Multifacility Ordered Median Problems on Networks: A Further Analysis

Jörg Kalcsics, Stefan Nickel
Fraunhofer Institut für Techno- und Wirtschaftsmathematik, Kaiserslautern, Germany

Justo Puerto
Universidad de Sevilla, Sevilla, Spain

In this paper, we address the ordered *p*-median problem, which includes as special cases most of the classical multifacility location problems discussed in the literature. Finite dominating sets (FDS) are known for particular instances of this problem: *p*-median, *p*-center, and *p*-centdian. We find an FDS for the ordered *p*-median problem. This set allows us to gain a better insight into the general FDS structure of network location problems. This FDS is later used to present the first polynomial time algorithm for *p*-facility ordered median problems on tree networks. This result is combined with some approximation algorithms to give an $O(\log M \log \log M)$ approximate solution of these problems on general networks, where *M* is the number of vertices. © 2003 Wiley Periodicals, Inc.

1. INTRODUCTION

One of the most important and well-developed branches in location theory are location problems on networks. Numerous surveys and textbooks give evidence of this fact. (See Mirchandani and Francis [13], Labbé et al. [12], Drezner [4], Puerto [17], Drezner and Hamacher [5] and references therein.) The starting point of this development may be considered the node-dominance result of Hakimi [7] and the extensions by Hooker et al. [8], which we will show to be essential. In a series of previous papers, a new type of objective function in location theory was introduced and analyzed (see Nickel and Puerto [15], Puerto and Fernández [18], Rodríguez-Chía et al. [20], Francis et al. [6], Nickel [14], and Kalcsics et al. [10]). In this paper, we will analyze the *p*-facility version of this new type of objective function called ordered median function which is a generalization of the most popular objective functions: median, center, centdian, *k*-centrum, among many others.

For networks, this new objective function was first defined in Nickel and Puerto [15] to prove many well-known results in a much easier way and to gain better insight into the geometrical structure of the network with respect to different criteria. In this paper, we discuss the conditions under which finite dominating sets (FDS) for the multifacility formulation of these problems can be derived. Recall that an FDS is a set which always contains an optimal solution of the problem. See Hooker et al. [8] for further details. Identifying a general FDS for this family of problems has important implications on the development of the same kind of efficient algorithms for all these problems simultaneously.

Let $\mathcal{N} = (\mathcal{G}, l)$ denote a network with underlying undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with node set $\mathcal{V} = \{v_1, \ldots, v_M\}$ and edge set $\mathcal{E} = \{e_1, \ldots, e_N\}$. For **an undirected** graph $\mathcal{G}$, an edge $e \in \mathcal{E}$ is defined as $e = [v_i, v_j] = [v_j, v_i]$, $v_i, v_j \in \mathcal{V}$.

Associated to each edge $e \in \mathcal{E}$, there is a positive length $l(e)$ defined by the function $l : \mathcal{E} \to \mathbb{R}_+$. By $d(v_i, v_j)$, we denote the length of a shortest path between v_i and v_j measured by l. Through $w : \mathcal{V} \to \mathbb{R}$, every vertex is assigned a nonnegative weight. For short, we write $w_i := w(v_i)$ for a node $v_i \in \mathcal{V}$.

A point on an undirected edge $e = [v_i, v_j]$ is defined as a pair $x = (e, t)$, $t \in [0, 1]$, with

$$d(v_k, x) = d(x, v_k) := \min\{d(v_k, v_i) + tl(e),$$
$$d(v_k, v_j) + (1 - t)l(e)\},$$

for any $v_k \in \mathcal{V}$. The set of all points of a network $\mathcal{N}$ is denoted by $\mathcal{P}(\mathcal{G})$. It should be noted that this set also

contains the nodes $\mathcal{V}$. The sets $(e, (t_1, t_2)) := \{(e, t) \in \mathcal{P}(\mathcal{G}) : t \in (t_1, t_2), t_1, t_2 \in [0, 1], t_1 < t_2\}$ are called open subedges of $e \in \mathcal{E}$. Analogously, the sets $(e, [t_1, t_2]) := \{(e, t) \in \mathcal{P}(\mathcal{G}) : t \in [t_1, t_2], t_1, t_2 \in [0, 1], t_1 \leq t_2\}$ are closed subedges of $e \in \mathcal{E}$. Furthermore, $(e, (0, 1))$ is called the interior of edge e.

Let $p \geq 2$ be an integer. Then, for $X_p := \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$, we define the distance from a node $v_i \in \mathcal{V}$ to the set X_p as

$$d(v_i, X_p) = d(X_p, v_i) := \min_{k=1, \ldots, p} d(v_i, x_k).$$

A point $x = (e, t)$ on an edge $e = [v_i, v_j] \in \mathcal{E}$ is called a **bottleneck point** if there exists some node v_k with $w_k \neq 0$, such that

$$d(x, v_k) = d(x, v_i) + d(v_i, v_k) = d(x, v_j) + d(v_j, v_k).$$

Let BN_i denote the set of all bottleneck points of a node $v_i \in \mathcal{V}$ and let $\mathcal{BN} := \cup_{i=1}^{M} BN_i$ be the set of all bottleneck points of the graph.

For all $v_i, v_j \in \mathcal{V}$, $v_i \neq v_j$, $w_i w_j \neq 0$ define

$$EQ'_{ij} := \{x \in \mathcal{P}(\mathcal{G}) : w_i d(v_i, x) = w_j d(v_j, x)\}.$$

Let EQ_{ij} be the relative boundary of EQ'_{ij}, that is, the set of endpoints of the closed subedges forming the elements in EQ'_{ij}, and let $\mathcal{EQ} := \cup_{i,j,i \neq j} EQ_{ij}$. The points in $\mathcal{EQ}$ are called **Equilibrium points** of $\mathcal{N}$. To simplify the presentation, we will denote by $EQ_{ij}^{kl} \subseteq EQ_{ij}$ the equilibrium points of nodes v_i, v_j on edge $[v_k, v_l]$, for any $i, j \in \{1, \ldots, M\}$ and $k, l : [v_k, v_l] \in \mathcal{E}$. Note that $|EQ_{ij}^{kl}| \leq 2$. In the case that $|EQ_{ij}^{kl}| = 1$, we denote for the sake of simplicity by EQ_{ij}^{kl} also the only element of the set EQ_{ij}^{kl}.

For $X_p \subseteq \mathcal{P}(\mathcal{G})$, we define

$$d(X_p) := (w_1 d(v_1, X_p), \ldots, w_M d(v_M, X_p))$$

and

$$d_{\leq}(X_p) := (w_{(1)} d(v_{(1)}, X_p), \ldots, w_{(M)} d(v_{(M)}, X_p)),$$

with $(\cdot)$ being a permutation of the set $\{1, \ldots, M\}$ satisfying

$$w_{(1)} d(v_{(1)}, X_p) \leq w_{(2)} d(v_{(2)}, X_p) \leq \cdots \leq w_{(M)} d(v_{(M)}, X_p).$$

To simplify the notation, we will denote the entries $w_i d(v_i, X_p)$ and $w_{(i)} d(v_{(i)}, X_p)$ in the above vectors by $d_i(X_p)$ and $d_{(i)}(X_p)$, respectively.

The **ordered p-median problem** on $\mathcal{N}$ is defined as

$$\mathrm{OM}(\Lambda) := \min_{X_p \subset \mathcal{P}(\mathcal{G})} \mathrm{OM}_{\Lambda}(X_p),$$
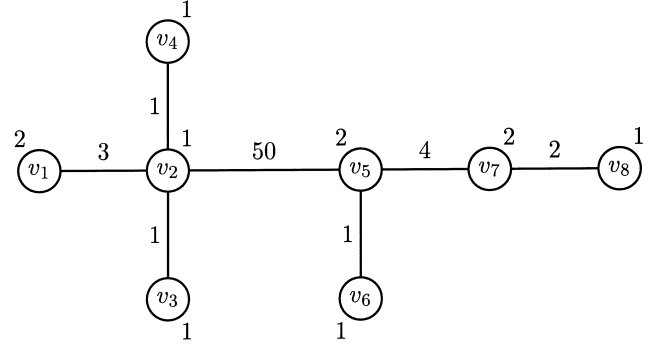

FIG. 1.   Tree network of Example 1.

with

$$\mathrm{OM}_{\Lambda}(X_p) := \langle \Lambda, d_{\leq}(X_p) \rangle = \sum_{i=1}^{M} \lambda_i d_{(i)}(X_p)$$

$$\text{and} \quad \Lambda = (\lambda_1, \ldots, \lambda_M) \in \mathbb{R}_{0+}^{M}. \quad (1)$$

This problem is the natural multifacility extension of the ordered median problem considered in Nickel and Puerto [15]. The function $\mathrm{OM}_{\Lambda}(X_p)$ is called the **ordered p-median function.** Note that this function is defined pointwise. In the context of continuous location theory, a similar objective function was introduced in Puerto and Fernández [18] and later studied in Francis et al. [6], Rodríguez-Chía et al. [20], and Puerto and Fernández [19].

The rest of the paper is organized as follows: In the next section, we study the ordered p-median problem with a special structure in the Λ-modeling weights. Under this assumption, we can identify an FDS for the problem, namely, the set of pseudo-equilibrium points (r-extremes). Section 3 uses the characterization of the FDS to present the first polynomial time algorithm for solving these problems on tree networks. This polynomial time algorithm for trees is combined with the general approximation algorithms of Bartal [1] and Charikar et al. [2] to obtain an $O(\log M \log \log M)$ approximate solution for the p-facility ordered median problem in general networks. The paper ends with some conclusions.

2. THE ORDERED p-MEDIAN PROBLEM

In Nickel and Puerto [15], it was proved that for $\lambda_1 \geq \cdots \geq \lambda_M$ the node set $\mathcal{V}$ constitutes an FDS for the ordered p-median problem and that for arbitrary $\Lambda \geq 0$, $\mathcal{V} \cup \mathcal{EQ}$ is an FDS for the single-facility ordered median problem. We demonstrate by a simple counterexample that this latter dominance result for the single-facility case does not hold for the p-facility case.

Example 1.   *Consider the tree network of Figure 1. Observe that $\mathcal{V} \cup \mathcal{EQ}$ is not an FDS for the ordered 2-median*   F1
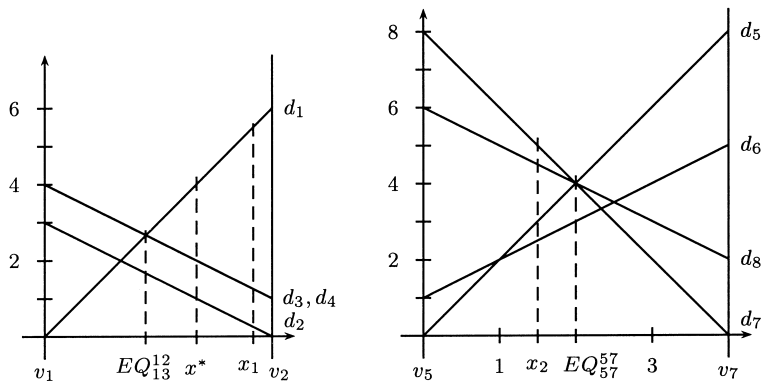
FIG. 2.   The distance functions along $[v_1, v_2]$ and $[v_5, v_7]$ of Example 1.

problem with $\Lambda = (0.2, 0.2, \ldots, 0.2, 1)$. *If we restrict* X_2 *to be in* $\mathcal{V} \cup \mathcal{EQ}$, *the optimal solution is given by*

$$X_2 = \left\{ EQ_{13}^{12} = \left([v_1, v_2], \frac{4}{9}\right), EQ_{57}^{57} = \left([v_5, v_7], \frac{1}{2}\right)\right\},$$

with objective value $\text{OM}_\Lambda(X_2) = 8\frac{2}{15} = 8.1\bar{3}$. *If we drop this restriction, we obtain a better solution, namely,*

$$X_2^* = \left\{ x^* = \left([v_1, v_2], \frac{2}{3}\right), EQ_{57}^{57} = \left([v_5, v_7], \frac{1}{2}\right)\right\},$$

with an optimal objective function value of 8.0 *(see also Fig.* **F2** *2). Note that* x^* *is neither an equilibrium point nor a vertex.*

Despite this negative result, we are able to characterize a polynomial-size FDS for an important class of ordered *p*-median problems. We note in passing that identifying an FDS is of great importance because it allows one to develop algorithms for these problems even if the solutions are not required to be at vertices. Let $1 \le k < M$ and $\Lambda^k = (a, \ldots, a, b \ldots, b) \in \mathbb{R}_{0+}^M$, where

$$a = \lambda_1 = \cdots = \lambda_k \ne \lambda_{k+1} = \cdots = \lambda_M = b.$$

Note that the Λ vectors corresponding to the center, centdian, or *k*-centrum problem are of this type. See Nickel and Puerto [15]. For $a \ge b$, the results in [15] ensure that $\mathcal{V}$ is an FDS for the problem. Thus, only the case with $a < b$ has to be further investigated.

Example 1 already points out the two main characteristics of a potential FDS for the case $a < b$. First, one of the solution points belongs to the set $\mathcal{V} \cup \mathcal{EQ} : EQ_{57}^{57}$. Second, the other solution point, x^*, is related to the one in $\mathcal{V} \cup \mathcal{EQ}$, namely, there exist two nodes v_1 and v_5 allocated to x^* and EQ_{57}^{57}, respectively, such that $d_1(x^*) = d_5(EQ_{57}^{57})$. In general, we will show in the following that there exists an optimal solution X_p^* such that

- One or more solution points belong to the set $\mathcal{V} \cup \mathcal{EQ}$, i.e. $X_p^* \cap (\mathcal{V} \cup \mathcal{EQ}) \ne \varnothing$, and

- For every point in $X_p^* \backslash (\mathcal{V} \cup \mathcal{EQ})$, there exists another solution point in $X_p^* \cap (\mathcal{V} \cup \mathcal{EQ})$ and two nodes allocated to each of the two points such that the weighted distances of these two nodes to their respective solution points are equal.

The important point is not just to prove the existence of a finite-size FDS for the problem but to identify it. This dominating set can then be used for developing algorithms to solve the problem while the existence itself is of limited use. For the sake of readability, the following proof of the FDS is split into a sequence of four results. Moreover, we first give an informal description of the main arguments of the proofs:

Denote by $\mathcal{M} := \{1, \ldots, M\}$ and let $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$. We define the following sets:

$$I_l := \{i \in \mathcal{M} | d(X_p, v_i) = d(x_l, v_i)\} \backslash \bigcup_{j=0}^{l-1} I_j, \quad l = 1, \ldots, p,$$

of the indices of nodes which are allocated to x_l, $l = 1, \ldots, p$, where $I_0 := \varnothing$. Note that ties are resolved by allocating nodes to solution points with smallest indices. The objective function $\text{OM}_\Lambda(X_p)$ can now be rewritten with respect to I_l as

$$\text{OM}_\Lambda(X_p) = \sum_{i=1}^{M} \lambda_i w_{(i)} d(X_p, v_{(i)})$$

$$= \underbrace{\sum_{i \in I_1} \lambda_i w_{(i)} d(x_1, v_{(i)})}_{=: f_1(x_1)} + \cdots + \underbrace{\sum_{i \in I_p} \lambda_i w_{(i)} d(x_p, v_{(i)})}_{=: f_p(x_p)}.$$

For a fixed permutation $(\cdot)$ and fixed allocations I_l, the functions f_l, $l \in \{1, \ldots, p\}$, are, as a sum of concave functions, also concave in x_l.

Now, consider the tree network given in Figure 1. Let X_2

$= \{x_1, x_2\}$ with $x_1 = ([v_1, v_2], \frac{11}{12})$ and $x_2 = ([v_5, v_7], \frac{3}{8})$ (see Fig. 2). The vector of ordered distance functions for this set of points is $d_{\leq}(X_2) = (0.25, 1.25, 1.25, 2.5, 3, 4.5, 5, 5.5)$. Using $\Lambda = (0.2, 0.2, \ldots, 0.2, 1)$, we get $OM_{\Lambda}(X_2) = 9.05$.

Starting from X_2, we try to obtain a better solution. From Figure 2, observe that fixing x_2 and moving x_1 on its edge a little to the left or right will neither change the order of the distance functions in the vector $d_{\leq}(X_2)$ nor the allocation of nodes to the solution points. Hence, the permutation $(\cdot)$ and the index sets I_1 and I_2 remain the same and $OM_{\Lambda}(\cdot)$ *is concave with respect to x_1*. As a result, we can find a descent direction and obtain a better solution.

The formal argument is as follows: Define for $t \in \mathbb{R}$: $x_1(t) := ([v_1, v_2], \frac{11}{12} + t)$ and $X_2(t) := \{x_1(t), x_2\}$. Let $d_i(t) := w_i d(X_2(t), v_i)$, $i \in \mathcal{M}$. The vector of distance functions with respect to t is $d(X_2(t)) = (d_i(t))_{i=1, \ldots, M} = (5.5 + 6t, 0.25 - 3t, 1.25 - 3t, 1.25 - 3t, 3, 2.5, 5, 4.5)$. For $t = 0$, we obtain $d_{\leq}(X_2(t)) = (0.25 - 3t, 1.25 - 3t, 1.25 - 3t, 2.5, 3, 4.5, 5, 5.5 + 6t)$. Now observe that the order of the distance functions does not change for $-\frac{1}{12} \leq t \leq \frac{1}{12}$. This means that we can move the point $x_1 \to x_1(t)$ by a small amount on its edge (to the left) without disturbing the permutation. Therefore, we can write the objective function as $OM_{\Lambda}(X_2) := 9.05 + 4.2t$, which is a concave function for $t \in [-\frac{1}{12}, \frac{1}{12}]$. Moreover, any value of t, $-\frac{1}{12} \leq t < 0$, will yield a lower objective value.

In the above example, the order of the distance functions did not change at all for $t \in [-\frac{1}{12}, \frac{1}{12}]$. But, obviously, even a change in the ordering of only the first $k - 1$ or last $k + 1$ vertices is not going to be a problem and we can still argue that the objective function value will not increase. The following lemma addresses the circumstances under which we can move a point while not increasing the objective function value and how far we can move it.

Lemma 1. *Let $\mathcal{N} = (\mathcal{G}, l)$ be an undirected network with nonnegative node weights, $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$, $\tilde{x} = (e, t) \in X_p$, with $e \in \mathcal{E}$ and $t \in [0, 1]$ an arbitrary solution point and $\Lambda^k = (a, \ldots, a, b, \ldots, b) \in \mathbb{R}_{0+}^M$. Then, there exists a point $x' = (e, t')$, $t' \in [0, 1]$, such that $OM_{\Lambda}(X_p')$ $\leq OM_{\Lambda}(X_p)$, where $X_p' := X_p \setminus \{\tilde{x}\} \cup \{x'\}$, and either*

$$x' \in \mathcal{V} \quad \text{**or**} \quad d_{(k)}(X_p') = d_{(k+1)}(X_p') \quad (2)$$

holds.

Proof. Let $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$ with $x_l = (e_l, s_l)$, $l = 1, \ldots, p$, such that X_p does not satisfy one of the relations in (2). W.l.o.g., let $x_1 = \tilde{x} \in \mathcal{P}(\mathcal{G})$.

Furthermore, we assume that

  (i) $d_{(n)}(X_p) \neq d_{(n+1)}(X_p)$ for all $n \in \{1, \ldots, M - 1\}$ and zz
  (ii) $\not\exists \ v_i \in \mathcal{V} : w_i d(v_i, x_1) = w_i d(v_i, x_l)$, that is, none of the nodes is at the same distance from x_1 as to another solution point $x_l \neq x_1$.

Define for $t : X_p(t) := \{x_1(t), x_2, \ldots, x_p\}$, where $x_1(t) := (e_1, s_1 + t)$. Let $T := [\underline{t}, \overline{t}]$ be an interval with $-s_1 \leq \underline{t} < 0 < \overline{t} \leq 1 - s_1$, such that (i) and (ii) hold for $X_p(t)$ for all $t \in T$. This interval exists since (i) and (ii) are satisfied for $t = 0$ and all distance functions $d_i(\cdot)$ are continuous on an edge.

Let $v_i := v_{(n)}$, for some $n = 1, \ldots, M$, be allocated to x_1, that is, $d_{(n)}(X_p) = d_i(x_1)$. Then, by the above assumptions on X_p and the definition of T, we have $d_{(n)}(X_p(t)) = d_i(x_1(t))$, $\forall t \in T$. For all nodes v_j allocated to $x_l \in X_p$, $x_l \neq x_1$, $d_{(r)}(X_p(t)) = d_j(x_l)$, for some $r \neq n$, is constant. Therefore, $d_{(\cdot)}(X_p(t))$ is either concave or constant with respect to $t \in T$, since $d_i(x_1(t))$ is concave on e_1.

In summary, we have that $d_i(X_p(t)) = d_{(n)}(X_p(t))$ is a concave function for $t \in T$, $n = 1, \ldots, M$. Moreover, since the inequality $d_{(n)}(X_p(t)) < d_{(n+1)}(X_p(t))$, $n = 1, \ldots, M - 1$, holds for all $t \in T$, it follows that the order of the distance functions does not change. As a result, $OM_{\Lambda}(X_p(t))$ is also concave in the interval T. Assume w.l.o.g. that the objective function is nonincreasing for $t < 0$. Hence, we may decrease $\underline{t}$ until either $x_1(t) \in \mathcal{V}$ or $d_{(k)}(X_p(t)) = d_{(k+1)}(X_p(t))$.

Now, we prove that the two assumptions made on X_p do not imply any loss of generality:

  (i) Let $n \in \{1, 2, \ldots, M - 1\}$ such that $d_i(X_p) = d_{(n)}(X_p) = d_{(n+1)}(X_p) = d_j(X_p)$, where $i := (n)$ and $j := (n + 1)$. Note that $n \neq k$, since, otherwise, X_p would satisfy (2). Hence, the elements $d_i(X_p)$ and $d_j(X_p)$ possibly swap positions in the vector $d_{\leq}(X_p(t))$, that is, $d_i(X_p(t)) \leq d_j(X_p(t))$ for $t \leq 0$ and $d_i(X_p(t)) > d_j(X_p(t))$ for $t > 0$. However, both functions $d_i(X_p)$ and $d_j(X_p)$ are still concave with respect to $t \in T$ and both are still multiplied by the same Λ value, a or b (since $n \neq k$). Therefore, this change has no influence on the concavity and the slope of the objective function $OM_{\Lambda}(X_p(t))$.

  (ii) Concerning the reallocation of nodes, let $v_i \in \mathcal{V}$ be a node such that $w_i d(v_i, x_1) = w_i d(v_i, x_l)$ holds for another solution point $x_l \neq x_1$ and x_1 is not the bottleneck point of this node on edge e_1. (Otherwise, the allocation will not change with respect to $x_1(t)$, $t \in T$.) One of the following two cases can occur:

1. v_i is allocated to $x_1(t)$ for $t \leq 0$ and to x_l, $l = 2, \ldots, p$, for $t > 0$. Thus, we have $d_i(X_p(t)) = d_i(x_1(t))$ for $t \leq 0$ on edge e_1 and $d_i(X_p(t)) = d_i(x_l)$, $t > 0$, on edge e_l. To be reallocated, the distance function of v_i on edge e_1, $d_i(x_1(t))$ has to be increasing for $t \leq 0$. After the change of allocations, we obtain $d_i(X_p(t)) = d_i(x_l)$ on e_l, which is constant with respect to t. (See the two leftmost edges of Fig. 3.) Thus, $d_i(X_p(t))$ is concave for $t \in T$.   **F3**

2. v_i is allocated to x_l for $t \leq 0$ and to $x_1(t)$ for $t > 0$. As above, $d_i(X_p(t))$ is concave for $t \in T$ since it is the minimum of a linear and a constant function (see the two edges on the right-hand side of Fig. 3).   ∎
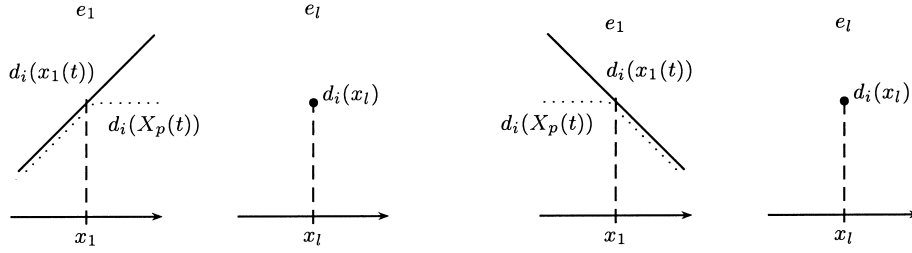
Note that the above result does not hold if one or more

FIG. 3. v_i changes its allocation from $x_1(t)$ to x_l, respectively, from x_l to $x_1(t)$.

node weights are negative. In this case, $d_i(X_p(t))$ may be convex with respect to t and, hence, the ordered median function may no longer be concave in the interval T.

From Lemma 1, it follows that we can move an arbitrary solution point on its edge either to the left or to the right without increasing the objective function value until a point is attained for which (2) holds. This is illustrated in the following example continuing the discussion preceding Lemma 1.

Example 2. *Consider the network of Example* 1 *and the initial solution* $X_2 = \{x_1 = ([v_1, v_2], \frac{11}{12}), x_2 = ([v_5, v_7], \frac{3}{8})\}$. *For* $\Lambda^7 = (0.2, 0.2, \ldots, 0.2, 1)$, *it is possible to improve the objective function value by moving the point* x_1 *to the left up to the point* $([v_1, v_2], \frac{10}{12})$. *This yields* $d_{(7)}(X_2(-\frac{1}{12})) = d_7(x_2)$ $= 5 = d_1(x_1(-\frac{1}{12})) = d_{(8)}(X_2(-\frac{1}{12}))$. *Hence,* (2) *holds.*

In the previous lemma, we could choose a single solution point and move it along its edge until condition (2) was fulfilled. Obviously, we can repeat this procedure for all points in X_p. This leads to the following corollary:

Corollary 2. *Let* $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$ *be a solution to the ordered p-median problem with nonnegative node weights,* $p \geq 2$ *and* $\Lambda^k \in \mathbb{R}_{0+}^M$, $1 \leq k \leq M - 1$. *Then, there exists a solution* X'_p *with* $OM_\Lambda(X'_p) \leq OM_\Lambda(X_p)$ *such that either* $X'_p \subseteq \mathcal{V}$ *or* $d_{(k)}(X'_p) = d_{(k+1)}(X'_p)$ *holds.*

Proof. Assume that $X_p \nsubseteq \mathcal{V}$ and $d_{(k)}(X_p) < d_{(k+1)}(X_p)$. Then, according to Lemma 1, we start by moving one solution point after the other until we obtain a new solution X'_p where either all solution points are nodes or finally $d_{(k)}(X'_p) = d_{(k+1)}(X'_p)$ holds. ∎

In the following, we will deal with the difficulties which occur when the kth and $(k + 1)^{st}$ vertices in the ordered vector of the distance functions have the same value. Resolving these difficulties will lead us to the proof of the FDS.

We first give a formalized description of the set of solution points addressed in the second part of the characterization of our FDS introduced at the beginning of this section. For every point in $X_p^* \backslash (\mathcal{V} \cup \mathcal{EQ})$, there exists another solution point in $X_p^* \cap (\mathcal{V} \cup \mathcal{EQ})$ and two nodes allocated to each of the two points such that the weighted

distances of these nodes to their respective solution points are equal. Define the set of ranges (canonical set of distances) $\mathcal{R} \subseteq \mathbb{R}_+$ by

$$\mathcal{R} := \{r \in \mathbb{R}_+ | \exists\, EQ \in EQ_{ij} : d_i(EQ) = r = d_j(EQ)$$
$$\text{or} \quad \exists\, v_i, v_j \in \mathcal{V}, v_i \neq v_j : r = w_i d(v_j, v_i)\}.$$

Ranges correspond to function values of equilibria or to node-to-node distances. In terms of the general characterization of the FDS, let R' be the set of ranges of the points in $X_p^* \cap (\mathcal{V} \cup \mathcal{EQ})$. Then, for every other solution point x not in this set there exists a node v_i allocated to x and a range $r \in R'$ such that $w_i d(x, v_i) = r$. This is now generalized as follows:

Definition 3. *Let* $\mathcal{N} = (\mathcal{G}, l)$ *be an undirected network with nonnegative node weights and the set* $\mathcal{R}$ *as defined above. A point* $x = (e, t)$ *is called an r-extreme point or pseudoequilibrium with range* $r \in \mathcal{R}$ *if there exists a node* $v_i \in \mathcal{V}$ *with* $r = w_i d(x, v_i)$.

Let us denote by $\mathcal{PEQ}$ the set of all pseudoequilibria with respect to all ranges $r \in \mathcal{R}$. Note that $\mathcal{V} \subseteq \mathcal{PEQ}$, which follows directly from the above definition, and also that $\mathcal{EQ} \subseteq \mathcal{PEQ}$, since every equilibrium $EQ \in EQ_{ij}$ of two nodes v_i and v_j is an r-extreme point with $r = d_i(EQ) = d_j(EQ)$. The above definition generalizes a concept introduced in Pérez-Brito et al. [16].

Example 3. *The set of ranges on the edge* $[v_1, v_2]$ *of the network given in Example* 1 *is* $\{0, 1, 2, 2.6, 3, 4, 6\}$ *(see Fig. 2). The point* $x^* = ([v_1, v_2], \frac{2}{3})$ *is a pseudoequilibrium of range* 4.

Our goal is to prove that $\mathcal{PEQ}$ is an FDS for the ordered p-median problem with $p \geq 2$ and $\Lambda^k \in \mathbb{R}_{0+}^M$, $1 \leq k \leq M - 1$. In addition, any optimal solution X_p^* must satisfy $X_p^* \cap (\mathcal{V} \cup \mathcal{EQ}) \neq \varnothing$. The first result proves the existence while the latter allows us to identify an FDS for any given problem.

Using Lemma 1, we could move an arbitrary solution point along its edge to the left or to the right until we have $d_{(k)}(X_p) = d_{(k+1)}(X_p)$. The goal is to find a method to continue this process without increasing the objective func-

tion value. If we are in this situation of equality, the idea is to move more than one solution point simultaneously, preserving the relationship $d_{(k)}(X_p) = d_{(k+1)}(X_p)$ and the permutation of the distance functions at the positions k and $k + 1$. In Example 2, we have $d_{(7)}(X_2(-\frac{1}{12})) = 5 = d_{(8)}(X_2(-\frac{1}{12}))$. Here, we can continue moving x_1 to the left if we simultaneously move x_2 to the right in such a way that $d_{(7)}(X_2(\cdot)) = d_{(8)}(X_2(\cdot))$ is preserved.

Before formalizing this approach, we introduce additional notation. Observe that we stop moving a point if, for two vertices $v_{(k)}$ and $v_{(k+1)}$, we had $r_k := d_{(k)}(X_p) = d_{(k+1)}(X_p)$. But, obviously, there may be more than two nodes allocated to solution points whose weighted distance to their respective points is r_k. Therefore, let $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$ be a solution to the ordered p-median problem with $X_p \not\subseteq \mathcal{V}$ and $d_{(k)}(X_p) = d_{(k+1)}(X_p) = r_k$, $r_k \in \mathcal{R}$. Define $\underline{n}$, $\bar{n}$ as the two indices with $1 \le \underline{n} \le k < k + 1 \le \bar{n} \le M$ such that

$$d_{(\underline{n}-1)}(X_p) < d_{(\underline{n})}(X_p) = \cdots = r_k = \cdots = d_{(\bar{n})}(X_p)$$
$$< d_{(\bar{n}+1)}(X_p),$$

where $d_{(0)}(X_p) := -\infty$ and $d_{(M+1)}(X_p) := +\infty$. Note that $\bar{n} - \underline{n} \ge 1$, that is, $\underline{n} < \bar{n}$, by the assumption on X_p.

Furthermore, define $X_L \subseteq X_p$ as the (sub)set of points of X_p such that for every $x_l \in X_L$ there exists a node $v_i \in \mathcal{V}$

$$\mathcal{R}_X := \begin{cases} \{r \in \mathbb{R} \mid \exists \, v_i \in \mathcal{V}, \, \exists \, v_j \in \mathcal{V}_X, \, v_i \neq v_j : r = w_i d(v_i, v_j)\} & \text{if } \mathcal{V}_X \neq \varnothing \\ \varnothing & \text{otherwise.} \end{cases}$$

Now for X_p let L, X_L, $\underline{n}$, and $\bar{n}$ be defined as above and assume that $\bar{n} - \underline{n} + 1 > L$.

Then, there exist $x_l \in X_L$ and $v_i, v_j \in \mathcal{V}$, $v_i \neq v_j$, such that $d_{(n_1)}(X_p) = d_i(x_l) = r_k = d_j(x_l) = d_{(n_2)}(X_p)$, where $(n_1) = i$, $(n_2) = j$, and v_i and v_j are both allocated to x_l (with respect to X_p). Thus, $x_l \in EQ_{ij}$ is an equilibrium of the two nodes v_i and v_j with range r_k. As a result, $r_k \in \mathcal{R}$ and all the points of the set X_L are r_k-extreme points. Let $x_m \in X_p \backslash (X_L \cap \mathcal{V}_X)$. Using the same arguments as in Lemma 1, we can fix all other solution points and just move $x_m \rightarrow x_m(t)$ on its edge until $x_m(t)$ is a node or $d_i(x_m(t)) =: r \in (\{r_k\} \mathcal{W} \mathcal{R}_X) \subseteq \mathcal{R}$ for some node v_i allocated to $x_m(t)$. In this case, $x_m(t)$ is a pseudoequilibrium with range r. This procedure can be applied to all solution points not belonging to $X_L \cup \mathcal{V}$.

It is also obvious that for those solution points $x_l \in \mathcal{V}_X \cap X_L$ the above procedure can be applied. Therefore, if $\bar{n} - \underline{n} + 1 > L$, we have $X_p \subseteq \mathcal{P}\mathcal{E}\mathcal{Q}$ and $X_p \cap (\mathcal{V} \cup \mathcal{E}\mathcal{Q}) \neq \varnothing$ and the desired result follows. ∎

Lemma 4 characterizes an FDS for the ordered p-median problem with $\Lambda^k = (a, \ldots, a, b \ldots, b)$ except for $\bar{n} - \underline{n} + 1 = L$. Dealing with this case will finally complete the identification of the FDS. Here, we really have to move

allocated to x_l with $r_k = d_i(x_l) = d_{(n)}(X_p)$ and $(n) = i$. W.l.o.g., we assume that $X_L := \{x_1, \ldots, x_L\}$ consists of the first L solution points of X_p, $1 \le L \le p$. Note that, by assumption, $L \ge 2$. Moreover, $\mathcal{M}_L := \{1, \ldots, L\}$.

We first state a lemma for the special case $\bar{n} - \underline{n} + 1 > L$, which means that there exists a solution point x which has at least two nodes, v_i and v_j, allocated to it with weighted distance r_k. In this case, $x \in X_p$ is an equilibrium of these two nodes, yielding $x \in EQ_{ij}$. Therefore, it is possible to prove the FDS using the results of Lemma 1.

Lemma 4. *Let* $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$ *be a solution to the ordered p-median problem with nonnegative node weights, $p \ge 2$ and $\Lambda^k \in \mathbb{R}_{0+}^M$, $1 \le k \le M - 1$. Then, there exists a solution X_p' with $OM_\Lambda(X_p') \le OM_\Lambda(X_p)$ such that either $\bar{n} - \underline{n} + 1 = L$ holds for the new solution or $X_p' \subseteq \mathcal{P}\mathcal{E}\mathcal{Q}$ with $X_p' \cap (\mathcal{V} \cup \mathcal{E}\mathcal{Q}) \neq \varnothing$.*

Proof. Let $\bar{X}_p$, $p \ge 2$, be a solution. Given $\bar{X}_p$, we know from Corollary 2 that there exists another solution $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$, $x_l = (e_l, s_l)$, $l = 1, \ldots, p$, with $OM_\Lambda(X_p) \le OM_\Lambda(\bar{X}_p)$ such that $X_p \subseteq \mathcal{V}$ or $d_{(k)}(X_p) = d_{(k+1)}(X_p) =: r_k$. Note that if the former case holds the desired result follows.

Let $\mathcal{V}_X := X_p \cap \mathcal{V}$ and define $\mathcal{R}_X$ as the set of ranges of the nodes in $\mathcal{V}_X$, that is:

solution points simultaneously in order to find a nonascent direction for the objective function.

Theorem 5. *The ordered p-median problem with nonnegative node weights, $p \ge 2$ and $\Lambda^k \in \mathbb{R}_{0+}^M$, $1 \le k \le M - 1$, always has an optimal solution $X_p^* \subseteq \mathcal{P}(\mathcal{G})$ in the set $\mathcal{P}\mathcal{E}\mathcal{Q}$. Moreover, $X_p^* \cap (\mathcal{V} \cup \mathcal{E}\mathcal{Q}) \neq \varnothing$.*

Proof. Let $\bar{X}_p$, $p \ge 2$, be an optimal solution. We know from Lemma 4 that there exists another optimal solution $X_p = \{x_1, \ldots, x_p\} \subseteq \mathcal{P}(\mathcal{G})$, $x_l = (e_l, s_l)$, $l = 1, \ldots, p$, with $OM_\Lambda(X_p) = OM_\Lambda(\bar{X}_p)$ such that either $\bar{n} - \underline{n} + 1 = L$ holds for the new solution or $X_p \subseteq \mathcal{P}\mathcal{E}\mathcal{Q}$ with $X_p \cap (\mathcal{V} \cup \mathcal{E}\mathcal{Q}) \neq \varnothing$. Note that if the latter case holds the desired result follows. Consider for X_p the elements $\mathcal{V}_X$, $\mathcal{R}_X$, L, X_L, $\underline{n}$, and $\bar{n}$ as defined above. Now, we analyze the case $\bar{n} - \underline{n} + 1 = L$. Observe that $\mathcal{V}_X \cap X_L = \varnothing$ (see the proof of Lemma 4). Thus, for every $x_l \in X_L$ there exists a unique $v_{i_l} \in \mathcal{V}$ allocated to x_l with $d_{i_l}(x_l) = r_k$. First, we assume that

(i) $d_{(n)}(X_p) \neq d_{(n+1)}(X_p)$ for all $n \in \{1, \ldots, M\} \backslash (\{M\} \cup \{\underline{n}, \ldots, \bar{n}\})$,

(ii) None of the solution points $x_l \in X_L$ is a bottleneck point of some node $v_i \in \mathcal{V}$, and

(iii) $\nexists\ v_i \in \mathcal{V} : w_i d(v_i, x_{l_1}) = w_i d(v_i, x_{l_2})$, that is, no node is at the same distance from two solution points $x_{l_1}, x_{l_2} \in X_p$.

Define for $t \in \mathbb{R} : X_p(t) := \{ x_1(t/\nabla_1), \ldots, x_L(t/\nabla_L), x_{L+1}, \ldots, x_p \}$, where $\nabla_l := \pm w_{i_l} l(e_l)$ is the slope of the distance function $d_{i_l}(\cdot)$ of node v_{i_l} at the point x_l on edge e_l, and $x_l(t/\nabla_l) := (e_l, s_l + (t/\nabla_l))$, $l = 1, \ldots, L$. Note that the distance functions $d_{i_l}(\cdot)$ are all linear in a sufficiently small interval around the points x_l. Otherwise, x_l would be a bottleneck point of node v_{i_l}.

Let $T := [\underline{t}, \bar{t}] \in \mathbb{R}$ be an interval with $\underline{t} < 0 < \bar{t}$ and $s_l + (t/\nabla_l) \in (0, 1)$, $\forall l \in \mathcal{M}_L$, $t \in T$, such that (i) and (iii) hold for $X_p(t)$ and (ii) for $x_l(t/\nabla_t)$, $1 \leq l \leq L$, for all $t \in T$. This interval exists since (i), (ii), and (iii) hold for $t = 0$ and all distance functions $d_i(\cdot) = w_i d(\cdot, v_i)$ are continuous on any edge.

Let $n \notin \{\underline{n}, \ldots, \bar{n}\}$ and $v_i = v_{(n)}$. By the above assumptions on X_p and the definition of T, we have that $d_{(n)}(X_p(t)) = d_i(x_l(t/\nabla_l))$, for all $t \in T$, if v_i is allocated to $x_l \in X_L$ and $d_{(n)}(X_p(t)) = d_i(x_l)$ if v_i is allocated to a solution point in $X_p \backslash X_L$. In both cases, $d_{(n)}(X_p(t))$ is linear with respect to $t \in T$.

On the other hand, let $n_1, n_2 \in \{\underline{n}, \ldots, \bar{n}\}$, $n_1 \neq n_2$, and let node $v_{(n_1)} =: v_i$ and node $v_{(n_2)} =: v_j$ be allocated to x_{l_1} and x_{l_2}, respectively. By the definition of T, we have that $x_{l_1}, x_{l_2} \in X_L$. Note that $x_{l_1} \neq x_{l_2}$, since $\bar{n} - \underline{n} + 1 = L$. Thus,

$$d_i \left(x_{l_1} \left(\frac{t}{\nabla_{l_1}} \right) \right) = d_i(x_{l_1}) + \mathrm{sgn}(\nabla_{l_1}) w_i \frac{t l(e_{l_1})}{\nabla_{l_1}}$$

$$= d_{(n_1)}(X_p) + t = r_k + t = d_{(n_2)}(X_p) + t$$

$$= d_j(x_{l_2}) + \mathrm{sgn}(\nabla_{l_2}) w_j \frac{t l(e_{l_2})}{\nabla_{l_2}}$$

$$= d_j \left(x_{l_2} \left(\frac{t}{\nabla_{l_2}} \right) \right) \quad \text{for all } t \in T,$$

where $\mathrm{sgn}(x)$ is the sign function of x. Hence, $d_i(x_{l_1}(t/\nabla_{l_1})) = d_{(n_1)}(X_p(t)) = d_{(n_2)}(X_p(t)) = d_j(x_{l_2}(t/\nabla_{l_2}))$, $\forall t \in T$, and we are increasing or decreasing r_k by $|t|$. This means that we are simultaneously moving each solution point $x_l \in X_L$ on e_l by $x_l \to x_l(t/\nabla_t)$, while preserving the relationship $d_{(\underline{n})}(X_p(t)) = \cdots = d_{(\bar{n})}(X_p(t))$. See Figure 4 for an example with $\underline{n} = k$ and $\bar{n} = k + 1$. As a result, all entries $d_{(n)}(X_p(t))$ are linear with respect to $t \in T$. This fact together with the assumption that $d_{(n)}(X_p(t)) \neq d_{(n+1)}(X_p(t))$, $\forall t \in T$, $n \in \{1, \ldots, M\} \backslash (\{M\} \cup \{\underline{n}, \ldots, \bar{n}\})$, implies that the objective function $\mathrm{OM}_\Lambda(X_p(t))$ is also linear with respect to $t \in T$ and, hence, constant over T, since $X_p = X_p(0)$ is already optimal. Consequently, any $X_p(t)$ with $t \in T$ is also optimal.

In summary, the objective function $\mathrm{OM}_\Lambda(X_p(t))$ is constant over the interval T and we can either decrease $\underline{t}$ or increase $\bar{t}$ by an arbitrarily small value without changing the
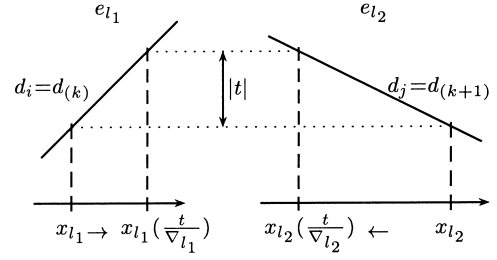

FIG. 4. Simultaneous movement of x_{l_1} and x_{l_2} on e_{l_1} and e_{l_2}.

objective function value $\mathrm{OM}_\Lambda(X_p(t))$. Assume w.l.o.g. that we increase $\bar{t}$. Then one of the following two cases can occur:

$r_k + \bar{t} \in \mathcal{R}_X$: Then, either $x_l(\bar{t}/\nabla_l) = v_i \in \mathcal{V}$ is a node for some $l \in \mathcal{M}_L$ or $x_l(\bar{t}/\nabla_l)$ is an equilibrium EQ_{ij} of two nodes v_i and v_j which are both allocated to $x_l(\bar{t}/\nabla_l)$ such that $d_{(n_1)}(X_p(\bar{t})) = d_i(x_l(\bar{t}/\nabla_l)) = r_k + \bar{t} = d_j(x_l(\bar{t}/\nabla_l)) = d_{(n_2)}(X_p(\bar{t}))$, where $v_i = v_{(n_1)}$ and $v_j = v_{(n_2)}$. Then, all the remaining points $x_l(\bar{t}/\nabla_l)$, $l \in \mathcal{M}_L$, are pseudoequilibria with range $r_k + \bar{t}$. In the latter case, we extend $\mathcal{R}_X$ by the ranges of v_i. Now, we can, again, as already described above, move the remaining solution points $x_m \in X_p \backslash X_L$ independently from each other until we obtain a new optimal solution $X_p^* \in \mathcal{PEQ}$ such that all solution points are pseudoequilibria with respect to a range of one of the points in $X_p^* \cap (\mathcal{V} \cup \mathcal{EQ})$.

$r_k + \bar{t} \notin \mathcal{R}_X$: In this case, a solution point $x_l \in X_p \backslash X_L$ must exist together with a node v_{i_l} allocated to x_l [with respect to $X_p(\bar{t})$] such that $d_{(n)}(X_p(\bar{t})) = d_{i_l}(x_l) = r_k + \bar{t}$, where $v_{i_l} = v_{(n)}$. We redefine $X_L := X_L \cup \{x_l\}$ and also $X_p(\bar{t}) := \{ x_1(t/\nabla_1), \ldots, x_{L+1}(t/\nabla_{L+1}), x_{L+2}, \ldots, x_p \}$, where w.l.o.g. $l = L + 1$. Then, we can apply the same argument as above in order to move the $L + 1$ solution points simultaneously.

Now, we show that the assumptions (i)–(iii) previously made for X_p do not imply any loss of generality:

(i) As in Lemma 1, a possible swap of the elements in the vector $d_\leq(X_p)$ has no influence on the slope of the objective function $\mathrm{OM}_\Lambda(X_p(t))$.

(ii) If $x_l \in X_L$ would be a bottleneck point of some node v_i allocated to x_l, then the distance function of this node $d_i(x_l(t/\nabla_l)) = d_{(n)}(X_p(t))$, $v_i = v_{(n)}$, would be concave with respect to t and, therefore, also $\mathrm{OM}_\Lambda(X_p(t))$, that is, we could find a descent direction for t, which contradicts the assumption that X_p is optimal.

(iii) Let $v_i \in \mathcal{V}$ be a node such that $w_i d(v_i, x_{l_1}) = w_i d(v_i, x_{l_2})$, that is, v_i possibly changes its allocation between the solution points x_{l_1} and x_{l_2} for $t < 0$ or $t > 0$, where one or both points are in X_L, w.l.o.g. $x_{l_1} \in X_L$. (Otherwise, x_{l_1} and x_{l_2} are fixed with respect to $t \in T$.) But, similar to Lemma 1, a reallocation of $v_i \in \mathcal{V}$ from, w.l.o.g., the solution point on e_{l_1} to the solution point on edge e_{l_2} can only occur if the distance function of node v_i on e_{l_1}, that is, $d_i(x_{l_1}(t/\nabla_1))$, has, with

F4

respect to t, a greater slope than has $d_i(x_2(t/\nabla_2))$ on edge e_2. Hence, the function $d_{(n)}(X_p(t)) = d_i(X_p(t))$, $v_{(n)} = v_i$, is concave over T for some $n = 1, \ldots, M$, which leads again to a contradiction to the optimality of X_p. ∎

Example 4. *Consider the situation of Example* 2. *We have* $\underline{n} = 7$ *and* $\bar{n} = 8$. *Hence,* $\bar{n} - \underline{n} + 1 = 2 = L$. *Starting with* $X_2 = \{x_1, x_2\}$, *where* $x_1 = ([v_1, v_2], \frac{10}{12})$ *and* $x_2 = ([v_5, v_7], \frac{3}{8})$, *we have* $i_1 = 1$ *and* $i_2 = 7$. *Define* $X_2(t) := \{x_1(t/6), x_2(t/-8)\}$. *If* t *is negative we strictly decrease the objective function value. Finally, for* $t = -1$, *we obtain the optimal solution* $X_2^* = \{x^* = ([v_1, v_2], \frac{2}{3}), EQ_{57}^{57} = ([v_5, v_7], \frac{1}{2})\}$.

Since $\mathcal{PEQ}$ is an FDS for the ordered p-median problem with $\Lambda = (a, \ldots, a, b \ldots, b)$, a natural question that arises now refers to the number of elements contained in the set $\mathcal{PEQ}$. Taking $K = |\mathcal{EQ}|$ and $M = |\mathcal{V}|$, we obtain a range r for every equilibrium and every pair of nodes u, $v \in \mathcal{V}$, $u \neq v$, yielding in total $|\mathcal{R}| = O(K + M^2)$ ranges. Since every distance function $d_i(\cdot)$ can assume a value $r \in \mathcal{R}$ in at most two points on an edge $e \in \mathcal{E}$, we have $O(NM)$ r-extreme points and as a result $|\mathcal{PEQ}| = O(NM(K + M^2))$.

From the above results, it is possible to devise an algorithm to solve the ordered p-median problem exactly. The new algorithm generalizes the one proposed by Pérez-Brito et al. [16] and will be presented in the next section.

2.1. An Algorithm for Solving the Ordered p-Median Problem with Λ^k-Vector

By Theorem 5, there always exists an optimal solution in which one of the points (e.g., x_p) is a node or an equilibrium point. From the proof of the theorem, it follows that all the other solution points are either nodes or pseudoequilibria with respect to the range of the equilibrium or one of the ranges of the node(s). Hence, we first compute the set of equilibria $\mathcal{EQ}$, then the ranges $\mathcal{R}$, and afterward the r-extreme points for every $r \in \mathcal{R}$. The latter must be saved with a reference to r in a set $\mathcal{PEQ}[r]$. Next, we choose a candidate x_p from the set $\mathcal{V} \cup \mathcal{EQ}$. If $x_p \in EQ_{ij}$ is an equilibrium of range $r = d_i(x_p) = d_j(x_p)$, then the objective function value $OM_\Lambda(X_{p-1} \cup \{x_p\})$ is determined for all $p - 1$ subsets $X_{p-1} = \{x_1, \ldots, x_{p-1}\}$ of $\mathcal{V} \cup \mathcal{PEQ}[r]$.

If $x_p = v_i$ is a node, then the set $\mathcal{R}_{v_i}$ is computed (see proof of Theorem 5). Furthermore, for all subsets $X_{p-1} = \{x_1, \ldots, x_{p-1}\}$ of $\mathcal{V} \cup \{\mathcal{PEQ}[r] | r \in \mathcal{R}_{v_i}\}$, the objective function value $OM_\Lambda(X_{p-1} \cup \{x_p\})$ must be determined.

A summary of the steps required to find an optimal solution of the ordered p-median problem with $\Lambda^k \in \mathbb{R}_{0+}^M$, $1 \leq k \leq M - 1$ is given below:

Algorithm 2.1

Computation of an optimal solution set X_p^*
Input: *Network* $\mathcal{N} = (\mathcal{G}, l)$, *distance-matrix* D, $p \geq 2$, *and a vector* $\Lambda^r = (a, \ldots, a, b, \ldots, b)$, $1 \leq r \leq M - 1$
Output: *An optimal solution set* X_p^*
1. *Initialization*
   Let $X_p^* := \emptyset$, $res := +\infty$
2. *First compute* $\mathcal{EQ}$, *then the set of ranges* $\mathcal{R}$, *and based on these sets, determine for every* $r \in \mathcal{R}$ *the* r-*extreme points and save them with a reference to* r *in a set* $\mathcal{PEQ}[r]$.
3. **FORALL** $EQ \in \mathcal{EQ}$ **DO**
   Let $x_p := EQ \in EQ_{ij}$ *and compute the range* r *of the equilibrium, i.e.,* $r := d_i(EQ) = d_j(EQ)$.
   FORALL $X_{p-1} = \{x_1, \ldots, x_{p-1}\} \subseteq \mathcal{V} \cup \mathcal{PEQ}[r]$ **DO**
      Compute $OM_\Lambda(X_p)$, *where* $X_p := X_{p-1} \cup \{x_p\}$.
      IF $OM_\Lambda(X_p) < res$ **THEN**
         $X_p^* := \{X_p\}$, $res := OM_\Lambda(X_p^*)$
4. **FORALL** $v_i \in \mathcal{V}$ **DO**
   Let $x_p := v_i$ *and compute the set* $\mathcal{R}_{v_i}$ *of all ranges of the node.*
   FORALL $X_{p-1} = \{x_1, \ldots, x_{p-1}\} \subseteq \mathcal{V} \cup \{\mathcal{PEQ}[r] | r \in \mathcal{R}_{v_i}\}$ **DO**
      Compute $OM_\Lambda(X_p)$, *where* $X_p := X_{p-1} \cup \{x_p\}$.
      IF $OM_\Lambda(X_p) < res$ **THEN**
         $X_p^* := \{X_p\}$, $res := OM_\Lambda(X_p^*)$
5. **RETURN** X_p^*

The above algorithm has complexity $O(pN^{p-1}M^p\log M(K + M^p))$. To show this, observe that the computation of the equilibria set $\mathcal{EQ}$ is possible in $O((NM + K)\log M)$ steps. If we integrate the computation of the range of an equilibrium and a node in the line-sweep algorithm, then the complexity for obtaining the set $\mathcal{R}$ is $O(K + M^2)$, where it is possible to compute $\mathcal{PEQ}$ in $O(NM(K + M^2))$. In Steps 3 and 4, we have to evaluate the objective function for a node or an equilibrium for all subsets of size $p - 1$ of $\mathcal{V} \cup \mathcal{PEQ}[r]$ and $\mathcal{V} \cup \{\mathcal{PEQ}[r] : r \in \mathcal{R}_{v_i}\}$, respectively. In the first case, we have $O(\binom{M + MN}{p-1}) = O((MN)^{p-1})$ and, in the second case, $O(\binom{M + (M-1)MN}{p-1}) = O((M^2N)^{p-1})$ different subsets ($|\mathcal{R}_{v_i}| = M - 1$). Since the evaluation of the objective function takes $O(pM \log M)$ time [because it is no longer possible to compute the ordered vectors $d_\leq(X_p)$ a priori in the line-sweep algorithm], we obtain for Steps 3 and 4 the complexity

$$O(pM \log M(K(MN)^{p-1} + M(M^2N)^{p-1}))$$
$$= O(pN^{p-1}M^p\log M(K + M^p)),$$

which is the total complexity of the algorithm. It is clear that these problems are *NP*-hard because the p-median and p-center problems are particular instances. Due to this reason, we may have to apply approximation algorithms in order to solve the problem.

The recent papers by Charikar et al. [2], Jain and Vazirani [9], and Charikar and Guha [3] provide constant-factor approximation algorithms for the p-median problem. At this point, it is not yet clear whether the techniques in those papers can be applied (with the necessary adjustments) to the ordered p-median problem with Λ-modeling weights.

In the next section, we show how to solve the ordered p-median problem with Λ-modeling weights on tree graphs in polynomial time. This development is important because it can be applied to approximate such problems in general graphs. In Bartal [1] and Charikar et al. [2], $O(\log M \log \log M)$ approximation algorithms are given for the p-median problem. These algorithms are based on solving p-median problems on a family of trees. (The general network metric is approximated by the tree metrics.) The same approach can be applied to the ordered p-median problem with Λ-modeling weights. Therefore, polynomial time algorithms for solving that problem on trees are useful to derive $O(\log M \log \log M)$ approximating solutions for ordered p-median problems on general networks.

3. A POLYNOMIAL ALGORITHM FOR THE ORDERED p-MEDIAN PROBLEM ON TREE NETWORKS

In this section, we solve the ordered p-median problem with at most two types of Λ-weights $[\Lambda = (a, \overset{M-s}{\cdots}, a, b, \overset{s}{\cdots}, b)]$ in polynomial time on a tree. To do that, we adapt and modify the dynamic algorithm for the p-centdian model on a tree proposed in Tamir et al. [24]. Assume that we are given a tree $\mathcal{T}$ with $|\mathcal{V}| = M$.

Using the discretizing result, it is clear that the optimal ordered p-median set can be restricted w.l.o.g. to the set $Y = \mathcal{PE2}$.

Once we restrict to trees, the set Y is of cardinality $O(M^4)$ (notice that $|\mathcal{E}| = M - 1$). Computing and augmenting these points into the node set of $\mathcal{T}$ has complexity $O(M^4)$ by the procedure in Kim et al. [11]. Let $\mathcal{T}^a$ denote the augmented tree with the node set Y. Each point in Y is called a *seminode*. In particular, a node in $\mathcal{V}$ is also a seminode.

Suppose now that the given tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$, $|\mathcal{V}| = M$ and $|\mathcal{E}| = M - 1$, is rooted at some distinguished node, say, v_1. For each pair of nodes v_i, v_j, we say that v_i is a *descendant* of v_j if v_j is on the unique path connecting v_i to the root v_1. If v_i is a descendant of v_j and v_i is connected to v_j with an edge, then v_i is a *child* of v_j and v_j is the (unique) *father* of v_i. If a node has no children, it is called a *leaf* of the tree.

As shown in Tamir [22], we can now assume w.l.o.g. that the original tree is a binary tree, where each nonleaf node v_j has exactly two children, $v_{j(1)}$ and $v_{j(2)}$. The former is called the *left child,* and the latter, the *right child*. For each node v_j, $\mathcal{V}_j$ will denote the set of its descendants.

Once the tree $\mathcal{T}^a$ has been obtained, a second prepro-cessing phase similar to the used in Tamir et al. [24] is performed. For each node v_j, we compute and sort the distances from v_j to all seminodes in $\mathcal{T}^a$. Let this sequence be denoted by $L_j = \{r_j^1, \ldots, r_j^m\}$, where $r_j^i \leq r_j^{i+1}$, $i = 1, \ldots, m - 1$, and $r_j^1 = 0$. We can assume w.l.o.g. that there is a one-to-one correspondence between the elements in L_j and the seminodes in Y (Tamir et al. [24]). The seminode corresponding to r_j^i is denoted by y_j^i, $i = 1, \ldots, m$.

We note that the total computational effort of this phase is $O(M^6)$ and can be achieved by using the centroid decom-position approach as in Kim et al. [11] or the procedure described in Tamir [22].

For each node v_j, an integer $q = 0, 1, \ldots, p$, $r_j^i \in L_j$, an integer $l = 0, 1, 2 \ldots, s$, and c being a weighted distance from any node to a seminode, let $G(v_j, q, r_j^i, l, c)$ be the optimal value of the subproblem defined on the subtree $\mathcal{T}_j$, given that a total of at least one and at most q seminodes (service centers) can be selected in $\mathcal{T}_j$. More-over, we assume that at least one of them is located in $\{y_j^1, \ldots, y_j^i\} \cap Y_j$, that exactly l vertices are associated to b λ-weights, and that the minimal distance allowed for an element with a b λ-weight is c (in the above subproblem, we implicitly assume no interaction between the seminodes in $\mathcal{T}_j$ and the rest of the seminodes in $\mathcal{T}$). The function $G(v_j, q, r, l, c)$ is computed only for $q \leq |\mathcal{V}_j|$, where $\mathcal{V}_j$ is the node set of $\mathcal{T}_j$, $l \leq \min(s, |\mathcal{V}_j|)$ (notice that a larger l would not be possible), and if $l > 0$, then $c \leq \max\{w_k d(v_k, y) | v_k \in \mathcal{V}_j$ and $y \in Y_j\}$. Also, for each node v_j we define

$$G(v_j, 0, r, 0, c) = +\infty.$$

Analogously, $G(v_j, q, r, l, c) = +\infty$ for any combination of parameters that leads to an infeasible configuration.

Similarly, for each node v_j, an integer $q = 0, 1, \ldots, p$, $r_j \in L_j$, an integer $l = 0, 1, 2 \ldots, s$, and c being a weighted distance from any node to a seminode, we define $F(v_j, q, r_j, l, c)$ as the optimal value of the subproblem defined in $\mathcal{T}_j$ satisfying the following conditions:

1. A total of q service centers can be located in $\mathcal{T}_j$;
2. There are already some selected seminodes in $Y\backslash Y_j$ and the closest among them to v_j is at a distance r_j;
3. There are exactly $l \leq \min\{|\mathcal{V}_j|, s\}$ vertices with b λ-weight in $\mathcal{T}_j$; and
4. c is the minimal weighted distance allowed for a weighted distance with a b λ-weight.

Obviously, the function F is only computed for those r_j^i that correspond to $y_j^i \in Y\backslash Y_j$.

The algorithm computes the function G and F at all the leaves of $\mathcal{T}$ and then, recursively, proceeding from the leaves to the root, computes these functions at all nodes of $\mathcal{T}$. The optimal value of the problem will be given by

$$\min_c G(v_1, p, r_1^m, s, c),$$

where v_1 is the root of the tree.

Define

$$f_j(r, l, c) = \begin{cases} ar & \text{if } r < c \\ br & \text{if } r \geq c \text{ and } l < s \\ +\infty & \text{if } r \geq c \text{ and } (l = s \text{ or } l = 0) \end{cases}$$

and

$$g_j(r, l, c) = \begin{cases} ar & \text{if } r < c \\ br & \text{if } r \geq c \text{ and } l > 0 \\ +\infty & \text{if } r \geq c \text{ and } l = 0. \end{cases}$$

Let v_j be a leaf of $\mathcal{T}$. Then,

$$G(v_j, 0, r_j^i, 0, c) = +\infty, \qquad i = 1, 2, \ldots, m, \qquad c \neq 0$$

$$G(v_j, 1, r_j^1, 0, c) = 0, \qquad c \neq 0,$$

$$G(v_j, 1, r_j^i, l, c) = +\infty \text{ otherwise.}$$

For each $i = 1, \ldots, m$ such that $y_j^i \in Y \backslash Y_j$,

$$F(v_j, 0, r_j^i, 0, c) = \begin{cases} ar_j^i & \text{if } r_j^i < c \\ +\infty & \text{if } r_j^i \geq c \end{cases}$$

$$F(v_j, 0, r_j^i, 1, c) = \begin{cases} br_j^i & \text{if } r_j^i \geq c \\ +\infty & \text{if } r_j^i < c \end{cases}$$

$$F(v_j, 1, r_j^i, 0, c) = 0$$

$$F(v_j, 1, r_j^i, 1, c) = +\infty.$$

Let v_j be a nonleaf node in V, and let $v_{j(1)}$ and $v_{j(2)}$ be its left and right children, respectively. The element r_j^1 corresponds to v_j. In addition, it corresponds to a pair of elements, say $r_{j(1)}^k \in L_{j(1)}$ and $r_{j(2)}^t \in L_{j(2)}$, respectively. Then,

$$G(v_j, q, r_j^1, l, c) = \min \left\{ \begin{array}{l} \min\limits_{\substack{q_1+q_2=(q-1)^+ \\ l_1+l_2=l}} \{F(v_{j(1)}, q_1, r_{j(1)}^k, l_1, c) + F(v_{j(2)}, q_2, r_{j(2)}^t, l_2, c)\}; \\[2em] \min\limits_{\substack{q_1+q_2=(q-2)^+ \\ l_1+l_2=l \\ y_{j(1)}^i \in (v_{j(1)}, v_j)}} \{F(v_{j(1)}, q_1, r_{j(1)}^i, l_1, c) + F(v_{j(2)}, q_2, r_{j(2)}^t, l_2, c); \\[2em] \min\limits_{\substack{q_1+q_2=(q-2)^+ \\ l_1+l_2=l \\ y_{j(2)}^i \in (v_{j(2)}, v_j)}} \{F(v_{j(1)}, q_1, r_{j(1)}^k, l_1, c) + F(v_{j(2)}, q_2, r_{j(2)}^i, l_2, c)\}\}, \end{array} \right.$$

where for any number a, we denote by $a^+ = \max(0, a)$.

For $i = 2, \ldots, m$ consider r_j^i. If $y_j^i \in Y \backslash Y_j$, then

$$G(v_j, q, r_j^i, l, c) = G(v_j, q, r_j^{i-1}, l, c).$$

If $y_j^i \in Y_{j(1)}$, then it corresponds to $r_{j(1)}^k \in L_{j(1)}$ and to $r_{j(2)}^t \in L_{j(2)}$. If $y_j^i \in Y_{j(1)}$, we can compute G in the following way:

$$G(v_j, q, r_j^i, l, c) = \min\{G(v_j, q, r_j^{i-1}, l, c), g_j(r_j^i, l, c)$$
$$+ \min_{\substack{q_1+q_2=q \\ 1 \leq q_1 \leq |V_{j(1)}| \\ q_2 \leq |V_{j(2)}| \\ l_1+l_2=\begin{cases} l & \text{if } r_j^i < c \\ (l-1)^+ & \text{if } r_j^i \geq c \end{cases} \\ l_i \leq \min\{l, |V_{j(i)}|\}, i=1,2}} \{G(v_{j(1)}, q_1, r_{j(1)}^k, l_1, c)$$
$$+ F(v_{j(2)}, q_2, r_{j(2)}^k, l_2, c)\}\}.$$

If $y_j^i \in (v_j, v_{j(1)})$ and $y_j^i \neq v_{j(1)}$, then

$$G(v_j, q, r_j^i, l, c) = \min\{G(v_j, q, r_j^{i-1}, l, c), g_j(r_j^i, l, c)$$
$$+ \min_{\substack{q_1+q_2=q-1 \\ 1 \leq q_1 \leq |V_{j(1)}| \\ q_2 \leq |V_{j(2)}| \\ l_1+l_2=\begin{cases} l & \text{if } r_j^i < c \\ (l-1)^+ & \text{if } r_j^i \geq c \end{cases} \\ l_i \leq \min\{l, |V_{j(i)}|\}, i=1,2}} \{F(v_{j(1)}, q_1, r_{j(1)}^k, l_1, c)$$
$$+ F(v_{j(2)}, q_2, r_{j(2)}^k, l_2, c)\}\}.$$

Analog formulas can be derived for $y_j^i \in Y_{j(2)}$ with the obvious changes.

Once the function G is obtained, we compute the function F. Let y_j^i be a seminode in $Y \backslash Y_j$. Thus, y_j^i corresponds to some elements, say $r_{j(1)}^k \in L_{j(1)}$ and $r_{j(2)}^t \in L_{j(2)}$. Therefore,

$$F(v_j, q, r_j^i, l, c) = \min\{G(v_j, q, r_j^i, l, c), f_j(r_j^i, l, c)$$

$$+ \min_{\substack{q_1+q_2=q \\ q_1 \leq |V_{j(1)}| \\ q_2 \leq |V_{j(2)}| \\ l_1+l_2 = \begin{cases} l & \text{if } r_j^i < c \\ (l-1)^+ & \text{if } r_j^i \geq c \end{cases} \\ l_i \leq \min\{l, |V_{j(i)}|\}, \ i=1, 2}} \{F(v_{j(1)}, q_1, r_{j(1)}^k, l_1, c)$$

$$+ F(v_{j(2)}, q_2, r_{j(2)}^k, l_2, c)\}\Big\}.$$

Complexity: It is clear that the complexity required to evaluate the functions G and F depends on the cardinality of the FDS for this problem. If $a \geq b$, then $\mathcal{V}$ is an FDS with cardinality $O(M)$ (see Nickel and Puerto [15]); else, if $a < b$, then $\mathcal{PEQ}$ is an FDS with cardinality $O(M^4)$. Therefore, it follows directly from the recursive equations that the effort to compute the function G at a given node v_j, for all relevant values of q, r, l, and c, is $O(p^2 M s^2 M^2)$ in case $a \geq b$ and $O(p^2 (M^4) s^2 M (M^4))$ if $a < b$. Therefore, the overall complexity of the algorithm is clearly

$$\begin{cases} O(p^2 s^2 M^3) & \text{if } a \geq b \\ O(M(psM^4)^2) & \text{if } a < b. \end{cases}$$

However, it is easy to verify that the analysis in Tamir [22] can also be applied to the above model to improve the complexity to

$$\begin{cases} O(ps^2 M^3) & \text{if } a \geq b \\ O(pM(sM^4)^2) & \text{if } a < b. \end{cases}$$

It is also worth noting that for $a < b$ and $s = 1$ the algorithm reduces to the one given in Tamir et al. [24] and the complexity is $O(pM^6)$. Moreover, if $s = 0$, the problem reduces to the p-median problem and the complexity is $O(pM^2)$ by the algorithm by Tamir [22]. Finally, recently and independently, Tamir [23] presented another algorithm for the p-facility k-centrum whose complexity for the discrete problem is $O((\min(k, p))kpM^5)$. Note that for the continuous version the number of possible candidates is $O(M^3)$ instead of $O(M)$ and, therefore, the overall complexity is at least $O((\min(k, p))kpM^8)$.

Since the p-centdian complexity is $O(pM^6)$, the complexity of the algorithm for the more general ordered p-median problem is not too large.

4. CONCLUSIONS

From a methodological point of view, the question for which classes of Λ vectors we obtain an FDS is of great importance. In this paper, we identify a polynomial size FDS for the multifacility ordered median problem. This set allows us to develop an exact algorithm for the problem. The algorithm has exponential complexity since the problem is NP-hard on general networks. Moreover, we give the first polynomial time algorithm for the problem on trees. This algorithm was used to develop a polynomial time $O(\log M \log \log M)$ approximate solution algorithm for the problem on general networks based on some approximation algorithms by Bartal [1] and Charikar et al. [2].

REFERENCES

[1] Y. Bartal, On approximating arbitrary metrics by tree metrics, Proc 30-th Ann Symp on Foundation of the Computer Sciences, 1998, pp. 161–168.

[2] M. Charikar, C. Chekuri, A. Goel, and S. Guha, Rounding via trees: Deterministic approximation algorithms for group Steiner trees and k-median, Proc 30-th Ann Symp on Foundation of the Computer Sciences, 1998, pp. 114–123.

[3] M. Charikar and S. Guha, Improved combinatorial algorithms for the facility location and k-median problems, Proc 31-st Ann Symp on Foundation of the Computer Sciences, 1999, pp. 378–388.

[4] Z. Drezner (Editor), Facility location. A survey of applications and methods, Springer, 1995. **AQ: 1**

[5] Z. Drezner and H.W. Hamacher (Editors), Facility location: Applications and theory. Springer, 2002. **AQ: 1**

[6] R.L. Francis, T.J. Lowe, and A. Tamir, Aggregation error bounds for a class of location models, Oper Res 48 (2000), 294–307.

[7] S.L. Hakimi, Optimum location of switching centers and the absolute centers and medians of a graph, Oper Res 12 (1964), 450–459.

[8] J.N. Hooker, R.S. Garfinkel, and C.K. Chen, Finite dominating sets for network location problems, Oper Res 39 (1991), 100–118.

[9] K. Jain and V.V. Vazirani, Primal-dual approximation algorithms for the metric facility facility location problem and k-median problems, Proc 31-st Ann Symp on Foundation of the Computer Sciences, 1999.

[10] J. Kalcsics, S. Nickel, J. Puerto, and A. Tamir, Algorithmic results for ordered median problems defined on networks and the plane, Oper Res Lett 30 (2002), 149–158.

[11] T.U. Kim, L.J. Lowe, A. Tamir, and J.E. Ward, On the location of a tree-shaped facility, Networks 28 (1996), 167–175.

[12] M. Labbé, D. Peeters, and J.-F. Thisse, "Location on networks," Handbooks in operations research and management science, Vol. 8: Network routing, C.L. Monma, M.O. Ball, T.L. Magnanti, and G.L. Nehmhauser (Editors), Elsevier, 1995. **AQ: 1**

[13] P.B. Mirchandani and R.L. Francis (Editors), Discrete location theory, Wiley, New York, 1990.

[14] S. Nickel, "Discrete ordered Weber problems," Operations research proceedings 2000, Springer-Verlag, Berlin, 2001, pp. 71–78.

[15] S. Nickel and J. Puerto, A unified approach to network location problems, Networks 34 (1999), 283–290.

[16] D. Pérez-Brito, J.A. Moreno-Pérez, and I. Rodríguez-Martín, Finite dominating set for the p-facility cent-dian network location problem, Stud Locat Anal 11 (1997), 27–40.

[17] J. Puerto (Editor), Lecturas en teoría de localización, Serie Ciencias 46, Secretariado de Publicaciones de la Universidad de Sevilla, Sevilla, Spain, 1996 (in Spanish).

[18] J. Puerto and F.R. Fernández, The symmetrical single facility location problem, Technical report, Prepublicacion de la Facultad de Matematicas, Universidad de Sevilla, Sevilla, Spain, 1995.

[19] J. Puerto and F.R. Fernández, Geometrical properties of the symmetrical single facility location problem, J Nonlin Convex Anal 1 (2000), 321–342.

[20] A.M. Rodríguez-Chía, S. Nickel, J. Puerto, and F.R. Fernández, A flexible approach to location problems, Math Methods Oper Res 51 (2000), 69–89.

[21] P. Slater, Center to centroids in graphs, J Graph Theory 2 (1978), 209–222.

[22] A. Tamir, An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs, Oper Res Lett 19 (1996), 59–64.

[23] A. Tamir, The k-centrum multi-facility location problem, Discr Appl Math 109 (2000), 292–307.

[24] A. Tamir, D. Pérez-Brito, and J.A. Moreno-Pérez, A polynomial algorithm for the p-centdian problem on a tree, Networks 32 (1998), 255–262.

AQ1   Please give city of publication.