Action Selection in Customer Value Optimization: An Approach Based on Covariate-Dependent Markov Decision Processes

Angi Rösch and Harald Schmidbauer

***Abstract*— Typical methods in CRM marketing include action selection on the basis of Markov Decision Processes with fixed transition probabilities on the one hand, and scoring customers separately in pre-defined segments on the other. This points to a gap in the usual methodology insofar as customer scoring implies the explicit use of customer-specific information (covariates), while transition probabilities of Markov chains are conceived of as averages, without reference to the peculiarities of the customer to be addressed. Trying to unite both approaches, we suggest a model for customer transitions which allows transition probabilities to depend on covariates. Our model can be seen as an effort to focus on one-to-one marketing methods, permitting customer-specific action selection with the overall goal of customer value optimization. We show how to maximize the objective function subject to budget constraints. Our approach is motivated by the needs of a major European insurer. A numerical example with a realistic structure illustrates the capabilities of our approach.**

I. Introduction

A. The problem setting

Marketing in CRM typically involves different forms of customer contacts and marketing interventions, e.g. catalogs, simple or sophisticated mailing, product- or relationship-oriented actions. Traditional segmental marketing approaches deal with the allocation of marketing efforts to different segments of customers, where segments are built on customers' characteristics, past customer relationship, or heterogeneity with respect to response (e.g. [3]; an overview can be found in [7]). On the other hand, a "one-to-one marketing" has been promoted as the ultimate form of CRM ([4]), demanding different treatment for different customers. Personalized marketing interventions could explicitly take into account personal responsiveness. In ([7]), it is shown that the heterogeneity of response can be partially explained by customer's characteristics and past behavior. In order to optimize next-period change in profitability at the customer level, the authors propose a hierarchical model for the shift in gross profit due to marketing allocations to the customer, introducing a customer-specific response parameter vector which is estimated from a linear model with covariates.

This research project was presented on the 2009 International Conference on Data Mining DMIN'09, held in Las Vegas, July 13-16.

The paper is published in: Robert Stahlbock, Sven F. Crone, Stefan Lessmann (eds.), *Proceedings of the 2009 International Conference on Data Mining DMIN'09*, Las Vegas, July 13-16. CSREA Press, USA, 189-192. ISBN: 1-60132-099-X.

Angi Rösch is with the FOM University of Applied Sciences, Study Centres Munich, Germany, and Taian, China (email: angi@angi-stat.com).

Harald Schmidbauer is with the Department of Business Administration, Bilgi University, Istanbul, Turkey (email: harald@bilgi.edu.tr).

In recent years, Markov chain approaches have increasingly gained popularity in CRM marketing, introduced by ([5]). They can accommodate situations of both retention and migration of customers in a probabilistic way, and enable future prospect for customer relationship in terms of expected customer value and customer equity. The idea of a Markov chain approach is to model the route of a typical customer across customer segments (states) from one time period to the next as governed by transition probabilities. Probabilities are supposed to depend on the current state of the relationship only (this property is called the Markov property). They can be estimated from transactional data. Marketing-action-specific data can then be used to model the intermediate- and long-term impacts of marketing in Markov Decision Processes and to find the optimal allocation of marketing interventions to customer segments ([8], [3]; [1]).

B. Our Objective

In this paper we present an approach for marketing action selection which is optimal with respect to expected intermediate- or long-term customer value at the customer level rather than for customer segments. To this goal we employ a covariate-dependent Markov Decision Process. It incorporates transition probabilities that are based on the customer's characteristics and past behavior, in particular being appropriate to shed light on the customer's inclination to respond to different types of mailing contact or at least implying a moderating effect on it. Transition probabilities are estimated using multinomial logistic regression models. As we have in mind an application to a major European insurance company, we estimate the risk of loss which diminishes customer value.

C. Available Data

Apart from transactional customer data, our present study requires realistic company data on demographic and past behavioral customer characteristics, e.g. age, relationship duration, type and number of contracts, loss events, form and number of mailing contacts. Microgeographic data on living, social and economic environment add to the set of covariate data.

D. Outlook

This paper is organized as follows. Section II introduces the model which we use in our study. How we proceed to optimize personalized customer values with respect to action selection and empirical findings are provided in Sections III and IV. Finally, Section V gives a brief discussion of our approach. — All computations were carried out in R [6].

II. THE MODEL

We relate the observed behavior of a customer to the outcome of a stochastic model which is governed by latent behavioral attitudes and marketing actions. We suppose that each time epoch t from a finite set $\mathcal{T} = \{0, 1, 2, \ldots, T-1\}$ is a decision epoch for both the customer and the marketer on how to proceed until $t+1$. Thereon, we define a Markov Decision Process (MDP) featuring the following:

A. States

Customers are classified according to states which reflect the hierarchy of potential needs, e.g. standard customer, standard customer plus simple product, and plus comprehensive product respectively. By introducing an additional state for new or former customers, we can take into account situations of both retention and migration of customers.

Then, at time epoch t, a customer may decide to move from state s to s' from a finite set $\mathcal{S}$ of n states. Her decision may be driven not only by personal preferences but also by the marketing action she experienced at that time.

B. Actions

At each time epoch t, the marketer decides about which actions to applied to each of her customers. Action $a \in \mathcal{A}$ is selected according to its desirability, which depends on the targeted customer state. The action set $\mathcal{A}$ comprises three basic categories: no action, simple mailing, sophisticated mailing.

C. Transition probabilities

Suppose that customer k is in state s at time epoch t. Then, the probability that she will switch to state s' at time epoch $t+1$ under the regime of marketing action a is denoted by

$$p_{t,k}(s, s'|a, X_{t,k,s,a}).$$

The characteristic feature of our approach is that we allow transition probabilities to depend on customer-specific covariates $X_{t,k,s,a}$, specifying the covariates of a customer k who is exposed to action a when being in state s at time t, to switch to s' until time $t+1$. Estimates of transition probabilities are obtained using a multinomial logistic regression model with mean function

$$p_{t,k}(s, s'|a, X_{t,k,s,a}) = \frac{\exp(X_{t,k,s,a} \cdot \beta)}{1 + \exp(X_{t,k,s,a} \cdot \beta)} \quad (1)$$

Then, for a given sequence of actions a, the sequence of states which customer k visits within the time-horizon $\mathcal{T}$ is the realization of a Markov Decision Process with state space $\mathcal{S}$, action space $\mathcal{A}^n$ and transition probabilities $p_{t,k}(s, s'|a, X_{t,k,s,a})$.

D. Reward function

The magnitude of reward obtained from the application of action a to a customer depends on the target state, and may as well be affected by customer attributes and the outgoing state, e.g. elder customers or customers in certain states receive special offers. The expected reward by customer k moving from state s to s' under the regime of action a at time epoch t is denoted by

$$R_{t,k}(s, s'|a).$$

This may include expected damages produced by an insurant when applying the model in the context of insurance industry.

E. Costs

The expected reward generated by marketing intervention will be diminished by costs. Marketing costs are action-specific, as well in the sense that sophisticated mailings use to be more costly than simple mailings, while no action causes no costs. Let

$$c_t(a)$$

be the costs of action a applied to a customer which is canvassed at time epoch t.

F. Customer value generated by actions

Application of action a generates the following expected value of customer k which was in state s at time epoch t:

$$CV_{t,k}(s|a) = \sum_{s'} p_{t,k}(s, s'|a, X_{t,k,s,a}) R_{t,k}(s, s'|a) - c_t(a)$$

G. Objective function

Let $\alpha = (a_{t,k})_k$, for each decision epoch t, define a mapping from customers to actions. Then, the intermediate-term expected value of customer k which is in state s_t at time epoch t, given a policy α and a finite horizon of length T, is defined as

$$CV^{\alpha}_{t,k}(s_t) = \sum_{t'=t}^{T-1} CV_{t',k}(s'_t|a_{t,k}). \quad (2)$$

The optimal policy α is defined as the policy maximizing the intermediate-term expected value (2) for each customer.

III. OPTIMIZATION STRATEGY

A. Estimation of the MDP

Using past customer data concerning transition behavior and covariates, we use multinomial logistic regression models to estimate transition probabilities from one state to another in a Markov Decision Process.

B. Constraint optimization

Let marketing costs be given for each action and customer, and rewards in case of success. The intermediate-term values of customers, given a policy α will be optimized under the following constraints:

- The total marketing costs at a given time epoch t must not exceed a given budget C:

$$\sum_k c_t(a_{t,k}) \leq C$$

- Initial condition on customers: The number of customers in state s at time epoch 0 equals:

$$\sum_a \mathcal{N}_0(s|a),$$

where $\mathcal{N}_t(s|a)$ denotes the number of customers in state s at time epoch t experiencing action a.

- The total number of customers across all states is constant over time, i.e. the total number of customers moving to state s' generated by some action a at time epoch t equals the number of customers in state s' at time epoch $t+1$:

$$\sum_{a,s} \sum_{k \in \mathcal{N}_t(s|a)} p_{t,k}(s, s'|a) = \sum_a \mathcal{N}_{t+1}(s'|a).$$

IV. A Numerical Illustration

To illustrate our model, we give a fairly simple but realistic example and compare our approach to a standard approach, which neglects some of the information available. We assume that each customer is in one of three categories (states) at a given time. The state-dependent rewards (i.e. the revenue the customer contributes during one period) are given by

$$R(1) = 10, \quad R(2) = 20, \quad R(3) = 30.$$

Customers now in state s will be in state s' in the next period with a transition probability which depends on a customer-specific covariate x_k, which can be thought of as summarizing customer information, for example the score of a principal component analysis.

The transition probabilities for a customer k with covariate x_k are assumed to satisfy

$$\ln \frac{p_k(s, s'|a, x_k)}{p_k(s, 1|a, x_k)} = \beta_{0ss'} + \beta_{1ss'} x_k + \beta_{2ss'} a \tag{3}$$

for $s = 1, 2, 3$ and $s' = 2, 3$. (This is a multinomial logistic regression model with state 1 as baseline category.) Here, a indicates the action:

$$a = \begin{cases} 1 & \text{if customer } k \text{ is selected for mailing,} \\ 0 & \text{otherwise.} \end{cases}$$

(For simplicity, we only allow a dichotomous action in this illustration.) This specification leads to probabilities as given in equation (1). For a numerical illustration, the parameter values are:

$$\begin{array}{rclrclrcl}
\beta_{012} &=& -2, & \beta_{112} &=& 0.6, & \beta_{212} &=& 1.5, \\
\beta_{013} &=& -3, & \beta_{113} &=& 1.0, & \beta_{213} &=& 1.5, \\
\beta_{022} &=& 1, & \beta_{122} &=& -0.2, & \beta_{222} &=& 2.0, \\
\beta_{023} &=& -1, & \beta_{123} &=& 0.8, & \beta_{223} &=& 2.0, \\
\beta_{032} &=& 0, & \beta_{132} &=& 0.5, & \beta_{232} &=& 1.5, \\
\beta_{033} &=& 0, & \beta_{133} &=& 2.4, & \beta_{233} &=& 1.5.
\end{array}$$

(All these parameters can be estimated from customer-specific data in real-world applications.) It is desirable to make customers move to state 3, where they contribute the highest amount to total revenue. The transition probability to a higher category can be increased by setting $a = 1$, i.e. by applying a marketing measure (mailing). This will incur a cost of $c = 3$ for each customer selected for mailing. The goal is to select customers for mailing such that the next periods's expected total revenue (from all customers combined) is maximized.

Our example assumes that there are initially 10000 customers in each category. Their covariates were taken as simulated values from a standard normal distribution. The idea to solve this optimization problem is as follows. To begin with, assume that no marketing measure is applied at all. The 30000 customers will then create a certain expected total revenue. The first customer among the 30000 to be selected for mailing will be the one for whom the expected benefit of the action, that is: expected revenue with mailing, minus the sum of expected revenue without mailing and the cost of mailing, is the largest. This selection procedure can then be continued until further mailing is not meaningful anymore (because the marginal benefit for another mail is negative) or the marketing budget is exhausted.

How does the model outlined in this paper compare to a standard approach of action selection in this area? A typical model in customer value optimization practice uses fixed transition probabilities and separate customer scoring. For a comparison with our model, we first calculated from our dataset the matrix of average transition probabilities $p(s, s')$ between states, as well as expected customer values $CV(s)$ with respect to initial state s, generated without action, which define the baseline setting for estimating the benefit of mailing actions by customer scoring:

$$\begin{pmatrix} p(11) & p(12) & p(13) \\ p(21) & p(22) & p(23) \\ p(31) & p(32) & p(33) \end{pmatrix} = \begin{pmatrix} 0.819 & 0.123 & 0.058 \\ 0.235 & 0.644 & 0.121 \\ 0.338 & 0.265 & 0.398 \end{pmatrix}$$

$$\begin{pmatrix} CV(1) \\ CV(2) \\ CV(3) \end{pmatrix} = \begin{pmatrix} 12.38 \\ 18.86 \\ 20.60 \end{pmatrix}$$

We fitted a logistic regression model to mailing simulation data of next states. The target variable was defined to be 1 if a customer had switched from state 1 to 2, from 2 to 3, or had stayed with state 3, and it was set to 0 in all other cases. The values of the covariate were the same as those used in our model, but we added the initial state level information. As usual in customer scoring, the score values are interpreted as probabilities of switching to the target state.

For each customer k, the expected customer value conditional on selection for mailing $CV_k(s|a)$ was defined as score probability times reward from the target state plus converse probability times weighted reward from other states, less mailing costs. Customers are selected according to magnitude of benefit, which was calculated as $CV_k(s|a)$ compared to baseline expected customer value $CV(s)$.

Figure 1 shows the total expected revenues dependent on the number of mails — for the typical model (dotted line) in comparison with our model (solid line). Both lines reach a peak at about the same number of mails, that is, both models agree with respect to the optimal number of mails. However, our model outperforms the standard approach, which does not use state-specific transition probabilities based on a Markov Decision Process.

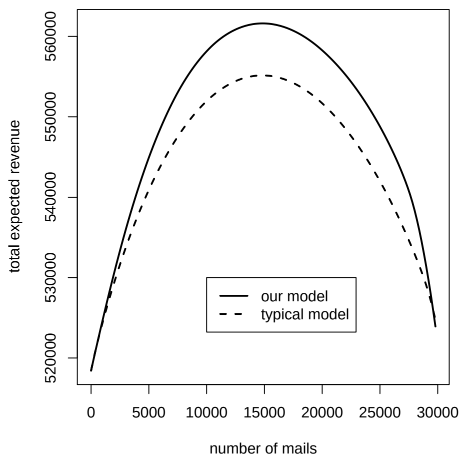

Fig. 1. Total expected revenue by number of mails: two models

V. Discussion

The transition probabilities in our model depend on covariates. Including covariates permits us to retain a simple state-based structure of the model while incorporating a potentially vast amount of information on each customer. This permits customer-specific action selection to optimize overall expected customer value. A numerical illustration with a realistic structure shows that this model can lead to higher expected revenues than the traditional approach, which separates the analysis of customer transition behavior and customer scoring.

References

[1] W.-K. Ching, M. K. Ng, *Markov Chains. Models Algorithms and Applications*, New York: Springer, 2006.

[2] S. Gupta, D. Hanssens, B. Hardie, W. Kahn, V. Kumar, N. Lin, N. Ravishanker, S. Sriram, "Modeling customer lifetime value," it IBM J. Res. & Dev., vol. 51, pp. 421–431, May/July 2007.

[3] A. Labbi, C. Berrospi, "Optimizing marketing planning and budgeting using Markov decision processes: An airline case study," *Journal of Interactive Marketing*, vol. 14, pp. 43–55, Spring 2000.

[4] D. Peppers, M. Rogers, *The One-to-One Future Revisited*, Norwalk, CT: Peppers & Rogers Group, March 2004.

[5] P. E. Pfeifer, R. L. Carraway, "Modeling customer relationships as Markov Chains," *Journal of Interactive Marketing*, vol. 14, pp. 43–55, Spring 2000.

[6] R Development Core team, *R: A language and environment for statistical computing*, R foundation for statistical computing, Vienna: 2009. URL http://www.R-project.org.

[7] R. T. Rust, P. C. Verhoef, "Optimizing the marketing interventions mix in intermediate-term CRM," *Marketing Science*, vol. 24, pp. 477–489, Summer 2005.

[8] G. Tirenni, A. Labbi, A. Elisseeff, C. Berrospi, "Efficient allocation of marketing resources using dynamic programming," *Proc. SIAM International Conference on Data Mining*, Newport Beach, CA, 2005, pp. 581–585.