On the Connected Minimum Sum of Radii Problem

Hyung-Chan An
Yonsei University, Seoul, Republic of Korea

Mong-Jen Kao
National Yang-Ming Chiao-Tung University, Hsinchu, Taiwan

Abstract

In this paper, we consider the study for the *connected minimum sum of radii* problem. In this problem, we are given as input a metric defined on a set of *facilities* and *clients*, along with some *cost parameters*. The objective is to open a subset of facilities, assign every client to an open facilitiy, and connect open facilities using a Steiner tree so that the weighted (by cost parameters) sum of the maximum assignment distance of each facility and the Steiner tree cost is minimized. This problem introduces the *min-sum radii objective*, an objective function that is widely considered in the clustering literature, to the connected facility location problem, a well-studied network design/clustering problem. This problem is useful in communication network design on a shared medium, or energy optimization of mobile wireless chargers.

We present both a constant-factor approximation algorithm and hardness results for this problem. Our algorithm is based on rounding an LP relaxation that jointly models the min-sum of radii problem and the rooted Steiner tree problem. To round the solution we use a careful clustering procedure that guarantees that every open facility has a *proxy client* nearby. This allows a reinterpretation for part of the LP solution as a fractional rooted Steiner tree. Combined with a cost filtering technique, this yields a 5.542-approximation algorithm.

2012 ACM Subject Classification Theory of computation → Facility location and clustering

Keywords and phrases connected minimum sum of radii, minimum sum of radii, connected facility location, approximation algorithms, Steiner trees

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2024.7

Funding This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. RS-2021-II212068, Artificial Intelligence Innovation Hub). This work was supported by the National Science and Technology Council (NSTC), Taiwan, grants 112-2628-E-A49-017-MY3, 113-2628-E-A49-017-MY3, and 113-2634-F-A49-001-MBK.

Acknowledgements The authors made equal contributions. We thank the anonymous reviewers for their helpful comments.

1 Introduction

Connected facility location is a joint optimization problem that combines network design with clustering, and it has wide applications in the design of communication networks [1, 8, 23]. In this problem, we are given as input a metric on a set of nodes (some of which are called *facilities* and some *clients*) in addition to the *opening cost* of each facility and a *connectivity cost parameter* M. The goal is to open some facilities, assign every client to an open facility, and finally connect the open facilities with a Steiner tree whose terminals are the open facilities. The cost of a solution is defined as the total assignment distance between each client and the facility it is assigned to, plus the total opening costs of the open facilities and the cost of the Steiner tree scaled by M. This problem is particularly useful in the design of a communication network where a *central core* is formed by connecting *core nodes* together and individual *endnodes* are assigned to one of the core nodes [1, 8, 23]. There exists an extensive volume of research on this problem: in addition to the problem described

© Hyung-Chan An and Mong-Jen Kao;
licensed under Creative Commons License CC-BY 4.0

above [8, 10, 14, 23], a variant where opening a facility itself does not incur opening cost has also been investigated [15, 23]. In addition to the facility location problem, connectivity constraints have been introduced in other classical problems including dominating set (see, e.g., [7, 13, 21, 22]).

The facility location problem is a *min-sum* optimization problem, in that it minimizes the *sum* of the assignment distances. Yet, this is not always the objective one is most interested in in practice. For example, when the endnodes are connected via broadcasting on a shared medium, one may be more interested in the longest assignment distance [11]. On the extreme in this direction is the *k-center* problem [17], which minimizes the maximum assignment distance in the *entire* solution. This unfortunately tends to yield a less desirable clustering since the maximum assignment distance by itself determines the objective. In order to avoid the *dissection effect* [16], we can use the sum of radii in lieu of maximum radii as the objective function [6]. Under this objective, however, there exists a trivial optimal solution when there is no distinction between facilities and clients – one can open and assign every node to itself, resulting in the zero sum of radii – and therefore the min-sum radii problem was previously studied usually under a cardinality constraint on the number of open facilities [2, 3, 6, 9, 11].

However, the connected facility location problem was little studied under this min-sum radii objective. This paper proposes to study the *connected minimum sum of radii* problem, aiming at addressing this gap. Our problem takes as input the *assignment cost parameter* of each facility in addition to the *connectivity cost parameter* M and a metric on facilities and clients. The goal still is to open some facilities, assign every client to an open facility, and connect the open facilities. The problem however differs from connected facility location in its objective function, which is now defined as the sum of radii, i.e., the sum of the longest assignment distance of each facility, plus the Steiner tree cost connecting the open facilities. The radii and the Steiner tree cost are respectively scaled by the assignment cost parameters and the connectivity cost parameter.

A sample application that well illustrates this problem is wireless charging of sensors. Consider a set of sensors distributed over a region, which are charged by a wireless charger that moves between charging spots to charge near sensors [19, 24]. The wireless charging energy is proportional to the maximum distance to a sensor being charged, and the proposed problem well reflects this setting. The connected minimum sum of radii problem also arises when we want to broadcast messages to a set of sensors. Suppose we install a set of mutually connected stations each of which broadcasts messages over the air to nearby sensors. The total communication cost will then depend on the over-the-air broadcast range of each station and their mutual connection cost.

Our results and techniques

In this paper, we propose to study the connected minimum sum of radii problem, present an approximation algorithm for it, and show its NP-hardness. Our main result is the following theorem. While this paper primarily considers the version of the problem that opening a facility itself does not incur a fixed opening cost, Theorem 1 immediately extends to the version with opening cost as well, without affecting the final approximation ratio.

▶ **Theorem 1.** *There is a polynomial-time algorithm that computes a* 5.542*-approximation solution for the connected minimum sum of radii problem.*

The algorithm we present is an LP-rounding algorithm that is partially based on a greedy clustering of fractionally open facilities. Greedy clustering approach was previously used to handle the (non-connected) minimum sum of radii problem [9]. In this paper, we propose that we use a carefully designed new clustering procedure to ensure that each open facility always has a “proxy client” nearby.

After clustering, the LP solution can be reinterpreted as a fractional solution to a rooted Steiner tree instance whose terminals are the proxy clients, at the expense of a slight increase in the cost. This fractional solution is then rounded using any LP-based algorithm with a good approximation ratio for the Steiner tree problem, such as the LP-rounding algorithm of Jain [18] or the primal-dual algorithm of Goemans and Williamson [12]. Finally, to obtain the desired approximation ratio, we compare this solution against a trivial solution that opens a single guessed facility, and output the better between the two solutions.

We will complement the above result by showing that the problem is NP-hard.

▶ **Theorem 2.** *The connected minimum sum of radii problem is NP-hard.*

Organization of this paper

The rest of this paper is organized as follows. In Section 2 we provide a formal definition of the connected minimum sum of radii problem and the notation we will be using throughout this paper. In Section 3 we present our approximation algorithm. We establish the approximation guarantee in Section 4 and present the hardness results in Section 5.

2 Preliminaries

We begin with a formal definition on the connected minimum sum of radii problem. In this problem, we are given a set F of facilities, a set D of clients, a distance metric d defined over $F \cup D$ and two additional parameters $m\colon F \to \mathbb{Q}_{\geq 0}$ and $M \in \mathbb{Q}_{\geq 0}$.

A feasible solution consists of a tuple (S, ρ, T), where $S \subseteq F$ is a subset of facilities, $\rho\colon S \to \mathbb{Q}_{\geq 0}$ is the set of respective radii for the facilities in S such that all clients in D are covered, i.e., for any $j \in D$, there always exists some $i \in S$ such that $d(j,i) \leq \rho_i$, and T is a Steiner tree with terminal set S and Steiner nodes $F \cup D$.

The objective is to minimize the sum of radii of the clusters in S, each weighted by the parameters $m_i\,|_{i\in S}$, plus the total length of the Steiner tree weighted by M, i.e.,

$$\sum_{i\in S} m_i \cdot \rho_i \;+\; \sum_{e\in T} M \cdot d_e.$$

Note that, provided that $M \neq 0$, we may assume without loss of generality that $M = 1$, for otherwise we can scale m_i for all $i \in F$ uniformly. For the rest of this paper we will take this assumption that $M = 1$.

We also note that, our algorithm and the analysis can be modified in a straightforward way to work for the extreme case that $M = 0$.

Notations

We additionally use the following notation in this paper. We use $V := F \cup D$ to denote the set of vertices in the given metric space and $E := \{\,(u,v) \mid u,v \in V\,\}$ to denote the set of possible edges when considering the corresponding metric graph. For any $U \subseteq V$, we use $\delta(U)$ to denote the set of edges in the cut $(U, \bar{U})$ with respect to the metric graph.

For any $i \in F$ and any $r \in \mathbb{Q}_{\geq 0}$, we use $B(i,r)$ to denote the set of clients that belong in the ball centered at i with radius r, i.e,

$$B(i,r) := \{\, j \in D \;\mid\; d(i,j) \leq r \,\}.$$

For each $i \in F$, we use $R_i := \{d(i,j) \mid j \in D\}$ to denote the set of "meaningful" radii for i.

3 Approximation Algorithm

In this section, we present our algorithm for the connected minimum sum of radii problem. Let $\mu \geq 1$ be a parameter to be determined.

Our algorithm starts by guessing a facility $t \in F$ that is opened in an optimal solution $(S^{\mathsf{opt}}, \rho^{\mathsf{opt}}, T^{\mathsf{opt}})$ with the minimum m_t value, i.e., $t = \operatorname{argmin}_{i \in S^{\mathsf{opt}}} m_i$. For each candidate guess t, the algorithm generates two solutions $(S^{\mathrm{I}}_t, \rho^{\mathrm{I}}_t, T^{\mathrm{I}}_t)$ and $(S^{\mathrm{II}}_t, \rho^{\mathrm{II}}_t, T^{\mathrm{II}}_t)$. When this process ends, the one with the smallest cost is output as the approximation solution. In the following we describe how the solutions are generated for each guess $t \in F$. To simplify the notations, the dependency on t will be omitted when there is no ambiguity in the context.

The first solution $(S^{\mathrm{I}}, \rho^{\mathrm{I}}, T^{\mathrm{I}})$ is a trivial one with $S^{\mathrm{I}} := \{t\}$, i.e., t is the only open facility. Naturally, $T^{\mathrm{I}} = \emptyset$ and $\rho^{\mathrm{I}}_t = \max_{j \in D} d(t,j)$ in this solution. The cost of this solution is hence $m_t \cdot \max_{j \in D} d(t,j)$.

To obtain the second solution $(S^{\mathrm{II}}, \rho^{\mathrm{II}}, T^{\mathrm{II}})$, let $F_\mu := \{i \in F \mid m_i \geq \mu\}$. We use the following LP relaxation for a further restricted scenario for which $S^{\mathsf{opt}} \subseteq F_\mu$, i.e., the (unknown) referenced optimal solution only uses facilities in F_μ. This unusual setting will become clear in the analysis.

$$\begin{aligned}
\text{minimize} \quad & \sum_{i \in F_\mu, r \in R_i} m_i \cdot r \cdot x_{i,r} + \sum_{e \in E} d_e \cdot y_e \\
\text{subject to} \quad & \sum_{i \in F_\mu} z_{i,j} \geq 1, && \forall j \in D, \\
& \sum_{r \in R_i : j \in B(i,r)} x_{i,r} \geq z_{i,j}, && \forall i \in F_\mu, j \in D, \\
& \sum_{e \in \delta(U)} y_e \geq \sum_{i \in F_\mu \cap U} z_{i,j}, && \forall j \in D, U \subseteq V \setminus \{t\}, \qquad (1) \\
& x, y, z \geq 0.
\end{aligned}$$

We have three sets of indicator variables in the above LP.

- $x_{i,r}$ for each (i,r) pair with $i \in F_\mu$ and $r \in R_i$.
- y_e for each edge $e \in E$.
- $z_{i,j}$ for the assignment of client $j \in D$ to the facility $i \in F_\mu$.

The first constraint requires that any client in D has to be assigned to at least one facility in F_μ. The second constraint demands that, in order for a client j to be assigned to facility i, j must be contained in an opened ball centered at i. The third constraint models the connectivity requirement between the opened facilities via the assignment variables $z_{i,j}$ and the predetermined sink t. Note that the constraints of this LP does not require that t is opened but rather use it to ensure the connectivity between the opened facilities.

Note that the last set of inequalities can be separated by finding a minimum j-t cut. We solve the LP in polynomial time to obtain an optimal fractional solution $(x^\star, y^\star, z^\star)$. In the following we describe our rounding procedure to obtain the second solution $(S^{\mathrm{II}}, \rho^{\mathrm{II}}, T^{\mathrm{II}})$. The rounding procedure consists of two parts. In the first part, we select a set of facilities along with their respective radii to be opened. In the second part, we compute a Steiner tree for the opened facilities.

Opening facilities

Let $\mathcal{B}_0 := \{(i,r) \mid x^\star_{i,r} > 0\}$ be the support of $x^\star$. Let $\mathcal{G}$ be a bipartite graph with partite sets $\mathcal{B}_0$ and D, where $(i,r) \in \mathcal{B}_0$ and $j \in D$ are adjacent if and only if $j \in B(i,r)$. For any $j \in D$ and any $B^\star \subseteq \mathcal{B}_0$, let $\Delta_{\mathcal{B}^\star}(j)$ denote the minimum distance between j and any vertex in $B^\star$: i.e.,

$$\Delta_{\mathcal{B}^\star}(j) := \min\{|P| \mid P \text{ is a path in } \mathcal{G} \text{ between } j \text{ and some } y \in \mathcal{B}^\star\},$$

where $|P|$ denotes the number of edges on P. Note that we define $\Delta_\emptyset(j) := +\infty$ for all $j \in D$.

Algorithm 1 Determining open facilities and their proxy clients.

1: $S^{\text{II}} \leftarrow \emptyset$; $\mathcal{B}^\star \leftarrow \emptyset$
2: **while** $\exists j \in D$ with $\Delta_{\mathcal{B}^\star}(j) \geq 5$ **do**
3: $\quad \bar{\mathcal{B}} := \{(i,r) \in \mathcal{B}_0 \mid$ there exists some j that is adjacent to (i,r) and $\Delta_{\mathcal{B}^\star}(j) \geq 5\}$
4: $\quad (i^\star, r^\star) \in \arg\max_{(i,r)\in\bar{\mathcal{B}}} r$
5: $\quad$ let $\pi_{i^\star}$ be some $j \in D$ such that (i,r) and j are adjacent and $\Delta_{\mathcal{B}^\star}(j) \geq 5$
6: $\quad \mathcal{B}^\star \leftarrow \mathcal{B}^\star \cup \{(i^\star, r^\star)\}$; $S^{\text{II}} \leftarrow S^{\text{II}} \cup \{i^\star\}$; $\rho^{\text{II}}_{i^\star} = 3r^\star$

Consider Algorithm 1 that returns S^{II}, ρ^{II}, and $\{\pi_i\}_{i\in S^{\text{II}}}$. It additionally maintains $\mathcal{B}^\star$, which denotes the set of (i,r) pairs to be rounded up, and π_i for each $i \in S^{\text{II}}$ which denotes the representative *proxy client* we pick for facility i. It is to ensure the existence of these proxy clients why we use Algorithm 1 as opposed to a simple greedy clustering.

Initially, $S^{\text{II}} := \emptyset$ and $\mathcal{B}^\star := \emptyset$. In each iteration, the algorithm considers the set of (i,r) pairs in $\mathcal{B}_0$ that are adjacent to some client $j \in D$ in $\mathcal{G}$ with $\Delta_{\mathcal{B}^\star}(j) \geq 5$. Among all such (i,r) pairs, the algorithm picks the one with the largest r. Let the pair be (i^*, r^*) and let π_{i^*} be the witness client with $j \in D$ with $\Delta_{\mathcal{B}^\star}(j) \geq 5$.

The algorithm puts i^* in S^{II}, sets $\rho^{\text{II}}_{i^*}$ to be $3r^\star$, and adds (i^*, r^*) to $\mathcal{B}_0$. Then the algorithm iterates until $\Delta_{\mathcal{B}^\star}(j) < 5$ holds for all $j \in D$.

The following two observations show that this algorithm is well-defined. First, Observation 3 shows that the set $\bar{\mathcal{B}}$ at Step 3 of Algorithm 1 is always nonempty.

▶ **Observation 3.** *For all $j \in D$, there exists some $(i,r) \in \mathcal{B}_0$ such that $j \in B(i,r)$.*

Proof. From the feasibility of $(x^\star, y^\star, z^\star)$, there exists some $i \in F_\mu$ such that $z^\star_{i,j} > 0$, and this in turn implies that there exists some $r \in R_i$ such that $j \in B(i,r)$ and $x^\star_{i,r} > 0$. ◀

The following observation shows that ρ^{II} is unambiguously defined by Algorithm 1.

▶ **Observation 4.** *Step 4 of Algorithm 1 never chooses the same facility more than once.*

Proof. Suppose towards contradiction that Step 4 chooses (i, r_1) at some point and (i, r_2) at a later point during the execution of Algorithm 1 for some $r_1 \neq r_2$.

Suppose $r_1 < r_2$. Consider the moment the algorithm chooses (i, r_1). This implies that there exists some $j \in B(i, r_1)$ such that $\Delta_{\mathcal{B}^\star}(j) \geq 5$; since $B(i, r_1) \subseteq B(i, r_2)$, this implies $(i, r_2) \in \bar{\mathcal{B}}$, a contradiction to the design of the algorithm.

Suppose $r_1 > r_2$. Consider the moment the algorithm chooses (i, r_2). Since $B(i, r_2) \subseteq B(i, r_1)$ and $(i, r_1) \in \mathcal{B}^\star$, we have $\Delta_{\mathcal{B}^\star}(j) \leq 1$. Hence $(i, r_2) \notin \bar{\mathcal{B}}$ and cannot be picked in Step 4. ◀

The following lemma summarizes one of the key properties our algorithm aims to have.

▶ **Lemma 5.** *For any* $(i^\star, r^\star) \in \mathcal{B}^\star$ *and any* $(i', r') \in \mathcal{B}_0$ *such that* $\pi_{i^\star}$ *and* (i', r') *are adjacent in* $\mathcal{G}$*, we have* $r' \leq r^\star$*.*

Proof. Consider the moment the algorithm chooses $(i^\star, r^\star)$. Step 5 of the algorithm guarantees that $\Delta_{\mathcal{B}^\star}(\pi_{i^\star}) \geq 5$ and therefore $(i', r') \in \bar{\mathcal{B}}$. Since the algorithm chose $(i^\star, r^\star)$ over (i', r'), it shows that $r' \leq r^\star$. ◀

Connecting the opened facilities

To obtain the second solution $(S^{\mathrm{II}}, \rho^{\mathrm{II}}, T^{\mathrm{II}})$, it remains to build the Steiner tree T^{II}. Consider the following LP relaxation for the Steiner tree problem with vertex set $V := F \cup D$, edge set E, terminal set $W \subseteq V$, and a given root $t \in W$.

$$\begin{aligned} \text{minimize} \quad & \sum_{e \in E} d_e \cdot h_e \\ \text{subject to} \quad & \sum_{e \in \delta(U)} h_e \geq 1, \qquad \forall U \subseteq V \setminus \{t\} \text{ with } U \cap W \neq \emptyset, \\ & h \geq 0. \end{aligned} \tag{2}$$

Note that the costs of the edges in this relaxation are defined by d_e. In the following we construct a feasible solution for the above LP relaxation, where the set of terminals W is chosen as the set of proxy clients $\{\pi_i \mid i \in S^{\mathrm{II}}\}$.

We can assume without loss of generality on the variables $z^\star$ that, for all $j \in D$, $\sum_{i \in F_\mu} z^\star_{i,j} = 1$ for otherwise we can scale down $z^\star_{i,j}$ for all $i \in F_\mu$ simultaneously to make it so without losing the feasibility of the resulting solution. Construct a vector $p \in \mathbb{R}^E$ by setting

$$p_{(\pi_i, i')} := \begin{cases} z^\star_{i', \pi_i}, & \text{for all } i \in S^{\mathrm{II}} \text{ and } i' \in F_\mu, \\ 0, & \text{otherwise.} \end{cases}$$

Intuitively, in the above construction we fractionally wire π_i for each $i \in S^{\mathrm{II}}$ to all the facilities that fractionally covers π_i in $z^\star$. Since all the facilities are fractionally connected to the sink t in $y^\star$ by the LP constraint (1), it follows by the above construction that $y^\star + p$ fractionally connects the representative proxy client π_i to t for all $i \in S^{\mathrm{II}}$. Hence $y^\star + p$ is a feasible solution to (2).

Although we can use any LP-based algorithm for the Steiner tree problem at this point, let us assume that we use the LP-rounding algorithm of Jain [18] on this solution to construct a Steiner tree T_{pre} for the set of representative proxy clients in $W := \{\pi_i \mid i \in S^{\mathrm{II}}\}$. To obtain the desired Steiner tree T^{II}, we add edges (i, π_i) for all $i \in S^{\mathrm{II}}$ to T_{pre}.

The following lemma, which formally verifies the feasibility of $y^\star + p$ for (2), shows that the algorithm for this part is also well-defined, and a valid Steiner tree for W is produced.

▶ **Lemma 6.** $y^\star + p$ *is a feasible with respect to the constraint* (2)*.*

Proof. Consider an arbitrary terminal $\pi_{i^\star} \in W$ and an arbitrary set $U \subseteq V \setminus \{t\}$ such that $\pi_{i^\star} \in U$. We have

$$\begin{aligned} \sum_{e \in \delta(U)} (y^\star_e + p_e) \quad &\geq \quad \sum_{e \in \delta(U)} y^\star_e + \sum_{i' \in F_\mu \setminus U} p_{(\pi_{i^\star}, i')} \\ &\geq \quad \sum_{i' \in F_\mu \cap U} z^\star_{i', \pi_{i^\star}} + \sum_{i' \in F_\mu \setminus U} z^\star_{i', \pi_{i^\star}} \;=\; \sum_{i' \in F_\mu} z^\star_{i', \pi_{i^\star}} \;=\; 1, \end{aligned}$$

where the second inequality follows from the feasibility of $y^\star$ and the construction of p and the last equality follows from the construction of the above algorithm. ◀

4 Analysis

In this section, we show that our algorithm is an approximation algorithm for the connected minimum sum of radii problem and establish the approximation guarantee.

Feasibility of the solutions

Consider each guess $t \in F$. It is clear that $(S_t^{\mathrm{I}}, \rho_t^{\mathrm{I}}, T_t^{\mathrm{I}})$ is a feasible solution. In the following we show that $(S_t^{\mathrm{II}}, \rho_t^{\mathrm{II}}, T_t^{\mathrm{II}})$ is also feasible.

By Lemma 6 and the correctness of Jain's rounding algorithm [18], T_t^{II} is indeed a Steiner tree for S_t^{II}. Hence, it suffices to prove the following lemma, which implies that, for all $j \in D$, there always exists some opened facility $i \in S^{\mathrm{II}}$ such that $d(i,j) \le \rho_i^{\mathrm{II}}$.

▶ **Lemma 7.** *For all $j \in D$, there exists some $(i^\star, r^\star) \in \mathcal{B}^\star$ such that $d(i^\star, j) \le 3r^\star$.*

Proof. Note that we have $\Delta_{\mathcal{B}^\star}(j) = +\infty$ at the beginning and $\Delta_{\mathcal{B}^\star}(j) < 5$ at the end of the execution of Algorithm 1. Consider the iteration at which $\Delta_{\mathcal{B}^\star}(j)$ becomes smaller than 5 for the first time and let $(i^\star, r^\star)$ be the ball chosen at Step 4 during this iteration. Since $\mathcal{G}$ is bipartite, $\Delta_{\mathcal{B}^\star}(j)$ becomes 1 or 3 at this iteration. If it becomes 1, this implies $j \in B(i^\star, r^\star)$ and there is nothing to prove. If $\Delta_{\mathcal{B}^\star}(j)$ becomes 3, this implies that there exists a path of length three between $(i^\star, r^\star)$ and j in $\mathcal{G}$; let $(i^\star, r^\star) - j' - (i', r') - j$ denote this path. At the beginning of this iteration, $\Delta_{\mathcal{B}^\star}(j)$ was no smaller than 5 and therefore $(i', r') \in \bar{\mathcal{B}}$. Since the algorithm chose $(i^\star, r^\star)$ over (i', r'), we have $r^\star \ge r'$, yielding $d(i^\star, j) \le d(i^\star, j') + d(j', i') + d(i', j) \le r^\star + r' + r' \le 3r^\star$. ◀

Approximation Guarantee

In the following we establish the approximation guarantee. Let $(S^{\mathsf{opt}}, r^{\mathsf{opt}}, T^{\mathsf{opt}})$ be an optimal solution and OPT denote its cost.

If $|S^{\mathsf{opt}}| = 1$, then the facility in S^{opt} will be iterated by the algorithm. Denote this facility by t^*. Then $(S_{t^*}^{\mathrm{I}}, \rho_{t^*}^{\mathrm{I}}, T_{t^*}^{\mathrm{I}})$ is an optimal solution and there is nothing to prove.

In the following we assume that $|S^{\mathsf{opt}}| \ge 2$. Since the algorithm iterates over all possible guesses, we assume without loss of generality that t is the facility with the smallest m_t value in S^{opt}, i.e.,

$$t \in S^{\mathsf{opt}} \quad \text{and} \quad t = \operatorname{argmin}_{i \in S^{\mathsf{opt}}} m_i.$$

Depending on whether or not $t \in F_\mu$, we further consider two cases. The following lemma shows that $(S_t^{\mathrm{I}}, \rho_t^{\mathrm{I}}, T_t^{\mathrm{I}})$ is a μ-approximation solution if $t \notin F_\mu$.

▶ **Lemma 8.** *If $t \in S^{\mathsf{opt}}$ and $t \notin F_\mu$, then*

$$OPT \ \ge\ \frac{1}{\mu} \cdot m_t \cdot \max_{j \in D} d(t, j).$$

Proof. We have $m_t < \mu$ by the assumption. Let $j^o := \arg\max_{j \in D} d(t, j)$ be the client that defines the radius ρ_t^{I}. Let i' be a facility in S^{opt} with $d(i', j^o) \le \rho_{i'}^{\mathsf{opt}}$. We have

$$
\begin{aligned}
\mathsf{OPT} &= \sum_{i \in S^{\mathsf{opt}}} m_i \cdot \rho_i^{\mathsf{opt}} \;+\; \sum_{e \in T^{\mathsf{opt}}} d_e \\
&\geq m_{i'} \cdot \rho_{i'}^{\mathsf{opt}} \;+\; d(t, i') \\
&\geq m_t \cdot d(i', j^o) \;+\; \frac{1}{\mu} m_t \cdot d(t, i') \\
&\geq \frac{1}{\mu} \cdot m_t \cdot d(t, j^o) \;\;=\;\; \frac{1}{\mu} \cdot m_t \cdot \max_{j \in D} d(t, j),
\end{aligned}
$$

where in the last inequality we apply the triangle inequality and the fact that $\mu \geq 1$. ◀

It remains to consider the case that $t \in F_\mu$, which in particular implies that $S^{\mathsf{opt}} \subseteq F_\mu$. We prove in the following that $(S_t^{\mathrm{II}}, \rho_t^{\mathrm{II}}, T_t^{\mathrm{II}})$ is a $(5 + \frac{3}{\mu})$-approximation solution in this case.

Since $S^{\mathsf{opt}} \subseteq F_\mu$, it follows that the LP (1) admits $(S^{\mathsf{opt}}, r^{\mathsf{opt}}, T^{\mathsf{opt}})$ as a feasible solution. Hence the cost of the fractional solution $(x^\star, y^\star, z^\star)$ provides a lower-bound for OPT. Similarly to the facility location problem [4] and the minimum sum of radii problem [9], we use the dual optimal solution to bound the cost of the rounded solution via complementary slackness. Consider the dual LP of the LP (1), which we provide below, and let $(\alpha^\star, \beta^\star, \gamma^\star, \lambda^\star)$ be an optimal solution for it.

$$
\begin{aligned}
\text{maximize} \quad & \sum_{j \in D} \alpha_j \\
\text{subject to} \quad & \sum_{j \in B(i,r)} \gamma_{i,j} \;\leq\; m_i \cdot r, && \forall i \in F_\mu, r \in R_i, \\
& \alpha_j - \sum_{U \subseteq V \setminus \{t\} : i \in U} \beta_{j,U} \;\leq\; \gamma_{i,j}, && \forall i \in F_\mu, j \in D, \\
& \sum_{j \in D} \sum_{U \subseteq V \setminus \{t\} : e \in \delta(U)} \beta_{j,U} \;\leq\; d_e, && \forall e \in E, \\
& \alpha, \beta, \gamma \geq 0.
\end{aligned}
\tag{3}
$$

The following lemma bounds the weighted cost of a facility in term of the dual values of the clients contained within. Intuitively, it follows from standard complementary slackness conditions between $(x^\star, y^\star, z^\star)$ and $(\alpha^\star, \beta^\star, \gamma^\star, \lambda^\star)$.

▶ **Lemma 9.** *For any $i \in F_\mu$ and any $r \in R_i$, we have $x^\star_{i,r} > 0$ implies that $m_i \cdot r \leq \sum_{j \in B(i,r)} \alpha^\star_j$.*

Proof. From the complementary slackness condition, $x^\star_{i,r} > 0$ implies

$$
\sum_{j \in B(i,r)} \gamma^\star_{i,j} \;=\; m_i \cdot r. \tag{4}
$$

Consider an arbitrary $j \in B(i,r)$. If $\gamma^\star_{i,j} > 0$, we have from the complementary slackness that

$$
z^\star_{i,j} \;=\; \sum_{r' \in R_i : j \in B(i,r')} x^\star_{i,r'} \;>\; x^\star_{i,r} \;>\; 0.
$$

By complementary slackness condition again this implies

$$\gamma^\star_{i,j} \;=\; \alpha^\star_j - \sum_{U\subseteq V\setminus\{t\}:i\in U} \beta^\star_{j,U} \;\le\; \alpha^\star_j. \tag{5}$$

On the other hand, if $\gamma^\star_{i,j} = 0$, it trivially holds that

$$\gamma^\star_{i,j} \le \alpha^\star_j. \tag{6}$$

Combining (5) and (6) with (4) yields $m_i \cdot r \;\le\; \sum_{j\in B(i,r)} \alpha^\star_j$. ◀

Consider any $(i_1, r_1), (i_2, r_2) \in \mathcal{B}^\star$ such that $i_1 \neq i_2$. By the design of the rounding procedure in the first part of the algorithm, we always have that $B(i_1, r_1)$ and $B(i_2, r_2)$ are disjoint. Hence, combining this fact with Lemma 9, the total weighted facility cost can be bounded as

$$\begin{aligned}\sum_{i\in S^{\mathrm{II}}_t} m_i \cdot \rho^{\mathrm{II}}_i \;&=\; \sum_{(i^\star,r^\star)\in\mathcal{B}^\star} 3\cdot m_i \cdot r^\star \\ &\le\; \sum_{(i^\star,r^\star)\in\mathcal{B}^\star} \left(3\cdot \sum_{j\in B(i^\star,r^\star)} \alpha^\star_j\right) \;=\; 3\cdot\sum_{j\in D}\alpha^\star_j \;\le\; 3\cdot\mathsf{OPT}.\end{aligned} \tag{7}$$

In the following we consider the cost incurred by the Steiner tree T^{II}_t. We have the following lemma regarding the value of the solution $y^\star + p$ with respect to LP (2).

▶ **Lemma 10.**

$$\sum_{e\in E} d_e \cdot (y^\star_e + p_e) \;\le\; \left(1+\frac{1}{\mu}\right)\cdot\mathsf{OPT}.$$

Proof. By the construction of p we have

$$\sum_{e\in E} d_e\cdot y^\star_e + \sum_{e\in E} d_e\cdot p_e \;\le\; \mathsf{OPT} \;+\; \sum_{i^\star\in S^{\mathrm{II}}}\sum_{i'\in F_\mu} d(i',\pi_{i^\star})\cdot z^\star_{i',\pi_{i^\star}}. \tag{8}$$

For any (i, j) such that $z^\star_{i,j} > 0$, the feasibility of $(x^\star, y^\star, z^\star)$ implies that there must exist some $r \in R_i$ such that $j \in B(i, r)$ and $x^\star_{i,r} > 0$. This yields

$$\begin{aligned}\sum_{i^\star\in S^{\mathrm{II}}}\sum_{i'\in F_\mu} d(i',\pi_{i^\star})\cdot z^\star_{i',\pi_{i^\star}} \;&=\; \sum_{(i^\star,r^\star)\in\mathcal{B}^\star}\sum_{i'\in F_\mu} d(i',\pi_{i^\star})\cdot z^\star_{i',\pi_{i^\star}} \\ &\le\; \sum_{(i^\star,r^\star)\in\mathcal{B}^\star}\sum_{i'\in F_\mu} r^\star\cdot z^\star_{i',\pi_{i^\star}} \\ &=\; \sum_{(i^\star,r^\star)\in\mathcal{B}^\star} r^\star \;\le\; \sum_{i^\star\in S^{\mathrm{II}}} \frac{m_{i^\star}}{\mu}\cdot r^\star \;\le\; \frac{1}{\mu}\cdot\mathsf{OPT},\end{aligned}$$

where the first inequality follows from Lemma 5 and the fact that $z^\star_{i'\pi_{i^\star}} > 0$ implies that there exists some $r' \in R_{i'}$ such that $\pi_{i^\star} \in B(i', r')$ and $x^\star_{i',r'} > 0$, the second equality follows from the construction in the second part of the algorithm, the second inequality from $S^{\mathrm{II}} \subseteq F_\mu$ which implies that $m_i \ge \mu$ for all $i \in F_\mu$, and the last inequality follows from (7). ◀

By the design of the algorithm for constructing T_t^{II}, the bound in Lemma 10, and the fact that Jain's rounding algorithm gives a 2-approximation [18], we have

$$\begin{aligned}\sum_{e \in T_t^{\mathrm{II}}} d_e &= \sum_{e \in T_{\mathsf{pre}}} d_e + \sum_{i^\star \in S_t^{\mathrm{II}}} d(i^\star, \pi_{i^\star}) \\ &\leq \left(2 + \frac{2}{\mu}\right) \cdot \mathsf{OPT} + \sum_{(i^\star, r^\star) \in \mathcal{B}^\star} r^\star \\ &\leq \left(2 + \frac{2}{\mu}\right) \cdot \mathsf{OPT} + \sum_{(i^\star, r^\star) \in \mathcal{B}^\star} \frac{m_{i^\star}}{\mu} \cdot r^\star \leq \left(2 + \frac{3}{\mu}\right) \cdot \mathsf{OPT}, \end{aligned} \tag{9}$$

where in the second last inequality we use the fact that $m_{i^\star} \geq \mu$ for all $i^\star \in S_t^{\mathrm{II}}$ and in last inequality we apply Inequality (7). Combining Inequalities (7) and (9), we obtain

$$\sum_{i \in S_t^{\mathrm{II}}} m_i \cdot \rho_i^{\mathrm{II}} + \sum_{e \in T_t^{\mathrm{II}}} d_e \leq \left(5 + \frac{3}{\mu}\right) \cdot \mathsf{OPT}.$$

This proves the following theorem. Choosing $\mu := \frac{5+\sqrt{37}}{2} < 5.542$ yields a μ-approximation algorithm.

▶ **Theorem 11.** *The given algorithm is a* $\max\left(\mu, 5 + \frac{3}{\mu}\right)$*-approximation algorithm.*

5 NP-hardness Results

In this section, we prove Theorem 2 by showing that the problem remains NP-hard even for two special cases. First, the following theorem shows that this problem remains NP-hard even when we only allow clusters with zero radii.

▶ **Theorem 12.** *The* connected minimum sum of radii *problem is NP-hard when* $m_i = +\infty$ *for all* $i \in F$ *and* $M = 1$*.*

Proof. We give a reduction from the Metric Steiner Tree problem, which is known to be NP-complete [20]. Construct an instance of the connected minimum sum of radii problem where the terminals in the Steiner tree instance become facilities and clients at the same time. Observe that an optimal solution to this instance opens all terminals, set their radii to zeroes, and takes a Steiner tree connecting them. ◀

On the other hand, the following theorem shows that the NP-hardness remain true even when no connection between opened facilities is required. The proof closely follows the NP-hardness proof of the (non-connected) minimum sum of radii problem [11]; but we present the full proof here for the sake of completeness.

▶ **Theorem 13.** *The* connected minimum sum of radii *problem is NP-hard when* $m_i = 1$ *for all* $i \in F$ *and* $M = 0$*.*

Proof. We give a reduction from 3SAT [5]. Consider an instance of 3SAT with n variables $x_1, \ldots, x_n$ and k clauses $C_1, \ldots, C_k$. We construct an instance of the connected minimum sum of radii problem as follows.

Let $F := \{x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_n, \bar{x}_n\}$ and $D := \{C_1, \ldots, C_k, v_1, \ldots, v_n\}$. To define a metric on $V := F \cup D$, consider a weighted graph on the vertex set V, where we have an edge (x_i, C_j) (or $(\bar{x}_i, C_j)$, respectively) of weight 2^{i-1} if and only if C_j contains x_i (or $\bar{x}_i$). We also add edges (x_i, v_i) and $(\bar{x}_i, v_i)$ of weight 2^{i-1} for all $i = 1, \ldots, n$. The metric d is then defined as the shortest path metric on this weighted graph. We claim that the optimal solution value to this constructed instance is at most $\sum_{i=1}^{n} 2^{i-1} = 2^n - 1$ if and only if the 3SAT instance is satisfiable.

Suppose that the 3SAT instance is satisfiable. Fix a satisfying assignment. For each variable x_i, we open x_i (or $\bar{x}_i$, respectively) if x_i is `true` (or `false`) under the fixed assignment and set its radius to 2^{i-1}. This yields a solution of value $2^n - 1$ in which every client can be assigned.

Conversely, suppose that there exists a solution to the constructed instance of the connected minimum sum of radii problem whose value is at most $2^n - 1$. Fix such a solution. Suppose towards contradiction that there exists some k such that neither x_k nor $\bar{x}_k$ is open with radius at least 2^{k-1}. Let $k^\star$ be the largest such k. Then there must exist some ℓ such that $v_{k^\star}$ is assigned to x_ℓ or $\bar{x}_\ell$. Note that $d(v_{k^\star}, x_\ell) = d(v_{k^\star}, \bar{x}_\ell) \geq 2 \cdot 2^{k^\star-1} + 2^{\ell-1}$ since every edge incident with $x_{k^\star}$ or $\bar{x}_{k^\star}$ is of weight $2^{k^\star-1}$ and every edge incident with x_ℓ or $\bar{x}_\ell$ is of weight $2^{\ell-1}$. If $\ell > k^\star$, the total cost of the solution must be at least

$$\sum_{i \in \{k^\star+1, \ldots, n\} \setminus \{\ell\}} 2^{i-1} + (2 \cdot 2^{k^\star-1} + 2^{\ell-1}) > 2^n - 1,$$

which leads to contradiction. If $\ell < k^\star$, the total cost of the solution must be at least

$$\sum_{i \in \{k^\star+1, \ldots, n\}} 2^{i-1} + (2 \cdot 2^{k^\star-1} + 2^{\ell-1}) > 2^n - 1,$$

leading to contradiction again.

We thus have that, for all $i = 1, \ldots, n$, x_i or $\bar{x}_i$ (or both) is open with radius at least 2^{i-1}. Since $2^n - 1 = \sum_{i=1}^{n} 2^{i-1}$, this implies that exactly one of x_i and $\bar{x}_i$ is open with radius exactly 2^{i-1} for all $i = 1, \ldots, n$. (Note that opening with zero radius is useless.) Consider a truth value assignment that sets x_i to `true` if x_i is open, and `false` otherwise. Observe that this is a satisfying assignment. ◀

References

1 Matthew Andrews and Lisa Zhang. The access network design problem. In *39th Annual Symposium on Foundations of Computer Science, FOCS '98, November 8-11, 1998, Palo Alto, California, USA*, pages 40–59. IEEE Computer Society, 1998. `doi:10.1109/SFCS.1998.743427`.

2 Moritz Buchem, Katja Ettmayr, Hugo K. K. Rosado, and Andreas Wiese. A $(3+\varepsilon)$-approximation algorithm for the minimum sum of radii problem with outliers and extensions for generalized lower bounds. In David P. Woodruff, editor, *Proceedings of the 2024 ACM-SIAM Symposium on Discrete Algorithms, SODA 2024, Alexandria, VA, USA, January 7-10, 2024*, pages 1738–1765. SIAM, 2024. `doi:10.1137/1.9781611977912.69`.

3 Moses Charikar and Rina Panigrahy. Clustering to minimize the sum of cluster diameters. In Jeffrey Scott Vitter, Paul G. Spirakis, and Mihalis Yannakakis, editors, *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 1–10. ACM, 2001. `doi:10.1145/380752.380753`.

4 Fabián A. Chudak and David B. Shmoys. Improved approximation algorithms for the uncapacitated facility location problem. *SIAM Journal on Computing*, 33(1):1–25, 2003. `doi:10.1137/S0097539703405754`.

5 Stephen A. Cook. The complexity of theorem-proving procedures. In Michael A. Harrison, Ranan B. Banerji, and Jeffrey D. Ullman, editors, *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing, May 3-5, 1971, Shaker Heights, Ohio, USA*, pages 151–158. ACM, 1971. `doi:10.1145/800157.805047`.

6 Srinivas Doddi, Madhav V. Marathe, S. S. Ravi, David Scot Taylor, and Peter Widmayer. Approximation algorithms for clustering to minimize the sum of diameters. In Magnús M. Halldórsson, editor, *Algorithm Theory - SWAT 2000, 7th Scandinavian Workshop on Algorithm Theory, Bergen, Norway, July 5-7, 2000, Proceedings*, volume 1851 of *Lecture Notes in Computer Science*, pages 237–250. Springer, 2000. doi:10.1007/3-540-44985-X_22.

7 Ding-Zhu Du and Peng-Jun Wan. *Connected Dominating Set: Theory and Applications*. Springer Publishing Company, Incorporated, 2012.

8 Friedrich Eisenbrand, Fabrizio Grandoni, Thomas Rothvoß, and Guido Schäfer. Approximating connected facility location problems via random facility sampling and core detouring. In Shang-Hua Teng, editor, *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*, pages 1174–1183. SIAM, 2008. URL: http://dl.acm.org/citation.cfm?id=1347082.1347210.

9 Zachary Friggstad and Mahya Jamshidian. Improved polynomial-time approximations for clustering with minimum sum of radii or diameters. In Shiri Chechik, Gonzalo Navarro, Eva Rotenberg, and Grzegorz Herman, editors, *30th Annual European Symposium on Algorithms, ESA 2022, September 5-9, 2022, Berlin/Potsdam, Germany*, volume 244 of *LIPIcs*, pages 56:1–56:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICS.ESA.2022.56.

10 Zachary Friggstad, Mohsen Rezapour, and Mohammad R. Salavatipour. Approximating connected facility location with lower and upper bounds via LP rounding. In Rasmus Pagh, editor, *15th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2016, June 22-24, 2016, Reykjavik, Iceland*, volume 53 of *LIPIcs*, pages 1:1–1:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPICS.SWAT.2016.1.

11 Matt Gibson, Gaurav Kanade, Erik Krohn, Imran A. Pirwani, and Kasturi R. Varadarajan. On metric clustering to minimize the sum of radii. *Algorithmica*, 57(3):484–498, 2010. doi:10.1007/S00453-009-9282-7.

12 Michel X. Goemans and David P. Williamson. A general approximation technique for constrained forest problems. *SIAM J. Comput.*, 24(2):296–317, 1995. doi:10.1137/S0097539793242618.

13 Sudipto Guha and Samir Khuller. Approximation algorithms for connected dominating sets. *Algorithmica*, 20(4):374–387, 1998. doi:10.1007/PL00009201.

14 Anupam Gupta, Jon M. Kleinberg, Amit Kumar, Rajeev Rastogi, and Bülent Yener. Provisioning a virtual private network: a network design problem for multicommodity flow. In Jeffrey Scott Vitter, Paul G. Spirakis, and Mihalis Yannakakis, editors, *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 389–398. ACM, 2001. doi:10.1145/380752.380830.

15 Anupam Gupta, Amit Kumar, and Tim Roughgarden. Simpler and better approximation algorithms for network design. In Lawrence L. Larmore and Michel X. Goemans, editors, *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 365–372. ACM, 2003. doi:10.1145/780542.780597.

16 Pierre Hansen and Brigitte Jaumard. Cluster analysis and mathematical programming. *Math. Program.*, 79:191–215, 1997. doi:10.1007/BF02614317.

17 D. S. Hochbaum and D. B. Shmoys. A best possible heuristic for the k-center problem. *Mathematics of Operations Research*, 10:180–184, 1985. doi:10.1287/MOOR.10.2.180.

18 Kamal Jain. A factor 2 approximation algorithm for the generalized steiner network problem. *Comb.*, 21(1):39–60, 2001. doi:10.1007/S004930170004.

19 Riheng Jia, Jinhao Wu, Jianfeng Lu, Minglu Li, Feilong Lin, and Zhonglong Zheng. Energy saving in heterogeneous wireless rechargeable sensor networks. In *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications, London, United Kingdom, May 2-5, 2022*, pages 1838–1847. IEEE, 2022. doi:10.1109/INFOCOM48880.2022.9796770.

20 Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller and James W. Thatcher, editors, *Proceedings of a symposium on the Complexity of Computer Computations, held March 20-22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, USA*, The IBM Research Symposia Series, pages 85–103. Plenum Press, New York, 1972. `doi:10.1007/978-1-4684-2001-2_9`.

21 Samir Khuller, Manish Purohit, and Kanthi K. Sarpatwar. Analyzing the optimal neighborhood: Algorithms for budgeted and partial connected dominating set problems. In Chandra Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1702–1713. SIAM, 2014. `doi:10.1137/1.9781611973402.123`.

22 Samir Khuller and Sheng Yang. Revisiting connected dominating sets: An almost optimal local information algorithm. *Algorithmica*, 81(6):2592–2605, 2019. `doi:10.1007/S00453-019-00545-0`.

23 Chaitanya Swamy and Amit Kumar. Primal-dual algorithms for connected facility location problems. *Algorithmica*, 40(4):245–269, 2004. `doi:10.1007/S00453-004-1112-3`.

24 Wenzheng Xu, Weifa Liang, Xiaohua Jia, Haibin Kan, Yinlong Xu, and Xinming Zhang. Minimizing the maximum charging delay of multiple mobile chargers under the multi-node energy charging scheme. *IEEE Trans. Mob. Comput.*, 20(5):1846–1861, 2021. `doi:10.1109/TMC.2020.2973979`.