SIAM J. COMPUT.
Vol. 13, No. 1, February 1984

0097-5397/84/1301-0014 $01.25/0

ON THE COMPLEXITY OF SOME COMMON GEOMETRIC LOCATION PROBLEMS*

NIMROD MEGIDDO† AND KENNETH J. SUPOWIT‡

Abstract. Given n demand points in the plane, the p-center problem is to find p supply points (anywhere in the plane) so as to minimize the maximum distance from a demand point to its respective nearest supply point. The p-median problem is to minimize the sum of distances from demand points to their respective nearest supply points. We prove that the p-center and the p-median problems relative to both the Euclidean and the rectilinear metrics are NP-hard. In fact, we prove that it is NP-hard even to approximate the p-center problems sufficiently closely. The reductions are from 3-satisfiability.

Key words. computational geometry, facility location, p-center problem, p-median problem, NP-hardness, approximation of NP-hard problems

1. Introduction. The goal of the present paper is to prove the NP-hardness of the following common problems in geometric location theory:

P1. *Euclidean p-center problem*: Given a set $X=\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$ of points in the plane, find a set $S=\{(z_1, t_1), (z_2, t_2), \cdots, (z_p, t_p)\}$ of p points so as to *minimize*

$$\max_{1\leq i\leq n} \min_{1\leq j\leq p} \{(x_i-z_j)^2+(y_i-t_j)^2\}.$$

Intuitively, we wish to minimize the radius R such that the points (x_i, y_i) $(i=1, 2, \cdots, n)$ can be enclosed by p circles of radius R.

P2. *Rectilinear p-center problem*: Following the notation of Problem 1, we wish to minimize

$$\max_{1\leq i\leq n} \min_{1\leq i\leq p} \{|x_i-z_j|+|y_i-t_j|\}.$$

In other words, we wish to minimize the number A such that all the points (x_i, y_i) $(i=1, 2, \cdots, n)$ can be enclosed within p squares of area A, the edges of each square forming angles of 45° with the axes.

P3. *Euclidean p-median problem*: Following the notation of Problem 1, we wish to minimize

$$\sum_{i=1}^{n} \min_{1\leq j\leq p} \{\sqrt{(x_i-z_j)^2+(y_i-t_j)^2}\}.$$

P4. *Rectilinear p-median problem*: Here we wish to minimize

$$\sum_{i=1}^{n} \min_{1\leq j\leq p} \{|x_i-z_j|+|y_i-t_j|\}.$$

We will prove that in fact it is NP-hard even to approximate P1 to within about 15% and P2 to within 50%.

We note that the analogous problems on the real line (instead of the plane) are both "easy": the p-center problem on the real line is solvable in $O(n \log n)$ time [10], [2], and the p-median problem on the real line is solvable in $O(n^2p)$ time [11].

* Received by the editors May 21, 1982, and in final form January 14, 1983.

† Department of Statistics, Tel Aviv University, Tel Aviv, Israel. The research of this author was partially supported by the National Science Foundation under grant ECS8121741 at Northwestern University, Evanston, Illinois 60201.

‡ Hewlett-Packard Laboratories, Palo Alto, California 94304.

The 1-median problem is also known as the Weber problem. No known algorithm finds the exact coordinates of the median (see [12] for a discussion of this difficulty).

The graphic counterparts of the p-center and the p-median problems are easily shown to be NP-hard [6], [7], using a reduction from Minimum Dominating Set. It is usually more complicated to prove NP-hardness of a geometric problem than of its graphic counterpart (see [4] and [14]).

Shamos [16] conjectures that P1 is NP-hard. Papadimitriou [15] proved NP-hardness of a different Euclidean p-median problem, namely, that in which the points (z_j, t_j) $(j = 1, 2, \cdots, p)$ must be selected from the set X. He mentioned the NP-hardness of both our Problems P1 and P3 as open. Previous versions of the proofs in the present paper were given in [18] and [8].

2. An overview of the proofs. In each of the proofs we establish a reduction from 3-satisfiability [5]. Formally, given a boolean expression

$$E = E_1 \wedge E_2 \wedge \cdots \wedge E_m,$$

where $E_j = x_j \vee y_j \vee z_j$ $(\{x_j, y_j, z_j\} \subseteq \{u_1, \bar{u}_1, u_2, \bar{u}_2, \cdots, u_q, \bar{u}_q\})$, the 3-satisfiability problem is to decide whether there exists a set $S \subseteq \{u_1, \bar{u}_1, u_2, \bar{u}_2, \cdots, u_q, \bar{u}_q\}$ such that

$$S \cap \{x_j, y_j, z_j\} \neq \emptyset \qquad (j = 1, 2, \cdots, m),$$

and

$$|S \cap \{u_i, \bar{u}_i\}| = 1 \qquad (i = 1, 2, \cdots, q).$$

The reduction from 3-satisfiability to a geometric problem will be established as follows. Each variable u_i $(i = 1, 2, \cdots, q)$ will be represented by a "circuit" of objects (e.g. circles, squares, points) in the plane. There will be essentially two different ways to partition the objects of the circuit so that the solution of the location problem is close to optimal. These two different partitions correspond to the choice of truth value for u_i. The clauses E_j $(j = 1, 2, \cdots, m)$ are represented by "clause configurations" which determine how the different circuits meet each other. A clause configuration relates the property that a clause is satisfied to the property that a partition is efficient from the point of view of the location problem.

Circuits must cross each other, without interfering with each other's properties; this requires that we design the "junctions" carefully. A schematic view of the circuits and their relations to the clause configurations is shown in Fig. 1. The details for each of the four problems are given in the succeeding sections.

3. The Euclidean p-center problem. We shall establish the NP-hardness of the Euclidean p-center problem by proving the following problem to be NP-hard:

Circle Covering. Given n unit circles in the plane and an integer $p > 0$, decide whether there exist p points such that each circle contains at least one point (we say that a circle *contains* a point if the point lies on, or in the interior of, the circle).

We now reduce 3-satisfiability to Circle Covering. In the reduction each variable u_i will be represented by a circuit of circles (see Fig. 2) $C^i = \{C^i_0, C^i_1, \cdots, C^i_{r_i}\}$, where $C^i_0 = C^i_{r_i}$, r_i is even and $C^i_k \cap C^i_l \neq \emptyset$ if and only if $|k - l| \leqq 1 \pmod{r_i}$. We say that a set of points Z *covers* a set of circles C if each circle in C contains at least one point in Z. Thus, at least $r_i/2$ points are required to cover C^i. There are essentially two different ways to cover all the circles, namely, either all the points belong to $C^i_k \cap C^i_{k+1}$ for $k = 0, 2, 4, \cdots$, or all belong to $C^i_k \cap C^i_{k+1}$ for $k = 1, 3, 5, \cdots$. In the former case, corresponding to the assignment of "true" to u_i, the points are called *true points*; in

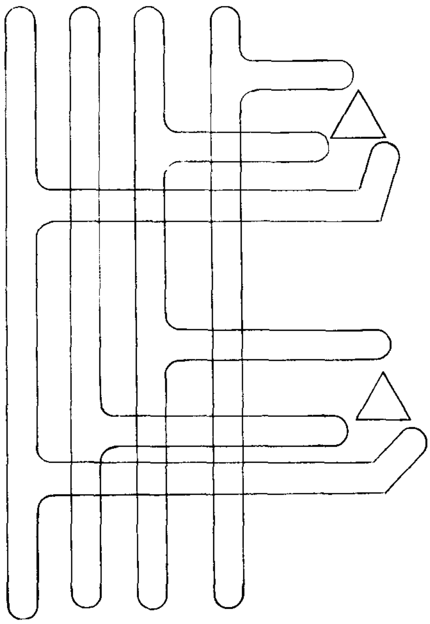

FIG. 1. *A schematic view of the reductions.*

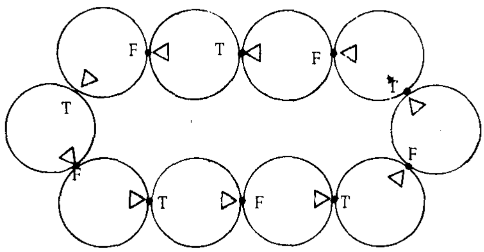

FIG. 2. *A circuit in the reduction for circle covering.*

the latter case, corresponding to the assignment of "false" to u_i, the points are called *false points.* We note that circuits may have to cross each other; we specify later how the "junctions" are designed.

Each clause E_j is represented in the reduction by a configuration of four circles as shown in Fig. 3. Specifically, there is one central circle that intersects the intersection of every two other circles of the configuration. However, the intersection of all four circles is empty. These properties imply that two points are both necessary and sufficient to cover all four circles, namely, one point to cover the central circle and two other circles, and another point to cover the remaining circle. Denote the central circle by D^j and the other three by D^j_x, D^j_y and D^j_z, corresponding to the literals x_j, y_j and z_j, respectively. The circle D^j_x, for example, intersects precisely two circles C^i_k, C^i_{k+1}, where i is such that $x_j \in \{u_i, \bar{u}_i\}$. Moreover, $D^j_x \cap C^i_k \cap C^i_{k+1} \neq \varnothing$. If $x_j = u_i$ then k is even; otherwise $(x_j = \bar{u}_i)$ k is odd. Thus, if the assignment of a truth value to u_i implies that some point in $C^i_k \cap C^i_{k+1}$ is selected, then this point may be selected so as to belong to D^j_x as well. It thus follows that if the overall truth assignment satisfies E_j, then at least one of the circles D^j_x, D^j_y, D^j_z can be covered by a true or a false point. The circle D^j can never be covered by such a point; thus we need precisely one more point per satisfied clause to guarantee that all the corresponding clause configurations are covered.

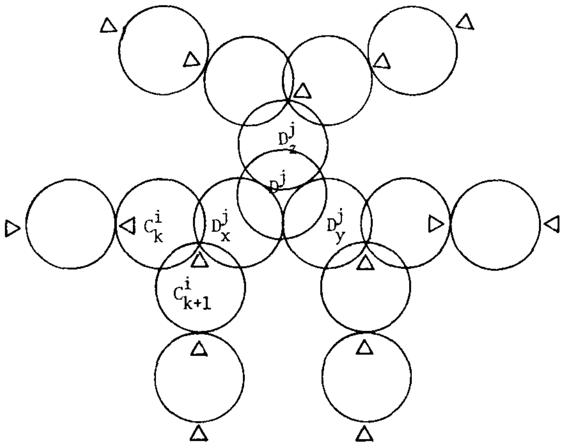

FIG. 3. *A clause configuration in the reduction for circle covering.*

We now discuss junctions. In each junction, a vertical segment of one circuit crosses a horizontal segment of another; the exact structure is shown in Fig. 4. Formally, a junction common to the circuits corresponding to u_i and to u_j has the following characteristics: Suppose that the circuits meet at a circle C^i_k (of the circuit corresponding to u_i) which is identical with a circle C^j_l (of the circuit corresponding to u_j). We insist that both k and l be odd numbers. This ensures that the segments of circuits between consecutive junctions have equal numbers of true points and of false points.

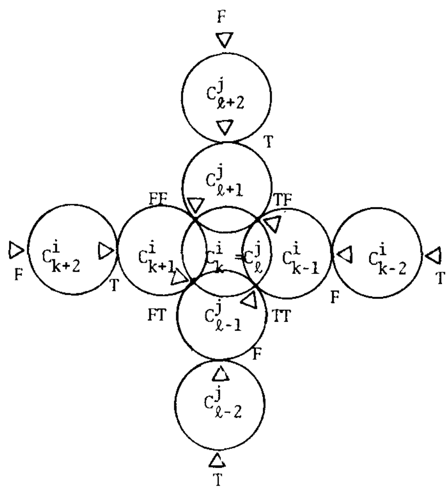

FIG. 4. *A junction in the reduction for circle covering.*

Furthermore, the junction is designed so that the central circle $C^i_k = C^j_l$ intersects the following *nonempty* sets: $C^i_{k-1} \cap C^j_{l-1}$, $C^i_{k-1} \cap C^j_{l+1}$, $C^i_{k+1} \cap C^j_{l-1}$ and $C^i_{k+1} \cap C^j_{l+1}$. Note that the requirement $C^i_{k-1} \cap C^i_{k+1} = C^j_{l-1} \cap C^j_{l+1} = \varnothing$ is satisfied. These facts imply that one point may be saved at each junction. Specifically, a point of the intersection $C^i_{k+1} \cap C^i_k \cap C^j_{l-1}$, for example, is both a false point for the circuit of u_i and a true point for the circuit of u_j. Thus if we assign false to u_i and true to u_j then the points marked with arrows in Fig. 4 constitute a cover for the two circuits that complies with this truth assignment. We denote the number of junctions by J.

Letting

$$p=\sum_{i=1}^{q}\frac{r_i}{2}+m-J,$$

we claim that E is satisfiable if and only if there exists a set Z of p points covering our entire structure. First, assume that E is satisfied by a truth assignment τ. For $i=1,2,\cdots,q$, if $\tau(u_i)=$ true, then include in Z the true points of the circuit for u_i; otherwise include in Z the false points of the circuit for u_i. Since one point is saved per junction, Z so far contains $\sum_{i=1}^{q} r_i/2-J$ points. Since each clause is satisfied, include in Z only one more point per clause in order to cover the central circle as well as the at most two more circles of the clause configuration not covered by a true or a false point. Thus Z contains p points and covers the complete structure.

To prove the converse, let Z be a set of p points that covers the entire structure. We will construct a truth value τ satisfying E. To that end, we will count the number of points available and conclude that the points selected for covering each circuit are either all true points or all false points. Consider a *segment* of a circuit between two consecutive junctions, i.e. a maximal set of circles of the form $\{C^i_{k+2}, C^i_{k+3},\cdots, C^i_{l-2}\}$, where k and l are odd and for each s, $k+2\leqq s\leqq l-2$, C^i_s is not involved in any junction. It follows that the length of each segment is odd, and that segments are pairwise disjoint. Furthermore, there are $2J$ of them, since each junction touches four segments and each segment touches two junctions. The total length of the segments is equal to

$$\left(\sum r_i-J\right)-5J=\sum r_i-6J.$$

Since there are $2J$ segments, each of odd length, it follows that

$$\frac{1}{2}\left(\sum r_i-6J-2J\right)+2J=\sum\frac{r_i}{2}-2J$$

points are required to cover all of them. It is easy to verify that within each segment, except for at most one point, either all the selected points are true, or all are false (provided that all of the segments are covered with no more than $\sum r_i/2-2J$ points). We are therefore left with only

$$p-\left(\sum\frac{r_i}{2}-2J\right)=m+J$$

more points with which to cover the rest of the circles (i.e. the junctions and the clause configurations), possibly with the aid of the previous points. However, the central circle of each junction and the central circle of each clause configuration are not covered by any point that covers a circle of a segment; hence we need to allocate *precisely one* more point for each junction and for each clause configuration. However, this implies that the aid from the previous points must be organized carefully.

Consider first the junctions. Each junction consists of five circles. The point which covers the central circle can cover at most two more circles; thus two circles per junction need to be covered with the aid of points covering the segments. However, each segment may assist in covering at most one circle of the junction; hence each segment must assist by covering precisely one such circle. Moreover, the assistance at each junction must come from segments of two distinct circuits, since the point that covers the central circle of the junction also covers two more circles belonging to distinct circuits. It finally follows that at each junction the two segments of the same circuit involved must be covered with the same truth-value type of cover. This implies that the points selected for the cover induce an assignment τ of truth value to the literals.

Now consider any clause configuration. The point that covers the central circle cannot cover all three other circles; hence at least one of the other three is covered with the aid of points covering the segments. The particular way of constructing the clause configuration ensures that its corresponding clause is satisfied by τ. This completes the proof that 3-satisfiability is reducible to Circle Covering. It is easy to verify that the reduction is polynomial.

An interesting consequence of this reduction applies to the NP-hardness of finding an approximate solution to the p-center problem. Suppose that instead of unit circles we draw circles of radius $R \geqq 1$ centered at the centers of the circles used in the reduction. We claim that as long as $R < 2/\sqrt{3}$, the same intersection relations hold among the circles; that is, the minimum number of points required to cover all the circles is independent of R when $1 \leqq R < 2/\sqrt{3}$ (if $R = 2/\sqrt{3}$ then the intersection of all the four circles in a clause configuration is nonempty and hence p points suffice even if E is not satisfiable). It follows that an approximate solution to the p-center problem less than $2/\sqrt{3}$ times the optimal is necessarily an optimal solution. This implies that if $\mathrm{P} \neq \mathrm{NP}$ then no polynomial-time algorithm for the p-center problem always gives a solution less than $2/\sqrt{3} \approx 1.15$ times the optimal; in other words, it is NP-hard to approximate the p-center problem with a relative error of less than about 15%.

4. The rectilinear *p*-center problem. We now prove the NP-hardness of the rectilinear p-center problem, which is more surprising that the NP-hardness of the Euclidean problem because of the following facts:

(i) The rectilinear 1-center problem is trivially solvable in linear time [3], while the Euclidean problem requires much more sophisticated tools [1], [13], [17], [9].

(ii) The rectilinear problem seems to decompose into two one-dimensional problems. This is true in the case $p = 1$, but turns out to be false in general, in view of our NP-hardness result.

As in the Euclidean case, we will consider a "covering" problem. In the present case, instead of circles, we deal with squares that are all identically oriented; without loss of generality we may assume that their boundaries are parallel to the axes.

Square Covering. Given n unit squares in the plane, each of whose boundaries are parallel to the axes, and an integer $p > 0$, decide whether there exist p points such that each square contains at least one point.

As an aside, we note that our squares have the property of interval intersection, namely, a set of such squares has a nonempty intersection if and only if every two squares in the set intersect. We now define the concept of a square-intersection graph, which extends the notion of an interval-intersection graph. Specifically, an undirected graph is a *square-intersection graph* if there is a one-to-one correspondence between the vertices of the graph and a set of squares in the plane (whose boundaries are parallel to the axes) such that two squares intersect if and only if their corresponding vertices are linked with an edge. Obviously, two such squares intersect if and only if their intervals of projection on the axes intersect. However, the problem of minimum cover by cliques is easily seen to be polynomial on an interval-intersection graph and, as follows from our results, NP-hard on a square-intersection graph.

The proof of NP-hardness for square covering is an adaptation of that for circle covering. First, variable u_i is represented by a circuit $\{S_0^i, S_1^i, \cdots, S_{r_i}^i\}$, of squares as is shown in Fig. 5; r_i is even. Two squares of a circuit intersect if and only if they are adjacent. A clause configuration is shown in Fig. 6. The intersection of two squares may have positive measure. Each clause E_j is represented by a *single* square S_j that

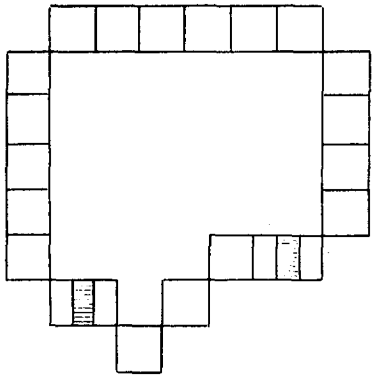

FIG. 5. *A circuit in the reduction for square covering.*

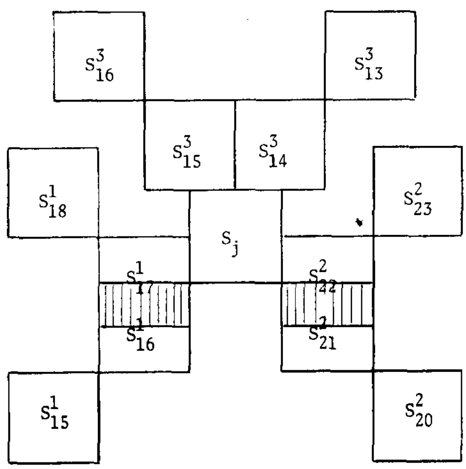

FIG. 6. *A clause configuration in the reduction for square covering.*

touches the three circuits involved (this differs from the proof for circle covering, in which each clause was represented by *four* extra circles; see Fig. 3). A junction is illustrated in Fig. 7. In a junction we simply coalesce a square of one circuit with a square of the other circuit. The coalesced squares must each be odd indexed in its respective circuit so that there will be an *odd* number of squares between consecutive junctions of a circuit. The four corners of the junction square correspond to the four combinations of possible truth values for the corresponding variables. Thus, in order to cover all the circuits, we need $p = \sum r_i/2 - J$ points, where J is the number of junctions. If E is satisfiable then p points suffice for covering the clause configurations as well as the circuits if the type of cover in each circuit is chosen appropriately. Conversely, suppose that p points suffice to cover the entire structure. Then consider segments of circuits between consecutive junctions: $\{S^i_{k+2}, S^i_{k+3}, \cdots, S^i_{l-2}\}$, where S^i_k and S^i_l are consecutive junction squares in the circuit for u_i. As in the Euclidean case, we need $\sum r_i/2 - 2J$ points in order to cover all these segments. We are therefore left with only J more points with which to cover the junctions and the clause configurations. By arguments analogous to those used in the Euclidean case, it follows that the cover with p points induces truth values that satisfy E.

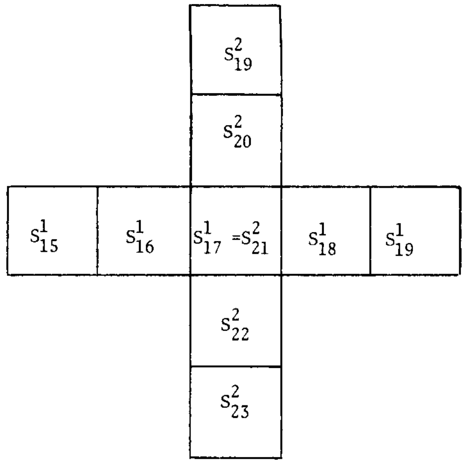

FIG. 7. *A junction in the reduction for square covering.*

An example of the entire structure of squares in the reduction is illustrated in Fig. 8.

As in the Euclidean case, the problem of finding an approximate solution is NP-hard. More precisely, the intersection relations among squares remain the same even if we enlarge their sizes by any factor less than $\frac{3}{2}$. Indeed, if the two unit squares of a circuit which touch a square representing a clause have 50% overlap (see Fig. 6), then by inflating each square by a factor of $\frac{3}{2}$ we obtain nonempty intersections of squares belonging to distinct circuits. This implies that, assuming $P \neq NP$, no polynomial-time algorithm for the rectilinear problem always gives solutions less than $\frac{3}{2}$ times the optimal.

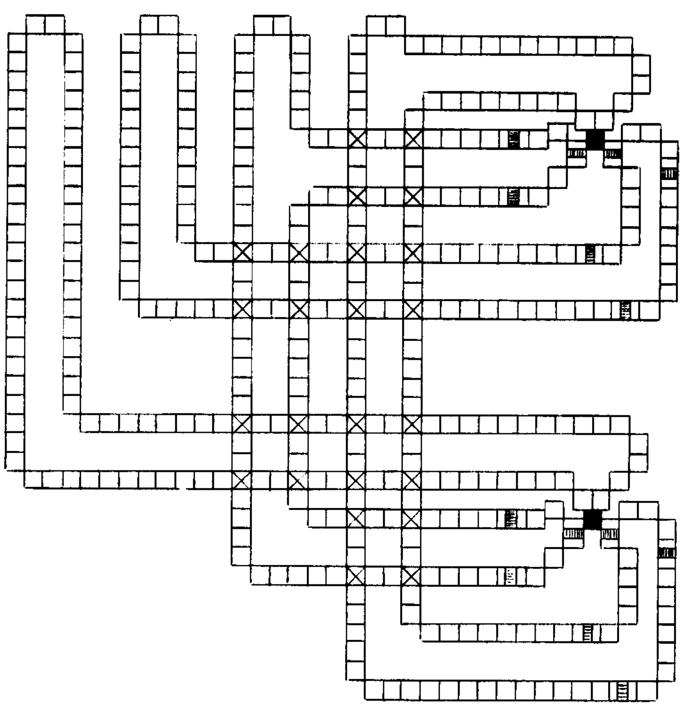

FIG. 8. *The complete structure for a sample square-covering reduction.*

5. The rectilinear p-median problem. We reduce 3-satisfiability to the rectilinear p-median problem. Here clause E_j is represented by a single point P_j^*, while variable u_i is represented by a circuit C^i of points $\{P_0^i, P_1^i, \cdots, P_{r_i}^i\}$, such that $P_0^i = P_{r_i}^i$, and $r_i \equiv 0 \pmod 3$. Denoting the rectilinear distance by d, we also require that if $k \equiv 0 \pmod 3$ then $d(P_k^i, P_{k+1}^i) = 1$, and that otherwise $d(P_k^i, P_{k+1}^i) = b \gg 1$. Moreover, if $|k-l| > 1 \pmod{r_i}$ then $d(P_k^i, P_l^i) > b$. An example of a circuit is shown in Fig. 9.

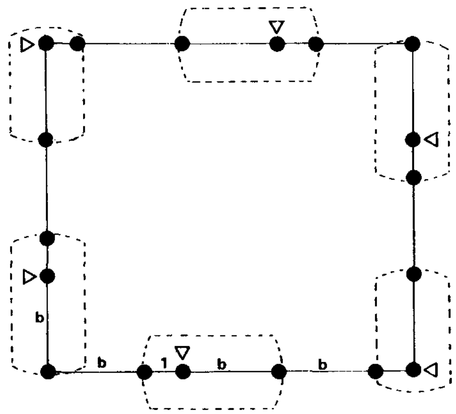

FIG. 9. *A circuit in the reduction for rectilinear p-median. The triangles indicate points P_k^i such that $k \equiv 0 \pmod 3$; hence the dotted lines indicate a true partition.*

Consider the p-median problem on the circuit C^i with $p = r_i/3$. The circuit may be partitioned into triples of points of the form (x, y, z), where $d(x, y) = 1$, $d(y, z) = b$ and $d(x, z) = b + 1$. We claim that the optimal solution of the p-median problem on C^i is obtained by partitioning into such triples and allocating one point per triple. Formally, let $f(S)$ denote the minimum of the 1-median problem on a subset S of a circuit C^i. The p-median problem on C^i may be rephrased as: Partition C^i into p sets $S_1, S_2, \cdots, S_p$ so as to minimize $\sum f(S_j)$. Now consider the function g, where

$$g(1) = 0, \qquad g(2) = 1, \qquad g(k) = (k-2)b + 1 \quad \text{for } k \geqq 3.$$

It can be verified that for every subset S of C^i, $f(S) \geqq g(|S|)$. Considering the problem of minimizing $\sum g(s_j)$ subject to $\sum s_j = 3p$, we observe that the optimal solution is $s_j = 3$, $j = 1, 2, \cdots, p$. Thus, the value $p(b+1)$ is a lower bound on the minimum of $\sum f(S_j)$. Note that there are two different partitions that yield the same value of $p(b+1)$. These are the partitions into triples $P_{k-1}^i, P_k^i, P_{k+1}^i$ where either $k \equiv 0 \pmod 3$ for every k (this is called the *true partition*) or $k \equiv 1 \pmod 3$ for every k (this is called the *false partition*). The solution points coincide with the middle points P_k^i of the triples. Note that every other selection of solution points does not achieve the lower bound of $(b+1)r_i/3$ on a circuit.

We now discuss the clause configurations. For each clause $E_j = x_j \vee y_j \vee z_j$, we allocate one point P_j^* that is situated at a distance of b from three points, one from each circuit related to the clause E_j (see Fig. 10). The point of a circuit that is nearest to P_j^* is chosen according to the relation of the corresponding literal to E_j. For example, if $x_j = u_i$, then the point P_k^i of the circuit of u_i that is nearest to P_j^* is such that $k \equiv 0 \pmod 3$; if $x_j = \bar{u}_i$ then $k \equiv 1 \pmod 3$. The second nearest point is at a distance of $b+1$ from P_j^* and each other point is at a distance of at least $2b$. Suppose that S is a set of points that contains precisely one "clause point" P_j^* and 0 or more circuit points (we have not yet introduced junctions). Consider the minimum sum of distances $f(S)$ in the 1-median problem on the set S. We claim that $f(S) \geqq h(|S|)$, where

$$h(1) = 0, \quad h(2) = b, \quad h(3) = b + 1, \quad h(k) = (k-2)(b+1) - 1 \quad \text{for } k \geqq 4.$$

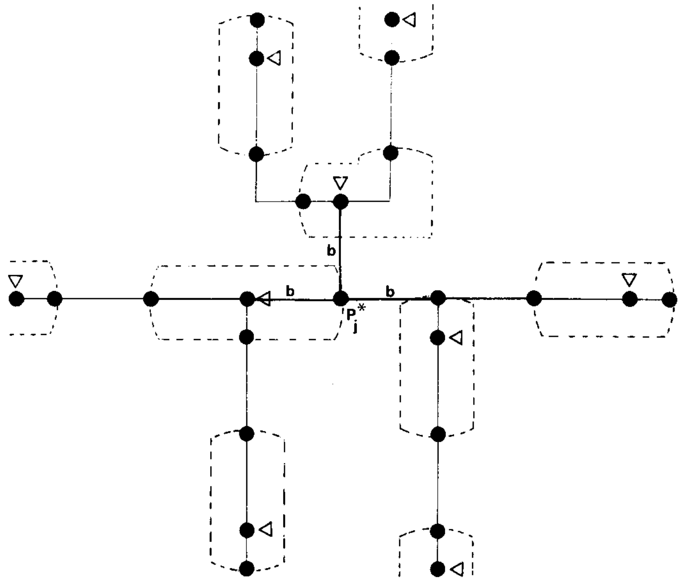

FIG. 10. *A clause configuration in the reduction for rectilinear p-median.*

In particular, the minimum $f(S)$ when $|S|=4$ is attained at sets of the form $S=\{P^i_{l-1}, P^i_l, P^i_{l+1}, P^*_j\}$, where P^i_l is closest to P^*_j, in which case $f(S)=2b+1$ (the solution point coincides with P^i_l). Our strategy is to enforce that if the truth values are chosen so that E_j is satisfied, then there is a solution point at a distance of b from P^*_j.

We now describe the junctions. As in the previous problems, a junction is established when a vertical piece of one circuit meets a horizontal piece of another. A junction of the circuits C^i and C^j is shown in Fig. 11. The junction occurs at a unit

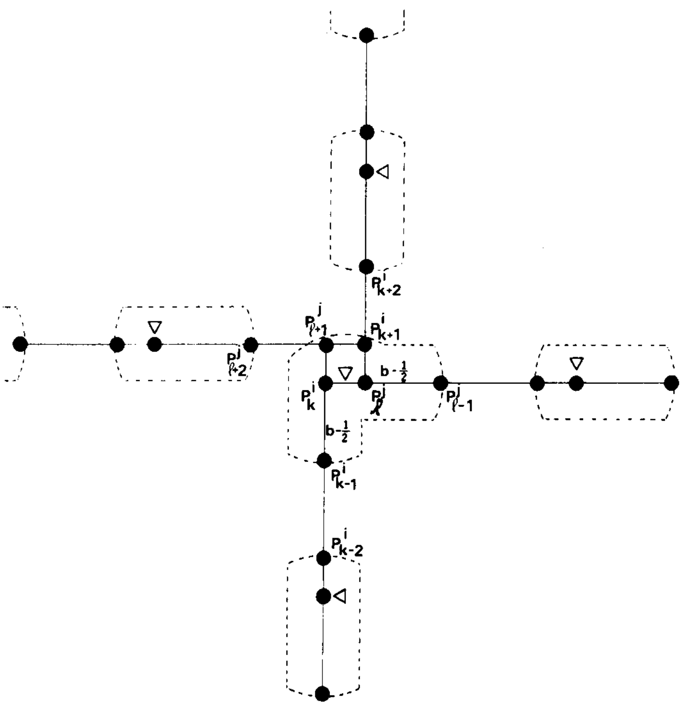

FIG. 11. *A junction in the reduction for rectilinear p-median.*

square whose vertices are the points P_k^i, P_l^j, P_{k+1}^i, P_{l+1}^j such that $k, l \equiv 0 \pmod 3$. The points P_{k-1}^i, P_{l-1}^j, P_{k+2}^i, P_{l+2}^j are each at a distance of $b-\frac{1}{2}$ from some vertex of this square. Note that the sum of rectilinear distances from the vertices of a unit square to any point on the square is equal to 4. Suppose that S is a set of points of our structure that contains at least one junction point. Then $f(S) \geqq e(|S|)$, where

$$e(1)=0, \quad e(2)=1, \quad e(3)=2, \quad e(4)=4, \quad e(5)=b+3.5, \quad e(6)=2b+4,$$
$$e(7)=3b+4.5, \quad e(8)=4b+6.$$

Consider now the p-median problem on the entire structure we have defined, where $p=\sum r_i/3-J$. Note first that there is a set of p solution points such that each of the points P_k^i with $k \equiv 2 \pmod 3$ and each clause point P_j^* has a solution point at a distance not greater than $b+1$ from it, while each other point has a solution point at a distance not greater than 1 from it. Thus, if b is sufficiently large, then an optimal solution for the p-median problem must yield a total distance less than $(\sum r_i/3+m)b+A$, where A is some constant independent of b. By considering the segments between consecutive junctions (as in the previous problems), we find that we must allocate precisely one solution point per junction.

More formally, let

$$T=(b+1)\sum \frac{r_i}{3}+mb+2J.$$

First, assume that E is satisfied by a truth assignment τ; we will construct a solution to our p-median problem of value T. τ induces a set of solution points at locations on the circuits and on one edge per junction square such that every point P_k^i ($k \equiv 2 \pmod 3$) has a solution point at a distance of b from it (see Fig. 11). The same is true for the clause points P_j^*. We thus manage to have a total distance of

$$\left(\sum \frac{r_i}{3}+m\right)b+4J+\left(\sum \frac{r_i}{3}-2J\right)1=(b+1)\sum \frac{r_i}{3}+mb+2J=T.$$

To show the converse, assume that there is a set Z of p points such that

$$\sum_{i=1}^{n} \min_{(z_1,z_2)\in Z} \{|x_i-z_1|+|y_i-z_2|\}=T.$$

Note that the p-median problem amounts to partitioning our set of $n=\sum r_i+m$ points into $p=\sum r_i/3-J$ sets $S_1, S_2, \cdots, S_p$ and then solving a 1-median problem of each S_j. We know that the partition is into sets of the following types:

(i) m sets each containing precisely one clause point and no junction points. If S is of this type then $f(S) \geqq h(|S|)$.

(ii) J sets containing four junction points (one whole junction) and no clause points. If S is of this type then $f(S) \geqq e(|S|)$.

(iii) $p-m-J$ sets containing neither junction points nor clause points. These sets satisfy $f(S) \geqq g(|S|)$.

Now consider the optimization problem of finding a vector $s=(s_1, s_2, \cdots, s_p)$ so as to minimize

$$\sum_{j=1}^{m} h(s_j)+\sum_{j=m+1}^{m+J} e(s_j)+\sum_{j=m+J+1}^{p} g(s_j)$$

subject to $\sum s_j = n$. Consider Table 1, showing "marginal costs". It follows from Table 1 that by letting

$$s_j = \begin{cases} 4, & 1 \leqq j \leqq m, \\ 6, & m+1 \leqq j \leqq m+J, \\ 3, & m+J+1 \leqq j \leqq p, \end{cases}$$

TABLE 1

k	2	3	4	5	6	7
$h(k)-h(k-1)$	b	1	b	$b+1$	$b+1$	$b+1$
$e(k)-e(k-1)$	1	1	2	$b-\frac{1}{2}$	$b+\frac{1}{2}$	$b+\frac{1}{2}$
$g(k)-g(k-1)$	1	b	$b+1$	$b+1$	$b+1$	$b+1$

we obtain an optimal solution to the optimization problem. The value of this solution is

$$\begin{aligned} h(4)m+e(6)J+g(3)(p-m-J) &= m(2b+1)+J(2b+4)+(p-m-J)(b+1) \\ &= \left(\sum \frac{r_i}{3}\right)(b+1)+mb+2J = T, \end{aligned}$$

which is, of course, a lower bound on the solution of the p-median problem. We have already seen that it is realizable if E is satisfiable. Hence the solution induced by the truth values is optimal. Moreover, if this bound is realizable then necessarily every set of four points of type (i) must have a total distance of $2b+1$, and this is possible if and only if each clause point has a solution point that reflects the fact that the clause is satisfied. The characteristic that a solution with total distance of T must induce truth values is established by considering the segments of circuits between junctions as in the previous problems. In summary, E is satisfiable if and only if the optimal solution of the p-median problem is T.

6. The Euclidean p-median problem. The proof that the Euclidean p-median problem is NP-hard is very similar to that for the rectilinear case. Note that if the angles of the polygon corresponding to a circuit are greater than or equal to 120°, then solution points around the polygon coincide with circuit points (see Fig. 12). A clause configuration is shown in Fig. 13. We note that the 1-median problem on a set of three points of a circuit P^i_k, P^i_{k+1}, P^i_{k+2} (where $k \equiv 0 \pmod 3$ or $k \equiv 1 \pmod 3$) has an optimal value of $b+1$. Moreover, the function g of the preceding section is

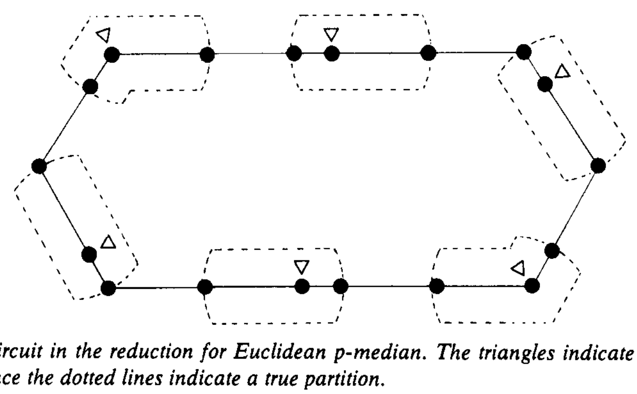

FIG. 12. *A circuit in the reduction for Euclidean p-median. The triangles indicate points P^i_k such that $k \equiv 0 \pmod 3$; hence the dotted lines indicate a true partition.*

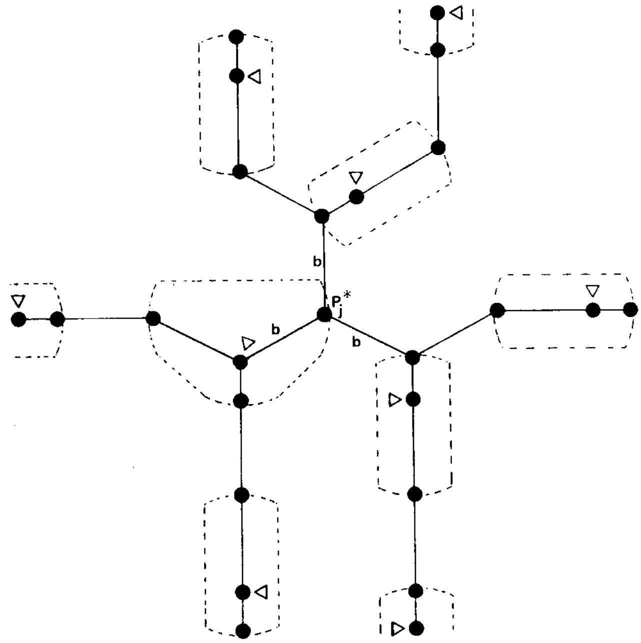

FIG. 13. *A clause configuration in the reduction for Euclidean p-median.*

valid for lower bounding in the present section. The same is true for the function h, i.e. lower bounding for a set S that contains a clause point (there are sets of four points for which $f(S)=h(4)=2b+1$; see Fig. 13).

The situation with junctions (shown in Fig. 14) is a little more delicate in the present section. Points P_k^i, P_{k+1}^i, P_l^j and P_{l+1}^j form a unit square. The distance between a point P_{k-1}^i ($k \equiv 0 \pmod 3$) and a corner of the square equals b. We now need to revise the definition of the lower-bounding function $e(k)$ of the preceding section. We define $e(k)$ to be the minimum of the optimal solutions of 1-median problems on sets S such that $|S|=k$ and S contains four junction points. Then $e(4)=2\sqrt{2}$.

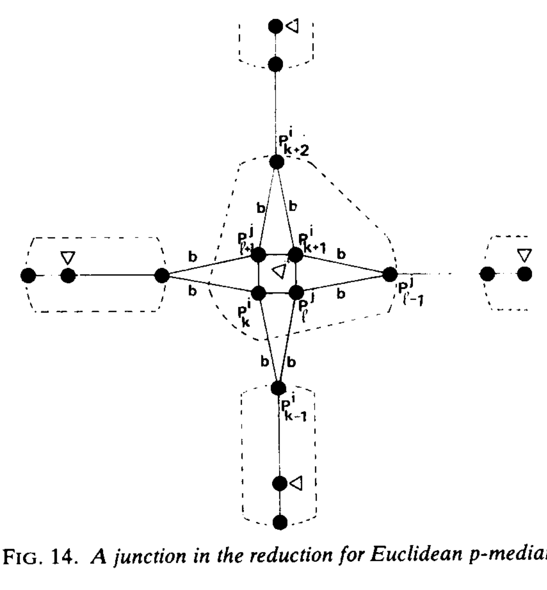

FIG. 14. *A junction in the reduction for Euclidean p-median.*

The problems corresponding to $e(5)$, $e(6)$ and $e(7)$ are shown in Fig. 15(a, b, c). It follows that, for b sufficiently large,

$$b+2\sqrt{2} \leqq e(5) \leqq b+2\sqrt{2}+0.5.$$

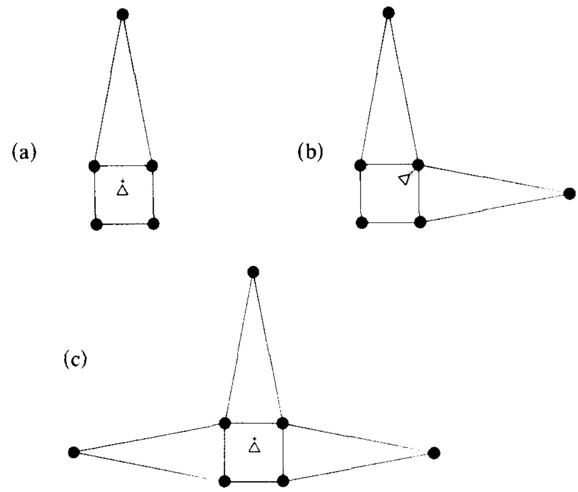

FIG. 15. *Covering points near a junction for Euclidean p-median.*

Furthermore, as $b \to \infty$,

$$e(b) \to 2b+\sqrt{2}+2$$

and

$$e(7) \approx e(5)+2b+1.$$

These facts are sufficient for deducing that an optimal solution to the optimization problem of minimizing

$$\sum_{j=1}^{m} h(s_j)+\sum_{j=m+1}^{m+J} e(s_j)+\sum_{j=m+J+1}^{p} g(s_j)$$

(subject to $\sum_{j=1}^{p} s_j = n$) is the same as in the preceding section, i.e., $s_j=4$ $(j=1,2,\cdots,m)$, $s_j=6$ $(j=m+1, m+2,\cdots,m+J)$, $s_j=3$ $(j=m+J+1, m+J+2,\cdots,p)$. The value of the optimal solution is asymptotically equal to

$$m(2b+1)+J(2b+\sqrt{2}+2)+(p-m-J)(b+1)=\left(\sum \frac{r_i}{3}\right)(b+1)+mb+J\sqrt{2}.$$

The optimal value is realizable in the p-median problem if and only if E is satisfiable.

REFERENCES

[1] J. ELZINGA AND D. W. HEARN, *Geometrical solutions for some minimax location problems*, Transportation Sci., 6 (1972), pp. 379–394.

[2] G. N. FREDERICKSON AND D. B. JOHNSON, *Optimal algorithms for generating quantile information in $X+Y$ and matrices with sorted columns*, in Proc. 13th Annual Conference on Information Science and Systems, Johns Hopkins Univ., Baltimore, MD, 1979, pp. 47–52.

[3] R. L. FRANCIS AND J. A. WHITE, *Facility Layout and Location*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[4] M. R. GAREY, R. L. GRAHAM AND D. S. JOHNSON, *Some* NP-*complete geometric problems*, in Proc. 8th ACM Symposium on the Theory of Computing, Association for Computing Machinery, New York, 1976, pp. 10–22.

[5] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of* NP-*completeness*, W. H. Freeman, San Francisco, 1979.

[6] O. KARIV AND S. L. HAKIMI, *An algorithmic approach to network location problems, Part* I: *The p-centers*, SIAM J. Appl. Math., 37 (1979), pp. 513–538.

[7] ———, *An algorithmic approach to network location problems, Part* II: *p-medians*, SIAM J. Appl. Math., 37 (1979), pp. 539–560.

[8] N. MEGIDDO, *On some planar location problems*, Dept. of Statistics, Tel Aviv Univ., Tel Aviv, October 1981.

[9] ———, *Linear-time algorithms for linear programming in* R^3 *and related problems*, this Journal, 12 (1983), pp. 759–776.

[10] N. MEGIDDO, A. TAMIR, E. ZEMEL AND R. CHANDRASEKARAN, *An* $O(n \log^2 n)$ *algorithm for the k-th longest path in a tree, with applications to location problems*, this Journal, 10 (1981), pp. 328–337.

[11] N. MEGIDDO, E. ZEMEL AND S. L. HAKIMI, *The maximum coverage location problem*, SIAM J. Alg. Discr. Meth., 4 (1983), pp. 253–261.

[12] Z. A. MELZAK, *A Companion to Concrete Mathematics*, Wiley Interscience, New York, 1973.

[13] K. P. K. NAIR AND R. CHANDRASEKARAN, *Optimal location of a single service center of certain types*, Naval Res. Logist. Quart., 18 (1971), pp. 503–510.

[14] C. H. PAPADIMITRIOU, *The Euclidean traveling salesman problem is* NP-*complete*, Theoret. Comput. Sci., 4 (1977), pp. 237–244.

[15] ———, *Worst-case and probabilistic analysis of a geometric location problem*, this Journal, 10 (1981), pp. 542–557.

[16] M. I. SHAMOS, *Computational Geometry*, Doctoral dissertation, Dept. of Computer Science, Yale Univ., New Haven, CT, 1978.

[17] M. I. SHAMOS AND D. HOEY, *Closest-point problems*, in Proc. 16th Annual IEEE Symposium on the Foundations of Computer Science, Institute of Electrical and Electronics Engineers, New York, 1975, pp. 151–162.

[18] K. J. SUPOWIT, *Topics in computational geometry*, Doctoral dissertation, Dept. of Computer Science, Univ. of Illinois, Urbana–Champaign, IL, April 1981.