European Journal of Operational Research 210 (2011) 670–683

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Optimization models for targeted offers in direct marketing: Exact and heuristic algorithms

Fabrice Talla Nobibon *, Roel Leus, Frits C.R. Spieksma

Operations Research Group, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 29 September 2009
Accepted 9 October 2010
Available online 20 October 2010

Keywords:
Direct marketing campaign
Integer programming
Branch-and-price algorithm
Non-approximability
Heuristics
Tabu search

ABSTRACT

This paper presents an optimization model for the selection of sets of clients that will receive an offer for one or more products during a promotion campaign. We show that the problem is strongly NP-hard and that it is unlikely that a constant-factor approximation algorithm can be proposed for solving this problem. We propose an alternative set-covering formulation and develop a branch-and-price algorithm to solve it. We also describe eight heuristics to approximate an optimal solution, including a depth-first branch-and-price heuristic and a tabu-search algorithm. We perform extensive computational experiments both with the exact as well as with the heuristic algorithms. Based on our experiments, we suggest the use of optimal algorithms for small and medium-size instances, while heuristics (especially tabu search and branch-and-price-based routines) are preferable for large instances and when time is an important factor.

1. Introduction

In this paper, we examine the development of optimization models that decide on the details of a direct-marketing campaign that will make targeted offers to customers. Such campaigns are fundamental marketing tools for improving the economic profit of a firm, either by acquiring new customers or by generating additional revenue from existing customers [20]. The former action is called "acquisition" while the latter is "retention" [33]. In this paper, we are concerned with the latter case: campaigns that generate additional revenue by offering new products to existing customers. This study is justified by at least two practical facts. Reinartz et al. [33] point out that "When firms trade off between expenditures for acquisition and those for retention, a suboptimal allocation of retention expenditures will have a greater impact on long-term customer profitability than will suboptimal acquisition expenditures". Moreover, models and methods used for data analysis are more suited for retention [19] since more information is available. Retention boosts the customer lifetime value, which is defined by Kumar et al. [21] to be "the sum of cumulated cash flows – discounted using the weighted average cost of capital – of a customer over his or her entire lifetime with the firm". Customer lifetime value usually serves as a metric for a ranking or segmentation of the firm's customers [34]. During the last decades, the advances in data analysis coupled with the availability of customer data have pushed firms to develop a more customer-oriented strategy. Nowadays, such a strategy is globally accepted, but its practical implementation is far from being accomplished. This implementation delay is observed both in business-to-business and business-to-consumer settings, and is particularly pronounced in financial institutions such as large banks and insurance companies, which often have a large number of customers with full data available but may lack sophisticated tools that efficiently take into account these advantages in decision making [10]. Although retail banks have access to the deepest veins of customer data of any industry in the world [12], Bernstel [3] reveals that many banks are not using the full power of their customer databases.

Optimal product targeting models examine "Which products should be targeted to which customers to maximize profits, under the constraints that only a limited number can be targeted to each customer, and each product has a minimum sales target" [8,19]. Such problems are essentially characterized by two steps, which are "data analysis" and "problem formulation and solution". The first step, which is mainly statistical, has received increasing attention with the advances in data analysis, and is often referred to by the term "marketing response models", which are models intended to help scholars and managers understand how consumers

* Corresponding author. Tel.: +32 16326960.
E-mail addresses: Fabrice.TallaNobibon@econ.kuleuven.be (F. Talla Nobibon), Roel.Leus@econ.kuleuven.be (R. Leus), Frits.Spieksma@econ.kuleuven.be (F.C.R. Spieksma).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2010.10.019

respond to marketing activities [16]. In academia, these models were introduced in the early eighties (see, for instance, Guadagni and Little [14]), and in practice became popular in the nineties thanks to the scanner-data revolution. De Reyck and Degraeve [9] have developed a model in which a proxy is used for direct response rates, effectively reducing the effort spent on response modeling. Some improvements to this model were proposed by Tripathi and Nair [38]. Recently, Bhaskar et al. [4] have proposed a fuzzy mathematical-programming approach where fuzzy numbers are used to represent the output of the first step, allowing to take into account uncertainty. Their model, however, incorporates only a budget constraint and volume target constraints. An elaborate discussion on the inputs for direct-marketing models can be found in Bose and Chen [5].

This paper investigates the development of optimization models for targeted offers in a promotion campaign, based on integer programming. Motivation for studying this problem comes from a case occurring at FORTIS [17], until recently one of the leading banks in Belgium. More generally, Stone and Jacobs [36] point at direct marketing's recent growth in business categories such as banks, investment and insurance companies. We aim to maximize the profit subject to business constraints such as the campaign's return-on-investment hurdle rate that must be met (the hurdle rate is the minimum acceptable rate of return that management will accept for the campaign), a limitation on the funding available for each product, a restriction on the maximum number of possible offers to a client, and a minimum-quantity commitment (MQC) on the number of units of a product to be offered in order for that product to be part of the campaign.

The MQC constraint has been briefly mentioned by Cohen [8] as a technical issue for an application in a bank. However, he did not explicitly incorporate it into his model. Our model does take into account this constraint, making it an extension of the model used by Cohen. We also study the case where the budget of a product is managed by a business unit. These two constraints make our model substantially different from that proposed e.g., by Cohen. Contrary to Cohen, our model also considers a more general setting where a client can receive more than one offer. With respect to solution procedures, Cohen restricts himself to LP-based solutions for aggregated data, while our most important contributions are computational, in that we propose several exact and heuristic algorithms. In our formulation, we also impose a more general version of the MQC constraint, allowing the fixed minimum quantity to depend on the product, which distinguishes the constraint from comparable MQC restrictions studied in the analysis of transportation problems [25], bottleneck problems [26], and assignment problems [24]. Contrary to these references, our model also includes both the budget constraints and the hurdle-rate constraint, making it more difficult to solve.

In this paper, we present a basic integer-programming formulation for the problem at hand. We show the non-approximability of the problem, which makes the existence of an algorithm that will always provide a feasible solution and guarantee a specified proportion of optimal profit in polynomial time, highly unlikely. We next present a set-covering formulation and develop a branch-and-price algorithm for solving it. A dynamic-programming algorithm and a 2-approximation algorithm are presented for solving the pricing problem, which is closely related to the k-item knapsack problem. The size of instances that can be solved optimally using this algorithm allows its efficient use for business-to-business promotion campaigns (which have moderate size and high variable and fixed costs) and for sampling approaches in financial institutions [8]. We then present eight heuristics to approximate an optimal solution, which can be used for large instances and hence for business-to-consumer promotion campaigns. These heuristics are either variants of the algorithms used in practice for application in a bank (see [17,19]) or developed based on the structure of the problem.

The contributions of this paper are fourfold:

(1) The formulation of the product targeting problem as a mixed-integer programming (MIP) problem including more business constraints.
(2) An observation concerning the non-approximability of the problem.
(3) The reformulation of the problem as a set-covering model, which is solved via branch-and-price; a special cardinality-constrained knapsack problem is obtained as a pricing problem, which we solve by adapting well-known results concerning the knapsack problem.
(4) The development of efficient heuristics, including a depth-first branch-and-price heuristic and a tabu-search algorithm, that outperform methods used in practice.

This paper is organized as follows. Section 2 describes the basic integer-programming formulation for selecting clients that will receive targeted offers. The complexity of the problem makes it very difficult to produce optimal solutions by straightforward use of a MIP solver. We propose an alternative formulation, called the set-covering formulation, in Section 3 and develop a branch-and-price algorithm to solve it. In Section 4, we describe eight heuristics to approximate an optimal solution. Experimental results are presented in Section 5 both for the exact as well as for the heuristic algorithms. Finally, a number of possible extensions of the studied model are presented in Section 6, followed by a summary and conclusions in Section 7.

2. Basic formulation

The objective of a direct-marketing promotion campaign is to find a way to achieve a maximum profit by offering n different products to m customers while taking into account various business constraints. We incorporate the following restrictions: the return-on-investment hurdle rate must be met for the campaign, the budget allocated to each product is limited, an upper bound is imposed on the number of products that can be offered to each client and there is also an MQC constraint for each product. We define the parameter r_{ij} as the probability that client i reacts positively to an offer of product j (or the probability that product j is the next product bought by client i [19]) and DFV_{ij} as the return to the firm when client i responds positively to the offer of product j. The latter is termed the Delta Financial Value by FORTIS [17]. These two parameters are the basis for the computation of customer lifetime value [34]. Practically, these parameters are estimated using response models based on historical data [8,19,32] and are assumed to be available within the firm. We denote by p_{ij} the expected return to the firm (revenue) from the offer of product j to client i, so $p_{ij} = r_{ij}DFV_{ij}$. In the remainder of this paper, we will mostly use p_{ij}. Further, there is a variable cost c_{ij} associated with the offer of product j to client i, the upper bound M_i of offers that can be made to a client i (this quantity is related to the status of the client), the minimum-quantity commitment bound O_j associated with product j, the budget B_j allocated to the product j, a fixed cost f_j needed if product j is used for the campaign and finally the corporate hurdle rate R. The value of R is dependent on the firm and the riskiness of the investment. In practice, most firms use their weighted average cost of capital (WACC) as an estimation of R [6]. We define the decision variables $y_j \in \{0,1\}$, equal to 1 if product j is used during the campaign, 0 otherwise, and $x_{ij} \in \{0,1\}$, which is equal to 1 if product j is offered to client i and 0 otherwise. A basic formulation for the product targeting problem can be expressed as: