African Journal of Business Management Vol. 5(14), pp. 5567-5572, 18 July, 2011
Available online at http://www.academicjournals.org/AJBM
DOI: 10.5897/AJBM11.290
ISSN 1993-8233 ©2011 Academic Journals

Full Length Research Paper

Determination of customer value measurement model RFM index weights

Liu Wei-jiang[1*], Duan Shu-yong[1], Yang Xue[2] and Wang Xiao-feng[2]

[1]Business School, Jilin University, China.
[2]Population Resource and Environment Research Institute, Jilin University, China.

Accepted 22 April, 2011

How to measure customer value and segment customers in order to do target marketing is a problem that many enterprises concern about. As a method of measure customer value, recency, frequency, monetary (RFM) had applied in many enterprises. Usually index weights of RFM model were determined by analytic hierarchy process (AHP) through experts' scores, this method was affected by experts' subjectivity. This paper presented a method of starting from enterprise' sale data, constrained by the fact the sum of the index weights were 1. According to rule of the index weights of RFM under the biggest response rate were the enterprise's best RFM index weights, objectively determined RFM index weights. This method made the determination of RFM index weights conform to the enterprise objective situation, thus it provided a feasible method for the enterprise customer management.

Key words: RFM model, analytic hierarchy process, customer value.

INTRODUCTION

Customers are obviously the source of revenue for any enterprise and winning new customers or encouraging precious customers to purchase more is the primary objective of every enterprise. In this process, the 20/80 principle (Schmittlein, 1993) first postulated by Italian economist Pareto (1906) is still valid. This principle states that most enterprise profits come from a few valuable customers, in fact 80% of the revenue comes from only 20 of the customers. Hence, identifying the most valuable customers of an enterprise and based on their purchasing characteristics and then to recommend other goods suiting their preferences in order to enhance future purchases for enterprise is a continually pursued goal.

Bob Stone (1994) suggested using recency, frequency, monetary (RFM) model to assess customer value. Bob Stone (1994) pointed out that in a period time, customers who purchase more recently are more inclined to buy again than those purchased earlier. Customers who purchase often are more inclined to buy again than those purchased infrequent. Customers who spend more are more inclined to buy again than those who spend less. This indicates that a relationship between customers' past purchasing behavior and their future purchase probabilities exists. Many scholars and business experts as well as most businesses would like to be able to predict and/or encourage future sales based on previous customer activity. Bob thinks the weights of various indexes in the RFM model are different and are dependent on industry characteristics. Hughes (1994) thinks that the various index weights in the RFM model are consistent cross industries. Currently there is a lot of literatures (Satty, 1980; Liu et al., 2005; Shen et al., 2009; Deng et al., 2008) on the use of the analytic hierarchy process (AHP) to determine the weight of each index in the RFM model, but this method depend on the expert's experience and is therefore subjective. Experts can be chosen from administrators, market consultant, or customers in order to obtain various representative aspects on customer preferences, but the results will all be subjective based on the expert's individual thinking. Because different index weights on RFM make enterprise have different segment customer methods, thus, how to get proper index weights on FRM is very important for enterprises to find most valuable customers and do effective customer management so that improve enterprise's

*Corresponding author. E-mail: liuweijiang0922@gmail.com Tel: 0086-431-88941670

economic revenue.

The purpose of this paper is to propose an objective method to determine RFM index weights based on customers' purchasing data. First qualitative analysis of the R, F, M relationship of RFM model; second constrained by the fact the sum of the indexes weight is 1, on the basis of various combinations of the three index weights are tried using different step gradients of 0.1(Greater precise might be obtained by using smaller step gradients, such as o.o1 and so on); third according to rule of the index weights of RFM under the biggest response rate were the enterprise's best RFM index weights, objectively determined RFM index weights. Finally, based on the weights thus determined, the value of each customer to the enterprise is calculated by RFM index weighted algebra.

Enterprises can segment customers according to customer value by clustering and provide a feasible method to do target marketing.

METHODOLOGY

Determination of RFM index weights

Data standardization

In the RFM model, R indicates the day/week/year from the date of the customer's last purchase to statistical date. Statistically, the larger this value (the greater the lapse time) the lower the probability the customer will buy again. F is a frequency that the customer purchases products during the statistical cycle. The higher the frequency means the higher the possibility the customer will make future purchases. M is the value for the total purchases by a customer in the statistical cycle. The higher the M Value the higher the probability the customer will buy again. In the RFM model, R, F, and M have different units and F, M have a positive impact on future purchases while R has a negative impact. In order to facilitate the analysis the RFM data must be standardized (Sohrabi et al., 2007) as follows:

$$R' = \frac{R^L - R}{R^L - R^S} \quad (1)$$

Where R' is the value after standardization, R is the customer's recency value for the selected statistical cycle, R^L, R^S are the maximum and minimum values respectively for the all customers activity during the selected statistical cycle.

For Frequency value F, the standardization is determined by:

$$F' = \frac{F - F^S}{F^L - F^S} \quad (2)$$

Where F' is the value after standardization, F is the customer's buying frequency during the selected statistical cycle F^L, F^S are the maximum and minimum values respectively for all customers during the selected statistical cycle.

For the monetary value M, the standardization is determined by:

$$M' = \frac{M - M^S}{M^L - M^S} \quad (3)$$

Where M' is the value after standardization, M is the customer's purchasing amount during the statistical cycle , M^L, M^S is the maximum and minimum value for the purchase amount of all customers during the selected statistical cycle. For convenience, in the following content, the standardize results will be represented as R, F, and M.

RESULTS

As previously stated a customer's buying history (RFM) is directly related to future purchases. Determining the weights for the RFM in order to identify the most valued customers is important. Therefore, qualitative analysis of R, F and M is required to understand the relationship between each index.

Firstly, this paper introduces the concept of customer response rate. This response rate is determined as the ratio or proportion of the customer number of a previous given subdivision customers appear in future high-end purchasing customers(customers who have spent a lot of money in the future time) and the customer number of the previous given subdivision consumers. This can be described by the following formula:

$$\text{Response rate} = \frac{A \cap B}{A} \quad (4)$$

Where A is the number of a previous given subdivision customers, B is the number of future high-end customers.

Secondly, in order to verify that the enterprise's weights are suitable for different statistical cycles so a series of cycles are analyzed. Given a particular retail enterprise's sales data for January 1, 2007 to June 30, 2009, 24068 people's purchase history is used for this analysis. A Statistical cycle is defined as a six month period, hence, the selected sales data has 5 statistical cycles.

The first cycle is January 1, 2007 to June 30, 2007, the last cycle is January 1, 2009 to June 30, 2009. For each statistical cycle, the customer's buying situation for the first 5 months of cycle is considered the historical or past subdivision and the last month is considered the current or future division of the cycle for the purposes of establishing a response rate. In the sample the customers are divided in seven subdivisions(1-7, 1 means the lowest buying amount level, and 7 means the highest buying amount level) according to their buying amount (M)'s descending and each group has the same consumer number. Each of these seven subdivisions is further divided into seven more groups based on each group consumers' Frequency (F)'s descending order thus generating a total 49 subdivisions. Customers in the future set (the data in the last month of the cycle and it is also called the test set) will be ordered by their buying amount and are divided into 3 customer groups. Consumers in one consumer group with the largest purchase amount are called the most valuable customers or the high-end consumers. Now the 49 subdivision of customers

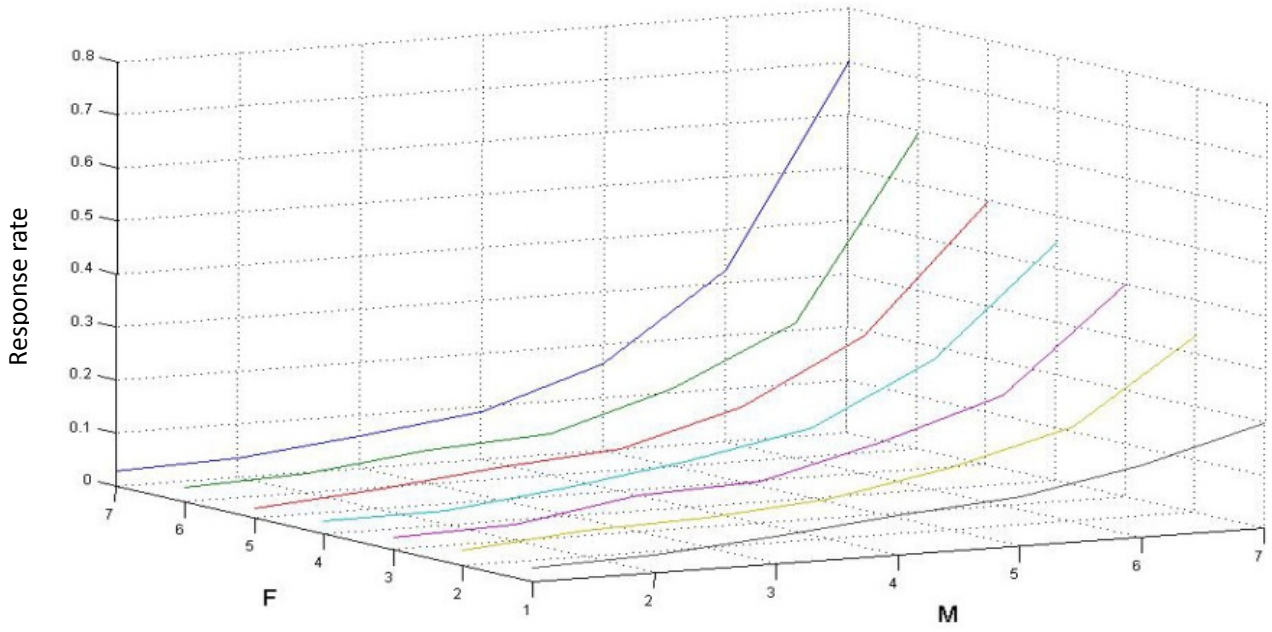

Figure 1. Relationship between response rate and customer's MF level.

is then processed against the test set to obtain a base response rate. The results of these calculations are shown in Figure 1.

In Figure 1, the X axis (M) represents the customers' purchase amount rating and the Y axis (F) represents the customers' frequency level and Z axis represents each segment customer's base response rate. The higher M's level, the greater the customer's purchase amount has been. The greater the F value the more frequently the customer has made purchases. The graph illustrates that frequency and purchase amount obviously have a positive relationship to the customers' future purchase level. When M and F has the highest response rate is the largest, or approximately 0.71. In same way that MF is determined, the MR relationship to the response rate is determined and graphed. The result is depicted in Figure 2.

Figure 2 illustrates the obvious fact that recency (R) and purchase amount (M) have a positive relationship to the customers' future purchase level. When M and R are both maximum, the response rate is the largest, the future purchase probability is the highest. Figure 2 indicates that the best MR's response rate is approximately 0.62.

The results of the analysis of the first statistic cycle indicate the response rate of MF which was found to be 0.71(Figure 1) is higher than the response rate of MR which was found to be 0.62 (Figure 2). That means the role of F is bigger than the role of R. In the same manner using the first period sales data can be calculated for the relationships between RM and response rate, relationships between RF and response rate. In doing so, it was found that the RM maximum response rate was 0.77 compared to the maximum response rate for RF that was 0.69. That means the role of M is bigger than the role of F. Therefore, the relationship between R, F and M is better understood. From the first statistical cycle it was determined that the role of M is the biggest, F second, R the least. This same analysis was applied to the remaining four statistical cycles to obtain the results presented in Tables 1 and 2.

Tables 1 and 2 demonstrate that although the response rates for MF and MR differ between statistical cycles, the general rule that for the priorities of M, F and R remain relatively the same: The role of M is the biggest, F second, R the least. More quantitative analysis is required to determine the proper index weights for the enterprise's RFM model.

RFM weight and customer value determination

Based on the above analysis the method for determining the RFM weights based on a given enterprise's sales data is presented in this section.

Since the sum of the three index weights equals 1, the proposed method sets each index weight range from 0.1 - 0.8 in steps of 0.1. Thus a total of 36 combination is formed, such as R = 0.1, F = 0.2, M = 0.7, but based on the response rate discussion above regarding the relationship of RFM weight, all combinations which do not satisfy the relationship M>F>R are deleted. Four possible combinations are left.

For each of these possible combinations of weights the

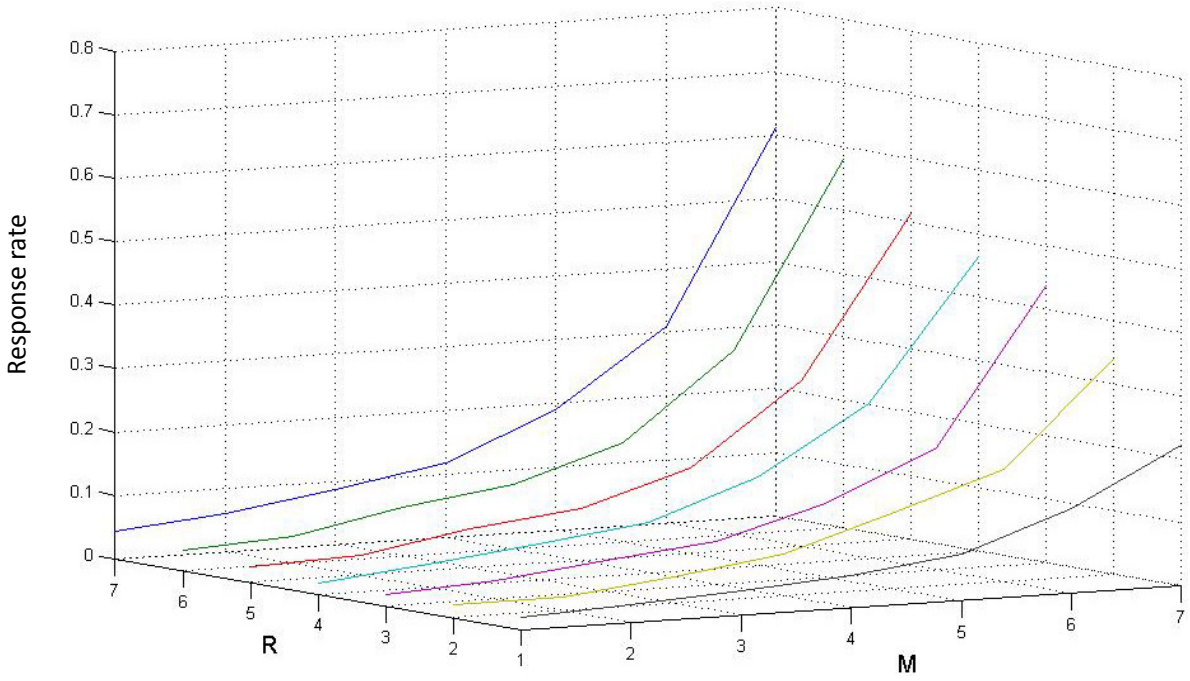

Figure 2. Relationship between response rate and customer's MR level.

Table 1. Maximum MF and MR response rates for all statistical cycles.

Statistical cycles	Maximum MF response rate	Maximum MR response rate
First statistical cycle	0.71	0.62
Second statistical cycle	0.77	0.70
Third statistical cycle	0.94	0.79
Fourth statistical cycle	0.78	0.78
Fifth statistical cycle	0.78	0.72

statistical cycles are again analyzed. For each of the six months of sales data the first five months are consider historical data and the last month is the test data as in the previous analysis. Then using the recency R, frequency F and the monetary M, the weighted value T is determine algebraically as:

$$T = a \times R + b \times F + c \times M \quad (5)$$

Where a, b, and c are corresponding weights for R, F and M, given a, b and c value, customer value in this five months is obtained corresponding to this weight algebra. For enterprises, they can segment customers according to this weight algebra. In this paper, after calculating customer value, these customer values are arranged in from large T value to small T value and divided customers into 5 groups and each group has the same customer number. At the same time, customers who are in last month of the cycle are arranged in M descending order and divided in three groups. After that this paper calculates response rate of the largest customer values group compared to the largest purchase amount of the customer group, then changing weights (one of four possible combinations), re-calculating T value and response rate under this new weights in this cycle. In all four combinations of weights, one combination which produces the maximum response rate is the enterprise's best weight for the RFM during this statistical cycle. In order to make this weight combination adapts to different statistical cycle, this method is then applied to the remaining statistical cycles to verify this conclusion.

The above method is applied against the five cycles defined earlier using the SQL Server2005 program to calculate corresponding response rate under four combinations of weight in every cycle, the result is in Figure 3. In Figure 3 the horizontal axis represents the four weight combinations (for example, r1f2m7 means RFM index weights are a=0.1, b=0.2, c=0.7) and the vertical axis represents the corresponding response rate. A different line (color) is used to represent the different

Table 2. Maximum RF and RM response rates for all statistical cycles.

Statistical cycles	Maximum RF response rate	Maximum RM response rate
First statistical cycle	0.69	0.77
Second statistical cycle	0.71	0.77
Third statistical cycle	0.68	0.82
Fourth statistical cycle	0.69	0.77
Fifth statistical cycle	0.63	0.77

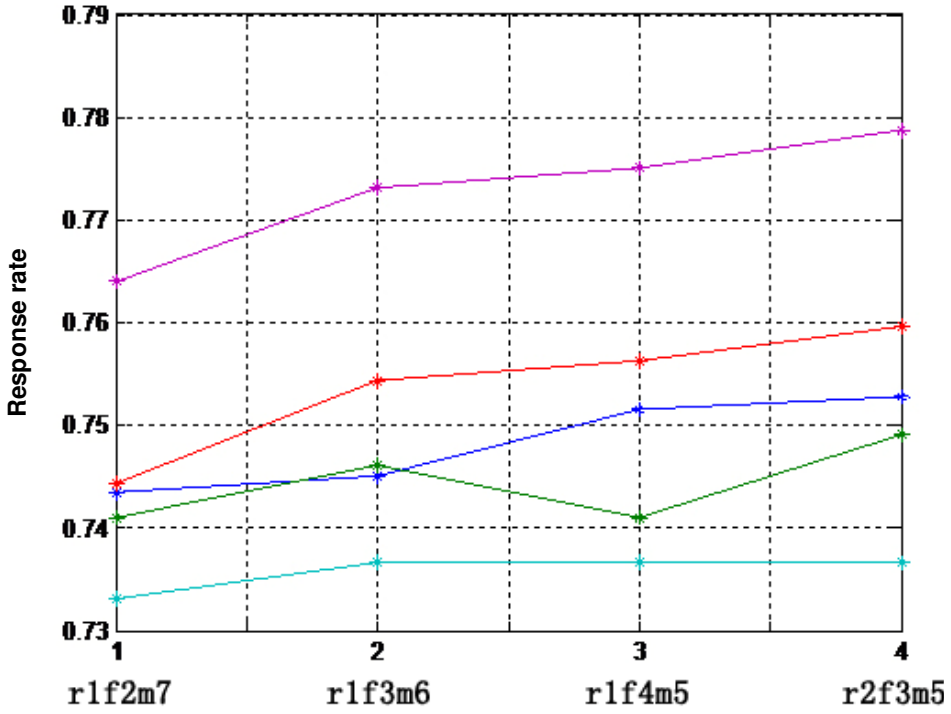

Figure 3. Response rates according to different weights in different statistical cycles.

statistical cycles (5 statistical cycles).

Figure 3 shows that there is some regularity in each batch but the r2f3m5 weights provides the best or optimum results. Hence this can be regarded as the enterprises RFM index weights. To improve accuracy a 0.01 gradient instead of a 0.1 gradient can be used. The numeric value of the optimum RFM may be different in this case but the principal is still valid.

Conclusion

In this application of the RFM an index weight is determined specifically for the enterprise as the basis evaluating customer value and classifying them according to potential sales. A method to find an enterprise's RFM weight from its sales data reduces or eliminates the subjectivity of the tradition analytic hierarchy process (AHP) that is biased by the expert's prejudices. Hence, a more credible result is objectively obtained from the enterprise's sales data and this method is not restricted by enterprise characteristics and can be applied to any other enterprises. Based on the objectively determined weights from the enterprise's sales records the customers can be segmented and the most valuable customer set be identified. Hence the overall enterprise management of their customer segment is achieved.

REFERENCES

Deng B, Shao PJ, Zhao D (2008). Data mining for needy students identify based on improved RFM model: A case study of university. Int. Conf. Inform. Manage. Innov. Manage. Ind. Eng., pp. 244-247.

Hughes AM (1994).Strategic database marketing. Chicago: Probus Publishing.

Liu DR, Shih YY (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value. Info. Manage.

42: 387-400.
Pareto V(1906). Manuale di economia politica. Piccola biblioteca scientific, Milan. Translated into English by Schwier AS (1971). Manual of political economy. MacMillan, London.
Satty TL (1980). The analytical hierarchy process: planning, priority setting, resource allocation. McGraw-Hill, New York.
Schmittlein DC, Cooper LG, Morrison DG (1993). Truth in concentration in the Land of 20/80 Laws. Mark. Sci., 12: 167-183.
Shen CC, Chuang HM (2009). A study on the applications of data mining techniques to enhance customer lifetime value. WSEAS Trans. Inform. Sci. Appl., 6: 319-328.
Sohrabi B, Khanlari A (2007).Customer lifetime value measurement based on RFM model. Iran. J. Account. Audit., Rev., 14: 7- 20.
Stone, Bob (1994). Successful direct marketing methods. 5th ed, NTC Business Books.