A SIMPLE HEURISTIC FOR THE p-CENTRE PROBLEM

M.E. DYER

Department of Mathematics and Statistics, Teesside Polytechnic, Middlesbrough, Cleveland TS1 3BA, UK

A.M. FRIEZE

Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, USA and Department of Computer Science and Statistics, Queen Mary College, London, UK

Received August 1984
Revised November 1984

We describe a simple heuristic for determining the p-centre of a finite set of weighted points in an arbitrary metric space. The computational effort is $O(np)$ for an n-point set. We show that the ratio of the objective function value of the heuristic solution to that of the optimum is bounded by $\min(3, 1+\alpha)$, where α is the maximum weight divided by the minimum weight of points in the set.

p-centre problem * heuristic solution

1. Introduction

Location problems of the type we consider here have been discussed by various authors. See, for example, [1,7]. To define the problem formally, let X be any (complete) metric space, with metric $d(\cdot, \cdot)$. Let V be any subset of X of (finite) size n, with positive weights $w(v)$ for each $v \in V$. Then the p-centre problem may be cast as:

$$\begin{aligned} &\text{Determine} && c_1, c_2, \ldots, c_p \quad \text{to} \\ &\text{minimize} && F(x_1, x_2, \ldots, x_p), \\ &\text{where} && F(x_1, x_2, \ldots, x_p) \\ &&& = \max_{v \in V} \min_{1 \leqslant i \leqslant p} w(v) d(v, x_i). \end{aligned} \tag{1.1}$$

We will write $F_0 = F(c_1, c_2, \ldots, c_p)$.

The most usual setting for this type of problem is in a finite-dimensional Euclidean space. In particular the planar case is of practical interest. However, it has been shown that the problem (1.1) is NP-hard even in the planar Euclidean case [3,8,9]. Thus heuristics are more likely to be of practical use in the general case. In Section 2 we describe an extremely simple heuristic, and prove that it guarantees $\min(3, 1+\alpha)$ times the minimum value of F in the worst case, where α is the maximum ratio between the weights of points in V. Note that in the case where all the weights $w(v)$ are equal, which is often called the *unweighted* case, this reduces to only twice the optimum, since then $\alpha = 1$. This heuristic requires only $O(np)$ distance evaluations, arithmetic operations and comparisons. The idea behind the heuristic is fairly intuitive. We select any point of largest weight for the first centre. Then we successively choose new centres so that the next centre chosen is the point which has the largest weighted distance from its nearest centre. We repeat this until we have the required number of centres. This has the useful property that the $(p+1)$-centre solution is a superset of the p-centre solution.

2. A p-centre heuristic

Consider the following heuristic for generating p points $v_1, v_2, \ldots, v_p$ of V.

Heuristic H_p

(0) Choose $v_1 \in V$ so that $w(v_1) =$

$\max_{v \in V} w(v)$. (If there are ties for the maximum, choose arbitrarily amongst these.) Set $D(v) \leftarrow w(v)d(v, v_1)$ for each $v \in V$.

(1) While $i < p$ do
   Determine v_{i+1} by
   $D(v_{i+1}) = \max_{v \in V} D(v)$.
   Set $D(v) \leftarrow \min\{D(v), w(v)d(v, v_{i+1})\}$ for all $v \in V$.

It is easy to verify that H_p requires $O(np)$ distance evaluations, arithmetic operations and comparisons. We take $v_1, v_2, \ldots, v_p$ as the p centres. Note that all $v_i \in V$. The quality of the heuristic solution is then measured by $F_{H_p} = F(v_1, v_2, \ldots, v_p)$. Let v_{p+1} denote the point chosen in a hypothetical pth repetition of the while-loop in H_p. (Thus $v_1, v_2, \ldots, v_{p+1}$ are the points which would be selected by H_{p+1}.) Let $v_{r(i)}$ be the point such that, when v_i is chosen during the $(i-1)$th repetition of the loop, $D(v_i) = w(v_i)d(v_i, v_{r(i)})$. Then it follows easily that these $D(v_i)$ values form a nonincreasing sequence, and that $F_{H_p} = D(v_{p+1})$. Let

$$\alpha = \max_{v \in V} w(v) / \min_{v \in V} w(v)$$

and

$$\delta = \min(3, 1 + \alpha).$$

We may now prove the following results relating the values of the heuristic and optimal solutions.

Theorem 1. $F_{H_p} \leqslant \delta F_0$.

Proof. The optimal centres are $c_1, c_2, \ldots, c_p$. Let

$$V_i = \left\{ v \colon d(v, c_i) = \min_j d(v, c_j) \right\}.$$

Clearly V_i partitions V and, for all $v \in V_i$, $w(v)d(v, c_i) \leqslant F_0$.

Now, by the pigeonhole principle, some V_i includes at least two of $v_1, v_2, \ldots, v_p + 1$. Thus suppose that V_k contains v_i and v_j, with $1 \leqslant i < j \leqslant p + 1$. Let $\beta = w(v_j)/w(v_i)$. Note that $\beta \leqslant \alpha$. Consider the following two cases.

(a) $\beta \leqslant 2$. Thus $\beta \leqslant \min(\alpha, 2)$, and hence $1 + \beta \leqslant \delta$. Note also that this case includes $i = 1$, since then, by choice of v_1, $\beta \leqslant 1$. Then, by the above observations,

$$F_{H_p} \leqslant w(v_j)d(v_j, v_{r(j)})$$
$$\leqslant w(v_j)d(v_j, v_i)$$
by the minimality of $v_{r(j)}$, and $i < j$,
$$\leqslant w(v_j)\{d(v_j, c_k) + d(v_i, c_k)\}$$
by the triangle inequality,
$$= w(v_j)d(v_j, c_k) + \beta w(v_i)d(v_i, c_k)$$
$$\leqslant (1 + \beta)F_0$$
$$\leqslant \delta F_0.$$

(b) $\beta > 2$. Then, clearly, $\alpha > 2$, and hence $\delta = 3$. Observe that we have $\delta > 2(1 + \beta^{-1})$. We may also assume $i > 1$. Let v_l $(l < i)$ be the closest point to v_j at the stage when v_i is chosen. Then

$$F_{H_p} \leqslant w(v_j)d(v_j, v_{r(j)})$$
$$\leqslant w(v_j)d(v_j, v_l)$$
by the minimality of $v_{r(j)}$,
$$\leqslant w(v_i)d(v_i, v_{r(i)})$$
otherwise v_j would have been chosen rather than v_i,
$$\leqslant w(v_i)d(v_i, v_l)$$
by the minimality of $v_{r(i)}$,
$$\leqslant w(v_i)\{d(v_i, v_j) + d(v_j, v_l)\}$$
by the triangle inequality.

Now suppose that $d(v_i, v_j) \leqslant d(v_j, v_l)$, then the above inequalities would yield $w(v_j)d(v_j, v_l) \leqslant 2w(v_i)d(v_j, v_l)$. Thus we would have $w(v_j) \leqslant 2w(v_i)$, which contradicts $\beta > 2$. Thus we must, in fact, have $d(v_i, v_j) > d(v_j, v_l)$, and hence

$$F_{H_p} < 2w(v_i)d(v_i, v_j)$$
$$\leqslant 2w(v_j)\{d(v_i, c_k) + d(v_j, c_k)\}$$
by the triangle inequality,
$$= 2\{w(v_i)d(v_i, c_k) + \beta^{-1}w(v_j)d(v_j, c_k)\}$$
$$\leqslant 2(1 + \beta^{-1})F_0$$
$$< \delta F_0.$$

This completes the proof.

To show that the above bound is tight, consider the four points on the line $\mathbb{R}$, with the usual distance measure, shown in Figure 1. Here we have $1 \leqslant w \leqslant 2$. It may be verified that we can

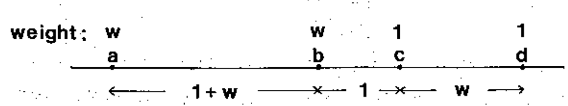

Fig. 1.

have, in the case $p = 2$,

$$F_{H_p} \geqslant F(a, d) = w(1 + w)$$

and $F_0 \leqslant F(a, c) = w$. Thus $F_{H_p}/F_0 \geqslant 1 + w = \delta$, where δ is as defined above. The replication of this configuration, with a large separation between each replication, can be used to show that the bounds are tight for all p.

3. Comments and conclusions

We have shown that the heuristic H_p gives solutions which are never more than δ times the optimum. This bound is obviously tightest when $\delta = 2$, i.e. $\alpha = 1$, the unweighted case. Now it is known [4,6] that it is NP-hard to produce solutions to the general unweighted p-centre problem which are within τ times the optimum for any $\tau < 2$. Thus, in a certain sense, the heuristic H_p is best possible. A 2 times optimal heuristic for the p-centre problem on graphs with the triangle inequality on their edge distances has been given by Hochbaum and Shmoys [4], for the vertex unweighted case. This is a special case of our unweighted problem. Their heuristic, which differs from ours, runs in time $O(|E| \log|E|)$ on a complete graph, which is $O(n^2 \log n)$ time on an n-vertex graph. Our heuristic tunes in time $O(np) = O(n^2)$ time in this case since, in adjacency matrix form, distance calculations on the graph can be done in constant time (in view of the triangle inequality). Actually Hochbaum and Shmoys use an adjacency list form, but an adjacency matrix can be prepared from such an input in $O(n^2)$ time. Hochbaum and Shmoys have also generalised their methods to certain types of weighted problems on graphs in [5].

The quality of solutions given by our heuristic can be improved if we can optimally solve, say, the 1-centre problem in X. For each v_i we determine the set N_i of its closest points, i.e. $v \in N_i$ if $d(v, v_i) = \min_j d(v, v_j)$. If we now solve the 1-centre problem on N_i to give a centre h_i, say, then clearly

$$\max_{v \in N_i} w(v) d(v, h_i) \leqslant \max_{v \in N_i} w(v) d(v, v_i)$$

and strict inequality is possible. Thus $F(h_1, h_2, \ldots, h_p)$ may be smaller than F_{H_p} (and is certainly no larger). This improvement appears particularly attractive in the Euclidean space $\mathbb{R}^d$, where it is known that, for fixed d, there is an $O(n)$ time algorithm for the 1-centre problem. (See [2].) Similarly, if we could efficiently solve the k-centre problem for any higher $k < p$, we might be able to obtain further improvements in quality.

Another possible source of improvement in solution quality would be to consider more than one choice of v_1 in Step (0) of the heuristic, since this choice may contain some arbitrariness, particularly in the unweighted case. Different choices may lead to better solutions. Here it may be noted also that the proof of Theorem 1 does not actually require that v_1 has the largest weight, but merely that no other point has more than double its weight.

In spite of the possibility of improvements in quality from these modifications, we can give no proof that we will do better than the worst-case bound. The NP-hardness results referred to above would, in fact, seem to imply that no such guarantee of improvement could be achieved except in special cases.

It may be observed that the generality of our heuristic allows it to be applied in spaces where the distance does not arise from a norm. For example, arbitrary geographical regions of the plane or sphere fall within its scope. As long as we can compute distances between the points of V, we can apply the heuristic. This generality has some interesting consequences. The heuristic selects only points of V, and only uses distance calculations between points of V. Thus the same set of points $\{v_i\}$ is δ times optimal, whether or not the centres are restricted to lie within V or not. In one case the appropriate matrix space is X, in the other it is V equipped with the metric of X. In fact, V equipped with this metric may be viewed as a complete graph with edge weights satisfying the triangle inequality. Thus, by using our heuristic to determine the δ times optimal solution on such a graph, we get the same performance guarantee in any metric space within which the graph can be embedded using a distance-preserving mapping.

Finally, the proof of Theorem 1 may be easily modified so that α is replaced by $\alpha' = \max_{1 \leqslant i < j \leqslant p+1} w(v_j)/w(v_i)$. In some cases this may provide a tighter posteriori bound on the optimum than the priori bound given by Theorem 1.

References

[1] Z. Drezner, "The p-centre problem – Heuristic and optimal algorithms", *J. Oper. Res. Soc.* **35**, 741–748 (1984).

[2] M.E. Dyer, "On a multidimensional search technique and its application to the Euclidean one-centre problem", Mathematics Department, Teesside Polytechnic, 1984.

[3] R.J. Fowler, M.S. Paterson and S.L. Tanimoto, "Optimal packing and covering in the plane are NP-complete", *Inform. Proc. Letters* **12**, 133–137 (1981).

[4] D.S. Hochbaum and D.B. Shmoys, "A best possible heuristic for the k-centre problem", University of California, Berkeley, 1983.

[5] D.S. Hochbaum and D.B. Shmoys, "Powers of graphs: A powerful approximation technique for bottleneck problems", in: *Proceedings of the 16th annual ACM Symposium on Theory of Computing, Washington, 1984.*

[6] W.L. Hsu and G.L. Nemhauser, "Easy and hard bottleneck location problems", *Discrete Appl. Math.* **1**, 209–216 (1979).

[7] O. Kariv and S.L. Hakimi, "An algorithm approach to network location problems I: The p-centers", *SIAM J. Appl. Math.* **37**, 513–538 (1979).

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

[9] N. Megiddo and K.J. Supowit, "On the complexity of some common geometric location problems", *SIAM J. Computing* **18**, 182–196 (1984).