Annals of Operations Research 111, 75–88, 2002

Optimal Positioning of a Mobile Service Unit on a Line

E. CARRIZOSA
Departamento de Estadística e Investigación Operativa, Facultad de Matemáticas, c/ Tarfia s/n, 41012-Sevilla, Spain

M. MUÑOZ-MÁRQUEZ manuel.munoz@uca.es
Departamento de Estadística e Investigación Operativa, Escuela Superior de Ingeniería de Cádiz, c/ Sacramento 82, 11002-Cádiz, Spain

J. PUERTO
Departamento de Estadística e Investigación Operativa, Facultad de Matemáticas, c/ Tarfia s/n, 41012-Sevilla, Spain

Abstract. In this paper we address the problem of locating a mobile response unit when demand is distributed according to a random variable on a line. Properties are proven which reduce the problem to locating a non-mobile facility, transforming the original optimization problem into an one-dimensional convex program.

In the special case of a discrete demand (a simple probability measure), an algorithm which runs in expected linear time is proposed.

Keywords: location, facilities planning and design, stochastic location, patrolling

1. Introduction

The single-facility location problem consists of determining the location for one server minimizing some performance measure which is a function of the distances from that server to the demand points, see [19] for further details.

The most popular model is the Weber problem, see [8], in which the objective function is the average distance between the location of the server and the demand, assumed to be concentrated on a finite set of points. The case of continuous demand has also been addressed, see, e.g., [4], by considering then as performance measure the expected distance between the server and the demand, the latter assumed to be random.

When the facility is not static but a patrolling unit, the decision variable is not any more the facility location but the patrolling area, modeled in [5,6,20] by means of a random variable.

When no constrains are imposed over the shape of the patrolling area, the optimal (in the sense of minimal average distance to demand) patrolling area consists of fixing the server at the location which is optimal for the static problem, see [6] for further details. This means that if one wished to locate a service which could be mobile, it is preferable to fix its position to maintain it patrolling.

If patrolling is required or the server must has a non-null area, some constrains must be added to the model. The case in which the patrolling area is an n-dimensional rectangle with fixed volume has been studied in [5]. The non-null measure case appears, for example, when one wants to locate an industrial park in a city or a chip in an integrated circuit. These examples show that, as a consequence of technological or infrastructure constrains, modeling the facility to be located as a region may be more accurate than modeling it as a single point.

This paper supposes that there are no congestion or queue in the system. The location problem with congestions has been addressed in [2,7,13,14].

The one-dimensional case has a richer structure than the general one, and there are a variety of situations such as disk arm, idle elevators, smoke alarms, patrol cars, service units along an oil pipeline, emergency vehicles in a highway, in which the demand is spread over a line, as in [1,13,15], these papers shows some particular instances of the above-mentioned result in [6].

Other extensions of the one-dimensional case can be found in [10,12,16,21], where some special policies are consider in order to reduce the seek time for read/write operations on disks.

In this paper we consider the one-dimensional case of the problem in [5]: the demand is distributed (following a random variable in the real line), and the server (e.g., a police patrol in a highway) has a patrolling area whose length cannot be lower than a given threshold value. This particular case enable further results and more powerful resolution procedures.

The remaining of the paper is organized as follows. Section 2 introduces the model, reviews existing results from [5], and develops general properties that exploit the one-dimensional nature of the problem. Then, section 3 addresses the particular case in which the demand is concentrated at a finite set of points, and an algorithm which runs in expected linear time is proposed. The last section is devoted to conclude the paper.

2. Finding optimal solutions: The general case

Let us consider the problem of locating a patrolling unit in the line, whose patrolling area has the form $I = [c-k, c+k]$, and minimizes its expected distance to the demand, distributed in the line following a random variable A with distribution function F, and finite first moment, i.e., $E(A) < \infty$.

For $k > 0$, the expected distance between A and I is

$$\begin{aligned}\overline{d}(c,k) &= \frac{1}{2k}\int_{c-k}^{c+k}\int |u-a|\,\mathrm{d}F(a)\,\mathrm{d}u\\ &= \int \frac{1}{2k}\int_{c-k}^{c+k} |u-a|\,\mathrm{d}u\,\mathrm{d}F(a).\end{aligned}$$

Note that $\overline{d}(c,k)$ is well defined, since A has, by assumption, finite first moment, see [18].

For the degenerate case in which the patrolling area reduces to the point c (i.e., when $k = 0$), then the average distance is given by

$$\overline{d}(c, 0) = \int |c - a| \, \mathrm{d}F(a). \tag{1}$$

The problem addressed consists of determining the patrolling area (interval) of length at least $2l$ ($l > 0$), minimizing its average distance to the demand. This yields the optimization problem

$$\begin{aligned} \min_{(c,k)\in\mathbb{R}\times\mathbb{R}^+} \quad & \overline{d}(c, k) \\ \text{s.t.} \quad & k \geqslant l. \end{aligned} \tag{$\mathrm{P}_\geqslant$}$$

The results in [5] reduce the resolution of ($\mathrm{P}_\geqslant$) to a univariate convex differentiable unconstrained problem, thus solvable by a large variety of algorithms available in the literature, see [11,17]. To do such reduction, define first ($\mathrm{P}_=$) as the problem of locating a patrolling unit, whose patrolling area has length equal to $2l$,

$$\min_{c\in\mathbb{R}} \overline{d}(c, l). \tag{$\mathrm{P}_=$}$$

It turns out that problem ($\mathrm{P}_\geqslant$) can be reduced to ($\mathrm{P}_=$):

Theorem 1 (Cf. theorem 5 of [5]). For each $c \in \mathbb{R}$, the function $\overline{d}(c, \cdot)$ is non-decreasing. In particular, if c^* solves ($\mathrm{P}_=$), then (c^*, l) solves ($\mathrm{P}_\geqslant$).

Moreover, as a direct rephrasing of theorem 8 of [5], we obtain that (partial) derivatives of $\overline{d}$ can be written as expected distances.

Theorem 2. For each $k > 0$, the function $\overline{d}(\cdot, k)$ is convex, differentiable and its derivative being is given by

$$\frac{\partial}{\partial c}\overline{d}(c, k) = \frac{\overline{d}(c + k, 0) - \overline{d}(c - k, 0)}{2k}$$

where $\overline{d}(c, 0)$ was introduced in (1).

The next step will be the derivation of optimality conditions for problem ($\mathrm{P}_=$). To do this, we first observe that the random variable which represents the patrolling area is $c + U$, where U is uniformly distributed in the segment $[-l, l]$. Hence, denoting A' as $A - U$, and then $|c + U - A| = |c - A'|$, the problem of determining the patrolling area $[c - l, c + l]$ minimizing the expected distance to A is equivalent to the problem of locating a point minimizing the expected distance to a transformed demand A'. In other words, ($\mathrm{P}_=$) – thus also ($\mathrm{P}_\geqslant$) – is reduced to a one-dimensional Weber problem with demand A'.

Lemma 1. The distribution function of $A - U$ is given by

$$G_l(c) = \frac{1}{2l}\int_{c-l}^{c+l} F(u)\,\mathrm{d}u.$$

Proof. Let $D = A - U$ and $\bar{p}$ its probability measure. By elementary probability calculus we have

$$\begin{aligned} G_l(c) &= \bar{p}(A - U \leqslant c) \\ &= \frac{1}{2l}\int_{-l}^{l} p(A - u \leqslant c)\,\mathrm{d}u = \frac{1}{2l}\int_{-l}^{l} p(A \leqslant c + u)\,\mathrm{d}u \\ &= \frac{1}{2l}\int_{-l}^{l} F(c+u)\,\mathrm{d}u = \frac{1}{2l}\int_{c-l}^{c+l} F(u)\,\mathrm{d}u. \end{aligned}$$

□

Since a point is optimal for the one-dimensional Weber problem if and only if it is a median of the distribution of the demand, see [3, p. 54], we have the following optimality condition.

Theorem 3. The function G_l is continuous. Thus c^* is an optimal solution if and only if it verifies

$$G_l(c^*) = \frac{1}{2l}\int_{c^*-l}^{c^*+l} F(u)\,\mathrm{d}u = \frac{1}{2}.$$

Proof. Since F is integrable (is bounded), its integral with upper and lower limit functions of k is continuous. Thus G_l is continuous, and a point c^* is a median if and only if $G_l(c^*) = 1/2$, which is the assertion. □

Remark 1. The distribution function G_l is a continuous function, so the equation $G_l(c^*) = 1/2$ always has at least one solution.

Remark 2. Since $\overline{d}(\cdot,)$ is a convex function, a point c is an optimal solution to ($\mathrm{P}_=$) if and only if its partial derivative with respect to c, given in theorem 2, is 0.

This condition is

$$\frac{\partial}{\partial c}\overline{d}(c,l) = \frac{\overline{d}(c+l,0) - \overline{d}(c-l,0)}{2l} = 0,$$

and it can be reduced to the one given in theorem 3.

In the following for each l consider $c(l)$ the optimal location center of the server, when the patrolling area has a length of $2l$. The next example shows how this equation can sometimes be solved.

Example 1. Let A be an exponential random vector and

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - \mathrm{e}^{-\lambda x} & \text{if } x \geqslant 0, \end{cases}$$

the distribution functions of the demand.

The optimality condition is $G_l(c) = 1/2$, where

$$\begin{aligned} G_l(c) &= \frac{1}{2l} \int_{c-l}^{c+l} F(u)\,\mathrm{d}u \\ &= \begin{cases} 0 & \text{if } c < -l, \\ \dfrac{1}{2l}\left(l + c + \dfrac{\mathrm{e}^{-\lambda(l+c)} - 1}{\lambda}\right) & \text{if } -l \leqslant c < l, \\ \dfrac{1}{2l}\left(2l + \dfrac{\mathrm{e}^{-\lambda(l+c)} - \mathrm{e}^{-\lambda(l-c)}}{\lambda}\right) & \text{if } c \geqslant l. \end{cases} \end{aligned}$$

Thus the function that gives the optimal location of the center as function of l is

$$c(l) = \begin{cases} c_1(l) & \text{if } c_1(l) \leqslant l, \\ c_2(l) & \text{if } c_1(l) \geqslant l, \end{cases}$$

where

$$\begin{aligned} c_1(l) &= \frac{W(-\mathrm{e}^{-\lambda l - 1}) + 1}{\lambda}, \\ c_2(l) &= \frac{1}{\lambda} \log \frac{\mathrm{e}^{2\lambda l} - 1}{\lambda l} - l, \end{aligned}$$

W is the Lambert's W, and $W(a)$ is the solution to $we^w = a$, see [9].

Figure 1 depicts the function c, for an exponential distribution with mean value of 1. Note that when l goes to infinity the optimal location of the service center approaches to 1 the mean value of A.

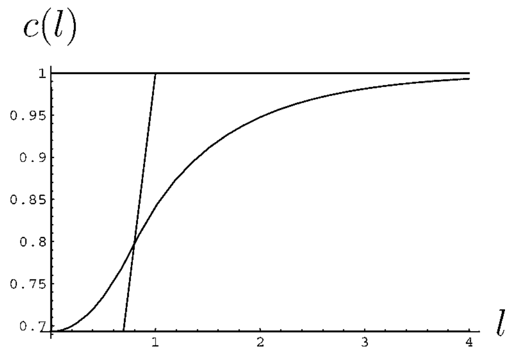

Figure 1. Function $c(l)$ for Exp(1).

The behavior of function c when l goes to infinity is a general property as shows the next theorem. Let $\mathrm{E}(A)$ be the mean value of A. One then has

Theorem 4.

$$\lim_{l \to +\infty} c(l) = \mathrm{E}(A).$$

Proof. The optimality condition given in remark 2 is $\frac{\partial}{\partial c}\overline{d}(c, l) = 0$, that is,

$$\begin{aligned}
\frac{\partial}{\partial c}\overline{d}(c, l) &= \frac{1}{2l}\int \big(|c+l-a| - |c-l-a|\big)\,\mathrm{d}F(a)\\
&= \frac{1}{2l}\bigg(\int_{-\infty}^{c-l} \big(|c+l-a| - |c-l-a|\big)\,\mathrm{d}F(a)\\
&\qquad + \int_{c-l}^{c+l} \big(|c+l-a| - |c-l-a|\big)\,\mathrm{d}F(a)\\
&\qquad + \int_{c+l}^{+\infty} \big(|c+l-a| - |c-l-a|\big)\,\mathrm{d}F(a)\bigg)\\
&= \frac{1}{2l}\bigg(\int_{-\infty}^{c-l} 2l\,\mathrm{d}F(a) + \int_{c-l}^{c+l} 2(c-a)\,\mathrm{d}F(a) + \int_{c+l}^{+\infty} -2l\,\mathrm{d}F(a)\bigg) = 0.
\end{aligned}$$

Since we are considering the case $l > 0$, this condition reduces to

$$\int_{-\infty}^{c-l} 2l\,\mathrm{d}F(a) + \int_{c-l}^{c+l} 2(c-a)\,\mathrm{d}F(a) + \int_{c+l}^{+\infty} -2l\,\mathrm{d}F(a) = 0. \tag{2}$$

By assumption, $\mathrm{E}(A)$ exists (and is finite), thus

$$\lim_{u \to +\infty} \int_{u}^{+\infty} a\,\mathrm{d}F(a) = 0.$$

Hence,

$$\begin{aligned}
\lim_{l \to +\infty} \int_{-\infty}^{c-l} 2l\,\mathrm{d}F(a) &\leqslant \lim_{l \to +\infty} \int_{-\infty}^{c-l} (c-a)\,\mathrm{d}F(a)\\
&= \lim_{l \to +\infty} \int_{l}^{+\infty} a\,\mathrm{d}F(a) = 0,
\end{aligned}$$

and

$$\lim_{l \to +\infty} \int_{c+l}^{+\infty} l\,\mathrm{d}F(a) \leqslant \lim_{l \to +\infty} \int_{c+l}^{+\infty} a\,\mathrm{d}F(a) = 0.$$

On the other hand,

$$\begin{aligned}
\lim_{l \to +\infty} \int_{c-l}^{c+l} (c-a)\,\mathrm{d}F(a) &= \lim_{l \to +\infty} \int_{c-l}^{c+l} c\,\mathrm{d}F(a) - \lim_{l \to +\infty} \int_{c-l}^{c+l} a\,\mathrm{d}F(a)\\
&= c - \mathrm{E}(A).
\end{aligned}$$

Hence, it follows that the limit of (2) is $c - \mathrm{E}(A) = 0$, thus

$$\lim_{l \to +\infty} c(l) = \bar{A}. \qquad \square$$

When the patrolling area has length $l = 0$, the optimal solution is the median of A, whereas for length l tending to infinity, the optimal interval has its center $c(l)$ at $\mathrm{E}(A)$, the mean of A. This fact is due that, when $l \to +\infty$, the optimality condition (2) converges to

$$\int_{-\infty}^{+\infty} 2(c - a)\,\mathrm{d}F(a),$$

which is the well-known optimality condition for the mean.

One may wonder if, for l varying from 0 to ∞, $c(l)$ describes a monotonic trajectory. The next example shows that this may not be the case.

Example 2. Let A_1, A_2, A_3, A_4 be random variables whose probability distributions are mixtures with the same weights of three normal distributions with means and variances:

Case	Means			Variances			Mixture Mean	Mixture Median
1	-1	1	1	1	1	1/16	1/3	0.79297
2	-1	1	2	1	1	1/16	2/3	1.00407
3	-1	1	3	1	1	1/16	1	1.00407
4	-1	1	4	1	1	1/16	2/3	1.00407

Figure 2 depicts the corresponding c functions. The problem is solved numerically with Mathematica 4.0.

Let $\mathcal{M}$ be the set of medians of A, which is known to be an (possibly degenerate) interval. The median of $A - U$ is related to the median of A, giving the next localization theorem.

Theorem 5. Any optimal solution to ($\mathrm{P}_=$) belongs to the interval $\mathcal{M} + [-l, l]$.

Proof. Let x be an optimal solution to ($\mathrm{P}_=$), and, by contradiction, suppose that $x \notin \mathcal{M} + [-l, l]$.

If $p(A \leqslant x - l) > 1/2$ then

$$G_l(x) = p(A - U \leqslant x) = p(A \leqslant x + U) \geqslant p(A \leqslant x - l) > 1/2$$

and x would not be optimal.

On the other hand, if $p(A \leqslant x + l) < 1/2$ then $G_l(x) < 1/2$, and again, x would not be optimal.

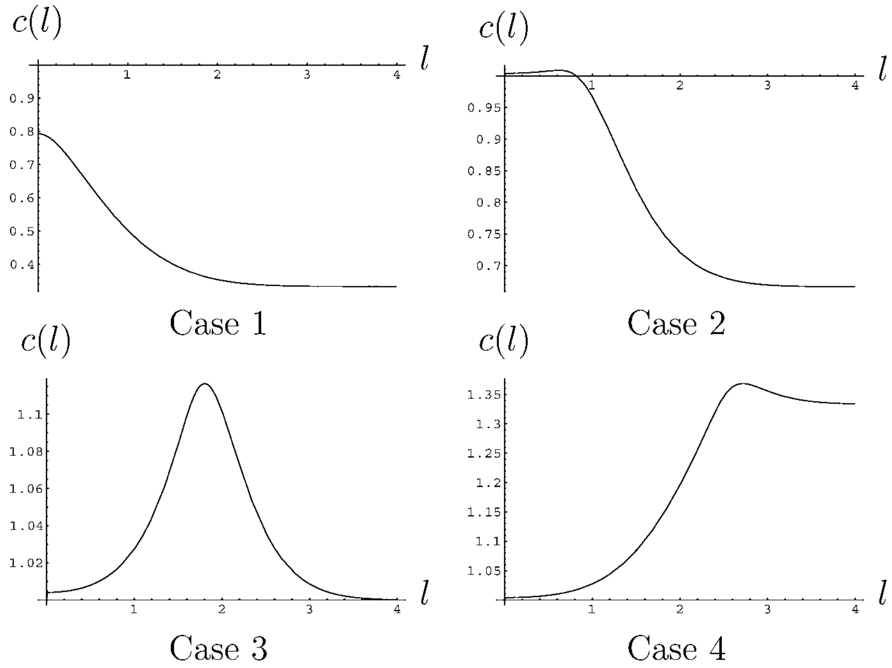

Figure 2. Functions c for mixtures in example 2.

The remaining x verifies that $p(A \leqslant x - l) \leqslant 1/2 \leqslant p(A \leqslant x + l)$. Thus $F(x - l) \leqslant 1/2 \leqslant F(x + l)$ and there exists a median in $x + [-l, l]$, which contradicts the assumption. □

The next theorem gives a new localization result assuming certain symmetry property for the probability distribution of A.

Theorem 6. Let m be a point such that

(1) $F(m^-) + F(m) = 1$, where $F(m^-) = \lim_{x \uparrow m} F(x)$.

(2) A is symmetrical with respect to m in the interval $[m - l, m + l]$.

Then m is an optimal solution to ($P_=$).

Proof. By the second assumption, one has that

$$p(m - u < A \leqslant m) = p(m \leqslant A < m + u) \quad \forall u \in [0, l].$$

Hence

$$F(m) - F(m - u) = F\big((m + u)^-\big) - F(m^-) \quad \forall u \in [0, l]$$

and

$$F(m^-) + F(m) = F(m - u) + F\big((m + u)^-\big) \quad \forall u \in [0, l].$$

Using this equation one obtains

$$
\begin{aligned}
G_l(m) &= \frac{1}{2l}\int_{-l}^{l} F(m-u)\,\mathrm{d}u \\
&= \frac{1}{2l}\left(\int_{-l}^{0} F(m-u)\,\mathrm{d}u + \int_{0}^{l} F(m-u)\,\mathrm{d}u\right) \\
&= \frac{1}{2l}\int_{0}^{l} \big(F(m-u) + F(m+u)\big)\,\mathrm{d}u \\
&= \frac{1}{2l}\int_{0}^{l} \big(F(m-u) + F\big((m+u)^-\big)\big)\,\mathrm{d}u \\
&= \frac{1}{2l}\int_{0}^{l} \big(F(m^-) + F(m)\big)\,\mathrm{d}u \\
&= \frac{1}{2}\big(F(m^-) + F(m)\big) = \frac{1}{2},
\end{aligned}
$$

which implies the optimality of m. □

Remark 3. This result implies that for every symmetrical distribution modeling the demand, the median of the distribution is the optimal solution of $(\mathrm{P}_=)$.

Remark 4. If F is continuous then $F(m) + F(m^-) = 1$ implies that m is a median, but this is not true in the general case.

3. Finding optimal solutions: The discrete case

In this section we particularize the results obtained in the preceding section to the discrete case in order to develop an efficient algorithm to solve the problem. Consider the case in which the demand is discrete: the random variable A takes the values $x_1, \ldots, x_m$ with probabilities $p_1, \ldots, p_m$.

We will give an algorithm running in expected linear time. To do this, the problem is reduced to finding the median of a new random variable, and then a modified version of a known algorithm is applied.

We must note that the algorithm that find the median in linear time in not applicable to the transformed demand distribution, because such distribution is not a discrete one. The others ways to compute the median of the transformed demand distribution need sort the values of the original demand distribution which means that they do not run linear time.

This algorithm can be used to find the median of a sample whose data are clustered in intervals, or to locate a point where the demand is a mixture of uniform distributions over equal length segments.

First of all, an explicit form for the function $G_l(c)$ is needed. This form is given in the next lemma with straight proof.

Lemma 2. Let ψ be the function given by $\psi(u) = \int_{-\infty}^{u} F(x)\,\mathrm{d}x$. Then

$$\psi(u) = u \sum_{x_i \leqslant u} p_i - \sum_{x_i \leqslant u} p_i x_i.$$

In order to obtain a simpler expression for ψ let y, w denote the following vectors:

$$\begin{array}{lll} y_i = x_i - l, & w_i = p_i, & i = 1, \ldots, m, \\ y_{m+i} = x_i + l, & w_{m+i} = -p_i, & i = 1, \ldots, m. \end{array}$$

With this notation the next theorem holds.

Theorem 7. One has

$$G_l(c) = \sum_{y_i \leqslant c} w_i (c - y_i).$$

Proof. Since

$$G_l(c) = \int_{c-l}^{c+l} F(u)\,\mathrm{d}u,$$

we have by lemma 2 that

$$\begin{aligned} G_l(c) &= \int_{c-l}^{c+l} F(u)\,\mathrm{d}u \\ &= \psi(c+l) - \psi(c-l) \\ &= \sum_{x_i \leqslant c+l, i=1,\ldots,m} p_i (c + l - x_i) - \sum_{x_i \leqslant c-l, i=m+1,\ldots,2m} p_i (c - l - x_i) \\ &= \sum_{y_i \leqslant c, i=1,\ldots,m} w_i (c - y_i) + \sum_{y_{m+i} \leqslant c, i=1,\ldots,m} w_i (c - y_i) \\ &= c \sum_{y_i \leqslant c, i=1,\ldots,2m} w_i - \sum_{y_i \leqslant c, i=1,\ldots,2m} w_i y_i. \end{aligned}$$

□

Now we will develop rules to phantom points which cannot be optimal solutions. To this end, we introduce further notation: For each triplet of vectors $Y, W, Z \in \mathbb{R}^{2m}$, and each subset I of $N = \{1, \ldots, 2m\}$, let Γ_I be given by

$$\Gamma_I(c, Y, W, Z) = c \sum_{Y_i \leqslant c, i \in I} W_i - \sum_{Y_i \leqslant c, i \in I} Z_i.$$

Remark 5. If z is the vector with coordinates $z_i = w_i y_i$, it follows that

$$G_l(c) = \frac{1}{2l} \Gamma_N(c, y, w, z).$$

The following two lemmas present some elimination rules.

Lemma 3. Given $I \subset N$ and Y_j, let $I' = \{i \in I\colon\ Y_i \leqslant Y_j\}$. One has

$$\Gamma_{I'}(c, Y, W, Z) = \Gamma_I(c, Y, W, Z) \quad \forall c \leqslant Y_j.$$

Proof. If $c \leqslant Y_j$ then $\{i \in I\colon\ Y_i \leqslant c\} = \{i \in I'\colon\ Y_i \leqslant c\}$. Hence, since the index sets are equal in both sums, we obtain that $\Gamma_I(c, Y, W, Z) = \Gamma_{I'}(c, Y, W, Z)$. □

Lemma 4. Given $I \subset N$ and Y_j, let I', W' and Z' be given by

$$I' = \{i \in I\colon\ Y_i > Y_j\} \cup \{j\},$$

$$W' = \begin{cases} W'_i = W_i, & \forall i \neq j, \\ W'_j = \displaystyle\sum_{Y_i \leqslant Y_j, i \in I} W_i, & \end{cases}$$

$$Z' = \begin{cases} Z'_i = Z_i, & \forall i \neq j, \\ Z'_j = \displaystyle\sum_{Y_i \leqslant Y_j, i \in I} Z_i. & \end{cases}$$

One has that

$$\Gamma_{I'}\big(c, Y, W', Z'\big) = \Gamma_I(c, Y, W, Z) \quad \forall c \geqslant Y_j.$$

Proof. Given $c \geqslant Y_j$, it follows from the definition that

$$\begin{aligned}
\Gamma_I(c, Y, W, Z) &= c \sum_{Y_i \leqslant c, i \in I} W_i - \sum_{Y_i \leqslant c, i \in I} Z_i \\
&= c\Bigg(\sum_{Y_i \leqslant Y_j, i \in I} W_i + \sum_{Y_j < Y_i \leqslant c, i \in I} W_i\Bigg) - \Bigg(\sum_{Y_i \leqslant Y_j, i \in I} Z_i + \sum_{Y_j < Y_i \leqslant c, i \in I} Z_i\Bigg) \\
&= c\Bigg(W'_j + \sum_{Y_j < Y_i \leqslant Y_j, i \in I} W_i\Bigg) - \Bigg(Z'_j + \sum_{Y_j < Y_i \leqslant c, i \in I} Z_i\Bigg) \\
&= c \sum_{Y_i \leqslant c, i \in I'} W'_i - \sum_{Y_i \leqslant c, i \in I'} Z'_i \\
&= \Gamma_{I'}\big(c, Y, W', Z'\big).
\end{aligned}$$

□

These lemmas lead to the algorithm presented in table 1. (#I denotes the cardinality of I.)

Remark 6. For the evaluation of $\Gamma_I(y_j, y, w, z)$ one only has to evaluate $\sum_{y_i \leqslant y_j, i \in I} w_i$ and $\sum_{y_i \leqslant y_j, i \in I} z_i$. Hence, the actualization of step 0 does not require additional time.

Table 1
Average linear time algorithm.

Step 0: Initialization
For $i = 1, \ldots, m$ $\{y_i \leftarrow x_i - l \;\; y_{m+i} \leftarrow x_i + l\}$
For $i = 1, \ldots, m$ $\{w_i \leftarrow p_i \;\; w_{m+i} \leftarrow -p_i + l\}$
For $i = 1, \ldots, 2m$ $z_i \leftarrow y_i w_i$
$I \leftarrow \{1, \ldots, 2m\}$

Step 1: Elimination
While $(\#I > 5)$ Repeat
Select at random $j \in I$
Evaluate $\Gamma_I(y_j, y, w, z)$
If $(\Gamma_I(y_j, y, w, z) < l)$ Then
$I \leftarrow \{i \in I\colon\ y_i > y_j\} \cup \{j\}$
$w_j \leftarrow \sum_{y_i \leqslant y_j, i \in I} w_i$
$z_j \leftarrow \sum_{y_i \leqslant y_j, i \in I} z_i$
End If
If $(\Gamma_I(y_j, y, w, z) > l)$ Then $I \leftarrow \{i \in I\colon\ y_i \leqslant y_j\}$
End Repeat

Step 2: Solve
Sort $\{y_i\colon\ i \in I\} \to y'$
For each y'_i, y'_{i+1} Solve $\Gamma_I(c, y, w, z) = l$ within $[y'_i, y'_{i+1}]$

Remark 7. At every iteration of the procedure, the cardinality of the set $\{i \in I\colon \Gamma_I(y_i, y, w, z) = l\}$ is at most 3. Indeed, the cardinality of the set $\{y_i\colon\ \Gamma_I(y_i, y, w, z) = l\}$ is not greater than 2, and there is at most one index repeated.

Theorem 8. The algorithm given in table 1 gives all the optimal solutions to $(\mathrm{P}_=)$.

Proof. After step 0 one has $G_l(\cdot) = (1/2l)\Gamma_I(\cdot, y, w, z)$.

In step 1, given y_j, one of the following three cases hold:

1. $\Gamma_I(y_j, y, w, z) < l$. Then all the solutions to the equation $G_l(c) = 1/2$ are greater than y_j. Hence, by lemma 4, in the following step $G_l(c) = (1/2l)\Gamma_I(c, y, w, z)$ for all $c \geqslant y_j$.
2. $\Gamma_I(y_j, y, w, z) > l$. Then all the solutions to the equation $G_l(c) = 1/2$ are smaller than y_j. Hence, by lemma 3, in the following step $G_l(c) = (1/2l)\Gamma_I(c, y, w, z)$ for all $c \leqslant y_j$.
3. $\Gamma_I(y_j, y, w, z) = l$. In this case the algorithm does nothing.

Hence, after step 1, $G_l(y_i) = (1/2l)\Gamma_I(c, y, w, z)$, $\forall i \in I$, and then all the optimal solutions are found in step 2. □

Remark 8. The equation $\Gamma_I(c, y, w, z) = l$ can be solved by linear interpolation.

Now, we are going to study the complexity of the algorithm. This is a probabilistic algorithm so the complexity must be studied in average time. Note that the probabilistic behavior is due to the algorithm itself and not to the input.

Theorem 9. The algorithm in table 1 runs in expected linear time.

Proof. Let $B(l)$ be the worst case time for one execution of one elimination step when $\#I = l$. Firsts of all, note that the evaluation of Γ_I is done in linear time in l. Thus $B(l)$ is bounded by a linear function. Let $T(j)$ be the average time to finish the elimination step. Consider

$$\begin{aligned}\alpha &= \#\big\{i \in I\colon\ \Gamma(y_i, y, w, z) < l\big\},\\ \beta &= \#\big\{i \in I\colon\ \Gamma(y_i, y, w, z) > l\big\},\\ \gamma &= \#\big\{i \in I\colon\ \Gamma(y_i, y, w, z) = l\big\}.\end{aligned}$$

Thus

$$\begin{aligned}p\big(\Gamma(y_i, y, w, z) < l\big) &= \alpha/j,\\ p\big(\Gamma(y_i, y, w, z) > l\big) &= \beta/j,\\ p\big(\Gamma(y_i, y, w, z) = l\big) &= \gamma/j.\end{aligned}$$

If $\Gamma_I(y_j, y, w, z) < l$ then $\#\{i \in I\colon\ y_i \leqslant y_j\} - 1$ is the number of points removed, and this number is uniformly distributed between 0 and $\alpha - 1$. On the other hand, if $\Gamma_I(y_j, y, w, z) > l$ then the number of points removed is $\#\{i \in I\colon\ y_i > y_j\}$, which is uniformly distributed between 0 and $\beta - 1$.

This implies that

$$\begin{aligned}T(j) &= B(l) + \frac{\alpha}{j}\frac{\sum_{i=1}^{\alpha} T(j-i+1)}{\alpha} + \frac{\beta}{j}\frac{\sum_{i=1}^{\beta} T(j-i+1)}{\alpha} + \frac{\gamma}{j}T(j)\\ &= B(l) + \frac{1}{j}\sum_{i=1}^{\alpha} T(j-i+1) + \frac{1}{j}\sum_{i=1}^{\beta} T(j-i+1) + \frac{\gamma}{j}T(j).\end{aligned}$$

Taking into account that $\gamma \leqslant 3$ and $B(l)$ is bounded by a linear function in l, this equation implies that $T(l)$ is average linear time. $\square$

4. Conclusions

We have solved the problem of locating a patrolling unit, whose position will be uniformly distributed the patrolling area (a segment of length at least $2l > 0$) and minimizes its average distance to the demand points, assumed to be distributed on the real line according to an arbitrary random variable A with finite first moment.

The problem is reduced to a one-dimensional convex differentiable optimization program, whose derivative is written in terms of average distances: $\int \|\cdot - a\|\,\mathrm{d}F(a)$.

For the particular case of symmetrical distribution of the demand for requests, an explicit form of the solution is developed for any finite $l > 0$. Also the limit cases of length 0 and infinity are explicitly solved.

For the case in which the distribution of the demand is discrete, we present an algorithm that runs (in average) in linear time. It is worth nothing that this algorithm can also be used to find the median of a set of data clustered in segments.

References

[1] L.R. Anderson and R.A. Fontenot, Optimal positioning of service unit along a coordinate line, Transportation Science 26(4) (1992) 346–351.

[2] O. Berman and R.C. Larson, The median problem with congestion, Computers and Operations Research 9 (1982) 119–126.

[3] P.J. Bickel and K.A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics* (Holden-Day, 1977).

[4] E. Carrizosa, E. Conde, M. Muñoz-Márquez and J. Puerto, The generalized Weber problem with expected distances, RAIRO Recherche Opérationnelle/Operations Research 29 (1995) 35–57.

[5] E. Carrizosa, M. Muñoz-Márquez and J. Puerto, Location and shape of a rectangular facility in $\mathbb{R}^n$: Convexity properties, Mathematical Programming 83 (1998) 277–290.

[6] E. Carrizosa, M. Muñoz-Márquez and J. Puerto, A note on the optimal positioning of service units, Operations Research 46 (1998) 155–156.

[7] S.S. Chiu and R.C. Larson, Locating an n-server facility in a stochastic environment, Computers and Operations Research 12 (1985) 509–516.

[8] R.L. Francis, L.F. McGinnis and J.A. White, *Facility Layout and Location: An Analytical Approach* (Prentice-Hall, Englewood Cliffs, 1992).

[9] F.N. Fritsch, R.E. Schafer and W.P. Crowley, Solution of tracendental equation $w\mathrm{e}^w = x$, Communications of the ACM 16 (1973).

[10] Y. Gerchak and X. Lu, Optimal anticipatory position of a disk arm for queries of random length and location, INFOR 34(4) (1996) 251–262.

[11] J.B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Vols. 1 and 2 (Springer, Berlin, 1993).

[12] R.P. King, Disk arm movement in anticipation of future requests, ACM Transactions on Computer Systems 8 (1990) 214–229.

[13] R. Larson and A. Odoni, *Urban Operations Research* (Prentice-Hall, Englewood Cliffs, NJ, 1981).

[14] R.C. Larson, A hypercube queuing model for facility location and restricting in urban emergency services, Computers and Operations Research 1 (1974) 67–95.

[15] A. Levine, A patrol problem, Mathematical Magazine 59 (1986) 159–166.

[16] X. Lu and Y. Gerchak, Multiple anticipatory moves of a server on a line, Location Science 5(4) (1997) 269–287.

[17] D.E. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. (Addison-Wesley, 1989).

[18] M. Muñoz-Márquez, El problema de Weber regional, Ph.D. thesis, Universidad de Sevilla (1995).

[19] F. Plastria, Continuous location problems, in: *Facility Location. A Survey of Applications and Methods*, ed. Z. Drezner (Springer, 1995) chapter 11.

[20] A.M. Rodríguez Chía, Avances sobre el problema de localización contínua de un único centro, Ph.D. thesis, Universidad de Sevilla (1998).

[21] R.G. Vickson and Y. Gerchak, Optimal positioning of read/write heads in mirrored disks, Location Science 3(2) (1995) 125–132.