RAIRO. INFORMATIQUE THÉORIQUE

A. MARCHETTI-SPACCAMELA

M. TALAMO

Probabilistic analysis of two euclidean location problems

RAIRO. Informatique théorique, tome 17, nº 4 (1983), p. 387-395

<http://www.numdam.org/item?id=ITA_1983__17_4_387_0>

R.A.I.R.O. Informatique théorique/Theoretical Informatics
(vol. 17, n° 4, 1983, p. 387 à 395)

PROBABILISTIC ANALYSIS OF TWO EUCLIDEAN LOCATION PROBLEMS (*)

by A. MARCHETTI-SPACCAMELA, M. TALAMO (1)

Communicated by G. AUSIELLO

Abstract. — *In this paper two location problems in the plane are considered. We give two polynomial time algorithms that produce arbitrarily good solutions with probability* 1.

Résumé. — *Dans ce papier on a considéré deux problèmes sur le plan. On montre deux algorithmes caractérisés par complexité de temps polynômial qui produisent des solutions arbitrairement proches de la solution optimale avec probabilité* 1.

1. INTRODUCTION

Given n points in the plane we consider the problem of locating k centers so that the distance from each given point to its closest center is minimized. These location problems are often encountred in problems of emergency facility services, for example ambulance services, police stations [9]. More precisely we consider the following problems:

Problem P

Given a set of points $x = \{x_1, x_2, \ldots, x_n\}$ belonging to the unit square U in the plane and a number k, we wish to find a set C of centers belonging to U such that:

(i) cardinality of C is equal to k;

(ii) the following objective function is minimized:

$$z(P) = \max_i \{\min_j \operatorname{dist}(x_i, c_j)\ x_i \in X, c_j \in C\}.$$

(*) Received on May 1982, revised on September 1982.

(1) Dipartimento di Informatica e Sistemistica, Università di Roma, Via Eudossiana 18, 00184 Roma, Italy.

R.A.I.R.O. Informatique théorique/Theoretical Informatics, 0399-0540/1983/387/$ 5.00

problem Q

Given a set of points $X = \{x_1, x_2, \ldots, x_n\}$ belonging to the unit square U in the plane and a number k, we wish to find a set D of centers belonging to X such that:

(i) cardinality of D is equal to k;

(ii) the following objective function is minimized:

$$z(Q) = \max_i \{\min_j \operatorname{dist}(x_i, d_j)\, x_i \in X,\ d_j \in D\}.$$

The difference between the two problems lies in the fact that the set of centers in problem Q must be a subset of the given set of points X. Hence the optimal solution of problem Q is greater or equal to the optimal solution of Problem P.

As the associated decision problems are NP-complete ([5] and [8]) it follows that also problems P and Q are computationally intractable. Furthermore, the reductions used to prove the NP-hardness can be used also to show that these problems are not ε-approximable (*see* [4] for definition of ε-approximability).

These negative results suggest the idea of a different approach. In particular we use a probabilistic approach that has been used successfully for other NP-complete geometric problems; we assume that the points are scattered at random in the unit square and we study the probabilistic behaviour of polynomial time algorithms.

A first step in this direction has been performed by Frieze [3]; he considered the case when the number of centers is less than $p \log n$, for some p greater than zero, and the points are uniformly and independently distributed in the region. For both problems P and Q he exhibits a family of algorithms with the following properties:

for every $\varepsilon > 0$ there is an algorithm $A(\varepsilon)$ such that:

(i) $A(\varepsilon)$ runs in time $0(n^{1/\varepsilon})$;

(ii) $A(\varepsilon)$ finds an ε-approximate solution in probability.

Note that, when $k < p \log n$, then both problems P and Q are solvable in $O(n^{\log n})$ by complete enumeration (*see* [7]); this implies that, most probably, the problems are not NP-complete, since they are solvable by a subexponential algorithm.

In this paper we study the case when the number of centers is $k = O(n^a)$, for some $1 > a > 0$. We propose two simple heuristics based on partitioning the unit square into hexagons and we show that both algorithms are asimptotically optimal with probability one (which implies convergence in probability, *see* [1]); moreover the time complexity of the algorithms is $O(n \log n)$.

2. PROBABILISTIC ANALYSIS OF PROBLEM P

In order to tackle Problem P we first consider a continuum version P' of P.

Problem P'

Given a number k, we wish to partition the unit square U in k regions $V_1, V_2, \ldots, V_k$, with centers $c'_1, c'_2, \ldots, c'_k$ such that the following objective function is minimized:

$$z(P') = \max_i \{ \max_{y \in V_i} (\text{dist}(y, c'_i))\, y \in V_i \}.$$

The difference between problems P and P' lies in the fact that we must completely cover U. This implies that if $z^*(P)$ and $z^*(P')$ are the optimal solutions of P and P', respectively, then we have:

$$z^*(P') \geqq z^*(P).$$

It is easy to see that the optimal solution of Problem P' divides the unit square in convex regions. Let be pol the regular polygon of unit area with n sides and $w(n)$ the square of the radius of the circle that circumscribes pol.

We have that:

$$w(n) = (2/(n \text{ sen } 2\pi/n)).$$

For the proof of Theorem 1 we will use the following lemma that is an application of the Euler's formula on planar graphs (*see* [2]).

LEMMA 1: *Let* $\{R_1, R_2, R_3, \ldots, R_k\}$ *be a partition of the unit square in* k *convex polygons with* $n_1, n_2, \ldots, n_k$ *sides, respectively. Then we have* $\sum_{j=1}^{k} n_i \leqq 6k-2$.

THEOREM 1: *If* $k = n^a$ *then, for every* $0 < a < 1$, *the optimal solution* $z^*(P')$ *of* P' *satisfies, for every* $\varepsilon > 0$:

$$(1+\varepsilon)\, z(P') \geqq [w(6)/k]^{1/2}.$$

Proof: Let be given a polygon R_j with n_j sides and area A_j; then the radius of the minimum circle that includes R_j is at least:

$$(w(n_j)\, A)^{1/2}.$$

Now suppose that there is a partition $\{R_1, R_2, \ldots, R_k\}$ of the unit square and an $\varepsilon > 0$ such that:

$$(1+\varepsilon) \max_j \{ \max_{y \in R_j} (\mathrm{dist}(y, c_j)), c_j \text{ is the center of region } R_j \} \leqq [w(6)/k]^{1/2}.$$

This implies that for every $j = 1, 2, \ldots, k$

$$(1+\varepsilon)\ [w(n_j) A_j]^{1/2} \leqq [w(6)/k]^{1/2}$$

$$(1+\varepsilon)^2 A_j \leqq w(6)/kw(n_j).$$

As the total area is 1 we have that:

$$(1+\varepsilon)^2 \sum_{j=1}^{k} A_j = 1+\varepsilon' \leqq \sum_{j=1}^{k} w(6)/kw(n_j) = w(6) \sum_{j=1}^{k} 1/kw(n_j).$$

Since $1/w(n_j)$ is concave we have:

$$\sum_{j=1}^{k} 1/kw(n_j) \leqq 1/w\left(\sum_{j=1}^{k} n_j/k\right)$$

and since $1/w(n_j)$ is increasing, applying Lemma 2 we obtain:

$$1+\varepsilon' \leqq \sum_{j=1}^{k} w(6)/kw(n_j)$$

$$\leqq w(6)\left(1/w\left(\sum_{j=1}^{n} n_j/k\right)\right)$$

$$\leqq w(6)\ (1/w(6)) \leqq 1.$$

This contradiction proves the theorem.

Q.E.D.

Theorem 1 says that if we ignore the effect of the boundaries of U then the optimal solution for Problem P' is obtained by dividing U in regular hexagons; furthermore it says that the error introduced is negligible.

Now we turn to Problem P and we show that, with high probability, the difference between the optimal solutions of Problem P' and Problem P is asymptotically negligible. Let us divide the unit square U in h, small squares $u_1, u_2, \ldots, u_h$ each one with side $1/\sqrt{h}$ and area $1/h$, where $h = n^d$, for some $1 > d > 0$.

LEMMA 2: *Each square u contains at least one point of X with probability* 1.

Proof: Let us define the following events $E_j = \{$ in region j there is no point $\}$.
$E = \{$ there is an index m such that E_m *occurs* $\}$.

Since each point x is uniformly and independently distributed in U we have:

Prob $\{x_i \notin u_j\} = 1 - 1/h$ for every $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, h$

Prob $\{E_j\} =$ Prob $\{x_i \notin u_j, \forall i\} = (1 - 1/h)^n$ for every $j = 1, 2, \ldots, h$

$$\text{Prob}\{E\} = \text{Prob}\left\{\bigcup_{j=1}^{h} E_j\right\} \leqq \sum_{j=1}^{h} \text{Prob}\{E_j\} = h(1-1/h)^n.$$

Now, since $(1 - 1/h) \leqq e^{-1/h}$, we have, for n sufficiently large:

Prob $\{E\} < he^{-n/h} \leqq O(n^{-2})$,

the lemma follows applying the Borel-Cantelli lemma (*see* [1]). Q.E.D.

THEOREM 2: *If* $k = n^a$, $0 < a < 1$, *then, for every* $\varepsilon > 0$, $(1+\varepsilon) z^*(P) \geqq z^*(P')$ *with probability* 1.

Proof: Lemma 2 implies that, with probability 1, the difference between the optimal solution of Problem P' and Problem P is less than $\sqrt{2/h}$, the diagonal of the small square u_j. Hence applying Theorem 1 we have, for every $\varepsilon > 0$; and $d > a$:

$$(1+\varepsilon) z^*(P) \geqq z^*(P) + \sqrt{2/h} = z(P) + \sqrt{2}\, n^{-d/2} \geqq z^*(P')$$

with probability 1. Q.E.D.

Now we present an algorithm for the solution of Problem P, that divides the unit square in k regular hexagons and assigns a center to each hexagon. Note that in Problem P we do not need to cover the unit square completely as it has been done in Problem P', but we will show that the error introduced is asimptotically neglegibile.

Algorithm 1

input: k;

Sol $= \emptyset$;

Partition the plane in regular hexagons with area:

$$b = (1 + 12\sqrt{3}(1 + \sqrt{1+k}/6\sqrt{3})/k)/k;$$

for each hexagon with center c do

  if c is contained in U then:

    Sol: $=$ Sol $\cup \{c\}$

  else

    begin

find a point p in U such that
1: p belongs to U
and
2: there is no point of $h \in U$
such that $\text{dist}(x, p) \geqq (w(6)\, b)^{1/2}$;
Sol: = Sol $\cup \{p\}$
end;
while $|\text{Sol}| \leqq k$ do begin
choose a point p in U at random;
Sol: = Sol $\cup \{p\}$;
end;
output: Sol.

It is easy to see that the time complexity of the algorithm is $O(n \log n)$. Let $\hat{z}(P)$ be the value of the solution given by algorithm 1.

THEOREM 3: *For every* $\varepsilon > 0$ *we have that:*

$$\hat{z}(P) \leqq (1+\varepsilon)(w(6)/k)^{1/2}.$$

Proof: In order to show the correctness of the algorithm it is sufficient to prove that the number of centers used is, at most, k. As the side of the hexagon is $(bw(6))^{1/2}$ the number of hexagons on each side of the unit square that is not completely contained in U is, at most:

$$1/(bw(6)\ )^{1/2}.$$

On the other side the number of hexagons completely contained in U is, at most:

$$1/b.$$

Hence the total, number of hexagons used during step 2 is less than:

$$(1/b) + 4/(bw(6))^{1/2} \leqq k.$$

Now we evaluate the solution given by the algorithm: $z(P)$ is equal to the side of the regular hexagon with area b that is:

$$\hat{z}(P) = (w(6)\, b)^{1/2}.$$

By substituting the value of b we obtain for k sufficiently large:

$$\hat{z}(P) = (w(6)\, b)^{1/2} = (w(6)\ (1 + 12\sqrt{3}(1 + \sqrt{1+k}/6\sqrt{3})/k)/k)^{1/2}$$
$$\leqq (w(6)/k)^{1/2}(1+\varepsilon) \quad \text{for every } \varepsilon > 0.$$

Q.E.D.

COROLLARY 1: *Algorithm* 1 *finds a solution to problem P such that for every* $\varepsilon>0$:

$(1+\varepsilon) z^*(P) \geqq \hat{z}(P)$ *with probability* 1.

Proof: Given ε, $0<\varepsilon<1$, we have:

$$(1+\varepsilon/7) z^*(P) \geqq z^*(P') \quad \text{from theorem 2:}$$

$$(1+\varepsilon/7) z^*(P') \geqq (w(6)/k)^{1/2} \quad \text{from theorem 1}$$

$$(1+\varepsilon/7)(w(6)/k)^{1/2} \geqq \hat{z}(P) \quad \text{from theorem 3}$$

from the last three inequalities we obtain:

$$(1+\varepsilon) z^*(P) \geqq (1+\varepsilon/7)^3 z^*(P) \geqq \hat{z}(P).$$

Q.E.D.

The proof of the following corollary is trivial and is omitted.

COROLLARY 2: *For every* $\varepsilon>0$, *the optimal solution of P satisfies:*

$$(1+\varepsilon)[w(6)/k]^{1/2} \geqq z^*(P) \geqq (1-\varepsilon)[w(6)/k]^{1/2} \quad \textit{with probability } 1.$$

3. PROBABILISTIC ANALYSIS OF PROBLEM Q

In order to obtain a lower bound on problem Q it is sufficient to observe that the optimal solution of problem Q is greater or equal to the optimal solution of Problem P. Hence we obtain from Corollary 2 the following lemma.

LEMMA 3: *For every* $\varepsilon>0$, *the optimal solution of problem Q satisfies:*

$$z^*(Q) \geqq (1-\varepsilon)[w(6)/k] \quad \textit{with probability } 1.$$

In order to obtain an upper bound on the optimal solution of problem Q we consider the following algorithm.

Algorithm 2

```
input: X, k;
Apply algorithm 1 with input k obtaining Sol
Y:=∅;
for each i∈Sol do
     begin
     find y∈X such that the dist(y, i) is minimized;
     Y:=Y∪{y}
     end;
output: Y.
```

It is easy to see that algorithm 2 runs in $O(n \log n)$, in the worst case. Now let $\hat{z}(Q)$ be the value of the solution given by the algorithm.

THEOREM 4: *For every* $\varepsilon > 0$, *we have that:*

$$\hat{z}(Q) \leqq (1+\varepsilon)(w(6)/k)^{1/2}, \text{ with probability } 1.$$

Proof: By Lemma 2 we have:

$$\hat{z}(Q) \leqq \hat{z}(P) + \sqrt{2/h} \leqq (1+\varepsilon)\,\hat{z}(P),$$

where $h = n^d$, $0 < d < 1$.

The thesis follows applying Theorem 3.

Q.E.D.

The proofs of the following corollaries are trivial and are omitted.

COROLLARY 3: *Algorithm* 2 *finds a solution to problem* Q *such that for every* $\varepsilon > 0$:

$$(1+\varepsilon)\,\hat{z}(Q) > z^*(Q) \quad \textit{with probability } 1.$$

COROLLARY 4: *For every* $\varepsilon > 0$, *the optimal solution of problem* Q *satisfies:*

$$(1+\varepsilon)\,[w(6)/k] > z^*(Q) > (1-\varepsilon)\,[w(6)/k] \quad \textit{with probability } 1.$$

4. CONCLUSIONS

The technique of partitioning the plane into hexagons for designing efficient aproximate algorithms in geometric problems was first introduced by Papadimitriou [6]. He has considered the k-median problem that is the problem to choose k-centers in order to minimize the sum of the distances from each point to its closet center. He has shown that a simple heuristic that partitions the plane into hexagons and assigns a center to each hexagon constructs a solution that has relative error smaller than $\varepsilon > 0$ in probability.

Exploiting the same technique but achieving stronger results, in this paper we have examined the behaviour of two aproximate algorithms, that with probability 1 and relative error smaller than ε, for every $\varepsilon > 0$, find the solution of two different location problems.

REFERENCES

1. V. L. CHUNG, *A Course in Probability Theory*, Academic Press, 1974.
2. S. EVEN, *Graph Algorithms*, Computer Science Press, 1979.

3. A. M. FRIEZE, *Probabilistic Analysis of Some Euclidean Clustering Problems*, Disc. Appl. Mathem., 2, 1980.
4. M. S. GAREY and D. S. JOHNSON, *Computers and Intractability*, W. H. Freeman and Company, 1979.
5. A. MARCHETTI-SPACCAMELA, *The* P-*Center Problem in the Plane is* NP-*Complete*, Proc. 19-th Allerton Conference on Communication, Control and Computing, 1981.
6. C. H. PAPADIMITRIOU, *Worst-Case and Probabilistic Analysis of a Geometric Location Problem*, S.I.A.M. J. on Computing, Vol. 10, No. 3, 1981.
7. M. J. SHAMOS, *Computational Geometry*, Doct. Th., Yale Univ., New-Haven, 1978.
8. K. SUPOWIT, *Topics in Computational Geometry*, Doct. Th., Univ. of Illinois at Urbana Champaign, 1981.
9. C. TOREGAS, L. BERGMAN, R. REVELLE and R. SWAIN, *The location of Emergency Service Facilities*, Oper. Res., Vol. 19, No. 6, 1971.