Efficient coverage of the Grid for Surveillance in Sensor Networks

TEOFILO F. GONZALEZ and ANAND MEKA
Department of Computer Science
University of California
Santa Barbara, CA, 93106-5110
email: {teo,meka} @cs.ucsb.edu

ABSTRACT
Sensor networks are distributed systems constrained by power and memory resources that can span large geographic regions. We consider the problem of placement of a fixed number of base stations k, in a $N \times N$ grid of sensor nodes in order to aggregate and process the data. Every node transmits data to the closest base station via Manhattan routing. The communication cost is the length of the shortest path between the sensor node and the base station. In such an environment, we try to determine the optimal placement of base stations in order to optimize the overall communication overhead in transmitting data from sensor nodes to the base station. We first present a greedy heuristic which places the base stations in a specified fashion. Then, we give the lower bounds (for communication overhead) for any such placement. We also present an near optimal algorithm for such a placement. The experimental results show that our algorithms perform close to the optimal.

KEY WORDS
Sensor Networks, Approximation Algorithms, Coverage, k–medians.

1 Introduction

Governments are now planning to deploy large scale sensor networks in order to provide solutions to challenging problems: real-time traffic monitoring, safety monitoring (structural, fire and physical security monitoring), military sensing and tracking, seismic measurement, and wildfire tracking. Wireless Sensor Networks have merged as a new information gathering paradigm based on the collective effort of a large number of sensing nodes. In such networks, nodes act in response to environmental events and relay collected and possibly aggregated information with the help of a wireless network to the base stations. The goal of active *sensornet* research has been to lower communication in order to reduce power consumption and increase the lifetime of the network. We deal with the problem of optimal placement of k base stations in a $n \times n$ grid (2D-mesh) network. Every sensor node forwards data to the "closest" base station, where the distance is defined by the L_1 metric. In other words, we desire to locate k stations on the grid such that the sum of distances from each of nodes to the closest station is minimized. So, we are trying to solve the k-median problem for the grid setting where the distance measure is L_1.

We present related work and background in Section 2. In Section 3, we describe our algorithms and evaluate the performance of our algorithms in Section 4, and conclude in Section 5.

2 Related Work

The general problem that the paper addresses—k-median—has been extensively studied in classification and data mining. The *facility-location* problem is somewhat similar to the k-median problem: we are given a set $S = x_1, \ldots, x_n$ of n points together with the cost c_i of opening a facility at x_i. We are to find a set F of facilities to minimize the sum of distances from each points of S to the closest facility, with an additional cost of opening the facilities. Hochbaum [5] gives a $O(\log n)$ greedy approximation algorithm; Korupolu et. al [7] give a $(5+\epsilon)$-approximation for the problem. Shymos et. al [10] give a 3.16 approximation, improved to 2.41 by Guha and Khuller [4] finally improved to 1.74 by Chudak [3].

The k-median problem is NP-hard for arbitrary metric spaces. Lin and Vitter [8] use *filtering* to round fractional solutions of linear programming relaxations of the problem. They also [9] give a $2(1+\epsilon)$ approximation using $(1+1/\epsilon)k$ median locations. Jain and Vazrani [6] give a primal-dual 6-approximation for the capacitated version of the problem (the medians can be placed only in specific places). Charikar and Guha [1] refine it to a 4-approximation algorithm. Finally, they [2] present a first constant-factor (6.67) approximation to the general problem.

3 Problem

Given an $n \times n$ grid of sensors and a positive integer k, our problem is to place in k of the sensors a base station. Our objective function is to minimize the summation of the Manhattan distance from each sensor to its closest base station. The idea is that the the sensors will communicate with one of its closest base stations where their information will be processed and summarized. Then the bases stations will communicated with each other figure out what is the global state of the system. In this paper we will solve the first problem. Namely, given a grid (or lattice) where the sensors are located at the intersection of each tow i and column j and given a positive integer k, find a subset of k grid points $B \subseteq \{(i,j)|1 \leq i \leq n, 1 \leq j \leq n\}$ in such a way that $\sum_i \sum_j min_{(i',j') \in B} d\{(i,j),(i',j')\}$.

3.1 Lower Bounds

Suppose that we locate a base station at the center of the $n \times n$ grid. Then there is one sensor at a distance zero from it, 4 sensors at a distance 1 from it, 8 sensors ar a distance 2, $\ldots$, and $4i$ sensors at a distance i from it as long as $i \leq n/2$. Let us now define the *diamond of size s* for the base station (i,j) as the set of all the sensors located at a distance at most $s-1$ form (i,j). Clearly the number of such points, provided that (i,j) is at the center of the $n \times n$ grid and $s < n/2$ is $1+\sum_{j=2}^{s} 4(j-1)$. The objective function value of the a diamond of size s is therefore $\sum_{j=2}^{s} 4(j-1) = 2((s-1)s(2s-1))/3$.

Suppose now that n^2 is divisible by k, and that $n^2/k = 1 + \sum_{j=2}^{s} 4(j-1)$ for some positive integer s. Then the objective function value of an optimal solution must be at least equal to the objective function value of a diamond of size s multiplied by k.

Theorem: Given n and k an optimal solution to our problem is at least $k \cdot 2((s-1)s(2s-1))/3$, where $s = \frac{k+\sqrt{2kn^2-k^2}}{2k}$

Proof: The positive roots of $1+\sum_{j=2}^{s} 4(j-1) - n^2/k$ are a solution and the only positive root is given above. []

The above expression is

$$\frac{(n^2-k)\sqrt{2kn^2-k^2}}{3k}$$

3.2 Simple Upper Bound for a special case

In this section we find a simple lower bound for the case when there exists two positive integers a and b such that $k = a^2$ and $n = ba^2$.

In this case our solution is straight forward. We place the base stations in a subgrid of size $k \times k$ where each pair of base stations in each row is at a distance b from each other and at a distance $b/2$ from the left and right boundary of the sensor grid. The same relation holds for the columns.

The objective function of the solution that we propose is $\frac{n^3}{2\sqrt{k}} - n^2$.

Here is how to established the above equation. The objective function value of our solution is

$$\sum_{i=0}^{\frac{n}{2\sqrt{k}-1}} \sum_{j=0}^{\frac{n}{2\sqrt{k}-1}} i+j$$

Multiplying by $4k$ we obtain

$$\frac{n^3}{2\sqrt{k}} - n^2$$

Our lower bound from above is

$$\frac{(n^2-k)\sqrt{2kn^2-k^2}}{3k}$$

The upper bound divided by the lower bound is

$$\frac{6kn^2 - 3\sqrt{k}n^3)}{2(n^2-k)\sqrt{2n^2k-k^2}}$$

Substituting $n = ba^2$ and $k = a^2$ and simplifying we get

$$\frac{3a^4b^2(ab-2)}{2(a^2b^2-1)a^2\sqrt{2a^2b^2-1}}$$

Suppose that $n^2/k = a^2b^2 \geq c$, for some integer $c > 1$. Then, $a^2b^2 - 1 \geq (1-1/c)a^2b^2$ and $2a^2b^2 - 1 \geq (2-1/c)a^2b^2$. Replacing the above inequalities and $ab - 2 < ab$, the approximation bound becomes

$$\frac{3}{2(1-1/c)\sqrt{2-1/c}}$$

The following table gives the approximation bound for different values of $c = n^2/k$.

Table 1. Approximation bound for different values of c.

$c = n^2/k$	2	3	4	5	6	7	8	9
approx	2.45	1.74	1.51	1.40	1.33	1.28	1.25	1.23
$c = n^2/k$	10	11	12	13	14	16	19	26
approx	1.21	1.19	1.18	1.17	1.16	1.15	1.13	1.11

It also works for $c = 2$, but not for $c = 1$. Need to say something when c is small.

3.3 Upper Bound for the General Case

The case when $k \neq a^2$ for some positive integer a. In this case we use the next largest integer and leave some vacant base stations. In this case the approximation bound is ...

3.4 Improved Upper Bound

$\frac{n^2}{k} = 1$

Use the diamonds instead approximating the edges only.

3.5 Algorithms

write some stuff

4 Experimental Results

5 Conclusions

References

[1] M. Charikar and S. Guha. Improved combinatorial algorithms for the facility location and k-median problems. *IEEE Symposium on Foundations of Computer Science*, 1999.

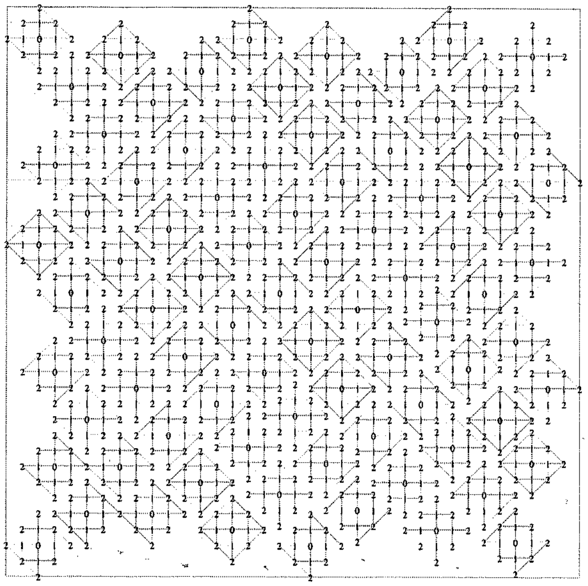
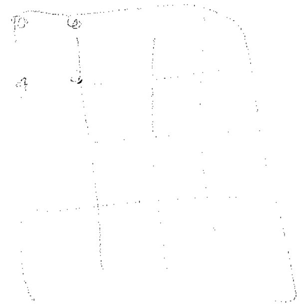


Figure 1. Routing Example



Romb
Rhombus

[2] M. Charikar, S. Guha, E.Tardos, and D. Shyamos. A constant factor approximation algorithm for the k-median problem. *ACM-STOC*, 1999.

[3] F. Chudak. Improved approximation algorithms for uncapacitated facility location. *LNCS*, 1998.

[4] S. Guha and S. Khuller. Greedy strikes back: Improved facility location algorithms. *ACM-SIAM*, 1998.

[5] D. S. Hochbaum. *Heuristics for the fixed cost median problem.* Math. Programming, 1982.

[6] K. Jain and V. V. Vazirani. *Primal-dual approximation algorithms for metric facility location and k-median problems.* IEEE Symposium on Foundations of Computer Science, 1999.

[7] M. R. Koruplou, C. G. Plaxton, and R. Rajaraman. Analysis of a local search heuristic for facility location problems. *ACM-SIAM*, 1998.

[8] J. H. Lin and J. S. Vitter. Approximation algorithms for geometric median problems. *Inform. Proc. Letters*, 1992.

[9] J. H. Lin and J. S. Vitter. *ϵ-approximation with minimum packing constraint violation.* ACM-STOC, 1992.

[10] D. B. Shymos, E. Tardos, and K. Aardal. Approximation algorithms for facility location problems. *ACM-STOC*, 1997.