Mathematical Programming manuscript No.
(will be inserted by the editor)

**Roberto Tadei · Nicoletta Ricciardi ·
and Guido Perboli**

The Stochastic p-Median Problem with Unknown Cost Probability Distribution

Received: date / Accepted: date

Abstract Consider a set of potential facility locations partitioned into non-empty disjoint subsets, called clusters. The cost for using a facility is a stochastic variable given by the sum of a deterministic cost associated to each cluster plus a random term with unknown probability distribution which represents the heterogeneity of the costs inside the cluster. The Stochastic p-Median Problem with Unknown Cost Probability Distribution consists in finding the location of p facilities, each belonging to one cluster, which minimizes the expected total cost. A large number of real-life situations can be modeled in such a way. Using the method of the asymptotic approximations derived from the extreme value theory the expected optimal value of the allocation variables is shown to be a multinomial Logit model. The optimal value of the location variables is then obtained by solving an integer deterministic nonlinear problem.

Keywords p-median · stochastic costs · extreme value theory · asymptotic approximations · multinomial Logit model

---

Roberto Tadei
Politecnico di Torino - Turin (ITALY)
Tel.: +39 011 0907032
Fax: +39 011 0907099
E-mail: roberto.tadei@polito.it

Nicoletta Ricciardi
"Sapienza" - Università di Roma - Rome (ITALY)
Tel.: +39 06 49910904
Fax: +39 06 4959241
E-mail: nicoletta.ricciardi@uniroma1.it

Guido Perboli
Politecnico di Torino - Turin (ITALY)
Tel.: +39 011 0907097
Fax: +39 011 0907099
E-mail: guido.perboli@polito.it

Mathematics Subject Classification (2000) 90B80 · 91B70 · 60G70 · 41A60

1 Introduction

Given a set of customers, a set of potential facility locations, partitioned into n nonempty disjoint subsets called clusters, and stochastic costs with unknown probability distribution, the Stochastic p-Median Problem with Unknown Cost Probability Distribution ($SpMP-UCPD$) consists in finding p, with $p \leq n$, facility locations, no more than one per cluster, which minimize the expected total cost. The stochastic costs are given by the sum of a deterministic cost associated to each cluster plus a random term, with unknown probability distribution, which represents the cost heterogeneity inside each cluster.

A large number of real-life situations can be satisfactorily modeled as $SpMP-UCPD$. For instance, this problem may arise in marketing, where a producer wishes to bring out a mix of p products which maximizes his profit (see Section 2 for details). Further, it may apply in any context where random utility choices are to be considered (e.g. transport planning, residential choice, service location, etc.).

Moreover, the $SpMP-UCPD$ may appear as a subproblem in many approaches for several combinatorial optimization problems, like stochastic scheduling and routing. Because of that, finding good methods for coping with this problem may help to better solve stochastic scheduling and routing problems also.

Various results on the p-median problem on stochastic networks are given in the literature ([19], [20], [21], [23] and [31]). When the stochastic network is a tree, Mirchandani and Oudjit [24] use a selective enumeration approach for solving the 2-median problem.

The uncertainty in the p-median problem concerns mainly the weights, which are considered as random variables. In particular, there are studies where the probability that the objective function exceeds a given threshold is minimized ([3], [4], [7]), others which analyze either the probability to have particular optimal solutions ([32], [10], [9]) or the uncertainty about future weights ([11], [8]). A different stream of studies concerns the uncertainty of the locations. In particular, future locations [15] or the number of facilities in the future [1] are considered.

In several papers dealing with uncertainty in the p-median problem, assumptions on the type of the cost probability distribution are given (either the probability distribution is known or a finite number of possible states, each occuring with non-zero probability, is assumed). Unfortunately, in many real-life situations the exact shape of this distribution is unknown.

Very good surveys on the p-median problem both deterministic and stochastic can be found in [22], [30] and [2]. For a more general and very recent survey on location analysis see [26].

In a very recent paper dedicated to the memory of Charles ReVelle [27] the authors identify some new fields which they claim to be particularly promising in location theory. They are:

- different structures of customer demand
- models ranging from gravity types to Logit functions (which appear to the authors to be most promising in the context of location modeling, see e.g. [25]), and
- congestion along the roads and at the facilities.

This paper involves both the first and the second field. In fact, it considers a random evaluation of the facilities by the customers (so a different structure of customer demand) proving that, under a quite mild and reasonable assumption on the shape of the unknown cost probability distribution and the number of the potential facility locations, the probability distribution of the minimum cost becomes a Gumbel (or double exponential) one. Moreover, using such a distribution, a multinomial Logit function for the allocation variables (i.e. the variables which allocate demand to supply) of the p-median problem is derived. We observe that the multinomial Logit function has been proved to be particularly suitable to describe the dispersion, due to the cost stochasticity, of the customer preferences among the facilities [6].

The remainder of the paper is organized as follows. In Section 2 the $SpMP-UCPD$ is introduced. In Section 3 the stochastic minimum cost for any customer is considered and its, still unknown, probability distribution is given. In Section 4 by applying the method of the asymptotic approximations it is proved that the above unknown distribution becomes a Gumbel (or double exponential) distribution. In Section 5 the expected optimal value of the allocation variables is shown to be a multinomial Logit model. An integer deterministic nonlinear problem to find the optimal location of the p facilities is derived in Section 6. Finally, the conclusions of our work are reported in Section 7.

2 Problem definition

Let V be the set of customers and U the set of potential facility locations, partitioned into n nonempty disjoint subsets $U_1, ..., U_n$, called clusters.

Let r_{ij}^k be the stochastic cost that customer $i \epsilon V$ pays for using facility $j \epsilon U_k, k = 1, ..., n$.

We assume

$$r_{ij}^k = c_k + \theta_{ij}, \quad i \epsilon V, j \epsilon U_k, k = 1, ..., n, \tag{1}$$

i.e., r_{ij}^k is the sum of c_k, a deterministic cost equal for any facility $j \epsilon U_k$ (and for any customer $i \epsilon V$), and θ_{ij}, a stochastic variable representing the heterogeneity of the cost paid by customer $i \epsilon V$ for using any facility j, i.e. a deviation (positive or negative) from the deterministic costs c_k.

In practice, as the number of customers and potential facility locations is usually very high, it is impossible to observe and measure the deviations θ_{ij}.

Since the random terms θ_{ij} are not observable, their probability distribution is in general unknown, and any tight assumption about it would be arbitrary.

Let us assume θ_{ij} are independent and identically distributed (i.i.d.) random variables with a common unknown probability distribution ([17], [16], [18]) given by

$$F(x) = Pr\{\theta_{ij} \geq x\},\ i\epsilon V, j\epsilon U_k, k = 1, ..., n \tag{2}$$

The independence of the random variables θ_{ij} is justified by the fact they derive from the heterogeneity of the customer costs, and the customers are independent each other; moreover, there is no reason to consider different distributions for these random variables, which implies the identity of the distribution.

Let x_{ij}^k be a continuous decision variable corresponding to the fraction of the demand of customers $i\epsilon V$ to facility $j\epsilon U_k$ $(0 \leq x_{ij}^k \leq 1)$; f_j^k the fixed cost associated with opening facility $j\epsilon U_k$; p the desired number of facilities to be opened; y_j^k a binary variable which takes value 1 if destination $j\epsilon U_k$ is open, 0 otherwise.

The $SpMP-UCPD$ may be formulated as follows

$$\min \sum_{j\epsilon U_k} \sum_{i\epsilon V} \sum_{k=1}^{n} r_{ij}^k x_{ij}^k + \sum_{k=1}^{n} \sum_{j\epsilon U_k} f_j^k y_j^k \tag{3}$$

subject to

$$\sum_{k=1}^{n} \sum_{j\epsilon U_k} x_{ij}^k = 1,\ \ i\epsilon V \tag{4}$$

$$\sum_{k=1}^{n} \sum_{j\epsilon U_k} y_j^k = p \tag{5}$$

$$\sum_{j\epsilon U_k} y_j^k \leq 1,\ \ k = 1, ..., n \tag{6}$$

$$\sum_{i\epsilon V} x_{ij}^k \leq |V|\, y_j^k,\ \ j\epsilon U_k, k = 1, ..., n \tag{7}$$

$$x_{ij}^k \geq 0,\ \ i\epsilon V,\ \ j\epsilon U_k, k = 1, ..., n \tag{8}$$

$$y_j^k \in \{0, 1\},\ \ j\epsilon U_k, k = 1, ..., n \tag{9}$$

The objective function (3) expresses the minimization of the total cost; constraints (4) ensure that the demand of each customer is satisfied; constraint (5) establishes that the number of facilities is p; constraints (6) guarantee that an optimal solution will contain no more than one facility per cluster; constraints (7) establish that no customer can use a not opened facility; constraints (8) provide lower bounds on the x_{ij}^k variables (it is worth noting that constraints (4) and (8) imply $0 \leq x_{ij}^k \leq 1$ $(i\epsilon V,\ j\epsilon U_k, k = 1, ..., n)$). Finally, (9) are the integrality constraints.

As stated in the introduction, $SpMP-UCPD$ models several combinatorial optimization problems. In the following, we will give some more examples of possible applications of $SpMP-UCPD$.

Let us consider a set of potential products to be brought out by a producer, in order to maximize his profit. A situation that arises often in practice is when the products can be characterized by a market price level. In this case, a typical marketing issue is to have products differentiated enough in their price. This can be obtained clustering them according to their market price and imposing to produce at most one product per cluster. The problem of choosing p products, one for each cluster, such that the producer's profit is maximized can be modeled by (3)-(9), setting $c_k = -p_k$ and $\theta_{ij} = -\eta_{ij}$, where p_k is the market price associated to cluster U_k, and η_{ij} represents a deviation from the common price p_k.

A second application arises in the design of transportation systems. In this case, the problem consists in delivering goods to customers, whose exact location inside given clusters is unknown. Examples of this type of problem are present, in particular, in the dispatching of cars in the automotive industry [29] and in the location of intermediate depots in Multi-Echelon Vehicle Routing Problems [13].

3 Solving the problem

Provided that, for any given solution y_j^k, the optimal solution in the x_{ij}^k variables of problem (3)-(9) is unique, it is obvious that the optimal solution $x_{il}^m, i \epsilon V$ of problem (3)-(9) becomes

$$x_{il}^m = \begin{cases} 1, \; if \;\; r_{il}^m = \min_{j\epsilon U_k, k=1,...,n} r_{ij}^k \\ 0, \; otherwise. \end{cases} \tag{10}$$

Let us first consider the probability distribution of the stochastic costs $r_{ij}^k, i\epsilon V, \;\; j\epsilon U_k, k = 1, ..., n$

$$\Pr\left\{r_{ij}^k \geq x\right\} = \Pr\left\{c_k + \theta_{ij} \geq x\right\} = \Pr\left\{\theta_{ij} \geq x - c_k\right\}. \tag{11}$$

The minimum cost $\overline{r}_i$ for a customer i is given by

$$\overline{r}_i = \min_{k=1,...,n} \left\{ \min_{j\epsilon U_k} \left\{ r_{ij}^k \right\} \right\} = \min_{k=1,...,n} \left\{ \min_{j\epsilon U_k} \left\{ c_k + \theta_{ij} \right\} \right\} = \min_{k=1,...,n} \left\{ c_k + \min_{j\epsilon U_k} \theta_{ij} \right\} \tag{12}$$

Let us define

$$\overline{\theta}_i^k = \min_{j\epsilon U_k} \theta_{ij}, \;\; i\epsilon V, k = 1, ..., n \tag{13}$$

The unknown probability distribution of the extreme (in this case the minimum) value $\overline{\theta}_i^k$ of the set of random variables θ_{ij} for any customer i is

$$H_k(x) = \Pr\left\{\overline{\theta}_i^k \geq x\right\}. \tag{14}$$

For (13), eq. (12) becomes

$$\overline{r}_i = \min_{k=1,...,n} \left\{ c_k + \overline{\theta}_i^k \right\} \tag{15}$$

and its probability distribution

$$G(x) = \Pr\{\overline{r}_i \geq x\} = \Pr\left\{\min_{k=1,\ldots,n}\left\{c_k + \overline{\theta}_i^k\right\} \geq x\right\}. \tag{16}$$

As

$$\min_{k=1,\ldots,n}\left\{c_k + \overline{\theta}_i^k\right\} \geq x \iff \left\{c_k + \overline{\theta}_i^k\right\} \geq x, \forall k = 1, ..., n \tag{17}$$

and the random variables $\overline{\theta}_i^k$ are independent, using (14) eq. (16) gives

$$G(x) = \Pr\left\{\min_{k=1,\ldots,n}\left\{c_k + \overline{\theta}_i^k\right\} \geq x\right\} = \prod_{k=1}^{n} \Pr\left\{\left\{c_k + \overline{\theta}_i^k\right\} \geq x\right\} =$$
$$\prod_{k=1}^{n} \Pr\left\{\overline{\theta}_i^k \geq x - c_k\right\} = \prod_{k=1}^{n} H_k(x - c_k). \tag{18}$$

In order to determine the unknown distributions H_k, $k = 1, ..., n$ let us consider that $\overline{\theta}_i^k \geq x \iff \theta_{ij} \geq x, \forall j \epsilon U_k$.

Using (2) one gets

$$H_k(x) = \Pr\left\{\overline{\theta}_i^k \geq x\right\} = \prod_{j\epsilon U_k} \Pr\{\theta_{ij} \geq x\} = [F(x)]^{Q_k} \tag{19}$$

where $Q_k = |U_k|$.

Substituting (19) into (18), one obtains

$$G(x) = \prod_{k=1}^{n} H_k(x - c_k) = \prod_{k=1}^{n} [F(x - c_k)]^{Q_k} \tag{20}$$

which is the distribution of the minimum cost $\overline{r}_i$ for a customer i as a function of three terms:

- the number of products in each cluster k, Q_k
- the deterministic cost, c_k
- the distribution of the random variables θ_{ij}, $F(x)$.

Unfortunately, the distribution (20) can not be used as it is because it implies the knowledge of the probability distribution $F(x)$, which is still unknown. In order to solve this problem, the *method of the asymptotic approximations* derived from the extreme value theory [12] will be used.

4 The asymptotic approximation of the probability distribution of the minimum cost

The method of the asymptotic approximations is based on the following observation: if under mild assumptions on the probability distribution $F(x)$ of the random terms θ_{ij}, the distribution of the stochastic variables r_{ij}^k (and then of their minimum $\overline{r}_i$) tends to a specific functional form as the system becomes large (i.e. when $|U_k|$, $k = 1, ..., n$ become large), then further specific knowledge of the probability distribution $F(x)$ is not needed.

The asymptotic theory of extreme values seems particularly tailored for this kind of problem, as it deals with the asymptotic behavior of maxima and minima of sequences of random variables. Because our problem actually deals with the minimum of a sequence of random variables, some results from that theory can then be used.

The only very mild assumption we take for the probability distribution $F(x)$ is that it is asymptotically exponential in its left tail, i.e. there is a constant $\beta > 0$ such that

$$\lim_{y \to -\infty} \frac{1 - F(x + y)}{1 - F(y)} = e^{\beta x}, \quad \forall x \epsilon \mathrm{IR} \tag{21}$$

Loosely speaking, this assumption states that $F(x)$ acts as an exponential function in its left tail. This property is widely used in the extreme value theory and defines the so-called *domain of attraction* of the double exponential distribution [12].

Moreover, many (actually, infinite) probability distributions show such a behavior, among them the following widely used distributions: Gumbel, Logistics, Gamma, and Generalized Exponential.

It will be proved that under assumption (21) $F(x)$ tends to a specific functional form as the number of potential facility locations in each cluster becomes large.

Consider the following simple but very important remark. If an arbitrary constant is added to the random variables θ_{ij} (and then to their minima $\overline{\theta}_i^k, k = 1, ..., n$), this leaves the solution of problem (3)-(9) unchanged.

As a specific choice, let us use the constant a_{Q_k} root of the equation

$$1 - F(a_{Q_k}) = 1/Q_k. \tag{22}$$

Replacing $\overline{\theta}_i^k$ with $\overline{\theta}_i^k - a_{Q_k}$ in (18) one has from (20)

$$G(x \mid Q_k) = \prod_{k=1}^{n} \left[F(x - c_k + a_{Q_k}) \right]^{Q_k}. \tag{23}$$

Let us assume that $Q_k, k = 1, ..., n$ are large enough to use $\lim_{Q_k \to \infty} G(x \mid Q_k)$ as an approximation of $G(x)$.

In order to obtain an explicit form of the probability distribution $G(x)$ consider the following

Theorem 1 *Under assumption (21)*

$$G(x) = \lim_{Q_k \to \infty} G(x \mid Q_k) = \exp\left(-Ae^{\beta x}\right), \tag{24}$$

where

$$A = \sum_{k=1}^{n} e^{-\beta c_k}. \tag{25}$$

Proof By (23) one has

$$G(x) = \lim_{Q_k \to \infty} G(x \mid Q_k) = \lim_{Q_k \to \infty} \prod_{k=1}^{n} \left[F\left(x - c_k + a_{Q_k}\right)\right]^{Q_k}. \tag{26}$$

As $\lim_{Q_k \to \infty} a_k = -\infty$, from assumption (21) one obtains

$$\lim_{Q_k \to \infty} \frac{1 - F(x - c_k + a_{Q_k})}{1 - F(a_{Q_k})} = e^{\beta(x - c_k)}. \tag{27}$$

By (22) and (27) one has

$$\lim_{Q_k \to \infty} F(x - c_k + a_{Q_k}) = \lim_{Q_k \to \infty} \left(1 - \frac{e^{\beta(x-c_k)}}{Q_k}\right) \tag{28}$$

and

$$\lim_{Q_k \to \infty} \left[F(x - c_k + a_{Q_k})\right]^{Q_k} = \lim_{Q_k \to \infty} \left(1 - \frac{e^{\beta(x-c_k)}}{Q_k}\right)^{Q_k} = \exp\left(-e^{\beta(x-c_k)}\right). \tag{29}$$

Substituting (29) into (26) one gets

$$G(x) = \prod_{k=1}^{n} \exp\left(-e^{\beta(x-c_k)}\right) = \exp\left(-Ae^{\beta x}\right). \tag{30}$$

□

A few words of comment on theorem 1 are worthwhile.

The main assumption is that $F(x)$ is an exponential function in its left tail. The second important assumption, which can be considered true in many applications of our concern, is that Q_k, $k = 1, ..., n$ are quite large, in order to justify the asymptotic approximation for $G(x)$.

The two above assumptions allow to calculate $G(x)$ as the asymptotic form of the probability distribution of the extreme value (the minimum cost in our problem) of a large set of i.i.d. random variables, which is consistent with the objective of using those facilities which minimize the expected total cost.

It is interesting to observe that consequently $G(x)$ becomes a Gumble (or double exponential) distribution [14], which will imply for the allocation variables x_{ij}^k of our problem a multinomial Logit form, as it will be show in Section 5.

5 Finding the expected optimal value of the allocation variables

The expected value of the minimum cost $\overline{r}_i$ for a customer i becomes

$$\widetilde{r}_i = \mathbb{E}\left(\overline{r}_i\right) = \int_{-\infty}^{+\infty} x dG(x) = -\int_{-\infty}^{+\infty} x \exp\left(-Ae^{\beta x}\right) Ae^{\beta x} \beta dx. \tag{31}$$

Substituting for $t = Ae^{\beta x}$ one gets

$$\begin{aligned} \widetilde{r}_i &= -1/\beta \int_0^{+\infty} \log(t/A) e^{-t} dt = \\ &= -1/\beta \int_0^{+\infty} e^{-t} \log t dt + 1/\beta \log A \int_0^{+\infty} e^{-t} dt = \\ &= \gamma/\beta + 1/\beta \log A = \\ &= 1/\beta(\gamma + \log A) \end{aligned} \tag{32}$$

where $\gamma = -\int_0^{+\infty} e^{-t} \log t \; dt \simeq 0.5772$ is the Euler constant.

By (32) and (25), the expected value of the minimum total cost $\widetilde{R}$ becomes

$$\begin{aligned} \widetilde{R} &= \sum_{i\epsilon V} \widetilde{r}_i = \sum_{i\epsilon V} 1/\beta(\gamma + \log A) = \\ &= |V| \, 1/\beta(\gamma + \log \sum_{k=1}^{n} e^{-\beta c_k}). \end{aligned} \tag{33}$$

We are now interested in finding the expected value $\widetilde{x}_{ij}^{k*}$ of the optimal solution x_{ij}^{k*} of the $SpMP - UCPD$ (3)-(9).

The following theorem holds

Theorem 2 *For any given solution y_j^k of problem (3)-(9), the expected value $\widetilde{x}_{ij}^{k*}$ of the optimal solution x_{ij}^{k*} is*

$$\widetilde{x}_{ij}^{k*} = \frac{e^{-\beta c_k}}{\sum_{l=1, l/\exists j \epsilon U_l, y_j^l = 1}^{n} e^{-\beta c_l}}, \quad i\epsilon V, \quad k = 1, ..., n/\exists j \epsilon U_k, y_j^k = 1. \tag{34}$$

Proof From the Total Probability Theorem [5] and assumptions (21), one obtains

$$\begin{aligned} \widetilde{x}_{ij}^{k*} &= \int_{-\infty}^{+\infty} Pr\left\{x < c_k \leq x + dx\right\} Pr\left\{c_s > x, \; \forall s \neq k\right\} = \\ &= \int_{-\infty}^{+\infty} \beta e^{\beta(x - c_k)} exp(-e^{\beta x} A) dx = \\ &= e^{-\beta c_k} \int_0^{+\infty} e^{-At} dt = \frac{e^{-\beta c_k}}{A} = \\ &= \frac{e^{-\beta c_k}}{\sum_{l=1}^{n} e^{-\beta c_l}} \end{aligned}$$

where $t = e^{\beta x}$.

$\square$

Let us observe that the formulation of $\widetilde{x}_{ij}^{k*}$ in (34) represents a multinomial Logit model. This model is widely used in choice theory and particularly suitable to describe the dispersion of customer preferences (due in our case to the cost stochasticity) among different alternatives (the facilities).

Moreover, it is interesting to note that in the literature [6] the multinomial Logit model is derived by assuming for the probability distribution of the random terms the following Gumbel (or double exponential) distribution

$$F(x) = \exp\left(-e^{\beta x}\right). \tag{35}$$

Here, vice versa, we have decided, as $F(x)$ is not observable, not to hypothesize any specific form for it, but only its asymptotic exponential behavior (see (21)), and assume that the number of alternatives in any cluster is quite large, which is easy observable and often true. As a result we get the distribution (24) for $G(x)$, which is actually a Gumbel distribution, but obtained *a posteriori* by theorem 1.

6 Finding the optimal solution of the location variables

Until now, we have assumed that the solution y_j^k (i.e., the location variables) of problem (3)-(9) is given. Now, we drop this assumption and look for such a solution.

As (34) is still valid, problem (3)-(9) has to be solved in y_j^k, only. It is easy to see that, using (33), the following integer deterministic nonlinear problem holds

$$\min_{\{y_j^k\}} |V|\, 1/\beta(\gamma + \log \sum_{k=1}^{n} y_j^k e^{-\beta c_k}) + \sum_{k=1}^{n} \sum_{j \epsilon U_k} f_j^k y_j^k \tag{36}$$

subject to

$$\sum_{k=1}^{n} \sum_{j \epsilon U_k} y_j^k = p \tag{37}$$

$$\sum_{j \epsilon U_k} y_j^k \leq 1, \quad k = 1, ..., n \tag{38}$$

$$y_j^k \in \{0, 1\}, \quad j \epsilon U_k, k = 1, ..., n. \tag{39}$$

By solving problem (36)-(39), a p-median solution y_j^{k*} and the corresponding expected value of the minimum total cost can be obtained.

It is easy to prove, e.g. by reduction from the Dominating Set problem, that problem (36)-(39) is NP-complete. Because of that, for solving real-life instances of this problem in a reasonable computational time, efficient heuristics are required. A particular efficient and effective one, developed for a similar problem, can be found in [28].

7 Conclusions

In this paper the p-median problem where the costs for using the facilities are stochastic with unknown probability distribution has been considered. It has been shown that, under a very mild and reasonable assumption on the shape of this unknown probability distribution, and assuming the number of potential facility locations quite large, some useful results of the extreme value theory become available.

In particular, by applying the method of the asymptotic approximations the above probability distribution becomes a Gumbel (or double exponential) distribution. This result is quite important because it allows us to derive for the expected optimal value of the x_{ij}^k variables (i.e. the variables which allocate demand to supply) a multinomial Logit model. This model is widely used in choice theory and particularly suitable to describe the dispersion of customer preferences (due in our case to the cost stochasticity) among different alternatives (the facilities).

In order to find the optimal value of the facility location variables an integer deterministic nonlinear problem derived from the original stochastic p-median problem has been introduced. This problem is NP-complete and for solving real-life size instances in a reasonable computational time heuristic algorithms available in the literature can be used.

8 Acknowledgments

This project has been partially supported by the Italian Ministry of University and Research, under the 2005 PRIN Project “Infomobility and Freight Distribution in Urban Areas”.

References

1. Berman, O., Drezner, Z.: The p-median problem under uncertainty. European Journal of Operational Research **doi: 10.1016/j.ejor.2007.05.045** (2007)
2. Berman, O., Krass, D.: Facility location problems with stochastic demands and congestion. In: Z. Drezner, H.W. Hamacher (eds.) Facility Location: Applications and Theory, pp. 329–371. Springer-Verlag, Berlin (2002)
3. Berman, O., Wang, J.: Probabilistic location problems with discrete demand weights. Networks **57**, 47–57 (2004)
4. Berman, O., Wang, J.: The 1-median and 1-antimedian with uniform distributed demand. Journal of the Operational Research Society (in press)
5. DeGroot, M., Schervish, M.: Probability and Statistics (Third Edition). Addison Wesley, Reading (2002)
6. Domencich, T., McFadden, D.: Urban travel dynamics: a behavioral analysis. North Holland, Amsterdam (1975)
7. Drezner, T., Drezner, Z., Shiode, S.: A threshold satisfying competitive location model. Journal of Regional Science **42**, 287–299 (2002)
8. Drezner, Z.: Sensitivity analysis of the optimal location of a facility. Naval Research Logistics Quaterly **32**, 209–224 (1985)
9. Drezner, Z., Shiode, S.: A distribution map for the one-median location problem on a network. European Journal of Operational Research **179**, 1266–1273 (2007)

10. Drezner, Z., Wesolowsky, G.O.: Optimum location probabilities in the l_p distance weber problem. Transportation Science **15**, 85–97 (1981)
11. Frank, H.: Optimum locations on graphs with correlated normal demands. Operations Research **15**, 552–557 (1966)
12. Galambos, J.: The asymptotic theory of extreme order statistics. John Wiley, New York (1978)
13. Gonzalez Feliu, J., Perboli, G., Tadei, R., Vigo, D.: The two-echelon capacitated vehicle routing problem. Transportation Science **forthcoming**
14. Gumbel, E.J.: Statistics of Extremes. Columbia University Press, New York (1958)
15. Juel, H.: Bounds in the generalized weber problem under locational uncertainty. Operations Research **29**, 1219–1227 (1981)
16. Leonardi, G.: The use of random utility theory in building location-allocation models. In: J.F. Thisse, H.G. Zoller (eds.) Locational analysis of public facilities, pp. 357–383. North Holland, Amsterdam (1983)
17. Leonardi, G.: The structure of random utility models in the light of the asymptotic theory of extremes. In: M. Florian (ed.) Transportation planning models, pp. 107–133. Elsevier, Amsterdam (1984)
18. Leonardi, G., Tadei, R.: Random utility demand models and service location. Regional Science and Urban Economics **14**, 399–431 (1984)
19. Louveaux, F.V.: Stochastic location analysis. Location Science **1**, 127–154 (1993)
20. Louveaux, F.V., Peeters, D.: A dual-based procedure for stochastic facility location. Operations Research **40**, 564–573 (1992)
21. Louveaux, F.V., Thisse, J.F.: Production and location on a network under uncertainty. Operations Research Letters **4**, 145–149 (1985)
22. Mirchandani, P.B.: The p-median problem and generalizations. In: P.B. Mirchandani, R.L. Francis (eds.) Discrete location theory, pp. 55–117. John Wiley, New York (1990)
23. Mirchandani, P.B., Odoni, A.R.: Location of medians on stochastic networks. Transportation Science **13**, 85–97 (1979)
24. Mirchandani, P.B., Oudjit, A.: Localizing 2-medians on probabilistic and deterministic tree networks. Networks **10**, 329–350 (1980)
25. Ng, L.F.Y., Tuan, C.: Location decisions of manufacturing fdi in china: Implications of china's wto accession. Journal of Asian Economics **14**, 51–72 (2003)
26. ReVelle, C.S., Eiselt, H.A.: Location analysis: a synthesis and survey. European Journal of Operational Research **165**, 1–19 (2006)
27. ReVelle, C.S., Eiselt, H.A., Daskin, M.S.: A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research **doi: 10.1016/j.ejor.2006.12.044** (2006)
28. Ricciardi, N., Tadei, R., Grosso, A.: Optimal facility location with random throughput costs. Computers and Operations Research **29**, 593–607 (2002)
29. Tadei, R., Perboli, G., Della Croce, F.: A heuristic algorithm for the autocarrier transportation problem. Transportation Science **36**(1), 55–62 (2002)
30. Tansel, B.C., Francis, R.L., Lowe, T.J.: Location on networks: a survey. Management Science **29**, 482–511 (1983)
31. Weaver, J.R., Church, R.L.: Computational procedures of location problems on stochastic networks. Transportation Science **17**, 168–180 (1983)
32. Wesolowsky, G.O.: Probabilistic weights in the one-dimensional facility location problem. Management Science **24**, 224–229 (1977)