European Journal of Operational Research 253 (2016) 514–523

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

A nonhomogeneous hidden Markov model of response dynamics and mailing optimization in direct marketing

Shaohui Ma*, Lu Hou, Wensong Yao, Baozhen Lee

School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang 212003, China

ARTICLE INFO

Article history:
Received 29 October 2013
Accepted 29 February 2016
Available online 8 March 2016

Keywords:
OR in marketing
HMM
POMDP
Customer lifetime value
Mailing optimization

ABSTRACT

Catalog firms mail billions of catalogs each year. To stay competitive, catalog managers need to maximize the return on these mailings by deciding who should receive a mail-order catalog. In this paper, we propose a two-step approach that allows firms to address the dynamic implications of mailing decisions, and to make efficient mailing decisions by maximizing the long-term value generated by customers. Specifically, we first propose a nonhomogeneous hidden Markov model (HMM) to capture the interactive dynamics between customers and mailings. In the second step, we use the parameters obtained from the HMM to determine the optimal mailing decisions using the Partial Observable Markov Decision Process (POMDP). Both the immediate and the long-term effects of mailings are accounted for. The mailing endogeneity that may result in biased parameter estimates is also corrected. We conduct an empirical study using six years of quarterly solicitation data derived from the well-known DMEF donation data set. All metrics used suggest that the proposed model fits the data well in terms of correct predictions and outperforms all other benchmark models. The simulative experimental results show that the proposed method for optimizing total accrued benefits outperforms the usual targeted-marketing methodology for optimizing each promotion in isolation. We also find that the sequential targeting rules acquired by our proposed methods are more cost-containment oriented in nature compared with the corresponding single-event targeting rules.

1. Introduction

Catalog firms mail billions of catalogs each year. To stay competitive, catalog managers need to maximize the return on these mailings by deciding who should receive a mail-order catalog. Traditional methods only attempt to maximize the benefit (equivalently, minimize the cost) of a single period mailing decision, and customers are typically selected for promotional mailings on the basis of the profits or revenues they are expected to generate on each mailing when viewed in isolation (Bult & Wansbeek, 2005). However, making a selection for only one period neglects the dynamics in the response to a mailing. For example, the expected profits obtained by mailing the current promotion to a certain customer might exceed the current cost of the mailing and might increase the profits generated by that customer in future mailings. More generally, marketing actions that are desirable from the perspective of maximizing customer value over time may sacrifice immediate rewards in anticipation of larger future revenues (DeSarbo & Ramaswamy, 1994).

We propose a two-step approach that allows firms to address the dynamic implications of mailing decisions and to make efficient mailing decisions by maximizing the long-term value generated by customers. Specifically, we first use a non-homogenous hidden Markov model (HMM) to capture the interactive dynamics between customers and mailings; the mailing endogeneity that may result in biased parameter estimates is corrected. Then, in the second step, to optimize the mailing decisions, we apply the Partial Observable Markov Decision Process (POMDP) to dynamically determine mailing numbers across customers and to maximize customer lifetime value (CLV). Seasonality is also considered in both HMM and the succeeding optimization model. In the proposed modeling framework, the customer is assumed to be in unobservable and unknown behavior states at any given point in time. These states represent a customer's latent psychological preference on keeping or ceasing doing business with the company. When the company sends mailings, the customer then makes a probabilistic transition to another state, possibly generating a reward by responding to the mailings. This process continues throughout the life of the customer's relationship with the company. By incorporating mailing intervention data into the procedure for estimating the transition matrices, one can obtain direct assessments of an intervention's effectiveness.

* Corresponding author. Tel.: +86 138 15179032.
E-mail address: msh@tju.edu.cn, shaohui.ma@hotmail.com (S. Ma).

http://dx.doi.org/10.1016/j.ejor.2016.02.055

To gain evidence for this concept, we test our approach on a donation data set from the Direct Marketing Educational Foundation (DMEF). The results of our experiments show that the proposed model fit the data well in terms of correct predictions and outperformed all other benchmark models. In terms of the cumulative profits obtained, our approach outperformed repeated applications of single-event targeting rules. We also observe that the sequential targeting rules acquired by our proposed methods were more effective with cost-containment when compared with the corresponding single-event targeting rules.

The remainder of this paper is organized as follows. Section 2 introduces existing studies on mailing optimization. Section 3 describes the proposed model and the estimation approach for mailing optimization. Section 4 provides an empirical study of the proposed modeling framework, including a summary of the parameter estimates and predictive validation, a discussion of the effects of the mailings, other covariates in terms of the transition probabilities of the HMM, and the results of the mailing optimization simulative experiments. Lastly, Section 5 provides a conclusion.

2. Literatures

Traditional methods for mailing decisions are based on response models that calculate the response probability for every customer in the database. The goal of response modeling is to identify customers who are likely to purchase a product on the basis of their purchase history and other information (Kang, Cho, & MacLachlan, 2012). Based on model predictions, catalog firms attempt to induce higher potential buyers to purchase the campaigned product using their mailed catalog. A large number of studies have been conducted with the objective of increasing response rate by improving the prediction algorithms used in response modeling (see Schröder and Hruschka, 2012 for a comprehensive review). Logistic regression has been widely employed as a base model because of its simplicity and availability (Kang et al., 2012). In addition to logistic regression, independent component analysis (Ahn, Choi, & Han, 2007), artificial neural networks (Baesens, Viaene, Van den Poel, Vanthienen, & Dedene, 2002; Kaefer, Heilman, & Ramenofsky, 2005), novelty detection (Lee & Cho, 2007), support vector machines (Shin & Cho, 2006), and decision trees (Coenen, Swinnen, Vanhoof, & Wets, 2000) were proposed from pattern recognition and data mining researchers.

In traditional methods, customers are selected for promotional mailings on the basis of the profits or revenues they are expected to generate on each mailing when viewed in isolation (Bose & Chen, 2009; Bult & Wansbeek, 2005). To maximize the expected profits for a given promotion, mailings should be sent only to customers whose predicted expected profits is positive when considering mailing costs. However, making a selection for only one period neglects the dynamics associated with the response to a mailing (Abe et al., 2002). For example, the expected profits obtained by mailing the current promotion to a certain customer might exceed the current cost of the mailing and might increase the profits generated by that customer in future mailings. More generally, marketing actions that are desirable from the perspective of maximizing customer value over time may sacrifice immediate rewards in anticipation of larger future revenues (DeSarbo & Ramaswamy, 1994).

A variety of optimal mailing models that address multiple solicitation problems have been proposed, beginning with the pioneering work of Bitran and Mondschein (1996) and continuing with important contributions by Elsner, Krafft, and Huchzermeier (2003, 2004), Gönül and Shi (1998), Pednault, Abe, and Zadrozny (2002), Piersma and Jonker (2004), Simester et al. (2006), and Neslin et al. (2013). All of these papers make significant contributions to a burgeoning literature on optimal mailing models. However, for most of these models, the definition of consumer behavior states is based on some variations of RFM segmentation. For example, Bitran and Mondschein (1996) and Gonul and Shi (1998) used Markov chain models to represent the evolution of customers' status characterized by their RFM values. Simester et al. (2006) proposed a classification and regression tree algorithm to segment the RFM state space and to determine whether to mail catalogs to all customers in a given state.

Despite RFM analysis often being used to identify customers to whom marketing should be targeted, this analysis usually simply shows who purchased from the company in the past. RFM states nothing about what is driving customers to make purchases and cannot be used to predict changes in customer behavior over time (e.g., seasonality). The method using RFM as representatives of customer relationships usually yields a large number of states and observations are often unevenly distributed across these states, making the state transition matrixes unreliably estimated. Additionally, customers are clustered into different groups assuming independent observations; transition matrices are then estimated assuming that each customer belongs to a cluster with probability 1. This procedure prevents uncertainty about cluster membership and transition probabilities from correctly propagating through the model.

We address the limitations by modeling the customer dynamics using a nonhomogeneous HMM. Because HMMs directly model the temporal aspect of the data, they can borrow strength across nearby observations when estimating model parameters and classifying observations to states. The Bayesian methods employed in this paper allow arbitrary functions of HMM parameters to be estimated when automatically accounting for parameter uncertainty. Our model can be applied to past, current, and new customers, and enables marketers to clearly identify the factors that drive spending decisions, thus enabling the marketer to identify new or up-and-coming markets. Our HMM can also be used to predict changes in customer behavior over time, thus enabling the marketer to respond quickly to an evolving market. Compared with the traditional Tobit response model (e.g., Van Diepen, Donkers, & Franses, 2009), HMM provides a more flexible modeling framework. The coefficients in the response function of a HMM are allowed to be state dependent.

Among the small but growing number of HMM applications in marketing (Kumar et al. 2011; Mark et al. 2013; Montoya, Netzer, & Jedidi, 2010; Netzer, Lattin, & Srinivasan, 2008; Romero, van der Lans, & Wierenga, 2013; Schweidel, Bradlow, & Fader, 2011), our approach is methodologically similar to Netzer et al. (2008) and Montoya et al. (2010). Netzer et al. (2008) was the first study to use the HMM framework to capture the dynamic interactions between marketing interventions and the state of customer relationships; however, their paper was not concerned with the problem of optimal marketing intervention. The key similarity of our study and Montoya et al. (2010) is that both studies further consider the optimal allocation of marketing resources. At the same time, some important differences exist. First, our application context differs from these two studies in that we look at the situation of a non-profit charity's optimal mailing decision. We consider the dynamic interactions between soliciting mailings and a donor's latent behavior state. All of the components of the HMM need to be designed according to this specific situation. For example, the state-dependent response functions in the HMM are deliberately designed for the sparse mailing responses using a binomial distribution model. We also need to consider seasonality in the model, which both Netzer et al. (2008) and Montoya et al. (2010) did not consider. Our substantive findings also enhance our understanding of the short-term and long-term effects of mailings in this research area. Second, we explicitly account for the endogeneity of mailings