

Resource Allocation Algorithms

Lecture 05

Lecture by Dr. Tami Tamir

Typeset by Steven Karas

2018-04-24

Last edited 18:15:55 2018-04-24

Disclaimer These lecture notes are based on the lecture for the course Resource Allocation Algorithms, taught by Dr. Tami Tamir at IDC Herzliyah in the spring semester of 2017/2018. Sections may be based on the lecture slides written by Dr. Tami Tamir.

Agenda

- Kahoot quiz
- Facility Locations

1 Facility Location

Facility location is the set of problems surrounding where facilities that provide services should be located so as to maximize some objective function (typically min max distance or min sum distance).

Antennas are a currently popular application of such algorithms¹.

There are some different questions we can ask, for example where facilities should be located, how many facilities are required. We can then apply many different constraints on the problem, such as differing costs, exclusive allocation, capacities, dynamic clients, etc. We can also have different location models, from planar models which we consider as a continuous optimization problem, to network models where there are a finite number of locations with differing costs.

Sometimes, these problems are easy, but other times they are \mathcal{NP} -hard, and sometimes approximate solutions are sufficient as the input data is estimated anyways.

1.1 Planar Location Models

A mathematical plane (typically embedded in a 3-dimensional space), using different distance metrics for the objective function.

- Euclidean distance
- Manhattan distance
- L_p norm²

Example objective functions:

- Single facility minisum - Establishing a single new facility to minimize sum of distances between clients to new and existing facilities.
- Single facility minimax - Establishing a single new facility to minimize max distance between clients to new and existing facilities.
- Multi-facility minisum - as above, but with multiple facilities
- Multi-facility minimax - as above, but with multiple facilities

¹Considerations for antennas are more complex than we can cover in this course, as they service asymmetric areas, frequency reuse, non-convex areas, etc.

²we will not discuss this metric in this course

1.1.1 Planar single facility - Minisum Manhattan distance

Given m existing customers with locations l_i and weight w_i , establish a single facility to minimize total distance traveled.

An optimal solution is given by algebraic derivation:

$$f(l_f) = \sum_{i=1}^m w_i d(l_f, l_i)$$

For Manhattan distance, distance is component-wise independent, so we can consider each one as a single optimization.

1.2 Graph Location Models

A better model for many applications, especially when the cost of modifying an existing network is sufficiently high. Sometimes a more generic model may provide quicker approximations, and guide future modifications of the network.

Path a path is a sequence of edges leading from one vertex to another.

Cycle a cycle is a nontrivial path that begins and ends at the same vertex.

Tree a connected, cycle-free graph. As a property, there is a unique path between any two vertices.

Distance the distance between two vertices is the length of the shortest path between them.

Example objective functions:

- Coverage - how many facilities need to be established to ensure all vertices are within a given distance of any facility.
- Center problems - k-centers. Minimize maximum distance between facilities and customers.
- Median models - k-medians. Minimize the sum of distances between customers and the closest facility.
- Fixed charge - Minimize sum of facility establishment and transportation costs
- Integrated production-distribution models - considers the total cost of production and shipping.
- Routing models - Vehicle Routing Problem (VRP) - deliver product from warehouses and return. TSP is a special case of this.

1.2.1 Coverage

How many facilities need to be established to cover all customers within some distance. Variants include per-customer requirements, restricting potential facility locations to only at vertices or allowing them along vertices.

There exists a \mathcal{P} -time algorithm to solve this for trees.

The generic problem is \mathcal{NP} -hard. We will prove this by reduction from DOMINATING SET. A dominating set in an undirected graph is a collection S of vertices with the property that every vertex $v \in G$ is either in S or there is an edge between v and some vertex in S .

Given a problem for DS, construct a new graph such that the weight of all vertices and edges is 1. Solving coverage for such a network gives a solution for DOMINATING SET.

1.2.2 k-center

Establish k facilities so as to minimize the maximum distance between a customer and its nearest facility (a minimax problem). We distinguish between vertex center which restricts facility locations to vertices, and absolute center which allows facilities to be located anywhere along the network.

1-Center in a general network Let x be any point on the network. The distance between x and the node of G which is farthest away from it is denoted as $m(x) = \max_{v \in V} d(x, v)$. Define a local center as a point x_e on an edge e if for every point x along the edge include both endpoints p, q , it holds that $m(x_e) \leq m(x)$. Define a vertex center as a vertex $v^* \in V$ if for every $v \in V$, $m(v^*) \leq m(v)$. Define an absolute center $x^* \in G$ as a point if for every point $x \in G$, it holds that $m(x^*) \leq m(x)$.

For vertex centers, find the all-pairs shortest paths, and select the vertex with the shortest distance to all other vertices.

It is easy to construct instances for which the absolute center is not along an edge adjacent to the vertex center.

For absolute centers, find all the local centers for each edge, and select the one that minimizes the maximal distance to all vertices. The algorithm for finding a local center can be found on slide 20.

We can prove that for a local center x_e on edge e between p and q with edge weight $c(p, q)$:

$$m(x_e) \geq \frac{m(p) + m(q) - c(p, q)}{2}$$

Let $d(x, v)$ be a distance metric from a point x to a vertex v such that:

$$\begin{aligned} d(x, p) &= x \\ d(x, q) &= c(p, q) - x \\ m(p) &\leq m(x) + x \\ m(q) &\leq m(x) + c(p, q) - x \\ m(q) + m(p) &\leq 2m(x) + x + c(p, q) - x \end{aligned}$$

We can use this theorem to reduce the number of edges that need to be considered for local centers.

General case Given a undirected network G with n vertices and a positive integer k . Let $d(X_k, n)$ be the minimum distance between any one of the points in X_k and the node $v \in G$. The K-CENTER problem is to find k distinct points on the graph $X_k = \{x_1, \dots, x_k\}$ that minimizes $\max_{v \in V} d(X_k, v)$.

If k is constant w.r.t. n , then we have $O(n^k)$ possible vertex centers. We can evaluate each of these possible vertex centers in turn (in polynomial time).

If k is a function of n problem is \mathcal{NP} -hard. The proof follows by reduction from DOMINATING SET, by noting that a dominating set of size k is a k -center with objective value of 1.

2-approximation The following is a 2-approximation that relaxes the objective function³ for instances that obey the triangle inequality. Start by choosing the first center arbitrarily, and at every step, choose the vertex that is furthest from the current centers to become a center. Continue until k centers have been chosen.

To prove approximation, first note that the sequence of distances from a new chosen center to the closest existing center is non-increasing. Consider the point that is furthest from the k chosen centers. We need to show that this point is at most $2 \cdot \text{OPT}$ from the closest center. Assume by contradiction that it is not. This means that distances between all centers are more than $2 \cdot \text{OPT}$. We have $k + 1$ points with distances greater than $2 \cdot \text{OPT}$ between every pair of centers and this point. Each point has a center of the optimal solution with less than OPT distance. There must exist a pair of points with the same center in the optimal solution (by pigeonhole). The distance between them is at most $2 \cdot \text{OPT}$. This contradicts our assumption, and therefore our algorithm is a 2-approximation.

An example was given for this was given on the board, but elided here because I feel it's trivial

Approximation hardness If $\mathcal{P} \neq \mathcal{NP}$, then there is no approximation algorithm for the K-CENTER problem that gives a $(2 - \epsilon)$ -approximation for any $\epsilon > 0$.

The proof sketch follows by reduction from DOMINATING SET. Let $G = (V, E); k$ be an instance of DOMINATING SET. Construct a complete graph $G' = (V, E' = V \times V)$ such that $d(u, v) = 1$ if $(u, v) \in E$ else 2. Note that G' satisfies the triangle inequality. Run such an algorithm and consider the objective value d of the selected set. If $d \leq 1$, then G has a DS of size k . If $d \geq 2$, then G does not have a DS of size k .

The rationale for this is that the approximation must select a set with $d < 2$, and it will use primarily edges that were in the original graph.

1.2.3 k-median

Minimize the sum of distances between customers and the closest facility.

Given an undirected network G with n vertices with vertex weights h_j , and a positive integer k . Let $d(X_k, j)$ be the minimum distance between any one of the points x_1, \dots, x_k and the node j . An optimal solution to K-MEDIAN minimizes:

³there are variants that relax the budget of k

$$J(X_k) = \sum_{i=1}^n h_j d(X_k, i)$$

I wasn't able to keep up with the rest of the lecture.

2 Next week

We will continue facility location next week.

References

- [1] Michael L Pinedo. *Scheduling: theory, algorithms, and systems*. Springer, 2016.