Computers & Industrial Engineering 94 (2016) 93–104

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Combining statistical learning with metaheuristics for the Multi-Depot Vehicle Routing Problem with market segmentation

Laura Calvet [a,*], Albert Ferrer [b], M. Isabel Gomes [c], Angel A. Juan [a], David Masip [a]

[a] Computer Science Department, Open University of Catalonia - IN3, Castelldefels, Spain
[b] Department of Mathematics, Technical University of Catalonia, Barcelona, Spain
[c] Centro de Matemática e Aplicações, FCT, Universidade Nova de Lisboa, Lisbon, Portugal

ARTICLE INFO

Article history:
Received 25 July 2015
Received in revised form 20 January 2016
Accepted 21 January 2016
Available online 28 January 2016

Keywords:
Multi-Depot Vehicle Routing Problem
Market segmentation applications
Hybrid algorithms
Statistical learning

ABSTRACT

In real-life logistics and distribution activities it is usual to face situations in which the distribution of goods has to be made from multiple warehouses or depots to the final customers. This problem is known as the Multi-Depot Vehicle Routing Problem (MDVRP), and it typically includes two sequential and correlated stages: (a) the assignment map of customers to depots, and (b) the corresponding design of the distribution routes. Most of the existing work in the literature has focused on minimizing distance-based distribution costs while satisfying a number of capacity constraints. However, no attention has been given so far to potential variations in demands due to the fitness of the customer-depot mapping in the case of heterogeneous depots. In this paper, we consider this realistic version of the problem in which the depots are heterogeneous in terms of their commercial offer and customers show different willingness to consume depending on how well the assigned depot fits their preferences. Thus, we assume that different customer-depot assignment maps will lead to different customer-expenditure levels. As a consequence, market-segmentation strategies need to be considered in order to increase sales and total income while accounting for the distribution costs. To solve this extension of the MDVRP, we propose a hybrid approach that combines statistical learning techniques with a metaheuristic framework. First, a set of predictive models is generated from historical data. These statistical models allow estimating the demand of any customer depending on the assigned depot. Then, the estimated expenditure of each customer is included as part of an enriched objective function as a way to better guide the stochastic local search inside the metaheuristic framework. A set of computational experiments contribute to illustrate our approach and how the extended MDVRP considered here differs in terms of the proposed solutions from the traditional one.

1. Introduction

In the distribution business, whenever a supplier operates from multiple warehouses or depots it needs to decide two things: (a) which set of customers will be served from each depot, i.e., the customer-depot assignment map; and (b) the vehicle routing plan for the given assignment map. This two-stage decision-making process is called the Multi-Depot Vehicle Routing Problem (MDVRP). During the last decades, researchers have extensively addressed different variants of this problem, among others those including heterogeneous fleets of vehicles, multiple products, simultaneous pick-up and delivery, etc. (Caceres, Arias, Guimarans, Riera, & Juan, 2015; Montoya-Torres, Lopez, Nieto, Felizzola, & Herazo-Padilla, 2015). The large majority of models aim at minimizing total distribution costs, which are often modeled by means of a distance-based cost function. Minimization of distribution costs has a major impact on the efficiency of any competitive shipping company. However, following the trend to consider richer and more realistic Vehicle Routing Problems (Barbucha, 2014; Ehmke, Campbell, & Urban, 2015; Taş, Jabali, & Woensel, 2014), it should be noticed that these costs represent only half of the equation, i.e.: if a distribution company wants to maximize its benefits, it has also to account for the expected incomes associated with different customer-to-depot assignment plans. Thus, retail centers (depots) belonging to the same organization may offer different products, trade credit policies, or complementary services, which often have a non-negligible impact on the customer's willingness to buy. Accordingly, under the existence of a diversity of depots and commercial offers, the customer-to-depot

* Corresponding author.
E-mail addresses: lcalvetl@uoc.edu (L. Calvet), alberto.ferrer@upc.edu (A. Ferrer), mirg@fct.unl.pt (M.I. Gomes), ajuanp@uoc.edu (A.A. Juan), dmasipr@uoc.edu (D. Masip).

http://dx.doi.org/10.1016/j.cie.2016.01.016
0360-8352/

assignment process should not only consider distribution costs but also expected sales or total income.

In order to increase sales revenue, companies use market segmentation strategies that allow grouping customers according to their features (preferences, rent, age range, etc.). Ideally, each group has homogenous features that allow the development of tailored strategies and actions oriented to increase the customer's willingness to buy, i.e., the fitness between his/her utility function and the commercial offer he/she is receiving. In this paper we address an extended version of the MDVRP that also includes market segmentation issues in order to maximize benefits (sales revenue minus distribution costs). Thus, in our model customer-to-depot assignation decisions are taken considering not only the traditional distance-based cost but also other customers' features in an attempt to increase the expected expenditure by providing a more adequate assignation. As a consequence of this, the assignment and routing solutions might be very different from the ones associated with the classical MDVRP. For instance, Fig. 1 shows two different solutions, with the shape of each customer representing the shape of its best-fit depot. The one on the left only considers distribution costs (to be minimized), while the one on the right considers expected benefits (to be maximized), i.e.: not only distribution costs but also additional revenue due to a 'smarter' customer-to-depot assignment. Notice that in the right-hand solution each depot tends to deliver those customers that share a similar shape, unless they are too far away so that the increase in distribution costs overshadows the potential increase in revenue. In the illustrative example of Fig. 1, it is estimated that customer j will spend 20 monetary units when assigned to depot 2 (left-hand solution). On the other hand, if this same customer is assigned to depot 1 (right-hand solution), it is estimated that his/her willingness to spend will increase up to 30 monetary units. Therefore, assigning customer j to depot 1 instead of to its closest depot (depot 2) will pay off as far as the increase in transportation costs will not exceed the marginal income attained (10 monetary units in this case).

Our solving approach is based on the combination of statistical predictive models with a metaheuristic framework. In short, the algorithm develops in two main steps. Firstly, supported by the company historical data concerning existent customers, new customers are assigned to depots. This step is preceded by a historical data analysis so that expected expenditure from new customers among depots is estimated throughout a multiple regression model. The regression model will capture the relationship between each customer's willingness to spend (response) as a function of several variables (predictors), including: the assigned depot as well as other customer's features (e.g.: preferences, rent, sex, age, etc.). In the second step, the routes associated to each customer-to-depot assignment map are built. Given the interdependency between both decisions (assignation and routing), our procedure is an iterative one. Different assignations are generated together with the routing decisions and the top best solutions will be saved and locally improved in the last step of the algorithm. The main contributions of our work are: (i) the description of an extended version of the MDVRP with heterogeneous depots, which can be considered a rich routing problem, (ii) the development of a methodology combining statistical learning and a metaheuristic for solving it, and (iii) an analysis of how the solutions found for the extended problem differ from those for the classical one in terms of both expected benefits and distribution costs for a set of instances artificially generated.

The rest of the paper is organized as follows: Section 2 formally describes the well-known Multi-Depot Vehicle Routing Problem and presents the extended version with heterogeneous depots, while Section 3 reviews works addressing the classical version. Section 4 discusses the importance of considering market segmentation. Section 5 provides an overview on our solving approach, while Section 6 offers some low-level details. The computational experiments and a discussion of the results are presented in Section 7. Lastly, the main contributions of this work are highlighted in the Conclusion section.

2. Mathematical formulation for the Multi-Depot Vehicle Routing Problem

The MDVRP may be formally described as an extension of the Capacitated Vehicle Routing Problem (CVRP) and it is defined as a complete directed graph $G = (V, E)$, where $V = \{V_d, V_c\}$ is the set of nodes including the depots, V_d, and the customers, V_c, and E is the set of edges or arcs connecting all nodes in V. Each customer i in V_c has a positive demand to be satisfied, q_i. Each edge in E has an associated cost $c_{ij} > 0$ and distance $d_{ij} > 0$ between customers i and j. The distance matrix $D := [d_{ij}]$ and the cost matrix $C := [c_{ij}]$ are square matrices of order $|V|$. Usually, both matrices are assumed to be symmetric (nevertheless, our approach could also be applied even in the case of non-symmetric distances or costs).

For the MDVRP, a solution is a customer-to-depot assignment map together with a set of routes covering all customers' demands. Each route starts at one depot in V_d, connects one or more customers in V_c, and ends at the same depot, without exceeding the capacity of the vehicle. The number of vehicles based at each depot may be fixed or unlimited. The former defines a harder problem, since it adds an additional constraint and there is also no guarantee that a feasible solution exists (Chao, Golden, & Wasil, 1993). The latter simplifies the modeling and solving.

As mentioned before, when adopting a marketing perspective, companies focus on market segmentation to group customers

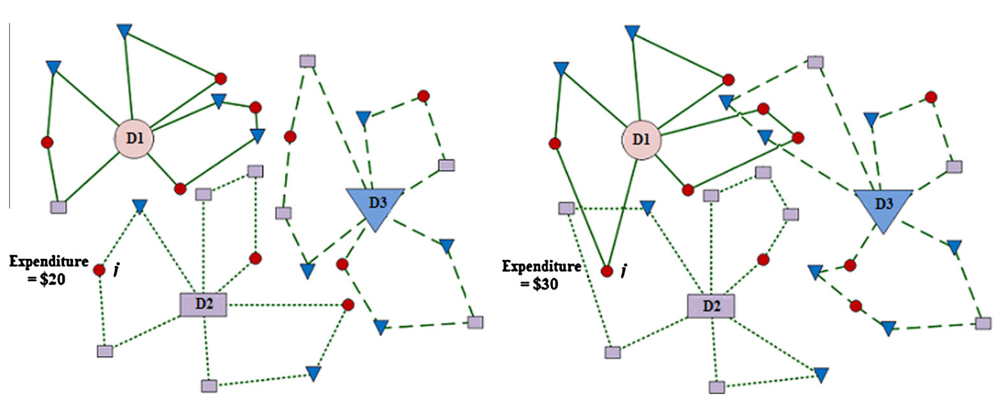

Fig. 1. Solutions for the classical MDVRP (left) and for the extended version (right).