A 1.488-approximation algorithm for the uncapacitated facility location problem

Shi Li

Department of computer science,
Princeton University, Princeton NJ 08544, USA
shili@cs.princeton.edu

Abstract. We present a 1.488 approximation algorithm for the metric uncapacitated facility location (UFL) problem. The previous best algorithm was due to Byrka [1]. By linearly combining two algorithms $A1(\gamma_f)$ for $\gamma_f = 1.6774$ and the (1.11,1.78)-approximation algorithm $A2$ proposed by Jain, Mahdian and Saberi [8], Byrka gave a 1.5 approximation algorithm for the UFL problem. We show that if γ_f is randomly selected from some distribution, the approximation ratio can be improved to 1.488. The algorithm cuts the gap with the 1.463 approximability lower bound by almost 1/3.

1 Introduction

In this paper, we present an improved algorithm for the uncapacitated facility location problem (UFL). In the UFL problem, we are given a set of potential facility locations $\mathcal{F}$, each $i \in \mathcal{F}$ with a cost c_i, a set of clients $\mathcal{C}$, and a metric d over $\mathcal{F} \cup \mathcal{C}$. The solution is a set $\mathcal{F}' \subset \mathcal{F}$ denoting the positions where facilities are open. We want to find a solution that minimizes the sum of the facility cost and the connection cost. The facility cost is the sum of c_i over all $i \in \mathcal{F}'$, and the connection cost is the sum of distances from all clients to their nearest facility in $\mathcal{F}'$.

The UFL problem is NP-hard and has received a lot of attention in the literature. Hochbaum [5] presented a greedy algorithm with $O(\log n)$ approximation guarantee. Shmoys, Tardos, and Aardal [12] used the techniques of Lin and Vitter [10] to give the first constant-approximation algorithm. The approximation ratio they achieved was 3.16. After that, a large number of constant approximation algorithms have been proposed([4], [9], [3], [7], [8], [11]). Up to now, the best known approximation ratio was 1.50, due to Byrka [1].

In terms of hardness result, Guha and Kuller [6] showed that there is no ρ approximation algorithm for the metric UFL problem for any $\rho < 1.463$, unless $\mathbf{NP} \subset \mathbf{DTIME}\left(n^{O(\log\log n)}\right)$. Jain et al. ([8]) generalized the result to show that no (λ_f, λ_c)-bifactor approximation algorithm exists for $\lambda_c < 1 + 2e^{-\lambda_f}$ unless $\mathbf{NP} \subset \mathbf{DTIME}\left(n^{O(\log\log n)}\right)$. Here, we say that an algorithm is a (λ_f, λ_c)-approximation algorithm if the solution has total cost $\lambda_f F^* + \lambda_c C^*$, where F^* and C^* are the facility and the connection cost of an optimal solution, respectively.

Built on the work of Byrka [1], we give a 1.488-approximation algorithm for the UFL problem. Byrka presented an algorithm $A1(\gamma_f)$ which gives the optimal bi-factor approximation $(\gamma_f, 1 + 2e^{-\gamma_f})$ for $\gamma_f \geq 1.6774$. By linearly combining $A1(1.6774)$ and the $(1.11, 1.7764)$-approximation algorithm $A2$ proposed by Jain, Mahdian and Saberi [8], Byrka was able to give a 1.5 approximation algorithm. We show that if γ_f is randomly selected from some distribution, a linear combination of $A1(\gamma_f)$ and $A2$ can give a 1.488 approximation.

Due to the hardness result, for every λ_f, there is an hard instance for the algorithm $A1(\gamma_f)$. Roughly speaking, we show that a fixed instance can not be hard for two different λ_f's. We give an bifactor approximation ratio for $A1(\lambda_f)$ that depends on the input instance. Then we introduce a 0-sum game that characterizes the approximation ratio of our algorithm. The game is between an algorithm player who plays a distribution of λ_f and an adversary who plays an instance. By giving an explicit strategy for the algorithm player, we show that the value of the game is at most 1.488.

We first review the algorithm $A1(\gamma)$ used in [1] in section 2, and then give our improvement in section 3.

2 Review of the algorithm $A1(\gamma)$ in [1]

In this section, we review the $(\gamma, 1 + 2e^{-\gamma})$-approximation algorithm $A1(\gamma)$ for $\gamma \geq 1.67736$ in [1].

In $A1(\gamma)$ we first solve the following natural LP for the UFL problem.

$$\min \sum_{i\in\mathcal{F}, j\in\mathcal{F}^C} c_{i,j}x_{i,j} + \sum_{i\in\mathcal{F}} f_i y_i$$

$$\text{s.t} \quad \begin{array}{l} \sum_{i\in\mathcal{F}} x_{i,j} = 1 \ \forall j \in \mathcal{C} \\ x_{i,j} - y_i \leq 0 \ \forall i \in \mathcal{F}, j \in \mathcal{C} \\ x_{i,j}, y_i \geq 0 \ \forall i \in \mathcal{F}, j \in \mathcal{C} \end{array}$$

If the y-variables are fixed, x-variables can be assigned greedily : each $j \in \mathcal{C}$ is connected to 1 fraction of the closest facilities. After obtaining a solution (x, y), we modify it by scaling the y-variables up by a constant $\gamma > 1$. Let $\overline{y}$ be the scaled y-variables. We reassign x-variables using the above greedy process to obtain a new solution $(\overline{x}, \overline{y})$. By splitting facilities if necessary, we can assume $x_{i,j} \in \{0, y_i\}$, $\overline{x}_{i,j} \in \{0, \overline{y}_i\}$ and $\overline{y} \leq 1$, for every $i \in \mathcal{C}, j \in \mathcal{F}$.

For some $\mathcal{F}' \subset \mathcal{F}$, define $\mathsf{vol}(\mathcal{F}') = \sum_{i\in\mathcal{F}'} \overline{y}_i$. For a client j, we say a facility i is one of his *close facilities* if it fractionally serves j in $(\overline{x}, \overline{y})$. If $\overline{x}_{i,j} = 0$, but i was serving client j in solution (x, y), then we say i is a *distant facility* of client j. Let $\mathcal{F}_j^C, \mathcal{F}_j^D$ to be the set of close facilities, distance facilities of j, respectively, and $\mathcal{F}_j = \mathcal{F}_j^C \cup \mathcal{F}_j^D$. Define $d_{av}^C(j), d_{av}^D(j), d_{av}(j)$ to be the average distance from j to $\mathcal{F}_j^C, \mathcal{F}_j^D, \mathcal{F}_j$, respectively. The average distances are with respect to the weights $\overline{y}$ (or equivalently, y). Thus, $d_{av}(j)$ is the connection cost of j in the fractional solution. Define $d_{max}^C(j)$ to be maximum distance from j to a facility in $\mathcal{F}_j^C$. It's easy to see the following facts :

1. $d_{av}^C(j) \leq d_{max}^C(j) \leq d_{av}^D(j), d_{av}^C(j) \leq d_{av}(j) \leq d_{av}^D(j), \forall j \in \mathcal{C}$;
2. $d_{av}(j) = \frac{1}{\gamma} d_{av}^C(j) + \frac{\gamma-1}{\gamma} d_{av}^D(j), \forall j \in \mathcal{C}$;
3. $\mathsf{vol}(\mathcal{F}_j^C) = 1, \mathsf{vol}(\mathcal{F}_j^D) = \gamma - 1, \mathsf{vol}(\mathcal{F}_j) = \gamma$.

We greedily select a subset of clients $\mathcal{C}'$ in the following way. Initially $\mathcal{C}'' = \mathcal{C}, \mathcal{C}' = \emptyset$. While $\mathcal{C}''$ is not empty, select the client j in $\mathcal{C}''$ with the minimum $d_{av}^C(j) + d_{max}^C(j)$, add j to $\mathcal{C}'$ and remove j and all clients j' satisfying $\mathcal{F}_j^C \cap \mathcal{F}_{j'}^C \neq \emptyset$ from $\mathcal{C}''$. So, $\mathcal{C}'$ has the following properties :

1. $\mathcal{F}_j^C \cap \mathcal{F}_{j'}^C = \emptyset, \forall j, j' \in \mathcal{C}, j \neq j'$;
2. For every $j \notin \mathcal{C}'$, there exists a $j' \in \mathcal{C}'$ such that $\mathcal{F}_j^C \cap \mathcal{F}_{j'}^C \neq \emptyset$ and $d_{av}^C(j') + d_{max}^C(j') \leq d_{av}^C(j) + d_{av}^C(j)$. This j' is called the *cluster center* of j.

We shall randomly rounding the fractional solution to an integral solution. For each $j \in \mathcal{C}'$, open one of his close facilities randomly with probabilities $\overline{y}_i$. For each facility i that is not a close facility of any client in $\mathcal{C}'$, open it independently with probability $\overline{y}_i$. Each client j is connected to its closest open facility, and let C_j be the connection cost of j.

It's easy to see that the expected facility cost of the solution is exactly γ times the facility cost in the fractional solution. If $j \in \mathcal{C}'$, $\mathbb{E}[C_j] = d_{av}^C(j) \leq d_{av}(j)$.

Byrka in [1] showed that

1. The probability that some facility in $\mathcal{F}_j^C$ is open is at least $1 - e^{-\mathsf{vol}(\mathcal{F}_j^C)} = 1 - e^{-1}$, and under the condition that this is true, the expected distance between j and the closest open facility in $\mathcal{F}_j^C$ is at most $d_{av}^C(j)$;
2. The probability that some facility in $\mathcal{F}_j^D$ is open is at least $1 - e^{-\mathsf{vol}(\mathcal{F}_j^D)} = 1 - e^{-(\gamma-1)}$, and under the condition that this is true, the expected distance between j and the closest open facility in $\mathcal{F}_j^D$ is at most $d_{av}^D(j)$;
3. For a client $j \notin \mathcal{C}'$, $d(j, \mathcal{F}_{j'}^C \backslash) \leq d_{av}^D(j) + d_{max}^C(j) + d_{av}^C(j)$, where j' is cluster center of j; or equivalently, under the condition that there is no open facility in $\mathcal{F}_j$, the expected distance between j and the unique open facility in $\mathcal{F}_{j'}^C$ is at most $d_{av}^D(j) + d_{max}^C(j) + d_{av}^C(j)$.

Since $d^C_{av}(j) \le d^D_{av}(j) \le d^D_{av}(j) + d^C_{max}(j) + d^C_{av}(j)$, $\mathbb{E}[C_j]$ is at most

$$\begin{aligned} & (1-e^{-1})d^C_{av}(j) + e^{-1}(1-e^{-(\gamma-1)})d^D_{av}(j) + e^{-1}e^{-(\gamma-1)}(d^D_{av}(j) + d^C_{max}(j) + d^C_{av}(j)) \\ &= (1-e^{-1}+e^{-\gamma})d^C_{av}(j) + e^{-1}d^D_{av}(j) + e^{-\gamma}d^C_{max}(j) \\ &\le (1-e^{-1}+e^{-\gamma})d^C_{av}(j) + (e^{-1}+e^{-\gamma})d^D_{av}(j) \end{aligned} \tag{1}$$

Notice that the connection cost of j in the fractional solution is $d_{av}(j) = \frac{1}{\gamma}d^C_{av}(j) + \frac{\gamma-1}{\gamma}d^D_{av}(j)$. We want to compute the maximum ratio between $(1-e^{-1}+e^{-\gamma})d^C_{av}(j) + (e^{-1}+e^{-\gamma})d^D_{av}(j)$ and $\frac{1}{\gamma}d^C_{av}(j) + \frac{\gamma-1}{\gamma}d^D_{av}(j)$. Since $d^C_{av}(j) \le d^D_{av}(j)$, the ratio is maximized when $d^C_{av}(j) = d^D_{av}(j)$ or $d^C_{av}(j) = 0$. For $\gamma \ge \gamma_0 \approx 1.67736$, the maximum ratio is achieved when $d^C_{av}(j) = d^D_{av}(j)$, in which case the ratio is $1 + 2e^{-\gamma}$. Thus, the algorithm $A1(1.67736)$ gives a $(1.67736, 1 + 2e^{-1.67736} \approx 1.37374)$ bi-factor approximation. [1]

Then, combining $A1(\gamma_0)$ and the $(1.11, 1.78)$-bi-factor approximation algorithm $A2$ due to [11] will give a 1.5-approximation.

3 A 1.488 approximation algorithm for the UFL problem

In this section, we give our new approximation algorithm for the UFL problem. Our algorithm is also based on the combination of the $A1(\gamma)$ and $A2$. However, instead of using $A1(\gamma)$ for a fixed γ, we randomly select γ from some distribution.

To understand why the approach can reduce the approximation ratio, we list some requirements that the upper bound given by (1) is tight.

First, the facilities in $\mathcal{F}_j$ have tiny weights. In other words, for every facility $\max_{i \in \mathcal{F}_j} \overline{y}_i$ tends to 0. Moreover, all these facilities were independently sampled in the algorithm. These conditions are necessary to tighten the $1 - e^{-1}$ ($e^{-1} - e^{-(\gamma-1)}$ resp.) upper bound for the probability that at least 1 facility in $\mathcal{F}^C_j$ (mF^D_j resp.) is open.

Second, the distances from j to all the facilities in $\mathcal{F}^C_j$ ($\mathcal{F}^D_j$ resp.) are the same. Otherwise, the expected distance from j to the closest open facility in $\mathcal{F}^C_j$ ($\mathcal{F}^D_j$ resp.), under the condition that it exists, is strictly smaller than $d^C_{av}(j)$ ($d^D_{av}(j)$ resp.).

Third, $d^C_{max}(j) = d^D_{av}(j)$. This is also required since we used $d^D_{av}(j)$ as an upper bound of $d^C_{max}(j)$ in the last inequality of (1).

To satisfy all the above conditions, the distances from j to $\mathcal{F}_j$ must be distributed as follows. $1/(\gamma + \epsilon)$ fraction of facilities in $\mathcal{F}_j$ have distances a to j, and the other $1 - 1/(\gamma + \epsilon)$ fraction have distances $b > a$ to j. For ϵ tends to 0, $d^C_{av}(j) = a$ and $d^C_{max}(j) = d^D_{av}(j) = b$.

However, for the above instance, if we replace γ with $\gamma + 1.01\epsilon$ (say), we have $d^C_{max}(j) = d^C_{av}(j)$ and thus can save a lot in the approximation ratio.

Thus, by using only two different γ s, we are already able to make an improvement. To give a better analysis, we first give a upper bound on the expected connection cost of j, in terms of the distribution of distances from j to $\mathcal{F}_j$, not just $d^C_{av}(j)$ and $d^D_{av}(j)$. Then, we consider a 0-sum game between player A and player B. Player A plays a distribution of γ such that the facility cost is not scaled up by too much, player B plays a distribution of distances from j to $\mathcal{F}_j$ such that $d_{av}(j) = 1$, and the value of the game is the expected connection cost of j. We show that the value of this game is small by giving an explicit strategy for player A.

We structure the remaining part of this section as follows. We give the upper bound for the expected connection cost of j in subsection 3.1, and the explicit strategy for player A in subsection 3.2.

3.1 Upper bounding the expected connection cost of a client

It suffices to consider a client j where $j \notin \mathcal{C}'$.(We can treat the case $j \in \mathcal{C}'$ as the case where $j \notin \mathcal{C}'$ and there is some $j' \in \mathcal{C}'$ such that $d(j, j') = 0$). We first give an upper bound on the average distance from j

[1] Byrka's analysis in [1] was a little bit different; it used some variables from the dual LP. Later in [2], Byrka et al. gave an analysis without using the dual LP, which is the one we cite in our paper.

to $\mathcal{F}_{j'}^C \backslash \mathcal{F}_j$. The bound and the proof are similar to the counterparts in [1], except that we made a slight improvement. The improvement only affects the final approximation ratio by a tiny amount; however, it is useful because it simplifies the analysis later.

Lemma 1. *For some $j \notin \mathcal{C}'$, let j' be the cluster center of j. So $j' \in \mathcal{C}'$, $\mathcal{F}_j^C \cap \mathcal{F}_{j'}^C \neq \emptyset$ and $d_{av}^C(j') + d_{max}^C(j') \leq d_{av}^C(j) + d_{max}^C(j)$. We have $d(j, \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) \leq (2-\gamma) d_{max}^C(j) + (\gamma - 1) d_{av}^D(j) + d_{max}^C(j') + d_{av}^C(j')$, for any $\gamma \geq 1$.*

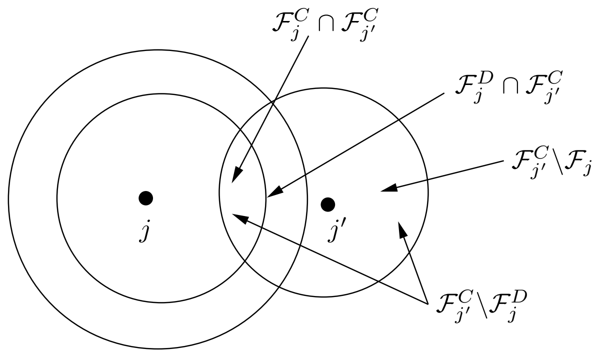

Fig. 1. Sets of facilities used in the proof

Proof. Figure 1 illustrates the sets of facilities we are going to use in the proof. If the distance between j and j' is at most $(2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + d_{av}^C(j')$, then the remaining $d_{max}^C(j')$ is enough for the distance from j' to any facility in $\mathcal{F}_{j'}^C$. So, we will assume

$$d(j, j') \geq (2-\gamma) d_{max}^C(j) + (\gamma - 1) d_{av}^D(j) + d_{av}^C(j') \tag{2}$$

(2) implies $d(j, j') \geq d_{max}^C(j) + d_{av}^C(j')$. Since $d(j, \mathcal{F}_j^C \cap \mathcal{F}_{j'}^C) \leq d_{max}^C(j)$, we have $d(j', \mathcal{F}_j^C \cap \mathcal{F}_{j'}^C) \geq d_{av}^C(j')$. If $d(j', \mathcal{F}_j^D \cap \mathcal{F}_{j'}^C) \geq d_{av}^C(j')$, then $d(j', \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) \leq d_{av}^C(j')$ and the lemma follows from the fact that $d(j, j') \leq d_{max}^C(j) + d_{max}^C(j') \leq (2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + d_{max}^C(j')$.

So, we can also assume

$$d(j', \mathcal{F}_j^D \cap \mathcal{F}_{j'}^C) = d_{av}^C(j') - z \tag{3}$$

for some positive z. Let $\hat{y} = \mathsf{vol}(\mathcal{F}_j^D \cap \mathcal{F}_{j'}^C)$. Notice that $\hat{y} \leq \max\{\gamma - 1, 1\}$. (3) implies

$$d(j', \mathcal{F}_{j'}^C \backslash \mathcal{F}_j^D) = d_{av}^C(j') + \frac{\hat{y}}{1-\hat{y}} z \tag{4}$$

From (2) and (3), we get

$$d(j, \mathcal{F}_j^D \cap \mathcal{F}_{j'}^C) \geq (2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + z = d_{av}^D(j) - (2-\gamma)(d_{av}^D(j) - d_{max}^C(j)) + z$$

This further implies

$$\begin{aligned} d_{max}^C(j) \quad &\leq \quad d(j, \mathcal{F}_j^D \backslash \mathcal{F}_{j'}^C) \leq d_{av}^D(j) - \frac{\hat{y}}{\gamma - 1 - \hat{y}} \left(z - (2-\gamma)(d_{av}^D(j) - d_{max}^C(j)) \right) \\ &\implies d_{av}^D(j) - d_{max}^C(j) \geq \frac{\hat{y}}{\gamma - 1 - \hat{y}} \left(z - (2-\gamma)(d_{av}^D(j) - d_{max}^C(j)) \right) \\ &\implies d_{av}^D(j) - d_{max}^C(j) \geq \frac{\hat{y}}{\gamma - 1 - \hat{y}} z \Big/ \left(1 + \frac{(2-\gamma)\hat{y}}{\gamma - 1 - \hat{y}} \right) \quad = \quad \frac{\hat{y} z}{(\gamma-1)(1-\hat{y})} \end{aligned} \tag{5}$$

Notice that we used the fact that $1 + \frac{(2-\gamma)\hat{y}}{\gamma - 1 - \hat{y}} \geq 0$. From (2) and (5), we get

$$\begin{aligned} d(j', \mathcal{F}_j^C \cap \mathcal{F}_{j'}^C) &\geq d(j,j') - d(j, \mathcal{F}_j^C \cap \mathcal{F}_{j'}^C) \geq (2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + d_{av}^C(j') - d_{max}^C(j) \\ &= (\gamma-1)(d_{av}^D(j) - d_{max}^C(j)) + d_{av}^C(j') \geq \frac{\hat{y}}{1-\hat{y}} z + d_{av}^C(j') \end{aligned}$$

Combining the above inequality and (4), we have

$$d(j', \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) \leq d_{av}^C(j') + \frac{\hat{y}}{1-\hat{y}} z$$

So,

$$\begin{aligned} d(j, \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) &\leq d_{max}^C(j) + d_{max}^C(j') + d(j', \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) \\ &\leq (2-\gamma) d_{max}^C(j) + (\gamma - 1)\left(d_{av}^D(j) - \frac{\hat{y} z}{(\gamma-1)(1-\hat{y})} \right) \\ &\quad + d_{max}^C(j') + d_{av}^C(j') + \frac{\hat{y}}{1-\hat{y}} z \\ &= (2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + d_{max}^C(j') + d_{av}^C(j') \end{aligned}$$

Lemma 2.

$$d(j, \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) \leq \gamma d_{av}(j) + (3-\gamma) d_{max}^C(j)$$

Proof. Noticing that $d_{max}^C(j') + d_{av}^C(j') \leq d_{max}^C(j) + d_{av}^C(j)$, the proof is straightforward.

$$\begin{aligned} d(j, \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) &\leq (2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + d_{max}^C(j') + d_{av}^C(j') \\ &\leq (2-\gamma) d_{max}^C(j) + (\gamma-1) d_{av}^D(j) + d_{max}^C(j) + d_{av}^C(j) \\ &= \gamma \left(\frac{1}{\gamma} d_{av}^C(j) + \frac{\gamma-1}{\gamma} d_{av}^D(j) \right) + (3-\gamma) d_{max}^C(j) \\ &= \gamma d_{av}(j) + (3-\gamma) d_{max}^C(j) \end{aligned}$$

For a client $j \in \mathcal{C}$, define $h_j : [0,1] \to \mathbb{R}^*$ to be the distribution of distances form j to $\mathcal{F}_j$ in the following way. Let $i_1, i_2, \cdots, i_m$ the facilities in $\mathcal{F}_j$, in the non-decreasing order of distances to j. Then $h_j(p) = d_{i_t, j}$, where t is the minimum number such that $\sum_{s=1}^{t} y_{i_s} \geq p$. Notice that h_j is defined using the y, not $\overline{y}$, and is thus independent of γ. Define $h(p) = \sum_{j \in \mathcal{C}} h_j(p)$. Observe the following facts :

1. h_js and h are non-decreasing functions;
2. $d_{av}(j) = \int_0^1 h_j(p)\mathbf{d}p, \qquad d_{av}^C(j) = \gamma \int_0^{1/\gamma} h_j(p)\mathbf{d}p,$

   $d_{max}^C(j) = h_j(1/\gamma), \qquad d_{av}^D(j) = \frac{\gamma}{\gamma-1} \int_{1/\gamma}^1 h_j(p)\mathbf{d}p.$

Lemma 3. *For any client j,*

$$\mathbb{E}(C_j) \leq \int_0^1 h_j(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} \left(\gamma \int_0^1 h_j(p)\mathbf{d}p + (3-\gamma) h_j\left(\frac{1}{\gamma}\right) \right)$$

Proof. Let $j' \in \mathcal{C}'$ be the cluster center of j. We connect j to the closest open facility in $\mathcal{F}_j \cup \mathcal{F}_{j'}^C$.

We can assume that facilities in $\mathcal{F}_j \backslash \mathcal{F}_{j'}^C$ are independently sampled; otherwise, the expected connection cost can only be smaller. Indeed, consider two distributions p_1 and p_2. The only difference between p_1 and p_2 is that, in p_1 two facilities i and i' are dependently sampled (with probability $\overline{y}_i$, i is open, with probability

$\overline{y}_{i'}$, i' is open, and with probability $1-\overline{y}_i-\overline{y}_{i'}$, none of them are open), while in p_2 they are independently sampled. W.L.O.G, assume $d(j,i) \leq d(j,i')$. We consider the distribution of the distance from j to the closest open facility in $\{i,i'\}$ (∞ if it does not exist). In p_1, the distribution is : with probability $\overline{y}_i$, we get $d(j,i)$; with probability $\overline{y}_{i'}$, we get $d(j,i')$ and with the remaining $1-\overline{y}_i-\overline{y}_{i'}$ probability, we get ∞. In p_2, the distribution is : with probability $\overline{y}_i$, we get $d(j,i)$, with probability $(1-\overline{y}_i)\overline{y}_{i'}$, we get $d(j,i')$ and with the remaining probability $(1-\overline{y}_i)(1-\overline{y}_{i'})$, we get ∞. So, the distribution for p_2 strictly dominates the distribution for p_1. So, the expected connection cost w.r.t p_2 is at least as large as the expected connection cost w.r.t p_1. This can be easily extended to more than 2 facilities.

Then, we perform the following sequence of operations :

1. Split the set $\mathcal{F}_{j'}^C$ into two subsets: $\mathcal{F}_{j'}^C \cap \mathcal{F}_j$, and $\mathcal{F}_{j'}^C \backslash \mathcal{F}_j$;
2. Scale up $\bar{y}$ values in $\mathcal{F}_{j'}^C \backslash \mathcal{F}_j$ so that the volume of $\mathcal{F}_{j'}^C \backslash \mathcal{F}_j$ becomes 1;
3. Assume $\mathcal{F}_{j'}^C \cap \mathcal{F}_j$ and $\mathcal{F}_{j'}^C \backslash \mathcal{F}_j$ are independently sampled.

We show that the sequence of operations do not change the expected connection cost of j. Indeed, consider distribution of the distance between j and the closest open facility in $\mathcal{F}_{j'}^C$. The distribution does not change after we performed the operations. This is true since $d_{max}(j, \mathcal{F}_{j'}^C \cap \mathcal{F}_j) \leq d_{min}(j, \mathcal{F}_{j'}^C \backslash \mathcal{F}_j)$, where d_{max} and d_{min} represents the maximum and the minimum distance from a client to a set of facilities, respectively.

Then again, we can assume facilities in $\mathcal{F}_{j'}^C \cap \mathcal{F}_j$ are independently sampled. Now, we are in the situation where, facilities in $\mathcal{F}_j$ are independently sampled, exact 1 facility in $\mathcal{F}_{j'}^C \backslash \mathcal{F}_j$ is open and the probabilities that facilities are open are proportional to their $\overline{y}$ values.

We split each facility $i \in \mathcal{F}_j$ into facilities with infinitely small $\overline{y}$ values. This can only increase the expected connection cost of j. Thus, the expected connection cost of j is upper bounded by

$$\begin{aligned}
&\int_0^1 h_j(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} d(j, \mathcal{F}_{j'}^C \backslash \mathcal{F}_j) \\
\leq &\int_0^1 h_j(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} \left(\gamma \int_0^1 h_j(p) \mathbf{d}p + (3-\gamma) h_j \left(\frac{1}{\gamma} \right) \right)
\end{aligned}$$

This concludes the proof.

Lemma 4. *The expected connection cost of the integral solution is*

$$\mathbb{E}(C) \leq \int_0^1 h(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} \left(\gamma \int_0^1 h(p) \mathbf{d}p + (3-\gamma) h \left(\frac{1}{\gamma} \right) \right)$$

Proof.

$$\begin{aligned}
\mathbb{E}(C) &\leq \sum_{j \in \mathcal{F}} \left(\int_0^1 h_j(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} \left(\gamma \int_0^1 h_j(p) \mathbf{d}p + (3-\gamma) h_j \left(\frac{1}{\gamma} \right) \right) \right) \\
&= \int_0^1 h(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} \left(\gamma \int_0^1 h(p) \mathbf{d}p + (3-\gamma) h \left(\frac{1}{\gamma} \right) \right)
\end{aligned}$$

3.2 An explicit distribution of γ

In this subsection, we give an explicit distribution of γ. This can be proved through a 0-sum game between an algorithm player and an h player.

For fixed h and γ, let's define

$$\alpha(\gamma, h) = \int_0^1 h(p) e^{-\gamma p} \gamma \mathbf{d}p + e^{-\gamma} \left(\gamma \int_0^1 h(p) \mathbf{d}p + (3-\gamma) h \left(\frac{1}{\gamma} \right) \right)$$

We can scale h so that $\int_0^1 h(p)\mathbf{d}p = 1$. Then,

$$\alpha(\gamma, h) = \int_0^1 h(p)e^{-\gamma p}\gamma\mathbf{d}p + e^{-\gamma}\left(\gamma + (3-\gamma)h\left(\frac{1}{\gamma}\right)\right)$$

Then, for a fixed h, the algorithm $A1(\gamma)$ will give a $(\gamma, \alpha(h,\gamma))$-bi-factor approximation. We consider a 0-sum game between an algorithm player A and an adversary B. The mixed strategy of player A is a pair (μ, θ), where $0 \le \theta \le 1$ and μ is $1-\theta$ times a probability density function for γ. So, $\theta + \int_1^\infty \mu(\gamma)\mathbf{d}\gamma = 1$. This strategy corresponds to running algorithm $A2$ with probability θ and running algorithm $A1(\gamma)$ with probability $\mu(\gamma)\mathbf{d}\gamma$. The mixed strategy for player B is a monotone non-negative function h over $[0,1]$ such that $\int_0^1 h(p)\mathbf{d}p = 1$. The value of the game is

$$\nu(\mu, \theta, h) = \max\left\{\int_1^2 \gamma\mu(\gamma)\mathbf{d}\gamma + 1.11\theta, \int_1^2 \alpha(\gamma, h)\mu(\gamma)\mathbf{d}\gamma + 1.78\theta\right\}$$

Then, our goal becomes finding a strategy for A that minimizes the value. For a fixed strategy (θ, μ) of player A, the best strategy of player B is a threshold function h_q, for some $0 \le q < 1$, where

$$h_q(p) = \begin{cases} 0 & p < q \\ \frac{1}{1-q} & p \ge q \end{cases}$$

In order to obtain the value of this game, we discretize the domain of μ into many small intervals divided by points $\{\gamma_i = 1 + i/n : 0 \le i \le n\}$, and restrict μ to be non-zero only in the center of each interval. Thus, the value of the game is approximately characterized by the following LP.

$$\begin{aligned} \min \beta \quad \text{s.t} \\ \frac{1}{n}\sum_{i=1}^n x_i + \theta \ge 1 \\ \frac{1}{n}\sum_{i=1}^n \frac{\gamma_{i-1}+\gamma_i}{2} x_i + 1.11\theta \le \beta \\ \frac{1}{n}\sum_{i=1}^n \alpha\left(\frac{\gamma_{i-1}+\gamma_i}{2}, h_q\right) x_i + 1.78\theta \le \beta \ \ \forall q \in [0,1) \\ x_1, x_2, \cdots, x_n, \theta \ge 0 \end{aligned}$$

Solving the above LP for $n = 500$ using Matlab, we get a mixed strategy for player A that achieves value 1.4879. The strategy is roughly of the following form. θ is about 0.2; μ has an impulse point at $\gamma \approx 1.5$, with impulse value 0.5; the remaining 0.3 weight is distributed between $\gamma = 1.5$ and $\gamma = 2$.

In light of the program generated solution, we give an analytical solution for the strategy of player A and show that the value of the game is at most 1.488. With probability θ_2, we run algorithm $A2$; with probability θ_1, we run algorithm $A1(\gamma)$ with $\gamma = \gamma_1$; and with probability $1 - \theta_2 - \theta_1$, we run algorithm $A1(\gamma)$ with γ randomly chosen between γ_1 and γ_2. So, the function μ is

$$\mu(p) = \theta_1\delta(p - \gamma_1) + aI_{\gamma_1,\gamma_2}(p)$$

where δ is the Dirac delta function, $a = \frac{1-\theta_1-\theta_2}{\gamma_2-\gamma_1}$, and I_{γ_1,γ_2} is the 1 if $\gamma_1 < \gamma < \gamma_2$ and 0 otherwise. (See figure 2.) The values of $\theta_1, \theta_2, \gamma_1, \gamma_2, a$ will be determined later.

The scaling factor for the facility cost will be

$$\lambda_f = \theta_1\lambda_1 + a(\gamma_2 - \gamma_1)\frac{\gamma_1+\gamma_2}{2} + 1.11\theta_2$$

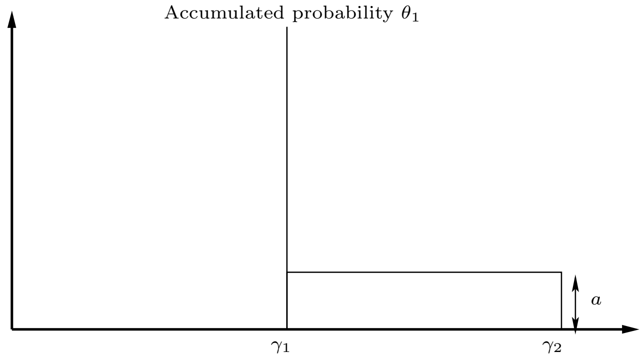

Fig. 2. The distribution of γ. With probability θ_1, we run algorithm $A1(\gamma_1)$; with probability θ_2, we run algorithm $A2$; with probability $1-\theta_1-\theta_2 = a(\gamma_2-\gamma_1)$, we run algorithm $A2(\gamma)$ with γ randomly selected from $[\gamma_1,\gamma_2]$.

Now, we consider the scaling factor λ_c for the connection cost. It's enough to consider the case where $h = h_q$ for some $0 \le q < 1$. When $h = h_q$,

$$\begin{aligned}
\lambda_c(q) &= \int_1^\infty \left(\int_0^1 e^{-\gamma p}\gamma h_q(p)\mathbf{d}p + e^{-\gamma}(\gamma+(3-\gamma)h_q(1/\gamma))\right)\mu(\gamma)\mathbf{d}\gamma + 1.78\theta_2 \\
&= \int_{\gamma_1}^{\gamma_2}\int_q^1 e^{-\gamma p}\gamma\frac{1}{1-q}\mathbf{d}pa\mathbf{d}\gamma + \int_{\gamma_1}^{\gamma_2} e^{-\gamma}\gamma a\mathbf{d}\gamma + \int_{\gamma_1}^{\gamma_2}(3-\gamma)h_q(1/\gamma)a\mathbf{d}\gamma \\
&\quad +\theta_1\int_q^1 e^{-\gamma_1 p}\gamma_1\frac{1}{1-q}\mathbf{d}p + \theta_1 e^{-\gamma_1}(\gamma_1+(3-\gamma_1)h_q(1/\gamma_1)) + 1.78\theta_2 \\
&= B_1(q)+B_2(q)+B_3(q)+1.78\theta_2
\end{aligned}$$

where

$$\begin{aligned}
B_1(q) &= \int_{\gamma_1}^{\gamma_2}\int_q^1 e^{-\gamma p}\gamma\frac{1}{1-q}\mathbf{d}pa\mathbf{d}\gamma + \int_{\gamma_1}^{\gamma_2} e^{-\gamma}\gamma a\mathbf{d}\gamma \\
&= \frac{a}{1-q}\int_{\gamma_1}^{\gamma_2}(e^{-\gamma q}-e^{-\gamma})\mathbf{d}\gamma - a(\gamma+1)e^{-\gamma}\Big|_{\gamma_1}^{\gamma_2} \\
&= \frac{a}{(1-q)q}(e^{-\gamma_1 q}-e^{-\gamma_2 q}) - \frac{a}{1-q}(e^{-\gamma_1}-e^{-\gamma_2}) + a((\gamma_1+1)e^{-\gamma_1}-(\gamma_2+1)e^{-\gamma_2})
\end{aligned}$$

$$\begin{aligned}
B_2(q) &= \int_{\gamma_1}^{\gamma_2}(3-\gamma)h_q(1/\gamma)a\mathbf{d}\gamma \\
&= \begin{cases}
\int_{\gamma_1}^{\gamma_2}(3-\gamma)\frac{1}{1-q}a\mathbf{d}\gamma = \frac{a}{1-q}\left((2-\gamma_1)e^{-\gamma_1}-(2-\gamma_2)e^{-\gamma_2}\right) & 0 \le q < 1/\gamma_2 \\
\int_{\gamma_1}^{1/q}(3-\gamma)\frac{1}{1-q}a\mathbf{d}\gamma = \frac{a}{1-q}\left((2-\gamma_1)e^{-\gamma_1}-(2-1/q)e^{-1/q}\right) & 1/\gamma_2 \le q \le 1/\gamma_1 \\
0 & 1/\gamma_1 < q < 1
\end{cases}
\end{aligned}$$

$$\begin{aligned}
B_3(q) &= \theta_1\int_q^1 e^{-\gamma_1 p}\gamma_1\frac{1}{1-q}\mathbf{d}p + \theta_1 e^{-\gamma_1}(\gamma_1+(3-\gamma_1)h_q(1/\gamma_1)) \\
&= \theta_1\left(\frac{1}{1-q}(e^{-\gamma_1 q}-e^{-\gamma_1}) + e^{-\gamma_1}\gamma_1 + e^{-\gamma_1}(3-\gamma_1)h_q(1/\gamma_1)\right) \\
&= \begin{cases}
\theta_1\left(\frac{1}{1-q}(e^{-\gamma_1 q}-e^{-\gamma_1}) + e^{-\gamma_1}\gamma_1 + \frac{e^{-\gamma_1}(3-\gamma_1)}{1-q}\right) & 0 \le q \le 1/\gamma_1 \\
\theta_1\left(\frac{1}{1-q}(e^{-\gamma_1 q}-e^{-\gamma_1}) + e^{-\gamma_1}\gamma_1\right) & 1/\gamma_1 < q < 1
\end{cases}
\end{aligned}$$

So, we have 3 cases :

1. $0 \le q < 1/\gamma_2$

$$\begin{aligned}
\lambda_c(q) &= \frac{a}{(1-q)q}(e^{-\gamma_1 q} - e^{-\gamma_2 q}) - \frac{a}{1-q}(e^{-\gamma_1} - e^{-\gamma_2}) \\
&\quad + a((\gamma_1+1)e^{-\gamma_1} - (\gamma_2+1)e^{-\gamma_2}) + \frac{a}{1-q}\left((2-\gamma_1)e^{-\gamma_1} - (2-\gamma_2)e^{-\gamma_2}\right) \\
&\quad + \theta_1(\frac{1}{1-q}(e^{-\gamma_1 q} - e^{-\gamma_1}) + e^{-\gamma_1}\gamma_1 + \frac{1}{1-q}e^{-\gamma_1}(3-\gamma_1)) + 1.78\theta_2 \\
&= \frac{a}{(1-q)q}(e^{-\gamma_1 q} - e^{-\gamma_2 q}) + \frac{A_1}{1-q} + \theta_1\frac{e^{-\gamma_1 q}}{1-q} + A_2 \\
\text{where } A_1 &= a(e^{-\gamma_1} - \gamma_1 e^{-\gamma_1} - e^{-\gamma_2} + \gamma_2 e^{-\gamma_2}) + 2\theta_1 e^{-\gamma_1} - \theta_1 e^{-\gamma_1}\gamma_1 \\
A_2 &= a((\gamma_1+1)e^{-\gamma_1} - (\gamma_2+1)e^{-\gamma_2}) + \theta_1 e^{-\gamma_1}\gamma_1 + 1.78\theta_2
\end{aligned}$$

2. $1/\gamma_2 \le q \le 1/\gamma_1$

$$\begin{aligned}
\lambda_c(q) &= \frac{a}{(1-q)q}(e^{-\gamma_1 q} - e^{-\gamma_2 q}) + \frac{A_1}{1-q} + \theta_1\frac{e^{-\gamma_1 q}}{1-q} + A_2 \\
&\quad + \frac{a}{1-q}\left((2-\gamma_2)e^{-\gamma_2} - (2-1/q)e^{-1/q}\right)
\end{aligned}$$

3. $1/\gamma_1 < q < 1$

$$\begin{aligned}
\lambda_c(q) &= \frac{a}{(1-q)q}(e^{-\gamma_1 q} - e^{-\gamma_2 q}) - \frac{a}{1-q}(e^{-\gamma_1} - e^{-\gamma_2}) \\
&\quad + a((\gamma_1+1)e^{-\gamma_1} - (\gamma_2+1)e^{-\gamma_2}) + \theta_1(\frac{1}{1-q}(e^{-\gamma_1 q} - e^{-\gamma_1}) + e^{-\gamma_1}\gamma_1) + 1.78\theta_2 \\
&= \frac{a}{(1-q)q}(e^{-\gamma_1 q} - e^{-\gamma_2 q}) + \frac{A_3}{1-q} + \theta_1\frac{e^{-\gamma_1 q}}{1-q} + A_2 \\
\text{where } A_3 &= a(-e^{-\gamma_1} + e^{-\gamma_2}) - \theta_1 e^{-\gamma_1}
\end{aligned}$$

Let $\gamma_1 = 1.479311, \gamma_2 = 2.016569, \theta_1 = 0.503357, \theta_2 \approx 0.195583, a = 0.560365$. Then,

$$\lambda_f = \theta_1\lambda_1 + a(\gamma_2 - \gamma_1)\frac{\gamma_1 + \gamma_2}{2} + 1.11\theta_2 \approx 1.487954$$

and $\lambda_c(q)$ has maximum value about 1.487989, achieved at $q = 0$(see figure 3).

This concludes our proof.

References

1. Jaroslaw Byrka. An optimal bifactor approximation algorithm for the metric uncapacitated facility location problem. In *APPROX '07/RANDOM '07: Proceedings of the 10th International Workshop on Approximation and the 11th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 29–43, Berlin, Heidelberg, 2007. Springer-Verlag.
2. Jaroslaw Byrka, MohammadReza Ghodsi, and Aravind Srinivasan. Lp-rounding algorithms for facility-location problems. *in arxiv1007.3611*, 2010.
3. Moses Charikar and Sudipto Guha. Improved combinatorial algorithms for the facility location and k-median problems. In *In Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 378–388, 1999.
4. Fabián A. Chudak and David B. Shmoys. Improved approximation algorithms for the uncapacitated facility location problem. *SIAM J. Comput.*, 33(1):1–25, 2004.

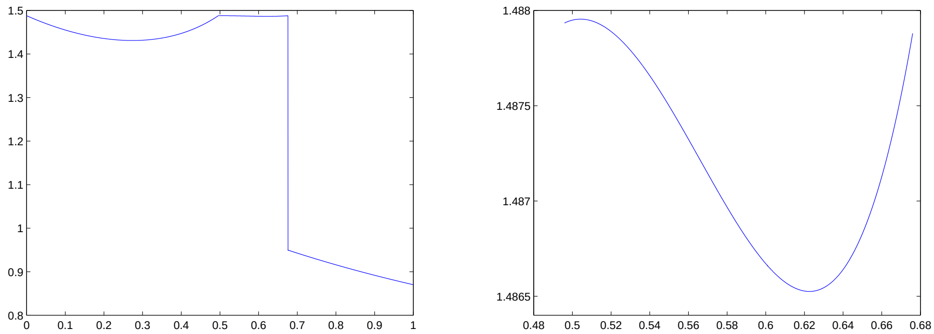

Fig. 3. The function $\lambda_c(q)$. The curve on the right-hand side is the function restricted to the interval $(1/\gamma_2, 1/\gamma_1)$.

5. D.S.Hochbaum. Heuristics for the fixed cost median problem. *Mathematical Programming*, 22:148–162, 1982.
6. Sudipto Guha and Samir Khuller. Greedy strikes back: Improved facility location algorithms. In *Journal of Algorithms*, pages 649–657, 1998.
7. Kamal Jain, Mohammad Mahdian, Evangelos Markakis, Amin Saberi, and Vijay V. Vazirani. Greedy facility location algorithms analyzed using dual fitting with factor-revealing lp. *J. ACM*, 50:795–824, November 2003.
8. Kamal Jain, Mohammad Mahdian, and Amin Saberi. A new greedy approach for facility location problems. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, STOC '02, pages 731–740, New York, NY, USA, 2002. ACM.
9. Madhukar R. Korupolu, C. Greg Plaxton, and Rajmohan Rajaraman. Analysis of a local search heuristic for facility location problems. In *Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*, SODA '98, pages 1–10, Philadelphia, PA, USA, 1998. Society for Industrial and Applied Mathematics.
10. Jyh-Han Lin and Jeffrey Scott Vitter. Approximation algorithms for geometric median problems. *Inf. Process. Lett.*, 44:245–249, December 1992.
11. Mohammad Mahdian, Yinyu Ye, and Jiawei Zhang. Approximation algorithms for metric facility location problems. *SIAM J. Comput.*, 36(2):411–432, 2006.
12. David B. Shmoys, Éva Tardos, and Karen Aardal. Approximation algorithms for facility location problems (extended abstract). In *STOC '97: Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 265–274, New York, NY, USA, 1997. ACM.