Analysis of a Local Search Heuristic for Facility Location Problems*

Madhukar R. Korupolu[1] *C. Greg Plaxton*[1] *Rajmohan Rajaraman*[2]

Abstract

In this paper, we study approximation algorithms for several NP-hard facility location problems. We prove that a simple local search heuristic yields polynomial-time constant-factor approximation bounds for the metric versions of the uncapacitated k-median problem and the uncapacitated facility location problem. (For the k-median problem, our algorithms require a constant-factor blowup in the parameter k.) This local search heuristic was first proposed several decades ago, and has been shown to exhibit good practical performance in empirical studies. We also extend the above results to obtain constant-factor approximation bounds for the metric versions of capacitated k-median and facility location problems.

1 Introduction

Suppose that you plan to build a new chain of widget stores in a given city, and you have identified potential store sites in a number of different neighborhoods. Further assume that the demand for widgets in each neighborhood of the city is known. If you want to build exactly k stores, where should you locate them in order to minimize the average traveling distance of your customers? If instead you are willing to build any number of stores, and the cost of building a store at each potential site is known, where should you build stores in order to minimize the sum of the total construction cost and the average traveling distance of your customers?

The preceding questions are examples of *facility location problems*. This paper is concerned with metric versions of the following specific facility location problems: (i) the uncapacitated k-median problem, (ii) the uncapacitated facility location problem, (iii) the capacitated k-median problem with unsplittable demands, (iv) the capacitated facility location problem with unsplittable demands, (v) the capacitated k-median problem with splittable demands, and (vi) the capacitated facility location problem with splittable demands. Formal definitions of these problems, all of which are known to be NP-hard, are given in Section 2.

Informally, the uncapacitated k-median (resp., facility location) problem corresponds to the first (resp., second) of the two questions stated above, where we assume that there is no upper bound on the demand that can be satisfied by a given store, and hence the traveling distance of a customer is just the distance from that customer's neighborhood to the nearest store. In the capacitated versions of these problems, there is an upper bound on the demand that can be satisfied by a given store, and we need to exhibit an assignment of customers to stores that minimizes the average traveling distance while respecting the capacity bound of each store. In assigning the demand of customers to stores, there are two natural variations to consider: (i) the demand of a customer must be met by a single store (unsplittable demands), and (ii) the demand of a customer may be divided across any number of stores (splittable demands).

The input to a facility location problem includes a distance matrix specifying the traveling distance from node i to node j, for all i and j. (The nodes correspond to the neighborhoods in our widget chain example.) In the *metric version* of a facility location problem, it is assumed that these distances are nonnegative, symmetric, and satisfy the triangle inequality. All of our results are for metric facility location problems.

The local search heuristic for facility location prob-

*This research was supported by the National Science Foundation under Grant No. CCR–9504145.

[1]Department of Computer Science, University of Texas, Austin, TX 78712. Email: {madhukar,plaxton}@cs.utexas.edu.

[2]DIMACS Center, Rutgers University, Piscataway, NJ 08854. Email: rraj@dimacs.rutgers.edu. This work was done while the author was at the University of Texas at Austin.

lems, described more formally in Section 3, is extremely straightforward. The idea is to start with any feasible solution (set of stores) and then to iteratively improve the solution by repeatedly moving to the best "neighboring" feasible solution, where one solution is a neighbor of another if it can be obtained by either adding a facility (store), deleting a facility, or changing the location of a facility. (Note that in the case of k-median problems, only the third operation is needed since all feasible solutions have exactly k facilities.) This heuristic was proposed by Kuehn and Hamburger [7], and was subsequently shown to exhibit good practical performance in empirical studies (see, e.g., [2, 12]).

For the k-median problems we consider (facility location variants (i), (iii), and (v) above), we define an (a, b)-*approximation algorithm* as a polynomial-time algorithm that computes a solution using at most bk facilities, whose cost is at most a times the cost of the optimal solution that uses at most k facilities. For facility location variants (ii), (iv), and (vi), we define an a-*approximation algorithm* as a polynomial-time algorithm that computes a solution with cost at most a times optimal. Using this terminology, and letting n denote the number of nodes in the input instance, we now summarize the previous approximation results known for the facility location problems considered in this paper: (i) Hochbaum [3] showed that a simple greedy algorithm is an $O(\log n)$-approximation algorithm for the general (i.e., not restricted to metric distances) uncapacitated facility location problem; (ii) Lin and Vitter [9] gave a $(1+\varepsilon, (1+1/\varepsilon)(\ln n+1))$-approximation algorithm for the general uncapacitated k-median problem (their algorithm can also be adapted to match Hochbaum's result); (iii) Lin and Vitter [8] gave a $(2(1+\varepsilon), 1+(1/\varepsilon))$-approximation algorithm for the metric k-median problem; (iv) Shmoys et al. [11] gave a 3.16-approximation algorithm for the metric uncapacitated facility location problem, a 7-approximation algorithm for the metric capacitated facility location problem with splittable demands that requires a blowup of 7/2 in the capacity of each facility, and a 9-approximation algorithm for the metric capacitated facility location problem with unsplittable demands that requires a blowup of 4 in the capacity of each facility.

The approximation algorithms of Lin and Vitter [8, 9] and Shmoys et al. [11] are based on careful rounding of fractional solutions to linear programming relaxations. Thus, these algorithms are substantially more sophisticated than the local search heuristic analyzed in this paper. (Note that known polynomial-time algorithms for linear programming [5, 6] are quite complicated.) Nevertheless, we prove that the local search heuristic yields simple algorithms for the uncapacitated k-median and the uncapacitated facility location problems with approximation bounds matching (to within constant factors) the corresponding results in (iii) and (iv) above. In particular, for the metric uncapacitated k-median problem, we prove that for any $\varepsilon > 0$, local search yields (a, b)-approximation bounds with $(a, b) = (1+\varepsilon, 10+17/\varepsilon)$ and $(a, b) = (\Theta(\varepsilon^{-3}), 3+\varepsilon)$. For the metric uncapacitated facility location problems, we prove that for any $\varepsilon > 0$, local search yields an $(9+\varepsilon)$-approximation algorithm.

Our results for the capacitated facility location problems are similar. For the metric capacitated k-median problem with splittable demands, we prove that for any $\varepsilon > 0$, local search yields (a, b)-approximation bounds with $(a, b) = (1+\varepsilon, 12+17/\varepsilon)$ and $(a, b) = (\Theta(\varepsilon^{-3}), 5+\varepsilon)$ with no blowup in the capacity of each facility. For the metric uncapacitated facility location problems with splittable demands, we prove that for any $\varepsilon > 0$, local search yields an $(18+\varepsilon)$-approximation algorithm with no blowup in the capacity of each facility. Using the reduction of [11] that is based on solving a special case of the generalized assignment problem [10], the above results for the capacitated problems with splittable demands imply constant-factor approximation bounds for the capacitated problems with unsplittable demands assuming a factor of 2 blowup in the capacity of each facility.

The facility location problems that we consider in this paper are closely related to a number of other optimization problems (e.g., the k-center problem) that have been studied in the computer science and operations research literature. We refer the reader to [1, 4] for more information about other problems related to facility location.

The rest of the paper is organized as follows. Section 2 formally defines the six facility location problems considered in this paper. Section 3 defines the local search heuristic that we analyze. Section 4 explains how to find an initial feasible solution. Sections 5 and 6 analyze the uncapacitated k-median and facility location problems, respectively. Section 7 analyzes the capacitated k-median problem with splittable demands. Due to space constraints, our analysis of the three remaining capacitated problems has been omitted; these results will appear in the full version of the paper.

2 Facility Location Problems

Let $N = \{1, 2, \ldots, n\}$ be a set of locations and $F \subseteq N$ be a set of locations at which we may open a facility.

Each location j in N has a demand d_j that must be shipped to j. For any two locations i and j, let c_{ij} denote the cost of shipping a unit of demand from i to j. We assume that the costs are nonnegative, symmetric, and satisfy the triangle inequality. In each problem studied in this paper, we wish to determine a set of open facilities and an assignment of locations to open facilities such that some objective function is minimized.

In the *uncapacitated k-median problem*, we seek a set of at most k open facilities and an assignment of locations to open facilities such that the shipping cost of the solution is minimized, where the shipping cost associated with a a set S of open facilites and an assignment $\sigma : N \mapsto S$ is given by $\sum_{j \in N} d_j c_{j\sigma(j)}$. Given a set S of open facilities, an assignment that minimizes the shipping cost may be obtained by assigning each location $j \in N$ to the closest open facility in S. Thus, any solution to the uncapacitated k-median problem is completely characterized by the set of open facilities. For any set S of open facilities, let $C(S)$ denote the shipping cost obtained by assigning each location to its closest open facility.

In the *uncapacitated facility location problem*, for each location i in F, we are given a nonnegative cost f_i of opening a facility at i. The problem is to determine a set of open facilities and an assignment of locations to facilities such that the sum of the cost of opening the facilities and the shipping cost is minimized. (The shipping cost is defined as in the uncapacitated k-median problem.) As in the uncapacitated k-median problem, given a set S of open facilities, an assignment that minimizes the total cost is to assign each location $j \in N$ to the closest open facility in S. Therefore, any solution to this problem is characterized by the set of open facilities. Given any solution S of open facilities, we let $C_s(S)$, $C_f(S)$, and $C(S)$ denote the shipping cost, facility cost, and the total cost, respectively, obtained by assigning each location to its closest open facility. (Formally, $C_f(S) = \sum_{i \in S} f_i$ and $C(S) = C_s(S) + C_f(S)$.)

We now define *capacitated* variants of the above two problems in which there is a bound M on the total demand that can be shipped from any facility. The capacitated k-median (resp., facility location) problem is defined in the same manner as the uncapacitated k-median (resp., facility location) problem except that any solution must satisfy the additional constraint imposed by the capacity M. That is, a solution given by a set S of open facilities and an assignment σ is feasible only if the capacity constraints are respected at all the facilities in S. Two natural variants arise here: (i) *splittable demands*, where the demand of each location can be split across more than one facility, and (ii) *unsplittable demands*, where the demand of each location has to be shipped from a single facility. For the capacitated problems with splittable demands, given a set S of open facilities, an assignment is given by a function $\sigma : N \times S \mapsto \mathbf{R}$ with $\sigma(i,j)$ denoting the amount of demand shipped from facility j to location i. For the capacitated problems with unsplittable demands, an assignment is given by a function $\sigma : N \mapsto S$ as in the uncapacitated problems.

In the uncapacitated problems, given a set of open facilities, an optimal assignment is obtained by simply assigning each location to its closest open facility. In a capacitated problem, however, such an assignment may violate the capacity constraint. Fortunately, for the capacitated problems with splittable demands, given a set S of open facilities, an optimal assignment can be computed in polynomial time by solving an appropriately defined instance of the transportation problem [11]. Thus, any solution for the capacitated variants with splittable demands can be completely characterized by the set S of open facilites. For the capacitated problems with unsplittable demands, however, it is NP-hard to compute an optimal assignment associated with a given set S of open facilites. Therefore, any solution to the capacitated problems with unsplittable demands needs to specify a feasible assignment together with the set of open facilities.

The following notational conventions are used in the remainder of the paper. Given a set S of open facilites and an assignment σ, let $D_i(S, \sigma)$ denote the total demand shipped from a facility $i \in S$ under assignment σ. For the uncapacitated problems, given a set S of open facilities and an assignment σ, let $N_i(S, \sigma)$ denote the set of locations assigned to a facility $i \in S$. Where there is no ambiguity, the parameter σ will be omitted. Finally, we define the notion of a ball around a location. For each i in N, let $B(i, r)$ denote the set of locations j such that c_{ij} is at most r. Two balls are said to *overlap* if and only if they have a non-empty intersection.

3 The Local Search Paradigm

All of the algorithms considered in this paper are based on the following framework. We start with an arbitrary feasible solution and repeatedly refine the solution by performing a local search step. We show that within a polynomial number of local search steps, we arrive at a solution achieving the desired approximation factor. The problem of finding an initial feasible solution is discussed in Section 4. We now describe the local search step.

Let us consider the uncapacitated problems and the capacitated problems with splittable demands. Recall that any solution to these problems is completely characterized by the associated set S of open facilities. For the k-median problem, we define the *neighborhood* of S as $\{T \subseteq F : |S \setminus T| = |T \setminus S| = 1\}$. For the facility location problem, we define the neighborhood of S as $\{T \subseteq F : |S \setminus T| \leq 1 \text{ and } |T \setminus S| \leq 1\}$. Given a current solution corresponding to a set S of open facilities, the local search step sets the new solution to be a minimum-cost set T in the neighborhood of S. Since the neighborhood contains $O(n^2)$ solutions, a local search step can be performed in polynomial time.

The main motivation behind using the local search step is the following key property of the neighborhood of any solution, which is proved in Theorems 5.1, 6.1, and 7.1 for the uncapacitated k-median problem, the uncapacitated facility location problem, and the capacitated k-median problem with splittable demands, respectively. Let S^* be an optimal solution. Then, for a sufficiently large constant ε, given any solution S such that $C(S) > (1+\varepsilon)C(S^*)$ (for the k-median problems, we also require that $|S| = (1+\alpha)k$ for some sufficiently large constant α), there exists a solution T in the neighborhood of S such that $C(T) \leq C(S)(1 - 1/p(n))$, where $p(n)$ is a fixed polynomial in n. Thus, if we start with a solution S_0 and perform $p(n)\log(C(S_0)/C(S^*))$ local search steps, we will arrive at a solution that has cost at most $(1+\varepsilon)C(S^*)$. Since $\log(C(S_0))$ is polynomial in the input size, the number of local search steps involved is also polynomial in the input size. Moreover, each local search step takes polynomial time. We thus obtain a polynomial-time algorithm achieving an approximation factor of $(1+\varepsilon)$.

4 Finding an Initial Feasible Solution

In this section we show how to obtain an initial feasible solution for the uncapacitated problems and the capacitated problems with splittable demands. The task of selecting a valid set of open facilities is trivial. For the facility location problems, we select an arbitrary subset S of F. For the k-median problems, we require that $|S| = (1+\alpha)k$, for some sufficiently large constant α. In the extreme case where the total number of available facilities, $|F|$, is less than $(1+\alpha)k$, our algorithm trivially returns F as the solution. Hence throughout this paper, we assume without loss of generality that $|F|$ is at least $(1+\alpha)k$. Given this assumption, we obtain an initial feasible solution for the k-median problems by selecting an arbitrary set S of $(1+\alpha)k$ facilities.

As mentioned in Section 2, the set S of open facilities chosen above completely specifies the initial feasible solution. This is because we can obtain an optimal assignment for the initial solution by simply assigning each location to its closest open facility in the case of the uncapacitated problems, and by solving an appropriately defined instance of the transportation problem in the case of the capacitated problems with splittable demands.

5 Uncapacitated k-Median

Let S^* denote the optimal solution to the given instance of the k-median problem. Throughout this section, we assume that $p(n)$ is a polynomial in n and that α, β, γ, and μ are positive constants satisfying $\gamma < \alpha$, $\mu < 1$, and

$$\frac{\beta}{2} - \frac{k(1+\beta)}{2p(n)} - \frac{1}{\gamma(1-\mu)} \quad > \quad \frac{1+\beta}{\mu(\alpha-\gamma)}. \quad (1)$$

One possible choice, for example, is $\alpha = 9$, $\beta = 9$, $\mu = \frac{1}{2}$, $\gamma = 2$ and $p(n) = 10k$. The main result of this section is that, given a solution S of $(1+\alpha)k$ facilities with cost greater than $(1+\beta)C(S^*)$, we can perform a swapping of facilities to get another solution with the same size but significantly reduced cost.

Theorem 5.1 (Swapping of Facilities) *Let S be any subset of F such that $|S| = (1+\alpha)k$ and $C(S) > (1+\beta)C(S^*)$. Then there exist $u \in S$ and $v \in F$ such that $C(S) - C(S+v-u) \geq \frac{C(S)}{p(n)}$.*

We prove the theorem in two stages. First, in Lemma 5.1 we show that we can add a facility to S and get a significant reduction in the cost. Then, in Lemma 5.2, we show that we can drop a facility from S without increasing the cost too much. Combining these two lemmas then yields the theorem.

Lemma 5.1 (Adding a Facility) *Let S be any subset of F such that $C(S) > (1+\beta)C(S^*)$. Then there exists $v \in F$ such that $C(S) - C(S+v) \geq \frac{\beta C(S)}{(1+\beta)k}$.*

Proof: Let σ and σ^* be the optimal assignments for S and S^* respectively. We have

$$\begin{aligned} C(S) - C(S^*) &= \sum_{j \in N} d_j(c_{j\sigma(j)} - c_{j\sigma^*(j)}) \\ &= \sum_{i \in S^*} \sum_{j \in N_i(S^*)} d_j(c_{j\sigma(j)} - c_{ji}). \end{aligned}$$

By a simple averaging argument, it follows that there exists $v \in S^*$ for which

$$\sum_{j \in N_v(S^*)} d_j(c_{j\sigma(j)} - c_{jv}) \geq \frac{C(S) - C(S^*)}{k}.$$

Now consider the set $S+v$ along with the assignment σ' defined as follows: Assign $\sigma'(j)$ to v if $j \in N_v(S^*)$, and to $\sigma(j)$ otherwise. Using the fact that $C(S+v,\sigma') \geq C(S+v)$, we have

$$\begin{aligned} C(S) - C(S+v) &\geq C(S) - C(S+v,\sigma') \\ &= \sum_{j \in N_v(S^*)} d_j (c_{j\sigma(j)} - c_{jv}) \\ &\geq \frac{C(S) - C(S^*)}{k} \\ &\geq \frac{\beta C(S)}{(1+\beta)k}, \end{aligned}$$

as required. ■

Lemma 5.2 (Dropping a Facility) *Let S be any subset of F such that $|S| = (1+\alpha)k$ and $C(S) > (1+\beta)C(S^*)$. Then there exists $u \in S$ such that $C(S-u) - C(S) \leq \frac{\beta C(S)}{(1+\beta)k} - \frac{C(S)}{p(n)}$.*

Proof: Let $L = \frac{\beta C(S)}{(1+\beta)k} - \frac{C(S)}{p(n)}$. Let σ be an optimal assignment for S. For each facility i in S, we will be interested in the ball $B(i, r_i)$ where $r_i = \frac{L}{2D_i(S)}$. (Recall that $D_i(S)$ is the total demand shipped from facility i.) For the sake of brevity, let B_i denote the ball $B(i, r_i)$. We now consider two cases depending on whether any two of the B_i's overlap.

Suppose there exist facilities m and ℓ in S such that B_m and B_ℓ overlap. Without loss of generality, assume that $r_m \geq r_\ell$. Consider the set $S-m$ along with the following assignment σ': Assign $\sigma'(j)$ to ℓ if $\sigma(j) = m$, and to $\sigma(j)$ otherwise. Using the fact that $C(S-m,\sigma') \geq C(S-m)$, we have

$$\begin{aligned} C(S-m) - C(S) &\leq C(S-m,\sigma') - C(S) \\ &\leq \sum_{j \in N_m(S)} d_j (c_{\ell j} - c_{mj}) \\ &\leq \sum_{j \in N_m(S)} 2 d_j r_m \\ &= 2D_m(S) r_m \\ &= L. \end{aligned}$$

(The third equation follows from the triangle inequality and the fact that B_m and B_ℓ overlap, i.e., $c_{\ell j} - c_{mj} \leq c_{\ell m} \leq 2r_m$.) Thus, m satisfies the property stated in the lemma.

For the remaining case in which none of the B_i's overlap, we derive a contradiction by showing that $C(S) < (1+\beta)C(S^*)$. Let $\mu < 1$ be a positive constant. For a facility $i \in S$, let $Q_i(S)$ denote the total demand shipped from i to locations inside $B(i, \mu r_i)$. In other words,

$$Q_i(S) = \sum_{j \in N_i(S) \cap B(i,\mu r_i)} d_j.$$

Let $S' = \{i \in S : Q_i(S)(1-\mu)r_i \geq \frac{C(S)}{(1+\beta)\gamma k}\}$. We show in Lemma 5.3 that $|S'| \geq (1+\gamma)k$. Hence it follows that the set $S'' = \{i \in S' : B(i,r_i) \cap S^* = \emptyset\}$ has at least γk facilities (since $|S^*| = k$ and none of the B_i's overlap).

We now place a lower bound on $C(S^*)$ by considering the demands of only those locations that lie inside $B(i,\mu r_i)$ for $i \in S''$. Note that for each such location $j \in B(i, \mu r_i)$, the closest facility of S^* is at least $(1-\mu)r_i$ away.

$$\begin{aligned} C(S^*) &\geq \sum_{i \in S''} \sum_{j \in B(i,\mu r_i)} d_j (1-\mu) r_i \\ &\geq \sum_{i \in S''} Q_i(S)(1-\mu) r_i \\ &\geq \sum_{i \in S''} \frac{C(S)}{(1+\beta)\gamma k} \\ &= |S''| \frac{C(S)}{(1+\beta)\gamma k} \\ &\geq \frac{C(S)}{(1+\beta)}. \end{aligned}$$

(The third equation holds since $S'' \subseteq S'$. The fifth equation holds because $|S''| \geq \gamma k$.) Since our hypothesis was that $C(S) > (1+\beta)C(S^*)$, we have a contradiction. ■

Lemma 5.3 *Let $Q_i(S)$ and S' be as defined in Lemma 5.2. Then $|S'| \geq (1+\gamma)k$.*

Proof: Suppose $|S'| < (1+\gamma)k$. We obtain a contradictory lower bound on $C(S)$ by considering the shipping costs from the facilities in $S \setminus S'$ only. Note that for each facility i, a demand of $D_i(S) - Q_i(S)$ needs to be shipped to locations that are at least μr_i away. Recall that $D_i(S) r_i = L/2$ and for each $i \in S \setminus S'$, $Q_i(S)(1-\mu)r_i < \frac{C(S)}{(1+\beta)\gamma k}$. Therefore,

$$\begin{aligned} C(S) &\geq \sum_{i \in S \setminus S'} (D_i(S) - Q_i(S))\, \mu r_i \\ &\geq \sum_{i \in S \setminus S'} \mu \left(\frac{L}{2} - \frac{C(S)}{(1+\beta)(1-\mu)\gamma k} \right) \\ &\geq (\alpha - \gamma) k \mu \left(\frac{L}{2} - \frac{C(S)}{(1+\beta)(1-\mu)\gamma k} \right) \\ &\geq \frac{(\alpha-\gamma)\mu C(S)}{1+\beta} \left(\frac{\beta}{2} - \frac{k(1+\beta)}{2p(n)} - \frac{1}{\gamma(1-\mu)} \right) \\ &> C(S). \end{aligned}$$

(The third equation uses the fact that $|S \setminus S'| \geq (\alpha - \gamma)k$. The fourth equation is obtained by substituting the value of L. The fifth equation follows from Equation 1.) ■

Proof of Theorem 5.1: We first apply Lemma 5.2 to drop a facility u from S. We then apply Lemma 5.1 to add a facility v to $S-u$. We have

$$C(S) - C(S-u) \geq \frac{C(S)}{p(n)} - \frac{\beta C(S)}{(1+\beta)k}$$

and

$$C(S-u) - C(S-u+v) \geq \frac{\beta C(S-u)}{(1+\beta)k}.$$

Using the fact that $C(S-u) \geq C(S)$ and adding the above two inequalities, the theorem follows. ■

The corollary below is related to the question of how small the constants α and β can be. Note that any choice of α and β subject to the constraint of Equation 1 is valid. It is shown below that β can be made arbitrarily small, while on the other hand, α can be made arbitrarily close to 2.

Corollary 5.1.1 *For any constant $\varepsilon > 0$, there is a constant c such that the following can be computed in polynomial time: (i) a solution S with $C(S) \leq (1+\varepsilon)C(S^*)$ and $|S| \leq (1+c)k$, and (ii) a solution S with $|S| \leq (3+\varepsilon)k$ and $C(S) \leq (1+c)C(S^*)$.*

Proof: Let ε_0 be a sufficiently small positive constant and let c_0 be a sufficiently large positive constant. It can be verified that the following choices of the constants satisfy the constraint of Equation 1: (i) $\beta = \varepsilon$, $\mu = \frac{1}{2}$, $\gamma = \frac{8}{\varepsilon}$, $c = \alpha = 9 + \frac{17}{\varepsilon}$ and $p(n) = \frac{20k(1+\varepsilon)}{\varepsilon}$ and (ii) For $\varepsilon \leq \varepsilon_0$, $\alpha = 2+\varepsilon$, $\mu = 1 - \frac{\varepsilon}{4}$, $\gamma = \frac{\varepsilon}{4}$, $c = \beta = \frac{c_0}{\varepsilon^3}$ and $p(n) = (1+\beta)k$ and (iii) For $\varepsilon \geq \varepsilon_0$, $\alpha = 2+\varepsilon$, $\mu = 1 - \frac{\varepsilon_0}{4}$, $\gamma = \frac{\varepsilon_0}{4}$, $c = \beta = \frac{c_0}{\varepsilon_0^3}$ and $p(n) = (1+\beta)k$.

The first part of the claim follows by applying Theorem 5.1 with the first set of constants and using the local search algorithm described in Section 3. For the second part of the claim, if $\varepsilon \leq \varepsilon_0$ we use the second set of constants, and if $\varepsilon > \varepsilon_0$, we use the third set of constants. ■

6 Uncapacitated Facility Location

Let S^* denote the optimal solution for the given instance of the uncapacitated facility location problem. The main result of this section is Theorem 6.1, which shows that a local search step gives a significant improvement in the cost of the current solution if the cost of the current solution is sufficiently larger than the optimal cost. Throughout this section, we assume that $p(n)$ is a polynomial in n and that α, β, γ, δ, ε, and μ are arbitrary positive constants satisfying $\mu < 1$, $\delta < 1$, $\alpha\beta < 1$, $\gamma < 1$,

$$\begin{aligned}
\beta\mu(1-\delta)(1-\gamma) &\geq 2, && (2)\\
\frac{(1-\gamma)(1-\alpha\beta)(1+\varepsilon)}{1+\alpha} &> \frac{2}{(1-\mu)(1-\delta)}, && (3)\\
\frac{\varepsilon}{1+\varepsilon} - \frac{1}{1+\alpha} &\geq \frac{n}{p(n)}, \text{ and} && (4)\\
\frac{\gamma\delta(1-\alpha\beta)}{(1+\alpha)} &\geq \frac{n}{p(n)}. && (5)
\end{aligned}$$

One possible set of values for the constants is the following: $\alpha = 1/14$, $\beta = 7$, $\gamma = 1/8$, $\delta = 1/8$, $\varepsilon = 15$, $\mu = 1/2$, and $p(n) = 250n$.

Theorem 6.1 *Let S be any subset of F such that $C(S) > (1+\varepsilon)C(S^*)$. Then there exists $T \subseteq F$ such that $|S \setminus T| \leq 1$, $|T \setminus S| \leq 1$, and $C(S) - C(T) \geq C(S)/p(n)$.*

We consider two cases depending on whether the shipping cost of S, $C_s(S)$, is greater than α times the facility cost $C_f(S)$. For the former case, we show in Lemma 6.1 that we can add a facility to S and get a good improvement. For the latter case, we show in Lemma 6.3 that we can either drop a facility from S or substitute a facility in S, and obtain a good improvement.

Lemma 6.1 (Adding a Facility) *Let S be any subset of F. If $C(S) > (1+\varepsilon)C(S^*)$ and $C_s(S) > \alpha C_f(S)$, then there exists $v \in F$ such that $C(S) - C(S+v) \geq \frac{C(S)}{n}(\frac{\varepsilon}{1+\varepsilon} - \frac{1}{1+\alpha})$.*

Proof: We are given that $C(S) - C(S^*) > \varepsilon C(S)/(1+\varepsilon)$ and $C_f(S) < C(S)/(1+\alpha)$. Hence

$$\begin{aligned}
C_s(S) - C(S^*) &= C(S) - C(S^*) - C_f(S)\\
&\geq C(S)\left(\frac{\varepsilon}{1+\varepsilon} - \frac{1}{1+\alpha}\right).
\end{aligned}$$

If σ and σ^* denote the assignments associated with solutions S and S^*, respectively, then the quantity $C_s(S) - C(S^*)$ can be expanded as

$$\begin{aligned}
& C_s(S) - C_s(S^*) - C_f(S^*)\\
=\ & -C_f(S^*) + \sum_{j\in N} d_j(c_{j\sigma(j)} - c_{j\sigma^*(j)})\\
=\ & \sum_{i\in S^*} -f_i + \sum_{i\in S^*}\sum_{j\in N_i(S^*)} d_j(c_{j\sigma(j)} - c_{ji})\\
=\ & \sum_{i\in S^*}\left(-f_i + \sum_{j\in N_i(S^*)} d_j(c_{j\sigma(j)} - c_{ji})\right).
\end{aligned}$$

By a simple averaging argument, it follows that there exists a facility v in S^* such that

$$\begin{aligned} & -f_v + \sum_{j \in N_v(S^*)} d_j(c_{j\sigma(j)} - c_{jv}) \\ \geq \; & (C_s(S) - C(S^*))/n \\ \geq \; & \frac{C(S)}{n}\left(\frac{\varepsilon}{1+\varepsilon} - \frac{1}{1+\alpha}\right). \end{aligned} \tag{6}$$

Now consider the solution $T = S + v$ along with the assignment σ' defined as follows: If j is in $N_v(S^*)$, then set $\sigma'(j)$ to v; otherwise set $\sigma'(j)$ to $\sigma(j)$. Since $C(T, \sigma')$ is at least $C(T)$, Equation 6 implies that

$$\begin{aligned} C(S) - C(T) & \geq C(S) - C(T, \sigma') \\ & \geq \frac{C(S)}{n}\left(\frac{\varepsilon}{1+\varepsilon} - \frac{1}{1+\alpha}\right), \end{aligned}$$

thus establishing the desired claim. ■

We are now left with the case where $C_s(S)$ is at most $\alpha C_f(S)$. Before proceeding to handle this case in Lemma 6.3, we introduce some notation and technical machinery. For each facility i in S, we let g_i denote the shipping cost associated with i; i.e., $g_i = \sum_{j \in N_i(S)} d_j c_{ij}$. For any set A of facilities, we let g_A denote $\sum_{i \in A} g_i$. Note that $C_s(S) = g_S$.

Let S be a solution such that $C_s(S) \leq \alpha C_f(S)$. Our proof can be informally described as follows. We first identify a "useful" subset U of S by throwing away those facilities i which either have $f_i < \beta g_i$ or have a "small" f_i, and ensure that $C_f(U)$ is sufficiently large. Then in Lemma 6.3, we restrict our attention to the facilities in U.

The set U is identified in two stages as follows. Let P be the set of facilities i in S such that $f_i \geq \beta g_i$. We claim that the facility cost associated with P is at least a constant fraction of the facility cost associated with S. Since $C_f(S) \geq \frac{C_s(S)}{\alpha}$, we have

$$\begin{aligned} C_f(S) & \geq \frac{1}{\alpha}\left(\sum_{i \in P} g_i + \sum_{i \in S \setminus P} g_i\right) \\ & \geq \frac{1}{\alpha}\sum_{i \in P} g_i + \frac{\sum_{i \in S \setminus P} f_i}{\alpha\beta} \\ & \geq \frac{g_P}{\alpha} + \frac{C_f(S \setminus P)}{\alpha\beta}. \end{aligned} \tag{7}$$

(For the second equation, we note that for each facility i in $S \setminus P$, we have $f_i < \beta g_i$.)

Since $C_f(S) = C_f(P) + C_f(S \setminus P)$, we have

$$\begin{aligned} C_f(P) & = C_f(S) - C_f(S \setminus P) \\ & \geq C_f(S) - \alpha\beta C_f(S) + \beta g_P \\ & = \beta g_P + (1 - \alpha\beta)C_f(S). \end{aligned}$$

(The second equation follows from Equation 7.) Since $C_f(S) \geq C_s(S)/\alpha$, we also have $C_f(S) \geq C(S)/(1+\alpha)$. Hence, we obtain $C_f(P) \geq \beta g_P + (1-\alpha\beta)C(S)/(1+\alpha)$.

We now consider the set of facilities in P whose facility cost is at least a γ fraction of the average facility cost in P. Let U denote the set of facilities i in P such that f_i is at least $\gamma C_f(P)/|P|$. It is easy to derive that $C_f(U)$ is at least $(1-\gamma)C_f(P)$. Thus, U satisfies

$$\frac{C_f(U)}{(1-\gamma)} \geq \left(\beta g_P + \frac{(1-\alpha\beta)C(S)}{(1+\alpha)}\right) \tag{8}$$

and for all i in U,

$$f_i \geq \beta g_i, \text{ and} \tag{9}$$

$$f_i \geq \frac{\gamma}{|P|}\left(\beta g_P + \frac{(1-\alpha\beta)C(S)}{(1+\alpha)}\right). \tag{10}$$

For each facility i in U, we will be concerned with the ball $B(i, r_i)$ where

$$r_i = \frac{(1-\delta)f_i}{2(\sum_{j \in N_i(S)} d_j)}.$$

Lemma 6.2 *Let $\mu < 1$ be a positive constant and let i be any facility in U. If A denotes the set of locations in $N_i(S)$ that lie in $B(i, \mu r_i)$, then*

$$\sum_{j \in A} d_j r_i \geq \left(\frac{(1-\delta)f_i}{2} - \frac{g_i}{\mu}\right).$$

Proof: Fix any i in U. By the definition of A, locations in $N_i(S) \setminus A$ are assigned to i but lie outside $B(i, \mu r_i)$. Therefore, we have:

$$\begin{aligned} & \sum_{j \in A} d_j r_i \\ = \; & \left(\sum_{j \in N_i(S)} d_j r_i - \sum_{j \in N_i(S) \setminus A} d_j r_i\right) \\ \geq \; & \left(\frac{(1-\delta)f_i}{2} - \frac{1}{\mu}\sum_{j \in N_i(S) \setminus A} d_j c_{ij}\right) \\ \geq \; & \left(\frac{(1-\delta)f_i}{2} - \frac{g_i}{\mu}\right). \end{aligned}$$

(In the first equation, we replace the set A by the difference of two disjoint sets $N_i(S)$ and $N_i(S) \setminus A$. For the second equation, we use the definition of r_i and note that $c_{ij} \geq \mu r_i$ for every j in $N_i(S) \setminus A$. The last equation is derived by observing that $\sum_{j \in N_i(S) \setminus A} d_j c_{ij} \leq \sum_{j \in N_i(S)} d_j c_{ij} \leq g_i$.) ■

Lemma 6.3 (Dropping/Swapping of Facilities) *Let S be any subset of F such that $C(S) > (1+\varepsilon)C(S^*)$ and $C_s(S) \leq \alpha C_f(S)$. Then there exists a $T \subseteq F$ such that $|S \setminus T| \leq 1$, $|T \setminus S| \leq 1$ and $C(S) - C(T) \geq \frac{C(S)\gamma\delta(1-\alpha\beta)}{(1+\alpha)n}$.*

Proof: Consider the subset U of S that satisfies Equations 9 and 10. Consider the set of balls $B_i = B(i, r_i)$ for all i in U. We consider two cases depending on whether any pair of the B_i's overlap.

Let us assume first that there exist two facilities k and ℓ such that B_k and B_ℓ overlap. Without loss of generality, assume that $r_k \geq r_\ell$. Let T equal the solution $S - k$. We now show that T satisfies the stated property. The facility cost of T, $C_f(T)$, equals $C_f(S) - f_k$. We next consider the shipping cost of T. Let σ denote the optimal assignment associated with the solution S. Let σ' be the following assignment of locations to facilities in T: For each location j, if $\sigma(j) \neq k$, then $\sigma'(j)$ equals $\sigma(j)$; otherwise $\sigma'(j) = \ell$. Since $C(T)$ is at most $C(T, \sigma')$, we can place a lower bound on $C(S) - C(T)$ as follows:

$$\begin{aligned}
C(S) - C(T) &\geq f_k - \sum_{j \in N_k(S)} d_j(c_{\ell j} - c_{kj}) \\
&\geq f_k - \sum_{j \in N_k(S)} 2 d_j r_k \\
&\geq f_k - (1-\delta) f_k \\
&= \delta f_k \\
&\geq \frac{\delta\gamma(1-\alpha\beta)C(S)}{(1+\alpha)|P|} \\
&\geq \frac{\delta\gamma(1-\alpha\beta)C(S)}{(1+\alpha)n}.
\end{aligned}$$

(The second equation holds since the balls B_k and B_ℓ overlap and $r_k \geq r_\ell$. The third equation follows from the definition of r_k. The fifth equation follows from Equation 10. The last equation holds since $|P| \leq n$.) This establishes the desired claim in the case that two of the B_i's overlap.

For the remainder of the proof, we assume that none of the B_i's overlap. We again consider two cases. For the first case, we assume that there exists a facility i in U such that B_i contains a facility i^* (from $F \setminus S$) with $f_{i^*} \leq (1-\delta)f_i/2$. In this case, we set T to be $S - i + i^*$. For each location j in $N_i(S)$, the shipping cost associated with j in solution T is at most c_{ji^*}. We now place a lower bound on $C(S) - C(T)$ as follows.

$$\begin{aligned}
C(S) - C(T) &\geq f_i - f_{i^*} + \sum_{j \in N_i(S)} d_j(c_{ji} - c_{ji^*}) \\
&\geq f_i - (1-\delta)f_i/2 - \sum_{j \in N_i(S)} d_j r_i \\
&\geq f_i - (1-\delta)f_i/2 - (1-\delta)f_i/2 \\
&= \delta f_i \\
&\geq \frac{\delta\gamma(1-\alpha\beta)C(S)}{(1+\alpha)|P|} \\
&\geq \frac{\delta\gamma(1-\alpha\beta)C(S)}{(1+\alpha)n}.
\end{aligned}$$

(The second equation follows from the upper bound on f_{i^*} and the fact that i^* is in B_i. The third equation follows from the definition of r_i. The fifth equation follows from Equation 10. The last equation holds since $|P| \leq n$.) Thus, T satisfies the property stated in the lemma.

We are left with the case in which none of the B_i's overlap and for each i in U, either $B_i \cap S^*$ is $\emptyset$ or for all i^* in $B_i \cap S^*$, $f_{i^*} > (1-\delta)f_i/2$. In this case, we derive a contradiction by showing that $C(S) < (1+\varepsilon)C(S^*)$.

We obtain a lower bound on $C(S^*)$ as follows. Let U_1 denote the set of facilities i in U for which $B_i \cap S^*$ is empty. Let U_2 denote $U \setminus U_1$. For each i in U_1, since $B_i \cap S^*$ is empty, solution S^* assigns locations in B_i to facilities outside B_i. In particular, locations in $B(i, \mu r_i) \cap N_i(S)$ (which is a subset of B_i) are assigned by S^* to facilities outside B_i. Thus, for each location j in $B(i, \mu r_i) \cap N_i(S)$, the associated shipping cost is at least $d_j(1-\mu)r_i$. By Lemma 6.2, we obtain that the total shipping cost associated with locations in $B(i, \mu r_i) \cap N_i(S)$ is at least $(1-\mu)((1-\delta)f_i/2 - g_i/\mu)$. Summing over all facilities in U_1, we obtain the following lower bound on the shipping cost of S^*:

$$\frac{C_s(S^*)}{1-\mu} \geq \sum_{i \in U_1} \left(\frac{(1-\delta)f_i}{2} - \frac{g_i}{\mu} \right). \tag{11}$$

We now turn to the facilities in U_2. Since the balls B_i associated with the facilities in U_2 do not overlap, for each i in U_2 there exists a unique i^* in S such that $f_{i^*} > (1-\delta)f_i/2$. Summing over all facilities in U_2, we obtain the following lower bound on the facility cost of S^*:

$$C_f(S^*) \geq (1-\delta) \sum_{i \in U_2} f_i/2. \tag{12}$$

Using Equations 11 and 12, we obtain

$$\begin{aligned}
C(S^*) &\geq (1-\mu) \sum_{i \in U_1} \left(\frac{(1-\delta)f_i}{2} - \frac{g_i}{\mu} \right) + \\
&\quad (1-\delta) \sum_{i \in U_2} f_i/2 \\
&\geq (1-\mu) \sum_{i \in U} \left(\frac{(1-\delta)f_i}{2} - \frac{g_i}{\mu} \right) \\
&= (1-\mu) \left(\frac{(1-\delta)C_f(U)}{2} - \frac{g_U}{\mu} \right)
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{(1-\mu)(1-\delta)(1-\gamma)(1-\alpha\beta)C(S)}{2(1+\alpha)} + \\
&\quad (1-\mu)\left(\frac{(1-\delta)(1-\gamma)\beta g_P}{2} - \frac{g_U}{\mu}\right) \\
&\geq \frac{(1-\mu)(1-\delta)(1-\gamma)(1-\alpha\beta)C(S)}{2(1+\alpha)} \\
&> C(S)/(1+\varepsilon).
\end{aligned}
$$

(The fourth equation follows from Equation 8; the fifth equation follows from $g_P \geq g_U$ and Equation 2; the last equation follows from Equation 3.) We have thus derived a contradiction in the final case, completing the proof of the lemma. ∎

Proof of Theorem 6.1: We consider two cases: $C_s(S) > \alpha C_s(S)$ and $C_s(S) \leq \alpha C_s(S)$. In the first case, the result follows from Lemma 6.1 and Equation 4. In the second case, the result follows from Lemma 6.3 and Equation 5. ∎

Corollary 6.1.1 *Given any constant ε', we can compute a solution S in polynomial time such that $C(S)$ is at most $(9+\varepsilon')C(S^*)$.*

Proof: Given any ε', the following values of the parameters α, β, γ, δ, ε, μ, and $p(n)$ satisfy the desired inequalities: $\alpha = 1/8+\varepsilon'/(256n)$, $\beta = 4+\varepsilon'/(8n)$, $\gamma = \varepsilon'/(256n)$, $\delta = \varepsilon'/(256n)$, $\varepsilon = 8+\varepsilon'$ $\mu = 1/2+\varepsilon'/(8n)$, and $p(n) = n/\min\{\varepsilon/(1+\varepsilon) - 1/(1+\alpha), \gamma\delta(1-\alpha\beta)/(1+\alpha)\}$. The claim now follows from Theorem 6.1. ∎

7 Capacitated k-Median Problem with Splittable Demands

In this section, we consider the capacitated k-median problem with splittable demands and show that, as in the uncapacitated case, each step of the local search heuristic yields a significant improvement in cost whenever the current cost is sufficiently higher than the optimal cost.

Recall that given a set S of open facilities, the corresponding optimal (splittable) assignment of demands to facilities can be computed in polynomial time by a reduction to the transportation problem. Hence, the solution to the capacitated problem with splittable demands is entirely characterized by the set S of open facilities.

Let S^* denote an optimal solution for the given instance of the capacitated k-median problem with splittable demands. Throughout this section, we assume that $p(n)$ is a polynomial in n and that α, β, γ, and μ are positive constants satisfying $\gamma < \alpha$, $\mu < 1$ and

$$\frac{\beta}{2} - \frac{k(1+\beta)}{2p(n)} - \frac{1}{(\gamma-2)(1-\mu)} > \frac{1+\beta}{\mu(\alpha-\gamma)}. \quad (13)$$

The latter constraint is a slight modification of the constraint of Equation 1. One possible choice for the parameters is $\alpha = 11$, $\beta = 9$, $\gamma = 4$, $\mu = \frac{1}{2}$, and $p(n) = 10k$.

Theorem 7.1 (Swapping Facilities) *Let S be a feasible solution such that $|S| = (1+\alpha)k$ and $C(S) > (1+\beta)C(S^*)$. Then there exist $u \in S$ and $v \in F$ such that $C(S) - C(S+v-u) \geq \frac{C(S)}{p(n)}$.*

As before, we proceed in two stages. Lemma 7.1 shows that we can add a facility to S to get a significant improvement in the cost, and Lemma 7.2 shows that we can drop a facility from S without incurring a large increase in the cost. The proof of the following lemma related to adding a facility is similar to that of Lemma 5.1. The details will appear in the full version of the paper.

Lemma 7.1 (Adding a Facility) *Let S be a feasible solution such that $C(S) > (1+\beta)C(S^*)$. Then there exists $v \in F$ such that $C(S) - C(S+v) \geq \frac{\beta C(S)}{(1+\beta)k}$.* ∎

Lemma 7.2 (Dropping a Facility) *Let S be a feasible solution such that $|S| = (1+\alpha)k$ and $C(S) > (1+\beta)C(S^*)$. Then there exists a $u \in S$ such that $C(S-u) - C(S) \leq \frac{\beta C(S)}{(1+\beta)k} - \frac{C(S)}{p(n)}$.*

Proof: A facility $i \in S$ is *light* (under the corresponding optimal assignment σ) if the total demand shipped from i is at most $M/2$. Otherwise, it is *heavy*. As $\sum_{j\in N} d_j \leq kM$, there can be at most $2k$ heavy facilities.

Let L denote $\frac{\beta C(S)}{(1+\beta)k} - \frac{C(S)}{p(n)}$. As in Lemma 5.2, we define B_i to be the ball $B(i, r_i)$, where r_i equals $L/2D_i(S)$. We call B_i *light* if i is a light facility. We first consider the case where there exist two light balls B_m and B_ℓ that overlap. Without loss of generality, assume that $r_m \geq r_\ell$. Consider the set $S - m$ along with the assignment σ', which is same as σ, except that the demand that was previously assigned to m is now assigned to ℓ. Note that no capacity constraints are violated in σ', since m and ℓ are light facilities, We thus have,

$$
\begin{aligned}
C(S-m) - C(S) &\leq C(S-m, \sigma') - C(S) \\
&\leq D_m(S)c_{\ell m} \\
&\leq 2D_m(S)r_m = L,
\end{aligned}
$$

yielding the desired claim.

For the remaining case in which no two light balls overlap, we derive a contradiction as in Lemma 5.2. Let $\mu < 1$ be an arbitrary real number. For a facility

$i \in S$, as before, we let $Q_i(S)$ denote the total demand shipped from i to locations inside $B(i, \mu r_i)$. In other words,

$$Q_i(S) = \sum_{j \in B(i,\mu r_i)} \sigma(j,i).$$

Let $S' = \{i \in S : Q_i(S)(1-\mu)r_i \geq \frac{C(S)}{(1+\beta)(\gamma-2)k}\}$. Using Equation 13 and an argument very similar to that of Lemma 5.3, it follows that $|S'| \geq (1+\gamma)k$. Therefore the number of light facilities in S' is at least $(1+\gamma-2)k$. Hence it follows that the set $S'' = \{i \in S' : i \text{ is light and } B(i, r_i) \cap S^* = \emptyset\}$ has at least $(\gamma-2)k$ facilities (since $|S^*| = k$ and none of the light B_i's overlap). We now place a lower bound on $C(S^*)$ by considering the demands of only those locations that lie inside $B(i, \mu r_i)$ for $i \in S''$. Note that for each such location $j \in B(i, \mu r_i)$, the closest facility of S^* is at least $(1-\mu)r_i$ away.

$$\begin{aligned} C(S^*) &\geq \sum_{i \in S''} \sum_{j \in B(i,\mu r_i)} d_j (1-\mu) r_i \\ &\geq \sum_{i \in S''} Q_i(S)(1-\mu) r_i \\ &\geq |S''| \frac{C(S)}{(1+\beta)(\gamma-2)k} \\ &\geq \frac{C(S)}{(1+\beta)} \end{aligned}$$

(The third equation follows from the definition of $Q_i(S)$ and the fact that S'' is a subset of S'. The fourth equation holds because $|S''| \geq (\gamma-2)k$.) Since our hypothesis was that $C(S) > (1+\beta)C(S^*)$, we have a contradiction. ∎

The proof of Theorem 7.1 follows from Lemmas 7.2 and 7.1, using essentially the same arguments as in the proof of Theorem 5.1.

Corollary 7.1.1 *For any constant $\varepsilon > 0$, there is a constant c such that the following can be computed in polynomial time: (i) a solution S with $C(S) \leq (1+\varepsilon)C(S^*)$ and $|S| \leq (1+c)k$, and (ii) a solution S with $|S| \leq (5+\varepsilon)k$ and $C(S) \leq (1+c)C(S^*)$.*

Proof: It can be verified that the following sets of constants satisfy the constraint of Equation 13.

1. Take $c = \alpha = 11 + 17/\varepsilon$, $\beta = \varepsilon$, $\gamma = 2 + 8/\varepsilon$, $\mu = 1/2$, and $p(n) = \frac{20k(1+\varepsilon)}{\varepsilon}$.

2. For $\varepsilon \leq \varepsilon_0$, where ε_0 is a sufficiently small constant, we set $\alpha = 4 + \varepsilon$, $c = \beta = c_0/\varepsilon^3$, $\gamma = 2 + \varepsilon/4$, $\mu = 1 - \varepsilon/4$, and $p(n) = k(1+\beta)$. (Here c_0 is a constant that is chosen sufficiently large.)

   For $\varepsilon > \varepsilon_0$, we set $\alpha = 4 + \varepsilon$, $c = \beta = c_0/\varepsilon_0^3$, $\gamma = 2 + \varepsilon_0/4$, $\mu = 1 - \varepsilon_0/4$, and $p(n) = k(1+\beta)$.

The results follow by using these constants in Theorem 7.1 and applying the algorithm outlined in Section 3. ∎

References

[1] G. Cornuéjols, G. L. Nemhauser, and L. A. Wolsey. The uncapacitated facility location problem. In P. Mirchandani and R. Francis, editors, *Discrete Location Theory*, pages 119–171. John Wiley and Sons, Inc., New York, 1990.

[2] G. Diehr. An algorithm for the p-median problem. Technical Report No. 191, Western Management Science Institute, UCLA, June 1972.

[3] D. S. Hochbaum. Heuristics for the fixed cost median problem. *Mathematical Programming*, 22:148–162, 1982.

[4] D. S. Hochbaum. Various notions of approximations: Good, better, best, and more. In D. S. Hochbaum, editor, *Approximation Algorithms for NP-hard Problems*, pages 346–398. PWS Publishing Company, Boston, MA, 1995.

[5] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[6] L. G. Khachiyan. A polynomial algorithm for linear programming. *Soviet Math. Doklady*, 20:191–194, 1979.

[7] A. A. Kuehn and M. J. Hamburger. A heuristic program for locating warehouses. *Management Science*, 9:643–666, 1963.

[8] J.-H. Lin and J. S. Vitter. Approximation algorithms for geometric median problems. *Information Processing Letters*, 44:245–249, 1992.

[9] J.-H. Lin and J. S. Vitter. ε-approximations with minimum packing constraint violation. In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 771–782, May 1992.

[10] D. B. Shmoys and É. Tardos. An approximation algorithm for the generalized assignment problem. *Mathematical Programming*, 62:461–474, 1993.

[11] D. B. Shmoys, É. Tardos, and K. Aardal. Approximation algorithms for facility location problems. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, pages 265–274, May 1997.

[12] M. B. Teitz and P. Bart. Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations Research*, 16:955–961, 1968.