

Research Intelligent Precision Marketing of E-commerce Based on the Big Data

Jianhui Zhang¹ & Junxuan Zhu¹

¹ School of Management, Shanghai University of Engineering Science, Shanghai, China

Correspondence: Jianhui Zhang, School of Management, Shanghai University of Engineering Science, Shanghai 201600, China. E-mail: 498162197@qq.com

Received: December 4, 2013

Accepted: December 25, 2013

Online Published: February 15, 2014

doi:10.5430/jms.v5n1p33

URL: <http://dx.doi.org/10.5430/jms.v5n1p33>

Project supported by the Graduate Innovation of Shanghai University of Engineering Science (No.A-0903-13-01043).

Abstract

This paper analyzed and summarized the development path of electronic commerce marketing based on the big data; the related aspects of intelligent precision marketing framework has been designed combined with smart technology; and describes its functional structure and operational processes. Taking into account the differences between e-commerce and traditional retail industry; constructed RFMA model combined with characterizes of the electricity suppliers, by means of k-means clustering to achieve the client's "precision" division. Finally, verified the model of marketing by a c2c transaction data, clarify this model could develop precise marketing strategies to deal with the challenges posed by big data has some significance.

Keywords: big data, electronic commerce, precision marketing

1. Introduction

McKinsey assessed the western industrial data showed that make effective use of big data could improve margins for most retailers, even more than 60% (Bughin, J., Chui, M., & Manyika, J., 2010). The real value of big data is the huge potential commercial value. So, how to use these data to make management decisions are becoming the focus of business. As to electricity suppliers, how to integrate and accurately analyze these consumer behavior's data and the results are used to make marketing decisions to improve marketing efficiency, has become an important strategic objective of its development. However, the current application of intelligence technology mostly concentrates in the customer relationship management, market segmentation and dynamic pricing, etc., although some studies use smart technology in marketing, but most concentrates in the power of marketing applications, however, most studies are general theoretical, lack of a complete application analysis and empirical research to make full use of intelligence technology in marketing decision. As Xu weiting et al. (Ping, Shenzhen, & Yang, Xie, 2013), established a "smart marketing" theoretical system based on the advanced intelligent network, and expounded changes of intelligent marketing concepts and models. Chen xin (Hughes, MA., 1996), et al, proposed a cross-platform intelligent marketing model based on the open-end funds (F * OPIM), broaden the fund's operations span the lateral and vertical fields, trying to combine with the consultation delivery, fund marketing and after-sales.

Therefore, construct a framework of intelligent precision marketing model considering the characteristics of electricity suppliers, and combining intelligence technologies and related clustering algorithm. Empirical studies have shown that the customer segmentation based on RFMA is more accurate, and help companies to identify potential customers in the customer base, in order to minimize the corresponding errors in marketing decisions.

2. The Development Path of E-commerce Marketing in the Big Data Condition

With the development of e-commerce, the growth of internet subscriber and the diversification of network terminals make the data volume, complexity and generate speed far beyond the traditional data forms. As data scale continues to expand and consumer's needs continue to improve, intelligence has become the value carrier for development of e-commerce marketing in the future. Many studies show that the application of intelligence technology could process large amounts of data, and these data can be analyzed to obtain relevant knowledge and support decisions. In the past

ten years, the application of data mining is extensively expanded and achieved remarkable results in marketing, sales, human resources, e-commerce and other business areas (Yan, Pingzhi, Xun, Guo & Zeng, China, 2013). Therefore, it is foreseeable to e-commerce marketing path, and it will towards fast, accurate and stable path to evolve. Fast, that is how to accurately analyze data by use of smart technology, and adjust marketing strategies fast and dynamically. In the era of the internet and big data, customer behaviors are fast changeable and dynamic, companies need to quickly adapt to such changes and make timely adjustment of marketing strategies. The value of big data is accurately records the track of consumer's information, access to consumer's real behaviors, attitudes and responses to information. Accuracy, that is able to accurately define the consumer groups, the information contact point, and to know exactly marketing actions. Stability refers to the shorter information cycle caused by changes of dynamic behavior of customers. It requires companies steady use the data to cope with mistakes caused by information not considered comprehensively.

3. The Framework of Intelligent Precision Marketing Based on the RFMA

The main process of intelligent precision marketing is analysis the behavior data of consumer by use of modern intelligence technology, and establish a database of customer segmentation, then make marketing decisions to different type of customers; finally achieved effective integration both objective data and active feedback through customer's feedback, and continuously improve the level and efficiency of precision marketing.

Analyze the data of customer's behavior combined with the improved RFM model and K-means clustering algorithm, and shape different type of customers. Finally, make different strategies for each type of customers. The framework is divided into three layers from top to bottom: the data layer, analysis layer and decision-making layer (Figure 1). Three layers are closely linked and helpful to achieve precision marketing to e-commerce suppliers.

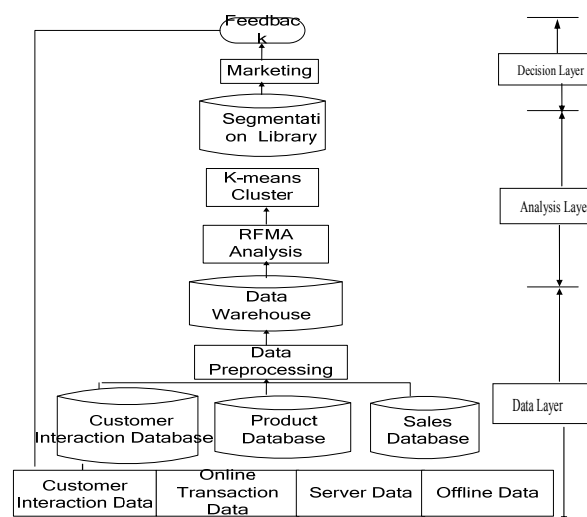


Figure 1. Framework of intelligent precision marketing based RFMA model

3.1 Data Layer

Data layer is responsible for receiving real-time data, combined with basic customer information, business information and consumer attributes and other data which is obtained from various sources, and preprocess these data, load the data warehouse for the next phase of data mining and intelligent analysis to prepare.

3.1.1 Determining Data Source

The data sources of this model include server data, online transaction data, customer interaction information, offline data and other relevant data, etc. (Wan, Hong, & Wan, Yan, 2010). There are two categories of server data, web logs and proxy server-side data; online transaction data are mainly e-commerce site information, user information, purchasing information and other data stored in the traditional relational database. Such as commodity prices, sales and the characteristics of goods, etc. Customer interaction information includes customer registration and feedback information. Customers enter and submit relevant information on the screen via the web, and then send to the server,

that is customer registration information. Such as customer's personal data, product information ordered by customer, and some of the issues and requirements raised by customers, etc. Feedback information mainly refers to the data which is searched and queried online and the evaluation information purchased. Offline data refers to the inventory, cost, and logistics information in e-commerce enterprises. (Wang, Feifei, & Li, Jing, 2012)

3.1.2 Data Preprocessing

E-commerce enterprise's data is large size and high dimensional (Ying Ho, Yuho Chung, & Kinnam Lau, 2010, June) based on the big data. The raw data collected from each of the data sources probably exist missing values, repeat, incomplete, etc, which requires data preprocessing. Convert the data to meet the application requirements by cleaning and switching (Ping, Shenzhen, & Yang, Xie, 2013), thereby reduce the dimensionality of the data and improve the data's quality. The main task in data preprocessing is about to organize the user's recording information, and form the data set for the next step, and load the data warehouse.

3.2 Analysis Layer

The main task of analysis layer is analysis, processing, and access to useful knowledge by using intelligence technologies and the related algorithms for supporting marketing decisions. This layer is the core of the entire marketing framework, analyzing a lot of information about consumer behavior based on the RFMA model and K-means clustering, determine the type of customers and form multiple market segments, and then provide the marketing decision for the next stage.

3.2.1 The Background of RFM Model

RFM model is not involved in personal privacy, and it's easy to access purchase data. It's applicable to the traditional retail industries which provide a variety of commodities, low price commodities relatively and commodities purchased repeatedly. The basic idea is about to determine customer value as the following three important behavioral indicators (Hughes, MA., 1996): R represents client's time span from the last purchase to now, the lower the R, the higher the customer value (Wu, C., & Chen, H.L., 2000); F stands for customer's purchase frequency, which means that the number of transactions in a period of time, the higher the F, the higher the customer value (Lemon, K.N., White, T.B., & Winer, R.S., 2002); M represents the amount of consumption in a period of time, the higher the M, the higher the customer value.

3.2.2 The Improved Basis of RFM Model

Most domestic scholars get a more accurate customer segments in e-commerce mainly by improving the algorithm, there is few realistic description to the model, and marketing strategies are still at the surface of the stage (Rule, Attaining, Wang, Jiaqiang, Tu, Huan, & etc., 2012). Foreign scholars believe that the model of customer segmentation directly impact on the accuracy of data mining technology (Maia, M., & Almeida, J., 2008). The high amount of consumption is not necessary to high customer value in a period of time. Although the high consumption could increase business sales, it's not enough to explain customer's satisfaction and lasting consumption attitude. That is not enough to explain customer's real value. For example, as to some consumers having the fewer amounts of consumption and the lower purchase frequency, but the evaluation is higher relatively, they are probably potential customers that companies need to focus on. So, using RFM model's clustering may conclude wrong results.

Combined with the real situation of e-commerce, e-commerce enterprises have complete records customer's behavior; it's easy to get those data in this situation. At the same time, there is a clear difference between e-commerce and traditional retail, the system of e-commerce will timely produce evaluations in different levels. On the contrary, it's rare to get quantitative evaluations timely in traditional retail after shopping. Therefore, the new model considers evaluation as the basis of segmentation. Because of customer evaluation is an important indicator.

3.2.3 Clustering

Generally, clustering is used to divide market and customer groups and other studies. It can cluster the similar individuals into the same category; individuals with large differences were clustered into different classes, and ultimately achieve intra-class individuals with a more high homogeneity (Liu, Heroic, Bong, 2006). Using clustering on sales data and customer purchase information for analysis is helpful to identify the different type of customers, in order to develop appropriate marketing strategies. In this paper, K-means clustering is used to divide different type of customers, guaranteed for the implementation of next stage of precision marketing decisions.

3.3 Decision-Making Layer

According to the RFMA model, customers were divided into three classes: loyal customers, potential customers and the upcoming loss of customers. Make decisions and develop appropriate marketing strategies to each type of

customer. After the implementation of marketing strategy to each type of customer, actively interact with customers by deeply understanding the customer's experience and the recommendations of the product or service. Feedback data circulate and constantly update in the marketing framework, and promote product or service's innovation or improvement. The formulation and implementation of marketing's strategy has always based on the customer's experience, which is consistent with 4C (Consumer, Cost, Convenience Communication) theory which customer oriented philosophy.

4. Model Applications

4.1 Data Description

Extract randomly sample of 50 customers in purchase record data from a C2C transaction data in 2012. Extracting record fields for the customer ID, transaction time, amount of sales and customer evaluation score. Do further preprocess to those data, remove the invalid data, such as transaction amount is negative, valid records are 732. Counting F, R, M and A. F represents each customer purchase within a year), R represents the time span from the last purchase to now), M represents the purchase quantity and A presents the score of evaluation. Get a 50 * 4 matrix, one part of the results are as follows:

$$X = \begin{pmatrix} 202 & 17 & 4028 & 4.6 \\ 81 & 12 & 1894 & 4.7 \\ 294 & 16 & 3874 & 4.6 \\ 70 & 17 & 3990 & 4.5 \\ 46 & 8 & 2368 & 4.8 \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

4.2 Data Standardization

In order to avoid greater errors caused by clustering results of different units of measurement. So we need to normalize for each type of data. This paper used poor standardization, to obtain a normalized value of each index ZR, ZF, ZM, ZA, each value is normalized to the range [0,1], elimination of the influence of dimension. The formula as below (1): i represents the ith customer, j corresponding to each customer that R, F, M, A.

$$X'_{ij} = \frac{X_{ij} - \min \{X_{ij}\}}{\max \{X_{ij}\} - \min \{X_{ij}\}} \quad (1)$$

4.3 Clustering Results of RFMA Model

Taking the scale and the number of customers into account, take the value of K is 3, avoiding the results of classification are too detailed, the understanding deviation are too large to each groups. Clustering results are shown in Table 1 as below:

Table 1. Final cluster centers of RFMA model

Clustering Properties	1	2	3
ZR	0.2242	0.0439	0.4856
ZF	0.2415	0.5594	0.5130
ZM	0.1546	0.5148	0.3788
ZA	0.7000	0.7230	0.1330

Table 2. The number of customers per cluster in RFMA model

Cluster	1	16.000
	2	13.000
	3	21.000
Valid	50.000	
Missing	0.000	

Looking at the results from the RFMA model, the number of customers in three types are 16, 13, 21. Among them, the second class of customers have the highest evaluation, the highest amount of consumption, the highest frequency of purchase and more long time from last purchase time to now, are the company's loyal customers; the first class of customers have the less amount of consumption and the lower frequency, the shorter last purchase time from now but the evaluation is higher, are the potential customers in enterprises, take appropriate marketing to stimulate their desire to buy as for this type of customer, they can be transformed into loyal customers; the third class of customers have the higher amount of consumption, the longest last purchase time from now, but the lowest evaluated, are the upcoming loss of the customer. Loyalty customer proportion (approximately 26%) and brought to the enterprise in line with the proportion of sales profits Pareto principle from the table.

4.4 RFM Model Clustering Results

Also taking the sample data $K = 3$ based on the RFM model, clustering results as shown in Tables 3 and 4:

Table 3. Final cluster centers of RFM model

Clustering Properties	1	2	3
ZR	0.1150	0.1627	0.6718
ZF	0.2775	0.6203	0.4351
ZM	0.2202	0.5190	0.2940

Table 4. The number of customers per cluster in RFM model

	1	19.000
	2	14.000
Cluster	3	17.000
Valid		50.000
Missing		0.000

In the clustering results of RFM model; the numbers of customers in three types are 19, 14, and 17. Among them, the second class of customers have the highest amount of consumption, the highest frequency, the shortest last purchase time from now on, are the company's loyal customers; the first class of customers have the less amount of consumption, the lower frequency, the shorter last purchase time from now on, are the temporary customers; the third class of customers have the higher amount of consumption, the higher frequency, the longer last purchase time from now, is about the upcoming loss of customers.

5. Conclusions and Appropriate Marketing Decisions

Compared with the two clustering results, the number of potential customers and loyal customer is reduced after the model was improved, the number of the loss of customers are increased. This is because do not consider the customer evaluation, the less the amount of consumption, lower frequency, the shorter last purchase time from now on, customers are likely to be judged as temporary clients; but considering the evaluation, having high evaluation in this category will be classified as potential customers. Similarly, if we do not take evaluation into account, the amount of consumption is higher, the higher the frequency, the longer last purchase time from now on, customers are probably judged as potential customers; But after considering the evaluation of customer, these people are the upcoming loss of customers. Therefore, the classification of e-commerce is more "accurate" after adding customer evaluation, it can help companies found theses potential customers who are spending less money, purchasing lower frequency but giving the high evaluation, these customers are probably the "petty bourgeoisie" class or student, this group pursuit "affordable" in reality, since their income is limited, so the amount of consumption and the purchase frequency are relatively low, which is also more consistent with the reality of the situation. But they have a higher degree of satisfaction to company's products or services, e-commerce suppliers, as to such groups, can reduce commodity prices to increase its price in their mind, so as to stimulate and maintain their desire to purchase, they eventually will develop into loyalty customers.

Similarly, clustering analysis based on RFMA model can also help companies to identify the customer groups which is the higher amount of consumption, the higher frequency within a certain time, the longer last purchase time from now, the lower evaluation, it is not necessary to continue to uphold the relationship with these customer. This group is pursuit fashion, brand, and occasionally shopping online and have high-income, which is also consistent with the real situation. We can completely abandon customers who evaluate too low to avoid waste of corporate marketing resources. This marketing decision like BCG matrix for the "skinny dog" type of market strategy adopted.

6. Epilogues

Big Data era has arrived; the research in precision marketing has just begun. As the accumulation of large amounts of data, intelligence marketing will be more and more attention. Intelligent Precision Marketing emphasized to quickly grasp the market information by using intelligence technology, profoundly analysis customer's behavior; and use the Internet to communicate directly with customers, avoiding the intermediate links, shorten the marketing channels to promote marketing costs reduced, so the customer's purchase cost will be lower; while intelligent precision marketing can also achieve two-way interactive communication with customers. All these features indicate that intelligent precision marketing will promote e-commerce marketing transfer from the traditional 4P marketing concept gradually to 4C marketing. However, studies show that the model of intelligent precision marketing based RFMA can make up the lack of qualitative analysis in traditional precision marketing, to help companies identify potential customers, and has a certain practical significance to achieve customers' precise positioning.

References

- Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly*, (8), 1-14.
- Hughes, MA (1996). Boosting response with RFM. *American Demographics*, (5), 4-9.
- Lemon, K.N., White, T.B., & Winer, R.S. (2002). Dynamic customer relationship management: incorporating future considerations into the service retention decision. *Journal of Marketing*, 66(1), 1-14. <http://dx.doi.org/10.1509/jmkg.66.1.1.18447>
- Liu, Heroic, Bong. (2006). Customer Segmentation Research. *Management Engineering*, (01), 53-57.
- Maia, M., & Almeida, J. (2008). Almeida, V. Identifying user behavior in online social networks // Proceedings of the 1st Workshop on Social Network Systems. New York: ACM, 1-6.
- Ping, Shenzhen, & Yang, Xie. (2013). Qun-based Enterprise Information Factory Business Intelligence Data Management Research. *Information Science*, 3, 102.
- Rule, Attaining, Wang, Jiaqiang, Tu, Huan, & etc. (2012). RFM model based on improved e-commerce customer segmentation. *Journal of Computer Applications*, (05), 1439-1442.
- Wan, Hong, & Wan, Yan. (2010). Commercial banking products designed smart marketing model. *Chinese Scientific & Technical Information*, (23), 264.
- Wang, Feifei, & Li, Jing. (2012). Based on data mining of e-commerce dynamic pricing model. *Chinese information industry*, (2), 56-59.
- Wu, C., & Chen, H.L. (2000). Counting your customers: counting customer' instore decisions, Interpurchase time and repurchasing behavior. *European Journal of Operational Research*, 127(1), 109-119. [http://dx.doi.org/10.1016/S0377-2217\(99\)00326-4](http://dx.doi.org/10.1016/S0377-2217(99)00326-4)
- Yan, Pingzhi, Xun, Guo, & Zeng, China. (2013). Army and other large data context several topics at the forefront of Business Management. *Management Science*, 16(10301), 1-9.
- Ying Ho, Yuho Chung, & Kinnam Lau. (2010, June). Unfolding large-scale marketing data. *International Journal of Research in Marketing*, 27(2), 119-132. <http://dx.doi.org/10.1016/j.ijresmar.2009.12.009>