

Study of Customer Segmentation for Auto Services Companies Based on RFM Model

Wang Hu Zhang Jing

School of Management, Wuhan University of Technology, Wuhan, P.R.China, 430070

(E-mail: wanghu61@126.com, zj123456789zj@yahoo.com.cn)

Abstract This paper aims to explore the applicability of the RFM (Recentness ,Frequency ,Monetary) model in the customer segmentation of auto services companies, for which it obtains the weight of each index through the method of analytic hierarchy process (AHP) and segments the customers with K-means clustering method. This paper divides customers into several segments by comparing customer lifetime value calculated with weights of indexes in the model. It is proved by a case study that the model and methods proposed in the paper can be applied to effectively solve the problem in customers segmentation of auto services companies.

Key words customer lifetime value, customer segmentation, RFM model, K-means clustering method

1 Introduction

Customer Lifetime Value (CLV) is one of the most important indexes in customer segmentation. When it was used to divide customers into groups, companies can decide how to make full use of their limited resources to customer service effectively by segmenting their customers into different groups based on CLV, and then design and implement different customer retention strategies to realize the maximization of their total profits^{[1][2]}. Nowadays, The methods of how to select the indexes of models and calculate yardstick needed is limited in the research field of methods of customer segmentation based on CLV^{[3][4]}. Some academicians introduced RFM (Recentness , Frequency , Monetary) model which is used frequently in the field of database marketing to be used to evaluate the value of customers indirectly instead of the method of evaluate directly. They also pointed out that the importance of each index in RFM model is different when using it in different industries. Aimed at customers of auto services companies, the paper tries to use clustering method based on RFM model and calculates the value of CLV of every cluster of customers to segment and analyze their consumption behavior characteristics. At the end of the paper, service records of year 2007 of an auto services company are used to carry through empirical study.

2 RFM Model for Auto after Sales Service Companies and Weights of Each Index

2.1 RFM Model for Auto After Sales Service Companies

RFM model is a smart and useful model used frequently to customer or market classification. It is composed of three indexes namely Recentness、 Frequency and Monetary. The three indexes can be used to sales forecasting^[5]. RFM model is also appropriate for customer segmentation of auto services companies and consumption behaviors of customers are considered to be the basis of the model in the paper. The comparison of the meaning of each index in RFM model in general and in auto industry is showed in Table 1.

2.2 The Weight Analysis of Each Index in RFM Model

The method of AHP (analytic hierarchy process) is applied to calculate the weight of each index in RFM model. The process of calculating is described below according to the calculating process of AHP^[6].

First, compare indexes between every two indexes. Decision-making group are asked to use the 1-9 comparable scale method to compare importance of every two indexes and draw a group of judgment matrixes as the format shown in Table 2;

Table 1 Comparison of the Meaning of Each Index in RFM Model in General Situation and in Auto Industry

Model	R (Recentness)	F (Frequency)	M (Monetary)
General RFM model	The number of days between the date of customers' last purchase in the company and the end of measurement period	The time of customers purchasing the commodities in the company among measurement period	The total amount of money for purchasing the commodities in the company among measurement period
RFM model in auto industry	The number of days between the date of customers' last purchase for service in 4s stores and the end of measurement period	The time of customers purchasing for service in 4s store among measurement period	The total amount of money for purchasing service in 4s store among measurement period

Table2 The Format of Judgment Matrix Drawn by Members of Decision-making Group

	R	F	M
R	1	a_{12}	a_{13}
F	$a_{21}(1/a_{12})$	1	a_{23}
M	$a_{31}(1/a_{13})$	$a_{32}(1/a_{23})$	1

The selection of professors for decision-making group is critical for the precise of calculated result and they must be selected scientifically and discreetly^[7]. After getting every judgment matrix of members, the method of calculating average values is used to get the final judgment matrix A. Secondly, have consistency check for judgment matrix A. C.R. (consistency ratio) is used to measure the degree of inconsistency, and the steps are as follows:

Step 1, calculate C.I. (consistency index), the formula is as follows:

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \dots\dots\dots (1)$$

In this formula, λ_{\max} presents maximum eigenvalue and n is matrix order of A.

Step 2, find appropriate value of R.I. (random index) in random index table;

Step 3, calculate C.R. (consistency ratio) and have consistency judgment with following formula:

$$C.R. = \frac{C.I.}{R.I.} \dots\dots\dots (2)$$

When C.R.is less than 0.1, the consistency of judgment matrix A is acceptable and it can be used to calculate weights of R, F and M. Otherwise, the requirement of consistency can not be satisfied in A and should be obtained and calculated again until it meets the need.

The last step, calculate weight of each index which are represented as w_R 、 w_F and w_M according to judgment matrix A.

3 Customer Segmentation Based on K-means Clustering Method and the Value of CLV

K-means clustering method^[8], with weighted R、 F and M as indexes, is used to segment customers according to the characteristics of R、 F and M, and then calculate the value of CLV of each cluster to segment customers for the second time. General steps are listed as follows:

- 1) Calculate weight of each index w_R 、 w_F 、 w_M with method of AHP;
- 2) Standardize value of each index and weigh them with their own weights calculated in step 1):

Because of the difference of units for each index, the data should be standardized to one uniform measuring unit so as to avoid the dependence on units. For index R, because of its negative influence to CLV, the paper standardizes R with formula (3) as follows:

$$R' = \frac{R^L - R}{R^L - R^S} \dots\dots\dots (3)$$

Because of positive influences to CLV, the other two indexes F and M are standardized in formula (4):

$$F' = \frac{F - F^S}{F^L - F^S}, M' = \frac{M - M^S}{M^L - M^S} \dots\dots\dots(4)$$

Then weigh the three indexes as follows:

$$R'' = w_R \times R', F'' = w_F \times F', M'' = w_M \times M' \dots\dots\dots(5)$$

Where R'', F'' and M'' respectively represent weighed R, F and M indexes. The values of R'', F'' and M'' are used for clustering analysis in the paper.

3) Decide the appropriate number of clusters of customers *m*;

4) Using K-means clustering method to cluster records of weighed R, F and M indexes to get *m* kinds of customers. First, use clustering method to clean unusual records. After clustering, if one record is not included in any one cluster, it is considered as an unusual record and can be thrown off. Then clustering cleaned records into *m* clusters with SPSS;

5) Comparing average value of R'', F'', M'' of every cluster with overall average value of R'', F'', M''. There are two results for each group of comparing: more (or equal) than or less than. Recording the comparing results of them;

6) Analyzing the characteristics of each kind of customers according to comparing results. For example, the customer is loyal to or departs from the company till now etc. And then define customer styles according to the characteristics.

7) After analyzing the characters and defining customer style, a quantized value is still needed to have a more explicit comparison between each cluster of customers^[4]. That quantized value is CLV. The calculate formula for value of CLV is as follows:

$$CLV_j = M_{R''} + M_{F''} + M_{M''} \dots\dots\dots(6)$$

where $M_{R''}$, $M_{F''}$ and $M_{M''}$ respectively represent average value of R'', F'', M'' of every cluster and *j* represents the code of each cluster.

4 Empirical Study Based on the Service Records of an Auto Service Company

4.1 Customer Segmentation Based on RFM Model and K-means Clustering Method

Service records of year 2007 of an auto service company are used in empirical study here, totally 25645 pieces. According to the meaning of indexes in RFM model, 5821 pieces of records are obtained to be used later. The sample of obtained records is shown in Table 4.

Weight of each index is calculated by means of AHP. In the paper decision-making group include one auto after sales service manager, two desk clerks of after sales operation and two customers with high loyalty. After drawing out each judgment matrix of them, final judgment matrix A is obtained as Table 3 shows. Then consistency check is done for the matrix. The result is that the degree of consistency can be accepted.

Table 3 Final Judgment Matrix A

	R	F	M
R	1	0.71	0.46
F	1.41	1	0.85
M	2.18	1.18	1

Weights of the three indexes are calculated according to A. The results are as follows:

$$w_R = 0.221, w_F = 0.341, w_M = 0.438$$

Through the result it was found that M is the most influential factor to the value of CLV. Then weighed value of each index R'', F'' and M'' are calculated according to formula (5). After those steps, final standardized data are obtained and showed in Table 4. Thereinto, three figures after decimal point of the value of R'', F'' and M'' are retained in the table.

Table 4 Original Service Records and Final Data Used in Clustering Analysis

No. of car owners	R(day)	F(time)	M(Yuan)	R''	f''	M''
AOW04090467	14	20	18060	0.192	0.000	0.047
AOH04002829	21	20	8780	0.174	0.000	0.000
AOW04100487	5	21	25170	0.216	0.009	0.083
AOW05030162	88	21	21506	0.000	0.009	0.065
AOH04000379	20	21	18322	0.177	0.009	0.049
AOH04002033	8	21	14742	0.208	0.009	0.030
AOW06090030	27	22	19764	0.159	0.018	0.056
...

Average value of each index for total standardized data group is showed in Table 5:

Table 5 Average Value of Each Index for Total Standardized Data Group

<i>Index</i>	<i>R''</i>	<i>F''</i>	<i>M''</i>
Average value	0.185	0.100	0.106

The characteristics of each cluster will be obtained through the comparison of the three indexes between every cluster and total data group. The result of comparison of every unique index includes two results: more than (or equal) and less than, so it is reasonable to cluster the records into $8(2 \times 2 \times 2)$ clusters. After cleaning the records, the last 5486 pieces are left for clustering analysis. The 5486 pieces of records are clustered into 8 clusters with K-means clustering method using SPSS. Then compare average value of R'' , F'' and M'' of each cluster with them of total standardized data group showed in Table 5. For index F and M the sign \uparrow is used to represent that the former value is more than or equal to the latter one and \downarrow to represent the opposite meaning. For R , because of its negative influence to the value of R'' , contrary signs are used to label the comparison result of R . According to the definition below, a theoretical clustering result is obtained as shown in Table 6, and the diagram obtained through SPSS is as Figure 1.

The sign $R\downarrow F\uparrow M\uparrow$ represents that the value of R of this cluster of customers is less than general average value. Besides, the time of coming and purchasing services in the company and the sum of money paid for the service of this kind of customers are both more than average value. It can be concluded that this kind of customers have come and purchased services in the company recently, they frequently come to accept service and the money they spent is higher than average in measurement period. Through the analysis, this kind of customers is considered as the company's most important and valuable customers now. The company should think highly of those customers and try to maintain good relationship with them and keep them as the company's most important and valued customers in the future. So this kind of customers is defined as maintaining customers. Analyzing other 3 kinds of customers with the same theory, the paper defines them as developing customers whom their value should be increased in the future; non-value customers that won't brought value to the company now or in the future; detaining customers that that the company consider them as valued customers now but have the trend of decreasing of their value for the company, and the company should think about the reason for their leaving, and try to maintain them and develop the relation with them to make them become more valuable customers in the future.

Table 6 Eight Clusters of Customers After Clustering Analysis

Cluster	1	2	3	4	5	6	7	8
$M_{R''}$	0.192	0.188	0.209	0.000	0.156	0.221	0.196	0.198
$M_{F''}$	0.206	0.020	0.045	0.009	0.106	0.249	0.272	0.341
$M_{M''}$	0.127	0.027	0.079	0.065	0.152	0.438	0.243	0.173
Amount	23	1977	156	2	2375	104	9	840
Character	$R\downarrow F\uparrow M\uparrow$	$R\downarrow F\downarrow M\downarrow$	$R\downarrow F\downarrow M\downarrow$	$R\uparrow F\downarrow M\downarrow$	$R\uparrow F\uparrow M\uparrow$	$R\downarrow F\uparrow M\uparrow$	$R\downarrow F\uparrow M\uparrow$	$R\downarrow F\uparrow M\uparrow$
Definition of clusters	maintaining customers	developing customers	developing customers	non-value customers	detaining customers	maintaining customers	maintaining customers	maintaining customers
Value of CLV	0.524	0.234	0.333	0.074	0.414	0.908	0.711	0.712
Order of CLV	4	7	6	8	5	1	3	2

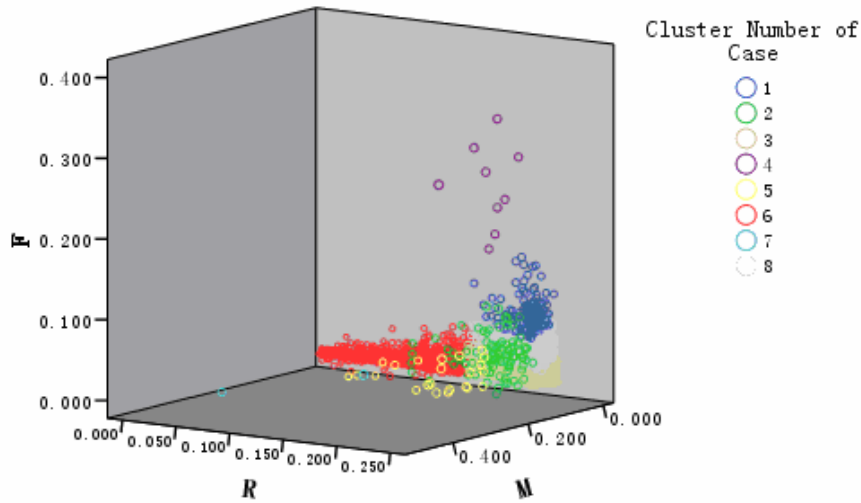


Figure 1 The Diagram for the Result of Clustering

4.2 CLV Value Analysis

After the classification below, customer's value and importance of each cluster to the company are not known yet. Besides, as shown in Table 7, cluster of No. 1、6、7 and 8 are all defined as maintaining customers, that is to say, all of them are the company's most valuable customers now. At this time, farther classification is necessary to the four clusters. The company can be helped to analyze the gap between different customers quantitatively and enact more feasible customer policies by ordering the customers for the second time depending on the value of CLV. So the value of CLV is calculated according to formula (6) and the clusters of customers are reordered as showed in Table 7. It indicates that the value of CLV of cluster 6 is the highest, cluster 7 and 8 take the second place and cluster 4 is the lowest one. Through comprehensive consideration of two steps of classification, it was found that customers belonging to cluster 6 is the company's most valuable ones and their needs should be considered first considering limited resource of the company. We can also find that both cluster 2 and 3 are developing customers, but value of CLV of cluster 2 is more than cluster 3, so the former should be developed preferentially.

5 Conclusions

Customer segmentation is one of the fundamental principles for companies to successfully formulate customer retention strategies. This paper shows the process of customer segmentation (customers of auto service companies) with RFM model of both short-term method and CLV of long-term predict method, analyze basic characteristics and consult different customer policies of each cluster in order to try to obtain maximum profit. Due to space limitation, the studies about the ontological characteristics of customers belonging to different cluster are not included in the paper and may be studied in the future.

References

- [1] Lin Yan, Cai Qiming. Study of Customer Resources based on CLV [J]. Forum of Statistics and Information, 2004, (11):31-33 (In Chinese)
- [2] Chen Mingliang, Li Huaizu. Study on Value Segmentation and Retention Strategies of Customer [J]. Group Technology & Production Modernization, 2001, 18(4):23-27 (In Chinese)
- [3] Etzion O., Fisher A.. CLV: a modeling approach for customer Life-time evaluation in e-commerce domains, with an application and case study for online auctions[C]. Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service(IEEE 04), 2004: 1-8
- [4] Jiang Guorui, Liu Pei, Huang Tiyun. Research of customer value segmentation based on AHP methods[J]. Computer Engineering and Applications, 2007, 43(8):238-241 (In Chinese)
- [5] Stone B., Jacobs R.. Successful Direct Marketing Methods[M]. McGraw Hill Publishing, 2001: 158-160
- [6] Tomas Rytel. Critical Insights of Customer Valuation Methods in the Relationship Marketing

- Paradigm[J]. Ekonomike, 2006(74):52-63
- [7] Sargeant A.. Customer Lifetime Value and Marketing Strategy: How to Forge the Link [J]. The Marketing Review, 2001, Vol. 1:427-440
- [8] Gummesson, E.. Return on Relationships (ROR): the Value of Relationship Marketing and CRM in Business-to-Business context [J]. Journal of Business and Industrial Marketing, 2004, Vol.19:136-148