Fifth International Conference on Fuzzy Systems and Knowledge Discovery

A Data Mining Approach to Customer Segment Based on Customer Value

Yun Chen, Chuan Fu, Hanhong Zhu
(Shanghai University of Finance and Economics, Shanghai, P.R.China, 200433)
chenyun@mail.shufe.edu.cn

Abstract

Customer segmentation is the basic issue for an effective CRM, and the methods and indicators of segmentation will impact on the results of applications. Many literatures have researched the application of data mining technology in customer segmentation, and achieved sound effectives. But most of them only segment customer by single data mining technology from a special view, rather than from systematical framework. Although previous segment methods may segment customer base into different groups,, it is unable to identify customer's capability to create profit directly. In this paper, a customer segmentation framework based on data mining is put forward, and a new customer segmentation method is constituted based on Customer Lifetime Value. The method introduced above has been applied to a dataset from a Frequent Flyer Program (FFP) of an airline in China.

1. Introduction

According to Newell (2000) CRM is a useful tool in terms of identifying the right customer groups and helping to decide which customers to jettison. Clemons (2000) estimates there may be a tenfold difference between the most profitable customers and average. Researches have demonstrated that the implementation of CRM generates better company performances when managers focus on maximizing the value of the customer (Gupta, Sunil, Donald R. Lehmann, and Jennifer A. Stuart, 2004). The more understanding of the customers, the more value of them is acquired.

Generally, customer segmentation methods mostly include experience description method, traditional statistical methods, and non-statistical methods (Per Vagn Freytag, et al, 2001, Lei-da Chen et al., 2000). Non-statistical methods, e.g. data mining technology, mainly are arisen methods in customer segmentation (Agnes Nairn, and Paul Bottomley, 2003, 2003, Jon Kleinberg, et al., 2004).These literatures use single data mining technology to analyze single business issue, and have got some good results. But these applications have one obvious shortage that is those segmentation methods may be able to segment customer into certain groups by different attributes, which are indirect indicators rather than direct one.

Our research efforts are aimed at the synthesis of the customer value concept by segmenting customer by CLV, a direct indicator representing monetary value of the customer. This paper seeks to formally infuse the customer value concept into customer relationship management by developing an integrated customer segmenting framework. It not only provides the way to guiding data mining technology used in customer segmentation, but also realizes a new customer segment model called Customer Segmentation Method Based on CLV (CSMBC). Firm can segment customer effectively and manage different customer group separately based on each group's value.

2. Customer Lifetime Value Concept

Within the scope of this paper, Customer Value was defined as from a supplier-oriented point of view as the customer's economic value to the enterprise, which is a definition different from the frequently employed demand-oriented view of customer value as the enterprise's or its products' value to the customer (Cornelsen 2000, Staat 2002). The Customer Lifetime Value (CLV) represents an application of the principles of contemporary finance to the evaluation of customer relations (Day/Fahey 1988; Doyle 2000). The CLV measures the profit streams of a customer across the entire customer life cycle. The model is aimed at the assignment of a profitability figure to the customer which is based on all prospective and directly attributable in-payments and out-payments. This procedure also accounts for effects that go beyond customer's own transactions, for example referring the products to other potential customers through word of mouth activities. Although a considerable number of CLV models have been developed so far, no generally accepted, superior approach exists (Jackson 1992).

When developing a customer management strategy, company needs to know how to evaluate customer value, especially the potential capability to create profit. In general, a CLV model has three components: customer's value over time, customer's length of service and a discounting factor. Each component can be calculated or estimated separately. When modeling CLV in the context of a customer relationship management, there is an additional issue, which is the need to calculate a customer's CLV before and after the company effort.

978-0-7695-3305-6/08 $25.00 © 2008 IEEE
DOI 10.1109/FSKD.2008.548

Here is the formula to calculate customer lifetime value provided by Jackson (1992):

$$CLV = \sum_{i=1}^{n} (R_i - C_i)(1+d)^{-i} \quad (1)$$

Where R_i is the revenue from existing customers during the i th period, and the C_i is the cost in the i th period, i presents the number of time in consideration to calculate CLV, it could be week, month, or some other scale of time that makes the most business sense. The d is the discount rate.

Paul D. Berger and Nadal. Nasr (1998) developed Jackson's model by introducing the ration of customer retaining and replacing the revenue and cost of customer with the function, $\pi(t)$. So the CLV model could be written as following:

$$CLV = \sum_{i=1}^{n} \pi(t) \times \gamma \times (1+d)^{-i} \quad (2)$$

Now, the challenge we face is how to generate forward-looking forecasts the parameter of γ and the value of $\pi(t)$ for each period time. At the heart of such effort to calculate the CLV turn to develop a model of customer purchasing that accurately characterizes depend on buyer behavior, which will be explained in details in the third section.

3. CLV Models

The Pareto/NBD model that Schmittlein, Morrison, and Colombo (1987) developed a popular and powerful model in explaining the flow of transactions in a noncontractual setting and Reinartz and Kumar (2000, 2003) provided excellent illustrations. This model is based on the following general assumptions about the repeat buying process:

1) Customers go through two stages in their "lifetime" with a specific firm: They are active for some period of time, and then they become permanently inactive.

2) While customers are in active stage, they can place orders whenever they want. The number of orders a customer places in any given time period (e.g., week, month) appears to vary randomly around his or her underlying average rate.

3) When a customer becomes inactive is unobserved by the firm. The only indicator of this change in status is an unexpectedly long time since the customer's transaction.

4) The inclination for customers to "drop out" of their relationship with the company is heterogeneous. In other words, some customers are expected to become inactive much sooner than others.

5) Customer purchase rates (while a customer is active) and drop out rates vary independently across customers.

To develop our model, we assume that monetary value is independent of the underlying transaction process. Although this may seem counterintuitive (e.g., frequent buyers might be expected to spend less per transaction than infrequent buyers), our analysis lends support for the independence assumption. This suggests that the Value per Transaction can be factored out, and we can focus on forecasting the Probability of Purchase (PP) and the "flow" of future transactions (discounted to yield a present value). We can then rescale this number of Expected Number of Transactions (ENT) by a monetary value "multiplier" to yield an overall estimate of lifetime value:

CLV = VT×PP×ENT.

We first develop sub-model for PP, and then we introduce the sub-mode for ENT. The only customer-level information that this model requires is recency and frequency. The notation used to represent this information is (X = x, t, T), where x is the number of transactions observed in the time interval (0, T] and t ($0 < t \le T$) is the time of the last transaction. The information about the recency and frequency are sufficient statistics for an individual customer's purchasing recorder. Schmittlein, Morrison, and Colombo (1987) derive expressions for several managerially relevant quantities, including E[x(t)], the expected number of transactions in a time period of length t, which is central to computing the expected transaction volume for the whole customer base over time.

$$E[x^* \mid r,\alpha,s,\beta,X=(x,t_1,t_2),t^*] = P[T>t_2 \mid r,\alpha,s,\beta,X=(x,t_1,t_2)] \cdot E[r^*,\alpha^*,s^*,\beta^*,t^*] \quad (3)$$

Where r,α,s,β are system parameters fot customer base, which can be calculated by Matlab program; x and t were mentioned above; t_2 is the expect time period; and

$$r^* = r + x, \alpha^* = \alpha + t_2,\ s^* = s,\ \beta^* = \beta + t_2.$$

P("active"|X= x, t, T) the probability that a customer with observed behavior (X = x, t, T) is still active at time T.

$$P[\tau > T \mid r,\alpha,s,\beta] = \int_0^\infty \int_0^\infty P[\tau > T \mid \lambda,\mu] \cdot f(\lambda,\mu \mid r,\alpha,s,\beta,x,t,T) d\lambda d\mu \quad (4)$$

Where the weights λ,μ are particular value for customer base, and the $f(\lambda,\mu)$ is updated distribution of λ and μ given the observed purchase pattern.

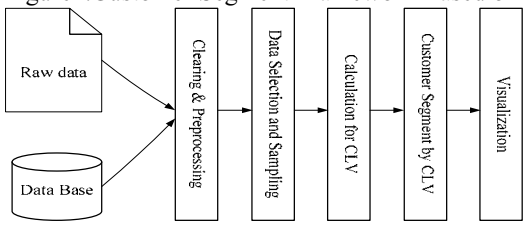
These calculation functions motioned above are based on the Pareto/NBD model, a more logical approach would be switch from discrete-time formulation to continuous-time formulation (as is often done in standard financial analyses) over an infinite time horizon.

4. Customer Segment Framework

As practitioners are enthusiastically seeking out groups of profitable customers with high CLV, some academics are beginning to question whether segments are actually stable entities and whether they really exist at all(Jon Kleinberg, et al., 2004). The customer segmentation based on data mining can solve this problem because the ways could study from new information that input afterward and get new rules. It provides completely support to the dynamic management process of customer attaining, retention and win-back, and building the mapping relationship among customer attributes, dispersive and continuous attributes. Setting each customer attribute as a dimension and setting each customer as a particle, the whole customers in an enterprise can form a multidimensional space, which has been defined as the attribute space of the customer.

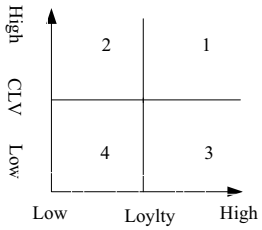
In general, data mining is a process of searching previously unknown but meaningful information, such as decision making patterns and trends, by sifting through large data sets and utilizing a combination of pattern-recognition, model building, and validation techniques. In this paper, we chose clustering from the various data mining methods due to its visual appeal and simplicity in segmenting object into group. To construct cluster customer base, we followed five key steps as shown in Figure 1 and described below.

Figure1.Customer Segment Framework Based on CLV



Segmenting markets by consumption patterns can be quite insightful for understanding the customer base. Differentiated customer relationship management strategies are suit for various user groups, as showed in figure 2. By classifying customer accounts based on CLV, company can develop effective and profitable strategies for retaining and upgrading customers.

Figure2. Customer Segment Model Based on CLV



Distinctly, cluster 1 is composed of high CLV who are in the stable churn phase with high loyalty. These people should be satisfied with the products or services and intend to maintain a long-relationship with the firm. They are cash cows to the company and it is not necessary to inject huge cost to retain them, but only some regular contact. Cluster 2 contains high CLV at risk, and company should pay a lot of attention to them who may become tomorrow's stars. They are probably people who are hesitating over the products or services and have not built up loyalty to the firm yet. Therefore, firm need to put more efforts on customers in order to set up long relationship with them. Loyal but low CLV customers constitute cluster 3. They may feel satisfactory with the products or services but their economic capability is comparatively low. They are part of customer base of a company and adequate cost is necessary to keep them from attracted by competitor. Cluster 4 is made up of customers whose value is low and their faith hasn't been constructed. These people probably dissatisfy the products or services or are sensitive to price. Firm needn't allocate much resource to them because they may contribute negative influence to the firm.

Generally, this sort of analysis is convenient to the understanding of customer behavior with different value of CLV. This effective strategy classifies customers via CLV to design differentiated customer relationship management for different customer segmentation. In sum, the CLV model can be beneficial to customer relationship management activities. It is useful to think about how to hold high CLV customers, upgrade light and medium users, build customer loyalty, understand buying motives to meet/exceed expectations, use appropriate selling strategies for each targeted usage group, win back lost customers, and learn why nonusers are not responding to a value proposition.

5. Empirical Analyses

5.1 Data Source

The statistical and data−mining models discussed above can be applied to the Frequent Flyer Program (FFP) of airlines in whose information resides in internal company databases producing every flight and travel record. These data include travel detail data, which describes the travel that when and where the passenger takes the plane from one city to another with the certain ticket price. For the purposes of the CLV modeling reported in this article, we obtain a data extract from the warehouse where each customer record includes flight record and other demographic information, such as age, income, gender, marriage status etc.

A sample of 6,470 active passengers was randomly selected from the entire customer base from an airline in china. All these passengers flight record includes forty

months, from January 2003 to April 2006. During the first twelve month, only 3290 passenger took the aircraft, so only these customers were kept as the analysis sample. Each passenger his three parameters (x, t and T). The x means the passenger's total number of flight during the observe time windows, and t is the point for the last time to take plane. T indicates 40 if we calculate with transaction interval by month.

5.2 Calculate CLV

In this paper, the four parameters estimated from customer base is $\gamma = 0.389$, α=283 , $s = 2.777$, β=3360. The value of γ is small indicates that the heterogeneous among passenger is vary clear. The average number of take plane is $\gamma / \alpha = 0.136$, and the ratio of churn for the passengers accords to an exponentially distribution duration with the parameter, $e = s / \beta$ = 0.083. Table 1 shows the probability of flight, expect number of flight, and the CLV for each passenger.

Table 1. Flight Record, PP, ENT and CLV for FFP

CARDNO	x,t,T	PP	ENT	CLV
4926853	0,0,24	0.156	0.023	4634
1779713	5,7,24	0.014	0.8386	3844
4954676	3,7,24	0.063	0.342	2153
872833	8,13,24	0.485	0.568	2305
…	…	…	…	…
2356001	1,6,24	0.182	0.034	5115
5753576	6,24,24	0.876	0.763	487
4953321	1,14,24	0.501	0.072	1121
5737723	6,7,24	0.006	0.653	1390

5.3 Result of Clustering

A K-means cluster analysis was performed. According to former literatures, we cluster the customer sample respectively with CLV, t and x. Generally, the parameter must be appointed in the beginning with K-means. Now according with advanced experience, the parameter K is set as 2-6. Using the software of SPSS, finally four clusters are gained (The result presented in Table 2).

Table 2. Clustering Results

Cluster ID	Number of Each Cluster
1	768
2	2139
3	329
4	54
Valid	3290
Missing	0

Due to the information in table 2，it is clear that more than seven hundreds of people are assigned into cluster 1, and the cluster 2, 2130, is largest group in 4 clusters. Furthermore, the sum of number for cluster 1 and 2 is close to 3000, 88.4% of the data sample. These are the customers who are probably giving some of their business to competitors or with little demand for the service actually. While, there are 329 persons belongs to cluster3, and only 54 persons is tagged by cluster 4. The fact is that both cluster 3 and 4 are the top group customer with excellent loyalty and of high profitability for the firm.

5.4 Segments Analysis

With the analysis on all kinds of features, we can find the rules and patterns as showed in Table 3.

Table 3.Customer Cluster Features

Cluster ID	Feature		
	CLV	Time to no show	Frequency to Take Plane
1	low	long	Few
2	low	short	Few
3	middle	short	not often
4	high	very short	Often

Cluster1: these customers share some important similar features, such as low value of CLV, long time no show (8-10 months) and seldom travel by air (1-2 times every year). As shown in table 2, the number of these customers reaches almost one quarter in the data sample. Because of their low CLV and weak motivation to travel by plane, it is no necessary for airlines to take any customer relationship management action for this group.

Cluster2: generally, it could be inferred that these customers in cluster 2 may be the new arrivals (3-4 month ago) for FFP program from their low CLV but short time to no show. Airlines should take these customers carefully during the recognize stage, and find out the right customer who has the great patient to be developed in the future.

Cluster3: these customers have taken the airplane recently, 3-4 months ago. Due to their frequency to travel by air higher comparatively, 7-10 times every year, it is reasonable to expect they have great probability to take plane for next trip. They satisfy with present services will keep them stay in FFP and choose the same provider in the future. So company need understand these customers' real feelings and expectations deeply, and distribute more marketing resource on them.

Cluster4: these customers are the most valuable customers for airlines, and the transaction record shows that they travel by air very frequently, more than 20 times annually. Company should concern the last transaction time because the risk to churn increase if the interval becomes longer. To keep and retain these valuable

customers, company need allocate more resources to them and encourage them to use and enjoy the service as many as possible.

6. Conclusions

A key purpose of marketing is to identify the customers or segment them with the greatest value-creating potential and target them successfully with corresponding marketing strategies to reduce the risk of these high lifetime value customers defecting to competitors (Andrew Banasiewicz, 2004). Segmenting customer is the basic work of data mining according to known historic segmentation information. The training data used to construct segment forecast mode can be historic data or exogenous data that are gained from experience or survey.

For an enterprise, how to use data mining technology, and how to choose the proper segment indicators are very important, which result will impact the company implement customer relationship management strategy directly. To answer this question, this paper proposes a segmentation framework based on data mining and constructs a segmentation methods based on CLV. By clustering customer with the discrete-time and continuous-time segment indicators, firm can make different relationship management action for each segment.

The segment method in this paper has been applied to the civil aviation industry, and it also may be used in other industries such as telecommunication, finance service etc.

7. References

[1] Adrian Sargeant, "Customer lifetime value and marketing strategy: how to forge the link", The Marketing Review, Jan, 2001, pp.427-440.

[2] Agnes Nairn, and Paul Bottomley, "Cluster analysis procedures in the CRM era", International Journal of Market Research,2003, Vol. 45 Quarter 2.

[3] Andrew Banasiewicz, "Acquiring high value, retainable customers", Database Marketing & Customer Strategy Management, Jan, 2004, pp.21–31.

[4] Berger, P.D. and Nasr, N.I, "Customer lifetime value: marketing models and applications", Journal of Interactive Marketing, Dec, 1998, pp.17-30.

[5] Clemons, E., "Gathering the nectar", Understanding CRM: Financial Times, No. Spring Supplement, 2000, pp.24-7.

[6] Fraley, C. and Raftery, A. E., "Model-based clustering, discriminate analysis, and density estimation". Journal of the American Statistical Association 2002, pp. 611-631.

[7] Jon Kleinberg, Christos Papadimitriou, Prabhakar Raghavan, "Segmentation Problems". Journal of the ACM, March, 2004, Vol. 51, No. 2, pp. 263–280.

[8] Michael J etc, "Knowledge management and data mining for marketing", Decision Support System, May, 2001, pp. 127-137.

[9] Per Vagn Freytag, et al, "Business to business market segmentation", Industrial Marketing Management, 2001, 30(6), pp.473-486.

[10] Verhoef P.C., Spring P.N., Hoekstra J.C. The commercial use of segmentation and predictive modeling techniques for database marketing in the Netherlands. Decision Support Systems, 2 003,34(4):pp.471-481.

[11] Wedel, M., W.A. Kamakura and U. Bockenholt, "Marketing data, models and decisions," International Journal of Research in Marketing, 2000, 17(2-3), pp.203-208.

[12] Werner Reinartz, Manfred Krafft, Wayne D. Hoyer, "The Customer relationship management process: its measurement and impact on performance", Journal of Marketing Research, Aug, 2004, 293 Vol. XLI, pp.293–305.