Metric facility location

Uriel Feige

Lectures 3 and 4, April 10–17, 2019

Abstract

These are note that I wrote for my own use, and are rather sketchy.

1 Introduction

In the metric facility location problem, there is a set F of facilities, a set D of clients, facility $i \in F$ has opening cost f_i, client $j \in D$ has connection cost c_{ij} to facility $i \in F$. The c_{ij} are distances that obey the triangle inequality. The goal is to open a set of facilities and connect every client to an open facility, while minimizing the total cost.

We have seen in the previous semester approximation algorithms for metric facility location based on an LP. Deterministic rounding gave an approximation ratio of 4 (Chapter 4.5 in [3]), and randomized rounding gave an approximation ratio of 3 (Chapter 5.8 in [3]).

Let us first recall the LP relaxation for facility location, its dual, and an interpretation for the dual.

Primal LP.
minimize $\sum f_i y_i + \sum c_{ij} x_{ij}$ **subject to:**

- $\sum_{i \in F} x_{ij} = 1$ for every client $j \in D$.
- $x_{ij} \leq y_i$ for every client $j \in D$ and facility $j \in F$.
- $x_{ij} \geq 0$ for every client $j \in D$ and facility $j \in F$.
- $y_i \geq 0$ for every facility $j \in F$.

Dual LP.
maximize $\sum v_j$ **subject to:**

- $\sum_j w_{ij} \leq f_i$ for every client $j \in D$ and facility $j \in F$.
- $v_j \leq w_{ij} + c_{ij}$ for every client $j \in D$ and facility $j \in F$.
- $w_{ij} \geq 0$ for every client $j \in D$ and facility $j \in F$.
- $v_j \geq 0$ for every client $j \in D$.

One may interpret w_{ij} as the share of the opening cost that facility i intends to charge j, and v_j as the price that a client is willing to pay in order to cover its connection cost and facility opening share. No facility is allowed to overcharge, and every client makes the choice minimizing his payment.

2 A deterministic primal dual algorithm

We present a deterministic primal dual algorithm (taken from Chapter 7.6 in [3]) that gives an approximation ratio of 3.

Every client j raises its budget v_j in a uniform rate. With respect to every facility i independently, the client computes a partition of the budget, c_{ij} towards the connection cost, and $w_{ij} = \min[v_j - c_{ij}, 0]$ towards the opening cost. A facility i becomes tight once $\sum_j w_{ij} = f_i$. At that moment it is opened and the clients contributing towards opening it are connected to it. We stop increasing the budget of client i at the moment that it becomes connected, either by contributing towards opening a tight facility, or because it can afford the connection cost to an already open facility.

Let F_1 be the set of facilities that are open at the end (when all clients get connected). A client j might pay shares of the opening costs of several facilities, if each of them on its own can be supported by j's budget. We say that two facilities in F are independent if there is no client who shares (a strictly positive payment) in the opening costs of both of them.

Of the open facilities F_1, keep only a maximal independent subset F_2. For facilities in $F_1 \setminus F_2$, their original clients are each moved to its nearest open facility.

The dual solution is feasible. A facility i is opened at the moment that $\sum_j w_{ij} = f_i$, and thereafter no client j can change the dual variable w_{ij}. A client j is connected at the moment that it can afford it, and thereafter, his cost to any facility does not change.

In the primal solution, the clients originally paying for the opening of F_2 have a budget of v_j that cover both their connection cost and the opening cost.

For a client j who was originally connected only to $F_1 \setminus F_2$, he need not contribute anything towards opening costs. Let i be j's original facility. If i was closed then it was not independent of some facility i'. Hence there is a client j' that shared cost both with i and i'. Hence $v_{j'} \leq v_j$ (this is an inequality because $v_{j'}$ stopped increasing no later than at the time that i was opened, whereas v_j increased until the point when j was assigned to i), and $c_{ij'} \leq v_{j'}$, $c_{i'j'} \leq v_{j'}$, and $c_{ij} \leq v(j)$. The triangle inequality implies that $c_{i'j} \leq 3v_j$.

3 Randomized rounding of the LP

We present a randomized rounding algorithm (taken from Chapter 12.1 in [3]) for the LP, that gives an approximation ratio of $1 + \frac{2}{e}$.

We recall the following about pairs of optimal primal and dual solutions to the LP.

Lemma 1 *For pairs of optimal primal and dual solutions to the LP, $x^*_{ij} > 0$ implies that $c_{ij} \le v^*_j$.*

Proof. By complementary slackness for x_{ij}, and nonnegativity of w_{ij}. ■

We assume that optimal solutions are *complete*, meaning that $x^*_{i,j} > 0$ implies that $x^*_{i,j} = y^*_i$. This can be assumed w.l.o.g., because an incomplete solution can be transformed into a complete solution for an equivalent instance, in which we make $|D|$ identical copies (in terms of opening costs and connection costs) of facility i. After sorting the clients in order of increasing x^*_{ij}, we set $y^*_{ij} = x^*_{ij} - x^*_{i,(j-1)}$, and connect client j (if $x^*_{ij} > 0$) to the first j copies.

For client j, let $C^*_j = \sum_{i \in F} c_{ij} x^*_{ij}$ denote its expected assignment cost.

We say that a client j and a facility i are *neighbors* if $x^*_{ij} > 0$.

The randomized rounding algorithm is as follows.

- Initially, all clients and all facilities are *active*. Clients can change status to *assigned* or *discarded*. Facilities can change status to *open* or *discarded*.
- If there is an active client, let j be the client minimizing $v^*_j + C^*_j$.
- Open one facility i at random with probability x^*_{ij}. Assign j to i. Discard all neighboring facilities of j and their neighboring clients.
- When no active client remains, if there still are active facilities, open each active facility i independently with probability y^*_i.
- Connect each client to its nearest open facility.

A **key observation** is that the set of active clients and the client considered in step t depend only on the step number, and not on the outcomes of random decisions in previous steps.

Lemma 2 *The expected opening cost of the rounded solution is not more than that of the LP.*

Proof. Every facility i had exactly one step it which it is considered, and at that step it is opened with probability y^*_i. ■

Lemma 3 *There is a* contention resolution *rule (see Section 4) that ensures that for every client j and every neighboring facility i, the probability that j ends up connected to i is at least $(1 - \frac{1}{e})x^*_{ij}$.*

Proof. Here we use the fact that the instance is complete.

If j was ever a minimizing client, no neighboring facility of j was considered in an earlier round, and the lemma holds. Hence we may assume that j was never a minimizing client.

The set of facilities considered in every step is independent of random decisions (due to the key observation). Use bounds for fair contention resolution from Section 4 (where p_i is the probability that a facility that is a neighbor of j is opened at step i) to show that there is probability at least $(1 - \frac{1}{e})y_i^*$ to connect to each neighboring facility i. ∎

As a consequence of Lemma 3 we can assume the following:

- For each client, there is probability at least $1 - \frac{1}{e}$ of connecting it to a neighboring facility.
- Conditioned on connecting client j to neighboring facility open, its expected connection cost is at most C_j^*.

Lemma 4 *Conditioned on not having a neighboring facility open, the expected connection cost of client j is at most $2v_j^* + C_j^*$.*

Proof. Client j can be connected to a facility i' opened by the first j' that is a distance 2 neighbor of j. Note that $v_{j'}^* + C_{j'}^* \leq v_j^* + C_j^*$. Let S be the set of facilities in the intersection of the neighborhoods of j and j'. The assignment cost of j to any facility in S is at most v_j^*. For j', there are two cases to consider.

- The expected assignment cost of j' to S is at most $C_{j'}^*$. Use the fact that the assignment cost of j' to i' is at most $v_{j'}^*$.
- The expected assignment cost of j' to S is more than $C_{j'}^*$. Use the fact that the expected assignment cost of j' to random $i' \notin S$ is at most $C_{j'}^*$, and the assignment cost of j' to any facility in S is at most $v_{j'}^*$.

In any case, using the triangle inequality through a random vertex in S, the expected assignment cost of j is at most

$$v_j^* + v_{j'}^* + C_{j'}^* \leq 2v_j^* + C_j^*$$

∎

If follows that the total assignment cost is at most

$$\sum_{j \in D} (1 - \frac{1}{e})C_j^* + \frac{1}{e}(2v_j^* + C_j^*) = \sum_{j \in D} (C_j^* + \frac{2}{e}v_j^*)$$

Hence the expected cost of the solution is at most

$$\sum_{j \in D} (C_j^* + \frac{2}{e}v_j^*) + \sum_{i \in F} f_i y_i^* \leq (1 + \frac{2}{e})LP^*$$

4 Fair contention resolution

Suppose that there are k candidates, and we need to select at most one of them. Each of the candidates i is *eligible* independently with probability p_i. (The source of the problem is the independence. When the events of being eligible are mutually exclusive, as is the case for opening facilities within a single step of the algorithm of Section 3, the contention resolution problem does not arise.) If no candidate is eligible, no candidate is selected. If only one candidate is eligible, we select that candidate. If more than one candidate is eligible, we have a choice which of them to select. The rule used for this choice is referred to as a *contention resolution* scheme, and it may be randomized, and may depend on the values p_i of the eligible candidates and of the non-eligible candidates. The contention resolution scheme that we use affects the end probability z_i with which candidate i is selected. A contention resolution scheme is ρ-fair if for every candidate, $z_i \geq \rho p_i$. A contention resolution scheme is *fair* if it is ρ-fair for the highest possible ρ. This ρ equalizes for all candidates the conditional probability of being chosen, given that they are eligible.

As an example, suppose that there are two candidates with $p_1 = p > \frac{1}{2}$ and $p_2 = 1 - p$. A possible contention resolution rule to use when both candidates are eligible is to choose each one of them with probability $\frac{1}{2}$. This results in $z_1 = p^2 + \frac{1}{2}p(1-p) = \frac{1+p}{2}p_1$ and $z_2 = (1-p)^2 + \frac{1}{2}p(1-p) = \frac{2-p}{2}p_2$. In this case $\rho = \min[\frac{1+p}{2}, \frac{2-p}{2}] = \frac{2-p}{2} \leq \frac{3}{4}$ (when $p \geq \frac{1}{2}$). A fair contention resolution rule (when both candidates are eligible) would choose candidate 1 w.p. $1-p$ and candidate 2 w.p. p. This results in $z_1 = p^2 + p(1-p)^2 = (1-p+p^2)p_1$ and $z_2 = (1-p)^2 + p^2(1-p) = (1-p+p^2)p_2$. In this case $\rho = 1-p+p^2 \geq \frac{3}{4}$ (when $p \geq \frac{1}{2}$).

In general, the value of the best ρ is the solution to the following linear program. In the LP, the p_i are constants, and the z_i as variables. In addition, there are variables z_{iS} for every set S of candidates, and candidate $i \in S$, specifying the probability that i is chosen given that S is eligible. For a set S of candidates, $p(S) = \prod_{j \in S} p_j \prod_{j \notin S}(1-p_j)$ denotes the probability that S is the set of eligible candidates.

Maximize ρ subject to:

- $z_i \geq \rho p_i$ for every candidate i.
- $\sum_{i \in S} z_{iS} = 1$ for very set S of candidates.
- $z_i = \sum_{S | i \in S} p(S) z_{iS}$ for every candidate i.
- $z_{iS} \geq 0$ for every set S of candidates, and candidate $i \in S$.

The optimal value of ρ is

$$\rho^* = \frac{1 - \prod_i (1-p_i)}{\sum_i p_i}$$

and a way of achieving it without solving the LP is presented in [1]. Here we present a weaker bound that suffices for our purpose.

Lemma 5 *If* $\sum p_i = 1$ *then there is a contention resolution rule that ensures* $\rho \geq 1 - \frac{1}{e}$.

Proof. Assume for simplicity that the p_i are rational. Let q be an integer for which $p_i q$ is integral for all i. We view the situation as if there are q symmetric mini-candidates, each eligible independently with probability $\frac{1}{q}$. Such a situation is equivalent to one in which each candidate i is represented by $p_i q$ mini-candidates, and conditioned on being eligible (which happens with probability p_i), each of its mini-candidates is eligible independently with probability $\frac{1}{qp_i}$. Resolve conflicts among eligible mini-candidates (of all candidates) uniformly at random. The winning candidate is the one whose mini-candidate was selected (if there were eligible mini-candidates).

The probability that there is at least one eligible mini-candidate is exactly $1-(1-\frac{1}{q})^q > 1-\frac{1}{e}$. Conditioned on there being an eligible mini-candidate, each mini-candidate is equally likely to be chosen (by symmetry). Hence candidate i is chosen with probability at least $p_i q \frac{1-\frac{1}{e}}{q} = (1-\frac{1}{e})p_i$. (This proof approach suffices for the lemma but is not tight, because there might not be any eligible mini-candidates even if there are eligible candidates.) ■

5 Towards optimal approximation ratios

There are approximation ratios better than $1 + \frac{2}{e}$ known for metric facility location. One such algorithm is given in [2], a paper which is the subject of a homework assignment. There have been further improvements, but we do not yet know the best possible approximation ratio.

The known hardness of approximation result is a ratio of roughly 1.463 (see Chapter 16.2 in [3]). Here is a high level sketch of how it is derived. There is a natural reduction from set cover, in which sets are facilities, items are clients, the connection cost from a set to its items is 1 , and 3 to other items. For the set cover instance, it is NP-hard to distinguish between the case that k sets cover all items, and the case that for every k', any collection of k' sets leaves a fraction of $(\frac{k-1}{k})^{k'}$ items uncovered. Making the opening cost of each facility $\gamma\frac{n}{k}$, choose the γ that gives the largest gap between *yes* instances and *no* instances.

References

[1] Uriel Feige, Jan Vondrk: The Submodular Welfare Problem with Demand Queries. Theory of Computing 6(1): 247–290 (2010).

[2] Kamal Jain, Mohammad Mahdian, Evangelos Markakis, Amin Saberi, Vijay V. Vazirani: Greedy facility location algorithms analyzed using dual fitting with factor-revealing LP. J. ACM 50(6): 795–824 (2003)

[3] David P. Williamson, David B. Shmoys: The Design of Approximation Algorithms. Cambridge University Press 2011.