee University
#1, Hoeki-Dong, Dongdaemoon, Seoul, 130-701, South Korea
Tel : 82-2-961-9355 Fax : 82-2-967-0788 e-mail : jaek@khu.ac.kr

[*]Corresponding author

[b] Graduate School of Management, Korea Advanced Institute of Science and Technology
207-43 Cheongryangri-Dong, Dongdaemun, Seoul, 130-012, South Korea
Tel : 82-2-958-3670 Fax : 82-2-958-3604 e-mail : hssong@kgsm.kaist.ac.kr

Abstract

Understanding and adapting to changes of customer behavior in internet shopping mall is an important aspect to survive in continuously changing environment. This paper develops a methodology based on decision tree algorithms to detect changes of customer behavior automatically from customer profiles and sales data at different time snapshots. We first define three types of changes as emerging pattern, unexpected change and the added / perished rule. Then, it is developed similarity and difference measures for rule matching to detect all types of change. Finally, the degree of change is developed to evaluate the amount of change. A Korean internet shopping mall case is evaluated to represent the performance of our methodology. And practical business implications for this methodology are also provided.

Keywords : Data Mining, Decision Tree, Change Analysis, Internet Shopping Mall

Introduction

Understanding and adapting to changes of customer needs is an important aspect of surviving in a continuously changing environment. Recent results of interview which had been performed by a consulting company show the importance of knowing the change of their customer needs. The result shows that about 25 % of world-wide 500 companies responded changing customer needs as a primary reason for changing their business strategy. In this paper, we develop a methodology for detecting the change of customer behavior as a means of identifying the change of customer needs. More specifically, this paper aims to give answers about the following questions; Which customer group's purchasing amount are gradually increasing? Which customer groups are churned over the years? Whether the preference of specific customer groups are changed over time?

Our suggested methodology is based on data mining methodology, especially the decision tree analysis. Data mining is the process of exploration and analysis of large quantities of data in order to discover meaningful patterns and rules. In this context, data mining gives great potential to discover automatically the change of customer behavior. Therefore, we use data mining techniques to detect the change of customer behavior. But most data mining techniques such as association rules, decision trees and neural networks cannot be applied alone to answer the above research questions, because they cannot handle dynamic situations well.

Liu et al.(2000) suggests a change mining methodology based on decision trees to find interesting relationships among a large set of data items. However, their research can not give an answer about how much change has been occurred. It is important to detect what kind of changes are occurred, but it is also important to differentiate which change is more serious. Song et al. (2001) developed a change detection procedure and a measure for evaluating the amount of change based on association rule mining. Change detection based on association rule mining can be used to identify changes of customer behavior in unstructured and ill-defined situation because of unsupervised learning feature of association rule mining. However, decision tree analysis in change detection problem can be used to more structured situation in which the manager have specific research question and also it detects the change of classification criteria in a dynamically changing environment. The application of association rule mining and decision tree are different, but the measure for evaluating the amount of change of Song et al. (2001)'s research is adapted in this research with modification.

Detected changes by decision tree analysis are usefully applied to plan various niche-marketing campaigns. For example, in a shop, if a manager can find out that a criteria of a certain customers' group to choose a product has changed from price into design, then he/she will modify existing merchandising strategy for such a group of customers. The methodology suggested in this paper detects changes automatically from customer profiles and sales data at different periods of time. The most common approach to discover changes between two datasets is to generate decision trees from each dataset and directly compare the rules from decision trees by rule matching. But this is not a simple process because of the following reasons. First, some rules cannot be easily compared due to different rule structures. Second, even with matched rules, it is difficult to know what kind of change has occurred and how much change has occurred. To simplify these difficulties, we first define three types of changes as emerging pattern, unexpected change and the added / perished rule. Then we develop similarity and difference measures for rule matching to detect all types of change from different time snapshot data. Finally, the degree of change is evaluated to detect significantly changed rules.

The content of our research is summarized as follows. First, a post analysis of decision tree is suggested to detect the change of classification criteria. Second, the suggested methodology is applied to the real-world internet shopping mall case. Third, the methodology is evaluated whether it generates meaningful rules for the detection of customer behavior change. Finally the business implication of this methodology is discussed.

Problem

In this section, we examine all possible types of change based on past research and business requirements (Dong & Li, 1999; Lanquillon, 1999; Liu & Hsu, 1996; Liu et al., 1997; Padmanabhan & Tuzhilin, 1999; Song et al., 2001; Suzuki, 1997). After that, each type of change and change detection problem are defined.

Let's define the following notation.

D^t, D^{t+k} : datasets at time t, t+k

R^t, R^{t+k} : discovered decision tree rulesets at time t, t+k

r_i^t, r_j^{t+k} : each rule from corresponding ruleset R^t, R^{t+k} ,where $i = 1,2,\Lambda,|R^t|$, $j = 1,2,\Lambda,|R^{t+k}|$

$Sup^t(r_i)$: support of r_i in time t dataset

Dong and Li (1999) introduced *Emerging Patterns* concept that captures significant changes and differences between datasets. Emerging patterns are defined as itemsets whose supports increase significantly from one dataset to another. More specifically, emerging patterns are itemsets whose growth rates are larger than a given threshold value.

Definition [1] Emerging Patterns

For rule r_j^{t+k}, if the following two conditions are met, then we call it the rule of Emerging Pattern with respect to r_i^t.

(1) Conditional and consequent parts are the same between r_i^t, r_j^{t+k}

(2) Supports of two rules are significantly different

Example [1]

r_i^t : If Income = High, Age = High , Then SalesMonuments = High (Support = 0.1)

r_j^{t+k} : If Income = High, Age = High, Then SalesMonuments = High (Support = 0.13)

In this case, r_j^{t+k} is the emerging pattern with respect to r_i^t if we specify minimum growth rate to be 0.2. This is because the two rules have same rule structure and their growth rate is 0.3.

The other type of change is unexpectedness which is found from many studies about discovering interesting patterns (Liu & Hsu, 1996; Liu et al., 1997; Padmanabhan & Tuzhilin, 1999; Silberschatz & Tuzhilin, 1996; Suzuki, 1997). A rule and a belief are "different" if either the consequents of the rule and the belief are "similar" but the conditions are "far apart," or the consequents are "far apart" but the conditions are "similar", where "similarity" and "difference" are defined syntactically based exclusively on the structure of the rules.

Definition [2] Unexpected Changes (or Unexpected Consequent Changes)

r_j^{t+k} is unexpected change with respect to r_i^t if the conditional parts of r_i^t, r_j^{t+k} are similar, but the consequent parts of the two rules are quite different.

Example [2]

r_i^t : If Income = High, Age = High ,Then SalesMonuments = High

r_j^{t+k} : If Income = High, Age = High ,Then SalesMonuments = Low.

In this case, r_j^{t+k} is unexpected consequent change with respect to r_i^t since the conditional parts of r_i^t, r_j^{t+k} are similar, but the consequent parts of the two rules are quite different.

Other types of change are added rules and perished rules (Lanquillon, 1999). An added rule is a newly arisen rule which could not be found in the past and a perished rule is a disappeared rule which can be found only in the past but not the present.

Definition [3] Added rules / Perished rules

r_j^{t+k} is an added rule if all the conditions and consequents are quite different from any of r_i^t in R^t and r_i^t is a perished rule if all the conditions and consequents are quite different from any of r_j^{t+k} in R^{t+k}.

We used the terms "similar" and "quite different" in the above definitions. Those terms are used to compare two rules in syntactic aspects and to judge degree of similarity and difference. But the terms "similar" and "quite different" are quite subjective and different from each individual. Therefore we define *Rule Matching Threshold (RMT)* which can be differently determined by individual user. Finally, we define *the degree of change* as the measure of how much change has occurred. The degree of change has to be evaluated differently by each type of change because of different characteristics. The main way of evaluating degree of change will be explained in the next section. Now, the change detection problem is defined as follows using the above definitions of each change type.

Definition [4] Change detection problem
The change detection problem consists of finding all emerging patterns, unexpected changes and added/perished rules between datasets which are collected from different periods and ranking the changed rules in each type by the degree of change.

Methodology

Overall procedure

The change detection problem is defined at previous section. Now we suggest the methodology for the change detection problem. The methodology consists of the following three phases in Figure 1.

In phase I, two rulesets are generated from each dataset by using Decision Tree Analysis (Liu et al., 2000). For such a purpose, we present two basic approaches to change mining in the decision tree model; new decision tree, and same attribute and same cut. In phase II, the changed ruleset is generated by using the rule matching method which compares two rules selected from each ruleset. Our rule matching method can detect all types of changed rules including emerging patterns, unexpected changes, added and perished rules. In phase III, various changed rules detected in phase II are ranked according to the predefined degree of change that is a measure to evaluate how much change has occurred. In case of emerging patterns, the growth or decrease rate of the *support* value of the changed rule is computed and then they are sorted by the absolute value of the rate. Second, the case of unexpected change is defined using the exception rule concept (Hussain et al., 2000; Suzuki, 1997). Third, in the case of added or perished rules, we use the support value and maximum similarity value of new or disappeared rules as the degree of change.

Discovery of rule change in decision tree

Each path of a decision tree becomes a rule. In this paper, two approaches are used to change mining in the decision tree model : new decision tree, and same attribute and same cut. The first approach modifies the original tree structure, which makes the comparison with the original structure difficult. The second approach is easy to compare the two decision tree, but the support values of the same rule of each tree might be different. Note that the basic decision tree algorithm we use in our study is C4.5. We have modified it in various places for change mining purposes.

Same attribute and same cut: A decision tree is developed with the dataset of time t. At time t+k, not only the same attribute as in the tree of time t at each step of partitioning are used, but also the same cut point in the tree of time t. If the algorithm has not reached the leaf node of a particular branch of the tree, the tree of time t+k needs to go beyond the depth of the corresponding branch in the tree of time t. Rulesets of time t and time t+k are generated based on decision trees of time t and time t+k respectively. In the case of same attribute and same cut in time t+k, a same precedure is performed except a decision tree of time t+k is developed first.

New decision tree : Decision trees are generated using the dataset of time t and that of time t+k. Two rulesets are generated from two decision trees of time t and time t+k. Based on this methods, added rules or perished rules are usually found

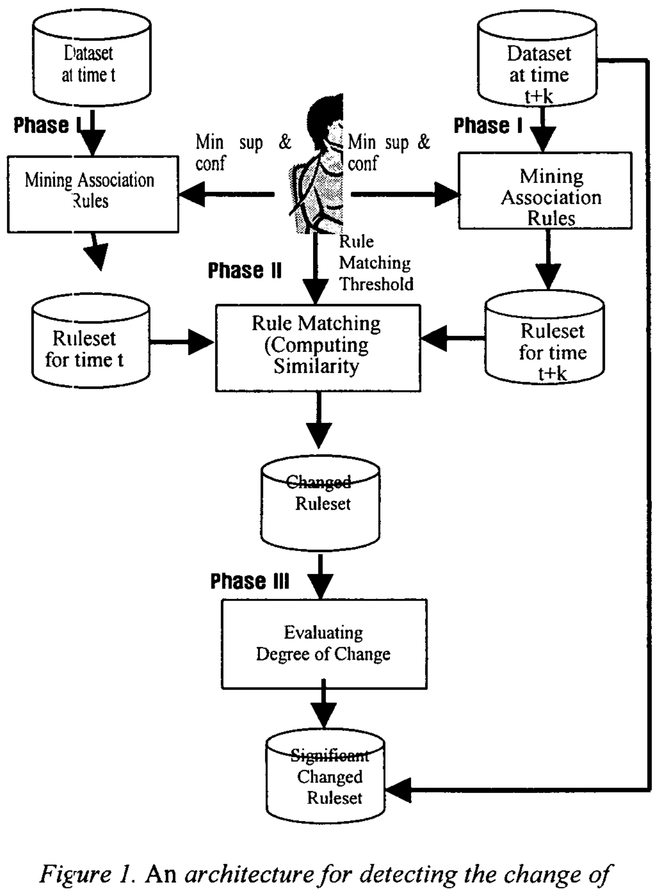

Figure 1. An *architecture for detecting the change of customer behavior*

Discovery of changed rule

In this phase, various types of changed rules are detected using the rule matching method. The inputs of phase II are discovered rulesets at time t and t+k which are generated from the two methods of the previous section, and the Rule Matching Threshold (RMT) which is specified by the user. Phase II is composed of the following three steps.

Step [1] Calculate the maximum similarity value for each rule in time t and t+k.

Step [2] For each rule r_i^t, calculate the difference measures between r_i^t, r_j^{t+k}.

Step [3] Classify the type of change for the rules using the maximum similarity value and the difference measures.

For the explanation of each step, some notations are briefly defined.

δ_{ij} : Difference measure. Degree of difference between r_i^t and r_j^{t+k} $(-1 \le \delta_{ij} \le 1)$

s_{ij} : Similarity measure. Degree of similarity between r_i^t and r_j^{t+k} $(0 \le s_{ij} \le 1)$

λ_{ij} : Degree of attribute match of the conditional parts,

$$\lambda_{ij} = \frac{|A_{ij}|}{\max(|X_i^t|, |X_j^{t+k}|)}$$

c_{ij} : Degree of attribute match of the consequent parts,

$$c_{ij} = \begin{cases} 1, & \text{if same consequent attribute} \\ 0, & \text{otherwise} \end{cases}$$

$|A_{ij}|$: Number of attributes common to both conditional parts of r_i^t and r_j^{t+k}.

$|X_i^t|$: Number of attributes in the conditional parts of r_i^t.

$|X_j^{t+k}|$: Number of attributes in the conditional parts of r_j^{t+k}.

x_{ijk} : Degree of value match of the kth matching attribute in A_{ij}, $x_{ijk} = \begin{cases} 1, & \text{if same value} \\ 0, & \text{otherwise} \end{cases}$

y_{ij} : Degree of value match of the consequent attribute,

$$y_{ij} = \begin{cases} 1, & \text{if same value} \\ 0, & \text{otherwise} \end{cases}$$

Now we provide a *similarity measure* as follows, adapted from the study of Liu and Hsu (1996).

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

In s_{ij}, $\dfrac{\lambda_{ij} \times \sum_{k \in A_{ij}} x_{ijk}}{|A_{ij}|}$ represents a similarity of conditional part, and $c_{ij} \times y_{ij}$ represents a similarity of consequent part between r_i^t and r_j^{t+k}. If the conditional and consequent parts between r_i^t and r_j^{t+k} are the same, then the degree of similarity becomes 1. The similarity measure can take any value between 0 and 1. To detect added and perished rules, the *maximum similarity value* is provided as follows.

$s_i = \max(s_{i1}, s_{i2}, \cdots, s_{i|R_{t+k}|})$; *Maximum Similarity Value of* r_i^t

$s_j = \max(s_{1j}, s_{2j}, \cdots, s_{|R_t|j})$; *Maximum Similarity Value of* r_j^{t+k}

The maximum similarity value indicates whether the rule is added or perished. If $s_i < RMT$, then r_i^t is recognized as a perished rule. If $s_j < RMT$, then the rule r_j^{t+k} becomes an added rule.

Example [3] Assume the following rules are generated from the decision tree of time t and that of time t+k.

r_1^t : If Income = High, Then SalesMonuments = High

r_2^t : If Age = High, Preference = Price , Then SalesMonuments = High

r_1^{t+k} : If Income = High , Then SalesMonuments = High

r_2^{t+k} : If Age = High , Then SalesMonuments = High

r_3^{t+k} : If Income = High, Preference = Price , Then SalesMonuments = Low

We can compute the similarity measure between r_2^t, r_2^{t+k} and the maximum similarity value of r_2^t as follows.

$$s_{22} = \frac{\frac{1}{2} \times 1 \times 1 \times 1}{1} = 0.5, \qquad s_2^t = \max(0, 0.5, 0) = 0.5$$

In the same manner, we can compute the maximum similarity value of each rule.

$s_1^t = \max(1,0,0) = 1 \quad s_2^t = \max(0,0.5,0) = 0.5$

$s_1^{t+k} = \max(1,0) = 1 \quad s_2^{t+k} = \max(0,0.5) = 0.5 \quad s_3^{t+k} = \max(0,0) = 0$

If RMT is specified to be 0.4, then we can conclude that only r_3^{t+k} is an added rule.

As we can see from example 3, the maximum similarity value in step [1] is used to discover added rules or perished rules. The purpose of step [2] is to detect unexpected changes and emerging patterns. Unexpected change consists of unexpected condition and unexpected consequents. To detect unexpected change, a *difference measure* is provided as follows.

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

The step [3] classifies the rules as three types of change. To classify the type of change, additional computation is needed. For example, although r_j^{t+k} is judged to be an unexpected change with regard to r_i^t by the difference measure, we cannot conclude directly whether it is an unexpected change or not. Because r_j^{t+k} can be an emerging pattern with regard to r_m^t which has the same structure with r_j^{t+k}. In this case, r_j^{t+k} should be classified into an emerging pattern and not to be classified as an unexpected change. As we cannot conclude based on δ_{ij} alone whether r_j^{t+k} is an unexpected change or an emerging pattern, we provide the following modified difference measure.

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

The fact that s_i (or s_j) is equal to 1 means that the same rule exists in another ruleset. That means r_j^{t+k} is likely to be classified into an emerging pattern. If δ_{ij}' is greater than the pre-specified RMT, then the rule r_j^{t+k} is concluded to be an unexpected change with respect to r_i^t.

Example [4]

r_1^t : If Income = High, Preference = Price, Then SalesMonuments = Low

r_2^t : If Age = High, Preference = Price , Then SalesMonuments = High

r_1^{t+k} : If Income = High , Then SalesMonuments = High

r_2^{t+k} : If Age = High , Then SalesMonuments = High

r_3^{t+k} : If Income = High, Preference = Price , Then SalesMonuments = Low

With the above ruleset, we can compute the difference and modified difference measure between r_2^t and r_3^{t+k} as follows; $\delta_{23} = 0.5$ $\quad \delta_{23}' = 0.5 - \max(0.5, 1) = -0.5$.

If we specify that RMT is equal to 0.4, we cannot conclude that r_3^{t+k} is an unexpected consequent change with respect to r_2^t because r_3^{t+k} has a more similar rule structure with r_1^t rather than r_2^t. Therefore, we can conclude that r_3^{t+k} is an emerging pattern of r_1^t. And r_3^{t+k} is not thought to be an unexpected consequent change with respect to r_1^t. Table 1. summarizes the value of each measure for each type of change.

Evaluating the degree of change

All the changed rules have to be ranked by the degree of change. The measure of the degree of change, α_{ij} , is summarized in each type of change. Please refer Kim (2001) for the detail.

$$\alpha_{ij} = \begin{cases} \dfrac{Sup^{t+k}(r_i) - Sup^t(r_i)}{Sup^t(r_i)} & ,emerging \ \ pattern \ \ case \\ \dfrac{Sup^{t+k}(r_{i \cap j})}{Sup^{t+k}(r_j)} & ,unexpected \ \ change \ \ case \\ (1 - s_i) \times Sup^t(r_i) & ,perished \ \ rule \ \ case \\ (1 - s_j) \times Sup^{t+k}(r_j) & ,added \ \ rule \ \ case \end{cases}$$

Experiments and Applications

The case study has been conducted to evaluate how well the procedure performs its intended task of detecting significant changes. The dataset is prepared from a Korean online shopping mall which sells various consumer goods. The dataset contains customer profiles and purchasing history such as age, job, sex, address, registration year, cyber money, number of purchases, number of visits, payment method during one year, and total purchase amount. We prepared two dataset to detect significant changes of purchasing behavior by their customers. The first dataset contains profiles and purchasing history information of certain customers who had bought more than one cosmetics from Feb/1/2000 to Jun/30/2000. The second dataset contains the same information but includes customers who had made one additional purchase of cosmetics from Jul/1/2000 to Jan/5/2001. After preprocessing the data for cleansing and discretization, we built the decision tree. The splitting criterion is chi-square test and significance level is set up as 2.0. The procedure to detect all the change is explained according to three categories.

Table 1. Value of measure for each type of change

ype of hange	Value of measure to classify
Emerging Pattern	$\delta_{ij} = 0, \ (\sum_{k \in A} x_{ijk} > 0 \ or \ y_{ij} > 0 \ or \ \lambda_{ij} > 0)$
Unexpected Consequent	$\delta_{ij} > 0, \quad \delta_{ij}' \geq RMT$
Unexpected Condition	$\delta_{ij} < 0, \quad \delta_{ij}' \geq RMT$
Added Rule	$s_j < RMT$
Perished Rule	$s_i < RMT$

Same attribute and same cut at time t

We processed the data at time t for cleansing and discretization and built the decision tree and corresponding rules with SAS 8.0 enterprise miner program. And with the same attribute and same cut, we built the decision tree and corresponding rules of dataset at time t+k. Then, we compared the rules of time t and t+k, and discovered the emerging pattern, unexpected changes, and added/perished rules. We evaluated the changed rules to find what rules were changed more. Figure 2 and Figure 3 show the decision trees of time t and time t+k.

The emerging patterns, unexpected changes, and added/perished rules are shown at Table 2. of the same attribute and the same cut at time t is as follows in Table 2. The specific significant emerging patterns, unexpected changes, and added/perished rules are summarized in Table 3, 4, and 5.

From changed rule (1) in Table 3, we can see the rapid decrease of customers with high purchasing rate (96 % decrease rate) in sales for customer segment that are mostly

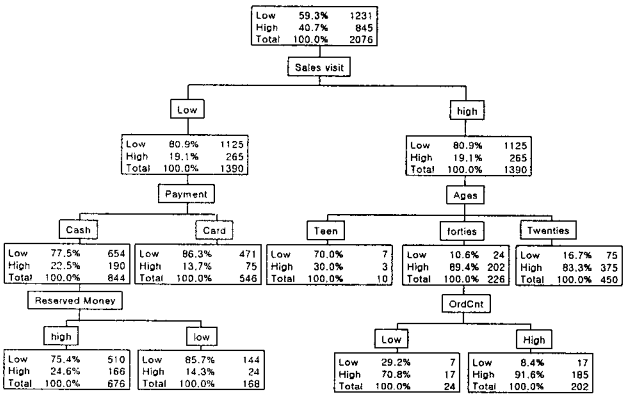

Figure 2. Decision Tree of dataset at time t in the same attribute and the same cut

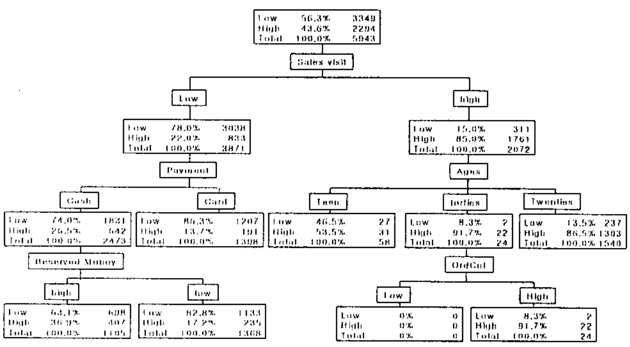

Figure 3. Decision Tree of dataset at time t+k in the same attribute and the same cut

Table 2. Number of changed rules in the same attribute and the same cut at time t

Type of change	Number of changed rules	Number of significant changed rules
Emerging Patterns	4	3 (Degree of change > 0.1)
Unexpected Changes	1	1 (Degree of change > 0.1)
Added/Perished Rules	1	1 (Degree of change > 0.001)

Table 3. Significant *Emerging Patterns (Degree of change > 0.1) in the same attribute and the same cut at time t*

No	Rules	n(t)	$Sup^t(r_i)$	n(t+k)	$Sup^{t+k}(r_i)$
1	if salesvisit = high, age = forties , order count = high then salesmonument = high	185	0.09	22	0.00
2	if salesvisit = high, age = twenties then salesmonument = high	375	0.18	1303	0.22
3	if sale visit = low, payment = card then salesmonument = low	471	0.23	1207	0.20

Table 4. Significant Unexpected Changes (Degree of change > 0.1) in the same attribute and the same cut at time t

	r_i^t	r_j^{t+k}	δ_{ij}	δ'_{ij}	α_{ij}
1	salesvisit = high, Ages = teen then salesmonument	salesvisit = high, Ages = teen then salesmonument = high support=0.005	0.5	0.5	0.6

Table 5. Significant Perished Rules (Degree of change > 0.001) in the same attribute and the same cut at time t

	r_i^t	MSV	Support	α_{ij}
1	if salesvisit = high, age = forties , order count = low then salesmonument = high	0	0.008	0.008

forties years old, order frequently, and visit the mall frequently. Because of large decreasing rate, that segment of customers will be disappeared in the near future. Therefore special careful consideraion of those customers will be needed. From changed rule (2) in Table 3, we can find the rapid growth of customer groups with high purchasing rate in sales who are twenties and visit the mall frequently. This tendency is obviously desirable, therefore a marketing campaign to invoke the revisiting of those customers should be developed. With regard to unexpected changes of Table 4, the sales for customers who are in their teens and visit the mall frequently are low from the first dataset. But in the second dataset, we can see that the sales for those customers are high. It means that the importance of such customers is gradually increasing. Therefore, a modification of the existing marketing strategy and plan for those customers is required. Finally, one perished rule is found in Table 5. From Feb. to Jun. in 2000, the sales of frequently visiting customers who are in their forties and order infrequently were high. But after Jun. 2000, we can not find those rules anymore, therefore we should decide that additional services and products for elders are to be developed or not.

Same attribute and same cut at time t+k

Likewise the case of the same attribute and same cut at time t, we built the decision tree and rules of dataset at time t+k with SAS 8.0 enterprise miner program. And with the same attribute and same cut of the tree, we built the decision tree and rules of dataset at time t. Then, we compared two ruleset of time t+k and t to discovere the emerging pattern, unexpected changes, and added/perished rules. Each decision tree and changed rules are omitted.

New decision tree at time t and time t+k

Two decision trees are built with the data set of time t and time t+K independently using SAS 8.0 enterprise miner program. And two rulesets are constructed along with the pathway of each decision tree. Then, it is compared the rules of time t and t+k and discovered the emerging pattern, unexpected changes, and added/perished rules. Compared to the previous cases, new decision trees at each time generate more added rules and perished rules.

Business implications

In this section, we summarize the opportunities of using this methodology and provide various applications in practical business perspectives. First, in macro aspects, business mangers can follow the changing trends using change detection methodology. They need to analyze their customer's changing behaviors in order to provide products and services that suit the changing needs of the customers (Liu et al., 2000). For example, if a manager finds the trend that the age of customer groups for certain product is decreasing, then he/she can develop additional services and product specifications for them.

On the other hand, undesirable changes can be properly controlled. If the manager detects the decreasing trend of purchase of certain products, then he/she can examine the reason and establish a reaction plan to prevent that trend. Second, in micro aspects, it is possible for a business manager to understand customer needs more deeply and design additional niche marketing campaigns based on the results of suggested methodology. Knowing the purchasing history of a certain customer groups can give a better understanding of the behavior of them. For example, although the current satisfaction level of two different customer groups are both "low" for a certain product, the satisfaction level of customer group A was "high" and that of the other group B was "low" in the past year. In this case, the manager can start a churn analysis to examine the decline of group A. To give another example, if a manager finds a certain customer group, the size of which is very small but has a large growth rate, then a direct mailing campaign is planned for customer groups who have similar properties.

Change detection is more suitable in domains where the environment is relatively dynamic and there is much human intervention. Besides understanding customer behavior change, another most promising application is analyzing the effectiveness of a marketing campaign. Using this methodology, if the manager classifies customer segments before and after a campaign, he/she can evaluate the effectiveness of his/her marketing campaign by seeing whether it operated correctly with respect to original intention.

Even in a static environment, our change detection methodology can be used. The most common area in a static environment is the comparison between two categorical dataset.

Conclusions

In this paper, we developed a methodology that detects changes of customer behavior automatically from customer profiles and sales data at different time snapshots. For this purpose, we suggested the methods to find rulesets from decision trees and defined types of change as emerging pattern, unexpected change and the added / perished rule. Then we developed similarity and difference measures for rule matching to detect all types of change in syntactic aspects. Additionally, the degree of change is evaluated to detect significantly changed rules in semantic aspects. We also suggested practical applications and opportunities to use for our methodology.

Our suggested methodology focuses on finding the changes of customer segments between different time snapshot datasets. But it can be used in another applications, too. The methodology can be applied for comparison of categorical data as well as dynamically changed data. With regard to the number of target datasets which should be compared, the methodology is also applied to three or more datasets to be compared over time.

As a further research area, we are to setup the campaign management plan based on our suggested methodology. And it will be also interesting to check the effectiveness of the campaign. We believe that the change detection problem will become more and more important as more data mining applications are implemented.

References

[1] Dong, G., and 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 (KDD-99)*, 43-52.

[2] Hussain, F., Liu, H., Suzuki, E., and Lu, H. 2000. Exception Rule Mining with a Relative Interestingness Measure. In *Proceedings of Pacific Asia Conference on Knowledge Discovery in Databases (PAKDD)*, 86-97.

[3] Kim, T. S. 2001. A Study on Detecting the Change of Customer Behavior: Application to Internet Shopping Mall. *MS Thesis, Graduate School of Business Administration, Kyung Hee University, Seoul, Korea.*

[4] Lanquillon, C. 1999. Information filtering in changing domains. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI99)*, 41-48.

[5] Li, J., Dong, G., and Ramamohanarao, K. 2000. Making Use of the Most Expressive Jumping Emerging Patterns for Classification. In *Proceedings of Pacific Asia Conference on Knowledge Discovery in Databases (PAKDD)*, 220-232.

[6] Liu, B., and Hsu, W. 1996. Post-Analysis of Learned Rules. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96)*, 828-834.

[7] Liu, B., Hsu, W., and Chen, S. 1997. Using general impressions to analyze discovered classification rules. In *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*, 31-36.

[8] Liu, B., Hsu, W., Han, H. S., and Xia, Y. 2000. Mining Changes for Real-Life Applications. In *Publishing in the Second International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2000)*, Y. Kambayashi, M.K. Mohania, A.M. Tjoa (eds.), 337-346.

[9] Padmanabhan, B., and Tuzhilin, A. 1999. Unexpectedness as a measure of interestingness in knowledge discovery. *Decision Support Systems* 27: 303-318.

[10] Silberschatz, A., and Tuzhilin, A. 1996. What makes patterns interesting in knowledge discovery systems. *IEEE Trans. On Knowledge and Data Engineering* 8(6): 970-974.

[11] 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), To be appeared.

[12] Suzuki, E. 1997. Autonomous discovery of reliable exception rules. In *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*, 259-262.