The probabilistic customer's choice rule with a threshold attraction value: effect on the location of competitive facilities in the plane

José Fernández*

Dpt. of Statistics and Operations Research, University of Murcia, Murcia, Spain

Boglárka G.- Tóth

Institute of Informatics, University of Szeged, Szeged, Hungary

Juana L. Redondo and Pilar M. Ortigosa

Dpt. of Informatics, University of Almería, ceiA3, Almería, Spain

Abstract

The classical probabilistic choice rule assumes that customers patronize *all* the existing facilities. As this assumption may not be appropriate in some cases, in this paper a variant is investigated, in which a customer only patronizes those facilities for which he/she feels an attraction greater than or equal to a threshold value. Implicitly, this implies that there may be some unmet demand. We apply this modified rule to the problem of locating a single new facility in the plane. A comparison of the location decisions derived from the modified rule with those obtained with the classical proportional choice rule when solving the location model reveals that the profit that the locating chain may lose if an inadequate choice rule is employed may be quite high in some instances.

Keywords:
Patronizing behavior, Customer choice rule, Competition, Continuous

*Corresponding author

Email addresses: `josefdez@um.es` (José Fernández), `boglarka@inf.u-szeged.hu` (Boglárka G.- Tóth), `{jlredondo, ortigosa}@ual.es` (Juana L. Redondo and Pilar M. Ortigosa)

URL: `http://www.um.es/geloca/gio/josemain.html` (José Fernández)

Location, Interval branch-and-bound method, Evolutionary algorithm, Computational study

1. Introduction

The patronizing behavior of customers is one of the core inputs in many business and economic indicators. This is the case for the estimation of the market share captured by the facilities in a competitive environment. When there exist several facilities offering the same product, the way in which customers decide to spend their buying power among them may determine the success or the failure of the facilities. Consumer behavior is a function of the attraction that consumers feel for the facilities. This attraction is the result of several factors, but the two most important forces are the location of the facilities and their quality.

Deciding the location and the quality of the new facilities to be opened is a hard task. There are many competitive location models in literature for that purpose. See for instance Eiselt and Laporte (1996), Eiselt et al. (1993), Plastria (2001) and the references therein. Depending on the location space, the number of facilities to be located, the type of competition, the demand and the way the attraction is measured, different location models have been proposed. Recently, in some of them, the quality of the new facility is also included as a variable to be determined (see Aboolian et al. (2007), Drezner et al. (2012), Fernández et al. (2007), Redondo et al. (2012, 2009b), Saidani et al. (2012)).

Of course, the patronizing behavior of customers is also taken into account in location models, and this is the topic this paper is devoted to. Two choice rules are usually employed in literature. In the *deterministic rule* consumers are assumed to patronize only one facility, the one for which customers feel more attraction. However, this hypothesis has not found much empirical support, except in areas where shopping opportunities are limited and transportation is difficult. On the contrary, in the *probabilistic rule*, consumers patronize all the existing facilities, although the amount spent at each of them is different, and depends on the attraction of the consumers towards the particular facility. In general, this second rule has proved to approximate the market share captured by the facilities more accurately than other alternatives.

However, does a customer really go to *all* the available facilities offering the product to satisfy his/her needs? To what extent does this assumption have an influence on the selection of the location and the quality of the new facilities to be located? This is the research question studied in this paper. If the attraction that a demand point feels towards a facility is strictly positive, then, no matter how small that attraction is, according to the probabilistic rule the facility will capture part of the demand of the demand point. This may be unrealistic in some cases.

The idea of patronizing only some of the existing facilities is not new. For example, in marketing literature, some theories of consumer behavior involve thresholds or discontinuities. For instance, in Gilbride and Allenby (2004) the authors investigate consumers' use of screening rules as part of a discrete-choice model and they present a model that accommodates conjunctive, disjunctive, and compensatory screening rules. However, in the location literature this topic has been hardly investigated. We are only aware of three location papers. In the first one, the so-called Pareto-Huff rule (see Peeters and Plastria (1992)) was introduced. According to it, a customer patronizes a more distant facility only if that facility has a higher quality. The second paper Suárez-Vega et al. (2007) introduces another variant, in which customers only patronize a facility provided that it offers a minimum attraction to them. Both papers deal with *network* location problems. The third paper is Ghosh and Craig (1991), where the related concept of reservation distance is used in a *discrete* location model. In all the three papers it is assumed that the quality of the new facilities to be located are fixed and given beforehand.

In this paper we follow the idea in Suárez-Vega et al. (2007), i.e., customers at each demand point have associated a minimum level of attraction in order to patronize a facility, and then they share their buying power among the facilities that pass this threshold. However, this paper differs from Ghosh and Craig (1991), Peeters and Plastria (1992) and Suárez-Vega et al. (2007) in two key points: (i) we analyze location problems in the plane, and (ii) the quality of the new facilities is not given beforehand, but it is considered to be an additional variable of the problem to be determined. Both things make the problem much more challenging from the optimization point of view. It also differs from Peeters and Plastria (1992) and Suárez-Vega et al. (2007) in two additional points: (iii) the function of the distance used in the attraction function does not have to be necessarily concave, and (iv) the locating chain does not have to be a new entrant in the market, it may already own some of the existing facilities. The variant of the probabilistic rule used in this paper, which will be referred to as *partially probabilistic choice rule* hereinafter, is in many cases a more realistic assumption than that of the probabilistic rule. Notice that when the attraction threshold is very low, this rule is similar to the proportional one. For higher threshold values this procedure may coincide with the deterministic rule. For intermediate threshold values, the customers' choices may be different from both cases.

The use of attraction thresholds is related to the constrained choice-set concept used in Thill (2000) for the analysis of non-cooperative competition between multi-branch networks when consumers have heterogeneous preferences. The implications of thresholds has also been investigated in some discrete choice models studies from different perspectives, considering them as minimum perceptible changes in attributes required for a decision maker (DM) to differentiate among alternatives (see for example Cantillo et al. (2010, 2006), Cantillo and Ortúzar (2006)). Those studies reveal that where perception thresholds exist in the population, the use of models without them leads to errors in estimation and prediction. However, in those papers the purpose is to analyze how the thresholds affect the decisions made by DMs, whereas in this paper we assume that perception thresholds exist in the customers and want to investigate its final influence in the location of the new facility.

In the location literature the use of thresholds is usually related to the distance. For example, in the location of some emergency facilities, the emergency facility is assumed not to provide a good response to a demand point if it is at a distance greater than a given threshold (see for instance Holmes et al. (1972) or Li et al. (2011)). In the location of some undesirable facilities, the undesirable effects produced by the facility are supposed to be negligible for a demand point if the facility is further than a given threshold distance (see for instance Plastria and Carrizosa (1999) or Yapicioglu et al. (2007)). In covering location problems, a demand point is assumed to be covered by a facility if their interdistance is less than a given threshold (see Berman et al. (2010) and references therein). There are also competitive location papers where thresholds are used. For instance, there are location models in which the facilities to be located are required to capture a minimum level of demand (see for instance Colomé et al. (2003)). When the deterministic choice rule is used, the location of the competitive facilities is fixed, and only the location is to be determined, then profit changes only occur at particular threshold qualities (see Plastria and Carrizosa (2004)). There are also competitive location models where two facilities with the same quality are regarded as similar for a customer if the difference between the distances from the customer to the facilities is less than a given threshold (see Kress and Pesch (2012) for a survey of that type of models on networks). The requirement of minimum level of attraction is also suggested in Arenoe et al. (2015) (apart from Thill (2000) and Suárez-Vega et al. (2007)).

Notice that if for a given demand point none of the facilities attains the minimum attraction level, then the demand at that demand point will not be served. Hence, in our model, it may exist some unmet demand. This feature also makes our model different from most of the models in the literature, where it is usually assumed that the whole demand is fully served. The most remarkable exceptions are the papers by Drezner and Drezner (2008, 2012), where the authors try to model the lost demand in a competitive environment similar to the one used in our paper (the probabilistic choice rule is explicitly considered in those two papers). However, the way in which lost demand is considered in those papers is different from the way it is done in our model. In those papers, (i) *all* customers are *partially* served (part of their demand is not served) and (ii) the demand actually served at any given demand point is served *from all* the existing facilities. Furthermore, (iii) in Drezner and Drezner (2008) an exponential distance decay function is assumed to model the disutility of the facility as the distance between the demand point and the facility increases. On the contrary, in our model (i) some customers are fully served, whereas others may be not served at all, (ii) the demand at a given demand point may be served by only some of the existing facilities (not necessarily all) and the facilities which serve a demand point depend on the demand point, and (iii) a general distance decay function can be used in our model (in the computational studies we use a particular one, namely, a power of the distance, but other types of distance decay functions could be used as well, and the algorithms proposed in this paper can also handle them).

As we will see, the partially probabilistic choice rule can be modeled in continuous location problems without major difficulties. And even though the resulting problems are usually hard-to-solve global optimization problems (perhaps one of the reasons for which this choice rule has not been used before in continuous problems), they can be handled with the techniques presented in this paper. The aim of this paper is to study whether the location and quality of the new facilities to be located suggested by the location models differ depending on the choice rule employed.

The paper is organized as follows. The classical probabilistic choice rule is reviewed in the next section in the context of a continuous competitive facility location and design model. The partially probabilistic choice rule is then introduced in the next section. Different approaches to solve the corresponding location problem are presented in Section 4. Some implementation issues related to the estimation of the parameters of the model are discussed in Section 5. The study of the influence of a particular customer choice rule in the decision about the optimal location and quality of a new facility is given in Section 6, whereas Section 7 shows a sensitivity analysis of the model to some of its parameters. Section 8 reports some computational studies to investigate the effectiveness and efficiency of the methods proposed. The paper ends with Section 9 offering some conclusions and pointing lines for future research.

2. The probabilistic choice rule

A demand point aggregates the demand of all the consumers that it represents. The probabilistic choice rule assumes that the demand at a demand point (or the buying power concentrated at it) is split probabilistically over *all* the facilities in the market proportionally to his/her attraction to each facility. Alternatively, it can be interpreted that customers patronize *all* the existing facilities (not just the most attractive to them), but the amount spent at each facility is proportional to the attraction that the customers feel towards the facility. This is the interpretation used throughout this paper, as well as in most of the location papers in the literature.

We will assume the following particular scenario (see Fernández et al. (2007) for more details). A single new facility is going to be located in a given region of the plane by a chain. There already exist other facilities around selling the same goods or product. Some of those facilities may belong to the locating chain. The demand to be served, known and fixed, is concentrated at some demand points, whose locations are known. The location and quality of the existing facilities are also known. The attraction function of a demand point towards a facility is modeled as perceived quality divided by perceived distance.

The maximization of the *profit* obtained by the chain after the location of the new facility is the objective to be achieved, to be understood as the income due to the market share captured by the chain minus its operational costs. The aim is to find both the location and the quality of the new facility to be located.

The following notation will be used in the mathematical formulation of the location models:

Indices

- i index of demand points, $i = 1, \ldots, i_{\max}$.
- j index of existing facilities, $j = 1, \ldots, j_{\max}$ (we assume that the first k of those $j_{\max}$ facilities belong to the locating chain ($0 \leq k < j_{\max}$).

Variables

- x location of the new facility, $x = (x_1, x_2)$.
- α quality of the new facility.

Data

- p_i location of demand point i.
- w_i demand (or buying power) concentrated at p_i.
- f_j location of existing facility j.
- d_{ij} distance between demand point p_i and facility f_j.
- α_j quality of facility f_j.
- $g_i(\cdot)$ a non-negative non-decreasing function.
- u_{ij} attraction that p_i feels for f_j. In this paper, $u_{ij} = \alpha_j / g_i(d_{ij})$.
- $d_i^{\min}$ minimum distance from p_i at which the new facility can be located.
- $\alpha_{\min}$ minimum level of quality.
- $\alpha_{\max}$ maximum level of quality.
- S region of the plane where the new facility can be located.
- G^{ex} fixed costs due to the existing chain-owned facilities.

Miscellaneous

- $d_i(x)$ distance between demand point p_i and the new facility.
- $u_{i0}(x, \alpha)$ attraction that p_i feels for the new facility; $u_{i0}(x, \alpha) = \alpha / g_i(d_i(x))$.

When a probabilistic rule is used, the market share captured by the chain is given by

$$M_P(x, \alpha) = \sum_{i=1}^{i_{\max}} w_i \frac{u_{i0}(x, \alpha) + \sum_{j=1}^{k} u_{ij}}{u_{i0}(x, \alpha) + \sum_{j=1}^{j_{\max}} u_{ij}}.$$

In order to determine the optimal location and quality for the new facility, the problem to be solved is

$$\left\{ \begin{array}{ll} \max & \Pi_P(x, \alpha) = F(M_P(x, \alpha)) - G(x, \alpha) - G^{ex} \\ \text{s.t.} & d_i(x) \geq d_i^{\min} \ \forall i \\ & \alpha \in [\alpha_{\min}, \alpha_{\max}] \\ & x \in S \subset \mathbb{R}^2 \end{array} \right. \tag{1}$$

where $F(\cdot)$ is a strictly increasing differentiable function which transforms the market share into expected sales, $G(x, \alpha)$ is a differentiable function which gives the operating cost of a facility located at x with quality α, and $\Pi_P(x, \alpha)$ is the profit obtained by the chain. The constraints $d_i(x) \geq d_i^{\min}$, with $d_i^{\min} > 0$, are included for technical reasons. They avoid the new facility to be located on a demand point. Notice that if the new facility is located exactly at a demand point, then $d_i(x) = 0$ and we will have divisions by 0 in the objective function. Apart from that, notice that the values $d_i^{\min}$ can be set arbitrarily small. Alternatively, the distance corrected function suggested in Drezner and Drezner (1997) could be used.

In this paper the function F is assumed to be linear, $F(M_P(x, \alpha)) = c \cdot M_p(x, \alpha)$, where c is the income per unit of goods sold. And for simplicity it is also assumed that c is the same across different facilities (which may not be true in the real world cases). Of course, other functions can be more suitable depending on the real problem considered (see Fernández et al. (2007)). Function $G(x, \alpha)$ should increase as x approaches one of the demand points, since it is rather likely that at around those locations the operational costs of the facility will be higher (due to the value of land and premises, which will make the cost of buying or renting the location higher). On the other hand, G will be in many cases a non-decreasing and convex function in the variable α, since the more quality we require of the facility, the higher the costs will be, at an increasing rate. But other more general functions could be considered if appropriate. In the problems solved in this paper we have assumed G to be separable, i.e. of the form $G(x, \alpha) = G_1(x) + G_2(\alpha)$, with $G_1(x) = \sum_{i=1}^{i_{\max}} \Phi_i(d_i(x))$, with $\Phi_i(d_i(x)) = w_i/((d_i(x))^{\varphi_{i0}} + \varphi_{i1})$, $\varphi_{i0}, \varphi_{i1} > 0$ given parameters, and $G_2(\alpha) = e^{\frac{\alpha}{\beta_0} + \beta_1} - e^{\beta_1}$, with $\beta_0 > 0$ and β_1 given values (see Fernández et al. (2007) for more details and other possible expressions). And the Euclidean distance has been used to measure distances between points.

3. The partially probabilistic choice rule

Although in some cases it may be true that the demand concentrated at the demand point is split among all the existing facilities (especially when the number of facilities is small and their attraction is high), there are other cases in which this is no longer true and the probabilistic choice rule may not represent customer behavior properly. We propose to use a modification, which will be called the *partially probabilistic* choice rule, according to which the demand concentrated at a demand point is split only among those facilities which have a minimum level of attraction, and the demand is split only among those facilities, proportionally to their attraction. Alternatively, it can be interpreted that customers patronize all the facilities which have a minimum level of attraction.

In order to give a mathematical formulation, we consider the same location scenario and notation as in the previous section. Let us denote $\overline{u}$ the minimum level of attraction that a facility must have for a customer to spend some of his/her buying power there. For simplicity, we assume that minimum level to be the same for all the demand points, but we could have a different value $\overline{u}_i$ for each demand point i. Let us define

$$\tilde{u}_{ij} = \begin{cases} u_{ij} & \text{if } u_{ij} \geq \overline{u} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\tilde{u}_{i0}(x,\alpha) = \begin{cases} u_{i0}(x,\alpha) & \text{if } u_{i0}(x,\alpha) \geq \overline{u} \\ 0 & \text{otherwise} \end{cases}$$

Then, the market share captured by the chain with the modified choice rule is

$$M_{PP}(x,\alpha) = \sum_{\{i:\max\{u_{i0}(x,\alpha),\max\{u_{ij}:j=1\ldots,j_{\max}\}\}\geq\overline{u}\}} w_i \frac{\tilde{u}_{i0}(x,\alpha) + \sum_{j=1}^{k} \tilde{u}_{ij}}{\tilde{u}_{i0}(x,\alpha) + \sum_{j=1}^{j_{\max}} \tilde{u}_{ij}}. \quad (2)$$

Notice that if for a given demand point i

$$\max\{u_{i0}(x,\alpha), \max\{u_{ij} : j = 1\ldots, j_{\max}\}\} < \overline{u},$$

then the demand at i is not served by any facility. Hence, in this model, some of the demand may be unmet. Notice that, in particular, this means that the model is suitable only for inessential goods.

The corresponding continuous competitive facility location and design problem is the same as (1), except that $M_P(x,\alpha)$ is replaced by $M_{PP}(x,\alpha)$ as given by (2). Accordingly, its objective function will be denoted by Π_{PP}.

It is known that (1) is a hard-to-solve global optimization problem, with many local maximum points which are not global optimal points (Fernández et al. 2007). The corresponding problem with the modified choice rule is even harder, as in addition to this, it also has many discontinuities. As an example, consider Figure 1, which gives the graphs of the objective function on the

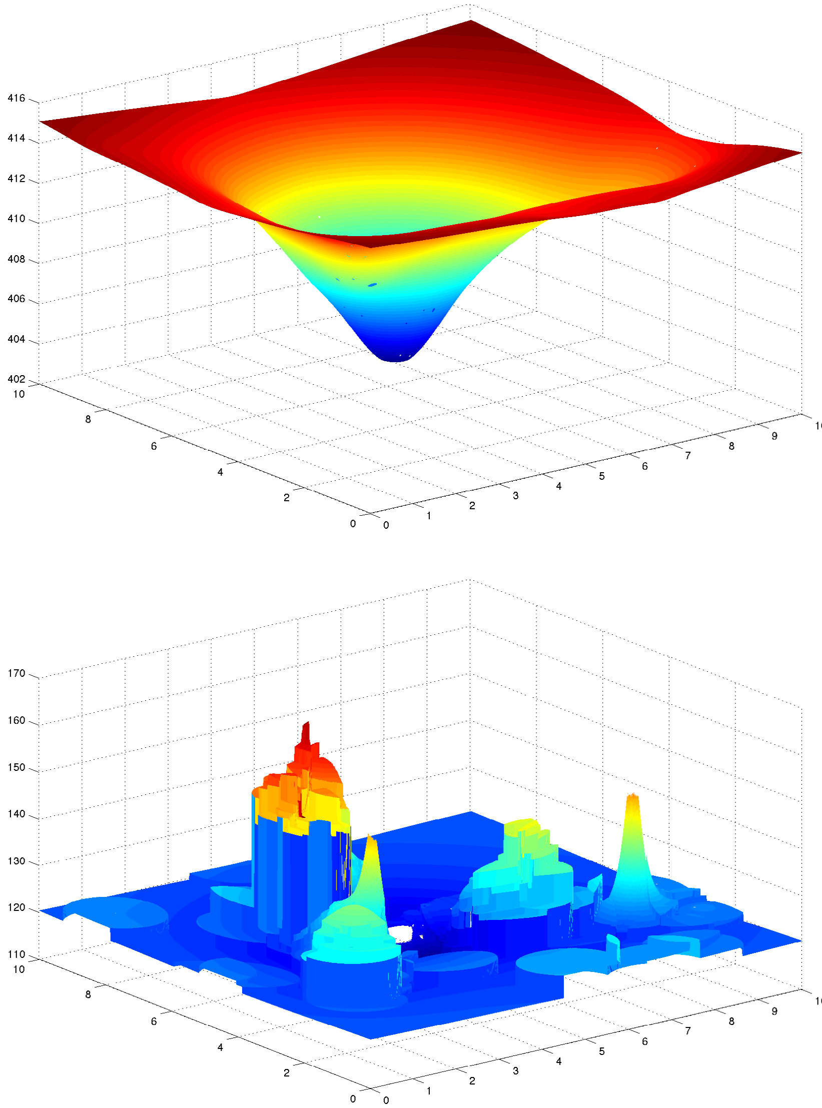

Figure 1: Objective function of an instance with setting ($i_{max} = 71, j_{max} = 5, k = 2$) when $\alpha = 1.545898$: in the top figure, using the probabilistic rule, in the bottom figure with the partially probabilistic choice rule with $\overline{u} = 2$.

location domain for a problem with setting ($i_{max} = 71, j_{max} = 5, k = 2$) when the variable α is fixed to its optimal value for the partially probabilistic problem: in the top figure, using the probabilistic choice rule, and in the

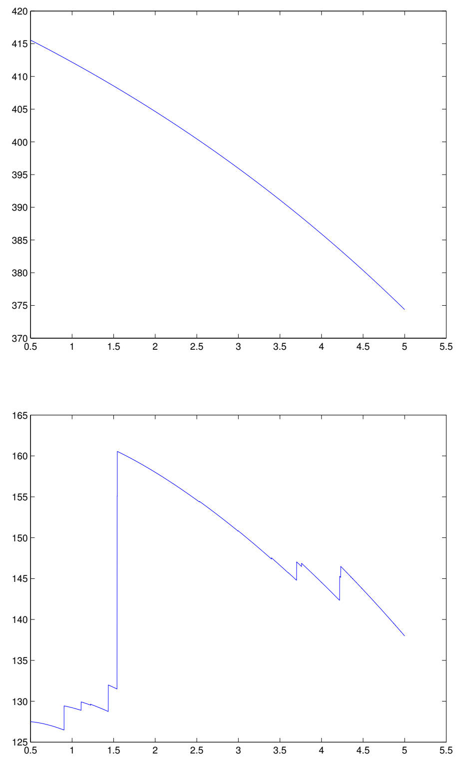

Figure 2: Objective function of an instance with setting ($i_{max} = 71, j_{max} = 5, k = 2$) when ($x_1 = 3.989257, x_2 = 7.065429$): in the top figure, using the probabilistic rule, in the bottom figure with the partially probabilistic choice rule with $\overline{u} = 2$.

bottom figure, with the partially probabilistic rule when $\overline{u} = 2$. The white holes in the graphs correspond to the forbidden regions around the demand points. As can be seen, both problems are highly nonlinear and require global optimization techniques to be solved, although the discontinuities of the problem with the partially probabilistic choice rule, due to the capture or loss of new customers, make it much more challenging. Something similar can be seen in Figure 2, where the graphs of the objective functions are shown when the location of the new facility is fixed at the optimal location of the partially probabilistic problem. As it can be seen, whereas for the probabilistic choice rule the function is differentiable and concave with regard to α (top picture), this is no longer true for the problem with the partially probabilistic choice rule (bottom picture).

4. Solving the partially probabilistic location model

Both rigorous (Fernández et al. 2007) and heuristic (or incomplete) methods (Redondo et al. 2009a, 2008) have been proposed to cope with the probabilistic model. By a rigorous method we mean a deterministic method which can obtain with certainty the global optimal solution of the problem within prescribed tolerances, even with approximate computations (see Neumaier (2004) for a taxonomy of global optimization strategies). Hence, in this section we only concentrate on the partially probabilistic model. Next, both rigorous and heuristic methods are suggested to cope with it.

4.1. A rigorous interval branch-and-bound method, iB&B

Among the rigorous methods to cope with global optimization problems, branch-and-bound (B&B) algorithms are among the most used. Their success relies on the goodness of the bounds they employ, although obtaining those bounds is sometimes a difficult task too. Interval analysis tools can be used both to compute bounds automatically (thanks to the use of *inclusion functions*, see the definition below) and to speed up the process by discarding suboptimal regions (using the so-called *discarding tests*). The interested reader will find the books by Hansen and Walster (2004) and Tóth and Fernández (2010) very helpful. Here we just introduce some notation and a few basic concepts used later on.

The set of compact intervals will be denoted by $\mathbb{IR}$, intervals in boldface letters, and lower and upper bounds of intervals by 'underlines' and 'overlines', respectively. The width of an interval $\boldsymbol{y} = [\underline{y}, \overline{y}] \in \mathbb{IR}$ will be denoted by $w(\boldsymbol{y}) = \overline{y} - \underline{y}$ and its relative width by $w_{relat}(\boldsymbol{y}) = w(\boldsymbol{y}) / \max\{1, \min\{|y| : y \in \boldsymbol{y}\}\}$. The width of an interval vector $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)^T \in \mathbb{IR}^n$ (also called a *'box'*) is to be understood as $w(\boldsymbol{y}) = \max\{w(\boldsymbol{y}_i) : i = 1, \ldots, n\}$.

Definition 1. *A function* $\boldsymbol{f} : \mathbb{IR}^n \to \mathbb{IR}$ *is said to be an* inclusion function *for* $f : \mathbb{R}^n \to \mathbb{R}$ *provided*

$$\{f(y) : y \in \boldsymbol{y}\} \subseteq \boldsymbol{f}(\boldsymbol{y})$$

for all boxes $\boldsymbol{y} \subset \mathbb{IR}^n$ *within the domain of* f.

Observe that if $\boldsymbol{f}$ is an inclusion function for f then we can directly obtain lower and upper bounds of f over any box $\boldsymbol{y}$ within the domain of f, hence its usefulness in B&B methods. There exist programming languages and libraries which have the interval arithmetic implemented (see URL (2017)), and they provide inclusion functions for predeclared functions (sin, exp,...). By using them, the *natural interval extension* or other techniques can be used to compute an inclusion function for other general functions (Hansen and Walster 2004, Tóth et al. 2007).

Interval B&B methods (iB&B in what follows) have been successfully applied to solve location problems (see for instance Fernández and Pelegrín (2001), Tóth and Fernández (2010) and the references therein). In particular, in Fernández et al. (2007) such an iB&B method was applied to solve the corresponding probabilistic problem described in Section 2. A similar method can handle the partially probabilistic model, thanks to the use of the interval tools employed to compute the bounds.

A point that deserves to be clarified here is how to compute an inclusion function $\tilde{\boldsymbol{u}}_{i0}(\boldsymbol{x},\boldsymbol{\alpha})$ for $\tilde{u}_{i0}(x,\alpha)$, as this function is given by an 'if' condition. This can be achieved by defining

$$\tilde{\boldsymbol{u}}_{i0}(\boldsymbol{x},\boldsymbol{\alpha}) = \begin{cases} 0 & \text{if } \overline{\boldsymbol{u}_{i0}(\boldsymbol{x},\boldsymbol{\alpha})} < \overline{u} \\ \boldsymbol{u}_{i0}(\boldsymbol{x},\boldsymbol{\alpha}) & \text{if } \underline{\boldsymbol{u}_{i0}(\boldsymbol{x},\boldsymbol{\alpha})} \geq \overline{u} \\ \left[0, \overline{\boldsymbol{u}_{i0}(\boldsymbol{x},\boldsymbol{\alpha})}\right] & \text{otherwise} \end{cases}$$

where $\boldsymbol{u}_{i0}(\boldsymbol{x},\boldsymbol{\alpha})$ is an inclusion function for $u_{i0}(x,\alpha)$. One can see that this is the cause of the discontinuity of the objective function, and it makes impossible to use any advanced tools that use differentiability or even continuity inside an iB&B. Therefore, only the basic discarding tests described in Fernández et al. (2007) can be used for the new problem, namely, the cut-off and the feasibility tests. It is still a valuable result that a discontinuous function can be optimized by a reliable method.

The output of the method is a list of 3-dimensional boxes, $\mathcal{L} = \{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_r\}$, which contain any global optimal solution. The best point found by the algorithm during its execution, $y^{best} = (x_1^{best}, x_2^{best}, \alpha^{best})$, is also offered as output. Let us denote by Π either Π_P or Π_{PP} depending on the problem to be solved. Notice that any point in $\bigcup_{l=1}^{r} \boldsymbol{y}_l$ will have its objective function value within the interval

$$[\min\{\underline{\boldsymbol{\Pi}(\boldsymbol{y}_l)} : l = 1,\ldots,r\}, \max\{\overline{\boldsymbol{\Pi}(\boldsymbol{y}_l)} : l = 1,\ldots,r\}]$$

although the optimal objective function value Π^* of our problem is within the narrower interval

$$[\underline{\boldsymbol{\Pi}(y^{best})}, \max\{\overline{\boldsymbol{\Pi}(\boldsymbol{y}_l)} : l = 1,\ldots,r\}].$$

The stopping rule sends a box $\boldsymbol{y}$ to the solution list $\mathcal{L}$ provided that

$$w(\boldsymbol{y}) < \epsilon \text{ or } w_{relat}(\boldsymbol{\Pi}(\boldsymbol{y})) < \epsilon.$$

In the computational studies carried out in this paper we have established that $\epsilon = 0.0001$.

4.2. An evolutionary algorithm

In Redondo et al. (2009a), the algorithm called UEGO was proposed for solving the corresponding model with the probabilistic choice rule. In fact, it has also been applied to other competitive location problems as well (see for instance Redondo et al. (2012)). UEGO can be classified as a global evolutionary optimization method. It applies procedures based on nature on a pool of M *independent* candidate solutions (individuals), which form a population. In other words, it applies methods, as reproduction or selection, to address the population towards the global (or local) optima. In this sense, UEGO can also be considered as a multimodal optimization algorithm, since each candidate solution in the population is intended to occupy a local or global optimum of the fitness landscape.

Each candidate solution in UEGO represents a subspace (in fact, a hypersphere) of the whole searching region by means of a radius. The main goal of the radius is to focus the optimization efforts on a particular sub-area. The solution is considered to be in the center of the corresponding subspace. During the optimization procedure, several candidate solutions with different radii can coexist simultaneously. Therefore, at the same stage of the optimization procedure, new promising regions are systematically analyzed (those with a bigger radius), while others are examined thoroughly (those with a smaller radius). The term "independent" signifies that a candidate solution has the ability of reproducing by itself, i.e. a new offspring can arise from a single individual independently of the rest of the population. It means that many peaks can be investigated in parallel and hence, the effects of the genetic drift can be prevented.

The radius associated to a candidate solution depends on the instant where it was created. Solutions created at the beginning of the search will examine big regions, while solutions created during the latest cycles will focus on smaller promising areas. More precisely, the radius of a candidate solution created at iteration i, is given by a decreasing exponential function that depends on the initial domain landscape, R_1, and an input parameter,

R_L, which represents the smallest radius. This mechanism has been designed to balance *exploration* and *exploitation*.

UEGO works as follows. Initially, a population of candidate solutions is created. The fitness is used to determine the relative merit of each individual. Once the initial population is obtained, an iterative process is carried out. At each cycle, UEGO mimics natural evolution on the population by applying procedures for the reproduction of individuals, the improvement of their offspring by means of a local optimizer, and the promotion of the best candidate solutions to the next generation. As UEGO applies a local search algorithm to each individual of the population, it can be also classified as a *memetic algorithm* (Moscato (1989)). In fact, only the local search procedure used within UEGO needs to be adapted for each particular problem.

UEGO has four user input parameters: (i) a maximum number of function evaluations for the whole optimization procedure, N; (ii) a maximum number of individuals in the population, M; (iii) a maximum number of cycles or iterations, L; and (iv) the radius of the smallest candidate solution R_L. The function evaluations, N, are distributed among the individuals in the population at each iteration, in such a way each one has a budget to generate new candidate solutions and to improve them. These budgets are mathematically computed by means of equations that depends on the previously mentioned input parameters. See Redondo et al. (2009a) for a more detailed description of the algorithm and its parameters.

In view of its success at solving different competitive location problems, UEGO has also been adapted to the problem with the partially probabilistic choice rule. The most important adjustment consisted of selecting an appropriate local optimizer. Initially, a Weiszfeld-like algorithm was considered as a local optimizer, following the lines in Redondo et al. (2009a). However, the obtained results were not as good as expected and it was rejected. The general purpose stochastic hill climber SASS (Solis and Wets 1981) was finally included in this work, but with some modifications, which are briefly described next.

Algorithm SASS is a derivative-free optimization algorithm that can be applied to maximize an arbitrary function over a bounded subset of $\mathbb{R}^N$, although internally SASS assumes that the range in which each variable is allowed to vary is the interval $[0, 1]$. Since this is not our case, when necessary we use functions to rescale (normalize) the variable values to the interval $[0, 1]$, and to invert (denormalize) this process. In SASS, the new points are generated using a Gaussian perturbation $\xi \in \mathbb{R}^3$ over the search point (x, α) and a normalized bias term $b \in \mathbb{R}^3$ to direct the search. The standard deviation σ specifies the size of the sphere that most likely contains the perturbation vector. In this work, its upper bound σ_{ub} should have the same value than the normalized radius of the caller solution. Then, the parameter σ_{ub} is also considered an argument of SASS. Hence, any single step taken by the optimizer is no longer than the radius of the calling candidate solution. Finally, the stopping rules are determined by a maximum number of iterations ($ic_{\max}$) and by the maximum number of consecutive failures ($Maxfcnt$). The use of SASS as the local optimizer within UEGO has also worked fine for other location problems with discontinuities (see for instance Redondo et al. (2013)).

The parameters that control UEGO have been tuned to this new problem by trying several combinations of parameter values on a reduced set of random problems. Finally, they have been set to $L = 15$, $R_L = 0.5$, $M = 100$ and $N = 3 \cdot 10^6$.

5. On the estimation of the parameters of the model

In order to apply the location model to a particular real problem, the parameters of the model have to be estimated. Next we discuss how that estimation can be carried out.

The distances d_{ij} between demand points and existing facilities can be obtained directly using road maps or the shortest distance capabilities of Geographic Information Systems (GIS) or Geographic Positioning Systems (GPS). Then, with those data, a distance predicting function (DPF) can be tailored to the geographical region under consideration (Brimberg and Love 1995, Fernández et al. 2002), which provides an estimation of the travel distance between two points, given their geographical coordinates. That DPF can be used as $d_i(x)$.

On the other hand, the function $g_i(d)$ modulating the distance is usually of the form $g_i(d) = d^\lambda$ or $g_i(d) = \exp(-\lambda d)$. For the simultaneous estimation of the parameter λ defining $g_i(d)$ and the qualities of the existing facilities, ordinary least squares could be used as proposed in Nakanishi and Cooper (1974, 1982). However, the strategies proposed in Drezner (2006) or Drezner and Drezner (2002) provide similar results and are easier to implement.

The functional forms of F and G should also be ascertained so that they reflect the reality as close as possible. Whereas for the income function F this seems to be an easy task (considering the experience of the managers of the locating chain), the functional form of G could be more tricky. The difficulty of the determination of those functions will depend on the particular type of facility to be installed, and on the ability and knowledge of the managers.

The estimation of the threshold values is also a critical issue, as we will see in Section 6. Using the probabilistic choice rule, the proportion of the demand concentrated at p_i captured by facility j is given by

$$p_{ij} = \frac{u_{ij}}{\sum_{k=1}^{j_{\max}} u_{ik}}.$$

If we assume that when p_{ij} is less than a predefined small value $\tilde{p}$ (for instance, let say $\tilde{p} = 0.05$), then it is not 'normal' for a customer at p_i to go to f_j, then we can consider that the threshold for demand point p_i, denoted by $\overline{u}_i$ (the minimum level of attraction that a facility must have for the customers at p_i to spend some of his/her buying power there), lies in the interval, $[\overline{u}_i^{lb}, \overline{u}_i^{ub}]$, where

$$\overline{u}_i^{lb} = \max_j \{u_{ij} : p_{ij} < \tilde{p}\}, \text{ and } \overline{u}_i^{ub} = \min_j \{u_{ij} : p_{ij} \geq \tilde{p}\}.$$

That interval can be computed easily using the available data and parameters. However, in order to choose the value $\overline{u}_i \in [\overline{u}_i^{lb}, \overline{u}_i^{ub}]$ that better fits the tastes of the customers at p_i more information is needed. Let us denote by f_{lb_i} the facility where $\overline{u}_i^{lb}$ is attained, and by f_{ub_i} the one for which $\overline{u}_i^{ub}$ is attained. A survey could be carried out among the customers at demand point i asking: (i) how much closer should the facility f_{lb_i} be so you would go there to buy goods or use its services?, and (ii) how much further should the facility f_{ub_i} be so you would not go there? The average distance of the answers of each of those questions is then computed, and from those two distances, two attraction thresholds for p_i can be computed. The mean value of those thresholds could be set as $\overline{u}_i$.

If a unique threshold value $\overline{u}$ is required for all the demand points, then the minimum of the $\overline{u}_i$ values, or its mean, or its weighted mean could be used.

6. Probabilistic choice rule versus partially probabilistic choice rule

To what extent is the difference between the probabilistic and the partially probabilistic choice rules important when deciding the location of the new facility? How much does it affect the optimal profit? We study these and other issues first using a quasi-real example of the location of a hypermarket, and then analyzing some random problems. In both studies we will solve the same location problems using both rules, and using the exact interval branch-and-bound method mentioned in Subsection 4.1, so as to have the optimal solutions and carry out a fair comparison.

6.1. A case study

In this subsection, we investigate a quasi-real example dealing with the location of a hypermarket in an area around the city of Murcia, in south-eastern Spain. In the study, we have considered a working radius of 25 km around Murcia. 632558 inhabitants live within the circle, distributed over $i_{\max} = 71$ population centers. Their population varies between 1138 and 178013 inhabitants. Each population center has been considered a demand point, with buying power proportional to its total population (one unit of buying power per 17800 inhabitants), resulting the demand's spectrum in $[0.064, 10]$. Their location and population can be seen in Figure 3: each demand point is shown as a gray circle (or a black dot) with a radius proportional to its buying power. The surrounding gray circles also show the forbidden regions around the cities. Five hypermarkets are present in the area: two from a first chain A (marked with a red $\bullet$), and three from another chain B (marked with a green $\times$). Figure 3 shows the location of each hypermarket. The feasible set S was considered the smallest rectangle containing all demand points (approximately a square centered at Murcia whose sides have a length close to 45 Km).

The coordinates of the population centers and the hypermarkets were re-scaled to an approximate square $([0, 10], [0, 10])$, in which the units correspond approximately to 4.5 Km. The radius of the forbidden circular area surrounding demand point p_i was set to $w_i/30$.

Approximate values for the quality parameters α_j for the five existing hypermarkets were obtained through a small survey among people who had visited all five hypermarkets. Each respondent was asked to rank the hypermarkets in increasing order of their perceived quality and to indicate the difference in quality between any two consecutive hypermarkets in their ranking by a score between 1 (small) and 4 (big). That information yielded individual marks for all hypermarkets, by starting from a lowest mark of 1 and adding each difference score to obtain the mark of the next hypermarket in their ranking. Finally, these marks were rescaled to the interval [3,4], because, according to all the respondents, in a proposed scale from 0.5 to 5, all considered hypermarkets have a quality above the mean (2.75). But they all still have a large margin for improvement. The qualities α_j considered are the average rescaled marks over the respondents. But see Section 5 for a better way to estimate the qualities of existing facilities in real problems. The quality of the new facility lies in the interval $[0.5, 5]$.

For the rest of the parameters, we used the following initial settings: the income per unit of good sold $c = 12$, in the location cost function G_1 we chose the parameter $\varphi_{i0} = 2$ for all i, while the value of φ_{i1} decreases as the population increases. Finally, the parameters of the quality cost function G_2 were initially set to $\beta_0 = 7$ and $\beta_1 = 3.75$. The interested reader is referred to Tóth et al. (2009) for more details about this example.

Three different scenarios have been considered:

1. Scenario 'small chain A': The locating chain is the small one, chain A, which owns $k = 2$ of the $j_{\max} = 5$ existing facilities.
2. Scenario 'large chain B': The locating chain is the greatest one, chain B, which owns $k = 3$ of the 5 existing facilities.
3. Scenario 'newcomer': The locating chain is new in the market ($k = 0$), and the new facility will have to compete with the 5 existing facilities.

The following notation will be used when analyzing the results: $\Pi_P(\cdot)$ is the objective function of location problem (1) when the probabilistic choice rule is employed, $\mathcal{L}_{\mathcal{P}}$ gives the output list of 3-dimensional boxes which contain any optimal solution, and (x_P^*, α_P^*) represents the best point found by the algorithm during the execution. The corresponding items for the problem with the partially probabilistic choice rule will be denoted by $\Pi_{PP}(\cdot)$, $\mathcal{L}_{\mathcal{PP}}$ and $(x_{PP}^*, \alpha_{PP}^*)$.

We will compute the Euclidean distance between the locations x_P^* and x_{PP}^*, which is denoted as $dist_{loc}$, as well as the absolute difference between the qualities α_P^* and α_{PP}^*, denoted as $dist_{qual}$, to measure the difference between the optimal solutions.

The objective function value of the partially probabilistic model before ($\Pi_{PP}(\text{before})$) and after ($\Pi_{PP}(x_{PP}^*, \alpha_{PP}^*)$) the location of the new facility will be shown, as well as the cost due to the new facility, $G(x_{PP}^*, \alpha_{PP}^*)$. We will compute the relative profit loss incurred when the probabilistic choice rule is assumed in a problem where the partially probabilistic rule should have been chosen,

$$lost(P|PP) = 100 \cdot (\Pi_{PP}(x_{PP}^*, \alpha_{PP}^*) - \Pi_{PP}(x_P^*, \alpha_P^*))/\Pi_{PP}(x_{PP}^*, \alpha_{PP}^*),$$

and the relative profit loss incurred when the partially probabilistic choice rule is assumed in a problem where the probabilistic rule should have been chosen,

$$lost(PP|P) = 100 \cdot (\Pi_P(x_P^*, \alpha_P^*) - \Pi_P(x_{PP}^*, \alpha_{PP}^*))/\Pi_P(x_P^*, \alpha_P^*),$$

to measure the cost of choosing the inadequate model for the chain as a whole.

Similarly, to measure the cost of choosing the inadequate model in the profit increment due to the new facility, the relative profit lost due to the new facility only when the probabilistic choice rule is assumed in a problem where the partially probabilistic rule should have been chosen,

$$lost(P|PP)_0 = 100 \cdot \frac{Incr\Pi_{PP}(x_{PP}^*, \alpha_{PP}^*) - Incr\Pi_{PP}(x_P^*, \alpha_P^*)}{Incr\Pi_{PP}(x_{PP}^*, \alpha_{PP}^*)},$$

will be computed, where $Incr\Pi_{PP}(x_{PP}^*, \alpha_{PP}^*) = \Pi_{PP}(x_{PP}^*, \alpha_{PP}^*) - \Pi_{PP}(\text{before})$ and $Incr\Pi_{PP}(x_P^*, \alpha_P^*) = \Pi_{PP}(x_P^*, \alpha_P^*) - \Pi_{PP}(\text{before})$. Analogously, $lost(PP|P)_0$ will be computed too. For the newcomer case these two values are not computed, as they coincide with $lost(P|PP)$ and $lost(PP|P)$, respectively.

As the demand actually served by the facilities may vary depending on the threshold value $\overline{u}$ employed (remember that we may have unmet demand), we will also compute the percentage of the total demand really served before ($\%W_{TB}$) and after ($\%W_{TA}$) the location of the new facility, the percentage of the total demand captured by the locating chain before ($\%W_{CB}$) and after ($\%W_{CA}$) the new facility is located, and the percentage of the total demand captured by the new facility ($\%W_{new}$). Notice that $\%W_{CB} + \%W_{new}$ may be different from $\%W_{CA}$, since the new facility may steal part of the demand to the existing chain-owned facilities (an effect known as cannibalization). The total demand in the region in all the cases is $W = \sum_{i=1}^{i_{\max}} w_i = 35.53$. For the newcomer case, $\%W_{CB}$ and $\%W_{CA}$ are omitted, as the locating chain has no existing facilities. For the sake of completeness, we also show the quality of the new facility for each value of $\overline{u}$.

The results obtained for each of the three scenarios are shown in tables 1, 2 and 3, respectively.

As we can see, in the scenario 'small chain A' (see Table 1), even for the threshold $\overline{u} = 0.5$ (a small value with which 95% of the total demand is still served) we can observe that both the location and the quality of the new facility to be located are different from those of the probabilistic choice

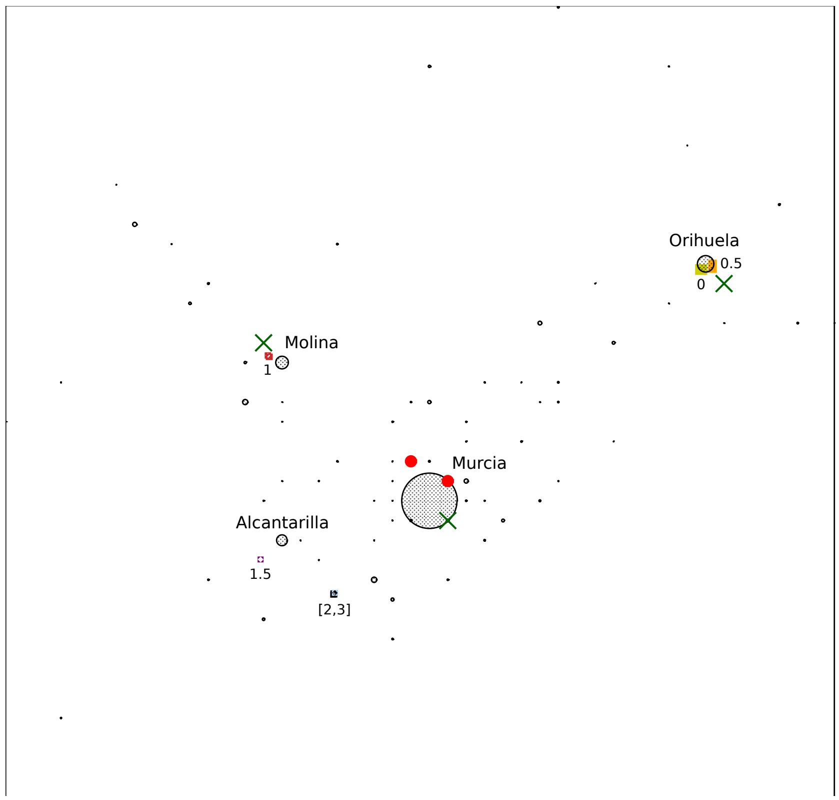

Figure 3: Case study: scenario small chain A.

rule (see Figure 3, where the optimal locations for the different values of $\overline{u}$ are drawn as small squares in different colors). However, the differences are not too important, as the optimal solution is still to locate the new facility close to Orihuela, the second most populated city, in the North-East of the map, where one of the facilities of the competing chain is set up (the chain already has two facilities in the surroundings of the most populated city, Murcia, so this is not the optimal place to expand the chain, due to the cannibalization). In fact, the differences in objective function value are almost negligible. It is for $\overline{u} = 1$ where we can observe a big change, both in location and quality. In this case, the optimal location changes to the third most populated city, Molina (in the North-West of the map, where another facility of the competing chain operates). The reason is that Molina is closer to Murcia, and the new facility not only competes against the competitor's

Table 1: Case study: differences in the solutions obtained by the probabilistic and the partially probabilistic choice rules for the scenario 'small chain A'.

$\overline{u}$	0	0.5	1.0	1.5	2.0	2.5	3.0
quality	1.45	1.24	4.70	0.70	1.49	1.86	2.24
$dist_{loc}$		0.16	5.31	6.35	5.88	5.88	5.90
$dist_{qual}$		0.22	3.24	0.76	0.03	0.41	0.78
Π_{PP}(before)		176.2	151.4	141.1	133.4	127.4	123.6
$\Pi_{PP}(x^*_{PP}, \alpha^*_{PP})$		197.0	182.6	166.9	162.6	157.1	153.6
$G(x^*_{PP}, \alpha^*_{PP})$		12.5	47.4	11.0	17.8	20.6	23.6
$lost(P\|PP)$		0.1	8.3	5.9	8.1	8.7	9.3
$lost(PP\|P)$		0.1	8.0	9.8	11.4	11.6	12.1
$lost(P\|PP)_0$		1.4	48.6	37.9	45.0	46.1	47.4
$lost(PP\|P)_0$		0.8	77.7	94.2	110.2	112.5	117.1
$\%W_{TB}$		95.0	83.3	76.9	76.2	72.2	66.5
$\%W_{TA}$		95.0	91.4	85.5	84.4	82.0	78.9
$\%W_{CB}$		41.3	35.5	33.1	31.3	29.9	29.0
$\%W_{CA}$		49.1	53.9	41.7	42.3	41.7	41.6
$\%W_{new}$		7.9	19.1	8.6	11.0	11.9	12.6

facility in Molina, but also attracts part of the demand from Murcia and Alcantarilla (the fourth most populated city), thanks to its high quality. Notice that when the facility is located in Orihuela, if $\overline{u} = 1$, even a high quality is not enough to attract demand from Murcia. As can be seen in Table 1, when $\overline{u} = 1$ an inadequate choice in the patronizing behavior of customers may lead to a considerable relative profit lost of around 8%. In fact, in this case the new facility captures 19.1% of the total demand, more than 1/3 of the demand captured by the chain. When $\overline{u} = 1.5$ we can observe another change in the location, which moves to the South-West, close to Alcantarilla and closer to Murcia; now the quality of the new facility is small, which reduces the costs, but still allows to capture most of the demand from Alcantarilla and its surroundings. Finally, for $\overline{u} = 2$, 2.5 and 3 a final change can be observed, where the location is the same for all those values, and only the quality changes. Now, although the new facility is at a similar distance from the most populated city, it is set up in a cheaper place and it is closer to the competitor's facility in Murcia, which allows to compete against it and to capture part of the demand from the South-West area (although at the cost of a higher quality.

Note that an increase in the market share captured does not necessarily means an increase in the profit, since the increment in the market share captured can be due to a higher quality or a better and more expensive location, and in both cases this implies a higher cost, and maybe the cost exceeds the income obtained from the market share captured. For instance, from $\overline{u}$ =0.5 to 1.0 the chain capture more market share (from 49.1% to 53.9%, which is an increase of 9.77%), but the profit decreases (from 197.0 to 182.6, a 7.30%) since the cost due to the new facility increases from 12.5 to 47.4, mainly due to the increment in the quality (from 0.22 to 3.24).

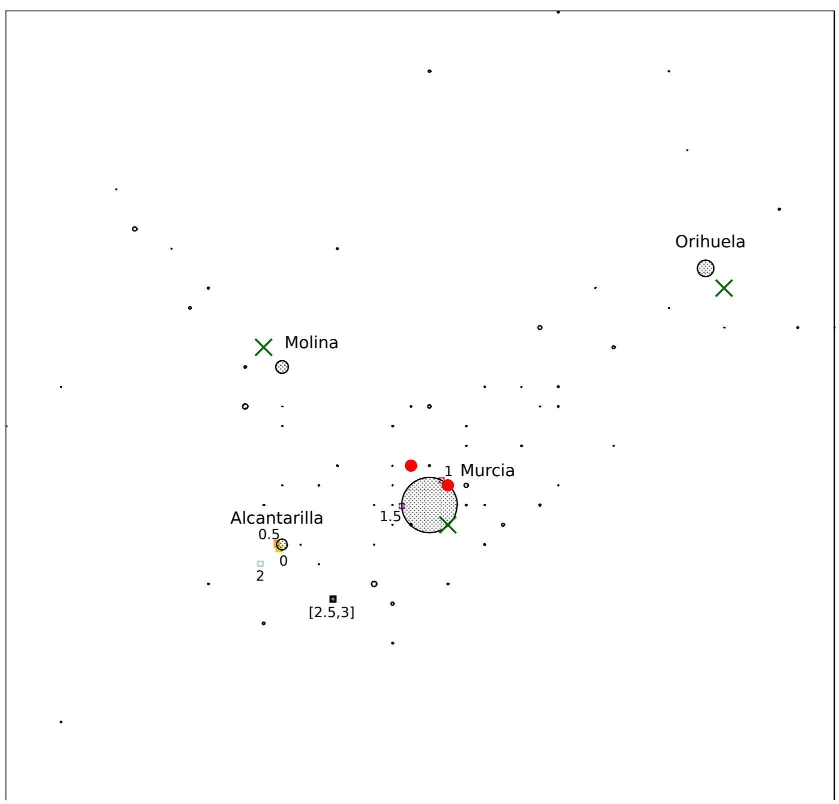

Figure 4: Case study: scenario large chain B.

In the scenario 'large chain B', for both the probabilistic choice rule ($\overline{u}$ =

Table 2: Case study: differences in the solutions obtained by the probabilistic and the partially probabilistic choice rules for the scenario 'large chain B'.

$\overline{u}$	0	0.5	1.0	1.5	2.0	2.5	3.0
quality	0.58	0.50	4.33	3.87	0.92	1.16	2.25
$dist_{loc}$		0.07	2.13	1.57	0.28	0.28	0.89
$dist_{qual}$		0.08	3.75	3.29	0.35	0.58	1.67
Π_{PP}(before)		229.0	203.7	186.7	191.6	180.6	159.9
$\Pi_{PP}(x^*_{PP}, \alpha^*_{PP})$		236.0	236.6	212.5	215.8	203.2	189.4
$G(x^*_{PP}, \alpha^*_{PP})$		10.7	48.8	43.8	12.6	14.2	23.6
$lost(P\|PP)$		0.2	7.2	4.0	4.3	4.5	8.5
$lost(PP\|P)$		0.0	1.6	0.8	1.3	1.4	3.8
$lost(P\|PP)_0$		5.4	51.9	33.3	38.2	40.5	54.4
$lost(PP\|P)_0$		1.2	61.9	30.8	51.2	52.3	147.6
$\%W_{TB}$		95.0	83.3	76.9	76.2	72.2	66.5
$\%W_{TA}$		95.7	91.5	84.6	84.9	80.9	78.9
$\%W_{CB}$		53.7	47.8	43.8	44.9	42.3	37.5
$\%W_{CA}$		57.9	66.9	60.1	53.6	51.0	50.0
$\%W_{new}$		7.4	25.1	22.0	8.6	8.6	12.6

0) and the partially probabilistic choice rule with threshold value $\overline{u} = 0.5$, the optimal solution is to locate the facility close to Alcantarilla, the fourth most populated city, with a very small quality: since there exist no other facilities around, the new facility captures most of its demand (see Figure 4) with a low cost.

However, when $\overline{u} = 1$ the optimal location moves to the surroundings of Murcia, the most populated city, between the two existing facilities of the competing chain, and this, despite the fact that the chain already owns a facility in the South-East of that city; to compete against them, a high quality is required. Notice that with $\overline{u}$ =0.5, the total market share served before the location of the new facility was 95.0%; whereas with $\overline{u}$ =1.0 it is only 83.3%. This means that part of the demand at Murcia (and other cities around) is not served by the existing facilities, and so, there is an opportunity for the new facility to capture that unserved demand. In fact, after the location of the new facility, the total demand served increases to 91.5%. As we can see, in this case, $lost(P|PP) = 7.2\%$, a very high value taking into account that the locating chain is dominant in the market and that 91.5% of the total demand is served after the location of the new facility. In this case, the new facility captures 25.1% of the total demand, whereas the chain, considering all the facilities, gets 66.9% of the total demand, i.e., the new facility captures 37.5% of the demand of the chain. This is, however, at the expense of suffering some cannibalization: notice that $\%W_{CB} + \%W_{new} = 72.9\%$, more than $\%W_{CA} = 66.9\%$. When $\overline{u} = 1.5$ the optimal location moves to another part of the city of Murcia, far from the existing competing facilities, which allows the reduce the quality, and hence, the costs, but still capturing a good part of the demand at Murcia, even though suffering some cannibalization. When $\overline{u}$ =2.0 the location changes again, close to Alcantarilla, again with a low quality. Since the costs are again much smaller, the profit increases, and this despite capturing a smaller percentage of the market. And finally, for $\overline{u} = 2.5$ and 3, a new location is obtained, similar to that of the small chain scenario, not too close to Murcia nor to Alcanrilla to reduce the location costs (see Figure 4), and with a medium quality. Notice that, as expected, the cannibalization decreases as $\overline{u}$ increases.

Table 3: Case study: differences in the solutions obtained by the probabilistic and the partially probabilistic choice rules for the scenario 'newcomer'.

$\overline{u}$	0	0.5	1.0	1.5	2.0	2.5	3.0
quality	5.00	5.00	4.92	4.16	4.37	5.00	2.24
$dist_{loc}$		0.10	0.67	0.04	0.01	0.09	1.29
$dist_{qual}$		0.00	0.08	0.84	0.62	0.00	2.76
Π_{PP}(before)		0.0	0.0	-0.0	0.0	0.0	0.0
$\Pi_{PP}(x^*_{PP}, \alpha^*_{PP})$		45.6	60.9	52.2	50.6	43.3	30.1
$G(x^*_{PP}, \alpha^*_{PP})$		56.7	55.7	46.8	49.3	56.6	23.6
$lost(P\|PP)$		0.9	11.8	3.7	4.1	45.0	17.4
$lost(PP\|P)$		0.7	5.8	3.8	2.3	10.2	64.5
$\%W_{TB}$		95.0	83.3	76.9	76.2	72.2	66.5
$\%W_{TA}$		95.7	91.5	84.6	83.4	80.8	78.9
$\%W_{new}$		24.0	27.3	23.2	23.4	23.4	12.6

Concerning the 'newcomer' case, for $\overline{u} = 0$ to 2.5, the optimal solution is

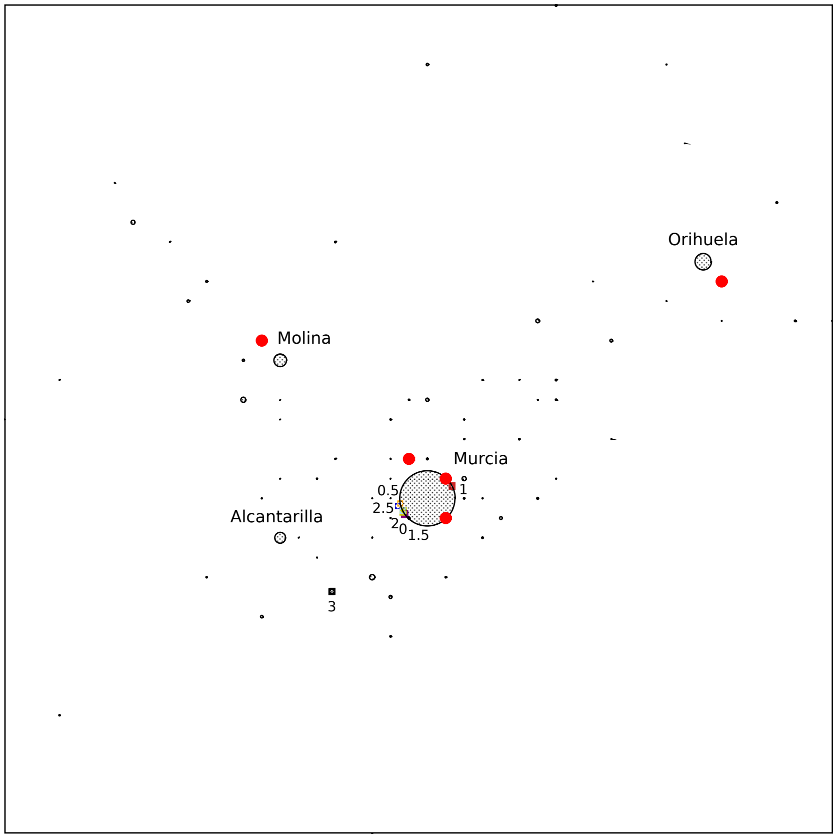

Figure 5: Case study: newcomer.

to locate the new facility in the surroundings of Murcia, the most populated city (see Figure 5). Even though there already exist three facilities around Murcia and there will be fierce competition, for the newcomer it is still the best option, as most of the demand is concentrated there. A high quality is required, though. It is important to highlight, however, that even with very small differences in location and quality, the relative profit loss incurred when an inadequate patronizing behavior of customers is assumed may be very high. See for instance the case $\overline{u} = 2.5$. In this case, the difference in location is very small ($dist_{loc} = 0.09$) and the difference in quality with regard to the probabilistic case is almost negligible. Nevertheless, $lost(PP|P) = 10.2\%$ and $lost(P|PP) = 45.0\%$. This clearly shows the big difference between the probabilistic and the partially probabilistic choice rule. For $\overline{u} = 3$ the relative profit lost is even higher, although in this case the location moves to a different place, the same as that of the small and large chain scenarios when $\overline{u}$ is 2.5 or more (not too close to Murcia nor to Alcanrilla to reduce the location costs), and the quality is also different from that of the probabilistic case. In this case, since $\overline{u}$ is very high, none of existing facilities can capture the demand from Alcantarilla and the cities around, and the new facility can capture it, and with a low quality and in a cheap location.

Summarizing, we can see that both the location and/or the quality of the facility to be located may change drastically as the threshold value varies, and even when those changes are very small, the relative profit loss incurred for the chain when an inadequate choice rule is employed may be very high. This is due to discontinuities of the objective function, see figures 1 and 2: every time a new demand point is captured or lost (and this may happen with a small change in the location and/or the quality), a jump in the objective function happens. This clearly shows that the selection of the choice rule in competitive location models should be made with care, and the assumption of the probabilistic choice rule commonly done in literature should only be considered when it is really the case.

6.2. Random problems

We have done a similar study on a set of 40 random problems, half of them with $n = 500$ demand points and the rest with 1000 demand points, although in this case, for the sake of brevity, only the differences in the location, quality and objective value have been analyzed. Several settings $(i_{\max}, j_{\max}, k)$ have been considered (see Table 4), and for each of them, five problems were generated by randomly choosing the parameters of the problems uniformly within the following intervals:

- $p_i, f_j \in S$,
- $\omega_i \in [0, 100/\sqrt{n}]$,
- $\alpha_j \in [0.4, 6]$,
- $G(x, \alpha) = \sum_{i=1}^{i_{\max}} \Phi_i(d_i(x)) + G_2(\alpha)$ where
    - $\Phi_i(d_i(x)) = w_i \frac{1}{(d_i(x))^{\varphi_{i0}} + \varphi_{i1}}$ with $\varphi_{i0} = \varphi_0 = 2, \varphi_{i1} \in [0.5, 1.5]$,
    - $G_2(\alpha) = e^{\frac{\alpha}{\beta_0} + \beta_1} - e^{\beta_1}$ with $\beta_0 \in [7, 9], \beta_1 \in [6, 6.5]$,

Table 4: Random problems: differences in the solutions obtained by the probabilistic and partially probabilistic choice rules.

$\overline{u}$	$dist_{loc}$	$dist_{qual}$	$lost(P\|PP)$	$lost(PP\|P)$	$lost_0(P\|PP)$	$lost_0(PP\|P)$
			(500,6,1)			
0.5	(1.5,2.9)	(0.0,0.0)	(9.5,13.2)	(2.2,5.6)	(14.6,21.7)	(4.0,8.7)
1.0	(1.6,2.2)	(0.1,0.2)	(21.1,30.7)	(3.4,4.7)	(32.0,44.7)	(7.0,10.5)
2.0	(4.0,7.4)	(0.1,0.2)	(25.3,31.0)	(7.0,11.6)	(42.5,58.3)	(14.2,24.0)
			(500,6,3)			
0.5	(2.7,5.3)	(0.1,0.4)	(7.1,12.0)	(2.6,4.0)	(20.3,35.0)	(15.5,39.0)
1.0	(3.2,6.1)	(0.1,0.3)	(14.3,20.2)	(2.2,4.6)	(34.2,66.0)	(14.2,44.3)
2.0	(3.5,6.9)	(0.2,0.3)	(20.0,27.8)	(2.6,4.9)	(48.5,60.0)	(15.8,47.2)
			(500,12,3)			
0.5	(1.3,3.7)	(0.1,0.3)	(7.3,18.3)	(0.8,2.1)	(18.3,39.2)	(4.2,14.9)
1.0	(2.0,3.4)	(0.1,0.1)	(12.0,22.3)	(1.7,2.3)	(25.9,33.5)	(6.7,9.8)
2.0	(1.8,4.0)	(0.5,2.0)	(18.8,34.0)	(2.6,5.1)	(39.5,50.3)	(9.5,14.4)
			(500,12,6)			
0.5	(1.6,3.7)	(0.0,0.0)	(5.8,8.8)	(1.0,2.2)	(16.7,26.7)	(6.4,16.7)
1.0	(2.3,6.2)	(0.0,0.1)	(10.0,15.1)	(2.7,8.5)	(27.8,38.5)	(13.1,38.3)
2.0	(2.6,3.8)	(0.2,0.3)	(14.6,20.3)	(2.7,4.1)	(44.6,53.9)	(14.8,28.2)
			(1000,12,3)			
0.5	(0.8,1.2)	(0.0,0.0)	(4.5,8.8)	(0.9,1.6)	(7.5,13.9)	(2.2,4.3)
1.0	(0.9,1.5)	(0.0,0.0)	(6.3,8.4)	(1.2,2.7)	(11.4,16.1)	(3.4,6.5)
2.0	(1.4,2.0)	(0.0,0.1)	(10.5,14.0)	(2.7,5.3)	(19.1,25.1)	(6.4,10.8)
			(1000,12,6)			
0.5	(3.8,9.0)	(0.4,1.8)	(4.5,9.2)	(1.3,3.4)	(15.0,28.1)	(7.2,19.0)
1.0	(4.1,7.8)	(0.0,0.0)	(9.9,20.0)	(1.4,2.1)	(28.2,50.7)	(8.4,13.7)
2.0	(3.9,7.7)	(0.1,0.2)	(11.9,24.8)	(1.8,3.6)	(37.3,66.7)	(11.6,26.2)
			(1000,25,6)			
0.5	(0.5,0.7)	(0.0,0.0)	(1.2,2.1)	(0.5,0.9)	(3.1,5.4)	(2.2,5.0)
1.0	(1.1,2.1)	(0.0,0.2)	(5.1,9.2)	(1.5,3.0)	(13.6,27.8)	(6.1,16.2)
2.0	(2.4,6.7)	(0.1,0.4)	(10.3,16.5)	(2.2,3.6)	(29.1,53.3)	(9.5,20.5)
			(1000,25,12)			
0.5	(3.6,8.5)	(0.4,0.8)	(1.5,2.2)	(0.6,1.1)	(12.7,29.8)	(6.7,9.4)
1.0	(4.3,8.6)	(0.3,0.7)	(5.0,8.4)	(1.0,1.5)	(31.6,63.0)	(12.1,23.1)
2.0	(5.8,10.3)	(0.4,1.0)	(8.0,10.3)	(1.1,2.4)	(41.7,66.3)	(13.6,29.7)

- $c \in [10, 11]$, the parameter for $F(M(x,\alpha)) = c \cdot M(x,\alpha)$,

The meaning of the columns in Table 4 correspond to those of tables 1, 2 and 3. For each setting and each threshold value $\overline{u}$, the average value of the five problems, followed by the corresponding maximum, are given.

As we can see, even for $\overline{u} = 0.5$ we have instances where the relative profit loss incurred when the probabilistic choice rule is assumed instead of the partially probabilistic choice rule is more than 18% (see the maximum value of $lost(P|PP)$ for the setting $(500, 12, 3)$). In fact, the mean is also rather high for some settings (see the results for (500,6,1)).

We can also see that there are some instances where the relative profit lost is high despite the location and quality being quite close to that of the probabilistic choice rule. See for instance the setting $(1000, 12, 3)$ when $\overline{u} = 1$: the difference in quality is negligible, and in location is rather small, but $lost(P|PP) = 6.3\%$ in average, and the maximum is 8.4%.

As in the example, the relative profit lost increases as the threshold value increases. Also, the figures of the relative profit lost due only to the new facility are higher.

7. Sensitivity analysis

We have also carried out a sensitivity analysis of some parameters of the model, in particular those of the cost function. Remember that the objective function of the partially probabilistic model is given by $\Pi_{PP}(x, \alpha) = F(M_{PP}(x, \alpha)) - G(x, \alpha) - G^{ex}$, and in our computational studies we have assumed that $G(x, \alpha) = G_1(x) + G_2(\alpha)$, where $G_1(x) = \sum_{i=1}^{i_{\max}} \Phi_i(d_i(x))$, with $\Phi_i(d_i(x)) = w_i/((d_i(x))^{\varphi_{i0}} + \varphi_{i1})$, $\varphi_{i0}, \varphi_{i1} > 0$ given parameters, and $G_2(\alpha) = e^{\frac{\alpha}{\beta_0}+\beta_1} - e^{\beta_1}$, with $\beta_0 > 0$ and β_1 given values. In our studies the same φ_{i0}value was assigned to each i. We have studied the influence of the parameters $\varphi_{i0}, \varphi_{i1}, \beta_0$ and β_1 in the optimal solution of the problem. Notice that this allows us to investigate how the costs due to the new facility affect the optimal solution of the problem, differentiating between the costs due to the location and those due to the quality of the facility.

We have proceed as follows. For each of the three basic scenarios considered in Subsection 6.1 and for each value of the threshold $\overline{u}$, we have solved 16 problems, by varying only one of those parameters up or down, and fixing the rest of the parameters to their original values. In particular, each parameter has been scaled down by multiplying it by 0.50 and 0.75, and scaled up by multiplying it by 1.33 and 2.00. Since similar conclusions can be inferred from the three scenarios and every threshold value, here we only present (see Table 5) the results corresponding to the 'large chain' scenario when $\overline{u} = 0.5, 1.5$ and 3.0. The rest of the tables can be obtained from the authors upon request. The first column in Table 5 gives the parameter that we have varied; the second one the scale factor applied to the parameter;

columns dist_x give the Euclidean distance between the optimal locations of the facility obtained with the original problem and the modified one; dist_q gives the corresponding variation in the quality, computed as quality of the modified problem minus quality of the original problem; and accordingly, diff Π_{PP} and diff G give the variation in the objective and cost functions, respectively.

Remember from Section 6.1 that for $\overline{u} = 0.5$ the optimal solution of the original problem was to locate the facility at Alcantarilla with $\alpha = 0.50$, for $\overline{u} = 1.5$ the location was at Murcia with $\alpha = 3.87$, and for $\overline{u} = 3.0$ at some place to the South between Alcantarilla and Murcia with $\alpha = 2.25$.

When β_0 decreases, $G_2(\alpha)$ increases. As a result, in the modified problems (see the lines for β_0 with mult=0.75 or 0.5), the quality of the new facility decreases (provided it is not at the minimum value 0.5, as it happens when $\overline{u} = 0.5$), which in turns implies that the location of the facility has also to be moved to a different place. Depending on the final combination of location and quality the total cost may increase or decrease, but the profit always decreases. On the other hand, when β_0 increases (see the lines for β_0 with mult=1.33 or 2.00), the opposite happens.

Something similar happens with β_1. When it decreases, so does G_2. Hence the quality of the new facilities increases, which in turn provokes a change in the location. And the optimal profit always increases (although the total cost may decrease or increase). The opposite happens when β_1 increases.

Concerning the parameter φ_{i0}, it has a smaller influence in the results than β_0 or β_1. Notice that G_2, as an exponential function, changes more than G_1 when varying its parameters. In addition to not changing too much, the shape of the modified location cost function is very similar to that of the original one, so the entire objective function does not change considerably. When φ_{i0} increases, the cost due to the location decreases. This provoke a change in the location, which in turn implies a change in the quality. And the final profit increases. The opposite happens when φ_{i0} decreases.

As for the parameter φ_{i1}, its influence in the objective function value is not too big, either. Again, as expected, when φ_{i1} decreases, the cost increases and hence Π_{PP} decreases, and viceversa.

8. Solving large problems: a computational study

As each choice rule may lead to a very different location and quality for the new facility to be located, it is important to have effective methods for solving the corresponding location problems. For the probabilistic choice rule we can find, in literature, both rigorous and heuristic methods (see Fernández et al. (2007), Redondo et al. (2009a)), and even parallel implementations of the heuristic proposed to cope with real-size problems (see Redondo et al. (2008)). Concerning the partially probabilistic choice rule, we have seen that the interval branch-and-bound algorithm iB&B described in Subsection 4.1 can manage small/medium size problems. It can also solve some large size problems, as we will see in this section, although heuristic procedures, like the one proposed in Subsection 4.2, are required in general. But what is the size of the largest problems that can be solved with iB&B? Is UEGO a trustworthy algorithm for solving competitive facility location and design problems when customers follow a partially probabilistic choice rule? These research questions are studied in this section.

All the computational studies have been carried out in a cluster with 18 nodes of shared memory and 8 GPUs. Each node has 16 cores (Intel Xeon E5 2650) and 64 GB of memory and 128 GB of solid-state drive. Each problem was run in one of the cores (one problem at a time). The algorithms have been implemented in C++. For the iB&B method we used the interval arithmetic in the PROFIL/BIAS library (Knüppel 1993).

Two new sets of location problems have been used to evaluate the performance of the algorithms. For the first one, 5 problems with $n = 5000$ demand points have been generated for each of the 4 settings considered, by randomly choosing its parameters uniformly within the intervals described in Subsection 6.2. Something similar was done to generate the second set, composed of 20 problems with $n = 10000$ demand points. The actual settings employed can be seen in tables 6 to 9. Furthermore, each problem was solved for the threshold values $\overline{u} = 0$ (i.e., with the classical probabilistic choice rule), 0.5, 1 and 2. It is important to highlight that the interval B&B algorithm applied to solve the problems with $\overline{u} = 0$ has been the same as the one applied to the problems with a strictly positive threshold (hence, only discarding tests without differentiability have been used), in order to study how the increase in the threshold value affects the algorithm. Notice, however, that using the method described in Fernández et al. (2007), which also includes discarding tests which make use of the differentiability of the objective function, more problems with $\overline{u} = 0$ could be solved, and faster.

The first important remark to make is that iB&B was not able to solve the problems for all the threshold values. This has been specified in the column labelled '$\overline{u}$' in tables 6 and 8, where the $\overline{u}$ values for which iB&B

was able to solve the problem are shown. As we can see, except for the first problem with setting $(n = 5000, j_{\max} = 50, k = 12)$, whenever iB&B was able to solve a given problem for a given value $\overline{u}'$, then it was also able to solve the problem for all the thresholds $\overline{u} \geq \overline{u}'$, which seems to suggest that the difficulty of the problem decreases as $\overline{u}$ increases. Note, however, that iB&B was not able to solve the problems with $\overline{u} = 0$, and just 5 out of 40 problems with $\overline{u} = 0.5$ were solved within the prescribed accuracy. In fact, as it can be seen in Table 8, half of the problems with $n = 10000$ demand points were not solved by iB&B for none of the threshold values (the computer ran out of memory). This clearly shows the difficulty of the problem at hand, and the need for a heuristic method to cope with large-size problems.

Tables 6 and 8 summarize the results obtained by the algorithms for the instances where iB&B was able to provide a solution. The first column refers to each of the five problems of each setting. Then we have seven columns related to iB&B. Column 'Time' gives the average time (in seconds) employed by iB&B in solving the problems with the threshold values shown in column '$\overline{u}$'. The next three columns indicate the corresponding average for the minimum, best and maximum objective value, as offered by iB&B. To be more precise, when solving a problem (with a given threshold value) with iB&B, we obtain a list of boxes $\mathcal{L}_{PP} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_r\}$ and the best point found by the algorithm y^{best}. $\min\{\underline{\boldsymbol{\Pi}_{PP}(\boldsymbol{y}_l)} : l = 1, \ldots, r\}$ (resp. $\max\{\overline{\boldsymbol{\Pi}_{PP}(\boldsymbol{y}_l)} : l = 1, \ldots, r\}$) gives the minimum (resp. maximum) objective function value at any point in the solution boxes. Column 'Π_{min}' shows the average of the lower limits, 'Π_{max}' that of the upper limits, and 'Π_{best}' the average of the $\underline{\boldsymbol{\Pi}_{PP}(y^{best})}$ values. Column 'Wid' indicates the average of the width of the intervals $[\min\{\underline{\boldsymbol{\Pi}_{PP}(\boldsymbol{y}_l)} : l = 1, \ldots, r\}, \max\{\overline{\boldsymbol{\Pi}_{PP}(\boldsymbol{y}_l)} : l = 1, \ldots, r\}]$ and 'S.D.', the corresponding standard deviation.

The next eight columns refer to the heuristic UEGO. Since it is a stochastic algorithm, each run may lead to a different solution. To take this fact into account, each problem has been solved 5 times with UEGO, and the following values have been computed: the average CPU time employed in solving the problem in the 5 runs, the maximum Euclidean distance between any pair of solutions, the minimum, average and maximum objective function values of the solutions and the corresponding standard deviation, the average reduction in CPU time as compared to the iB&B method and finally the percentage of successes (we say that UEGO has solved the problem successfully in a given run when the solution provided by UEGO is included in one of the solution boxes provided by iB&B). The figures in the columns corresponding to the UEGO algorithm give, in the same order, the averages of those values: in tables 6 and 8 when considering only the problems solved by iB&B, and in tables 7 and 9 considering the problems for all the values of $\overline{u}$. Notice that in tables 7 and 9 we omit columns 'T.R.' (average time reduction) and 'Succ' (average percentage of success), as the interval algorithm is not able to solve all the generated problems.

As we can see in tables 6 and 8, the best point found by the interval algorithm is (very close to) the optimal one, as its objective value Π_{best} is very close to the upper bound Π_{max} (remember that the optimal objective function value lies within the interval $[\Pi_{best}, \Pi_{max}]$). And even though the lower bound Π_{min} is not that close to Π_{best} (see column 'Wid'), the relative width of the interval $[\Pi_{min}, \Pi_{max}]$ is always less than 0.0025. As usual with branch-and-bound methods, the CPU time employed by iB&B is quite erratic: some problems are solved in just 314 seconds (see problem $(5000, 50, 25)_1$), whereas others require more than 68274 seconds (see problem $(5000, 100, 50)_3$). The CPU time employed by UEGO is also a bit erratic, but not as much as that of iB&B.

UEGO is rather robust. As we can see in tables 6 and 8, it has 100% success in all the problems solved by iB&B, i.e, in all the problems and in all the runs, the solution provided by UEGO was always included in one of the solution boxes provided by iB&B. In fact, we can see that the minimum, average and maximum objective function values obtained by UEGO are always about the same (see also column 'Dev'). Not only is the objective function value the same in all the runs, but the solution point offered by the algorithm is also the same, as can be seen in column 'Dist'. This is also true when considering all the problems, as can be checked in tables 7 and 9: in only 2 out of 40 settings is the average Dev greater than 0.1, and in only 4 settings is the average Dist greater than 0.1. Also notice that the worst objective function value found by UEGO in any of the runs is usually better than or equal to Π_{best}. And this using a tiny fraction of the CPU time employed by iB&B (see the values of the time reduction in column 'T.R.').

Hence, we can conclude that the rigorous iB&B can manage instances with up to 1000 demand points without difficulties, but it starts experiencing problems with instances with 5000 demand points. On the other hand, UEGO is a trustworthy heuristic method, able to solve problems with up to 10000 demand points without difficulties and requiring less than 13 minutes.

9. Conclusions and future research lines

An extension of the classical probabilistic choice rule has been introduced and studied. According to the modified rule, called partially probabilistic choice rule, a customer, in order to satisfy his/her demand, only patronizes those facilities for which he/she feels an attraction greater than or equal to a threshold value, and the demand is split among them proportionally to their attraction. Unlike most of the choice rules employed in literature, the threshold value implicitly implies that there may be some unmet demand. Hence, the model is suitable for inessential goods.

The influence of the choice rule in the location of competitive facilities has been analyzed. In particular, the problem of locating a single new facility in the *plane* has been considered. The corresponding location problem for profit maximization has been formulated, and a rigorous interval branch-and-bound method (iB&B), as well as a heuristic evolutionary algorithm (UEGO) have been proposed to cope with the problem.

According to the computational studies, the optimal location and quality of the new facility as well as the profit obtained by the chain and by the new facility may vary considerably depending on the customer choice rule employed. Hence, the selection of the choice rule to be used in real applications should be made with care. In particular, although the assumption of the probabilistic choice rule makes the problem more computationally tractable, it should only be used when customers really patronize *all* the existing facilities because even a small threshold value may lead to a very different solution.

As for the solution procedures, iB&B can solve problems with up to 1000 demand points rigorously. For larger instances, the use of the evolutionary algorithm UEGO is recommended, as it is robust method.

The extension of the single facility location model with partially probabilistic choice rule to the case of the location of more than one facility, and to the case where competitors react by locating new facilities too, will be analyzed in the future. The use of threshold values in other existing choice rules is another line for future research.

A modification of the partially probabilistic choice rule is required to handle essential goods, since in that case all the demand has to be served. But what if at a given demand point the attraction towards any of the facilities falls below the minimum threshold level? One possibility is that the demand at that demand point be served by the most attractive facility for that demand point (hence, all the customers at that demand point will go the same facility, the most attractive for them, similarly to the customers of the demand points for which there is only one facility with an attraction level above the threshold). This modified rule requires a deeper study, as the objective function should be reformulated.

As already highlighted in Serra et al. (1999), competitive location is a difficult field not only because it involves rather complex mathematical models, but also because customer behavior cannot easily be transcribed into neat equations. The model presented in this paper is just an approximation to the real world. More research is needed on this aspect of customer behavior modeling.

Acknowledgements

We thanks the referees for the their comments and suggestions, which have helped to improve the quality of the paper. The sensitivity analysis in Section 7 was suggested by the referees.

This research has been supported by grants from the Spanish Ministry of Economy and Competitiveness (MTM2015-70260-P, and TIN2015-66680-C2-1-R), the Hungarian National Research, Development and Innovation Office - NKFIH (OTKA grant PD115554), Fundación Séneca (The Agency of Science and Technology of the Region of Murcia, 19241/PI/14), Junta de Andalucía (P12-TIC301), in part financed by the European Regional Development Fund (ERDF). Juana López Redondo is a fellow of the Spanish 'Ramón y Cajal' contract program.

References

Aboolian, R., O. Berman, D. Krass. 2007. Competitive facility location and design problem. *European Journal of Operational Research* **182**(1) 40–62.

Arenoe, B., J.-P. I. van der Rest, P. Kattuman. 2015. Game theoretic pricing models in hotel revenue management: An equilibrium choice-based conjoint analysis approach. *Tourism Management* **51**(1) 96–102.

Berman, O., Z. Drezner, D. Krass. 2010. Generalized coverage: New developments in covering location models. *Computers & Operations Research* **37**(10) 1675–1687.

Brimberg, J., R.F. Love. 1995. Estimating distances. Z. Drezner, ed., *Facility location: a survey of applications and methods*. Springer Series in Operations Research and Financial Engineering, Springer, Berlin, 9–32.

Cantillo, V., J. Amaya, J. de D. Ortúzar. 2010. Thresholds and indifference in stated choice surveys. *Transportation Research Part B: Methodological* **44** 753–763.

Cantillo, V., B. Heydecker, J. de D. Ortúzar. 2006. A discrete choice model incorporating thresholds for perception in attribute values. *Transportation Research Part B: Methodological* **40**(9) 807–825.

Cantillo, V., J. de D. Ortúzar. 2006. Implications of thresholds in discrete choice modelling. *Transport Reviews* **26**(6) 667–691.

Colomé, R., H.R. Lourenco, D. Serra. 2003. A New Chance-Constrained Maximum Capture Location Problem. *Annals of Operations Research* **122**(1-4) 121–139.

Drezner, T. 2006. Derived attractiveness of shopping malls. *IMA Journal of Management Mathematics* **17**(4) 349–358.

Drezner, T., Z. Drezner. 1997. Replacing continuous demand with discrete demand in a competitive location model. *Naval Research Logistics* **44** 81–95.

Drezner, T., Z. Drezner. 2002. Validating the gravity-based competitive location model using inferred attractiveness. *Annals of Operations Research* **111** 227–237.

Drezner, T., Z. Drezner. 2008. Lost demand in a competitive environment. *Journal of the Opeartional Research Society* **59** 362–371.

Drezner, T., Z. Drezner. 2012. Modelling lost demand in competitive facility location. *Journal of the Operational Research Society* **63** 201–206.

Drezner, T., Z. Drezner, P. Kalczynski. 2012. Strategic competitive location: improving existing and establishing new facilities. *Journal of the Operational Research Society* **63**(12) 1720–1730.

Eiselt, H.A., G. Laporte. 1996. Sequential location problems. *European Journal of Operational Research* **96**(2) 217–231.

Eiselt, H.A., G. Laporte, J.F. Thisse. 1993. Competitive location models: a framework and bibliography. *Transportation Science* **27**(1) 44–54.

Fernández, J., P. Fernández, B. Pelegrín. 2002. Estimating actual distances by norm functions: a comparison between the $l_{k,p,\theta}$-norm and the $l_{b_1,b_2,\theta}$-norm and a study about the selection of the data set. *Computers and Operations Research* **29**(6) 609–623.

Fernández, J., B. Pelegrín. 2001. Using interval analysis for solving planar single-facility location problems: new discarding tests. *Journal of Global Optimization* **19**(1) 61–81.

Fernández, J., B. Pelegrín, F. Plastria, B. Tóth. 2007. Solving a Huff-like competitive location and design model for profit maximization in the plane. *European Journal of Operational Research* **179**(3) 1274–1287.

Ghosh, A., C.S. Craig. 1991. FRANSYS: a franchise distribution system location model. *Journal of Retailing* **67** 466–495.

Gilbride, T. J., G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science* **23**(3) 391–406.

Hansen, E., G. W. Walster. 2004. *Global optimization using interval analysis*. Second revised and expanded ed. Marcel Dekker.

Holmes, J., F.B. Williams, L.A. Brown. 1972. Facility location under a maximum travel restriction: an example using day care facilities. *Geographical Analysis* **4**(3) 258–266.

Knüppel, O. 1993. PROFIL/BIAS - A Fast Interval Library. *Computing* **53**(1) 277–287.

Kress, D., E. Pesch. 2012. Sequential competitive location on networks. *European Journal of Operational Research* **217**(3) 483–499.

Li, X., Z. Zhao, X. Zhu, T. Wyatt. 2011. Covering models and optimization techniques for emergency response facility location and planning: a review. *Mathematical Methods of Operations Research* **74**(3) 281–310.

Moscato, P. 1989. On evolution, search, optimization, genetic algorithms and martial arts: towards memetic algorithms. Tech. Rep. C3P 826, California Institute of Technology, Pasadena, CA.

Nakanishi, M., L.G. Cooper. 1974. Parameter estimate for multiplicative interactive choice model: Least square approach. *Journal of Marketing Research* **11** 303–311.

Nakanishi, M., L.G. Cooper. 1982. Simplified Estimation Procedures for MCI Models. *Marketing Science* **1**(3) 314–322.

Neumaier, A. 2004. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica 2004*. Cambridge University Press, 271–369.

Peeters, P.H., F. Plastria. 1992. Discretization results for the Huff and Pareto-Huff competitive location models on networks. *Top* **6** 247–260.

Plastria, F. 2001. Static competitive facility location: an overview of optimisation approaches. *European Journal of Operational Research* **129**(3) 461–470.

Plastria, F., E. Carrizosa. 1999. Undesirable facility location with minimal covering objectives. *European Journal of Operational Research* **119**(1) 158–180.

Plastria, F., E. Carrizosa. 2004. Optimal location and design of a competitive facility. *Mathematical Programming* **100**(2) 247–265.

Redondo, J.L., J. Fernández, A.G. Arrondo, I. García, P.M. Ortigosa. 2012. Fixed or variable demand? Does it matter when locating a facility? *Omega* **40**(1) 9–20.

Redondo, J.L., J. Fernández, A.G. Arrondo, I. García, P.M. Ortigosa. 2013. A two-level evolutionary algorithm for solving the facility location and design (1|1)-centroid problem on the plane with variable demand. *Journal of Global Optimization* **56**(3) 983–1005.

Redondo, J.L., J. Fernández, I. García, P.M. Ortigosa. 2008. Parallel algorithms for continuous competitive location problems. *Optimization Methods & Software* **23**(5) 779–791.

Redondo, J.L., J. Fernández, I. García, P.M. Ortigosa. 2009a. A robust and efficient global optimization algorithm for planar competitive location problems. *Annals of Operations Research* **167**(1) 87–106.

Redondo, J.L., J. Fernández, I. García, P.M. Ortigosa. 2009b. Solving the multiple competitive facilities location and design problem on the plane. *Evolutionary Computation* **17**(1) 21–53.

Saidani, N., F. Chu, H. Chen. 2012. Competitive facility location and design with reactions of competitors already in the market. *European Journal of Operational Research* **219**(1) 9–17.

Serra, D., H.A. Eiselt, G. Laporte, C. ReVelle. 1999. Market Capture Models under Various Customer Choice Rules. *Environment and Planning B* **26**(5) 141–150.

Solis, F.J., R.J.B. Wets. 1981. Minimization by random search techniques. *Mathematics of Operations Research* **6**(1) 19–30.

Suárez-Vega, R., D.R. Santos-Peñate, P. Dorta-González. 2007. The follower location problem with attraction thresholds. *Papers in Regional Science* **86**(1) 123–137.

Thill, J.C. 2000. Network competition and branch differentiation with consumer heterogeneity. *The Annals of Regional Science* **34**(3) 451–468.

Tóth, B., J. Fernández. 2010. *Interval methods for single and bi-objective optimization problems - applied to competitive facility location problems*. Lambert Academic Publishing, Saarbrücken.

Tóth, B., J. Fernández, T. Csendes. 2007. Empirical convergence speed of inclusion functions for facility location problems. *Journal of Computational and Applied Mathematics* **199** 384–389.

Tóth, B., F. Plastria, J. Fernández, B. Pelegrín. 2009. On the impact of spatial pattern, aggregation, and model parameters in planar Huff-like competitive location and design problems. *OR Spectrum* **31**(1) 601–627.

URL. 2017. Languages for interval analysis. Available at `http://www.cs.utep.edu/interval-comp/intlang.html`.

Yapicioglu, H., A.E. Smith, G. Dozier. 2007. Solving the semi-desirable facility location problem using bi-objective particle swarm. *European Journal of Operational Research* **177**(2) 733–749.

Table 5: Sensitivity analysis of the large chain scenario.

		$\overline{u} = 0.5$				$\overline{u} = 1.5$				$\overline{u} = 3.0$			
par	mult	dist_x	dist_q	diff Π_{PP}	diff G	dist_x	dist_q	diff Π_{PP}	diff G	dist_x	dist_q	diff Π_{PP}	diff G
β_0	0.50	0.00	0.00	-3.4	3.4	1.84	-3.18	-4.9	-27.9	1.12	-1.74	-14.0	-8.5
β_0	0.75	0.00	0.00	-1.1	1.1	1.84	-3.18	-1.6	-31.2	0.09	-0.39	-5.2	1.9
β_0	1.33	2.23	4.28	10.7	30.1	0.23	1.11	10.7	-1.3	0.03	0.02	4.5	-4.5
β_0	2.00	2.23	4.50	22.5	19.9	0.23	1.13	22.5	-13.1	1.32	2.75	12.1	6.5
β_1	0.50	2.23	4.50	34.0	8.5	0.23	1.13	34.0	-24.5	1.32	2.75	23.5	-4.9
β_1	0.75	2.23	4.50	23.4	19.0	0.23	1.13	23.4	-14.0	1.32	2.75	12.9	5.6
β_1	1.33	0.00	0.00	-7.8	7.8	1.90	-3.36	-10.4	-26.2	1.12	-1.74	-18.5	-4.0
β_1	2.00	0.00	0.00	-130.7	130.7	2.09	-3.37	-133.5	95.9	1.18	-1.75	-142.8	118.0
φ_{i0}	0.50	1.57	2.93	-4.1	32.7	0.00	0.00	-4.1	4.1	0.03	-0.01	-8.8	8.8
φ_{i0}	0.75	1.57	2.93	-4.1	32.7	0.00	0.00	-4.1	4.1	0.03	-0.01	-8.8	8.8
φ_{i0}	1.33	0.00	0.00	0.0	-0.0	0.00	0.00	0.0	0.0	0.00	0.00	0.0	-0.0
φ_{i0}	2.00	0.02	0.00	2.9	-2.9	1.84	-3.18	2.6	-35.4	0.01	0.00	2.4	-2.4
φ_{i1}	0.50	0.02	0.00	-2.7	2.7	1.84	-3.18	-2.2	-30.6	0.05	0.03	-2.1	2.1
φ_{i1}	0.75	0.01	0.00	-1.0	1.0	1.84	-3.18	-0.9	-31.9	0.02	0.01	-0.9	0.9
φ_{i1}	1.33	1.57	2.96	2.1	26.7	0.00	0.00	2.2	-2.2	0.02	-0.01	0.9	-0.9
φ_{i1}	2.00	1.57	2.96	4.6	24.2	0.00	0.00	4.6	-4.6	0.03	-0.01	2.1	-2.1

Table 6: iB&B and UEGO results for problems with $n = 5000$ demand points.

	iB&B							UEGO							
$(m,k)_{prob}$	$\overline{u}$	Time	Π_{min}	Π_{best}	Π_{max}	Wid	S.D.	Time	Dist	Min	Aver	Max	Dev	T.R	Succ
$(50,12)_1$	0.5,2.0	10376.2	7396.4	7413.1	7413.2	16.8	14.0	135.4	0.001	7413.1	7413.1	7413.2	0.024	99	100
$(50,12)_2$	0.5,1.0,2.0	1228.4	7201.6	7210.3	7210.4	8.9	6.9	122.9	0.000	7210.4	7210.4	7210.4	0.001	90	100
$(50,12)_3$	2.0	731.9	5935.6	5954.2	5954.3	18.6	0.0	121.4	0.000	5954.2	5954.2	5954.2	0.002	83	100
$(50,12)_4$	1.0,2.0	2977.3	7739.2	7748.9	7748.9	9.7	2.0	123.6	0.000	7748.9	7748.9	7748.9	0.003	96	100
$(50,12)_5$	1.0,2.0	1882.4	6458.2	6477.8	6477.9	19.6	5.7	128.4	0.000	6477.8	6477.8	6477.9	0.003	93	100
$(50,25)_1$	1.0,2.0	314.3	9362.1	9380.4	9380.5	18.3	4.2	143.4	0.001	9380.4	9380.5	9380.5	0.002	54	100
$(50,25)_2$	0.5,1.0,2.0	5892.4	13622.0	13633.2	13633.3	11.2	7.4	144.3	0.000	13633.2	13633.2	13633.2	0.004	98	100
$(50,25)_3$	1.0,2.0	537.4	12869.3	12888.2	12888.3	18.9	14.5	151.9	0.001	12888.2	12888.3	12888.3	0.002	70	100
$(50,25)_4$	0.5,1.0,2.0	1697.2	11324.2	11334.3	11334.4	10.1	6.0	137.9	0.000	11334.3	11334.3	11334.4	0.002	92	100
$(50,25)_5$	1.0,2.0	5079.9	11463.9	11478.4	11478.5	14.5	6.3	177.3	0.000	11478.5	11478.5	11478.5	0.001	96	100
$(100,25)_1$	2.0	5489.8	5084.4	5090.0	5090.1	5.6	0.0	177.1	0.000	5090.1	5090.1	5090.1	0.002	97	100
$(100,25)_2$	2.0	4076.7	5096.2	5101.4	5101.4	5.1	0.0	221.8	0.001	5101.4	5101.4	5101.4	0.003	95	100
$(100,25)_3$	2.0	1055.8	5636.6	5652.4	5652.5	15.8	0.0	192.1	0.000	5652.4	5652.4	5652.5	0.002	82	100
$(100,25)_4$	1.0,2.0	2942.4	7613.3	7628.2	7628.3	15.0	2.7	157.3	0.001	7628.3	7628.3	7628.3	0.002	95	100
$(100,25)_5$	2.0	19339.3	5117.5	5119.4	5119.4	1.8	0.0	223.8	0.000	5119.4	5119.4	5119.4	0.005	99	100
$(100,50)_1$	1.0,2.0	35746.0	16096.0	16102.7	16102.8	6.7	3.6	154.3	0.000	16102.7	16102.7	16102.7	0.001	100	100
$(100,50)_2$	0.5,1.0,2.0	8925.1	8965.7	8974.1	8974.2	8.5	10.3	247.5	0.000	8974.2	8974.2	8974.2	0.003	97	100
$(100,50)_3$	0.5,1.0,2.0	68274.5	9039.4	9042.8	9042.9	3.4	4.0	232.3	0.003	9042.8	9042.8	9042.9	0.002	100	100
$(100,50)_4$	2.0	855.3	14850.8	14868.0	14868.0	17.2	0.0	159.0	0.000	14868.0	14868.0	14868.0	0.001	82	100
$(100,50)_5$	2.0	6390.2	16524.3	16527.6	16527.7	3.3	0.0	223.3	0.000	16527.6	16527.6	16527.6	0.000	96	100

Table 7: UEGO results for problems with $n = 5000$ demand points.

$(m,k)_{prob}$	Time	Dist	Min	Aver	Max	Dev
$(50,12)_1$	168.1	0.019	7563.0	7563.0	7563.1	0.049
$(50,12)_2$	126.9	0.000	7220.4	7220.4	7220.4	0.001
$(50,12)_3$	132.7	0.001	6580.3	6580.3	6580.4	0.013
$(50,12)_4$	126.3	0.000	7976.9	7976.9	7976.9	0.002
$(50,12)_5$	190.1	0.015	6455.3	6455.3	6455.3	0.013
average	148.8	0.007	7159.2	7159.2	7159.2	0.016
$(50,25)_1$	206.2	0.000	9891.3	9891.3	9891.3	0.001
$(50,25)_2$	200.5	0.000	13645.7	13645.7	13645.7	0.003
$(50,25)_3$	279.3	0.005	13421.8	13421.8	13421.8	0.006
$(50,25)_4$	175.3	0.000	11390.6	11390.6	11390.6	0.002
$(50,25)_5$	270.4	0.000	11630.7	11630.7	11630.7	0.001
average	226.4	0.001	11996.0	11996.0	11996.0	0.003
$(100,25)_1$	319.1	0.039	5132.1	5132.2	5132.2	0.035
$(100,25)_2$	241.3	0.014	5175.9	5176.0	5176.0	0.008
$(100,25)_3$	288.9	0.057	5380.5	5380.6	5380.6	0.040
$(100,25)_4$	203.6	0.027	7543.6	7543.6	7543.6	0.004
$(100,25)_5$	423.8	0.028	4946.4	4946.4	4946.4	0.021
average	295.3	0.033	5635.7	5635.8	5635.8	0.022
$(100,50)_1$	221.7	0.007	16070.6	16070.6	16070.6	0.002
$(100,50)_2$	274.8	0.000	8975.2	8975.2	8975.2	0.002
$(100,50)_3$	325.6	0.004	9084.5	9084.5	9084.5	0.001
$(100,50)_4$	285.0	0.013	14853.1	14853.1	14853.2	0.036
$(100,50)_5$	373.1	0.009	16261.5	16261.5	16261.5	0.007
average	296.0	0.007	13049.0	13049.0	13049.0	0.010

Table 8: iB&B and UEGO results for problems with $n = 10000$ demand points.

	iB&B							UEGO							
$(m,k)_{prob}$	$\overline{u}$	Time	Π_{min}	Π_{best}	Π_{max}	Wid	S.D.	Time	Dist	Min	Aver	Max	Dev	T.R	Succ
$(100,25)_1$	2.0	5985.1	10215.9	10221.7	10221.8	5.8	0.0	381.0	0.018	10221.6	10221.7	10221.7	0.023	94	100
$(100,25)_2$	2.0	3678.8	9799.1	9807.6	9807.7	8.6	0.0	254.9	0.010	9807.5	9807.6	9807.6	0.015	93	100
$(100,25)_3$	2.0	1276.3	10903.9	10915.7	10915.8	11.9	0.0	251.0	0.001	10915.7	10915.7	10915.7	0.003	81	100
$(100,25)_4$															
$(100,25)_5$	2.0	504.2	9021.8	9041.5	9041.6	19.8	0.0	262.6	0.002	9041.5	9041.5	9041.5	0.005	46	100
$(100,50)_1$	2.0	13360.1	16768.8	16774.6	16774.6	5.7	0.0	570.1	0.001	16774.5	16774.6	16774.6	0.005	96	100
$(100,50)_2$	1.0,2.0	34013.3	16534.5	16537.5	16537.6	3.0	2.1	662.7	0.000	16537.4	16537.5	16537.5	0.046	98	100
$(100,50)_3$	2.0	3414.9	20728.3	20736.6	20736.7	8.3	0.0	276.3	0.000	20736.6	20736.7	20736.7	0.004	92	100
$(100,50)_4$	2.0	18150.6	16792.5	16800.2	16800.3	7.7	0.0	288.5	0.000	16800.2	16800.2	16800.3	0.035	98	100
$(100,50)_5$	1.0,2.0	344.7	18556.3	18573.2	18573.3	17.0	3.3	616.8	0.000	18573.3	18573.3	18573.3	0.004	73	100
$(200,50)_1$															
$(200,50)_2$															
$(200,50)_3$															
$(200,50)_4$															
$(200,50)_5$															
$(200,100)_1$															
$(200,100)_2$															
$(200,100)_3$	1.0,2.0	485.8	14179.3	14179.4	14179.4	0.0	0.0	783.4	0.000	14179.4	14179.4	14179.4	0.000	63	100
$(200,100)_4$															
$(200,100)_5$															

Table 9: UEGO results for problems with $n = 10000$ demand points.

$(m,k)_{prob}$	Time	Dist	Min	Aver	Max	Dev
$(100,25)_1$	763.1	0.010	9818.8	9818.8	9818.8	0.006
$(100,25)_2$	631.9	0.006	9794.8	9794.8	9794.8	0.012
$(100,25)_3$	629.2	0.046	10851.0	10851.1	10851.1	0.043
$(100,25)_4$	480.3	0.145	11424.5	11424.8	11425.0	0.165
$(100,25)_5$	534.0	0.002	7739.7	7739.7	7739.7	0.013
average	607.7	0.042	9925.8	9925.8	9925.9	0.048
$(100,50)_1$	729.4	0.002	15296.8	15296.8	15296.8	0.000
$(100,50)_2$	857.6	0.009	19886.6	19886.6	19886.6	0.012
$(100,50)_3$	435.3	0.010	18422.5	18422.6	18422.6	0.038
$(100,50)_4$	678.0	0.081	15446.0	15446.0	15446.0	0.013
$(100,50)_5$	837.9	0.000	16921.3	16921.3	16921.3	0.002
average	707.6	0.020	17194.6	17194.7	17194.7	0.013
$(200,50)_1$	527.3	0.000	6468.9	6468.9	6468.9	0.000
$(200,50)_2$	660.7	0.051	8108.3	8108.5	8108.5	0.074
$(200,50)_3$	679.8	0.108	7095.2	7095.3	7095.4	0.072
$(200,50)_4$	529.6	0.220	9303.7	9303.9	9304.0	0.110
$(200,50)_5$	691.1	0.105	7056.1	7056.1	7056.2	0.051
average	617.7	0.097	7606.4	7606.5	7606.6	0.061
$(200,100)_1$	903.4	0.098	18956.1	18956.1	18956.2	0.007
$(200,100)_2$	912.4	0.128	13691.1	13691.1	13691.2	0.037
$(200,100)_3$	884.9	0.000	12457.5	12457.5	12457.5	0.000
$(200,100)_4$	626.4	0.003	18128.4	18128.4	18128.4	0.001
$(200,100)_5$	594.2	0.059	19783.6	19783.6	19783.7	0.015
average	784.3	0.058	16603.3	16603.3	16603.4	0.012