

Study on Purchase Probability Model in CRM Systems

Jiayin Qi¹, Huaying Shu¹, and Huaizu Li²

1. Economics and Management School, Beijing University of Posts and Telecommunications, 10 Xi Tu Cheng Road, Haidian District, Beijing, China, 100876

mailto:ssfqjy@263.net,

shuhy@bupt.edu.cn,

WWW home page: <http://www.sem.bupt.cn>

2. Professor, School of Management

Xi'an Jiaotong University

Xi'an, Shannxi, China, 710049

Email: Hзли@xjtu.edu.cn

Abstract. Representative customer's purchase probability is the basis to analyze the purchase behavior of always-a-share customer's segment. Currently, analyzing the representative customer's purchase probability with the Dwyer model is quite complicated. Using uncertain reasoning, a backtracking Dwyer model and its algorithm are presented in this paper, which solves this problem in a more effective way. The work of this paper is helpful to design analytical CRM systems.

1 Introduction

Customer relationship management (CRM) has been one of the most exciting information system fields since the 1990s [1]. CRM systems are a set of enabling IT systems supporting a business strategy to build long term, profitable relationships with specific customers. The key objective of CRM systems is to provide customer with relevant information and deliver individualized products and services. Call center, sale force automation (AFA), marketing automation solution (MAS) and customer service and support (CSS) are the four software packages of CRM systems. The architecture of CRM systems is shown in Fig. 1.

Please use the following format when citing this chapter:

Qi, J., Shu, H., Li, H., 2006, in International Federation for Information Processing, Volume 205, Research and Practical Issues of Enterprise Information Systems, eds. Tjoa, A.M., Xu, L., Chaudhry, S., (Boston:Springer), pp.643-647.

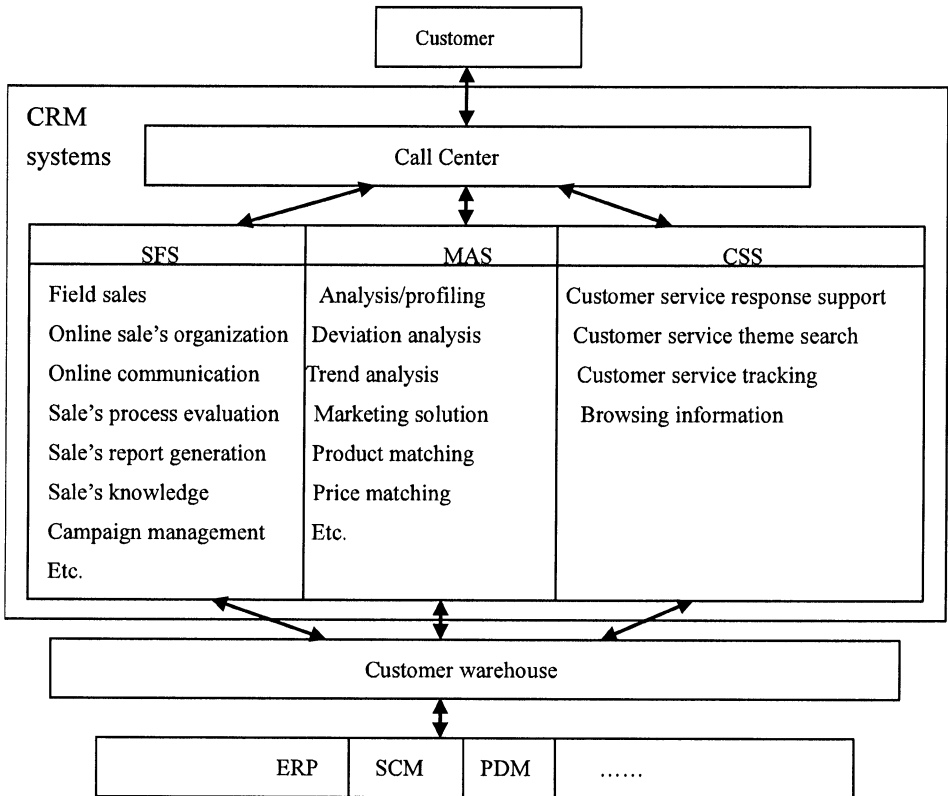


Fig. 1. CRM Architecture

Customer buying behavior analysis is one of the fundamental bases to design the functionalities of a CRM system. Customer's purchase behavior is one facet of customer profiling, the first step to analyze customers' profitability, and the base to provide marketing solution, organize sales activity and provide active customer service. Purchase probability is one of the most important variables to describe always-a-share customers' buying behavior and one of the most critical parameters [2] to calculate lifetime value of always-a-share customer [3-5]. If there is an efficient way to get purchase probability, it will improve the analytical functions of CRM systems.

Dwyer [4] obtained always-a-share customer's purchase probability through building a purchase decision-making tree. The path parameter of leaf nodes and the end condition of the tree were created through the customer's historical buying data. The possibility of next purchase depends on the late purchase of customers. Blattberg and Thomas [6] still use this method. The difference is that they use RFM (Recency, Frequency, Monetary) index to build purchase decision-making tree, while Dwyer only used "Recency" to construct the decision-making tree.

Blattberg’s result is more precise; however, it is fussier to apply. Generally, the principles of Dwyer and Blattberg and Thomas are the same.

As this method was first put forward by Dwyer, we call it the Dwyer model for convenience. Through the Dwyer model, the dynamic customer quantity change can be calculated. The ratio of dynamic customer quantity and the initial quantity of the customer segment is the representative customer’s purchase probability in each purchase period time. However, as the time variable t increases steadily, the Dwyer model becomes more and more complex. It is more and more difficult to get customer’s quantity of the t th purchase period simply and conveniently. So, it is difficult to get the representative customer’s purchase probability when t is a big.

The Dwyer model is used in most of the CLV researches to calculate representative customer’s purchase probability [2, 4, and 7]. This is the reason why we design a new method. With this new method, we can get the representative customer’s purchase probability much more easily and calculate always-a-share customer’s CLV more efficiently. Using uncertain reasoning, a back-tracking Dwyer model and its algorithm are presented in this paper. It is proved that the new method is more effective than the Dwyer model.

2 The Back-Tracking Dwyer Model

We assume that customer’s purchase behavior in the t th time period is effected by the pervious j periods’ buying behavior. If during the previous j periods, a customer didn’t buy anything, the purchase probability in the t th time period is zero.

For convenience, the following definitions should be made firstly:

Event A_t : a purchase behavior happens in the t th ($t = 1, 2, \dots$) time period; the probability of this event is $\Pr(A_t)$;

Event space Ω : Ω is the event space, with $A_t \subseteq \Omega$;

Event B_k : the last purchase behavior happened in $(t-k)$ th time period, the probability of this event is $\Pr(B_k)$, where $k < t, k = 1, 2, \dots$.

Under the condition of event B_k happening, the probability of event A_t happening is $\Pr(A_t | B_k)$. Given the fact that event B_k is a total division of event space Ω and $\bigcap_{k=1}^j B_k = \phi$, we can make the following equation based on the total probability formula:

$$\Pr(A_t) = \sum_{k=1}^j \Pr(B_k) \times \Pr(A_t | B_k) \tag{1}$$

The above $\Pr(B_k)$ can also be described in the following form:

$$\Pr(B_k) = \Pr(A_{t-k}) \times \Pr(\bar{A}_{t-k+1} | A_{t-k}) \times \Pr(\bar{A}_{t-k+2} | A_{t-k}, \bar{A}_{t-k+1}) \quad (2)$$

$$\times \dots \times \Pr(\bar{A}_{t-2} | A_{t-k}, \bar{A}_{t-k+1}, \dots, \bar{A}_{t-3}) \times \Pr(\bar{A}_{t-1} | A_{t-k}, \bar{A}_{t-k+1}, \dots, \bar{A}_{t-1})$$

Obviously,

$$\Pr(\bar{A}_{t-k+1} | A_{t-k}) = 1 - \Pr(A_{t-k+1} | B_1)$$

$$\Pr(\bar{A}_{t-k+2} | A_{t-k}, \bar{A}_{t-k+1}) = 1 - \Pr(A_{t-k+2} | B_2)$$

...

$$\Pr(\bar{A}_{t-2} | A_{t-k}, \bar{A}_{t-k+1}, \dots, \bar{A}_{t-3}) = 1 - \Pr(A_{t-2} | B_{k-2})$$

$$\Pr(\bar{A}_{t-1} | A_{t-k}, \bar{A}_{t-k+1}, \dots, \bar{A}_{t-2}) = 1 - \Pr(A_{t-1} | B_{k-1})$$

Thereby, we can build the dynamic model of representative customer’s purchase probability:

$$\left\{ \begin{array}{l} \Pr(A_t) = \sum_{k=1}^j \Pr(B_k) \times \Pr(A_t | B_k) \\ \Pr(B_k) = \Pr(A_{t-k}) \times \prod_{l=1}^{k-1} (1 - \Pr(A_{t-(j-k)} | B_l)) \end{array} \right\} \quad (4)$$

The way to make out $\Pr(A_{t-k})$ is the same as $\Pr(A_t)$. $\Pr(A_t | B_k)$ is a constant and estimated by experience. With the assumption that customer’s future purchase behavior is decided by the recent behavior, we can get customer purchase probability during any period using recursion and iterative algorithm based on the formula (4) if the initial purchase time record (the initial condition) is known.

The representative customer’s purchase probability of always-a-share in any time periods is easy to get if we design the arithmetic for formula (4). When the Dwyer model is built with RFM, the analysis of back-tracking Dwyer model with RFM is as same as the process above, namely back-tracking Dwyer model with Recency only.

3 Comparison

Comparing back-tracking Dwyer model with the Dwyer model, the new method is better than the old one for time-saving and space-saving in computer operation. Table 1 shows the details:

Table 1. Comparing back-tracking Dwyer model with Dwyer model

	Arithmetic Design	Data Structure
Back-tracking Dwyer model	Recursion and Iterative Arithmetic	Only search the efficient nodes of the tree
Dwyer model	Enumerative Arithmetic	Search all the nodes of the tree, including the non-effective ones

Considering the arithmetic design, we use recursion and iterative to compute the representative customer's purchase probability. It is time-saving and space-saving, and it is more efficiency. While the Dwyer method is an enumerative arithmetic. It is slow, and hard to finish. The non-effective nodes are referred to the ones which are dead before the observed time period. Our method ignores these nodes, but the Dwyer model only stops searching when the nodes terminate.

4 Conclusions

To always-a-share customer, back-tracking Dwyer model is presented in this paper to calculate the representative customer's purchase probability of customer segment. With this method, it is more convenient to predict customer's purchase behavior, calculate CLV and support the development of CRM systems.

Acknowledgements

Research of this paper is funded by the National Natural Science Foundation of China (Project No.70371056) and the Key Laboratory of Information Management & Information Economy of Ministry of Education of the People's Republic of China.

References

1. R. Ling and D.C. Yen, Customer relationship management: an analysis framework and implementation strategies, *Journal of Computer Information Systems* 82-97 (Spring 2001).
2. M. Calciu and F. Salerno, Customer value modeling: synthesis and extension proposals, *Journal of Targeting, Measurement and Analysis for Marketing* 11(2), 124-147 (2002).
3. B.B. Jackson, Building Customer Relationship that Last, *Harvard Business Review* 120-128 (November-December 1985).
4. F.R. Dwyer, Customer Lifetime Valuation to Support Marketing, *Decision Making* 11(4), 6-13 (Fall 1997).
5. J. Qi, H. Li, and H. Shu, A new method to design customer retention rate model, *The Journal Of Management Engineer* 10(4), 60~64 (2004).
6. R. Blattberg, G. Getz, and J. Thomas, Customer Equity Calculation. http://www.Customequity.Com/ce_in_depth/ce_calculations/acquisition_equity.html, 2001-10-24.
7. P.D. Berger and N.I. Nasr, Customer Lifetime Value: Marketing Models And Applications, *Journal of Interactive Marketing* 12(1), 17-30 (Winter 1998).