Approximation Algorithms for Aversion k-Clustering via Local k-Median*

Anupam Gupta[1], Guru Guruganesh[2], and Melanie Schmidt[3]

1 Computer Science Department, Carnegie Mellon University, Pittsburgh, USA
2 Computer Science Department, Carnegie Mellon University, Pittsburgh, USA
3 Institute of Computer Science, University of Bonn, Bonn, Germany

Abstract

In the *aversion k-clustering problem*, given a metric space, we want to cluster the points into k clusters. The cost incurred by each point is the distance to the furthest point in its cluster, and the cost of the clustering is the sum of all these per-point-costs. This problem is motivated by questions in generating automatic abstractions of extensive-form games.

We reduce this problem to a *"local" k-median problem* where each facility has a prescribed radius and can only connect to clients within that radius. Our main results is a constant-factor approximation algorithm for the aversion k-clustering problem via the local k-median problem. We use a primal-dual approach; our technical contribution is a non-local rounding step which we feel is of broader interest.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Approximation algorithms, clustering, k-median, primal-dual

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.66

1 Introduction

In this paper, we consider the following clustering problem: given a metric space (X, D) with a set of clients, and a parameter k, the goal is to group the clients into k clusters. For each client j, the cost incurred by j is its distance to the furthest client in its cluster. The cost of the clustering is the sum over all clients, more precisely, of the cost incurred by the client. We call this problem the *aversion k-clustering problem*.

This question is motivated by a problem in developing abstractions of extensive-form games. Since finding equilibria in large extensive form games is computationally expensive, one appealing approach if speeding things up is to develop an abstraction of this game. Since the abstraction is typically much smaller, existing algorithms can be used to solve them to find optimal strategies, which can be mapped back to the original game. However, there is often some loss in going to the abstraction. Recent work of Kroer and Sandholm [19] on automated abstraction algorithms proposed a following way to model this: since several states of the original game may be collapsed into a single state in the abstraction, the loss for each original state is its distance (in a suitably defined metric) to the furthest state that is collapsed with it. The overall loss is the sum of per-state losses. This is precisely the aversion k-clustering problem we study in this paper.

* This work was done in part when M. Schmidt was visiting the Computer Science Department at Carnegie Mellon University, supported by the German Academic Exchange Service (DAAD). A. Gupta and G. Guruganesh were supported in part by NSF awards CCF-1319811 and CCF-1536002

To the best of our knowledge, no prior approximation algorithms were known for the aversion k-clustering problem. Although it is related in spirit to several other clustering problems, it has some interesting and unique features. Indeed, something that makes this problem difficult is its high "sensitivity". To explain this, observe that in problems like k-median, if we re-assign a single client j to a new cluster C, loosely this changes the cost by the distance of the client to the new cluster. However, in aversion k-clustering, reassigning client j to a new cluster C may also significantly change the cost of all other clients in C, since j may become their new furthest client. This creates problems for most standard techniques used for facilty-location problems. Another facility-location problem with a similar high-sensitivity property is the *min-sum clustering* problem, for which only logarithmic approximations are known via HST embeddings and a non-trivial dynamic program [4, 5]. Since our objective is not linear (due to each client paying the distance to its furthest cluster-mate), we cannot even use tree embeddings.

The main result of the paper is the following:

▶ **Theorem 1.** *There is a constant-factor algorithm for the aversion k-clustering problem.*

1.1 Our Techniques

A few words about our techniques. To solve aversion k-clustering, we first move to a related problem that is more convenient to deal with: in the *local k-median problem*, each potential facility location in the metric space has a "range" R_i associated with it. Like in k-median, we need to open k facilities, to minimize the sum of distances from clients to their assigned facilities. However, we now additionally require that each client j is assigned to some facility i at distance at most R_i. This problem is NP-hard to approximate, but for our purpose it is sufficient to solve the relaxed version where clients can still connect at distance $O(\mathsf{R}_i)$.

The (relaxed) locality restriction causes many of the standard techniques for k-median, like local-search and LP-rounding, not to extend to this problem. (In fact, we do not know of a constact factor approximation algorithm for local k-median which only violates locaility constraints by a constant factor). However, we are successful in extending a primal-dual technique to the instances which arise from the aversion k-clustering problem. The following theorem is our main technical result, from which Theorem 1 follows immediately.

▶ **Theorem 2.** *There is an approximation algorithm for the local k-median problem that achieves a constant-factor approximation for instances arising from the aversion k-clustering problem. Its solutions violate the locality constraints by a constant factor.*

We use the primal-dual framework of Jain and Vazirani: we find two solutions that open k_1 and k_2 facilities (such that $k_1 < k < k_2$) such that the "average" of these two solutions has low cost and opens k facilities. This part of the analysis is well-understood by now and we omit details because of space limitations. We can view this average solution as a "well-behaved" LP solution, which we now have to round to integrally open k facilities.

The main problem with this rounding is the locality constraint — typical algorithms tend to round some fractional facility up to 1, round down close-by fractional facilities to zero to maintain the total facility mass at k, and reroute clients to the newly opened facility without increasing the cost by much. However, the locality constraint in our problem means that such simple rounding approaches fail. For example, the facility that we open may have a very small R_i value, and can only serve clients that are very close to it. However, the clients who want to be rerouted may be too far from this facility to satisfy the locality constraint, even if it is relaxed to $\gamma\mathsf{R}_i$ for some constant γ.

Our main technical contribution, and the novel ingredient of our rounding algorithm is a *non-local rounding approach*. We first transform the fractional solution so that its support is a forest. Then we partition this forest into carefully chosen subtrees, so that all the clients in each particular subtree can be reassigned simultaneously without violating the locality. Now choosing the least expensive of these subtrees to reassign gives us a solution with k facilities and a constant-factor approximation for instances arising from the aversion k-clustering problem. We feel that this non-local rounding will be useful in other contexts, and hence be of independent interest.

Related work. Approximation algorithms for facility location problems have been studied for a long time. Indeed, many approximation techniques have been developed while investigating these problems (see [22]). The problem closest to the *local k-median problem* is naturally the *metric k-median* problem. The first constant-factor for this problem was due to Charikar et al. [8] via rounding the LP; subsequently, primal-dual algorithms were given by Jain and Vazirani [17] and Charikar and Guha [7], a local-search algorithm was given by Arya et al. [2]. The recent approach of Li and Svensson [21] gave a $2.73+\varepsilon$-approximation, which was improved to $2.675+\varepsilon$ by Byrka et al. [6]. The current NP-hardness is a $1+2/e$-factor due to Jain, Mahdian, Saberi [16]. The related problem of *uncapacitated metric facility location* has constant-factor approximations via most approximation techniques: see, e.g., the book of [22]. The current best approximation factor is 1.488 due to Li [20], and the hardness is an 1.463-factor due to Guha and Khuller [13].

The k-median problem sums over each cluster, the sum of distances of clients to their cluster center. Instead of taking the sum of distances within each cluster, we could take the maximum distance within each cluster; this gives the *sum of cluster diameters* problem, for which a $O(1)$-factor is due to Charikar and Panigrahy [9]. And instead of summing diameters over the clusters, if we take the maximum diameter over all clusters, we get the *k-center* problem, for which a 2-approximation is due to Gonzalez [12], and Hochbaum and Shmoys [14], and a matching hardness is due to Hsu and Nemhauser [15].

Another related problem is the *min-sum clustering problem*, where we sum over the clusters of the distances between all pairs within the cluster. Bartal et al. [4] give a $O(\varepsilon^{-1}\log^{1+\varepsilon} n)$-approximation, which was recently improved to $O(\log n)$ by Behsaz [5]. There are easy examples where these problems differ from aversion k-clustering by arbitrarily large factors. Moreover, the non-linearity of our objective function means that we cannot use tree embedding results to even get a logarithmic approximation.

Our algorithm takes a primal-dual approach pioneered by Jain and Vazirani [17]; while solving the Lagrangian relaxation and getting a Lagrangian-multiplier preserving algorithm follows relatively easily, the main contribution is in the non-local rounding algorithm. This adds to the body of work exploring such primal-dual techniques, which include the work of Charikar and Panigrahy [9] to give a $O(1)$-factor approximation for the sum of cluster diameters, and Chuzhoy and Rabani [10] in their $O(1)$-factor bicriteria approximation for the capacitated k-median problem. Non-local roundings of a different flavor were also recently used for the capacitated k-center problem by Cygan et al. [11] and An et al. [1]. To the best of our understanding, our rounding technique is different from these previous works.

2 Preliminaries

Let (X, D) be a metric space and let $C \subseteq X$ be the set of *clients* and $F \subseteq X$ be the set of *facilities*. The *aversion k-clustering problem* is the task to partition C into a collection $\mathcal{C}$ of

k disjoint subsets $C_1, \dots, C_k$ with $C = \cup_{i=1}^k C_i$ such that

$$c_a(\mathcal{C}) := \sum_{\ell=1}^{k} \sum_{j_1 \in C_\ell} \max_{j_2 \in C_\ell} D(j_1, j_2) \tag{P1}$$

is minimized.

For the *local k-median problem*, we additionally get a *radius* (or *range*) R_i for every $i \in F$. We seek a set $\mathcal{F} \subseteq F$ of k facilities that minimizes

$$c_l(\mathcal{F}) := \sum_{j \in C} \min_{\substack{i \in \mathcal{F}, \\ D(i,j) \le \mathsf{R}_i}} D(i, j).$$

This differs from the classical k-median problem in that a client can only be assigned to a facility if it lies within the facility's radius. It is possible that there is no set of k facilities which can service all clients. If this is the case, we define the minimum clustering cost as infinity. In the following claim, we show that it is NP-hard to decide whether we are in this case or not.

▶ **Claim 1.** Deciding feasibility of a local k-median instance is NP-hard.

Proof. We use a well-known reduction from set cover. Let $\mathcal{S}$ be a set of sets over a universe $\mathcal{U}$. We construct a metric space that contains a facility i_S for every set $S \in \mathcal{S}$ and a client j_u for every element $u \in \mathcal{U}$. The distance between j_u and i_S is one if $u \in S$ and two otherwise. Observe that this is a metric. We set the radius R_{i_S} of all facilities to one. Observe that there is a feasible solution for this local k-median instance if and only if the set cover instance has a solution with at most k sets. Since deciding whether a set cover instance has a solution with at most k sets is NP-hard [18], it is also NP-hard to decide whether there is a feasible solution for the local k-median problem. ◀

Any approximation algorithm has to decide whether there is a feasible solution or not. Hence, we allow the locality constraint to be violated; i.e. a client may connect to a facility i if it is within a radius of $\gamma \mathsf{R}_i$ for a constant γ. We say a solution is a (γ, ψ) bicriteria solution if the solution violates the locality constraints by a factor of γ and has cost at most ψ times the optimal (with respect to the original problem).

3 Solving the aversion k-clustering problem via the local k-median problem

We show that the aversion k-clustering problem can be reduced to the local k-median problem by sacrificing a constant factor. The idea is to identify a cluster C_ℓ with a pair of points with largest distance and to use this information to represent clusters by an artificial facility with appropriate radius. More precisely, we define the following metric space. Set $F := \{p_{j_1 j_2} \mid j_1, j_2 \in C\}$ and refer to $p_{j_1 j_2}$ as the *midpoint* of clients j_1 and j_2. To extend D from C to $C \cup F$, we set

$$D(j_1, p_{j_1 j_2}) := D(j_2, p_{j_1 j_2}) = \frac{D(j_1, j_2)}{2} \text{ and } D(p_{j_1 j_2}, p_{j_1 j_2}) := 0$$

for all $j_1, j_2 \in C$. So far, no metric property is violated. Now imagine the incompletely defined metric as a weighted undirected graph G on the vertices $C \cup F$ where some edges are missing. Let D be defined as the shortest path metric in G. This is a metric by definition. It coincides with the previously defined distances since in a metric, the direct edge must be a shortest path. For the missing edges, we get that

$$D(j, p_{j_1 j_2}) = D(p_{j_1 j_2}, j) = \min\{D(j, j_1), D(j, j_2)\} + \frac{D(j_1, j_2)}{2} \tag{1}$$

for all $j \in C$: The point $p_{j_1 j_2}$ is only connected to j_1 and j_2, thus any path between j and $p_{j_1 j_2}$ has to travel over one of them. Since the edge lengths form a metric, the direct connection between j and j_1 or j_2 is shortest, so either $(j, j_1), (j_1, p_{j_1 j_2})$ or $(j, j_2), (j_2, p_{j_1 j_2})$ is a shortest path. Analogously, we get that

$$D(p_{j_1 j_2}, p_{j_3 j_4}) = D(j_1, j_2)/2 + D(j_3, j_4)/2 + \min\{D(j_1, j_3), D(j_1, j_4), D(j_2, j_3), D(j_2, j_4)\}$$

for all $j_1, j_2, j_3, j_4 \in C$. Finally, we define

$$\mathsf{R}_{p_{j_1 j_2}} := D(j_1, j_2)/2 \tag{2}$$

for all $p_{j_1 j_2} \in F$. Notice that our definition of F allows that $j_1 = j_2$. This ensures that singleton clusters can be expressed. Furthermore, notice that $\mathsf{R}_{p_{j_1 j_1}} = 0$, so the facility $p_{j_1 j_1}$ can only serve j_1 (or clients at the same location).

For any facility p_{j_1, j_2}, we will drop the reference to j_1 and j_2 when it is clear from the context. Hence, $p \in F$ refers to the midpoint of some two clients j_1 and j_2 and the radius of the facility R_p simply refers to half the distance between these points. Intuitively, each new "facility" corresponds to a midpoint of two clients in the original problem. These midpoints allow us to cast the current problem as a k-median problem with the addition of locality constraints placed on each facility.

We now show how solutions for the aversion k-clustering problem and (γ, α) bicriteria solutions for the local k-median problem are related. For a client j, let $F_j^\gamma := \{i \in F \mid D(i, j) \le \gamma \mathsf{R}_i\}$ be the set of facilities that j is allowed to connect to. The following integer linear program (ILP) which is a (natural) modification of the ILP proposed in [3] minimizes over all feasible (γ, α) bicriteria solutions.

$$\begin{aligned} \min \sum_{j \in C} \sum_{i \in F_j^\gamma} D(i,j) \cdot x_{i,j} & & \qquad (\mathrm{ILP}^\gamma) \\ \sum_{i \in F} y_i \le k & & \\ \sum_{i \in F_j^\gamma} x_{i,j} \ge 1 & \qquad \forall j \in C & \\ y_i - x_{i,j} \ge 0 & \qquad \forall j \in C, i \in F_j^\gamma & \\ x_{i,j}, y_i \in \{0,1\} & \qquad \forall j \in C, i \in F_j^\gamma & \end{aligned}$$

ILP$^\gamma$ has a variable y_i for each $i \in F$ that indicates whether the 'facility' i is opened, and a variable $x_{i,j}$ for any combination of an original point j and a facility $i \in F_j^\gamma$ that says whether j is connected to i.

Let (x, y) be any solution of ILP$^\gamma$ and let $c(x, y)$ be the cost of the solution. We relate the solutions of ILP$^\gamma$ to the problem (P1) by the following lemmas.

▶ **Lemma 3.** *Given a solution (x, y) of ILP$^\gamma$, there exists a solution $\mathcal{C} = \{C_l\}_{l=1}^k$ to* (P1) *which has cost no more than $c_a(\mathcal{C}) \le (\gamma + 1) c(x, y)$.*

Proof. Since (x, y) is an integral solution, let $\{p_1, \ldots, p_k\} \subseteq F$ denote the facilities which are opened. We define the cluster C_i to be the set of clients j such that $x_{p_i,j} = 1$. For any client $j \in C_i$, let $j' \in C_i$ be the client which maximizes $D(j, j')$. Since D is a metric, we know that $D(j, j') \leq D(p_i, j) + D(p_i, j')$. By the locality constraint, it holds that $D(p_i, j') \leq \gamma \mathsf{R}_{p_i}$. By definition of D and R_{p_i}, we know $D(p_i, j) \geq \mathsf{R}_{p_i}$, which implies $D(p_i, j') \leq \gamma D(p_i, j)$. Hence, $D(j, j') \leq (\gamma + 1)D(p_i, j)$. Summing this over all clients, we conclude that $c_a(\mathcal{C}) \leq (\gamma + 1)c(x, y)$. ◀

▶ **Lemma 4.** *Given a solution $\mathcal{C}$ to* (P1)*, we can construct a solution (x, y) to ILP$^\gamma$ (where $\gamma \geq 3$) which has cost $\frac{1}{2}c_a(\mathcal{C}) \leq c(x, y) \leq 2c_a(\mathcal{C})$.*

Proof. Fix a cluster C_i, let $j_1, j_2 \in C_i$ be two clients which maximize $D(j_1, j_2)$. Open facility $p_{j_1 j_2}$ and connect all clients in C_i to it. Notice that it holds $D(j, p_{j_1 j_2}) = \mathsf{R}_{p_{j_1 j_2}} + \min\{D(j, j_1), D(j, j_2)\} \leq 3\mathsf{R}_{p_{j_1 j_2}}$ because $D(j_1, j_2)$ is the maximum distance between two clients in C_i and because $D(j_1, j_2) = 2\mathsf{R}_{p_{j_1 j_2}}$. Thus, the solution is feasible for $\gamma = 3$.

For any client $j \in C_i$, let $j' \in C_i$ be the element which maximizes $D(j, j')$. For the first inequality notice that each client j will pay at least $D(j, p_{j_1 j_2}) \geq \frac{1}{2}D(j_1, j_2) \geq \frac{1}{2}D(j, j')$. Observe that $D(j, j') \geq \max\{D(j, j_1), D(j, j_2)\} \geq (D(j, j_1) + D(j, j_2))/2 \geq D(j_1, j_2)/2$. Combining this with the observation that $D(j, j') \geq \min\{D(j, j_1), D(j, j_2)\}$, we get that

$$D(p_{j_1 j_2}, j) = \min\{D(j, j_1), D(j, j_2)\} + D(j_1, j_2)/2 \leq 2 \cdot D(j, j').$$

Summing over all clients gives the second inequality. ◀

Let opt^γ_{ilp} be the optimal value for ILP$^\gamma$ and let opt_a be the value of an optimal solution for (P1). Lemma 4 implies that $opt^3_{ilp} \leq 2 \cdot opt_a$. Assuming we compute an ψ-approximate solution to the optimal ILP^3 solution that violates the locality constraint by an additional factor of ϱ. Then this solution costs at most $\psi \cdot opt^3_{ilp} \leq 2\psi \cdot opt_a$ and violates the locality constraints by 3ϱ. By Lemma 3, we can then construct a feasible solution for (P1) that costs at most $(3\varrho + 1) \cdot 2\psi \cdot opt_a$.

3.1 Good fractional solutions for the local k-median problem

Since problem ILP$^\gamma$ is NP-hard, we relax the integrality constraints to obtain a linear program. The only difference between the standard k-median relaxation and LP$^\gamma_{\mathcal{P}}$ is the locality constraint, i.e., each client j can only connect to facilities in F^γ_j.

$$\begin{aligned}
\min \quad & \sum_{i,j} D(i,j) x_{i,j} && (\text{LP}^\gamma_{\mathcal{P}}) \\
\text{s.t.} \quad & \sum_{i \in F^\gamma_j} x_{i,j} \geq 1 && \forall j \in C \\
& y_i - x_{i,j} \geq 0 && \forall j \in C, i \in F^\gamma_j \\
& \sum_{i \in F} -y_i \geq -k \\
& x, y \geq 0.
\end{aligned}$$

$$\begin{aligned}
\max \quad & \sum_{j} \alpha_j - kZ && (\text{LP}^\gamma_{\mathcal{D}}) \\
\text{s.t.} \quad & \\
& \alpha_j \leq D(i,j) + \beta_{i,j} && \forall j \in C, i \in F^\gamma_j \\
& \sum_{j : i \in F_j} \beta_{i,j} \leq Z && \forall i \in F \\
& \alpha, \beta, Z \geq 0.
\end{aligned}$$

The above LP is very similar to the LP for facility location and this fact was exploited by Jain and Vazirani to show that primal-dual solutions to the facility location problem can be transformed into the solutions for the k-median problem. Let LP-F$^\gamma_{\mathcal{P}}$ be the facility location variant of LP$^\gamma_{\mathcal{P}}$, and let LP-F$^\gamma_{\mathcal{D}}$ be its dual:

$$\min \sum_{i \in F, j \in C} D(i,j) x_{i,j} + \sum_{i \in F} f_i y_i \qquad (\text{LP-F}_{\mathcal{P}}^{\gamma})$$
$$\text{s.t.} \sum_{i \in F_j^{\gamma}} x_{i,j} \geq 1 \qquad \forall j \in C$$
$$y_i - x_{i,j} \geq 0 \qquad \forall j \in C, i \in F_j^{\gamma}$$
$$x, y \geq 0.$$

$$\max \sum_{j} \alpha_j \qquad (\text{LP-F}_{\mathcal{D}}^{\gamma})$$
$$\text{s.t.}$$
$$\alpha_j \leq D(i,j) + \beta_{i,j} \ \forall j \in C, i \in F_j^{\gamma}$$
$$\sum_{j : i \in F_j^{\gamma}} \beta_{i,j} \leq f_i \qquad \forall i \in F$$
$$\alpha, \beta \geq 0.$$

Augmenting ideas introduced by Jain and Vazirani [17], we obtain integer solutions to LP-F$_{\mathcal{P}}^{\gamma}$. This produces two solutions (x^1, y^1) and (x^2, y^2) that are nearly feasible for LP$_{\mathcal{P}}^{3\gamma}$, but $\sum_i y_i^1 = k_1 < k$ and $\sum_i y_i^2 = k_2 > k$. A suitable convex combination of the two is a feasible solution for LP$_{\mathcal{P}}^{3\gamma}$ and is a constant factor away from the optimal value of LP$_{\mathcal{P}}^{\gamma}$.

▶ **Lemma 5.** *Given any $\epsilon > 0$ and $\gamma > 0$, there exists a polynomial time algorithm which finds two feasible integer solutions $(x^1, y^1), (x^2, y^2)$ for LP-F$_{\mathcal{P}}^{3\gamma}$ with the following properties:*
1. *$\sum_i y_i^1 = k_1$ and $\sum_i y_i^2 = k_2$ for two integers $k_1 < k < k_2$.*
2. *Set $\rho = \frac{k_2 - k}{k_2 - k_1}$. The solution $(\hat{x}, \hat{y}) = \rho(x^1, y^1) + (1 - \rho)(x^2, y^2)$ is feasible for LP$_{\mathcal{P}}^{3\gamma}$ with cost at most $(3 + \epsilon)$ times the optimal solution to LP$_{\mathcal{P}}^{\gamma}$.*

Since the essential ideas behind this lemma use standard techniques, we omit the full proof because of space limitations. The main differences to the standard Jain-Vazirani primal-dual process are as follows: When finding the initial set of open facilities, we restrict clients to paying and connecting only to facilities whose radius they lie in. In the clean-up step, the Jain-Vazirani algorithm selects the finally open facilities by finding an arbitrary independent set of facilities in some graph. We use the freedom to choose any independent set and choose a set that ensures that clients that have to be reassigned (because their original facility was closed) can always be routed to an open facility with higher radius than their original facility.

3.2 Rounding

Given any two integer solutions (x_1, y_1) and (x_2, y_2) for LP-F$_{\mathcal{P}}^{\gamma}$, which open $A, B \subseteq F$ facilities, respectively, we define a weighted bipartite graph $G(x_1, y_1, x_2, y_2)$ as follows. The graph is defined on the vertex set with bipartitions A and B. We connect $i \in A$ to $i' \in B$ if there exists a client j such that $x_{i,j}^1 = 1$ and $x_{i',j}^2 = 1$. The weight of an edge (i, i') is the number of clients j which satisfy the above requirement.

▶ **Lemma 6.** *The following holds for local k-median instances that arise from the aversion k-clustering problem. Given two integer solutions $(x^1, y^1), (x^2, y^2)$ for LP-F$_{\mathcal{P}}^{\gamma}$ which open facilities $A, B \subseteq F$, respectively, we can construct solutions $(\tilde{x}^1, \tilde{y}^1), (\tilde{x}^2, \tilde{y}^2)$ that satisfy:*
1. *$(\tilde{x}^1, \tilde{y}^1)$ opens facilities A and $(\tilde{x}^2, \tilde{y}^2)$ opens facilities B.*
2. *If $(x_1, y_1), (x_2, y_2)$ are feasible for LP-F$_{\mathcal{P}}^{\gamma}$, then $(\tilde{x}^1, \tilde{y}^1), (\tilde{x}^2, \tilde{y}^2)$ are feasible solutions to LP-F$_{\mathcal{P}}^{3\gamma}$, and they satisfy $c(\tilde{x}^1, \tilde{y}^1) \leq 3\gamma c(x^1, y^1)$ and $c(\tilde{x}^2, \tilde{y}^2) \leq 3\gamma c(x^2, y^2)$.*
3. *The graph $G(\tilde{x}^1, \tilde{y}^1, \tilde{x}^2, \tilde{y}^2)$ is a forest.*

Proof. We will assume that all radii are distinct (we can ensure this, e.g., by adding a tiny amount of noise to all the radii, or by breaking ties consistently). We say that

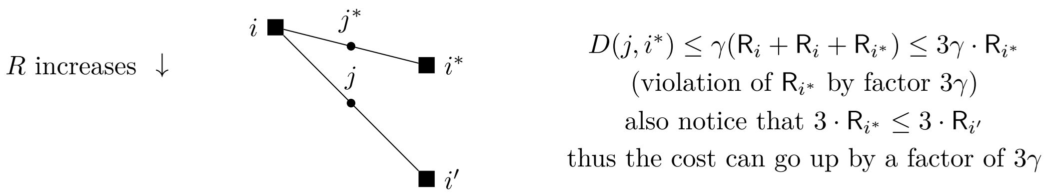

Figure 1 Removing all but one down edge for client i.

an edge $\{i, i'\}$ in $G(x^1, y^1, x^2, y^2)$ is a *down edge for* i if $\mathsf{R}_{i'} > \mathsf{R}_i$. For $i \in A \cup B$, let $\mathfrak{D}(i) := \{i' \mid \{i, i'\} \text{ is a down edge}\}$ be the set of facilities that are connected to i by down edges. Furthermore, for every $i \in A \cup B$, let i^* be a facility that minimizes $\{\mathsf{R}_i \mid i \in \mathfrak{D}(i)\}$, i.e., i^* is the endpoint of a 'highest' down edge. For each $i \in A \cup B$, we modify assignments as follows. For all clients $j \in C$ connected to i, and to some facility $i' \in \mathfrak{D}(i)$ in (x_1, y_1), (x_2, y_2), we reassign them to now connect to i and i^* in $(\tilde{x}^1, \tilde{y}^1), (\tilde{x}^2, \tilde{y}^2)$. Thus, for all clients $j \in C$ originally connected to i and i^*, the assignment does not change.

Let us calculate the costs of the resulting assignment. Let $i' \in \mathfrak{D}(i)$ be a facility with $i' \neq i^*$ and let j be a client that is reconnected from i' to i^*. Notice that $D(i, i^*) \leq \gamma \mathsf{R}_i + \gamma \mathsf{R}_{i^*}$ since at least one client lies within the (γ-expanded) radius of i and i^* simultaneously. We observe that $D(j, i^*) \leq D(j, i) + D(i, i^*) \leq \gamma(\mathsf{R}_i + \mathsf{R}_i + \mathsf{R}_{i^*}) \leq 3\gamma \cdot \mathsf{R}_{i^*}$ by the triangle inequality and by $\mathsf{R}_{i^*} \geq \mathsf{R}_i$. Thus, the new solution violates the locality constraint for j by a factor of at most 3. Since R_{i^*} is the smallest radius for all facilities in $\mathfrak{D}(i)$, it holds that $\mathsf{R}_{i^*} \leq \mathsf{R}_{i'}$. Thus, we also have $D(j, i^*) \leq 3\gamma \mathsf{R}_{i'}$. Moreover, since the instances arise from local k-median, equations (1) and (2) imply that $D(j, i') \geq \mathsf{R}_{i'}$. (This is the only part of the proof that relies on the local k-median instances arising from aversion k-clustering). Hence we have $D(j, i^*) \leq 3\gamma \mathsf{R}_{i'} \leq 3\gamma D(j, i')$. Thus the cost of each client j is increased by a factor of at most 3γ, which immediately proves Property 2. (Figure 1 visualizes this calculation).

We do not open or close any facilities, thus Property 1 is true. To see Property 3 holds, note that by the distinct radii assumption, any cycle would contain a facility with two down edges, which is no longer possible after the reassignment. ◀

Lemma 6 transforms our solution such that it corresponds to a forest T on the vertices $A \cup B$. We first assume that T is a tree and later deal with each connected component separately. We use the tree structure to define a depending rounding procedure to combine A and B into an integral solution C with low cost. It will be crucial to look at the difference between the number of vertices from B and A in subtrees of T.

▶ **Definition 7.** For any subtree of $T' \subseteq T$, we define the *deficiency* of T' to be $\mathsf{df}(T') = |B(T')| - |A(T')|$ where $B(T')$ (and $A(T')$) are the vertices from B (and A) in this subtree.

We start with $C = B$. Then we find a subtree T' with $\mathsf{df}(T') = 1$, i.e., one node more from B than from A. We want to close all facilities in $B(T')$, open all facilities in $A(T')$ and reconnect the affected clients. We want that the reassignment follows the assignments in $(\tilde{x}^1, \tilde{x}^2)$, so all facilities in A that are adjacent to $B(T')$ must be contained in $A(T')$. We then iterate this process with more subtrees until C has exactly k vertices. Since we gain one for every subtree, we need $t := k_2 - k$ subtrees until $|C|$ was reduced from k_2 to k. The subtrees must be disjoint on the B side, while the vertices from A can overlap. The following lemma shows that we can find a large set of suitable subtrees from which we can choose the cheapest t later. Figure 2 visualizes two examples.

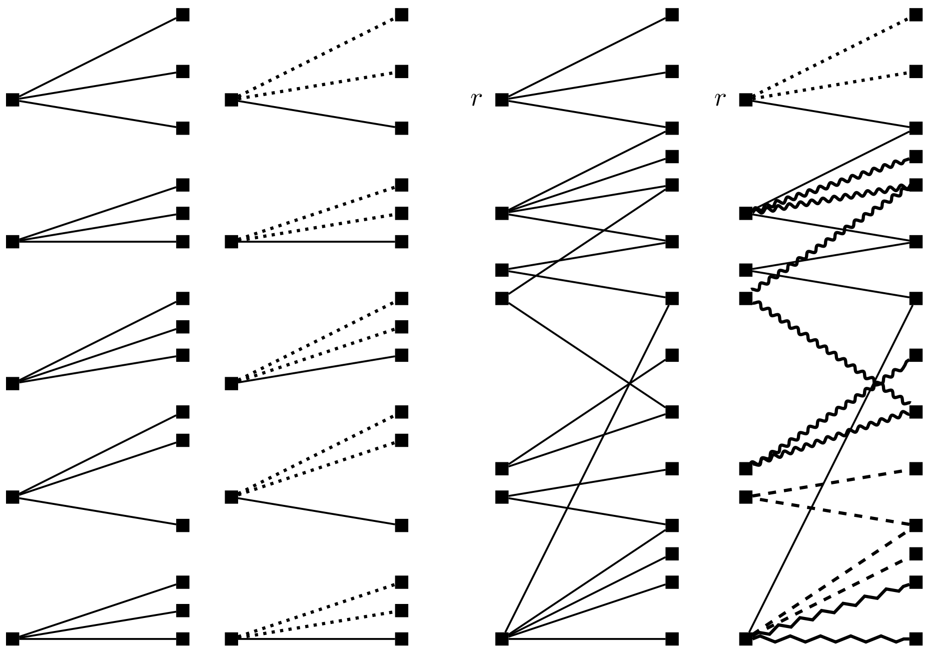

Figure 2 Examples on finding subtrees of $G(x^1, y^1, x^2, y^2)$ with $\mathsf{df}(T_i) = 1$. The two left pictures show a simple special case that also is a worst case for the number of subtrees: The deficiency of the shown forest F is $\mathsf{df}(F) = 2k = 10$ and we get $\mathsf{df}(F)/2 = 5$ subtrees by pairing the nodes from B. The two right pictures show a connected tree T with $\mathsf{df}(T) = 7$ and 4 subtrees with $\mathsf{df}(T_i) = 1$.

▶ **Lemma 8.** *Given any tree T with $\lceil \mathsf{df}(T)/2 \rceil = l$ and root $r \in A$, we can find l subtrees $T_1, \dots, T_l$ of T with*
1. $\mathsf{df}(T_i) = 1$
2. $B(T_i) \cap B(T_j) = \emptyset$
3. $A(\delta(B(T_i))) \subseteq A(T_i)$

where we use $\delta(X)$ to denote the set of edges from X to $\bar{X}$.

Proof. Let r be the root of T and $c_1, \dots, c_\nu$ be the children of r. Our proof will proceed with induction on the height of T. By *removing* a subtree T' we mean that we remove all edges that are in T' from T and all vertices except the root of T'.

Induction Hypothesis: There exist subtrees $T_1, \dots, T_z$, $z \in \mathbb{N}_0$, of T that satisfy:
1. Each subtree T_i is rooted at a vertex in A and satisfies that $\mathsf{df}(T_i) = 1$.
2. After removing $T_1, \dots, T_z$ from T, the following holds. If $r \in B$ then $\mathsf{df}(T) \le 1$. If $r \in A$ then $\mathsf{df}(T) \le 0$.

Base Case: T has height 0 or 1, i.e., it is a star. If $r \in B$, then $\mathsf{df}(T) \le 1$ because there is only one node from B. If $r \in A$, then we can remove the children in pairs until there are no pairs left. This is because the subtree consisting of r and any two of its children has deficiency 1. Therefore, each pair and the root will correspond to a subtree (one of T_i mentioned in the IH) that we remove. After removing them, T consists of only r or r and one node from B. In both cases, $\mathsf{df}(T) \le 0$.

Induction Step:

Case $r \in B$: By the induction hypothesis (IH), we can remove some subtrees to ensure that each subtree rooted at $c_1 \dots c_\nu$ will have deficiency at most 0. Since $\mathsf{df}(T) = \sum_{i=1}^{\nu} \mathsf{df}(T_{\text{rooted at } c_i}) + 1 \le 1$, we can conclude that this satisfies the first property in the

IH. Since we didn't remove any additional subtrees, the second property is vacuously satisfied.

Case $r \in A$: By the IH, we know that the subtree rooted at each child c_i has deficiency $\mathsf{df}(c_i) \leq 1$. Without loss of generality, let $c_1, \ldots, c_p$ be the children which have deficiency 1 and $c_{p+1} \ldots c_\nu$ have deficiency ≤ 0. If $p \leq 1$, then $\mathsf{df}(T) \leq 0$. If $p \geq 2$, then we remove pairs of children with positive deficiency. Observe that the subtree rooted at r containing only the children c_1 and c_2 has deficiency exactly 1. Hence, these satisfy the second property in the IH. We continue this process until there is at most 1 child which has positive deficiency, at which point the the first property is satisfied. This ensures that the induction step is satisfied.

Notice that each removed subtree has deficiency one. Since we keep the root, the deficiency decreases by two for each removed subtree. When $r \in A$ as assumed in the lemma, then $\mathsf{df}(T)$ is decreased to at most zero. Thus, at least $\lceil \mathsf{df}(T)/2 \rceil$ subtrees are removed. ◀

For a forest F consisting of trees $F_1, \ldots, F_x$, set $\mathsf{df}(F) := \sum_{j=1}^{x} \mathsf{df}(F_j)$. We can find $\lceil \mathsf{df}(F_j)/2 \rceil$ subtrees satisfying the properties of Lemma 8 for every every F_j. Thus, we get

$$\sum_{j=1}^{x} \left\lceil \frac{\mathsf{df}(F_i)}{2} \right\rceil \geq \left\lceil \frac{1}{2} \sum_{j=1}^{x} \mathsf{df}(F_i) \right\rceil = \lceil \mathsf{df}(F)/2 \rceil$$

subtrees, giving the following corollary.

▶ **Corollary 9.** *Given any forest F with $\lceil \mathsf{df}(F)/2 \rceil = l$, we can find l subtrees $T_1, \ldots, T_l$ of F satisfying the properties from Lemma 8.*

We now show Theorem 2. We are given an instance of the local k-median problem that arises from the aversion k-clustering problem. We know that the solutions for the local k-median problem that are induced by the aversion k-clustering instance are feasible for LP-$\mathrm{F}^3_{\mathcal{P}}$. Thus, we set $\gamma := 3$. Then we use Lemma 5 and Lemma 6 to get two solutions (x^1, y^1) and (x^2, y^2) so that the graph $G(x^1, y^1, x^2, y^2)$ is a forest, (x^1, y^1) opens k_1 facilities and (x^2, y^2) opens $k_2 \geq k_1$ facilities. Both Lemma 5 and Lemma 6 induce a factor of 3 in the radius violation, so (x^1, y^1) and (x^2, y^2) are feasible for LP-$\mathrm{F}^{9\gamma}_{\mathcal{P}}$. Furthermore, the intermediate solutions $(\hat{x}^1, \hat{y}^1)$ and $(\hat{x}^2, \hat{y}^2)$ coming from Lemma 5 have the property that for $\rho = (k_2 - k)/(k_2 - k_1)$, it holds that $\rho \cdot c(\hat{x}^1, \hat{y}^1) + (1 - \rho) \cdot (\hat{x}^2, \hat{y}^2) \leq (3 + \varepsilon) \cdot opt^{\gamma}_{l}$. Applying Lemma 6 increases the cost bound by a factor of 3γ. Thus, we know that

$$\rho \cdot c(x^1, y^1) + (1 - \rho) \cdot c(x^2, y^2) \leq (3 + \varepsilon) \cdot 3\gamma \cdot opt^{\gamma}_{l} := c^{mix}$$

for $\rho = (k_2 - k)/(k_2 - k_1)$. If (x^1, y^1) or (x^2, y^2) opens exactly k facilities, we are done. Otherwise, $k_1 < k < k_2$. If $\rho \geq 1/2$, simply output (x^1, y^1) which then costs $c(x^1, y^1) \leq 2\rho \cdot c(x^1, y^1) \leq 2c^{mix}$. We assume that this is not the case, i.e., $\rho < 1/2$.

We build a solution C and start with $C = B$. Using Corollary 9, we find $\frac{1}{2}(k_2 - k_1)$ subtrees $T_1, \ldots, T_\ell$ of $G(x^1, y^1, x^2, y^2)$. For each subtree T_i, we can reassign the clients from the facilities in $B(T_i)$ to the facilities in $A(T_i)$. We denote the connection cost for assigning the clients to $A(T_i)$ by $c(T_i)$. Notice that $c(x^1, y^1) \geq \sum_{s=1}^{\ell} c(T_i)$ because every edge of T can only appear in one subtree (since the $B(T_i)$ are pairwise disjoint). Thus, if we choose the $t = k_2 - k$ subtrees $T_{i_1}, \ldots, T_{i_t}$ with the cheapest $c(T_i)$, then

$$\sum_{z=1}^{t} c(T_{i_z}) \leq \frac{t}{\ell} \sum_{s=1}^{\ell} c(T_i) \leq \frac{t}{\ell} c(x^1, y^1) = \frac{k_2 - k}{\frac{1}{2}(k_2 - k_1)} c(x^1, y^1) = 2\rho \cdot c(x^1, y^1).$$

The cost of C starts at $c(x^2, y^2)$ and is increased by at most $2\rho \cdot c(x^1, y^1)$. Thus, the solution costs at most $2\rho \cdot c(x^2, y^2) + c(x^2, y^2) \le 2\rho \cdot c(x^2, y^2) + 2(1-\rho) \cdot c(x^2, y^2) \le 2 \cdot c^{mix}$ where we recall that $\rho < 1/2$. Thus, we get an integer solution of cost $2 \cdot (3+\varepsilon) \cdot 3\gamma \cdot opt_l^\gamma$ that is feasible for $\mathrm{LP}_{\mathcal{P}}^{9\gamma}$. That induces a solution for the aversion k-clustering instance that is a constant factor approximation as we described below Lemma 4.

3.3 Improving the Approximation Factor

To improve the final approximation ratio for the aversion k-clustering problem, we observe that the dual variables computed by the primal-dual algorithm can be directly related to the objective of the aversion k-clustering problem. We split each such dual: Let $\alpha_O(j)$ denote the amount that client j pays to open a facility (the subscript O stands for "open"). Using the terminology of Jain and Vazirani, we say a client j is directly connected to facility i if $\beta_{i,j} > 0$ and facility i is open. In this case, $\alpha_O(j) := \beta_{i,j}$. Otherwise, $\alpha_O(j) = 0$. Define $\alpha_C(j) := \alpha(j) - \alpha_O(j)$ (intuitively, this is the connection cost—the subscript C is for connection—that the client has paid for, but for indirectly connected clients we only know that $D(i,j) \ge \alpha_C(j) \ge (1/3)D(i,j)$ is true. For directly connected clients, $\alpha_C(j) = D(i,j)$.).

▶ **Lemma 10.** *At the end of the primal-dual algorithm, if client j connects to facility i, then $\alpha_C(j) \ge \mathsf{R}_i$.*

Proof. If j is directly connected to i, then it is immediate that $\alpha_C(j) = D(i,j) \ge \mathsf{R}_i$.

Suppose that j is indirectly connected to facility i'. In this case, let i be the facility that j was first connected to. Since j is indirectly connected, there has to be a client j' that has special edges to both i and i'. We use $t(i)$ and $t(i')$ to denote the times at which facilities i and i' were respectively opened. Notice that $\alpha_C(j) = \alpha(j)$ by definition of α_C for indirectly connected clients and that $\alpha(j) = t(i)$ because j was connected to i before.

Case $t(i) \ge t(i')$: In this case we know that $\alpha_C(j) = t(i) \ge t(i') \ge D(i',j') \ge \mathsf{R}_{i'}$.

Case $t(i) < t(i')$: Since j' has special edges to i and i', it had tight edges to both before either was opened, i.e., $D(i',j') \le t(i)$. Thus we can say $\alpha_C(j) = t(i) \ge D(i',j') \ge \mathsf{R}_{i'}$. ◀

Once again, we may assume that the Jain-Vazirani algorithm returns two solutions (x^1, y^1), (x^2, y^2) and their duals (α^1, Z^+) and (α^2, Z^-). It follows from Jain and Vazirani's analysis that the solutions have the following properties.

1. $\sum_i y_i^1 = k_1$ and $\sum_i y_i^2 = k_2$.
2. $\sum_j \alpha_C^1(j) = \sum_j \alpha_j^1 - k_1 \cdot Z^+$ and $\sum_j \alpha_C^2(j) = \sum_j \alpha_j^2 - k_2 \cdot Z^-$
3. $x_{i,j}^1 = 1$ or $x_{i,j}^2 = 1 \implies D(i,j) \le 3\gamma \mathsf{R}_i$
4. $|Z^+ - Z^-| \le \epsilon$
5. For $\rho = \frac{k_2 - k}{k_2 - k_1}$, $\rho(\alpha^1, Z^+) + (1-\rho)(\alpha^2, Z^-)$ is feasible for $\mathrm{LP}_{\mathcal{D}}^\gamma$.

For property 2, notice that $\alpha_C^1(j) = \alpha_j^1$ for indirectly connected clients, that $\alpha^1 C(j) = \alpha_j^1 - \beta_{\phi(j)j}$ for directly connected clients (where $\phi(j)$ is the center j is connected to) and that the sum of $\beta_{\phi(j)j}$ over all directly connected clients is just $k_1 Z^+$. The same holds for the second solution. Using property 5, we get that $\rho(\sum_j \alpha_j^1 - k \cdot Z^+) + (1-\rho)(\sum_j \alpha_j^2 - k_2 \cdot Z^-)$ is a lower bound for the optimal value of $\mathrm{LP}_{\mathcal{D}}^\gamma$ and thus also for the optimal value of $\mathrm{LP}_{\mathcal{P}}^\gamma$. Using property 2 and 4, this implies that $\rho(\sum_j \alpha_C^1(j)) + (1-\rho)(\sum_j \alpha_C^2(j)) \le opt(\mathrm{LP}_{\mathcal{P}}^\gamma) + \varepsilon$. We apply Lemma 6 to replace x^1 and x^2 to ensure that the resulting graph $G(x^1, y^1, x^2, y^2)$ is a forest. Note that the procedure only reassigns the clients to facilities with smaller radius than their currently connected facility. Hence, we can still assume that $x_{i,j}^1 = 1 \implies \alpha_C^1(j) \ge \mathsf{R}_i$

(similarly $x^2_{i,j} = 1 \implies \alpha^2_C(j) \geq \mathsf{R}_i$). However, we may now have solutions that violate the locality constraints by a factor of 9γ.

Now we use the procedure described in Lemma 8 to partition the graph $G(x^1, y^1, x^2, y^2)$ into subtrees $T_1, \ldots T_\ell$ with $\mathsf{df}(T_p) = 1$ for $p \in \{1, \ldots, \ell\}$ and $\ell = \frac{k_2 - k_1}{2}$. Each tree has the property that $A(\delta(B(T_p))) \subseteq A(T_p)$ for all $p \in \{1, \ldots, \ell\}$. Since each edge in this tree represents some set of clients, we use the notation $j \in T_p$ to denote that j is associated with an edge in T_p. Define the cost of the subtree T_p as $\sum_{j \in T_p} \alpha^1_C(j)$. We choose the $k_2 - k$ cheapest such trees. Since choosing all ℓ subtrees will result in a cost of $\sum_j \alpha^1_C(j)$, we can say that the cost of these chosen subtrees is at most $\frac{2(k_2-k)}{k_2-k_1} \sum_j \alpha^1_C(j) = 2\rho \sum_j \alpha^1_C(j)$.

Our rounded solution $(\hat{x}, \hat{y})$ opens all facilities from A that are part of the chosen subtrees and all facilities from B that are not part of any chosen subtree. Notice that since we open $k_2 - k$ subtrees and these satisfy $\mathsf{df}(T_p) = 1$, $\hat{y}$ opens exactly k facilities. The assignments of clients to facilities follow x^1 and x^2, respectively.

Analogously to Lemma 3, we construct a solution to the aversion k-clustering problem based on $\hat{x}, \hat{y}$. In this solution, each client assigned to a facility $p_i \in A$ pays at most

$$D(j, j') \leq D(p_i, j) + D(p_j, j') \leq 9\gamma \mathsf{R}_i + 9\gamma \mathsf{R}_i \leq (2 \cdot 9\gamma)\alpha^1_C(j)$$

where j' is the furthest away client among all that are assigned to p_i. Thus, by our choice of subtrees, all clients assigned to A pay at most $(2 \cdot 9\gamma)2\rho \sum_j \alpha^1_C(j)$ in total. The remaining clients pay at most $(2 \cdot 9\gamma)(\sum_j \alpha^2_C(j))$. As before, we can assume that $\rho \leq 1/2$. We conclude that the cost of $(\hat{x}, \hat{y})$ is bounded by

$$\begin{aligned}
&(2 \cdot 9\gamma)\big(2\rho \sum_j \alpha^1_C(j) + \sum_j \alpha^2_C(j)\big) \\
&\leq 2(2 \cdot 9\gamma)(1+\varepsilon)\big(\rho \sum_j \alpha^1_C(j) + (1-\rho) \sum_j \alpha^2_C(j)\big) \\
&\leq 2(2 \cdot 9\gamma)(1+\varepsilon)\mathit{opt}(\mathrm{LP}^\gamma_{\mathcal{P}}) \\
&\leq 2(2 \cdot 9\gamma)2(1+\varepsilon)\mathit{opt}_a
\end{aligned}$$

where opt_a is the optimal value for the aversion k-clustering instance and the last inequality follows by Lemma 4. Since $\gamma = 3$, the approximation factor is bounded by $216 + \varepsilon$.

4 Final Thoughts and Conclusions

This paper shows a $(216+\epsilon)$-approximation to the aversion k-clustering problem. Our results rely on achieving a constant factor bicriteria approximation for local k-median instances arising from aversion k-clustering problem. Lemma 6 is the only place in our proof where we use that the local k-median instances are generated from aversion k-clustering instances. It remains an open question if we can get a constant factor bicriteria approximation for arbitrary instances of local k-median.

Acknowledgments. We thank Christian Kroer for introducing us to the aversion k-clustering problem, and for enlightening discussions. We thank anonymous reviewers for constructive feedback, in particular for pointing out an oversight in Lemma 6 in an earlier draft.

References

1 Hyung-Chan An, Aditya Bhaskara, Chandra Chekuri, Shalmoli Gupta, Vivek Madan, and Ola Svensson. Centrality of trees for capacitated k-center. *Mathematical Programming*, 154(1-2):29–53, 2015.

2 Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. Local search heuristics for k-median and facility location problems. *SIAM Journal on Computing*, 33(3):544–562, 2004.

3 Michel Louis Balinski. On finding integer solutions to linear programs. Technical report, DTIC Document, 1964.

4 Yair Bartal, Moses Charikar, and Danny Raz. Approximating min-sum k-clustering in metric spaces. In *Proceedings of the 33rd STOC*, pages 11–20, 2001.

5 Babak Behsaz, Zachary Friggstad, Mohammad R. Salavatipour, and Rohit Sivakumar. Approximation algorithms for min-sum k-clustering and balanced k-median. In *Proceedings of the 42nd ICALP*, pages 116–128, 2015.

6 Jaroslaw Byrka, Thomas Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for k-median, and positive correlation in budgeted optimization. In *Proceedings of the 26th SODA*, pages 737–756, 2015.

7 Moses Charikar and Sudipto Guha. Improved combinatorial algorithms for facility location problems. *SIAM Journal on Computing*, 34(4):803–824, 2005.

8 Moses Charikar, Sudipto Guha, Éva Tardos, and David B. Shmoys. A constant-factor approximation algorithm for the k-median problem. *Journal of Computer and System Sciences*, 65(1):129–149, 2002.

9 Moses Charikar and Rina Panigrahy. Clustering to minimize the sum of cluster diameters. *Journal of Computer and System Sciences*, 68(2):417–441, 2004.

10 Julia Chuzhoy and Yuval Rabani. Approximating k-median with non-uniform capacities. In *Proceedings of the 16th SODA*, pages 952–958, 2005.

11 Marek Cygan, MohammadTaghi Hajiaghayi, and Samir Khuller. LP rounding for k-centers with non-uniform hard capacities. In *Proceedings of the 53rd FOCS*, pages 273–282, 2012.

12 Teofilo F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.

13 Sudipto Guha and Samir Khuller. Greedy strikes back: Improved facility location algorithms. *Journal of Algorithms*, 31(1):228–248, 1999.

14 Dorit S. Hochbaum and David B. Shmoys. A best possible heuristic for the k-center problem. *Mathematics of Operations Research*, 10:180–184, 1985.

15 Wen-Lian Hsu and George L. Nemhauser. Easy and hard bottleneck location problems. *Discrete Applied Mathematics*, 1:209–215, 1979.

16 Kamal Jain, Mohammad Mahdian, and Amin Saberi. A new greedy approach for facility location problems. In *Proceedings of the 34th STOC*, pages 731–740, 2002.

17 Kamal Jain and Vijay V. Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and lagrangian relaxation. *Journal of the ACM*, 48(2):274–296, 2001.

18 Richard M. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, pages 85–103, 1972.

19 Christian Kroer and Tuomas Sandholm. Extensive-Form Game Imperfect-Recall Abstractions With Bounds. *CoRR*, abs/1409.3302, 2014. also published at the Algorithmic Game Theory workshop at IJCAI, 2015.

20 Shi Li. A 1.488 approximation algorithm for the uncapacitated facility location problem. *Information and Computation*, 222:45–58, 2013.

21 Shi Li and Ola Svensson. Approximating k-median via pseudo-approximation. In *Proceedings of the 45th STOC*, pages 901–910, 2013.

22 David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011. Available at `http://www.designofapproxalgs.com`.