Available online at www.sciencedirect.com

ScienceDirect

Journal of Interactive Marketing 29 (2015) 27 – 37

JOURNAL OF INTERACTIVE MARKETING

www.elsevier.com/locate/intmar

Targeting High Value Customers While Under Resource Constraint: Partial Order Constrained Optimization with Genetic Algorithm

Geng Cui [a,b,*] & Man Leung Wong [c] & Xiang Wan [d]

[a] Dept. of Marketing and International Business, Lingnan University, Tuen Mun, Hong Kong
[b] Yunshan Scholar, School of Management, Guangdong University of Foreign Studies, Guangzhou, Guangdong Province 510006, China
[c] Dept. of Computing & Decision Sciences, Lingnan University, Tuen Mun, Hong Kong
[d] Department of Computer Science, Baptist University, Kowloon, Hong Kong

Available online 11 February 2015

Abstract

To maximize sales or profit given a fixed budget, direct marketing targets a preset top percentage of consumers who are the most likely to respond and purchase a greater amount. Existing forecasting models, however, largely ignore the resource constraint and render sup-optimal performance in maximizing profit given the budget constraint. This study proposes a model of partial order constrained optimization (POCO) using a penalty weight that represents the marginal penalty for selecting one more customer. Genetic algorithms as a tool of stochastic optimization help to select models that maximize the total sales at the top deciles of a customer list. The results of cross-validation with a direct marketing dataset indicate that the POCO model outperforms the competing methods in maximizing sales under the resource constraint and has distinctive advantages in augmenting the profitability of direct marketing.

Keywords: Direct marketing; Profit maximization; Partial order function; Constrained optimization; Genetic algorithms; Return on investment

Introduction

Resource constraint such as a limited budget is prevalent in marketing operations. How to maximize sales or profit given a limited budget is a common challenge for marketing managers, making constrained optimization a research topic of intensive interest. This problem is especially relevant and acute in direct mail marketing (Gönül and Shi 1998), where managers usually have a fixed budget for a campaign, thus can only contact a pre-set percentage of customers, e.g., the top percentile of a customer list (e.g., top 10%). Conventional models of customer selection and forecasting, however, are mostly based on the ordinary least square (OLS) approach, and do not consider such constraint or focus on the high value customers. While researchers have attempted to arrive at better estimates of response probabilities and expected profit (e.g., Rao and Steckel 1995), most customer selection models do not address the problem of maximizing sales or profit given a specific mailing depth. Since customer response probabilities may have a low or even negative correlation with the expected purchase amount, models with a high accuracy of predicting customer response may not generate maximum sales. Although such models can be calibrated for a certain budget or mailing depth, directly incorporating the resource constraint is preferred and can generate superior result (Prinzie and Van den Poel 2005).

These models following the ordinary least squares (OLS) approach typically produce a complete order of all cases based on their expected sales or profit. However, a complete order of all cases is neither necessary nor does it guarantee superior result because it focuses on fitting a model to the overall population using the conditional mean to generate the expected profit for the “average” customer, while most of a company’s profit usually comes from a small group of high value customers, which are ignored as outliers. Although a few studies have investigated

* Corresponding author at: Dept. of Marketing and International Business Lingnan University 8 Castle Peak Road Tuen Mun, Hong Kong.
E-mail address: gcui@ln.edu.hk (G. Cui).

http://dx.doi.org/10.1016/j.intmar.2014.09.001
1094-9968/

the problem of resource constraint in direct marketing and adopted the constrained optimization approach (Bhattacharyya 1999; Yan and Baldasare 2006), they have adopted different dependent variables and used objective functions that are neither effective for profit maximization nor efficient for model optimization. Thus, the lack of a plausible objective function and an efficient and robust solution to optimize the selection of models presents a significant challenge.

We formulate the problem of targeting the high value customers given a fixed budget as one of constrained optimization using the number of customers to be selected as the constraint. For an optimal solution that satisfies the constraint while maximizing the total profit, we propose a model of partial order constrained optimization (POCO) that directly integrates the mailing depth (i.e., a certain percentile) as the constraint and adopt a penalty weight that represents the marginal penalty for selecting one more customer from a list. This allows us to convert the problem of complete order into a simpler partial order optimization problem with soft constraint. To ensure the efficiency and robustness of the solution, we adopt an advanced genetic algorithm (GA) with Cauchy mutation as a method of stochastic optimization to estimate the parameters and to select the models that maximize revenue at the top percentiles of a customer list. We apply the POCO model to a large direct marketing dataset and compare our results with those of competing methods at different constraint levels. The results indicate that the POCO model outperforms the nonparametric boosting method of AdaC2 and the constrained DMAX model (Bhattacharyya 1999), proves to be a viable solution to the problem of constrained optimization in direct marketing. The POCO model can help managers augment the return on marketing investment under resource constraint.

Relevant Literature

Customer Selection and Forecasting

Despite the technological advances, direct mail remains a popular marketing vehicle due to its ease of production, personal touch and quick response. The primary objective of modeling consumer responses in direct marketing is to identify customers who are the most likely to respond. This requires the researchers to produce a greater number of true positives in the upper deciles in the testing or rollout data. Direct marketing forecasting has focused on predicting consumer purchase probability using a variety of consumer behavior and background variables. Sophisticated models include the Customer Response-Based Iterative Segmentation Procedures (DeSarbo and Ramaswamy 1994), tree models such as CART and CHAID (Haughton and Oulabi 1997), the beta-logistic model (Rao and Steckel 1995), and the hierarchical Bayes random-effects model (Allenby, Leone, and Jen 1999). When comparing the alternatives, a better model should correctly classify a greater percentage of responders at the top deciles of a customer list. Since customers with higher probabilities to respond may have a lower expected purchase amount, models with a high accuracy of predicting customer responses may not generate maximum sales or profit.

Another stream of research aims at improving the profitability of direct marketing. Bult and Wansbeek (1995) propose a profit maximization approach to direct marketing — sending catalogs to customers as long as their expected profit is greater than the cost or the expected marginal profit is greater than zero. This profit maximization approach, which includes all the customers with a positive expected marginal profit, is not realistic given a fixed budget. Rao and Steckel (1995) propose a model that combines the predicted probabilities and the expected profit to re-rank the customers to float the high value customer to the top of a file to improve profitability. Meanwhile, researchers have adopted the lifetime value (LTV) approach to select high value customers based on their long-term contribution of profit to a company in the future (Fader, Hardie, and Lee 2005; Venkatesan and Kumar 2004).

Recently, researchers have adopted joint distribution models such as neural networks and Bayesian networks for customer selection and revenue forecasting in direct marketing (Bose and Chen 2009; Cui, Wong, and Lui 2006; Zahavi and Levin 1997). A number of scholars proposed various machine learning methods such as genetic algorithms and evolutionary programming to optimize the performance of customer selection models (Bhattacharyya 1999; Cui, Wong, and Lui 2006; Jonker, Piersma, and Van den Poel 2004). Cui, Wong, and Wan (2012) proposed cost-sensitive learning that places greater weight on the high value customers in the customer selection process. Miguéis, Benoit, and Van den Poel (2014) applied the Bayesian quantile regression for selecting customers. Despite the low cost of contacting customers, low response rate is common for direct marketing campaigns, for instance, around 5% for catalog mailings, making improving profitability and return on investment a top priority. Even a small percentage of improvement in predictive accuracy can result in a significant increment in profitability.

To date, customer selection models including the OLS regression and its variants (logit) have largely followed the principle of mean squared error (MSE) and focus on the mean estimate for the entire population (Hao and Naiman 2007). Under a typical condition, i.e., when the errors are assumed to have precisely the same distribution, OLS regression is sufficient to describe the relationship between the covariates and the response distribution (Fitzenberger, Koenker, and Machado 2002). In addition, conditional-mean models lead to estimators that possess attractive statistical properties, are easy to calculate, and straightforward to interpret (Hao and Naiman 2007). OLS is efficient and powerful for generating the predicted profit of customers and can "generally" distinguish the high value customers from low value customers.

However, direct marketing typically has a low response rate. With a 5% response rate, for instance, most customers (i.e., 95%) are non-buyers with a negative profit. Among the small portion of buyers (e.g., 5%), a majority of them (e.g., 4%) contributes very small profit while the minority (1% or less) contributes high profit. In general, customer profits exhibit along-tailed, skewed distributions with the high-value customers as the outliers as shown in Fig. 1. This has been referred to as the so-called 80/20 rule in that 80% of a company's profit often comes from 20% of its customers. Thus, firms mostly depend on a small set of