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*

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.

© 2016 Elsevier Ltd. All rights reserved.

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

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 $c_{i,j} > 0$ and distance $d_{i,j} > 0$ between customers i and j. The distance matrix $D := [d_{i,j}]$ and the cost matrix $C := [c_{i,j}]$ 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
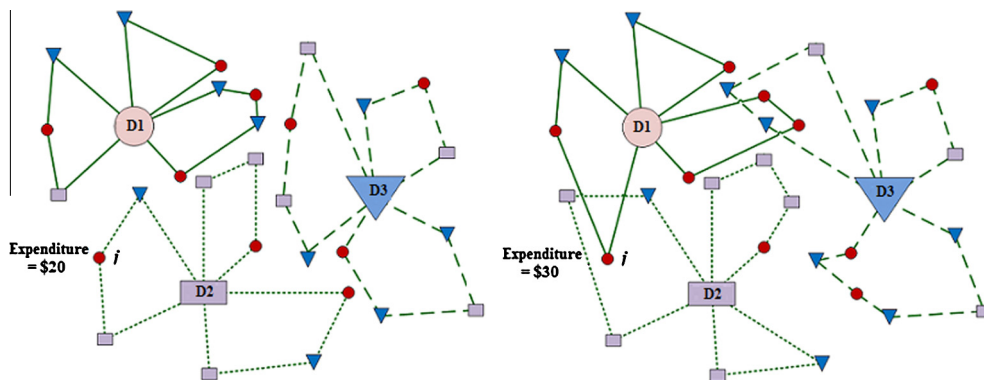

Fig. 1. Solutions for the classical MDVRP (left) and for the extended version (right).

according to their features and preferences. Considering the heterogeneity of markets, segmentation attempts to divide customers into subsets that behave in a similar way. Our extension of the MDVRP aims at assigning customers to depots based not only on distribution costs but also on customers' features and preferences. The goal is then to optimize expected benefits by considering both distribution costs and expected incomes.

To formally describe the mathematical model for the MDVRP with heterogeneous depots, we will first introduce a model for the CVRP problem, which is a particular case of the MDVRP when $|V_d| = 1$, i.e., $V_d = \{0\}$, and a model for the classical MDVRP.

2.1. Mathematical model for the MDVRP with one depot (CVRP)

In graph theory, a finite path, ϕ, of length r is a sequence of $r + 1$ vertices, $\{\alpha_0, \alpha_1, \ldots, \alpha_r\}$, together with a sequence of r arcs, $\{\phi^1, \phi^2, \ldots, \phi^r\}$, such that

$$\phi^k = (\alpha_{k-1}, \alpha_k), \quad k = 1, 2, \ldots, r.$$

Sometimes we will denote a finite path, ϕ, in the form:

$$\phi : \alpha_0 \to \alpha_1 \to \alpha_2 \to \cdots \to \alpha_{r-1} \to \alpha_r.$$

The vertex α_0 is called the start vertex and the vertex α_r is called the end vertex of the path. Both of them are called terminal vertices of the path. The other vertices in the path are internal vertices. A finite cycle is a path such that the start vertex and the end vertex are the same. Note that the choice of the start vertex in a cycle is arbitrary. A path with no repeated vertices is called a simple path, and a cycle with no repeated vertices or arcs aside from the necessary repetition of the start and the end vertex is a simple cycle.

Definition 2.1. In our context, a route, ρ, of order r is a simple finite cycle of length $r + 2$ in which the start vertex and the end vertex is the depot node 0,

$$\rho : 0 \to \alpha_1 \to \alpha_2 \to \cdots \to \alpha_{r-1} \to \alpha_r \to 0.$$

We denote, $\mathcal{R}$, the set of all routes of the complete directed graph G.

Notice that the cardinality of $\mathcal{R}$ is $|\mathcal{R}| = \sum_{k=1}^{n} P(n, k)$, where $P(n, k)$ represents the number of k-permutations of a set of n elements (or customers in our case). Notice that $|\mathcal{R}| = \sum_{k=1}^{n} P(n, k) \approx n!e$, where e represents the Euler's number, $e = \sum_{k=0}^{\infty} \frac{1}{k!}$.

Definition 2.2. Two routes are independent when they have no internal vertices in common, i.e., the only vertex in common is the depot node. A non-empty set of independents routes, $\mathcal{S} \subset \mathcal{R}$, is named a **complete system of routes** when every customer belongs to a route of $\mathcal{S}$. The set of all the complete system of routes of $\mathcal{R}$ is denoted by $\mathcal{CSR}$.

Notice that from now, in order to simplify the notation, when we write $\alpha \in \rho$, with $\rho \in \mathcal{S}$, and $\mathcal{S} \in \mathcal{CSR}$, we want to indicate that α is a node of the route ρ.

Traditionally, the cost of a route, c_ρ, and its distance, d_ρ, have been modeled as

$$c_\rho := c_{\alpha_r, \alpha_0} + \sum_{k=1}^{r} c_{\alpha_{k-1}, \alpha_k}, \quad d_\rho := d_{\alpha_r, \alpha_0} + \sum_{k=1}^{r} d_{\alpha_{k-1}, \alpha_k}.$$

Then, the optimization problem to be solved consists in finding a complete system of routes, $\mathcal{S}$, minimizing the total cost, $c_T := \sum_{\rho \in \mathcal{S}} c_\rho$ subject to the following constraints: the total demand served in each route $\rho \in \mathcal{S}$ does not exceed a maximum constant demand (or vehicles capacity) Q_{max}, $\sum_{\alpha \in \rho} q_\alpha \leqslant Q_{max}$, and the total distance of each route $\rho \in \mathcal{S}$ does not exceed a maximum constant distance D_{max}, $d_\rho \leqslant D_{max}$. Therefore, the optimization problem is

$$\begin{aligned} \text{minimize} \quad & c_T = \sum_{\rho \in \mathcal{S}} c_\rho \\ \text{subject to}: \quad & \sum_{\alpha \in \rho} q_\alpha \leqslant Q_{max}, \quad \rho \in \mathcal{S} \\ & d_\rho \leqslant D_{max}, \quad \rho \in \mathcal{S} \\ & \mathcal{S} \in \mathcal{CSR}. \end{aligned} \quad (1)$$

2.2. Mathematical model for the classical MDVRP

The extension to a MDVRP goes as follows: consider a complete directed graph $G = (V, E)$, where V is the disjoint union (also named a partition) of the set of nodes including the depots, V_d, and the set of nodes including customers V_c, $V := V_d \cup V_c$, and E is the set of edges connecting all nodes in V. Hereafter, $m := |V_d|$ will represent the number of depots. A feasible solution for the MDVRP is a partition of direct graphs $G_i = (V_i, E_i)$, $i = 1, \ldots, m$, obtained from G such that $V_i := \{0_i; v_1^i, \ldots, v_{m_i}^i\}$, for all $i = 1, \ldots, m$, with $0_i \in V_d$ and $v_j^i \in V_c$ for all $j = 1, \ldots, m_i$. Then, the optimization problem to solve consists in finding a family of complete system of routes, $\{\mathcal{S}_1, \ldots, \mathcal{S}_m\}$, minimizing the total cost, $c_T := \sum_{i=1}^{m} \sum_{\rho \in \mathcal{S}_i} c_\rho$ subject to the following constraints: the total demand served in each route $\rho \in \mathcal{S}_i, i = 1, \ldots, m$, does not exceed a maximum constant demand, Q_{max}, i.e., $\beta_\rho := \sum_{\alpha \in \rho} q_\alpha \leqslant Q_{max}$, for all $\rho \in \mathcal{S}_i, i = 1, \ldots, m$, and the total distance of each route $\rho \in \mathcal{S}$ does not exceed a maximum constant distance D_{max}, i.e., for all $\rho \in \mathcal{S}_i, d_\rho \leqslant D_{max}, i = 1, \ldots, m$. Therefore, the optimization problem is

$$\begin{aligned} \text{minimize} \quad & c_T = \sum_{i=1}^{m} \sum_{\rho \in \mathcal{S}_i} c_\rho \\ \text{subject to}: \quad & \beta_\rho \leqslant Q_{max}, \quad \rho \in \mathcal{S}_i, \ i = 1, \ldots, m, \\ & d_\rho \leqslant D_{max}, \quad \rho \in \mathcal{S}_i, \ i = 1, \ldots, m, \\ & \mathcal{S}_i \in \mathcal{CSR}, \quad i = 1, \ldots, m. \end{aligned} \quad (2)$$

2.3. Mathematical model for the MDVRP with heterogeneous depots

The heterogeneous version of the MDVRP analyzed in this paper does not assume depots are equal (homogeneous), which leads to consider customers' preferences. Then, demands will not be fixed parameters, but depend on the assignment map of customers to depots. Following a realistic approach, we assume demands are not known, but can be predicted relying on an historical database and information about new customers. In the heterogeneous case the assignation of the customers is not made in advance using the classical considerations of distance. Our procedure takes into account the combination of statistical predictive models with a metaheuristic, so three main steps must be considered.

(i) Analysis of the historical data so that expected expenditure from new customers among depots is estimated using a multiple regression model. The model captures the relationship between each customer's willingness to spend (response) as a function of several variables (predictors), which include the assigned depot as well as other customer's characteristics as preferences, rent, sex, age, and so on.

(ii) Assignation of the new customers to the depots supported by the company historical data with respect to the existent customers.

(iii) Routes are built, which are associated to each customer-to-depot assignment map.

Notice that revenue incomes are not considered in the model for the classical MDVRP because they do not depend on the assignation of customers to depots and, consequently, they are a

constant value. On the other hand, given the interdependency between both assignation and routing, the procedure is an iterative one. Different assignations are generated (see Fig. 1) then, together with the routing decisions. The top best solutions will be saved and locally improved in the last step of the algorithm in order to maximize the total benefit, b_T, obtained from the difference between the total income, $i_T := \sum_{i=1}^{m} \left(\sum_{\rho \in S_i} \beta_\rho \right)$ and the total cost $c_T := \sum_{i=1}^{m} \left(\sum_{\rho \in S_i} c_\rho \right)$.

$$b_T := i_T - c_T = \sum_{i=1}^{m} \sum_{\rho \in S_i} \left(\beta_\rho - c_\rho \right).$$

Thus, the optimization problem for the heterogeneous case can be described as

$$\begin{aligned} \text{maximize} \quad & b_T := \sum_{i=1}^{m} \sum_{\rho \in S_i} \left(\beta_\rho - c_\rho \right) \\ \text{subject to}: \quad & \beta_\rho \leqslant Q_{max}, \quad \rho \in S_i, \ i = 1, \dots, m, \\ & d_\rho \leqslant D_{max}, \quad \rho \in S_i, \ i = 1, \dots, m, \\ & S_i \in \mathcal{CSR}, \quad i = 1, \dots, m. \end{aligned} \tag{3}$$

3. Literature review on the classical MDVRP

The MDVRP has received a considerable amount of attention in the recent literature (Montoya-Torres et al., 2015). Tillman (1969) is usually referred as the first paper to address this problem. It considers a version where customer demands follow specific probability distributions, which is solved with an extension of the well-known CWS heuristic (Clarke & Wright, 1964). Most works may be classified according to the proposed approach: exact methods and heuristics/metaheuristics methods. The main difference is that the former guarantee the optimality of the solution found, while the latter usually provide a high-quality solution faster. Currently, hybrid approaches have received more attention. Ceselli, Righini, and Salani (2009) is an example of work employing an exact methodology. The authors describe a version of the Multi-Depot Heterogeneous Vehicle Routing Problem with Time Windows (MDHVRPTW) including diverse constraints. A column generation algorithm, in which the pricing problem is a resource-constrained elementary shortest-path problem, is implemented to solve real instances. Another methodology to solve the MDHVRPTW is proposed in Bettinelli, Ceselli, and Righini (2011). It describes a branch-and-cut-and-price algorithm, and different pricing and cutting techniques. More recently, Contardo and Martinelli (2014) have formulated the MDVRP employing a vehicle-flow and a set-partitioning formulation.

A higher number of published works rely on heuristics-based methodologies. For instance, Cordeau, Gendreau, and Laporte (1997) present a Tabu Search (TS) metaheuristic. In Salhi and Sari (1997), the authors propose a multi-level composite heuristic for addressing a MDVRP in which the vehicle fleet composition has to be determined. Nagy and Salhi (2005) consider the MDVRP with Pickups and Delivers. Several heuristics from the Vehicle Routing Problem (VRP) literature are adapted and some problem-specific are constructed. Metaheuristics are frequently implemented to solve real-size instances. The Simulated Annealing (SA) metaheuristic is chosen in Wu, Low, and Bai (2002) for solving the Multi-Depot Location-Routing Problem. Polacek, Hartl, and Doerner (2004) employ the Variable Neighborhood Search (VNS) metaheuristic for addressing the MDVRP with Time Windows (MDVRPTW). The MDVRP with a heterogeneous fleet of vehicles is faced in Salhi, Imran, and Wassan (2014), where an algorithm also based on the VNS metaheuristic is designed. Pisinger and Ropke (2007) tackle different variants of the VRP, including the

MDVRP, by transforming them into rich pickup and delivery models and developing an Adaptive Large Neighborhood Search methodology. A Genetic Algorithm (GA) is constructed in Ombuki-Berman and Hanshar (2009). Another population-based metaheuristic, the Path Relinking, is presented in Rahimi-Vahed, Crainic, Gendreau, and Rei (2013).

Regarding hybrid algorithms, Ho, Ho, Ji, and Lau (2008) introduce an algorithm relying on a GA. The initialization procedure consists in a distance-based grouping, the CWS heuristic is employed for routing, and the Nearest Neighbor Heuristic (NNH) for scheduling (i.e., sequencing each route in every depot). Another hybrid GA is developed in Vidal, Crainic, Gendreau, and Prins (2013) for addressing several rich VRPs, including the MDVRPTW. It has diversity management mechanisms, and employs geometric and structural problem decompositions for large instances. Mirabi, Fatemi-Ghomi, and Jolai (2010) describe a methodology combining a constructive heuristic search and improvement techniques. First, the nearest depot method, the CWS heuristic and the NNH are implemented for grouping, routing, and scheduling, respectively. The resulting solutions are improved by means of a deterministic, stochastic, or the SA metaheuristic. Yu, Yang, and Xie (2011) construct an algorithm based on the Ant Colony metaheuristic, applying a coarse-grain parallel strategy, an ant-weight strategy and mutation operation. Cordeau and Maischberger (2012) design a parallel Iterated Tabu Search heuristic which introduces the TS heuristic into the Iterated Local Search (ILS) framework, in order to ensure a broad exploration of the search space. The Particle Swarm Optimization (PSO) metaheuristic is proposed in Geetha, Vanathi, and Poonthalir (2012). It generates initial particles with the k-means algorithm and the NNH. Lahrichi et al. (2012) present a multi-thread cooperative search method called the Integrative Cooperative Search for multi-attribute combinatorial optimization problems. In Juan, Pascual, Guimarans, and Barrios (2014), the authors combine an ILS metaheuristic with biased-randomization techniques to solve the MDVRP. The same metaheuristic framework is proposed in Li, Pardalos, Sun, Pei, and Zhang (2015). In this case, an adaptive neighborhood selection mechanism is integrated for the MDVRP with simultaneous deliveries and pickups. Luo and Chen (2014a) develop an improved Shuffled Frog Leaping Algorithm (SFLA) and its multi-phase model for the MDVRP and the MDVRPTW. In order to improve the efficiency of the metaheuristic, a Power Law Extremal Optimization Neighborhood Search is used. The same problems are addressed in Luo and Chen (2014b), where a multi-phase modified SFLA is applied. It implements the k-means algorithm and presents cluster and global optimization procedures.

4. Importance of considering market segmentation

In a global and dynamic world, companies have to compete in order to build profitable and long-lived relationships with customers. Analyzing customer needs and desires, capabilities, social values, and objectives of a specific company as well as how these interrelate is a crucial area in business intelligence. During many decades mass market-based strategies had prevailed, which make profit from economies of scale, providing homogeneous goods and services for a vast number of customers. Technological developments and flexible manufacturing systems have boosted the customization of goods and services according to customer preferences (Datta, 1996; Liu, Kiang, & Brusco, 2012). Market segmentation is a key concept in this new approach.

Considering the heterogeneity of markets, segmentation attempts to divide customers into subsets that behave in the same way or have similar needs (Bennett, 1995). As a result, a better understanding of customer requirements is obtained, which may assist in the developing of marketing strategies as well as in the

efficient allocation of resources among markets and products (Wind, 1978). According to Foedermayr and Diamantopoulos (2008), the segmentation process includes the following stages (Fig. 2):

1. Market definition: The scope of the concept of market for a company is chosen. It should be broad enough to cover as many potential customers as possible, but also manageable.
2. Selection of segmentation variables or bases: These bases should be capable of diminishing the market heterogeneity and explaining why customers have different requirements and/or do not respond similarly to marketing campaigns. From the point of view of the company, they should be easy to obtain or infer in terms of cost and time, among others. The most popular are classified into the following groups (Kotler & Armstrong, 2011): (i) geographic bases (e.g., location); (ii) demographic bases (e.g., age, occupation, and education level); (iii) behavioral bases (e.g., purchase occasion, degree of usage, and degree of loyalty); and (iv) psychographic bases (customer activities and opinions).
3. Decision on segmentation method: A-priori versus post hoc methods, and descriptive versus predictive methods, are the criteria most commonly employed to classify segmentation methods (Foedermayr & Diamantopoulos, 2008). A-priori methods are based on intuitions and prior experience, and/or secondary data. While in post hoc methods the data analysis is what leads to the segments. In descriptive methods, no distinction is made between dependent and independent variables. The focus is on exploring the relation between the units of analysis and the variables. In contrast, predictive methods link a dependent variable (e.g., degree of loyalty) to a set of independent variables, and use this set to segment. There are plenty of techniques for segmentation, which includes: cross tabulation analysis, RFM analysis, k-means clustering, hierarchical clustering, self-organizing map (SOM), automatic interaction detection, classification and regression trees, logistic regression, support vector machine, linear regression, clusterwise regression, neural networks, finite mixture model, and metaheuristics, among others. For instance, McCarty and Hastak (2007) investigate RFM, decision trees, and logistic regression. Vellido, Lisboa, and Meehan (1999) present a strategy combining SOM and factor analysis before clustering. Another two-stage approach involving SOM is detailed in Kuo, Ho, and Hu (2014). These authors apply SOM to determine the number of clusters and the starting point, and the k-means algorithm to find the final solution. Huang, Tzeng, and Ong (2007) employ a support vector clustering algorithm. Fish, Barnes, and Aiken (1995) analyze the performance of artificial neural networks, in comparison with those of discriminant analysis and logistic regression. A case-based reasoning system is described in Chen, Wang, and Feng (2010). It implements GAs for selecting variables and instances.
4. Formation of market segments: The method selected in the previous step is applied to obtain a set of segments.
5. Profiling, evaluation, and final selection of target segments: A detailed analysis of the resulting segments and a selection of them are performed. There are several criteria to evaluate

market segments. Smith (1956), considered the first work to tackle this issue, highlights the characteristics of *identifiability*, which means that customers in a segment should have a similar profile, allowing for their identification, and *responsiveness*, i.e., customers in a segment should similarly respond to a marketing strategy. DeSarbo and DeSarbo (2007) gather the main criteria that have been proposed in the literature. Some examples are: *reachability*, *feasibility*, *profitability*, and *stability*.
6. Implementation: The next step is to translate the results of the previous work into specific strategies. This step involves decisions that depend on a large number of factors as relevant as company resources and ethics.
7. Segmentation strategy evaluation: Sales, profit, company expansion, reputation, and customer satisfaction may be used to evaluate a strategy. Although these steps could be sequentially followed, all are interconnected. Therefore, it is recommendable to allow the possibility to repeat previous steps in order to reconsider some selections.

As it has been shown, marketing segmentation has been extensively studied. It is an important topic of research due to its potential applications. New lines of research emerge from the development of data techniques, the gathering of empirical evidences, and the publication of new marketing theories, among others. Many challenges still remain to be faced such as reducing the gap between academic research and practitioner needs, studying implementation issues.

5. Overview of our approach

The MDVRP includes two sequential and correlated stages: (a) the assignment map of customers to depots; and (b) the corresponding design of distribution routes to satisfy all customers' demands. In order to assign customers, we take into account the heterogeneity of the depots. It can be considered a realistic approach, since depots belonging to the same organization usually have different characteristics related to products, trade credit policies, and complementary services, among others. The diversity of depots leads to consider customer preferences. Specifically, the willingness to consume (or expenditure) of each customer depends on how well the assigned depot fits his/her preferences. Market segmentation techniques are applied to identify subsets of customers with similar profiles and assign them to the particular depot that better fits their preferences, considering the restrictions of the problem. Accordingly, we propose to study the relationship between expenditure and customers' features from data of existent customers by employing statistical learning methodologies (e.g., prediction techniques). It will enable the assignation of new customers in such a way that the expected benefits (expected incomes minus distribution costs) is maximized. The phases of our approach are represented in Fig. 3 and described next:

1. Data collection. Our approach requires several inputs: database of historical sales, description of new customers, location of depots, vehicle maximum capacity, number of available vehicles at each depot, and maximum distribution costs per route. The sales database includes the following information for each
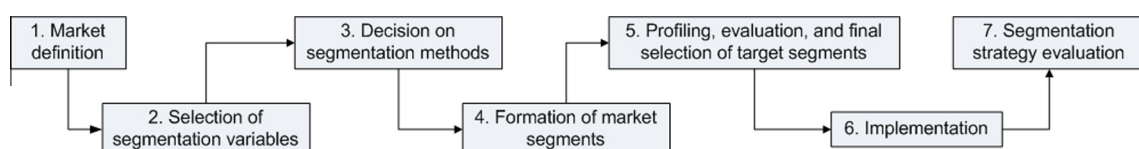

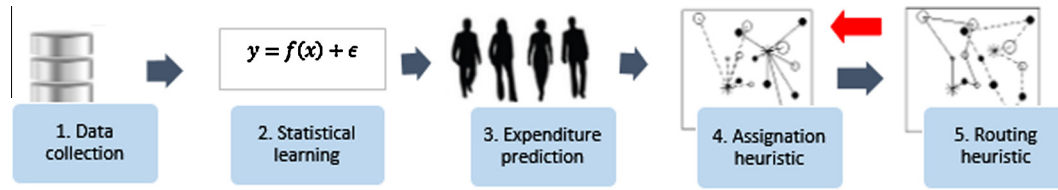**Fig. 2.** Scheme of the segmentation process.

Fig. 3. The proposed approach.

existent customer: personal features, geographical location, expenditure level, and depot to which he/she has been assigned (randomly or according to a metric not related to personal features such as distribution costs). The description of new customers gathers personal features and geographical locations. This information may be easily obtained, for instance, in e-commerce environments, where customers have to register and provide personal data before buying. After processing and analyzing this data, a company may assign a new client by redirecting him/her to a specific directory/website and offering goods from a given depot. Regarding the information of both existent and new customers, an initial selection of variables has to be performed by assessing which ones may be valuable. Besides explaining the differences of expenditures among depots, they should be easy to obtain, estimate or compute, and store.

2. Statistical learning. Given the database of existent customers, a statistical model exploring the relationship between customers' features and expenditure is performed for each group of customers assigned to a specific depot. Considering several groups, we allow the existence of a different trend in each one. A high number of methodologies are available to carry out regression analysis (Hastie, Tibshirani, & Friedman, 2001; Lantz, 2013). Probably, the most applied is Linear Regression (Montgomery, Peck, & Vining, 2012), which is easy to understand and interpret, highly relevant in the marketing literature, and has associated a relatively low risk of overfitting (i.e., the model describing noise). Neural Networks represent a popular alternative capable of capturing non-linear relationships. However, they are computationally more intensive, may overfit/underfit data more easily and are difficult to interpret. Support Vector Machines constitute another powerful black box approach, which is more robust and less prone to overfitting than Neural Networks. Its main disadvantage is that requires testing several combinations of kernels and model parameters. Model Trees combine Decision Trees with modeling of numeric data. It results in an approach that may fit some types of data better than linear regression and perform automatic feature selection. On the other hand, it may be difficult to determine the overall net effect of individual variables on the response.

3. Prediction of expenditure for new customers. Once a methodology has been selected and the different functions have been fitted, the expenditure is predicted for each new customer given his/her features if assigned to each depot. Here, it is assumed that the sample (set of existent customers) is representative of the population (market).

4. Assignment of customers to depots. In order to perform an efficient and feasible assignation, it is necessary not only to consider the predicted expenditure but also the distribution costs, the maximum number of vehicles per depot, and their capacity. Taking a decision for each customer individually may provide non-feasible and poor-quality solutions. Consequently, we present a global and iterative strategy where customers are selected one at a time to be assigned to a specific depot. It prioritizes the assignments of those customers that have

associated a relatively high expected benefits only for a particular depot, and is based on the procedure developed in Juan et al. (2014). In particular, the following steps are proposed:

- For each depot k and customer i,
  - Compute the expected benefits μ_i^k as the difference between the predicted expenditure p_i^k and the distribution costs c_i^k (computed as the cost of moving from k to i).
  - Compute the difference between the expected benefits of assigning i to k and the maximum expected benefits of assigning i to a depot l other than k, i.e.:

$$s_i^k = \mu_i^k - max_{l \in V_d \setminus \{k\}} \, \mu_i^l \quad \forall i \in V_c, \, \forall k \in V_d.$$

We refer to this measure as "marginal savings". Accordingly, s_i^k will be high in the case customer i reports relevant expected benefits only if assigned to k, low (in absolute terms) if the expected benefits are similar for k and at least one other depot, presenting both depots the highest expected benefits, and very low (negative) when there is at least one depot where the expected benefits are larger than those estimated for k.

- For each depot k, create a priority list of customers and sort it in descending order according to the marginal savings s_i^k.
- Create a list of unassigned customers. Then, select a depot and choose the next customer to assign from its priority list. Update the list of unassigned customers and repeat these steps while there are unassigned customers. Different policies may be applied to determine which depot selects the next customer, as: (i) allowing the depot with the highest remaining capacity to choose, (ii) using a round robin-based criterion, or (iii) selecting it randomly.

5. Routing. Having an assignment map, the MDVRP can be solved as a set of independent CVRPs. However, the most important challenge when addressing a MDVRP instance is the interrelation between assignation and routing. Therefore, algorithms are required to take the decisions associated to both phases 'simultaneously'. Thus, instead of finding an optimal or near-optimal solution for the customer-to-depot assignment phase and then use this unique solution as a starting point to solve the routing phase, an iteration process combines 'good' and fast computed solutions for the first stage with 'good' and fast computed solutions for the second one in order to find a near-optimal solution for the overall problem.

Note that our approach will be appropriate as long as the existent customers had been assigned randomly or based on a variable not related to personal features. If regression functions were estimated again after implementing this procedure (replacing existent customers by the new ones), the predictive model could be not valid anymore, since the groups of customers assigned to each depot may not be representative of all potential customers. At this point, a description of each resulting group may be performed. Accordingly, a new customer would be assigned to the closest group (considering standardized data, the Euclidean distance, and an average profile per group, for instance). In the described

approach, the statistical learning techniques and the metaheuristic are sequentially employed. There are other realistic versions of the problem that may be addressed by adapting our approach to integrate the statistical learning techniques inside the metaheuristic. For instance, consider a dynamic scenario in which the willingness of customers to spend varies as new customers are assigned to each depot (e.g., due to the decrease in the service's quality or in the number of available offers). In this case, the learning mechanism would iteratively run throughout the searching process in order to update each customer's willingness to spend after each assignment.

6. Detailed algorithm

This section describes some low-level details of the proposed approach. Fig. 4 summarizes it highlighting the main differences between the classical version of the problem and the proposed one.

Since the phase of data collection is company-specific, we will assume it has already been done. The second and the third phases are related to the development and use of predictive statistical learning models. First, the database of existent customers is split into two subsets: a training set, which will be used to build the models, and a test set, to assess their performance. These subsets are generated by means of random sampling: 75% of customers are assigned to the training set and 25% to the test set. Having different alternatives to explore the relationship between expenditure and customers' features, in our experiments (described later in this paper) we make use of three well-known methodologies: Multiple Linear Regression (MLR), Multi-layer Feedforward Network (MFN), and Model Tree.

- Regarding Multiple Linear Regression, given a database of customers with m features and $|V_d|$ depots, the models proposed may be described as follows:
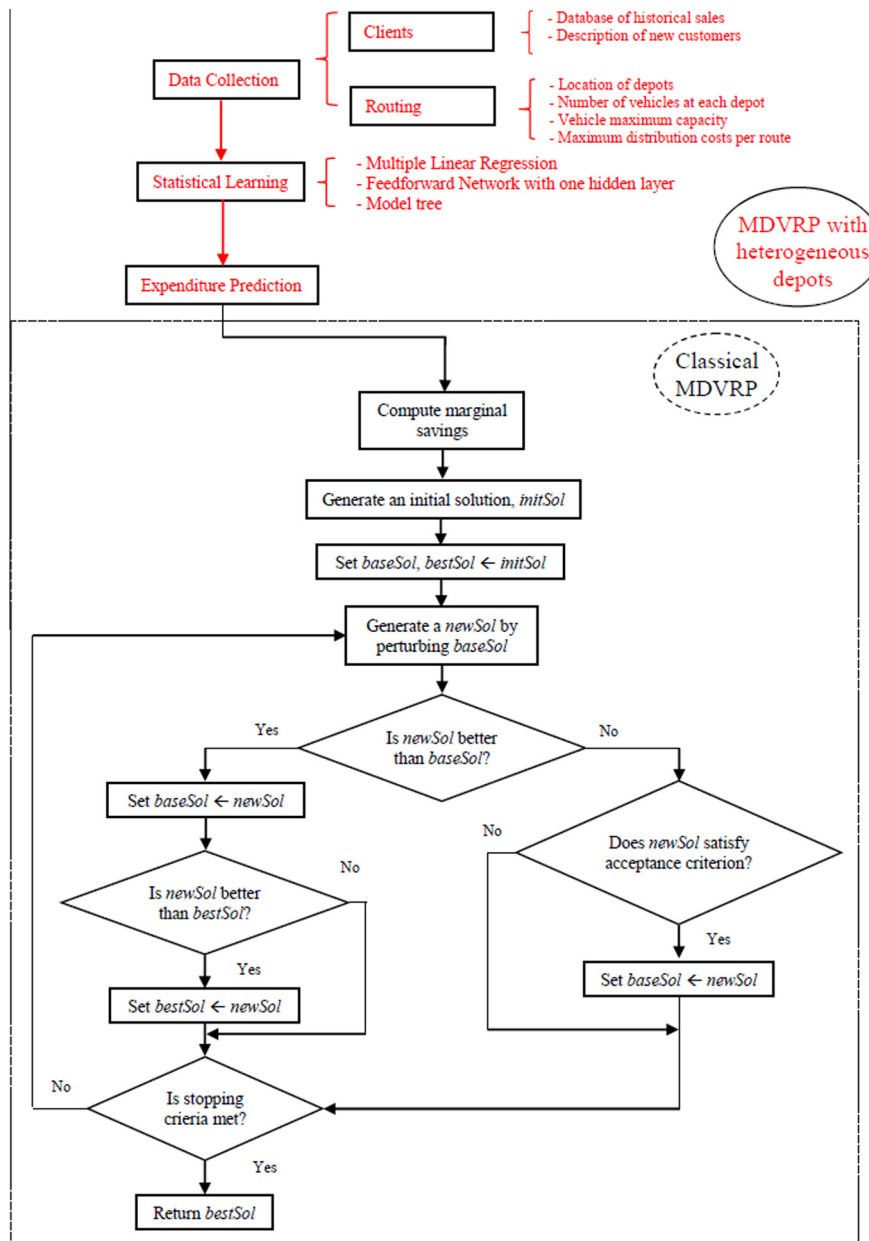

Fig. 4. Flow chart of our approach for solving the MDVRP with heterogeneous depots.

$$Exp_i = \beta_0^j + \beta_1^j \cdot f_{1i} + \beta_2^j \cdot f_{2i} + \ldots + \beta_m^j \cdot f_{mi} + \epsilon_i \quad \forall i \in V_c^j, \ \forall j \in V_d,$$

where $f_{1i}, \ldots, f_{mi}$ represent the features of customer i, $\beta_0^j, \ldots, \beta_m^j$ are the parameters of the model, Exp_i and ϵ_i denote the expenditure and an error term for customer i, and V_c^j is the set of customers assigned to depot j. The ordinary least squares method is applied to estimate the parameters, and the stepwise regression approach with a bidirectional elimination procedure is chosen to perform the variable selection.

- Regarding the Multi-layer Feedforward Network with one hidden layer, the generated models are:

$$Z_{li} = \sigma(\beta_0^{jl} + \beta_1^{jl} \cdot f_{1i} + \beta_2^{jl} \cdot f_{2i} + \cdots + \beta_m^{jl} \cdot f_{mi})$$
$$\forall i \in V_c^j, \ \forall j \in V_d, \ l = 1, \ldots, p,$$

$$Exp_i = \alpha_0^j + \alpha_1^j \cdot Z_{1i} + \cdots + \alpha_p^j \cdot Z_{pi} \quad \forall i \in V_c^j, \ \forall j \in V_d,$$

where σ is the sigmoid function and p the number of hidden units. The value of p $(4, 5, 6, 7,$ or $8)$ and the decay value for regularization $(0.2, 0.3, 0.4, 0.5$ or $0.6)$ are set using 10-fold cross validation based on the metric R^2 (Kuhn, 2008). The back propagation method is employed to estimate the parameters.

- The algorithm selected to implement a model tree is the standard M5P (Wang & Witten, 1996). Basically, it builds a decision-tree induction algorithm relying on a splitting criterion that minimizes the intra-subset variation in the class values down each branch. The pruning of the tree is performed back from each leaf. Instead of a constant value, the final solution for each leaf is a linear regression model considering the variables participating in decisions.

Different criteria can be employed to select one of the former statistical learning methodologies. The most common criteria are related to performance, easiness to apply and understand, required time, or any combination of the aforementioned properties. Considering the first one, we compute the Mean Squared Error (MSE) for each model (the number of models is the number of depots multiplied by the number of methodologies tested) using the same problem instance. The Total MSE (TMSE) is computed by aggregating the values of the models corresponding to the same methodology. In mathematical terms:

$$MSE^{aj} = \frac{1}{|V_c^j|} \sum_{\forall i \in V_c^j} \left(\widehat{Exp_i^a} - Exp_i \right)^2 \quad \forall a = 1, \ldots, o \quad \forall j \in V_d,$$

$$TMSE^a = \sum_{j=1}^{|V_d|} MSE^{aj} \quad \forall a = 1, \ldots, o,$$

where a represents the methodology assessed, and $\widehat{Exp_i^a}$ refers to the predicted expenditure for customer i employing the methodology a. In our experiments, for each instance we always select the methodology associated with the lowest *TMSE*. Thus, during the third phase, the expenditure that each new customer would make if he/she was assigned to each one of the depots is predicted using the selected methodology and the customer's features.

For the assignation and the routing phases, an existing methodology described in Juan et al. (2014) has been adapted. The authors propose an efficient algorithm based on an ILS metaheuristic framework (Lourenço, Martin, & Stützle, 2010), which is a popular choice for solving routing problems (see Cattaruzza, Absi, Feillet, & Vigo, 2014). This metaheuristic guides the search by interspersing exploration and intensification movements. Firstly, an initial solution is generated assigning customers to depots according to the marginal savings (only the distribution costs are considered) and

designing the routes by implementing the classical CWS heuristic (Clarke & Wright, 1964). Afterwards, an iterative procedure is started in which the base solution (the initial solution in the first iteration) is perturbed. If the new solution is better than the base solution, then the latter is replaced. In case no improvement is achieved, a Demon-based acceptance criterion (Talbi, 2009) is considered to avoid entrapment at local optimum. It allows movements that deteriorate the base solution with a higher frequency at the beginning, when a global search is required, and restricts them as the execution proceeds. These steps are repeated until a termination condition is met. Finally, the top best solutions are improved by means of a post optimization process, and the best one is returned. The described algorithm includes Biased Randomization techniques to further diversify the search (Juan et al., 2009). These techniques are introduced in traditionally deterministic steps in order to add biased randomization, which favors the generation of high-quality alternatives. In particular, they are implemented both in the assignation phase, to randomize the sorted priority list of customers of each depot in such a way that the reasoning behind the sorting is not erased but many orderings are provided, and in the routing phase, where the CWS heuristic is randomized.

7. Numerical experiments

An algorithm based on the described approach has been implemented and employed to solve a number of generated instances. The computational experiments compare the results of our approach for the analyzed version of the MDVRP and for the classical version (i.e., the one assuming homogeneous depots). This section provides the description of the instances and the tests carried out, as well as the numerical results and their analysis.

7.1. Set of instances

A total of 15 instances have been generated. Each of them consists in three datasets: the first two gather data concerning existent and new customers, respectively, and the third includes depots' locations and information related to restrictions. Regarding data of existent customers, four variables have been created: age (a discrete variable following a Uniform distribution with parameters 16 and 80), sex (a categorical variable with two equally probable values), estimated income (it follows a Normal distribution with a mean of 1500 and standard deviation of 300), and preferred article (a categorical variable including four equally probable values). Initially, each customer has been assigned to his/her closest depot, while the expenditure level has been determined by a given function that depends on the depot, the aforementioned variables and a white noise term. For a total of 100 new customers, the variables age, sex, estimated income and preferred article have been generated using the same distributions. Customers' and depots' locations have been randomly generated in a square of 100×100. In order to simplify the instances' generation, Euclidean distances are employed as distribution costs. Different values have been chosen for the number of depots, existent customers and vehicles, the maximum cost per route and vehicles' capacity. This information is shown in Table 1.

7.2. Test

Each instance has been adapted by modifying the expenditure of existent customers to analyze the following scenarios: (1) low ratio (LR), the ratio between average expenditure of existent customers and average distribution costs is similar; (2) medium ratio (MR), average expenditure is relatively higher than average

Table 1
Description of the generated instances.

Instance	Numb. depots	Numb. existent cust.	Numb. vehicles	Vehicle capacity	Max. cost
1	3	300	3	250	200
2	3	300	3	225	200
3	3	300	3	225	150
4	3	300	3	225	200
5	3	300	3	200	150
6	3	400	3	350	225
7	3	400	3	300	200
8	3	400	3	200	175
9	5	400	4	325	175
10	5	400	4	200	150
11	5	400	4	275	175
12	5	400	4	275	150
13	5	400	4	225	200
14	5	400	4	175	125
15	5	400	4	250	175

Fig. 5. Boxplot of the expected benefits for each scenario and version of the problem.

Fig. 6. Boxplot of the distribution costs and the expected incomes for the rich version of the problem.

distribution costs; and (3) high ratio (HR), average expenditure is much higher than average distribution costs. The target ratio has been reached multiplying expenditures by a coefficient. The resulting instances are available from the authors upon request. The analysis of these scenarios will allow us to compare the expected benefits (expected incomes, defined as the sum of predicted expenditures, minus distribution costs) associated to solutions considering only distribution costs and those taking into account also customers' preferences (predicted expenditure), thus exploring the consequences of having different weights of expenditure in the objective solution. For the first scenario, it is expected that the gap between distribution costs will be low (i.e., solutions are expected to be relatively similar). Likewise, it is expected that this gap will be higher as the ratio increases. Similarly, it is also expected that the higher the ratio, the higher the gap between the expected benefits of the solutions. The code has been implemented with Java and R – version 2.15.0 (Team, 2008) (packages: caret, MASS, nnet, and RWeka). A standard personal computer, Intel QuadCore i5 CPU at 3.2 GHz and 4 GB RAM with Windows XP, has been used to perform all tests. The ILS process runs for 4,000 iterations, and all executions are solved for 10 different seeds. Only the best values obtained after the 10 runs are reported.

7.3. Results and analysis

The results of the experiments carried out are summarized in Figs. 5 and 6. The boxplots in the first figure show the expected benefits per scenario and version of the problem: considering heterogeneous depots (rich) and assuming homogeneous ones (traditional). Even if the medians associated to each ratio level do not differ significantly, the third and second quartile values do present a higher value for the extended version of the problem. This behavior is caused by the long right tails of the corresponding distributions, which indicate that for some instances the rich version results in better solutions in terms of expected benefits. The second figure displays the variables in which expected benefits are decomposed per scenario considering the rich version. We observe that differences of expected benefits between scenarios are mainly due to differences between expected incomes.

Tables 2–4 provide a detailed description of the results. The information gathered in the tables is the following: instance name; methodology selected for prediction; distribution costs, expected incomes, expected benefits and time associated to the best solution found considering only distribution costs (classical MDVRP) and to the best solution found when maximizing expected benefit (MDVRP with heterogeneous depots); and gaps between distribution costs, expected incomes and expected benefits of both solutions. The average of each gap is also shown.
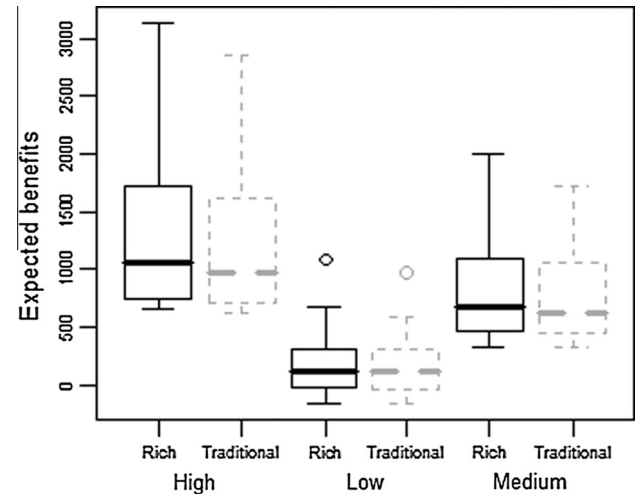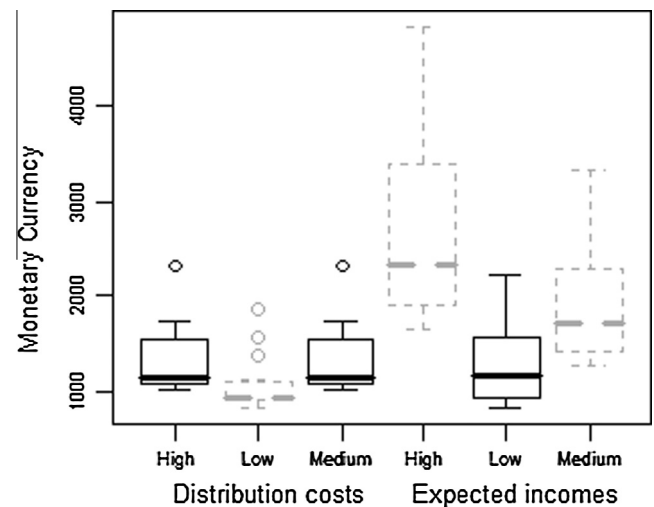
Given the flexibility of Feedforward Neural Networks to model relationships between variables, and despite the basic topology and parameter fine-tuning, and the medium size of the training set, they have been selected to solve more than half of the instances (57.8%). Multiple Linear Regression has provided the best TMSE in a high number of cases (31.1%). Although less frequently, the algorithm M5P has also been used in some instances (11.1%). Being an experiment for illustrative purposes, we show that different methodologies with particular strengths may be easily applied, but we do not aim to perform a comprehensive comparison among them.

The gaps related to the distribution costs and the expected incomes are strictly positive except in one case. It confirms the trade-off decision-makers face between both measures; that is to say, higher distribution costs are required to obtain an increase in expected incomes. Regarding the gap of expected benefits, it is strictly positive for all instances except for two where both solutions are equal. Therefore, attempting to achieve the highest benefits studying only distribution costs in instances with heterogeneous depots results in sub-optimal solutions. As expected, all average gaps increase with the ratio, i.e., the difference between solutions (in terms of distribution costs, expected incomes or expected benefits) is positively correlated to the

Table 2
Results obtained for 15 instances: scenario characterized by a low ratio.

Inst.	Meth.	Traditional (1)				Rich(2)				Gaps(2-1)		
		Dist. cost	Exp. Inc.	Exp. ben.	Time	Dist. cost	Exp. Inc.	Exp. ben.	Time	Dist. cost	Exp. Inc.	Exp. ben.
p01.1	MLR	898.6	961	62.4	82	930.6	1006	75.4	123	31.9	45.0	13.1
p02.1	M5P	834.3	943	108.7	112	834.5	947	112.6	335	0.1	4.0	3.9
p03.1	MFN	944.0	911	−33.0	143	964.4	939	−25.4	159	20.4	28.0	7.6
p04.1	MFN	891.8	852	−39.8	79	923.4	884	−39.4	165	31.6	32.0	0.4
p05.1	MFN	909.7	824	−85.7	189	914.4	829	−85.4	66	4.8	5.0	0.2
p06.1	MFN	868.5	1425	556.5	655	870.2	1429	558.8	613	1.7	4.0	2.3
p07.1	MFN	923.4	1073	149.6	103	925.7	1093	167.3	383	2.3	20.0	17.7
p08.1	M5P	898.2	867	−31.2	105	900.9	872	−28.9	122	2.7	5.0	2.3
p09.1	MLR	1039.2	2008	968.8	91	1127.5	2218	1090.5	33	88.3	210.0	121.7
p10.1	MFN	1029.6	1404	374.4	63	1062.5	1462	399.5	40	32.9	58.0	25.1
p11.1	MLR	880.7	1469	588.3	47	939.1	1609	669.9	464	58.4	140.0	81.6
p12.1	MFN	1858.4	1699	−159.4	108	1864.2	1709	−155.2	328	5.8	10.0	4.2
p13.1	MLR	1428.3	1495	66.7	437	1568.0	1691	123.0	144	139.6	196.0	56.4
p14.1	MFN	930.0	1163	233.0	43	930.0	1163	233.0	40	0.0	0.0	0.0
p15.1	M5P	1268.1	1401	132.9	374	1375.0	1512	137.0	59	107.0	111.0	4.0
									Average	35.2	57.9	22.7

Table 3
Results obtained for 15 instances: scenario characterized by a medium ratio.

Inst.	Meth.	Traditional (1)				Rich(2)				Gaps(2-1)		
		Dist. cost	Exp. Inc.	Exp. ben.	Time	Dist. cost	Exp. Inc.	Exp. ben.	Time	Dist. cost	Exp. Inc.	Exp. ben.
p01.2	MLR	925.3	1383	457.7	277	978.0	1483	505.0	173	52.7	100.0	47.3
p02.2	MLR	901.2	1334	432.8	301	921.9	1385	463.1	254	20.7	51.0	30.3
p03.2	MLR	959.3	1405	445.7	134	979.1	1438	458.9	89	19.8	33.0	13.2
p04.2	MFN	942.5	1280	337.5	124	947.8	1292	344.3	101	5.3	12.0	6.7
p05.2	MFN	919.0	1264	345.0	51	921.3	1269	347.8	221	2.3	5.0	2.7
p06.2	MFN	945.6	2103	1157.4	106	948.6	2122	1173.4	327	3.1	19.0	15.9
p07.2	MFN	962.8	1581	618.2	394	992.3	1617	624.7	139	29.5	36.0	6.5
p08.2	MFN	969.9	1302	332.1	300	969.9	1302	332.1	296	0.0	0.0	0.0
p09.2	MFN	1169.6	2897	1727.4	36	1336.1	3335	1998.9	173	166.5	438.0	271.5
p10.2	MFN	1165.1	2109	943.9	161	1222.9	2222	999.1	97	57.8	113.0	55.2
p11.2	MLR	1001.8	2212	1210.2	80	1054.4	2288	1233.7	253	52.5	76.0	23.5
p12.2	MFN	1050.0	2571	1521.0	75	1070.5	2620	1549.5	41	20.6	49.0	28.4
p13.2	MLR	1633.4	2178	544.6	106	1778.2	2446	667.8	270	144.8	268.0	123.2
p14.2	MFN	1020.2	1703	682.8	63	1026.8	1717	690.2	67	6.6	14.0	7.4
p15.2	M5P	1419.6	2090	670.4	69	1560.2	2257	696.8	106	140.5	167.0	26.5
									Average	48.2	92.1	43.9

Table 4
Results obtained for 15 instances: scenario characterized by a high ratio.

Inst.	Meth.	Traditional (1)				Rich(2)				Gaps(2-1)		
		Dist. cost	Exp. Inc.	Exp. ben.	Time	Dist. cost	Exp. Inc.	Exp. ben.	Time	Dist. cost	Exp. Inc.	Exp. ben.
p01.3	MLR	1060.3	1930	869.7	199	1153.7	2132	978.3	42	93.4	202.0	108.6
p02.3	M5P	1070.7	1803	732.3	253	1097.0	1864	767.0	174	26.3	61.0	34.7
p03.3	MFN	1042.7	1864	821.3	23	1067.1	1923	855.9	162	24.4	59.0	34.6
p04.3	MFN	1043.2	1701	657.8	54	1080.5	1755	674.5	393	37.2	54.0	16.8
p05.3	MFN	994.0	1621	627.0	174	1011.0	1657	646.0	68	17.0	36.0	19.0
p06.3	MFN	1068.1	2856	1787.9	109	1102.7	2906	1803.3	208	34.6	50.0	15.4
p07.3	MFN	1064.1	2115	1050.9	152	1081.2	2139	1057.8	71	17.1	24.0	6.9
p08.3	M5P	1069.6	1741	671.5	32	1069.6	1741	671.5	261	0.0	0.0	0.0
p09.3	MLR	1420.5	4269	2848.5	37	1690.6	4825	3134.4	138	270.1	556.0	285.9
p10.3	MFN	1434.8	2913	1478.2	113	1734.8	3396	1661.2	33	299.9	483.0	183.1
p11.3	MLR	1238.0	3020	1782.0	25	1486.3	3407	1920.7	265	248.3	387.0	138.7
p12.3	MFN	1195.7	3385	2189.3	37	1216.1	3452	2235.9	125	20.3	67.0	46.7
p13.3	MLR	1843.3	2801	957.7	79	2321.4	3387	1065.6	101	478.1	586.0	107.9
p14.3	MFN	1198.9	2297	1098.1	17	1251.0	2351	1100.0	23	52.1	54.0	1.9
p15.3	M5P	1416.0	2086	670.0	164	1595.5	2311	715.6	210	179.5	225.0	45.5
									Average	119.9	189.6	69.7

average expenditure for fixed average distribution costs. However, this rule does not apply for all cases. In some of them, despite the fact that the gap of expected incomes increases, so does the gap of distribution costs. As a consequence, the gap of expected benefit may be reduced.

8. Conclusions

This paper addresses an extension of the Multi-Depot Vehicle Routing Problem (MDVRP) in which heterogeneous depots are considered. The resolution of the classical MDVRP has two sequential

and interrelated stages: (a) the assignment of customers to depots, and (b) the corresponding design of distribution routes. Typically, the assignment map is generated by minimizing the total distance, which is intended to lead to the minimization of distribution costs. Implementing this approach, researchers assume that depots are homogeneous. However, this is an unrealistic assumption since several factors may result in differences between depots from a particular organization. We propose to take into account the existence of heterogeneous depots, which allows the consideration of customers' preferences. The customers' willingness to consume is affected by how well the assigned depot fits their preferences. Thus, the main contribution of this work is the development of a simple yet comprehensive metaheuristic-based approach including market segmentation issues in order to maximize expected benefits (expected sales incomes minus distribution costs).

The proposed methodology consists of five steps: (i) data collection, in which information basically related to existent customers that have been already served and new customers is gathered; (ii) statistical learning, where the relationship between customers' features and expenditure for different depots is studied employing existent customer data; (iii) expenditure prediction for new customers; (iv) assignment of new customers; and (v) routing. A set of computational experiments has been carried out in order to illustrate our methodology. A total of 15 instances have been artificially generated and analyzed considering three scenarios, which vary in the weight of the expenditure of existent customers. It has been shown how our approach differs from an approach based only on minimizing distribution costs when solving instances with heterogeneous depots. Our experiment also allows quantifying how the performance gap between both approaches increases as the weight of the expenditures is incremented.

Acknowledgements

References

Barbucha, D. (2014). A cooperative population learning algorithm for vehicle routing problem with time windows. *Neurocomputing, 146*, 210–229.

Bennett, P. (1995). *Dictionary of marketing terms* (2nd ed.). McGraw-Hill Contemporary.

Bettinelli, A., Ceselli, A., & Righini, G. (2011). A branch-and-cut-and-price algorithm for the multi-depot heterogeneous vehicle routing problem with time windows. *Transport Research Part C: Emerging Technologies, 19*, 723–740.

Caceres, J., Arias, P., Guimarans, D., Riera, D., & Juan, A. (2015). Rich vehicle routing problem: A survey. *ACM Computing Surveys, 47*, 115–129.

Cattaruzza, D., Absi, N., Feillet, D., & Vigo, D. (2014). An iterated local search for the multi-commodity multi-trip vehicle routing problem with time windows. *Computers & Operations Research, 51*, 257–267.

Ceselli, A., Righini, G., & Salani, M. (2009). A column generation algorithm for a rich vehicle routing problem. *Transport Science, 43*, 56–69.

Chao, I., Golden, B., & Wasil, E. (1993). A new heuristic for the multi-depot vehicle routing problem that improves upon best known solutions. *American Journal of Mathematical and Management Sciences, 13*, 371–406.

Chen, Y.-K., Wang, C.-Y., & Feng, Y.-Y. (2010). Application of a 3NN+1 based CBR system to segmentation of the note book computers market. *Expert Systems with Applications, 37*, 276–281.

Clarke, G., & Wright, J. (1964). Scheduling of vehicles from a central depot to a number of delivering points. *Operations Research, 12*, 568–581.

Contardo, C., & Martinelli, R. (2014). A new exact algorithm for the MDVRP under capacity and route length constraints. *Discrete Optimization, 12*, 129–146.

Cordeau, J., Gendreau, M., & Laporte, G. (1997). A tabu search heuristic for periodic and multidepot vehicle routing problems. *Networks, 30*, 105–119.

Cordeau, J., & Maischberger, M. (2012). A parallel iterated tabu search heuristic for vehicle routing problems. *Computers & Operations Research, 39*, 2033–2050.

Datta, Y. (1996). Market segmentation: An integrated framework. *Long Range Planning, 29*, 797–811.

DeSarbo, W., & DeSarbo, C. (2007). A generalized normative segmentation methodology employing conjoint analysis. In A. Gustafsson, A. Herrmann, & F. Huber (Eds.), *Conjoint measurement: Methods and application* (pp. 321–345). Springer Science & Business Media.

Ehmke, J., Campbell, A., & Urban, T. (2015). Ensuring service levels in routing problems with time windows and stochastic travel times. *European Journal of Operational Research, 240*, 539–550.

Fish, K., Barnes, J., & Aiken, M. (1995). Artificial neural networks: a new methodology for industrial market segmentation. *Industrial Marketing Management, 24*, 431–438.

Foedermayr, E., & Diamantopoulos, A. (2008). Market segmentation in practice: Review of empirical studies, methodological assessment, and agenda for future research. *Journal of Strategic Marketing, 16*, 223–265.

Geetha, S., Vanathi, P., & Poonthalir, G. (2012). Metaheuristic approach for the multi-depot vehicle routing problem. *Applied Artificial Intelligence, 26*, 878–901.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning. Springer series in statistics.* Springer New York Inc.

Ho, W., Ho, G., Ji, P., & Lau, H. (2008). A hybrid genetic algorithm for the multi-depot vehicle routing problem. *Engineering Applications of Artificial Intelligence, 21*, 548–557.

Huang, J.-J., Tzeng, G.-H., & Ong, C.-S. (2007). Marketing segmentation using support vector clustering. *Expert Systems with Applications, 32*, 313–317.

Juan, A., Faulin, J., Ruiz, R., Barrios, B., Gilibert, M., & Vilajosana, X. (2009). Using oriented random search to provide a set of alternative solutions to the capacitated vehicle routing problem. In *Operations research and cyber-infrastructure* (pp. 331–346). New York, USA: Springer.

Juan, A., Pascual, I., Guimarans, D., & Barrios, B. (2014). Combining biased randomization with iterated local search for solving the multi-depot vehicle routing problem. *International Transactions in Operational Research, 22*, 647–667.

Kotler, P., & Armstrong, G. (2011). *Principles of marketing* (14th ed.). Pearson Prentice Hall.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software, 28*.

Kuo, R., Ho, L., & Hu, C. (2014). Integration of self-organizing feature map and K-means algorithm for market segmentation. *Computers & Operations Research, 29*, 1475–1493.

Lahrichi, N., Crainic, T., Gendreau, M., Rei, W., Crişan, G., & Vidal, T. (2012). An integrative cooperative search framework for multi-decision-attribute combinatorial optimization. Technical Report No. 42 CIRRELT.

Lantz, B. (2013). *Machine learning with R.* Packt Publishing.

Li, J., Pardalos, P., Sun, H., Pei, J., & Zhang, Y. (2015). Iterated local search embedded adaptive neighborhood selection approach for the multi-depot vehicle routing problem with simultaneous deliveries and pickups. *Expert Systems with Applications, 42*, 3551–3561.

Liu, Y., Kiang, M., & Brusco, M. (2012). A unified framework for market segmentation and its applications. *Expert Systems with Applications, 39*, 10292–10302.

Lourenço, H., Martin, O., & Stützle, T. (2010). Iterated local search: Framework and applications. In M. Gendreau & J.-Y. Potvin (Eds.), *Handbook of metaheuristics. International series in operations research & management science* (Vol. 146, pp. 363–397). US: Springer.

Luo, J., & Chen, M.-R. (2014a). Improved shuffled frog leaping algorithm and its multi-phase model for multi-depot vehicle routing problem. *Expert Systems with Applications, 41*, 2535–2545.

Luo, J., & Chen, M.-R. (2014b). Multi-phase modified shuffled frog leaping algorithm with extremal optimization for the MDVRP and the MDVRPTW. *Computers & Industrial Engineering, 72*, 84–97.

McCarty, J., & Hastak, M. (2007). Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. *Journal of Business Research, 60*, 656–662.

Mirabi, M., Fatemi-Ghomi, S., & Jolai, F. (2010). Efficient stochastic hybrid heuristics for the multi-depot vehicle routing problem. *Robotics and Computer-Integrated Manufacturing, 26*, 564–569.

Montgomery, D., Peck, E., & Vining, G. (2012). *Introduction to linear regression analysis* (5th ed.). Wiley Publishing.

Montoya-Torres, J., Lopez, J., Nieto, S., Felizzola, H., & Herazo-Padilla, N. (2015). A literature review on the vehicle routing problem with multiple depots. *Computers & Industrial Engineering, 79*, 115–129.

Nagy, G., & Salhi, S. (2005). Heuristic algorithms for single and multiple depot vehicle routing problems with pickups and deliveries. *European Journal of Operational Research, 162*, 126–141.

Ombuki-Berman, B., & Hanshar, F. (2009). Using genetic algorithms for multi-depot vehicle routing. In F. Pereira & J. Tavares (Eds.), *Bio-inspired algorithms for the vehicle routing problem* (pp. 77–99). Berlin, Heidelberg: Springer.

Pisinger, D., & Ropke, S. (2007). A general heuristic for vehicle routing problems. *Computers & Operations Research, 34*, 2403–2435.

Polacek, M., Hartl, R., & Doerner, K. (2004). A variable neighborhood search for the MDVRP with time windows. *Journal of Heuristics, 10*, 613–627.

Rahimi-Vahed, A., Crainic, T., Gendreau, M., & Rei, W. (2013). A path relinking algorithm for a multi-depot periodic vehicle routing problem. *Journal of Heuristics, 19*, 497–524.

Salhi, S., Imran, A., & Wassan, N. (2014). The multi-depot vehicle routing problem with heterogeneous vehicle fleet: Formulation and a variable neighborhood search implementation. *Computers & Operations Research, 52*, 315–325.

Salhi, S., & Sari, M. (1997). A multi-level composite heuristic for the multi-depot vehicle fleet mix problem. *European Journal of Operational Research, 103*, 95–112.

Smith, W. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing, 21*, 3–8.

Taş, D., Jabali, O., & Woensel, T. (2014). A vehicle routing problem with flexible time windows. *Computers & Operations Research, 52*, 39–54.

Talbi, E. (2009). *Metaheuristics: From design to implementation*. New Jersey: John Wiley & Sons.

RDC Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Tillman, F. (1969). The multiple terminal delivery problem with probabilistic demands. *Distribution Science, 3*, 192–204.

Vellido, A., Lisboa, P., & Meehan, K. (1999). Segmentation of the on-line shopping market using neural networks. *Expert Systems with Applications, 17*, 303–314.

Vidal, T., Crainic, T., Gendreau, M., & Prins, C. (2013). A hybrid genetic algorithm with adaptive diversity management for a large class of vehicle routing problems with time windows. *Computers & Operations Research, 40*, 475–489.

Wang, Y., & Witten, I. (1996). Induction of model trees for predicting continuous classes. Technical Report Working paper 96/23 Hamilton, University of Waikato, New Zealand.

Wind, Y. (1978). Issues and advances in segmentation research. *Journal of Marketing Research, 15*, 317–337.

Wu, T.-H., Low, C., & Bai, J.-W. (2002). Heuristic solutions to multi-depot location-routing problems. *Computers & Operations Research, 29*, 1393–1415.

Yu, B., Yang, Z.-Z., & Xie, J.-X. (2011). A parallel improved ant colony optimization for multi-depot vehicle routing problem. *Journal of the Operational Research Society, 62*, 183–188.