European Journal of Operational Research 252 (2016) 204–212

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

A branch-and-bound algorithm for the maximum capture problem with random utilities

Alexandre S. Freire[b], Eduardo Moreno[a,*], Wilfredo F. Yushimito[a]

[a] *Faculty of Engineering and Sciences, Universidad Adolfo Ibañez, Santiago, Chile*
[b] *Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, Brazil*

ARTICLE INFO

Article history:
Received 3 November 2014
Accepted 15 December 2015
Available online 22 December 2015

Keywords:
Facility location
Branch and bound
Maximum capture
Random utility model

ABSTRACT

The MAXIMUM CAPTURE PROBLEM WITH RANDOM UTILITIES seeks to locate new facilities in a competitive market such that the captured demand of users is maximized, assuming that each individual chooses among all available facilities according to the well-know a random utility model namely the multinomial logit. The problem is complex mostly due to its integer nonlinear objective function. Currently, the most efficient approaches deal with this complexity by either using a nonlinear programing solver or reformulating the problem into a Mixed-Integer Linear Programing (MILP) model. In this paper, we show how the best MILP reformulation available in the literature can be strengthened by using tighter coefficients in some inequalities. We also introduce a new branch-and-bound algorithm based on a greedy approach for solving a relaxation of the original problem. Extensive computational experiments are presented, benchmarking the proposed approach with other linear and non-linear relaxations of the problem. The computational experiments show that our proposed algorithm is competitive with all other methods as there is no method which outperforms the others in all instances. We also show a large-scale real instance of the problem, which comes from an application in park-and-ride facility location, where our proposed branch-and-bound algorithm was the most effective method for solving this type of problem.

1. Introduction

In recent years, competitive facility location models have received considerable attention both due to their interesting theoretical aspects and their practical applications. These models extend conventional facility location models to a more complex scenario, in which (a) companies compete for their market share and (b) the choices of independent decision makers, such as customers, are considered. As an example, we can think of a company that wants to locate r new supermarkets in a geographical zone where some supermarkets are already located (the competitors). The competitive facility location problem consists of choosing, from a given set of available locations, the locations for these r new facilities such that the demand captured by them (i.e. market share) is maximized.

This problem can be traced back to Hotelling's (1929) optimal location of two competing facilities on a line segment, and it was later embedded within the location theory, initially by Slater (1975) and further developed by Hakimi (1983).

In general, the literature considers that customers choose among different alternatives based on a given utility function that depends on a set of facility attributes (e.g., distance, transportation costs and waiting times, among others). The first deterministic model was proposed by ReVelle (1986), in which customers choose the closest facility among different competitors. However, these models imply an "all or nothing" assignment, in which the demand of a given point is assigned entirely to one facility. An alternative approach is proposed in the gravity-based model (Huff, 1964; Reilly, 1931), in which the demand captured by a facility is proportional to the "attractiveness" of the facility and inversely proportional to a power of the distance. Drezner and Eiselt (2002) and Berman, Drezner, Drezner, and Krass (2009) provide a comprehensive review of these different models.

Another alternative approach to the "all or nothing" assignment is to estimate the market share obtained by each facility through a random utility model. In random utility models (e.g., logit or probit models; see McFadden, 1973 or Ben-Akiva and Lerman, 1985), the utilities of economic agents are essentially derived from their preferences among a set of discrete options. In this case, the

* Corresponding author. Tel.: +56 223311351; fax: +56 223311351.
E-mail addresses: afreire@ime.usp.br (A.S. Freire), eduardo.moreno@uai.cl (E. Moreno), wilfredo.yushimito@uai.cl (W.F. Yushimito).

http://dx.doi.org/10.1016/j.ejor.2015.12.026
0377-2217/

problem can be stated as follows: given a set of customers and their respective demands, a set of open facilities (the competitors), and a set of available locations, the problem is to locate r new facilities such that the expected market share captured by the new facilities is maximized, where the market share captured by each selected facility is estimated through a random utility model (e.g. logit). This problem is referred to as the MAXIMUM CAPTURE PROBLEM WITH RANDOM UTILITIES (or MCRU problem, for short) and it was first introduced by Benati and Hansen (2002) with the multinomial logit model (MNL) as the underlying random utility model. Recent applications for this model includes locating schools (Haase & Müller, 2012), preventive healthcare facilities (Haase & Müller, 2015; Zhang, Berman, & Verter, 2012), and siting park-and-ride facilities (Aros-Vera, Marianov, & Mitchell, 2013).

Since the multinomial logit model is nonlinear by nature, modeling the MCRU problem usually results in nonlinear integer programing models, which in general are difficult to solve. Benati and Hansen (2002) have proposed different approaches to address the problem, namely concave programing, integer fractional programing and submodular maximization. The computational analysis presented in the cited paper shows that the concave programing, which is basically a branch-and-bound algorithm with a concave relaxation of the problem as dual bound, behaves better than the other two approaches.

Alternatively, equivalent Mixed-Integer Linear Programing (MILP) formulations have been proposed by Benati and Hansen (2002), Haase (2009), Aros-Vera et al. (2013), and Zhang et al. (2012). These formulations were recently evaluated by Haase and Müller (2014) to provide a computational comparison of them, being the model by Haase (2009) the most efficient in practice.

In this paper, we show how the MILP model introduced by Haase (2009) can be strengthened by using tighter coefficients in a class of inequalities. We also introduce a greedy algorithm for solving a relaxation of the MCRU problem, which is embedded into a branch-and-bound (B&B) algorithm to compute dual bounds. The success of a B&B algorithm relies basically on finding a good threshold between the quality of the bounds and the computational effort needed to calculate them. In fact, the obtained dual bounds are not necessarily sharper than the ones given by the known linear and nonlinear formulations of the problem, but they can be calculated much faster than the others. This allows to explore more nodes of the B&B tree, which in many cases is more effective than spending too much time computing better bounds. Moreover, the proposed B&B algorithm can be easily implemented and does not make use of any external solver, while all the other methods considered here do.

To evaluate the algorithm, extensive computational results are obtained for instances from three different datasets, namely the UflLib repository, the randomly generated instances introduced by Haase and Müller (2014), and a relatively large size instance (82341 customers and 59 available locations) that comes from a real application in location of park-and-ride facilities in New York City (Holguín-Veras, Reilly, Aros-Vera, Yushimito, & Isa, 2012). The methods considered here for comparison are the concave programing approach introduced by Benati and Hansen (2002), the MILP formulation introduced by Haase (2009) (using the tighter coefficients proposed in this work) and the proposed B&B algorithm. Results show that the proposed B&B algorithm is competitive with other available methods on all instances, and the most efficient method for solving the large real instance mentioned above.

The remainder of this paper is organized as follows. In Section 2, we present the notation and definitions used throughout this paper. In Section 3, we present some mathematical formulations for the MCRU problem found in the literature and show how the MILP model introduced by Haase (2009) can be strengthened by using tighter coefficients in a class of inequalities. In Section 4, we introduce a new B&B algorithm for the MCRU problem. The computational results are presented in Section 5. Finally, in Section 6 we draw some conclusions and present opportunities for future work.

2. Problem description

In this section, we give a formal description of the MCRU problem. Before describing the problem itself, we first explain the behavioral rationale underlying the customers' decisions, in which the market share captured by a particular facility is based on the preferences of the customers, which results in a choice probability of selecting a particular facility.

2.1. Behavioral rationale and choice probability

Let S be a set of customers and H be a set of open facilities. Each customer $s \in S$ receives a utility $\tilde{u}_{sl}$ for choosing the facility $l \in H$. Assuming that customers behave rationally, each customer selects the facility that provides the highest utility value. That is, a customer $s \in S$ chooses a facility $l \in H$ if $\tilde{u}_{sl} \geq \tilde{u}_{sh}, \forall h \in H$.

In random utility theory, the utility $\tilde{u}_{sl}$ obtained by customer $s \in S$ choosing a facility $l \in H$ has two components: a deterministic part v_{sl} and a random term ϵ_{sl}, such that $\tilde{u}_{sl} = v_{sl} + \epsilon_{sl}$. The deterministic part is typically referred to as the systematic component, because it is composed of a set of observable attributes (e.g., distance and time), whereas the random components represent the non-observable attributes. The joint density of the random vector $\epsilon_s = \{\epsilon_{s1}, \ldots, \epsilon_{sl}\}$, denoted by $f(\epsilon_s)$, allows us to state the probability of choosing an alternative. According to McFadden (1973), whenever the elements in ϵ_s are identically and independently distributed, they have equal variability among cases, and $f(\epsilon_s)$ follows a Generalized Extreme Value (GEV) distribution (i.e., Gumbel distribution), the model is referred to as the multinomial logit model, and the probability that a customer s selects a facility l from the given set H of open facilities is given by the following equation:

$$p_{sl} = \frac{e^{v_{sl}}}{\sum_{h \in H} e^{v_{sh}}} \tag{1}$$

2.2. Problem description

In the MCRU problem, it is given a set L of available locations, a set A of open facilities (the competitors) and a set S of customers. For simplicity, we sometimes refer to an available location l, where a new facility can be located, as a facility itself. For each customer $s \in S$, it is given a positive demand d_s and a utility v_{sl} for choosing the facility located in $l \in L \cup A$. The objective is to choose a subset $L^* \subset L$ of r locations where new facilities can be located, such that the expected demand captured by the new facilities is maximized, which is given by

$$\sum_{s \in S} \sum_{l \in L^*} d_s p_{sl} \tag{2}$$

According to (1), the probability that a user $s \in S$ chooses a facility $l \in L^*$ is given by

$$p_{sl} = \frac{e^{v_{sl}}}{\sum_{h \in L^* \cup A} e^{v_{sh}}} \tag{3}$$

Note that, w.l.o.g., we can assume that there is a single open facility (i.e., $|A| = 1$). If there is more than one open facility (i.e., $|A| > 1$), we can represent them as a single facility a such that $v_{sa} = \log(\sum_{i \in A} e^{v_{si}})$. Hence, for simplicity, we assume that there is a single open facility a. Note that higher values of v_{sa} represent a problem with stronger incumbent competitors.