Expert Systems with Applications 36 (2009) 6091–6099

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A transaction pattern analysis system based on neural network

Tzu-Chuen Lu *, Kun-Yi Wu

Department of Information Management, Chaoyang University of Technology, 168, Jifong East Road, Wufong Township, Taichung 41349, Taiwan, ROC

ARTICLE INFO

Keywords:
Market segmentation
Clustering
Customer segmentation
Association rule mining
Neural network

ABSTRACT

Customer segmentation is a key element for target marketing or market segmentation. Although there are quite a lot of ways available for segmentation today, most of them emphasize numeric calculation instead of commercial goals. In this study, we propose an improved segmentation method called transaction pattern based customer segmentation with neural network (TPCSNN) based on customer's historical transaction patterns. First of all, it filters transaction data from database for records with typical patterns. Next, it reduces inter-group correlation coefficient and increases inner cluster density to achieve customer segmentation by iterative calculation. Then, it utilizes neural network to dig patterns of consumptive behaviors. The results can be used to segment new customers. By this way, customer segmentation can be implemented in very short time and costs little. Furthermore, the results of segmentation are also analyzed and explained in this study.

1. Introduction

Consumer market changes rapidly without any settled logic. In most occasions, you can find all kinds of demands in it. Customers' requirement will never be satisfied merely by one or two products. However, excessive products can be a burden or risk to the company's operation (Guha, Rastogi, & Shim, 2000; Hagan, Demuth, & Beale, 1996; Lee et al., 2000; Schiffman & Kanuk, 2000). Therefore, in order to satisfy various customer requirements within company's capacity, we need to split consumer market into several segmentations and find out appropriate marketing strategies for them (Changchien & Kuo, 2004; Changchien & Lu, 2001a; 2001b; Dennis et al., 2003; Kotler, 1994; Kotler & Armstrong, 1997).

The spirit of strategic marketing proposed by Kotler is STP: segmentation, targeting, and positioning (Kotler, 1994; Kotler & Armstrong, 1997). Based on some customer diversities, the complicated market in the reality can be separated into several small markets with similar properties. Among them, the companies can find their target markets and the positions. Such strategy is a golden rule even in today's business. In the recent years, Kotler proposed a new brand marketing mode create communicate deliver value target profit (CCDVTP). It tries to create new communication tunnels and deliver brand values. Then it conducts marketing with specific targets, and finally achieves profits. To stipulate for a marketing strategy, market segmentation is the first step (Kotler, 1994; Kotler & Armstrong, 1997).

Today, there are a lot of market segmentation methods available, but most of them are based on the existing segmentation methods, such as K-means, density-based spatial clustering of applications with noise (DBSCAN) and so on (Aldenderfer & Blashfied, 1984; Bezdek & Pal, 1998; Ester, Kriegel, Sander, & Xu, 1996; Filippone, Camastra, Masulli, & Rovetta, 2008; Guha et al., 2000; Guldemir & Senguar, 2006; Gunter & Bunke, 2003; Kumar & Patel, 2007; Liu & Samal, 2002). The users have to choose an appropriate clustering method based on the goals to be resolved or the characteristics of the database. After the decision, the users need to find suitable parameters for the clustering. It requires the users to be very familiar with the problem or the data characteristics to find the parameters and obtain the optimized result. As a matter of fact, this is a mission impossible for the normal companies. Therefore, in this study, we try to propose an easy and understandable method for the normal users, which can implement the segmentation rapidly and correctly. By this way, the business operation can be supported by the theory. Besides, most of the existing methods are not designed for business purpose. As a result, some adjustments to the data are required during the application. Such adjustment may make the results away from the target problems.

In 2004, ChangChien and Kuo proposed a customer segmentation method called transaction pattern based customer segmentation (TPCS) (Changchien & Kuo, 2004). Their customer segmentation method covers both marketing and business purposes. They utilize customers' historical transaction data and group the customers by similar transaction patterns. Meanwhile, they put marketing and business purposes into consideration. However, their method is incapable of analyzing or explaining the segmentation results. In this study, we try to improve the TPCS method and add extraction mechanics for the transaction data to the original

* Corresponding author. Tel.: +886 4 23323000; fax: +886 4 23742337.
E-mail address: tclu@cyut.edu.tw (T.-C. Lu).
URL: http://www.cyut.edu.tw/~tclu (T.-C. Lu).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.07.073

method, analyze the segmentation result. Furthermore, neural network technology is adopted in our improvement. It enables the users to obtain new customer segmentation quickly.

Here is a brief structure of this paper. Section 2 is a literature review, which introduces TPCS method and other relevant technologies. In Section 3, improved segmentation method proposed by us is introduced. In Section 4, we try to evaluate the improvements by simulation and actual data. The evaluation can be validated by the real system. Section 5 is a summary.

2. Related works

2.1. Transaction pattern based customer segmentation (TPCS)

TPCS looks for the patterns of consumer behaviors from the transaction data. The segmentation can be adjusted by the customer weighing, and can become customer oriented. This function avoids some single customer with transaction data in different segmentations. Then, the segmentation correlation matrix (SCM) is produced based on consumer behavior patterns and correlation coefficient to demonstrate the intersection degree among the segmentations. The segmentation can be further adjusted to achieve the lowest inter-segmentation correlation. After that, a density indicator is added to measure inner cluster correlation. By reducing inter-segmentation correlation and enhancing inner cluster density, customer segmentation can be achieved using a combination of merging and separation strategies. However, this method has two limitations:

1. During transaction rules mining, attentions are paid to items purchased instead of amount.
2. Transaction is a collection of items purchased. It is not connected with purchasing order. It also includes the situation of purchasing single item.

We are going to explain the steps of TPCS by 7 transaction records with 4 customers and 4 different items. Table 1a is the original transaction data. In column Item, commodities purchased by each customer are listed. A, B, C, and D are item codes. TID is transaction number. Above all, we convert this table into Table 1b. The symbol '1' stands for yes and '0' stands for no. Minimum support is adopted at this time to get rid of rare items.

Assume that the customers can be separated into cluster *X* and cluster *Y*, as shown in Table 2, we can calculate the customers' cluster correlation matrix for these two segmentations using the expected value and probability of the rules together with the permutation and combination of the four items, as shown in Table 3. Besides, if we assume that the profit for each product is 1, the $P(rule_i^X)$ in the table represents the occurrence probability of rule *i* in cluster *X*. $EV(rule_i^X)$ is the expected value. For example, rule 12 is 1100, it represents that the occurrence probability of transactions purchasing both A and B in cluster *X* is 1/7, the expected value is 2/7.

Then, we go ahead with the inner cluster density calculation. The number of data records is divided by the sum of distances between each transaction record and the cluster centroid. The result can be used as a density indicator for the inner cluster compactness. The distances are calculated as Euclidean distance. We can calculate the centroid for every cluster from Table 2. Table 4 shows the results.

From Table 4, we can find that the density of cluster *X* is 2.2857 and the density of cluster *Y* is 4.4998. The effect of two clusters is 42.4953. This algorithm pairs every two clusters and looks for the combination with worst effect for adjustment. If all possible combinations have been tested and none of them is better than the original cluster, the segmentation is finished.

2.2. Back-propagation network (BPN)

Although the segmentation can be implemented by TPCS effectively, this method has a no capability of explaining. Therefore, the BPN is adopted in the analysis and learning of segmentation (Hagan et al., 1996; Han & Kamber, 2001). BPN is a kind of supervised learning network and is widely used in many occasions (Wang & Huang, 2008). It has very good performance in diagnosis,

Table 1
The transaction records and the converted data

(a) Original transaction data

TID	Customer ID	Item
1	Alpha	A, B, D
2	Beta	A, B
3	Charlie	C, D
4	Delta	B, C, D
5	Alpha	A, B, C
6	Beta	A, B, D
7	Delta	C, D

(b) Converted data

Item TID	A	B	C	D
1	1	1	0	1
2	1	1	0	0
3	0	0	1	1
4	0	1	1	1
5	1	1	1	0
6	1	1	0	1
7	0	0	1	1
Count	4	5	4	5

Table 2
Initial segmentation results

TID item	A	B	C	D
(a) Cluster *X*				
1	1	1	0	1
2	1	1	0	0
3	1	1	1	0
4	1	1	0	1
(b) Cluster *Y*				
1	0	0	1	1
2	0	1	1	1
3	0	0	1	1

Table 3
Customer correlation matrix

Rule		Cluster			
		Cluster *X*		Cluster *Y*	
i	$rule_i$	$P(rule_i^X)$	$EV(rule_i^X)$	$P(rule_i^Y)$	$EV(rule_i^Y)$
1	0001	0	0	0	0
2	0010	0	0	0	0
3	0011	0	0	2/7	4/7
4	0100	0	0	0	0
5	0101	0	0	0	0
6	0110	0	0	0	0
7	0111	0	0	1/7	3/7
8	1000	0	0	0	0
9	1001	0	0	0	0
10	1010	0	0	0	0
11	1011	0	0	0	0
12	1100	1/7	2/7	0	0
13	1101	2/7	6/7	0	0
14	1110	1/7	3/7	0	0
15	1111	0	0	0	0
Correlation coefficient (CC)					−0.1721