Expert Systems with Applications 38 (2011) 7186–7191

Contents lists available at ScienceDirect

Expert Systems with Applications

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

# A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business

Der-Chiang Li [a,*], Wen-Li Dai [b], Wan-Ting Tseng [a,c]

[a] Department of Industrial and Information Management, National Cheng Kung University, 1st University Road, Tainan 70101, Taiwan
[b] Department of Information Management, Tainan University of Technology, Taiwan
[c] Nam Liong textile Enterprise Co., Ltd., Taiwan

ARTICLE INFO

*Keywords:*
Customer relationship management
Customer lifetime value
LRFM model
Data mining
Clustering analysis

ABSTRACT

In order to obtain comprehensive information about customers, this study aims to use a systematized analytic method to examine customers. This study uses LRFM customer relationship model, which consists of four dimensions: relation length (L), recent transaction time (R), buying frequency (F), and monetary (M), to carry out customer clusters. We proceed with this clustering analysis to classify customers in order to set discriminative marketing strategies. In addition, this study further employed a cross analysis over three predetermined dimensions: area, sales, and new/old characteristics to enhance the clustering analysis. The results obtained from the real textile business show that the customer groups formed using the four-factor (LRFM) clustering all has statistical significant differences, and with meaningful explanations in terms of marketing strategy. Thus, this study considers useful for discriminative customer relationship management.



## 1. Introduction

Customer behavior analysis is one way for firms to better understand the market and thus discover new business opportunity to pursue. Using systematic data analysis methods to understand and connect with customers becomes an important topic for customer relation management. Most companies have begun to realize that their customer databases are extremely important assets (Athanassopoulos, 2000; Jones, Mothersbaugh, & Beatty, 2000; Thomas, 2001); many use the databases to develop marketing strategy by carrying out the customer characteristic analysis (Kim, Suh, & Hwang, 2003).

This study uses sales transaction data as the basis for the work of knowledge discovery in database (KDD). It applies the LRFM customer relationship model (Chang & Tsay, 2004) to cluster customers into meaningful groups. Where four attributes are included as: (1) recent transaction time: referring to the time of the customer's last transaction; (2) frequency of buying; (3) monetary value: the total value bought during a period; and (4) relationship length.

In addition to carrying out customer classification, we also considers the grouping knowledge by cross analysis the found class information over three other pre-determined customer characteristics: (1) the factories divided into Taiwan-based and Mainland-based; (2) the sales divided into domestic, abroad, and Mainland areas; (3) the customers divided into new and old customers.

In all, this study aims to acquire customer information from: (1) using objective customer clustering technique to find the implicit group knowledge; (2) crossing analysis the objective clustering information over the three-dimension customer characteristic analysis to determine the customer behavioral tendencies. The results should enable us to suggest appropriate marketing strategies for the relevant firms.

## 2. Literature review

This section will survey past research concerning customer relationship management and customer value analysis.

### 2.1. Customer relationship management

Dong, Swain, and Berger (2007) proposes the maximization of customer equity, which is a core objective of customer–company relationship management. Doorn and Verhoef (2008) propose the key factors influencing customer satisfaction, and develop a dynamic model to explain customer loyalty. Krasnikov and Jayachandran (2008) find that marketing capability has a larger influence than research and development ability on enterprise performance and management strategy of customer relationship, and maintenance are the main ability of marketing. Kalakota and Robinson (1999) point out that customer relationship

* Corresponding author. Tel.: +886 6 2757575x53134; fax: +886 6 2362162.
*E-mail addresses:* lidc@mail.ncku.edu.tw (D.-C. Li), wenlit@ms21.hinet.net (W.-L. Dai), candy@namliong.com.tw (W.-T. Tseng).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.12.041

management regard as integrated actions of sales, marketing and service strategies. Tiwana (2001) notes that in markets with intense competition, globalization, high customer flowing rates and high customer obtaining costs, customer relationship could help enterprises improve customer satisfaction and promote customer loyalty.

Richards and Jones (2008) point out an intuition and general concept and claim that to increase customer relationship management should improve the business administration performance. Wong, Chan, Ngai, and Oswald (2009) discuss the relevant customer loyalty influencing factors, and suggest analyze tools to understand and improve customer loyalty strategies; King and Burgess (2008) point out some successes and failures factors influence customer relationship management; Yi and Gong (2008) point out the normal and abnormal behavioral problems influencing customer perception. Lin (2007) point out the customer satisfaction model and concept.

### 2.2. Customer value analysis

Kotler (2000) defines Customer Lifetime Value (abbreviated as CLV) as the profit net present value (NPV) that one can obtain in a customer's lifetime. Berger and Nada (1998) explain the importance of maintaining a customer by comparing customer lifetime value and the necessary cost of attracting a new customer; Kim, Jung, Suh, and Hwang (2006) define customer lifetime value as the net income amount of the business during the entire life cycle of a customer. He emphasizes long-term continued income and cost, instead of the profits from a specific trading activity. Brown (2000) proposes that not all customers are worth keeping, and uses value-based segment theory to determine the limitation resources and efforts to maintain a specific customer's loyalty. He claims that customer value analysis is the foundation of customer relationship management.

Mani, Drew, Betz, and Datta (1999) and Crowder, David, and Wojtek (2007) regard that the customer lifetime value is composed of two independent factors: tenure and value. They point out that CLV is an important concept in the work of customer classification, selection, and retention, because different strategies may apply to different customers.

Miglautsch (2000) and Kaymak (2001) use the RFM model as a way to measure customer lifetime value, and made extensive use of estimated customer value at present. Before carrying out database marketing, enterprises must focus research on the customers' historical trade records in order to obtain references for prediction and as the basis of decisions. Hughes (1994) defines the RFM model using three dimensions: (1) recent transaction time: referring to the time of the customer's last transaction and analyzing the time point distance; (2) buying frequency: how often a customer buy's products in some period, with a higher frequency representing larger demand and higher loyalty; (3) monetary value: the total value bought during a period, with a higher amount representing a greater contribution to the company.

Reinartz and Kumar (2000) propose the idea of customer relation length, and examine its influence on customer loyalty and profitability. They suggest increasing relation length to improve customer loyalty. Chang and Tsay (2004) later propose the LRFM model, mainly adding the customer relation length to RFM model. Benoit and den Poel (2009) led to an interest in understanding and estimating customer lifetime value and relation method. Glady, Baesens, and Croux (2009) propose the approach for predicting customer lifetime value with the Pareto/NBD model.

### 2.3. Knowledge discovery in database and data mining

Knowledge discovery in database (KDD) is a procedure that chooses suitable data for mining knowledge. Fayyad, Piatetsky-Shapiro, and Smyth (1996a,b,c) regard KDD as a method of repeated mining to obtain useful information, while Brachman and Anand (1996) present KDD as a knowledge strengthening process including complicated reciprocation, crossing analysis over time, between people and databases, and supported by a group of numerous and interconnected tools.

Sung and Sang (1998) classify data mining according to functions, as follows: (1) association analysis: put relevant things together; (2) classification: classify according to data attributes; (3) cluster analysis: classify similar data into the same classes; (4) prediction: make use of observation values in the past to estimate the future value of an attribute. Among these, the association, classification, and prediction functions are part of the supervised learning method, while cluster analysis is part of the non-supervised learning method. The main difference between supervised and non-supervised learning is that the former does not need to determine the data cluster in advance; it does the analysis directly and produces the clustering results.

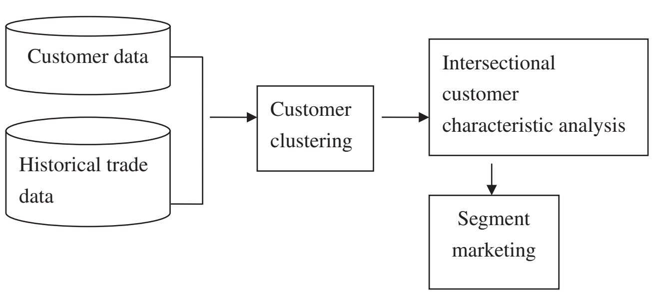


**Fig. 1.** Study frame.

## 3. Research method

This study uses existing data and unsupervised learning method to find implicit knowledge, and through the assistant experience in this professional field to determine significant characteristic attributes for customer relationship management.

### 3.1. Research framework

This study applies the LRFM model to the case study's customer data. The customers are grouped through the two-stage cluster

**Table 1**
Data form with seven attributes.

| | Field name | Data content |
|---|---|---|
| 1 | Area | The area divided into Taiwan and Mainland factories |
| 2 | Sales | The sales divided into domestic, abroad, and Mainland |
| 3 | New/old | The segmentation point regards a customer as old if it that began the trade before 2000 and new if it began the trade after |
| 4 | Transaction length | The interval is between the first and last exchange with a customer |
| 5 | Recent transaction time | From the last transaction time until now measured in years |
| 6 | Annual frequency | The average number of transactions a customer had per year from 2000 to 2006 |
| 7 | Average monetary value | The average monetary value is in each transaction from 2000 to 2006 |

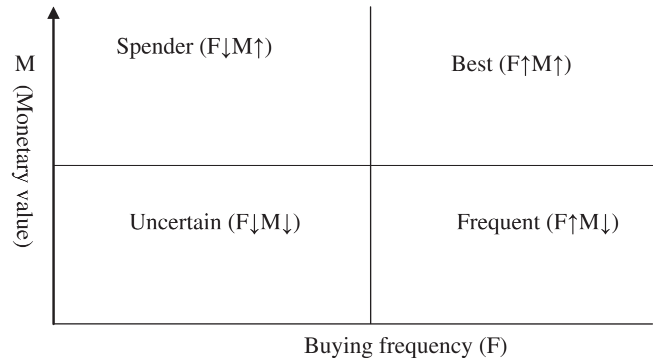


**Fig. 2.** Customer value matrixes (Marcus, 1998).

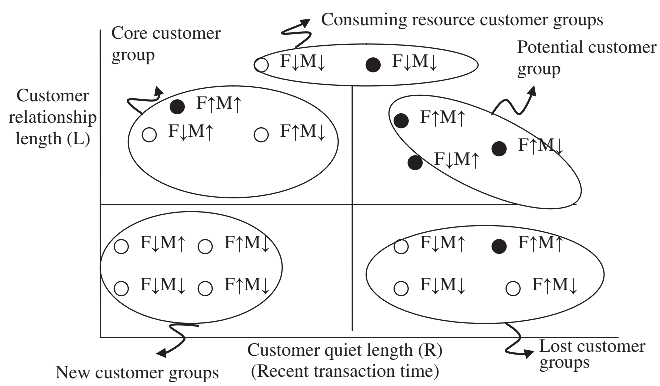


**Fig. 3.** Customer clustering on a customer loyalty matrix basis (Chang & Tsay, 2004).

**Table 2**
Sequence step concentration coefficients of Ward's method.

| Stage | Group | Coefficients | Coefficients difference |
|---|---|---|---|
| 470 | 7 | 419.628 | |
| 471 | 6 | 473.6721 | 54.04413 |
| 472 | 5 | 555.1429 | 81.47076 |
| 473 | 4 | 740.1191 | *184.9762* |
| 474 | 3 | 978.8008 | 238.6817 |
| 475 | 2 | 1306.34 | 327.5397 |
| 476 | 1 | 1719.488 | 413.1478 |

analysis. Where the first stage is the use of Ward's method to estimate the number of clusters, and the second stage will carry out a *K*-mean real clustering operation. The levels of customer value are determined according to the characteristics of each found group.

**Table 3**
Clustering results.

| Cluster | Number | % |
|---|---|---|
| 1 | 26 | 5.45 |
| 2 | 196 | 41.09 |
| 3 | 152 | 31.87 |
| 4 | 17 | 3.56 |
| 5 | 86 | 18.03 |
| Total | 477 | 100 |

Following this, an intersectional relation study executes between the predetermined customer characteristics and the separate groups to find out whether these characteristics are significant across groups. The study procedure shows in Fig. 1.

### *3.2. Data formation for establishing the LRFM model*

This study uses customer lifetime value as the quantitative indicator, and principally uses the LRFM (Chang & Tsay, 2004) model to do the measurement. Because the case study company belongs to the textile industry, where there is seasonal variance in demand, we use annual customer relationship length and recent transaction time unit to avoid this from affecting the analysis. The definition of the LRFM model used in this study shows in Table 1.

The source data is the real exchange sales data in the case company, which has 477 observed values collected in the file. In order to avoid periodic LRFM difference, we standardize the data first, and then transform the customer relationship length, recent exchange time, buying frequency, and monetary values into numerical values ($Z$) for the analysis. The data form presents in Table 1.

### *3.3. Clustering analysis*

This study applies a two-stage method of cluster analysis to the case company data to group customers. The first stage is a use of Ward's method to do clustering group estimation and determine group number, $k$; the second stage uses the $K$-mean clustering operation to separate the data into $k$ groups, conducted using a personal computer with a Pentium 4 processor and SPSS software.

### *3.4. Terms for group description*

Marcus (1998) proposes a customer value matrix, shows in Fig. 2, which uses customer buying frequency (F) and monetary value (M) as the two axes. Two other indicators are customer relationship length (L) and customer recent transaction time (R), these two indicators relate to customer loyalty, and therefore this is defined as the customer loyal matrix. Marcus (1998) claims that the longer a customer relationship, the higher the loyalty; and the

**Table 4**
ANOVA results through clustering.

| | | Total square | Degree of freedom | Average square | *F* test | Significance level | Multiple comparison (Scheffe) |
|---|---|---|---|---|---|---|---|
| Transaction length | Between | 294.006 | 4 | 73.502 | 189.589 | .000 | 1–3, 1–4, 2–3 |
| | Within | 182.989 | 472 | .388 | | | 2–4, 3–1, 3–2 |
| | Total | 476.996 | 476 | | | | 3–5, 4–5 |
| Recent transaction time | Between | 375.210 | 4 | 93.803 | 436.292 | .000 | 1–5, 2–5, 3–5 |
| | Within | 101.480 | 472 | .215 | | | 4–5 |
| | Total | 476.690 | 476 | | | | |
| Annual frequency | Between | 209.711 | 4 | 52.428 | 311.902 | .000 | 1–4, 2–4, 3–4 |
| | Within | 79.339 | 472 | .168 | | | 4–5 |
| | Total | 289.050 | 476 | | | | |
| Average monetary value | Between | 312.515 | 4 | 78.129 | 224.534 | .000 | 1–2, 1–3, 1–4 |
| | Within | 164.237 | 472 | .348 | | | 1–5, 2–3 |
| | Total | 476.752 | 476 | | | | |

shorter the recent transaction time, the greater the customer loyalty. Through buying frequency and monetary value one can form four quadrants in the first plane; and customer relation length and customer recent transaction time, one can form another four quadrants in the second plane. Consequentially, using the customer value and customer loyal matrices one can form 16 quadrants to explain the result of clustering.

This study refers to Sung and Sang's (1998) customer segment description and uses the up symbol (↑) to represent when the group's average value is larger than the total average value; and the down symbol (↓) to represent when the group's average value is smaller than the total average value.

Chang and Tsay (2004) further propose customer classification by summing the 16 groups to five kinds of customer groups, as Fig. 3 shows, including: (1) core customers: including high value loyal customers (LRFM, ↑↓↑↑), high frequency buying customers (LRFM, ↑↓↑↓), and platinum customers (LRFM, ↑↓↓↑); (2) potential customers: including potential loyal customers (LRFM, ↑↑↑↑), potential high frequency customers (LRFM, ↑↑↑↓), and potential consumption customers (LRFM, ↑↑↓↑); (3) lost customers: including high value lost customers (LRFM, ↓↑↑↑), frequency lost customers (LRFM, ↓↑↑↓), consumption lost customers (LRFM, ↓↑↓↑), and uncertain lost customers (LRFM, ↓↑↓↓); (4) new customer groups: including high value new customers (LRFM, ↓↓↑↑), frequency promotion customers (LRFM, ↓↓↑↓), spender promotion customers (LRFM, ↓↓↓↑), and uncertain new customers (LRFM, ↓↓↓↓); (5) consuming resource customers: including low consumption cost customers (LRFM, ↑↓↓↓), high consumption cost customers (LRFM, ↑↑↓↓).

### 3.5. Cross analysis over groups and characteristics

This study defined three other kinds of practically useful customer characteristics: (1) the area divided into Taiwan and Mainland factories; (2) the sales divided to domestic, abroad and Mainland; (3) the customers divided into new and old customers. With these individual characteristics, a cross analysis is made with the clustering results obtains in Section 3.3.

## 4. Experimental analysis

The case study company is a textile manufacturing business with over 30 years of trading history. The case company originally relied mainly on Taiwan based factories, and it began to enlarge and invested in Mainland production market 10 years ago. This study aims at finding answers for the following questions: (1) whether operating performance has any differences in different production areas (Taiwan/Mainland); (2) whether operating performance has any difference in different sale channels (domestic/international/Mainland); (3) whether operating performance has any differences in new or old customers (new/old).

### 4.1. Data clustering

The first stage is applying Ward's method of SPSS hierarchical cluster analysis to the standardized LRFM data of the case company. According to the sequence step of concentrate coefficients of Ward's method, as shows in Table 2, we can estimate the best number of clusters.

From Table 2 we find that the coefficient in stage 473, showing that it needs to increase to 184.9762 to concentrate. Since this is a large amount, we decide to stop further concentration and choose to divide the data into five groups. In fact, in Table 2, stages 473, 474, 475, 476 all have great gaps, and thus this study chooses five groups and proceeds to the second stage, the *K*-mean clustering operation.

**Table 5**
Descriptive statistics quantity after standardization (Z value) for the whole data set.

| | Number | Minimum | Maximum | Average | Standard deviation |
|---|---|---|---|---|---|
| Transaction length | 477 | −1.46684 | 2.68844 | −0.000134 | 1.0010456 |
| Recent transaction time | 477 | −.55591 | 3.18670 | 0.001165 | 1.0007249 |
| Annual frequency | 477 | −.38614 | 6.49171 | −0.028711 | .7792613 |
| Average monetary value | 477 | −.59980 | 7.36313 | 0.001043 | 1.0007896 |

**Table 6**
Descriptive statistics quantity of group 1.

| | Number | Minimum | Maximum | Average | Standard deviation | Symbol |
|---|---|---|---|---|---|---|
| Transaction length | 26 | −1.02944 | 1.15755 | −.2724056 | .6192869 | ↓ |
| Recent transaction time | 26 | −.55591 | .19261 | −.4407542 | .2754165 | ↓ |
| Annual frequency | 26 | −.38614 | .22432 | −.3230091 | .1159976 | ↓ |
| Average monetary value | 26 | 1.82570 | 7.36313 | 3.3401102 | 1.3765168 | ↑ |

**Table 7**
Group summary description.

| Group | Group name | LRFM index |
|---|---|---|
| 1 | Spender promotion (new) customers | Transaction length↓, Recent transaction time↓, belongs to new customer characteristic; Annual frequency↓, Average monetary value↑, means that it has good customer performance contribution but the frequency is lower |
| 2 | Uncertain new customers | Transaction length↓, Recent transaction time↓, belongs to the new customer characteristic; Annual frequency↓, Average monetary value↓, belongs to the uncertain customer characteristic |
| 3 | Long term low contribution customers | Transaction length↑, Recent transaction time↓, means there is a tight customer relationship; Annual frequency↓, Average monetary↓, means the contribution to the company small |
| 4 | Core customers | Transaction length↑, Recent transaction time↓, means there is a tight customer relationship; Annual frequency↑, Average monetary↓, Annual frequency is higher on average but average value monetary is smaller on average |
| 5 | Lost customers | Transaction length↓, Recent transaction time↑, means the customer and company has not exchanged recently: Annual frequency↓, Average monetary↓, belongs to the uncertain customer characteristic |

We use the *K*-mean method of SPSS to cluster the data. According to the results of the first stage, the cluster analysis has determined the five clusters, as shows in Table 3. A homogeneity examination of the variability is before carrying on the one-way ANOVA to test the cluster result. From Table 4, we can find the significance levels of the four factors: transaction length, recent transaction time, annual frequency, and average monetary are all smaller than 0.05, and thus all reach the significance level. In addition, we use the Scheffe method of posteriori comparison to do multiple tests in order to find the difference between a single factor and the average. This results show that each factor has significant differences.

By comparing the descriptive statistics quantity of each group with the statistics for whole data set, in Table 5, we use the terms defined in Section 3.4 to state the group description for the five groups found, shows in Table 7.

For example, Table 6 shows the descriptive statistics quantity for group 1; we can find the intersectional characteristics through customer loyalty and customer value matrixes to state the group description for group 1.

**Table 8**
Three characteristics and group observation cross table.

| Area group | | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|---|
| *(a)* | | | | | | | |
| Area | Taiwan | 26 | 133 | 140 | 11 | 61 | 371 |
| | % | 7.00 | 35.80 | 37.70 | 3.00 | 16.40 | 100.00 |
| | Mainland | 0 | 63 | 12 | 6 | 25 | 106 |
| | % | 0 | 59.40 | 11.30 | 5.70 | 23.60 | 100.00 |
| Total | | 26 | 196 | 152 | 17 | 86 | 477 |
| | % | 5.50 | 41.10 | 31.90 | 3.60 | 18.00 | 100.00 |
| *(b)* | | | | | | | |
| Sales group | | 1 | 2 | 3 | 4 | 5 | Total |
| Sales | Domestic | 1 | 44 | 118 | 11 | 38 | 212 |
| | % | 0.50 | 20.80 | 55.70 | 5.20 | 17.90 | 100.00 |
| | Abroad | 25 | 89 | 22 | 0 | 23 | 159 |
| | % | 15.70 | 56.00 | 13.80 | 0 | 14.50 | 100.00 |
| | Mainland | 0 | 63 | 12 | 6 | 25 | 106 |
| | % | 0 | 59.40 | 11.30 | 5.70 | 23.60 | 100.00 |
| Total | | 26 | 196 | 152 | 17 | 86 | 477 |
| | % | 5.50 | 41.10 | 31.90 | 3.60 | 18.00 | 100.00 |
| *(c)* | | | | | | | |
| New/old group | | 1 | 2 | 3 | 4 | 5 | Total |
| New/old customers | New | 8 | 15 | 140 | 12 | 49 | 224 |
| | % | 3.60 | 6.70 | 62.50 | 5.40 | 21.90 | 100.00 |
| | Old | 18 | 181 | 12 | 5 | 37 | 253 |
| | % | 7.10 | 71.50 | 4.70 | 2.00 | 14.60 | 100.00 |
| Total | | 26 | 196 | 152 | 17 | 86 | 477 |
| | % | 5.50 | 41.10 | 31.90 | 3.60 | 18.00 | 100.00 |

**Table 9**
Different area performance results.

| Group | Group name | Difference description |
|---|---|---|
| 1 | Spender promotion (new) customers | The potential customer ratio is 7.0% in Taiwan factories, which is larger than the 0% in Mainland, representing that all the group's factories are in Taiwan |
| 2 | Uncertain new customers | The uncertain new customer ratio is 59.4% in Mainland factories, which is larger than the 35.8% ratio in Taiwan factories, representing that many factories of this group's factories are in the Mainland |
| 3 | Long term low contribution customers | The long-term low contribution customer's ratio is 37.7% in Taiwan factories, which is larger than the 11.3% ratio in Mainland factories, representing that there are more long-term loyal customers in the Taiwan factories |
| 4 | Core customers | The core customer groups' ratio is 5.7% in Mainland factories, which is larger than that of 3.0% ratio in Taiwan factories |
| 5 | Lost customers | The lost customer groups' ratio is 16.4% in Taiwan factories, which is smaller than the 23.6% ratio in Mainland factories, representing that customer stability was good in the Taiwan factories |

**Table 10**
Different sales performance result.

| Group | Group name | Difference description |
|---|---|---|
| 1 | Spender promotion (new) customers | The potential customers' abroad sales ratio is 15.7% and domestic sales ratio is 0.5% in Taiwan; the sales ratio of potential customer is 0% in the Mainland, representing that the abroad sales ratio is higher in Taiwan |
| 2 | Uncertain new customers | The mainland sales ratio of uncertain new customers is 59.4%; the uncertain new customers' abroad sales ratio is 56.0%, and the domestic sales ratio is 20.8% in Taiwan |
| 3 | Long term low contribution customers | The long-term low contribution customers' domestic sale ratio is 55.7% and the abroad sale ratio is 13.8% in Taiwan; the Mainland sale ratio of long-term low contribution customers is 11.3% |
| 4 | Core customer | The core customers' domestic sales ratio is 5.2% and the abroad sale ratio is 0% in Taiwan; the Mainland core customer sales ratio is 5.7% |
| 5 | Lost customers | The lost customers' domestic sales ratio is 17.9% and the abroad sales ratio is 14.5% in Taiwan; Mainland sale of lost customer ratio is 23.6%<br>This means that the lost customer ratio of domestic and abroad sales in Taiwan is all smaller than that in the Mainland |

**Table 11**
Different new/old customer performance result.

| Group | Group name | Difference description |
|---|---|---|
| 1 | Spender promotion (new) customers | The new customer ratio is 7.1%, larger than the 3.6% for the old customer ratio in potential customer groups. It means that in this group the number of new customers is comparatively large |
| 2 | Uncertain new customers | The new customer ratio is 71.5%, larger than the 6.7% old customer ratio in the uncertain new customer group |
| 3 | Long term low contribution customers | The old customer ratio is 62.5% and larger than 4.7% new customer ratio in long-term low contribution customers groups |
| 4 | Core customers | The old customer ratio is 5.4%, larger than the 2.0% new customer ratio in the core customer group, meaning that long term and highly loyal customers have a high contribution for the company |
| 5 | Lost customers | The old customer ratio is 21.9%, which is larger than the 14.6% new customer ratio in the lost customers group, meaning that old customers' loyalty needs to be strengthening |

### *4.2. Cross analysis over groups and customer characteristics*

We use three dimensions: area, sales, and new/old as characteristics to intersect the five clusters for further relational analysis.

The three characteristics and group cross-table analysis shows in Table 8.

The dimensional characteristics analysis shows as follows:

Analysis 1: Area characteristic analysis.

The area dimension divides the factories into Taiwan-based and the Mainland-base, and the results shows in Table 9.

Analysis 2: Sales characteristic analysis.

The sales dimension is divided into domestic, abroad, and Mainland areas. The result shows in Table 10.

Analysis 3: New or old customer characteristic analysis

The study divides customer dimension into new and old customers to explore whether any differences exist for customers starting transactions at different time points. Table 11 shows the results.

## 5. Conclusion and discussion

The results of this study can provide a company with further understanding of customers for making segmented marketing strategies. After the preliminary clustering, the new customer is a group with the largest percentage (41.1%); meaning that the company operating performance is growing continuously and is worth giving an encouragement. In the long-term low contribution group and core customer group, both have lower than average transaction monetary value and are different only in the annual frequency of trades, meaning that the long-term customers in the transaction monetary contributing are worth strengthening.

From the cross characteristics analysis, we find that the Taiwan factories with domestic sales have more long-term customers with greater company loyalty, although monetary value and frequency are not high.

The cross customer characteristics analysis of objective clustering and the predetermined dimensions results provides multiple-dimensional information for us. In further study, it is worthy applying this model to other data with relevant attributes. On the method of clustering, this study uses Ward's and *K*-mean methods to cluster data, and further study can use other methods to compare with this study to find whether there are differences or not.

## References


Athanassopoulos, A. D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research, 47*(3), 191–207.

Benoit, D. F., & den Poel, D. V. (2009). Benefits of quantile regression for the analysis of customer lifetime value in a contractual setting: An application in financial services. *Expert Systems with Applications, 36*(7), 10475–10484.

Berger, P. D., & Nada, I. N. (1998). Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing, 12*(1), 17–30.

Brachman, R. J., & Anand, T. (1996). The process of knowledge discovery in databases. In *Advances in knowledge discovery and data mining* (pp. 37–57). Cambridge, MA: AAAI/MIT Press.

Brown, S. A. (2000). *Customer relationship management: A strategic imperative in the world of e-business*. Canada: John Wiley & Sons.

Chang, H. H., & Tsay, S. F. (2004). Integrating of SOM and *K*-mean in data mining clustering: An empirical study of CRM and profitability evaluation. *Journal of Information Management, 11*(4), 161–203.

Crowder, M., David, J. H., & Wojtek, K. (2007). On optimal intervention for customer lifetime value. *European Journal of Operational Research, 183*(3), 1550–1559.

Dong, W., Swain, S. D., & Berger, P. D. (2007). The role of channel quality in customer equity management. *Journal of Business Research, 60*(May), 1243–1252.

Doorn, J. V., & Verhoef, P. C. (2008). Critical incidents and the impact of satisfaction on customer share. *Journal of Marketing, 72*(July), 123–142.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996a). *Advances in knowledge discovery and data mining*. Menlo Park, CA, USA: American Association for Artificial Intelligence.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996b). Data mining and knowledge discovery in databases. *AI Magazine*, 37–54.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996c). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM, 39*(11), 27–34.

Glady, N., Baesens, B., & Croux, C. (2009). A modified Pareto/NBD approach for predicting customer lifetime value. *Expert Systems with Applications, 36*(2 (Part 1)), 2062–2071.

Hughes, A. M. (1994). *Strategic Database Marketing*. Probus Publishing: Chicago.

Jones, M. A., Mothersbaugh, D. L., & Beatty, S. E. (2000). Switching barriers and repurchase intentions in services. *Journal of Retailing, 76*(2), 259–374.

Kalakota, R., & Robinson, M. (1999). *E-business: Roadmap for success*. New Jersey: Addison-Wesley.

Kaymak, U. (2001). Fuzzy target selection using RFM variables. In *IFSA world congress and 20th NAFIPS international conference, joint 9th* (pp. 1038–1043).

Krasnikov, A., & Jayachandran, S. (2008). The relative impact of marketing, research-and-development, and operations capabilities on firm performance. *Journal of Marketing, 72*(July), 1–11.

Kim, J., Suh, E.-H., & Hwang, H. (2003). A model for evaluating the effectiveness of CRM using the balanced scorecard. *Journal of Interactive Marketing, 17*(2), 5–19.

Kim, S.-Y., Jung, T.-S., Suh, E.-H., & Hwang, H.-S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications, 31*(1), 101–107.

King, S. F., & Burgess, T. F. (2008). Understanding success and failure in customer relationship management. *Industrial Marketing Management, 37*, 421–431.

Kotler, P. (2000). *Marketing management*. New York: Prentice Hall International, Inc..

Lin, W.-B. (2007). The exploration of customer satisfaction model from a comprehensive perspective. *Expert Systems with Applications, 33*(1), 110–121.

Mani, D. R., Drew, J., Betz, A., & Datta, P. (1999). Statistics and data mining techniques for lifetime value modeling. In *Conference on knowledge discovery in data archive, proceedings of the fifth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 94–103).

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. *Journal of Database Marketing, 8*(1), 67–72.

Reinartz, W. J., & Kumar, V. (2000). On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. *Journal of Marketing, 64*(October), 17–35.

Richards, K. A., & Jones, E. (2008). Customer relationship management: Finding value drivers. *Industrial Marketing Management, 37*, 120–130.

Sung, H. H., & Sang, C. P. (1998). Application of data mining tools to hotel data mart on the Intranet for database marketing. *Expert Systems with Applications, 15*(1), 1–31.

Thomas, J. S. (2001). A methodology for linking customer acquisition to customer retention. *Journal of Marketing Research, 38*(2), 262–268.

Tiwana, A. (2001). *The essential guide to knowledge management: E-business and CRM application*. New Jersey: Prentice Hall.

Wong, Y. H., Chan, R. Y. K., Ngai, E. W. T., & Oswald, P. (2009). Is customer loyalty vulnerability-based? An empirical study of a Chinese capital-intensive manufacturing industry. *Industrial Marketing Management, 38*, 83–93.

Yi, Y., & Gong, T. (2008). The effects of customer justice perception and affect on customer citizenship behavior and customer dysfunctional behavior. *Industrial Marketing Management, 37*, 767–783.