Developing Strategies for Blooming Marketing of a Local Business – An Integrated Approach

Chung-Wei Yeh\(^1\) and Kee-Rong Wu\(^2\)\(^+\)

\(^1\) Dept. of Management Information System, Kao Yuan University, Kaohsiung, Taiwan 821
\(^2\) Dept. of Marine Engineering, National Kaohsiung Marine University, Kaohsiung, Taiwan 811

Abstract. This study presents a customized service approach by integrating techniques of recency, frequency and monetary (RFM) method and data mining to provide the decision makers for customer relationship management (CRM) strategies for a local mid-size chained store. With the mature and diverse computing technologies, to obtain consumer data for firms becomes much easier than ever before and to use these data to develop CRM strategies to significantly increase the transaction and generate more profit becomes more critical. To maintain loyal and valued customer relationships is a key part in the strategy design. Therefore, in this paper, we propose a framework for analyzing customer value and segmenting customers based on the RFM method. Strategies according to customer segments will be determined then. Association rules obtained by analyzing the group transaction data are also implemented into the strategies. This work provides the solutions via four steps: (1) Data preprocessing, (2) Customers grouping, (3) Data analysis and (4) Application. In this empirical study we demonstrate the proposed approach to construct a marketing application for organizations. This research offers managerial applications that can be used by marketers to increase profit.

Keywords: Marketing Strategies, RFM, Customer Loyalty, Data Mining.

1. Introduction

Over the past decade, information technologies (IT) have provided many competitive advantages such as agility, selectivity, intelligence, interactivity and also have enabled the business of collecting massive amount of instant data. Thus, incorporating IT and marketing strategy to strength the relationship between the business and customers has become one of the top priority tasks for the decision makers. For providing better decision making in business application, various data mining techniques have been developed to analyze massive amount of data, which are treated as the process of extracting implicit, unknown and potentially useful information from databases. Cluster analysis, unsupervised classification are used to segment similar data into clusters in data mining. One such technique is RFM method, proposed by Hughes [1], while R, F, and M stand for recency, frequency and monetary, separately. RFM has considered among the most popular segment technique used by decision makers [2] since it uses three consumer behavior variables to evaluate difference contribution for each consumer. Thus, data mining researches have been carried out based on RFM values in CRM domain. Since each consumer’s contribution is estimated by RFM, one can segment consumers into groups by similar values and then build predictive consumer modeling. While there could be other consumer-related factors, such as promotion or campaign, studies [3, 4] showed that the application of RFM could provide the prediction of the future consumer purchase considerably. Thus, the integration of RFM method with data mining techniques will benefit the business by CRM design.

In this study, we integrate profit factor with RFM and then segment customers into groups in term of profit and RFM variables, as group attributes, to present each customer’s value by its own recency, frequency, and monetary scores to represent customer value, then identified group members. Finally, we propose CRM group strategies accompanying with association rules from the transactions to improve the close relationships of enterprises with customers for the business. The remainder of this paper is organized as follows. Section 2 presents the related techniques for CRM. Section 3 describes the methodology and section...
4 presents an empirical case conducted in this study with respect to dataset and research framework. Section 5 ends the paper with conclusions.

2. Technologies

This section mainly explores the techniques for developing CRM strategies, including Data mining, customer relationship management, customer value analysis and the RFM method.

3. Data Mining

Data mining techniques have been applied to analyze customer behavior and characteristics to solve many business problems. Using six types of models, e.g. classification, regression, time series, clustering, association analysis, and sequence discovery, data mining is wildly used in many CRM-related researches [5,6]. The first two, classification and regression, are used to make predictions and have been carried out for various purposes in the CRM domain, while association and sequence discovery are used to describe behavior. Clustering can be used for either forecasting or description. Industrial companies can gain a competitive advantage by mining their databases for valuable transaction information. A survey [5] was conducted in the retail and service industries in Japan support that a firm’s relationship orientation had a positive effect on CRM implementation (data warehousing, data mining, using customer data for decision making). Research in [7] adopted decision tree to classify the customers and developed CRM strategies based on customer lifetime value. It suggested a model for customer base segmentation.

Other data mining techniques that are useful for the analysis of customer data are association rules and sequential patterns analyses. Researches on the utilization of association rules and sequential patterns have been conducted for various purposes in the CRM domain. Research in [8] has conducted market basket analysis to identify association rules and presented a framework for building behavioral profiles of individual users. It also has introduced an approach to expert-driven validation of a very large number of rules pertaining to these users.

3.1. Customer Relationship Management (CRM)

In brief, CRM is viewed as a holistic business and information systems strategy with the aim to enable organizations to establish a stronger customer relationship. CRM is to integrate the function of the related fields with customer in the enterprise, such as marketing, sales, services and technical support for customer needs, and it usually utilizes IT to help an enterprise managing relationships with customer in a systematic way, improving customer loyalty and increasing overall business profits. CRM is not narrowly defined as promotional marketing based on customer database. Its ultimate goal is to collect customer information with every customer interaction [9]. This concept presents a marketing relationship that each customer can be analyzed in order to understand future purchases and individual profit potential for the firm. Instead of developing marketing schemes for entire segments, it encourages firms to address individual customer needs and to analyze lifetime customer value. Enterprises can shorten sales cycle and increase customer loyalty to build better close relationships with customers and further add revenues by good CRM. Thus, a good CRM can help enterprises keeping existing customers and also attracting new ones. Enterprises apply some methods to effectively enhance customer relationships, which include customer value analysis, enterprise strategy, and positive service mechanisms. Moreover, enterprises also strengthen marketing and sales effectiveness in order to build good CRM.

3.2. Customer value analysis and RFM method

Customer value analysis is an analytic method for discovering customers’ characteristics and makes a further analysis of specific customers to abstract useful knowledge from large data. Thus, it is clear that enterprises apply value analysis to customers for identifying that who are the target customers and whose contribution is outstanding. When determining lifetime customer value, several factors are considered, including customer purchase history, contribution margin, and variable marketing costs. Each customer’s past profit is calculated by adding up past profit margins of all purchases and then subtracting the variable costs associated with that customer. Customers are often ranked by their lifetime customer values, allowing firms to distinguish the most valuable customers and to allocate necessary resources efficiently among them.
By taking the advantage of extracting characteristics of customers by using fewer criterions (a three-dimension) as cluster attributes, the RFM is one of the well-known customer value analysis methods. Moreover, from the consuming behavioral point of view, enterprises are intent via using RFM analysis to mine databases for knowing about customers who spend the most money and create the biggest value for the enterprises. The RFM analytic method is proposed by Hughes in 1994 [1] which differentiates important customers based on the values of three variables, i.e., recency of the last purchase (R), frequency (F) and money amount (M). The detail definitions are described as follows:

- **Interval of customer consumption (R).** R represents recency, which refers to the interval between the time that the latest purchasing behavior happens and present. The shorter the interval is, the bigger R is. It indicates that the bigger the value of R is, the more likely the corresponding customers are to make a new trade with companies.

- **Frequency of the purchases (F).** F represents frequency, which refers to the number of transactions over a certain period of time, for example, two times over one month. The many the frequency is, the bigger F is. Similar to recency, it indicates that the bigger the value of F is, the more likely the corresponding customers are to make a new trade with companies.

- **Monetary value of the purchases (M).** M represents monetary, which refers to consumption money amount spent on products or services over a certain period of time. The much the monetary is, the bigger the M is. It indicates that the bigger M is, the more likely the corresponding customers are to buy products from the same company again.

4. **Methodology**

This section presents the proposed approach and associated procedures for classifying customer value and CRM strategies. It consists of four essential phases: (1) Data preprocessing; (2) Customer grouping; (3) Data analysis; (4) Application. Figure 1 depicts the framework of our approach. We first collect data and redesign the database structure, generate program to perform customer grouping mainly by RFM application and profit factor for grouping. Procedures to execute CRM strategies are then designed according to the analysis and association rules. The executing steps are listed as follows.

Step 1: Data preprocessing. While preprocessing data, we delete the duplicated records, those with many missing or inaccurate values. Data from a quick survey is also collected on every transaction. The collected data including customer basic data, transaction data, product data and survey data are imported into SQL server database. The customer’s value is designed to be calculated monthly based on the data over the past two months. Data collection is constructed on relational database structure applied in the system. All collected source data will be preprocessed to delete the duplicated, missing or inaccurate value records and
will be transferred into SQL database. We design seven main tables (Customer, Trans_detail, Trans_date, Product, Product_price, category and Survey) to mine the purchase behaviors of customers.

Step 2: Customer grouping. By defining the data collecting cycle to be two months, we adopt RFM and WMLE method and calculate R, F and M values. We integrate Maximum Likelihood Estimation (MLE) and WMLE (Weighted Maximum Likelihood Estimation) [10] to analyze customer’s value tendency. The difference between MLE and WMLE is considered to determine the customer’s loyalty. MLE is an average interval of purchasing and provides estimates for the model’s parameters. WMLE stands for the weighted calculation of MLE. We define \( v \) as the value \( (\nu = \text{MLE-WMLE}) \) to be used to represent the customer’s value tendency. If \( v \) is greater than 0, which means the time interval is becoming shorter, otherwise, becoming longer. When time interval become shorter means the customer make purchasing more frequent, which is used as an indicator of his loyalty to the business in the proposed model. The sale profit of each product is programmed as a weight factor in the design as well. An attribute named “profit” calculated from Product_price table is another factor to calculate monthly profit. Four groups of customers are defined, loyal, loyal, potential, conservative and disloyal customers. We define customer grouping by customer’s value which is based on RFM and customer’s value tendency. As defined, R refers to the interval between the time that the latest purchasing behavior happens and present. F refers to the number of transactions over two months. Monetary refers to the average consumption money amount spent on products over the two months. P refers to the sale profit made by individual customer over two months. Programs are designed to segment customers into groups based on RFM method. For every customer, R, F, M and P values are calculated.

Step 3: Data analysis. Association rule in SQL Server is applied to generate purchasing rules for different groups, which basically adopt Apriori algorithm to mine association rules for specific purchase interest in this phase. The association rule algorithm is basically used to search the related items or features in database. The association rule is implication of the form \( X \rightarrow Y \), where \( X \subset I, Y \subset I \) and \( X \cap Y = \phi \). If \( D \) represents a set of transactions, then the rule \( X \rightarrow Y \) holds in the transaction set \( D \) according to two measurements, i.e., support and confidence. Support represents the rate of transactions in \( D \), that contains item set \( X \), and is denoted as \( \text{Sup}(X,D) \). Support is used to evaluate the statistical importance of \( D \). Thus, the rule \( X \rightarrow Y \) which has support \( \text{Sup}(X \cup Y,D) \) represents the rate of transactions in \( D \) containing \( X \cup Y \). Confidence represents the rate of transactions in \( D \), that contains \( X \) and \( Y \), and is denoted as \( \text{Conf}(X \rightarrow Y) \). The relationship between support and confidence is: \( \text{Conf}(X \rightarrow Y) = \frac{\text{Sup}(X \cup Y,D)}{\text{Sup}(X,D)} \). It means the probability of that the transaction includes \( X \), the chance that the transaction also contains \( Y \). Confidence is used to evaluate the level of certainty of the association rule \( X \rightarrow Y \). Given a set of transaction \( D \), association rule generates all transaction rules that have certain Minimum support \( (\text{Min sup}) \) and Minimum confidence \( (\text{Min conf}) \) [11], [12].

Step 4: Application. The CRM strategies for each group customers are generated by reference to the content of rules. Also, we reason for those not loyal customers and consult with business consideration to propose strategies for each group customers. In order to satisfy the needs of customers and enhance the relationship with customers for company, CRM strategies are developed in the application phase. Customers are separated into four groups in our approach. Loyal customers are considered to be those bring most profit to the company. The main strategy to the loyal customers is to provide them cash rebate. To the disloyal customers, system adopts customer purchase suggestion and promotion strategy. To design the potential customer strategy, customers are separated into three sub-groups in our approach based on the potential customer criteria in Fig. 2. For the potential customers, either their frequency or monetary or profit is less than its corresponding average. Thus we summarize three sub-groups – Premium Customers/busy (wealthy customers with no time for shopping), Premium Customers/needless (wealthy customers not fully satisfactory with our products) and regular customers (budgetary customers). To corresponding customer groups, five strategies are designed in the CRM, which are cash rebate, delivery service, collaborative marketing, customer purchase suggestion and promotion.

5. Empirical Study

The developed methodology is implemented in a local mid-size chained store - Gotop. The firm, located in Shih-Lin and Baitoe, near Taipei city, was founded in 1995 and is one of the popular mid-size chained
Data preprocessing: To build up the content of the mining database, we preprocess customer information from data collected between 1/1/2010 and 12/31/2012, showing a total of 267,204 transaction records which belonged to 5,268 customers. The membership customers count for 90% of the total sales, which indicates that the collected data can highly represent the customers of Gotop. The original data of Gotop were stored in Access. Seven main tables (Customer, Trans_detail, Trans_date, Product, Product_price, Category and Survey) were redesigned to mine the purchase behaviors of customers.

Customer grouping: After executing grouping programs, 1844 customers are loyal customers, 1501 customers are potential, 1043 customers are conservative and 880 customers are disloyal. Thus, the distribution of loyal, potential, conservative and disloyal are 35%, 28.5%, 19.8% and 16.7%, respectively.

Data Analysis: On phase three, ninety-four rules were generated from the association rule mining. There are 25, 18, 31, 20 rules created for four groups, respectively.

Application: The CRM strategies are developed in the application phase. For each of the four groups of customers, different strategies are performed based on the scheme shown in Fig. 2. Basically, we apply association rule algorithm to find out the four different group customers’ purchase patterns and predict their next purchases and then promote associated products which may attract their attention and encourage their purchasing intention. For example, if the probability of purchasing Kleenex toilet paper and meats is high, the system will generate a promotion scheme focusing on soft drinks. Besides, additional strategies are proposed for the four groups. For loyal customers, the cash rebate is designed to be 5% to 10% of each transaction amount. For the budgetary and disloyal customers, basically, promotion is the main strategy.

6. Conclusions

This paper proposes an approach for providing CRM strategies for a mid-size chained store. It aims at analysis of customer values by mix of RFM and other factor to segment the customers and design programs to analyze more precise customer grouping. The result of experiment shows the predictable performance of the proposed approach. In comparison with the average month profit data over year 2010 and 2011, the proposed approach achieves about 8% better performance in the monthly profit in 2012. The main contribution of the experimental result shows that (1) a dedicated analysis of customer grouping, purchase patterns and the generation of customer group strategies potentially benefit the mid-sized enterprises in the highly worldwide competition environment. (2) the proposed approach achieves 8% improvement in the monthly profit for Gotop for the past two years, as compared with the conventional mean.
The approach is currently implemented as a prototype. Further research will focus on two directions. (1) A dedicated customized mining system for the business is being developed; (2) The cash rebate is designed to be 5% to 10% of each transaction amount for loyal customers at the present time. An accurate profit calculation procedure is under taking to offer a more profitable cash rebate rate since the profit mostly comes from the loyal customers of Gotop.

7. 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 102 - 2221 - E - 244 - 005.

8. References