Discussion Paper No. 367

An $O((n \log P)^2)$ Algorithm for the Continuous P-Center Problem on a Tree

R. Chandrasekaran and A. Tamir

February 1979

Graduate School of Management
Northwestern University

Introduction

In this study we consider center location problems on undirected tree networks. Suppose that an undirected tree $T = T(N,A)$, with N and A denoting the set of all nodes and the set of all arcs respectively, is given. Each arc of T is associated with a positive number called the length of the arc. By a point on T we mean a point along any edge of T. In particular the nodes are points on T. Using the arcs lengths, we define the distance between two points x and y on T as the length of the unique path connecting x and y. This distance is denoted by $d(x,y)$.

In addition, a set, D, of points on T is specified. D which may be finite or infinite, represents the set of demand points. Assume that supply centers can be located anywhere on the tree. Given a number, P, the objective is to find locations for P supply points on T, such that the supremum of the distances of the demand points in D to their respective nearest supply centers is minimized.

Two special cases of the above model have been treated in the literature. The first corresponds to the case where demand occurs only at the nodes of T, i.e. $D = N$. Whenever $|D| < \infty$ one can also associate weights with the demand points and talk about minimizing the maximum of the weighted distances to the nearest supply centers. Efficient, polynomially bounded algorithms when $D = N$ whatever be P are given in [11,3], while the further specialization to the case where $P \leq 2$ is discussed in [4,6,7,8,9,12].

The second special case of the general model is the continuous case when $D = T$, i.e. each point of the tree is a demand point. This model is studied in [2], where it is solved in polynomial time .

The general model introduced above is related to the following P-center dispersion problem. A set, S, of points on the tree T is specified. Given an integer P, the objective is to locate P facilities at points in S such that these P facilities are as far from each other as possible.

The next theorem shows that the P-center dispersion problem is, in a sense, dual to the location problem introduced before, provided the two corresponding sets S and D are identical. Let D be a subset of T. Then it is convenient for the statement of the theorem to let $U_p = \{u_1, \ldots, u_p\}$ and $V_{p+1} = \{v_1, \ldots, v_{p+1}\}$ denote finite subsets of T, and to define

$$f_D(U_p) = \sup_{x \in D} \{\min_{u_i \in u_p} d(x, u_i)\} \quad \text{and}$$

$$g(V_{p+1}) = \min \{d(v_i, v_j)/2 : 1 \leq i < j \leq p + 1\}$$

<u>Theorem 1</u>:

For any subset D of a tree T,

$$\min \{f_D(U_p) : U_p \subseteq T\} = \sup \{g(V_{p+1}) : V_{p+1} \subseteq D\} .$$

The specialization of the theorem where D is a finite set is proved in [3], using the equality of the maximum anticlique and the minimum cardinality clique cover in perfect graphs. The continuous case where D is the entire tree is proved in [12]. In fact the proof of the above theorem is very similar to the one given in [12] for the case $D = T$, and we therefore omit the proof.

In this paper we focus on the case when D, the set of demand points is the entire tree. We show that for a given P the minimum value of the objective function of the P-center location problem is equal to $d(i,j)/2k$, where $d(i,j)$ is a distance between some pair of nodes, i and j, of T, and k is an integer satisfying $1 \leq k \leq P$. This result is then used to improve the algorithm of [2], yielding the bound of $O((n \log P)^2)$ for the continuous P-center location problem, i.e. $D = T$, on a tree $T(N,A)$ with n nodes. We also indicate how to improve the $O(n^2 \log n)$ bound of the algorithms of [11, 3] for the discrete P-center problem, i.e. $D = N$, to obtain an $O(n^2)$ time algorithm.

The Continuoue P-center Problem

In this section we consider the problem of locating P facilities on a tree network in order to minimize the maximum of the distances of the points on the network to their respective nearest facility. Using the notation presented above we want to find r(P) such that

$$r(P) = \min \{ f_T(U_P) \; ; \; U_P \subseteq T \} \tag{1}$$

Given a point x on T and $r > 0$, we define $N_r(x)$, the r- neighborhood of x, by $N_r(x) = \{y \in T : d(x,y) \leq r\}$. The location problem is then to find the minimum r such that P r- neighborhoods will cover the entire T. Similarly, given $r > 0$ we consider the reverse problem of covering the tree with minimum number of r- neighborhoods. This number is denoted by M(r). It is clear that M(r) is a monotone, nonincreasing, stepwise function, which is continuous from the right. r(P) is, therefore, the smallest r such that $M(r) \leq P$.

The algorithm of [2] for finding r(P) is based on an O(n) subroutine for finding M(r) for an arbitrary $r > 0$. (n is the number of nodes in T.)

In this section we show that $r(P) = d(i,j)/2k$ where d(i,j) is a distance between some pair of tips, i and j, of T and k is an integer satisfying $1 \leq k \leq P$. The latter combined with the monotonicity of M(r) will imply that the O(n) routine for finding M(r) is to be applied at most $\log(n^2 P)$ times, before r(P) is found.

To prove our claim on r(P) we will need the algorithm of [2] for finding M(r). Thus, for the sake of completeness we describe it here as well.

<u>Algorithm 1</u>

Suppose that the tree is rooted at some node and arranged in levels. The level of a node is defined as the number of arcs in the unique path connecting the node with the root. We start with the tips of the highest level and consider a maximal set of such tips having the same immediate predecessor, say node s. We will call such a set a cluster and denote it by $C(s)$. Let $\{(s,i)\}$ be the set of arcs connecting those tips to s. It is clear that if the length of any arc (s,i) is greater than or equal to 2r, a facility may be located at a point on (s,i) whose distance from the tip i is r, and the length of this arc is reduced by 2r. Thus we may assume that this has occurred, if necessary (and that $C(s)$ has been suitably reduced) and that the lengths of these arcs are at most 2r.

Let

$$\alpha = \min_{i \in C(s)} \{\bar{d}(i,s) : \bar{d}(i,s) > r\} = \bar{d}(i_1^*,s)$$

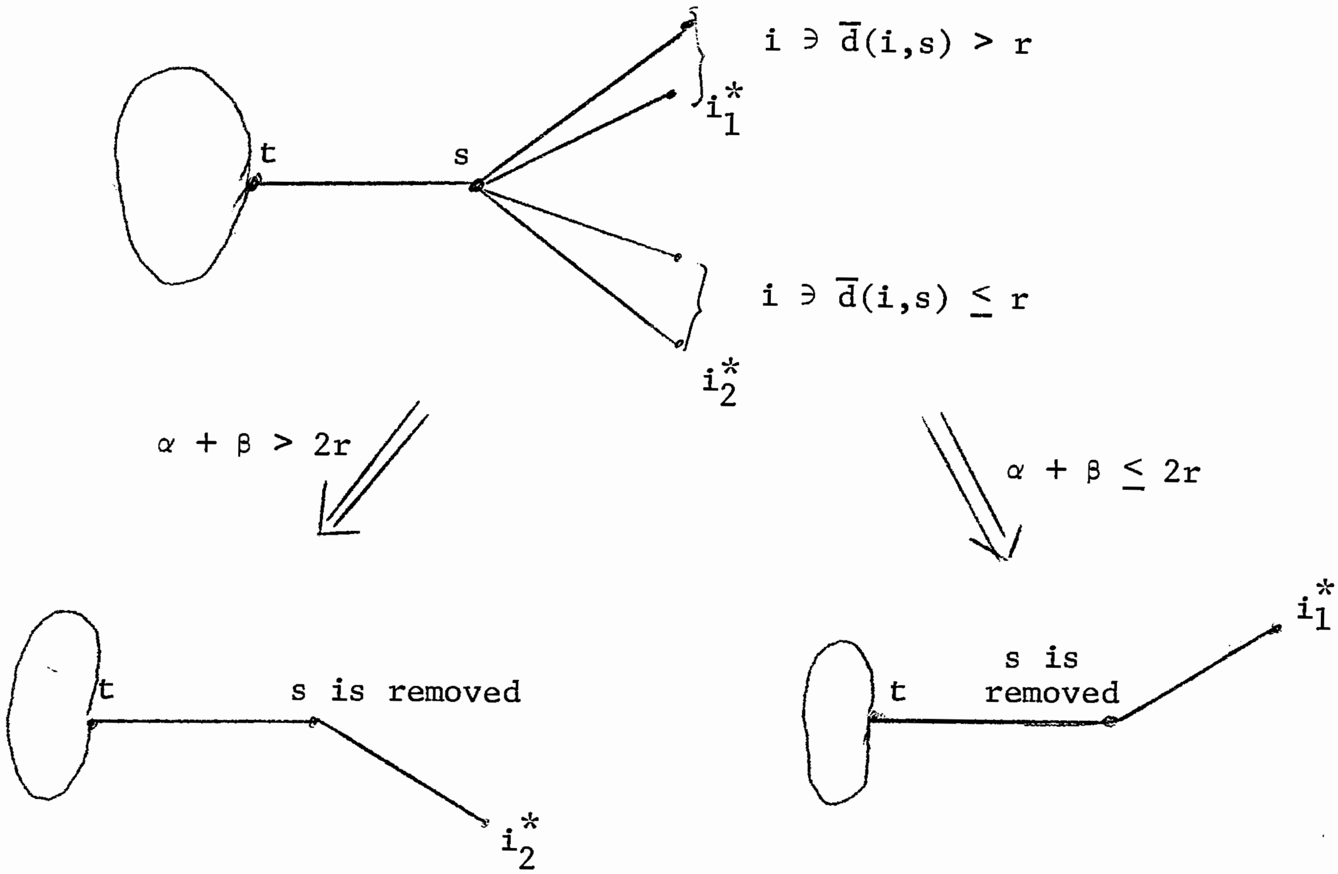
$$\beta = \max_{i \in C(s)} \{\bar{d}(i,s) : \bar{d}(i,s) \leq r\} = \bar{d}(i_2^*,s),$$

where $\bar{d}(i,s)$ is the reduced length of the arc (i,s). (If α or β are defined on empty sets they are set equal to ∞ and 0, respectively).

If $\alpha + \beta > 2r$, then locate a facility on (s,i) at a distance r from each tip i (of the reduced cluster) for which $\bar{d}(i,s) > r$, remove each spoke (s,i) in $C(s)$ except (s,i_2^*), and remove node s so that we have the case shown in Figure 1.

If $\alpha + \beta \leq 2r$, then locate a facility on (s,i) at a distance r from each tip i for which $\bar{d}(i,s) > r$, except i_1^*, remove all the spokes (s,i) except (s,i_1^*) and remove node s like in Figure 1. The process is then continued by considering a cluster of the highest level in the remaining tree.

It is clear that the above process will find $M(r)$ in $O(n)$ time.



$i \ni \bar{d}(i,s) > r$
i_1^*
t
s
$i \ni \bar{d}(i,s) \leq r$
i_2^*
$\alpha + \beta > 2r$
$\alpha + \beta \leq 2r$
t
s is removed
i_2^*
i_1^*
s is removed
t

Figure 1

Theorem 2

Let $r(P)$ be the solution to the continuous P-center problem, i.e. $r(P)$ is defined by (1). Then $r(P) = d(i,j)/2k$, where $d(i,j)$ is a distance between a pair of tips, i and j, of the tree T, and k is an integer, $1 \leq k \leq P$.

Proof

$r(P)$ is the minimum radius r, such that P r- neighborhoods will cover the entire tree T. We apply Algorithm 1 with $r = r(P)$ and show that if there is no integer multiple of $2r(P)$ which is equal to a distance between two tips of the tree then $r(P)$ is not minimal.

Assuming that T is rooted now and applying Algorithm 1, we see that at each step any tip of the reduced tree is at a distance of $2r(P)$ m from some tip of the original tree, with m being integer bounded by P. Therefore, if at some step the length of a pair of spokes from the same cluster is equal to $2r(P)$ the proof is complete. Hence, suppose that the latter has not occurred while applying the algorithm. For each point x on T, we define T_x as the tree rooted at x and consisting of all points y in T, such that x is on the unique simple path connecting y with the root of T. Let A be an arbitrary facility located during the application of the algorithm. (Observe that A is not necessarily an original node of T.) Let I(A) be the set of nodes of T which are immediate sons of A on T_A, and consider the set of rooted trees $\{T_i\}, i\in I(A)$. Then by our supposition there exists atmost one node, $i(A)$, in $I(A)$ such that the distance from A to a facility in $T_{i(A)}$ is exactly $2r$. The latter property holds for an arbitrary facility located by the algorithm. Thus, we can perturb the location of each facility A down towards $T_{i(A)}$ to obtain a cover of the tree T by the same number of neighborhoods, each having a radius smaller than $r(P)$. This contradicts the minimality of $r(P)$.

The above theorem implies that $r(P)$, the solution to the P-center problem can be found by applying Algorithm 1 $O(n^2P)$ times, thus yielding an $O(n^3P)$ bound for solving the model. Next we show a reduction of this bound which is based on the nature of the $O(n^2P)$ possible values for $r(P)$.

Due to the monotonicity property of $M(r)$, found by Algorithm 1, it is clear that if $M(\bar{r}) \leq P$ then $r(P) \leq \bar{r}$, and one can ignore all values of r greater than $\bar{r}$. Similarly, if $M(\bar{r}) > P$ we have $r(P) > \bar{r}$. Let R be the set of possible values for $r(P)$ as specified by Theorem 2. We start by finding the median of R, say r_1, and then apply Algorithm 1 to find $M(r_1)$. Comparing $M(r_1)$ and P we then eliminate half of the members of R from further discussion, remaining with the subset R_1. We then continue by finding the median of R_1, say r_2, computing $M(r_2)$, and so on. It is clear that Algorithm 1 is applied in this process $O(\log(n^2P))$ times.

We now show that the total effort of evaluating the sequence of medians $\{r_1, r_2, \ldots,\}$ is $O((n \log P)^2)$.

First, an effort of $O(n^2)$ yields the distances between all tips of T. For each such distance $d(i,j)$ the sequence $\{d(i,j)/2k\}$, $k = 1, \ldots, P$ is a monotone sequence. One can then apply the methods of [5,10] to find r_1 in $O(n^2 \log P)$ time. Successive applications of those methods for $q = \log P$ times will yield $r_1, r_2, \ldots, r_q$. By that time the remaining set of possible values, R_q, will contain $O(n^2)$ elements. Therefore, the remaining medians in the sequence are found in total effort of $O(n^2)$ using the linear time algorithm of [1]. By that we have demonstrated that the total effort of our procedure to find $r(P)$ is of order $O((n \log P)^2)$.

Finally, using the duality result presented in the introduction we observe that the P-center dispersion problem is also solvable in $O((n \log P)^2)$ time.

<u>Remarks</u>

1) The bound $O((n \log P)^2)$ given above for the continuous P-center problem can be further improved to $O(n^2 \cdot \min(\log n, \log P) \log P)$. To achieve the latter bound we have replaced the general algorithms of [5,10] by a special method that utilizes the structure of the set $\{d(i,j)/2k\}$, and finds the median of R_i in $O(n^2 \cdot \min(\log n, \log P))$.

2) The discrete P-center problem, i.e. the model where demand occurs only at the nodes of T, is solved in [11, 3] by an $O(n^2 \log n)$ algorithm. We indicate that this bound can be reduced to $O(n^2)$ for the method in [11]. The set R, of possible values for r(P) for the discrete problem is known to contain $O(n^2)$ elements. All these elements are computed in $O(n^2)$ total effort. Then, for each given r, an O(n) routine finding M(r), the minimum number of facilities covering all nodes, is given.

As was done above for the continuous P-center problem, one can generate the sequence of medians $\{r_1, r_2, \ldots\}$ and apply the procedure to find M(r) $O(\log(n^2)) = O(\log n)$ times. Since each time the cardinality of the remaining set R_i is cut by half, the linear time algorithm of [1] will generate the entire sequence of medians in total effort of $O(n^2)$. This latter term is then the dominating term yielding the bound $O(n^2)$ for the effort to find r(P).

1. Blum, M., R.W. Floyd, V.R. Pratt, R.L. Rivest and R.E. Tarjan, "Time bounds for selection," J. Compt. System Sci. 7 (1973), 448-461.

2. Chandrasekaran, R. and A. Daughety, "Problems of Location on Trees," Discussion Paper No. 357, Center for Mathematical Studies in Economics and Management, Northwestern University, 1978.

3. Chandrasekaran, R. and A. Tamir, "Polynomially bounded algorithms for locating P-centers on a tree," Discussion Paper No. 358, Center for Mathematical Studies in Economics and Management, Northwestern University, 1978.

4. Dearing, P.M. and R.L. Francis, "A Minimax Location Problem on a Network," Transportation Science, 8 (1974), 333-343.

5. Galil, Z. and N. Megiddo, "A fast selection algorithm and the distribution of effort," to appear in J. ACM (1979).

6. Goldman, A.J., "Minimax Location of a Facility in an Undirected Tree Graph," Transportation Science, 6 (1972), 407-418.

7. Hakimi, S.L., E.F. Schmeichel and J.G. Pierce, "On P-Centers in Networks," Transportation Science, 12 (1978), 1-15.

8. Handler, G.Y., "Minimax Location of a Facility in an Undirected Tree Graph," Transportation Science, 7 (1973), 287-293.

9. Handler, G.Y., "Finding Two-Centers of a Tree: The Continuous Case," Transportation Science, 12 (1978), 93-106.

10. Johnson, D.B. and T. Mizoguchi, "Selecting the Kth element in $X+Y$ and $X_1+X_2+\ldots+X_m$," SIAM J. Computing, 7 (1978), 147-153.

11. Kariv, O. and S.L. Hakimi, "An Algorithmic Approach to Network Location Problems," to appear in SIAM J. Appl. Math.

12. Shier, D.P. "A Min-Max Theorem for Network Location Problems on a Tree," Transportation Science, 11 (1977), 243-252.