

Customer Relationship Termination Problem for Beta-Geometric/Beta-Binomial Model of Customer Behavior

Y. Samimi^{*1}, A. Aghaie², H. Shahriari³

We deal with the relationship termination problem in the context of individual-level customer relationship management (CRM) and use a Markov decision process to determine the most appropriate occasion for termination of the relationship with a seemingly unprofitable customer. As a particular case, the beta-geometric/beta-binomial model is considered as the basis to define customer behavior and it is explained how to compute customer lifetime value when one needs to take account of the firm's choice as to whether to continue or terminate relationship with unprofitable customers. By numerical examples provided by simulation, it is shown how a stochastic dynamic programming approach can be adopted in order to obtain a more precise estimation of the customer lifetime value as a key criterion for resource allocation in CRM.

Keywords: *Customer relationship management, Customer lifetime value, Empirical Bayesian models, Beta-geometric/beta-binomial model, Stochastic dynamic programming.*

Manuscript received on 18/10/2012 and accepted for publication after revision on 15/12/2012.

1. Introduction

The customer relationship management (CRM) process at the customer-facing level has been defined as a systematic process to manage customer relationship initiation, maintenance, and termination in order to maximize the value of the relationship portfolio [11]. Although customer retention policies were often assumed as the foundation of CRM, customer relationship abandonment has been recently noticed as an important issue in CRM [2-5]. Specially, when we are not precisely aware of the customer churn time, a key question arises as to whether it is economical to keep spending on maintaining the relationship with a customer who has not shown up recently. Basically, relationship termination has been counted among the key customer intervention strategies for the management of the relationships which are not beneficial for the firm [13]. Despite the importance of this issue, the results of several exploratory studies such as Holmlund and Hobbs [7] in a business-to-business service setting and Helm et al. [6] in inter-organizational manufacturing

* Corresponding Author.

¹ Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran. E-mail: y_samimi@kntu.ac.ir

² Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran. E-mail: aaghaie@kntu.ac.ir

³ Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran. E-mail: hshahriari@kntu.ac.ir

setting have revealed that many firms rarely document a clear relationship-ending strategy, and in only a few cases such decision is often based on intuition rather analysis of objective data.

The termination of customer relationship has been categorized by Reinartz et al. [11] to be originated from customer causes for example when a customer no longer needs to use the current product, competitive reasons that result in customer churn, or internally intended reasons when the firm decides to abandon customer relationship. Regarding the last decision, which is indeed the focus of our attention here, a reasonable approach can be to analyze the cash flow of customer relationship in order to balance the profits of the relationship over the long run against the related marketing expenditures. Naturally, a negative present value of discounted cash flow can convince the manager not to continue the relationship with such a seemingly unprofitable customer. To this end, one inevitably needs to have a reliable prediction of the long-term net revenue of customer relationship or the so-called customer lifetime value (CLV). Although regression-based techniques have found a widespread use in such conditions to predict the relationship duration as well as the purchase volume over tenure, such investigation would result in biased results, since it might not have considered the flexibility of the firm's marketing decision making process in prediction of future customer behavior and value. Let it not remain unsaid that these models are generally established using historical revenue data, without taking account of the future potential of the customer or the effect of the firm's marketing policies to enhance the current level of the relationship. With respect to this issue, Sun et al. [14] and Rust and Chung [8] requested more attention to be paid when customer valuation techniques are employed.

Termination of customer relationship is always confronted with the risk of losing a potentially valuable customer. The future value of an apparently dead customer, on the one hand, and the collection of marketing costs spent periodically for the purpose of customer maintenance or promotion, on the other hand, should be evaluated and compared in order to make a reasonable choice in termination of a relationship. Hence, the decision of relationship termination has to be made on the basis of the future potential of a seemingly unprofitable customer. As recommended by Haenlein et al. [5], to incorporate the potential value of customer's future transactions, we need to use a dynamic programming approach enabling us to consider any likely transaction in the future and the effect of the customer value over the lifetime. For the same reason, using a well-established method such as real-option analysis (ROA) has proved very useful in estimating net present value of customer value over the remaining lifetime, since the decision to continue/terminate relationship with a customer always creates the choice of further investment in the relationship in subsequent periods. To this end, here we use a stochastic dynamic programming approach to solve customer relationship abandonment problem. The investigated problem here differs from Pfeifer and Carraway's case in which a Markov decision process is used based on a recency-frequency model of customer behavior [9]. Here, the underlying model of customer behavior is a well-established and a quite well-known model for customer-base analysis, called Beta-Geometric/Beta-Binomial (BG/BB) model, whose parameter state-space here is represented over a continuous bivariate space. Regarding applications of the dynamic programming approach and particularly Markov decision process (MDP) in CRM decision making problems, [9] and [1] may be referred as very interesting studies.

The remainder of our work is organized as follows. In the next section, we will outline the main assumptions and some key results of the Beta-Geometric/Beta-Binomial (BG/BB) model as the underlying model of customer behavior in our investigation. Section 3 proposes a detailed explanation of the dynamic programming approach to obtain an optimal solution for relationship termination problem. Numerical evaluation of the solution approach will be given in Section 4. Section 5 provides concluding remarks and some suggestions for future research studies.

2. Beta-Geometric/Beta-Binomial (BG/BB) Model

Probability models for customer-base analysis have been classified in four main categories on the basis of the business settings and customer relationship types [2]. The BG/BB model, as a hierarchical probability model that takes heterogeneity of customers into account, has been developed for non-contractual settings where customer transactions occur or recorded on a regular discrete basis. More specifically, using the BG/BB model one can represent two aspects of customer behavior: customer's purchasing, while he is still active, and time until each customer permanently becomes inactive. Five underlying assumptions of the BG/BB model, which has been introduced and illustrated by Fader et al. [3] and Fader et al. [4] as a very useful model in customer-base analysis, are as follows:

- (a) While active, the customer buys on any given transaction opportunity with probability p ,
- (b) Heterogeneity in p follows a beta distribution with the probability distribution function $f(p|\alpha, \beta) = p^{\alpha-1}(1-p)^{\beta-1}/B(\alpha, \beta)$, $0 < p < 1$.
- (c) An active customer becomes inactive at the beginning of the next transaction opportunity with probability q . This implies that the customer's (unobserved) lifetime is distributed across transaction opportunities according to a (shifted) geometric distribution.
- (d) Heterogeneity in q follows a beta distribution with probability distribution function $f(q|\gamma, \delta) = q^{\gamma-1}(1-q)^{\delta-1}/B(\gamma, \delta)$, $0 < q < 1$.
- (e) The transaction probability p and the dropout probability q vary independently across the customers.

Here, $B(a, b)$ denotes the beta function, defined as $\int_0^1 t^{a-1}(1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. The aforementioned studies have shown that to obtain maximum likelihood estimates of the model parameters $\alpha, \beta, \gamma, \delta$, the only required information about customer's behavior can be represented by the triple (n, m, x) , where n is the number of transaction opportunities, m denotes the last transaction opportunity in which a purchase has been made, and x is the number of purchases occurred during the first m periods.

By referring to the key results of the BG/BB model, it can be shown that for an active customer with the observed purchasing behavior, the present value of the discounted expected future cash flow or the so-called discounted expected residual transactions (DERT) is computed as follows:

$$\begin{aligned}
& DERT(d|\alpha, \beta, \gamma, \delta, n, m, x) \\
&= \int_0^1 \int_0^1 \frac{p(1+d)}{(d+q)} \\
&\times \frac{p^x(1-p)^{(n-x)}(1-q)^{n+1}}{L(p, q|n, m, x)} f_{posterior}(p, q|n, m, x, \alpha, \beta, \gamma, \delta) dpdq \\
&= \frac{B(\alpha+x+1, \beta+n-x)B(\gamma, \delta+n+1)}{B(\alpha, \beta)B(\gamma, \delta)} \\
&\times \frac{{}_2F_1\left(1, \delta+n+1; \gamma+\delta+n+1; \frac{1}{1+d}\right)}{L(\alpha, \beta, \gamma, \delta|n, m, x)}, \tag{1}
\end{aligned}$$

where d is the discount rate, ${}_2F_1(\cdot)$ is the Gaussian hyper-geometric function, $L(p, q|n, m, x)$ is the probability of observing the purchase history (n, m, x) , and $L(\alpha, \beta, \gamma, \delta|n, m, x)$ denotes the likelihood function for the purchase history (n, m, x) , with the expectation of $L(p, q|n, m, x)$ taken over the prior joint distribution of p and q . Furthermore, to convert the prior distribution of the parameters to the posterior one, we can use the following updating relationship:

$$f_{posterior}(p, q|n, m, x, \alpha, \beta, \gamma, \delta) = \frac{L(p, q|n, m, x)f_{prior}(p, q|n, m, x, \alpha, \beta, \gamma, \delta)}{L(\alpha, \beta, \gamma, \delta|n, m, x)}. \tag{2}$$

Accordingly, considering that a fixed marketing cost, denoted by M_c , is spending as the periodical customer maintenance expenditure, through modification of the above formula we can compute the mean present value of the customer relationship cash flow over the customer's residual lifetime as follows:

$$E(CRLV) = DERT(d|\alpha, \beta, \gamma, \delta, n, m, x) - M_c \times \frac{(1+d)}{d}. \tag{3}$$

It should be noted that the fixed periodical maintenance cost, M_c , is spending as long as the relationship with customer continues. In other words, since customer churn time is unobservable, the firm keeps spending on customer relationship until it decides not to continue the relationship. Hence, at the end of each period, we need to decide whether to continue or abandon the relationship with customer. As mentioned previously, to specify the optimal occasion to terminate customer relationship, the first approach may be to compute the expected value of future cash flow at the end of each period in order to find out whether or not it will be profitable to continue the relationship. This procedure uses the last observation of customer purchase behavior to complete the available information for obtaining the posterior distribution of p and q , and then computes $E(CRLV)$ using (2).

The above approach may lead to biased decisions on the basis of which the firm may stop the relationship with an already-active customer who has not made any purchase for a few number of periods. To prevent such an error, we need to take account of the potential of each seemingly dead

customer to make a new purchase in the future. To this end, rather than basing our decision only on the last updated distribution, we use a dynamic programming approach in which the possibility of the future transactions and naturally their immediate effect on customer behavioral state space are also taken into account.

3. A Dynamic Programming Approach

As mentioned before, the state of customer behavior can be specified by (n, m, x) from which a purchase action moves the customer to a new position of $(n + 1, n + 1, x + 1)$, while a no-purchase action leads to the new state of $(n + 1, m, x)$. Considering the fact that each state can be mapped into a specific joint distribution of the main parameters p and q , we can reach another representation of the initial state defining vector (n, m, x) based on the moments of the joint distribution of the parameters. In order not to restrict ourselves to a limited range of values (n, m, x) to be able to define customer behavioral state space, we prefer to use the secondary representation which can have a limited range and additionally can cover much larger number of states of customer behavior. This will be explained more precisely in this section.

Concerning the model of customer behavior outlined in the previous section, we deal with a continuous state space for the customer behavior parameters p and q over which the initial heterogeneity distribution is characterized by four hyper-parameters α, β, γ , and δ . After we observe the last customer action, which can be either a purchase or a no-purchase, the posterior joint distribution of the parameters p and q can be obtained by the updating procedure. In other words, with each observation, the joint probability distribution function (pdf) of p and q is revised. For the sake of simplicity, from the complete information about the posterior *pdf* we choose to use the expected value of the parameters p and q . Therefore, the updating mechanism, which runs after each new observation of customer behavior, shifts us from the current position to a new position whose coordinates are defined by a vector.

This fact that customer behavior can be mapped into the continuous square state space $(0,1) \times (0,1)$ of real numbers, and at the end of each period, after observing customer action, a transition occurs across this space, allows us to develop a transition matrix to define how a customer moves across the state space. More precisely, we aim to develop a Markov model of customer behavior which enables us to approximately represent the sequential shifts of the bivariate distribution of p and q . Subsequently, we are able to use the Markov decision process to determine the best appropriate action from the action set $\{continue, stop\}$ for all possible states of customer behavior.

To construct a Markov chain model of customer behavior, we need to define the state space as well as the pattern of transitions across states. To specify where a customer stands after each new observation, the continuous state space is divided into a given number of smaller parts whose sizes depend on the chosen precision to perform the approximation. For example, we may construct an $l_1 \times l_2$ mesh where l_1 and l_2 respectively show the number of divisions over respective axes of p

and q . Given the state of customer behavior, the next step involves calculation of the transition probabilities which are related to customer purchase/no-purchase action as well as specification of the state customer moves after the respective action. To obtain the transition probabilities, we need to know the specific information in the form of the initial state space representation (n, m, x) . For this reason, a reasonable approach can be to obtain the transition probabilities over the initial state space and then convert the resulted space into the secondary one based on the first moments of p and q . Accordingly, we chose the range of $\{0,1,2, \dots, 30\}$ for the parameter n , and considered all appropriate values for m and x . Transition probabilities to two new positions of $(n + 1, n + 1, x + 1)$ and $(n + 1, m, x)$ as the result of purchase/no-purchase actions were calculated and then the results were transferred into the meshed space of $(0,1) \times (0,1)$ which characterizes the expected values of the parameters p and q . It is also worth mentioning that the probability of purchase for an alive customer with purchase history of (n, m, x) could be provided by

$$\begin{aligned} & P(\text{purchase in } n + 1 | \alpha, \beta, \gamma, \delta, n, m, x) \\ &= \int_0^1 \int_0^1 \frac{p^{x+1}(1-p)^{(n-x)}(1-q)^{n+1}}{L(p, q | n, m, x)} f_{\text{posterior}}(p, q | n, m, x, \alpha, \beta, \gamma, \delta) dp dq \\ &= \frac{B(\alpha + x + 1, \beta + n - x)B(\gamma, \delta + n + 1)}{B(\alpha, \beta)B(\gamma, \delta)L(\alpha, \beta, \gamma, \delta | n, m, x)}. \end{aligned} \quad (4)$$

Recall that the objective function to specify the appropriate time for termination of customer relationship, which is denoted by T^* , can be formulated as

$$\max_{T^*} E \left\{ \sum_{t=n+1}^{T^*} [(1 - q_t)^t p_t \times R - M_c] / (1 + d)^{(t-n)} \right\}, \quad (5)$$

where R denotes the revenue earned from each customer purchase and M_c is the marketing cost that the firm incurs regularly to maintain the customer. The subscript t in p_t and q_t are used to describe the dynamic decision process which is caused by the Bayesian updating procedure to obtain the joint pdf of p and q at time t on the basis of the customer's last behavioral state and his respective action in period $t - 1$. By obtaining the Markov representation model of customer behavior, we are now able to use a Markov decision process in order to optimize the firm's action in relation to the customer behavior. To this end, the Bellman equation is used as a recursive solution approach to solve the stochastic dynamic programming problem as shown by

$$V_t(s) = \max_a \{0, E[\pi_t(s, a; p_t, q_t) + (1 + d)^{-1} V_{t+1}(s') | s, a]\}, \quad (6)$$

where in the above system of equations, $\pi_t(s, a; p_t, q_t)$ means the immediate profit earned after visiting state s and taking action a that can be either to continue or to terminate the relationship. The above equation which is in fact a representation of the so-called principle of optimality explains that in every state s , we need to choose the appropriate action, a , in such a manner that the expected

sum of the resulting immediate profit, $\pi_t(s, a; p_t, q_t)$, and the subsequent value, $V_{t+1}(s')$, which is itself the optimal value earned from occupying the next state, s' , are maximized [14]. As a matter of fact, $V_{t+1}(s')$ denotes the expected reward from period $t + 1$ onward.

To find the optimal stationary policy of the dynamic programming problem, we chose policy iteration procedure from several solution procedures proposed for the Markov decision process. The policy iteration method consists of two stages called policy evaluation and policy improvement which should be run sequentially until an optimality condition is satisfied [15]. In the first stage, we use the value determination equation (6), to obtain values assigned to the all states. Then, for each state, we assess whether there exists a different action which may result in a greater value compared to the current value obtained in the previous stage. To maximize the expected discounted reward, these two stages are repeated until there is no possible improvement for any state.

4. Numerical Evaluation

In this section, we set forth a numerical evaluation of the dynamic programming problem to decide whether it is profitable for a firm to continue its relationship with a customer. In a particular investigation, we choose the following set of values for the BG/BB model of customer behavior: $\alpha = 3$, $\beta = 1$, $\gamma = 1$, $\delta = 3$. R , M_c , and d , respectively the purchase revenue, the maintenance cost, and the discount rate, are set to 10, 1, and 3%. To solve the dynamic programming problem, we first discretized the continuous state space of $(0,1) \times (0,1)$ into 3600 small squares by dividing each axis into 60 equal parts. Then, we followed the procedure explained in Section 3 to obtain a 3600×3600 transition matrix. This matrix is constructed to represent the model of customer behavior in the form of a Markov chain model that could be used as the basis to formulate the Markov decision process. Accordingly, the solution procedure was run using the policy iteration method as described in the previous section. In comparison with the situation where the firm has no termination policy, the results show that nearly 80% improvement has been made in the total reward in this sense that the sum of the rewards across all states has increased from 6987 units for no-termination policy to 12569 units for the optimal policy.

Fig. 1 illustrates the optimal policy assignment with the mean probability of being inactive shown on the vertical axis. A similar figure, showing the optimal policy distributed with respect to the mean probability of purchase, indicates that the optimal result is much more sensitive to the mean probability of being active than to the expected value of the purchase parameter.

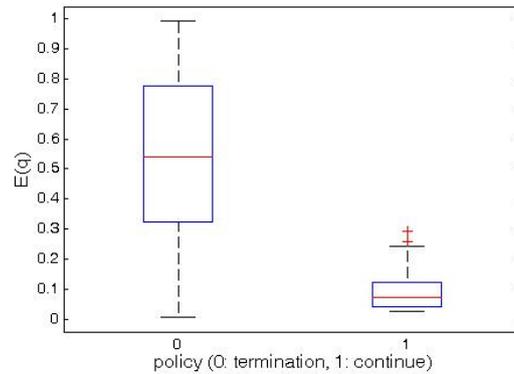


Figure 1. Representation of the mean probability of being inactive versus the optimal policy

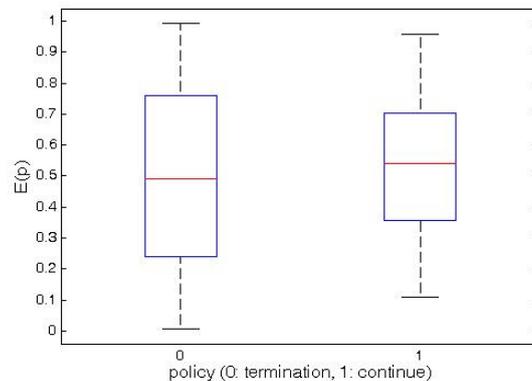


Figure 2. Representation of the mean probability of purchase versus the optimal policy

Table 1 shows the decision proposed by the MDP procedure for customers whose 10-period purchase history have been recorded in the form of the triplet (n, m, x) . Columns 5 and 6 of both halves of the table show the prediction of the customer value over his residual lifetime made respectively by the simple prediction model and the MDP approach. Accordingly, a value of 0 indicates that the MDP approach prescribes not to continue the relationship with a customer whose respective history has been shown; a value of 1 indicates continuation. Out of all possible cases, we just present cases whose simple prediction of the future cash flow of customer relationship result in a negative value indicating no profitability during the rest of the relationship.

Among 29 table rows, there exist numerous cases which show the different decision when one follows the optimal procedure in contrast to the simple prediction of $CRLV$ made based on the equation (3) using updated probability distribution at the end of period 10. It can be obviously inferred that when we do not take account of the future potential of making a purchase, the prediction reaches a remarkable underestimation of the value provided by the optimal value of the

stochastic dynamic programming problem. In the table, under “Optimal policy” column, 1 means to continue and 0 means to terminate the relationship.

5. Conclusions

The necessity of using methodological approaches for seller-initiated ending of customer relationship has been highly stressed by several authors [7]. We have considered a particular model of customer behavior called BG/BB model and investigated how one could obtain a biased prediction of customer value over his remaining lifetime when customer potential to make any additional transaction in the future is not taken into account. Sometimes, when traditional prediction models of customer lifetime value are used, the firm’s choice in terminating relationship with unprofitable customers is neglected. More precisely, this shows up all static models which do not consider the effect of the firm’s future marketing interventions on customer behavior. After applying a simple prediction model of customer lifetime value on the basis of the prior information and observations provided by the history of customer purchase behavior, we used a Markov decision process to accommodate the firm’s actions to the dynamics of customer behavioral state. A numerical evaluation showed that the result of the dynamic programming approach may considerably differ from that of the simple prediction model. As a result, besides being able to define the best occasion for termination of customer relationship, we can determine an accurate prediction of total value over the rest of customer’s lifetime.

There are a few limitations in our study which deserve more attention. First, as we only made a comparison between the MDP approach and no-termination strategy, a more thorough examination of some heuristic termination should provide more evidence on the capability of the MDP approach. Besides, we merely considered the BG/BB model as our basic model of customer behavior. However, more complex models of customer behavior, in which dynamics of customer behavior is explicitly taken into account, may be more relevant to real-world cases.

Finally, the problem investigated here may be extended to incorporate other marketing decisions such as promotion, advertisement, or pricing into a decision support system for the management of firm’s relationship with customers.

References

- [1] Ching, W.-K., Ng, M.K., Wong, K.-K. and Altman, E. (2004), Customer lifetime value: stochastic optimization approach, *Journal of the Operational Research Society*, 55, 860–868.
- [2] Fader, P.S. and Hardie, B.G.S. (2009), Probability models for customer-base analysis, *Journal of Interactive Marketing*, 23, 61–69.

- [3] Fader, P.S., Hardie, B.G.S. and Berger, P.D. (2004), Customer-Base Analysis with Discrete Time Transaction Data, Working Paper in Wharton School of the University of Pennsylvania.
- [4] Fader, P.S., Hardie, B.G.S. and Shang, J. (2010), Customer-base analysis in a discrete-time non-contractual setting, *Marketing Science*, 29(6), 1086-1108.
- [5] Haenlein, M., Kaplan, A.M. and Schoder, D. (2006), Valuing the real option of abandoning unprofitable customers when calculating customer lifetime value, *Journal of Marketing*, 70, 5–20.
- [6] Helm, S., Rolfes, L. and Gunter, B. (2006), Suppliers' willingness to end unprofitable customer relationships, an exploratory investigation in the German mechanical engineering sector, *European Journal of Marketing*, 40(3/4), 366-383.
- [7] Holmlund, M. and Hobbs, P. (2009), Seller-initiated relationship ending, an empirical study of professional business-to-business services, *Managing Service Quality*, 19(3), 266-285.
- [8] Kim, E. and Lee, B. (2009), Strategic use of analytical CRM in a market with network effects and switching costs: terminating unprofitable customer relationships, *Journal of Organizational Computing and Electronic Commerce*, 19(3), 153–172.
- [9] Pfeifer, E.P. and Carraway, R.L. (2000), Modeling customer relationships as Markov chains, *Journal of Interactive Marketing*, 14(2), 43-55.
- [10] Puterman, M.L. (2005), Markov Decision Processes and Discrete Stochastic Dynamic Programming, John Wiley & Sons, Inc., Hoboken, New Jersey.
- [11] Reinartz, W., Krafft, M. and Hoyer, W.D. (2004), The customer relationship management process: its measurement and impact on performance, *Journal of Marketing Research*, 41(3), 293-305.
- [12] Rust, R.T. and Chung, T.S. (2006), Marketing models of service and relationships, *Marketing Science*, 25(6), 560–580.
- [13] Sharma, A. (2008), Improving customer service and profitability through customer intervention in service relationships, *Journal of Relationship Marketing*, 7(4), 327-340.
- [14] Sun, B., Li, S. and Zhou, C. (2006), Adaptive learning and proactive customer relationship management, *Journal of Interactive Marketing*, 20(3/4), 82-96.
- [15] Winston, W.L. (1995), Operations Research: Applications and Algorithms, Third Edition, Belmont, Calif., Duxbury Press.