ON THE COMPLEXITY OF LOCATING LINEAR FACILITIES IN THE PLANE *

Nimrod MEGIDDO ** and Arie TAMIR

Statistics Department, Tel Aviv University, Tel Aviv, Israel

Received July 1982
Revised October 1982

We consider the computational complexity of linear facility location problems in the plane, i.e., given n demand points, one wishes to find r lines so as to minimize a certain objective-function reflecting the need of the points to be close to the lines. It is shown that it is NP-hard to find r lines so as to minimize any isotone function of the distances between given points and their respective nearest lines. The proofs establish NP-hardness in the strong sense. The results also apply to the situation where the demand is represented by r lines and the facilities by n single points.

p-line center, p-line median, planar location, NP-complete, strongly NP-complete

Most of the literature on location theory focuses on models where both the facilities and the users are represented by points in the appropriate space, e.g., points in the Euclidean plane, vertices of a given graph, etc. In this paper we consider a planar case where the facilities are modeled as straight lines. For example, we may view the problems faced by a planner who has to locate r (linear) segments of a new railroad system so as to minimize the average cost to the users who have to reach the tracks from a number of different small communities. Thus, a straight line or a line segment is of natural importance in this context. Sometimes such problems are *easier* than those with point facilities. For example, it is much easier to find a line, so as to minimize the sum of distances to it from a set of given points, than to find a single point with the same objective (see [7, 8]). Thus, it is interesting to find out that our problems are NP-hard. Our proof establishes NP-hardness of several different problems (e.g., replace 'sum of distances' by 'maximum distance' and also by the point-line duality we obtain more problems: point facilities and linear demand sets). We establish a non-trivial reduction from the 3-satisfiability problem to prove strong NP-hardness (see [3]). The complexity of the problems with point facilities and demands was investigated in [2,5,6,9].

Formally, we consider the following problem of locating linear facilities. Given p points (x_1, y_1), $\dots$, (x_p, y_p) in the plane, find a set of r straight lines $L_1, \dots, L_r$ so as to minimize $\sum_{i=1}^{p} \mathrm{Min}_{1 \leqslant j \leqslant r}$ $d(x_i, y_i;\ L_j)$, where $d(x_i, y_i;\ L_j)$ denotes the distance between the point (x_i, y_i) and the line L_j, relative to a certain metric on the plane.

A weighted version of the problem when $r = 1$ was solved in [7] relative to the Euclidean metric in $O(n^2 \log n)$ time and relative to the rectilinear metric in $O(n \log^2 n)$ time.

A graphic version of the problem is obtained by replacing the points by the vertices of a given graph and the r lines by edges or paths in the graph. The graphic version is easily verified to be NP-hard. However, testing whether some set of r edges ensures a zero value for the objective-function (i.e., testing whether r edges suffice to cover all vertices) can be accomplished in polynomial-time (via matching). In this paper we prove that the planar analogue (i.e., the question whether r straight lines suffice for covering p given points) is strongly NP-hard. In particular, this proves that

* The first author was supported in part by the National Science Foundation under Grant ECS-8121741.

** Currently visiting at Carnegie-Mellon University, Graduate School of Industrial Administration, Pittsburgh, PA 15213, U.S.A.

the planar problem formulated above is NP-hard not only for that specific objective-function but also for any isotone function of the distances between points and lines, relative to any non-trivial metric. Using the duality between points and lines, our proof also yields the NP-hardness of location problems with point facilities and linear demand sets.

Henceforth, we will be dealing with the following problems:

1. Point Covering (PC). A set of points $(x_1, y_1),\ldots,(x_p, y_p)$ (x_i, y_i rationals, $i = 1,\ldots,p$) is given. Find a collection of straight lines $\{l_1,\ldots, l_r\}$ of minimum cardinality, such that (x_i, y_i) lies on at least one l_j.

2. Line Covering (LC). A set of straight lines $L_1,\ldots, L_r$ is given. Find a set of points $\{(x_1, y_1),\ldots,(x_p, y_p)\}$ of minimum cardinality such that each L_j contains at least one (x_i, y_i).

In view of the renowned duality between lines and points (see [4], for example) which is discussed later, the two problems are obviously very close in nature.

We first mention briefly the trivial cases. First, PC is trivial when no three points are colinear, in which case $\lceil p/2 \rceil$ lines are necessary to cover all points. Analogously, LC is trivial when every subset of three lines has an empty intersection and there are no parallel lines.

If there may be parallel lines, but still no intersection points of three lines or more, then an optimal solution for LC can be easily computed as follows. Partition the set of lines into classes $R_1,\ldots, R_s$ such that two lines are parallel if and only if they belong to the same class R_i. Let $r_i = |R_i|$, $i = 1,\ldots, s$, and assume $r_1 \geqslant r_2 \geqslant ,\ldots, r_s \geqslant 1$. Now, select an arbitrary line from R_1 and an arbitrary line from R_2. The point of intersection of two selected lines will belong to the final solution. Next, drop the two selected lines from R_1 and R_2, rename the classes so as to conform with the requirement $r_i \geqslant r_j$ for $i < j$ and continue in the same manner. We observe that this is in fact a particular case of a well-known scheduling problem, namely minimal-length scheduling of unit-execution-time tasks with tree-structured precedence constraints (see [1, p. 54]), which is solvable by the 'level strategy'. The embedding of our problem in the scheduling problem is by viewing each line as a task where members of the same R_i form a chain and the different chains are disjoint. The number of machines is two and the interpretation is that at each time unit at most two lines can be processed and this is feasible if they are not parallel. Furthermore, the value of the optimal solution is simply $\max(r_1, \lceil r/2 \rceil$. This is easily proved by induction, distinguishing between the case $r_1 > r_3$ (where both r_1 and $\lceil r/2 \rceil$ decrease by one after the first time unit) and the case $r_1 = r_3$ (where only $\lceil r/2 \rceil$ decreases but $\lceil r/2 \rceil > r_1$, so that $\max(r_1, \lceil r/2 \rceil)$ decreases in any case).

We now turn to the NP-hardness of the problems in the general case. First it is easily verified that both problems are in NP.

We now reduce 3-satisfiability to PC. Let $E_1 \wedge \cdots \wedge E_m$ be an instance of 3-satisfiability, where $E_j = x_j \vee y_j \vee z_j$, $\{x_j, y_j, z_j\} \subset \{v_1, \bar{v}_1,\ldots, v_n, \bar{v}_n\}$, $j = 1,\ldots, m$. Assume $|\{v_i, \bar{v}_i\} \cap \{x_j, y_j, z_j\}| \leqslant 1$. The general idea of the reduction is as follows. We shall construct a set of $m + nm^2$ points, m corresponding to the clauses $E_1,\ldots, E_m$ and m^2 ones corresponding to each pair of variables $(v_i, \bar{v}_i)$. Also, a set of $2nm$ lines will be constructed with the following properties (for an example, see Fig. 1):

1. Each clause E_j is represented by a point P_j.
2. Each pair of variables $(v_i, \bar{v}_i)$ is presented by a grid of m^2 points P^i_{kl} $(1 \leqslant k, l \leqslant m)$.
3. For each i $(i = 1,\ldots, n)$ and j $(j = 1,\ldots, m)$, the points $P^i_{1j},\ldots, P^i_{mj}$ lie on a straight line denoted by L_{ij} and the points $P^i_{j1},\ldots, P^i_{jm}$ be on a straight line denoted by $\bar{L}_{ij}$.
4. Except for the lines L_{ij}, $\bar{L}_{ij}$ $(i = 1,\ldots, n;\ j = 1,\ldots, m)$, no other straight line of the plane contains more than two points of the set
$$\{P^i_{kl}: i = 1,\ldots, n;\ k = 1,\ldots, m;\ l = 1,\ldots, m\}$$
$$\cup \{P_1,\ldots, P_m\}.$$
5. For every j $(j = 1,\ldots, m)$ the point P_j lies on the line L_{ik} if and only if $j = k$ and $v_i \in \{x_j, y_j, z_j\}$ and P_j lies on $\bar{L}_{ik}$ if and only if $j = k$ and $\bar{v}_i \in \{x_j, y_j, z_j\}$.

The above five properties establish the reduction by the following argument. The points of the form P^i_{kl} cannot be covered by less than nm lines, since no straight line contains more than m of them and altogether they number nm^2. Moreover,

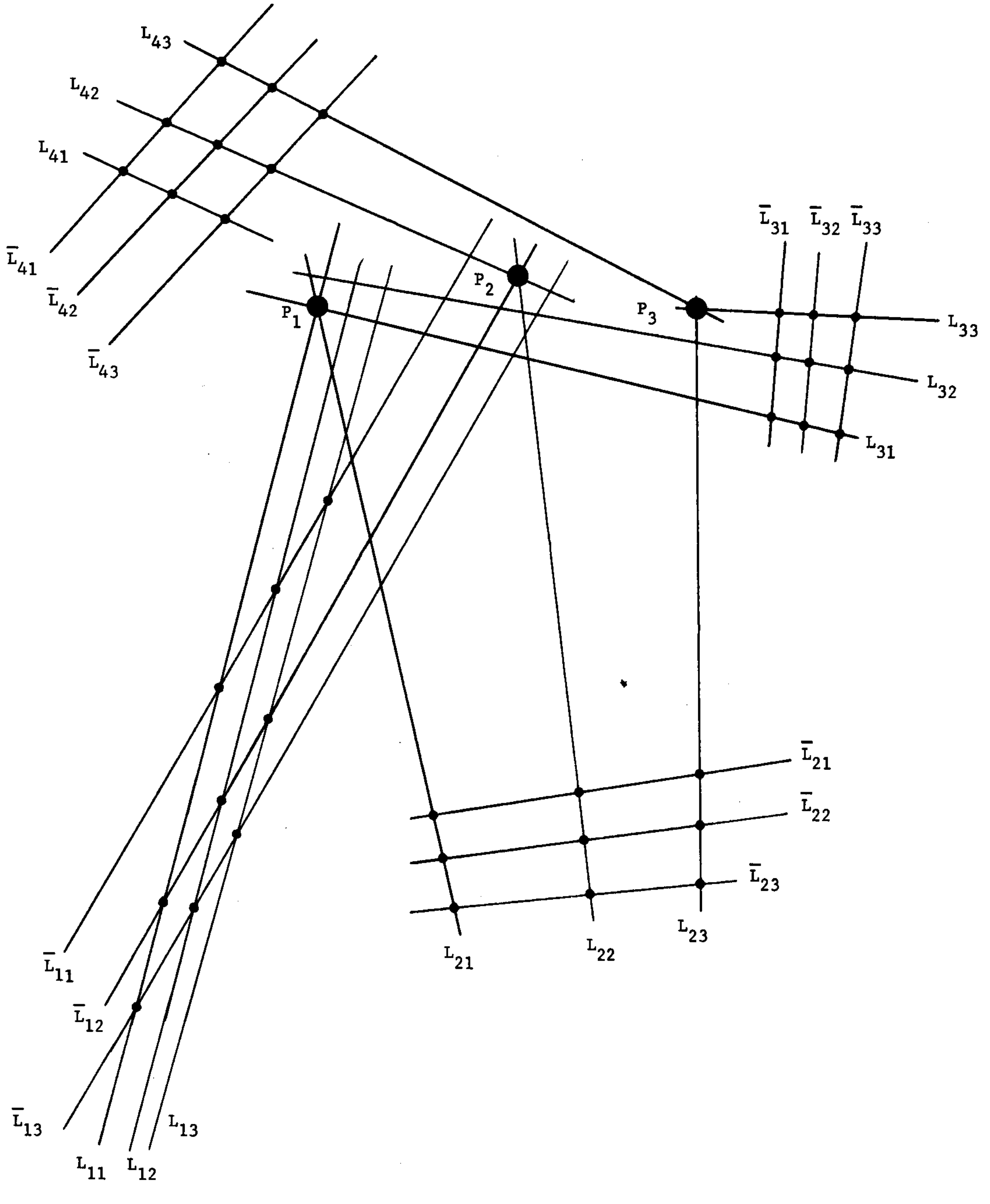

Fig. 1. Example: $E_1 = v_1 \vee v_2 \vee v_3$, $E_2 = \bar{v}_1 \vee v_2 \vee v_4$, $E_3 = v_2 \vee v_3 \vee v_4$.

to achieve that number, for every i ($i = 1,\dots,n$), the points P^i_{kl} ($1 \leqslant k,\ l \leqslant m$) must be covered either by the lines L_{ij} ($j = 1,\dots,m$) or by the lines $\bar{L}_{ij}$ ($j = 1,\dots,m$). No other collection of m lines can cover the collection of m^2 points P^i_{kl} ($1 \leqslant k,\ l \leqslant m$) (assuming $m > 2$). We claim that $E_1 \wedge \cdots \wedge E_m$ is satisfiable if and only if the entire collection of points $\{P_1,\dots,P_m\} \cup \{P^i_{kl}\colon i = 1,\dots,n;\ k = 1,\dots,m;\ l = 1,\dots,m\}$ can be covered by nm lines. For, the choice between $\{L_{ij}\}_{j=1}^m$ and $\{\bar{L}_{ij}\}_{j=1}^m$ for a given i simply corresponds to the assignment of a truth-value to $(v_i, \bar{v}_i)$. Specifically, for each i, v_i is true if and only if $\{L_{ij}\}$ is chosen to cover the m^2 points P^i_{kl}.

Finally, we have to discuss the actual construction of the points P_j and P^i_{kl}. We will construct points with rational coordinates, maintaining the numerators and the denominators separately. The numerical values of all the numerators and denominators will be bounded by a polynomial in m and n. First, let $P_j = (j, j^2), j = 1,\dots, m$. Thus, no three of the points $P_1,\dots, P_m$ are colinear. The construction of the points P^i_{kl} will be carried out with the aid of the lines L_{ij}, $\bar{L}_{ij}$ as follows. For each i $(i = 1,\dots, n)$, P^i_{kl} is the point of intersection of L_{ik} with $\bar{L}_{il}$. The lines L_{ij}, $\bar{L}_{ij}$ are successively constructed in the order $L_{11},\dots, L_{1m}, \bar{L}_{11},\dots, \bar{L}_{1m}, L_{21},\dots, L_{2m}, \bar{L}_{21},\dots$ so as to satisfy properties 3, 4, 5. When a specific line L_{ij} has to be constructed the following conditions should be satisfied:

(i) L_{ij} should contain P_i if and only if $v_i \in \{x_j, y_j, z_j\}$;

(ii) L_{ij} should not contain any previously constructed point of the form P_k (except possibly for P_j as explained before) or P^i_{kl};

(iii) L_{ij} should not coincide with any previously constructed line.

When a specific line $\bar{L}_{ij}$ has to be constructed the following conditions should be satisfied:

(i) $\bar{L}_{ij}$ should contain P_j if and only if $\bar{v}_i \in \{x_j, y_j, z_j\}$;

(ii) $\bar{L}_{ij}$ should not contain any previously constructed point (except possibly for P_j);

(iii) $\bar{L}_{ij}$ should not contain a point of intersection of two lines of the form L_{ik}, L_{il} (in order for the two points P^i_{kj}, P^i_{lj} to be distinct);

(iv) $\bar{L}_{ij}$ should not contain a point of intersection between some L_{ik} and another line which contains at least two previously constructed points (in order to satisfy condition 4; the intersection $\bar{L}_{ij}$ with L_{ik} becomes the point P^i_{kj});

(v) $\bar{L}_{ij}$ should not be parallel to any L_{ik}, in order to ensure the existence of the point P^i_{kj}.

Thus, a typical step is that a line has to be constructed so as to (possibly) contain one specified point of the P_j's and not any other point from a finite collection of 'forbidden' points, and also so as not to parallel any one of a finitely numbered lines. Suppose that we always construct the line whose slope is the integer closest to zero among the feasible slopes. The number of 'forbidden' slopes is obviously bounded by some polynomial in m and n and hence the slope of every constructed line is an integer whose absolute value is bounded by that polynomial. If the constructed line also crosses through one of the P_j's (whose coordinates are of the form (j, j^2)) then the coefficients of its equation will be polynomially bounded integers. Similarly, if the line should not cross through any P_j, then we may construct it so as to cross through an integer point, which is not forbideen, whose distance from the origin is minimal. It follows that the coordinates of such a point are polynomially bounded and hence all our constructed lines will have polynomially bounded integers as their coefficients. This implies that all the points P^i_{kl} will have coordinates which are rationals with polynomially bounded numerators and denominators. This establishes that PC is *strongly* NP-hard (see [3]).

To establish that LC is strongly NP-hard we reduce PC to LC by using the point-line duality argument. Specifically, given the points $(a_1, b_1),\dots,(a_n, b_n)$, we first find a translation $(a_i, b_i) = (a_i + a, b_i + b)$ that will assure that no two points are colinear with the origin. Next, we represent the point (a_i, b_i) by a line $a_i x + b_i y + 1 = 0$. Thus, two lines corresponding to two distinct points are not parallel and the main property is that points are colinear if and only if their corresponding lines all interesect at a single point.

References

[1] E.G. Coffman, Jr., *Computer and Job-Shop Scheduling Theory,* Wiley, New York (1976).

[2] R.J. Fowler, N.S. Paterson and S.L. Tanimoto, "Optimal packing and covering in the plane are NP-complete", *Information Processing Lett.* **12**, 133–137 (1981).

[3] M.G. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness,* Freeman, San Francisco (1979).

[4] W.C. Graustein, *Introduction to Higher Geometry,* Macmillan, New York (1946).

[5] S. Masuyama, T. Ibaraki and T. Hasegawa, "The computational complexity of the *m*-center problems on the plane", *Trans. IECE Japan* **E64**, 57–64 (1981).

[6] N. Megiddo and K.J. Supowit, "On the complexity of some common geometric location problems", manuscript (1982).

[7] N. Megiddo and A. Tamir, "Finding least-distances lines", *SIAM J. Alg. Disc. Meth.,* to appear.

[8] Z.A. Melzac, *Companion to Concrete Mathematics,* Wiley, New York (1973).

[9] C.H. Papadimitriou, "Worst-case and probabilistic analysis of a geometric location problem", *SIAM J. Comput.* **10**, 542–557 (1981).