Discrete Optimization

Approximation algorithms for hard capacitated k-facility location problems

CrossMark

Karen Aardal [a,b], Pieter L. van den Berg [a], Dion Gijswijt [a], Shanfei Li [a,c,*]

[a] *Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*
[b] *Centrum Wiskunde en Informatica, Science Park 123, 1098 XG Amsterdam, The Netherlands*
[c] *Faculty of Information System and Management, National University of Defense Technology, Yanwachi 137, 410073 Changsha, China*

ARTICLE INFO

ABSTRACT

We study the capacitated k-facility location problem, in which we are given a set of clients with demands, a set of facilities with capacities and a positive integer k. It costs f_i to open facility i, and c_{ij} for facility i to serve one unit of demand from client j. The objective is to open at most k facilities serving all the demands and satisfying the capacity constraints while minimizing the sum of service and opening costs.

In this paper, we give the first fully polynomial time approximation scheme (FPTAS) for the single-sink (single-client) capacitated k-facility location problem. Then, we show that the capacitated k-facility location problem with uniform capacities is solvable in polynomial time if the number of clients is fixed by reducing it to a collection of transportation problems. Third, we analyze the structure of extreme point solutions, and examine the efficiency of this structure in designing approximation algorithms for capacitated k-facility location problems. Finally, we extend our results to obtain an improved approximation algorithm for the capacitated facility location problem with uniform opening costs.

1. Introduction

In the capacitated k-facility location problem (CKFL), we are given a set D of clients and a set F of potential facilities (locations where we can potentially open a facility) in a metric space. Each facility $i \in F$ has a capacity s_i. Establishing facility i incurs an opening cost f_i. Each client j has a demand d_j that must be served. Shipping x_{ij} units from facility i to client j incurs service costs $c_{ij}x_{ij}$, where c_{ij} is proportional to the distance between i and j. The goal is to serve all the clients by using at most k facilities and satisfying the capacity constraints such that the total cost is minimized. In this paper, we consider the *hard* capacities, that is, we allow at most one facility to be opened at any location. (Note that in the *soft* capacities case multiple facilities can be opened in a single location, see for instance Mahdian, Ye, & Zhang, 2003; Shmoys, Tardos, & Aardal, 1997.)

CKFL can be formulated as the following mixed integer program (MIP), where variable x_{ij} indicates the amount of the demand of client j that is served by facility i, and y_i indicates whether facility i is open.

$$\min \quad \sum_{i \in F} \sum_{j \in D} c_{ij} x_{ij} + \sum_{i \in F} f_i y_i \tag{1}$$

subject to:
$$\sum_{i \in F} x_{ij} = d_j, \quad \forall j \in D, \tag{2}$$

$$\sum_{j \in D} x_{ij} \leq s_i y_i, \quad \forall i \in F, \tag{3}$$

$$\sum_{i \in F} y_i \leq k, \tag{4}$$

$$x_{ij} \geq 0, \quad \forall i \in F, \forall j \in D, \tag{5}$$

$$y_i \in \{0, 1\}, \quad \forall i \in F. \tag{6}$$

If we replace constraints (6) by
$$0 \leq y_i \leq 1, \forall i \in F, \tag{7}$$

we obtain the LP-relaxation of CKFL. Without loss of generality we assume that s_i, d_j are integral for all $i \in F, j \in D$.

CKFL is related to the capacitated k-median problem (CKM), which corresponds to the special case where $f_i = 0, \forall i \in F$. A constant factor approximation algorithm is still unknown for CKM, let alone CKFL. All the previous attempts with constant approximation ratios for these problems violate the capacity constraint, or cardinality constraint that at most k facilities are allowed to be used. We call these approximation algorithms *pseudo-approximation algorithms* when one of the constraints is violated up to a certain factor. Recently, Byrka, Fleszar, Rybicki, and Spoerhase (2013) gave a constant factor approximation algorithm for CKM with uniform capacities while violating the capacities by a factor $2 + \epsilon$, where $\epsilon > 0$ can be arbitrarily

* Corresponding author. Tel.: +31 15 27 84989.
E-mail addresses: K.I.Aardal@tudelft.nl (K. Aardal), P.L.vandenBerg@tudelft.nl
(P. L. van den Berg), D.C.Gijswijt@tudelft.nl (D. Gijswijt), shanfei.li@tudelft.nl,
lishanfei@nudt.edu.cn (S. Li).

Table 1
Approximation ratios for the capacitated facility location problem and its variants.

Problem	Ratio	Reference	Technique
CFL with hard uniform capacities	8	Korupolu, Plaxton, and Rajaraman (2000)	Local search
	5.83	Chudak and Williamson (2005)	Local search
	3	Aggarwal et al. (2013)	Local search
CFL with uniform opening costs	5	Levi, Shmoys, and Swamy (2012)	LP-rounding
	$4.562 + \varepsilon$	This paper	Combinatorial algorithm
CFL	8.53	Pál, Tardos, and Wexler (2001)	Local search
	7.88	Mahdian and Pál (2003)	Local search
	$3 + 2\sqrt{2} + \varepsilon$	Zhang, Chen, and Ye (2005)	Local search
	5	Bansal, Garg, and Gupta (2012)	Local search

small. Although most researchers believe that relaxing the cardinality constraint makes the problem simpler than relaxing the capacity constraint with respect to designing pseudo-approximation algorithms, the best known violation ratio for the cardinality constraint is still $5 + \epsilon$ (Korupolu et al., 2000) for CKM with uniform capacities. It seems that obtaining a better constant factor approximation algorithm by only violating the cardinality constraint has not received much attention yet.

In this paper, we give an improved pseudo-approximation algorithm for CKFL with uniform opening cost by using at most $2k$ facilities. To show the potential power of this algorithm, we improve the approximation ratio for the capacitated facility location problem with uniform opening costs (Levi et al., 2012), by combining this algorithm with a pseudo-approximation algorithm for the k-median problem derived from a bifactor approximation algorithm for the uncapacitated facility location problem (Charikar & Guha, 1999). That is, pseudo-approximation algorithms for capacitated k-facility location problems may be extended to get approximation algorithms for well-studied capacitated facility location problems. We believe that this technique has the potential to further improve approximation ratios for capacitated facility location problems.

In Section 2 we give the first fully polynomial time approximation scheme (FPTAS) for the single-sink capacitated k-facility location problem. In Section 3, we present a polynomial time algorithm for the uniform capacitated k-facility location problem with a fixed number of clients. In Section 4, we propose an improved pseudo-approximation algorithm for CKFL with uniform opening costs, and extend this algorithm to improve the approximation ratio for the capacitated facility location problem with uniform opening costs.

1.1. Related work

The k-facility location problem has already been studied since the early 1990s (Cornuéjols, Nemhauser, & Wolsey, 1990; Hsu, Lowe, & Tamir, 1997). It is a common generalization of the k-median problem (in which at most k facilities are allowed to be opened, and there is no opening costs) and the uncapacitated facility location problem, which are classical problems in computer science and operations research and have a wide variety of applications in clustering, data mining, logistics (Bradley, Fayyad, & Mangasarian, 1999; Jain & Dubes, 1988; Kuehn & Hamburger, 1963). Even the single-sink (single client) case is interesting. For instance, the product distribution problem, in which a company needs to deliver a given number units of a product by trucks every day within a short time, e.g., early in the morning, which implies that each truck can be used only once everyday. Each truck is characterized by three parameters: the unit transportation cost, the expense per usage, and the capacity. The objective is to select a set of trucks to satisfy these deliveries with minimum cost. This problem can be modeled as the single-sink capacitated facility location problem. See Herer, Rosenblatt, and Hefter (1996) for more applications of single-sink case, like the supplier selection problem. When (given) multiple products need to be delivered, the above problem can be modeled as the capacitated facility location problem with fixed

$m = |D|$. If we have a further constraint on the number of trucks that can be used, then it is a capacitated k-facility location problem.

For the uncapacitated k-facility location problem (UKFL), which is a special case of CKFL with $s_i = +\infty, \forall i \in F$, Charikar, Guha, Tardos, and Shmoys (1999) gave the first constant factor approximation algorithm with performance guarantee 9.8, by modifying their $6\frac{2}{3}$-approximation algorithm for the uncapacitated k-median problem. Later, the approximation ratio was improved by Jain and Vazirani (2001), who made use of a primal-dual scheme and Lagrangian relaxation techniques to obtain a 6-approximation algorithm. Jain, Mahdian, Markakis, Saberi, and Vazirani (2003) and Jain, Mahdian, and Saberi (2002) further improved the ratio to 4 by using a greedy approach and the so-called Lagrangian Multiplier Preserving property of the algorithms. The best known approximation algorithm for this problem, due to Zhang (2007), achieves a factor of $2 + \sqrt{3} + \epsilon$ using a local search technique for any $\epsilon > 0$. The k-median problem, as a special case of UKFL, was studied extensively (Archer, Rajagopalan, & Shmoys, 2003; Arya et al., 2001; Byrka, Pensyl, Rybicki, Srinivasan, & Trinh, 2014; Charikar & Guha, 1999; Charikar et al., 1999; Jain et al., 2002; Jain & Vazirani, 2001; Li & Svensson, 2013) and the best known approximation algorithm was recently given by Byrka et al. (2014) with approximation ratio $2.611 + \epsilon$ by improving the algorithm of Li and Svensson (2013). In addition, Edwards (2001) gave a 7.814-approximation algorithm for the multi-level uncapacitated k-facility location problem by extending the $6\frac{2}{3}$-approximation algorithm by Charikar et al. (1999) for the uncapacitated k-median problem.

Unfortunately, the capacitated k-facility location problem is much less understood although the presence of capacity constraints is natural in practice. The difficulty of the problem lies in the fact that two kinds of hard constraints appear together: the cardinality constraint, and the capacity constraints. This seems to result in hardness of the methods such as LP-rounding, primal-dual method used to solve the k-median problem, and even local search algorithms used to solve the capacitated facility location problem and the k-median problem.

The capacitated k-facility location problem is related to the capacitated facility location problem (CFL), which corresponds to a special case of CKFL with $k = n$. Most known approximation algorithms for CFL (see Table 1) are based on local search technique since the natural linear programming relaxation has an unbounded integrality gap for the general case (Pál et al., 2001). For uniform capacities, Korupolu et al. (2000) gave the first constant factor approximation algorithm, with a factor of 8. Later, this factor was improved to 5.83 by Chudak and Williamson (2005). The currently best approximation algorithm due to Aggarwal et al. (2013) has performance guarantee of 3. For nonuniform capacities, Pál et al. (2001) proposed the first constant factor approximation algorithm with a factor of 8.53. Later, Mahdian and Pál (2003) improved this factor to 7.88. Zhang et al. (2005) reduced this factor to $(3 + 2\sqrt{2} + \varepsilon)$ by introducing a multi-exchange operation, where $\varepsilon > 0$. The currently best known approximation algorithm, due to Bansal et al. (2012), achieves the approximation ratio 5.

Additionally, Levi et al. (2012) showed that the linear programming relaxation has a bounded integrality gap for CFL with uniform

Table 2
Approximation ratios for the capacitated k-facility location problem and its variants. "♯ open facilities" denotes the number of open facilities.

Problem	Ratio	Capacity blowup	♯ open facilities blowup	Reference	Remark
SCKFL	$1 + \epsilon$	None	None	This paper	$\epsilon > 0$
CKM with hard uniform capacities	$1 + \epsilon$	None	$5 + \frac{5}{\epsilon}$	Korupolu et al. (2000)	$\epsilon > 0$
	$1 + \frac{5}{\epsilon}$	None	$5 + \epsilon$	Korupolu et al. (2000)	$\epsilon > 0$
	16	3	None	Charikar et al. (1999)	
	$10\alpha + 6$	$2 + \frac{2}{(\alpha - 2)}$	None	Li (2014)	$\alpha \geq 4$
CKM with soft capacities	40	50	None	Chuzhoy and Rabani (2005)	
	$\frac{19.3(1+\epsilon)}{\epsilon^2}$	None	$1 + \epsilon$	Bartal, Charikar, and Raz (2001)	$\epsilon > 0$
CKM	$96 + \frac{180}{\epsilon}$	$3 + 3\epsilon$	None	Byrka et al. (2013)	$\epsilon > 0$
CKFL with hard uniform capacities	$32l^2 + 28l + 7$	$2 + \frac{3}{(l-1)}$	None	Byrka et al. (2013)	$l \in \mathbb{Z}_{\geq 2}$
CKFL with uniform opening costs	$7 + \epsilon$	None	2	This paper	$\epsilon > 0$

opening costs, and gave a 5-approximation algorithm for this case by an LP-rounding technique.

The capacitated k-median problem (CKM), which is a special case of CKFL, is already difficult to handle. The natural linear programming relaxation has an unbounded integrality gap (see Remark 1). We have to blow up the capacity or increase the number of opened facilities by a factor of at least 2 if we use the optimal cost of the natural LP relaxation as a lower bound to obtain an integral solution (Charikar et al., 1999). (See Table 2.)

For CKM with hard uniform capacities, Charikar et al. (1999) gave a constant factor approximation algorithm while violating the capacities within a constant factor 3 by LP-rounding. Recently, Byrka et al. (2013) improved this violation ratio to $2 + \epsilon$ by designing a $(32l^2 + 28l + 7)$-approximation algorithm increasing the capacity by a factor of $2 + \frac{3}{l-1}$, $l \in \{2, 3, 4, \ldots\}$. Li (2014) improved the approximation ratio $32l^2 + 28l + 7$ to $10\alpha + 6$ with violation ratio $2 + \frac{2}{\alpha - 1}$, $\alpha \geq 4$. Moreover, based on a local search technique, Korupolu et al. (2000) proposed a $(1 + 5/\epsilon)$-approximation algorithm by using at most $(5 + \epsilon)k$ facilities, and a $(1 + \epsilon)$-approximation algorithm by using at most $(5 + 5/\epsilon)k$ facilities. For CKM with hard non-uniform capacities, Byrka et al. (2013) gave a $(96 + 180/\epsilon)$-approximation algorithm while violating capacities by factor $3 + 3\epsilon$, $\epsilon > 0$.

For CKM with soft capacities, based on primal-dual and Lagrangian relaxation methods, Chuzhoy and Rabani (2005) presented a 40-approximation algorithm by violating the capacities within a constant factor of 50. Bartal et al. (2001) proposed a $19.3(1 + \epsilon)/\epsilon^2$-approximation algorithm ($\epsilon > 0$) by using at most $(1 + \epsilon)k$ facilities.

For CKFL with uniform hard capacities, the $(32l^2 + 28l + 7)$-approximation algorithm (Byrka et al., 2013) also works by increasing the capacity by a factor of $2 + \frac{3}{l-1}$, $l \in \{2, 3, 4, \ldots\}$.

Without violating any constraint, a constant factor approximation algorithm remains unknown even for the single-sink capacitated k-median problem in which $|D| = 1$, let alone the capacitated k-facility location problem.

1.2. Our contributions and techniques

(i) We show that the single-sink hard capacitated k-facility location problem, in which D contains exactly one client, is NP-hard even when $f_i = 0$, $i \in F$. We give the first FPTAS for SCKFL by extending the FPTAS for the knapsack problem. To the best of our knowledge, this is also the first FPTAS for the single-sink capacitated facility location problem, which answers a question by Görtz and Klose (2009).

(ii) For the hard capacitated k-facility location problem with uniform capacities, in which $s_i = s$, $i \in F$, we observe that for $|D| = 1$, it is easy to find an optimal solution. A natural question is whether we can extend this to the case $m := |D| > 1$. For general m, the answer is negative since even a special case of this problem, named the hard capacitated facility location problem with uniform capacities, is NP-hard (Chudak & Williamson,

2005; Korupolu et al., 2000). However, for fixed m, the question is still open. We give a positive answer for this setting by proposing a polynomial time algorithm that runs in time $O(n^m \cdot n^3)$, where $n = |F|$. Using the structure of the graph consisting of the fractional valued edges in any extreme solution, the problem is reduced to a number of transportation problems. This should be considered as a theoretical result. It will be impractical even for moderately large m.

(iii) We observe that the number of fractionally open facilities can be bounded by analyzing the rank of the constraint matrix corresponding to the tight constraints at a fractional extreme point solution. Then, we give approximation algorithms for two variants of the hard capacitated k-facility location problem based on this upper bound. Another example to show the potential power of the structure of extreme point solutions is that we can improve the previous best approximation ratio 5 obtained by Levi et al. (2012), and Bansal et al. (2012) for the capacitated facility location problem with uniform opening costs, by combining our technique with a pseudo-approximation algorithm for the k-median problem.

2. The single-sink capacitated k-facility location problem

In this section, we consider the single-sink capacitated k-facility location problem (SCKF). Since we only have one client with demand d, the formulation for the CKF is reduced to the following mixed integer program.

$$Z_{\text{MIP}} = \min \sum_{i \in F} (c_i x_i + f_i y_i) \tag{8}$$

subject to:

$$\sum_{i \in F} x_i = d, \tag{9}$$

$$\sum_{i \in F} y_i \leq k, \tag{10}$$

$$0 \leq x_i \leq s_i y_i, \qquad \forall i \in F, \tag{11}$$

$$y_i \in \{0, 1\}, \qquad \forall i \in F. \tag{12}$$

Again, the natural LP relaxation of SCKFL can be obtained by replacing constraints (12) by $0 \leq y_i \leq 1$, $\forall i \in F$. Let Z_{LP} denote the optimal cost of this LP relaxation.

Lemma 1. *The single-sink capacitated k-facility location problem is NP-hard even when $f_i = 0$ for all $i \in F$.*

Proof. Consider the case that $s_i > 1$, $c_i := 1 - \frac{1}{s_i}$ and $f_i = 0$ for all $i \in F$. We claim that

$$Z_{\text{MIP}} \leq d - k \iff \text{there exists } I \subseteq F \text{ with } |I| = k \text{ and } \sum_{i \in I} s_i = d. \tag{13}$$
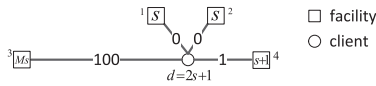
Fig. 1. An instance for SCKFL with unbounded integrality gap.

Indeed, for the objective value we find

$$\sum_{i \in F} c_i x_i = d - \sum_{i \in F} \frac{x_i}{s_i} = d - \sum_{i | y_i = 1} \frac{x_i}{s_i} \geq d - k,$$

where the last inequality holds because $x_i \leq s_i$ and $y_i = 1$ for at most k values of i. Equality holds if and only if $x_i = s_i$ for all $i \in F$ with $y_i = 1$ and $|\{i \in F \mid y_i = 1\}| = k$. That is, if and only if $\sum \{s_i \mid y_i = 1\} = d$ and $|\{i \in F \mid y_i = 1\}| = k$.

The claim above allows to reduce SUBSET-SUM, which is NP-hard (Garey & Johnson, 1979), to SCKFL as follows. Let positive integers $s_1, \ldots, s_n > 1$ and d form an instance of SUBSET-SUM. Now there exists a subset $I \subseteq \{1, 2, \ldots, n\}$ such that $\sum_{i \in I} s_i = d$ if and only if the objective value of SCKFL is at most $d - k$ for some $k \in \{1, \ldots, n\}$. □

Remark 1. The integrality gap $Z_{\mathrm{MIP}}/Z_{\mathrm{LP}}$ is unbounded. This implies that it is impossible to design a constant factor approximation algorithm for SCKFL based solely on this LP-relaxation.

Consider the instance shown in Fig. 1 with four facilities $F = \{1, 2, 3, 4\}$, $f = \{0, 0, 0, 0\}$, $c = \{0, 0, 100, 1\}$, $\{s_1, s_2, s_3, s_4\} = \{s, s, Ms, s + 1\}$, $d = 2s + 1$, $k = 2$ and $M \gg s \gg 100$.

For this instance, an optimal solution for the MIP is $y_1 = y_4 = 1$, $x_1 = s$, $x_4 = s + 1$ with the total cost $s + 1$. And an optimal solution for the LP-relaxation of this instance is $y_1 = 1$, $y_2 = \frac{Ms - s - 1}{(M-1)s}$, $y_3 = \frac{1}{(M-1)s}$, $x_1 = s$, $x_2 = \frac{Ms - s - 1}{M - 1}$, $x_3 = \frac{M}{M-1}$, with the total cost $\frac{100M}{M-1}$. So, we have $Z_{\mathrm{MIP}} = s + 1$ and $Z_{\mathrm{LP}} = \frac{100M}{M-1}$. Thus, $Z_{\mathrm{MIP}}/Z_{\mathrm{LP}} = \frac{s+1}{\frac{100M}{M-1}} > \frac{s+1}{200}$, which can be arbitrarily large.

In addition, a simple LP-rounding technique does not work for SCKFL. For this instance, a natural idea is to round y_3 to be 1, y_2 to be 0. It is clear that the objective value of the solution obtained by this simple rounding is still really large.

We aim to design a fully polynomial time approximation scheme (FPTAS) for SCKFL. Before introducing our algorithm, we present a key observation (Pál, Tardos, and Wexler gave a similar observation in the proof of Lemma 3.3 in Pál et al., 2001).

Observation 1. *For the single-sink capacitated k-facility location problem, there is an optimal solution (x^*, y^*) in which at most one open facility t^0 is not fully used, i.e., $x_i^* \in \{0, s_i\}$ for $i \neq t^0$.*

Without loss of generality we suppose that c_{ij} and f_i are integral for all $i \in F$, $j \in D$. Given t^0, which is allowed not to be fully used in an optimal integral solution (x^*, y^*), in order to solve SCKFL it is sufficient to solve the following problem for a given integer p:

$$\max \left\{ \sum_{i \in F'} s_i \mid F' \subseteq F \setminus \{t^0\}, |F'| \leq k - 1, \sum_{i \in F'} (c_i s_i + f_i) = p \right\}. \quad (14)$$

In words, we find for each total cost p a set of at most $k - 1$ facilities (not containing t^0) to open and use to full capacity, maximizing the total capacity. First, we will show that the above problem can be solved by dynamic programming, but may require exponential time in the size of the instance. Then, we use a rounding-and-scaling process to make sure that a feasible solution can be obtained (if it exists) in polynomial time while losing at most a factor of $1 + \epsilon$ to the optimal cost.

We can recursively solve problem (14) by dynamic programming. Without loss of generality, suppose $F = \{1, 2, \ldots, n\}$, where $n = |F|$. For nonnegative integers p, $0 \leq g \leq k - 1$ and $1 \leq b \leq n$

define

$$S_g(b, p) := \max \left\{ \sum_{i \in F'} s_i \mid F' \subseteq \{1, \ldots, b\} \setminus \{t^0\}, |F'| \leq g, \right.$$
$$\left. \sum_{i \in F'} (c_i s_i + f_i) = p \right\},$$

and let $F_g(b, p)$ be a corresponding optimal solution F'. If $\sum_{i \in F'} (c_i s_i + f_i) = p$ does not hold for any $F' \subseteq \{1, \ldots, b\} \setminus \{t^0\}$ with $|F'| \leq g$, we set $S_g(b, p) := -\infty$ and $F_g(b, p) := \emptyset$. Clearly, $S_g(0, 0) = 0$ and $S_g(0, p) = -\infty$ for $p > 0$. The other values $S_g(b, p)$, and the corresponding optimum solutions $F_g(b, p)$, can be computed recursively since if $t^0 = b + 1$, then

$$S_g(b + 1, p) = S_g(b, p)$$

for $0 < g \leq k - 1$; if $t^0 \neq b + 1$, then

$$S_g(b + 1, p) = \max \left(S_g(b, p), s_{b+1} + S_{g-1}(b, p - (f_{b+1} + c_{b+1}s_{b+1})) \right)$$

for $0 < g \leq k - 1$, where in the maximum the two values correspond to not opening and opening facility $b + 1$, respectively.

For computing the maximum in (14), it suffices to restrict to values $0 \leq p \leq (k - 1)\mathcal{P} \leq n\mathcal{P}$, where $\mathcal{P} = \max\{c_i s_i + f_i \mid i \in \{1, \ldots, n\} \setminus \{t^0\}\}$. Hence we can solve (14) in time $O(n^3 \mathcal{P})$ since $g \leq n$, $b \leq n$.

Since $\mathcal{P}$ may be exponential in the size of the input of SCKFL, the computing time could be non-polynomial. We overcome this difficulty by a scaling-and-rounding technique while losing at most a factor of $1 + \epsilon$ to the optimal cost. More precisely, we define $C_b^0 = \max\{c_i s_i + f_i \mid i \in \{1, \ldots, b\} \setminus \{t^0\}\}$ and $W = \frac{\epsilon C_b^0}{k}$, where $\epsilon > 0$. Then, for each $i \in \{1, \ldots, b\} \setminus \{t^0\}$, we scale and round the cost $c_i s_i + f_i$ to $\lfloor \frac{c_i s_i + f_i}{W} \rfloor$. Since t^0 may not be fully used, we just scale its opening and service cost without rounding. In this way, we can make sure that $\mathcal{P}$ is not exponential with respect to the scaled (and rounded) cost. Moreover, we need to try all the possibilities of t^0 since t^0 is not known in advance. (This is why we enumerate t from 1 to n in the Algorithm 1.) The resulting Algorithm 1 may be seen as a generalization of the FPTAS for the knapsack problem (with cardinality constraints) (Caprara, Kellerer, Pferschy, & Pisinger, 2000; Lawler, 1979). The total cost of the obtained solution is bounded by Theorem 1.

Assumption 1. *For each $i \in F$, $C_i > 0$, where $C_i := c_i s_i + f_i$. There are at least two open facilities in an optimal solution.*

Note that if $C_i = 0$ and $s_i < d$ we directly open i and serve demand s_i of the single client by i without increasing any cost. If $C_i = 0$ and $s_i \geq d$, the optimal total cost is 0. Moreover, if in an optimal solution only one facility is opened, we can easily obtain an optimal solution by opening a facility in $\{i \in F \mid s_i \geq d\}$ with $\min\{C_i \mid s_i \geq d, i \in F\}$. A precise description of the FPTAS is given in Algorithm 1.

Theorem 1. *Let SOL be the cost of the solution returned by Algorithm 1. Then, $SOL \leq (1 + \epsilon)Z_{\mathrm{MIP}}$. The running time of Algorithm 1 is $O(\frac{n^6}{\epsilon})$, for any $\epsilon > 0$.*

Proof. Suppose that (x^*, y^*) is an optimal solution in which at most one open facility is not fully used. Let t^0 be the open facility in (x^*, y^*) that is not fully used if it exists. Otherwise, let t^0 be some open facility in (x^*, y^*). Then, we define $F_1^* = \{i \in F \mid y_i^* = 1\} \setminus \{t^0\}$ as the set of opened facilities in (x^*, y^*) excluding t^0. By Assumption 1, we know $F_1^* \neq \emptyset$.

Let $C_i^0 = \max\{C_i \mid i \in F_1^*\}$. Note that $C_i^0 \leq Z_{\mathrm{MIP}}$. Moreover, let $i^0 = \max\{i \in F_1^* \mid C_i = C_i^0\}$. Thus, $C_{i^0} = C_i^0$.

Suppose that in iteration $t = t^0$, $r = i^0$ of Step 2 in Algorithm 1 we get a solution (x, y). Let $F_1 = \{i \in F \mid y_i = 1\} \setminus \{t^0\}$. Let $\mathrm{Cost}(x, y)$ and $\mathrm{Scaled\_cost}(x, y)$ be the original and scaled total cost of solution (x, y) respectively. So, $\mathrm{Cost}(x, y) = (\sum_{i \in F_1} C_i) + x_{t^0} c_{t^0} + f_{t^0}$, and $\mathrm{Scaled\_cost}(x, y) = (\sum_{i \in F_1} \bar{C}_i) + x_{t^0} \bar{c}_{t^0} + \bar{f}_{t^0}$, where the definition of $\bar{C}_i$ is given in Algorithm 1. We will show that $\mathrm{Cost}(x, y) \leq (1 + \epsilon)Z_{\mathrm{MIP}}$, which then implies $SOL \leq (1 + \epsilon)Z_{\mathrm{MIP}}$.

Algorithm 1 An FPTAS for the single sink capacitated k-facility location problem

Input Finite set F of facilities, costs $c \in \mathbb{Z}_{\geq 0}^F$, costs $f \in \mathbb{Z}_{\geq 0}^F$, demand $d \in \mathbb{Z}_{\geq 0}$, capacities $s \in \mathbb{Z}_{\geq 0}^F$, integer $1 \leq k \leq n := |F|, \epsilon > 0$.

Output A feasible solution (x, y) that is within a factor $1 + \epsilon$ of optimum, if a feasible solution exists.

Description

1. Suppose $F = \{1, 2, \ldots, n\}$. Let $C_i := c_i s_i + f_i, i \in F$.
2. for $t = 1$ to n do

    for $r = 1$ to n do

    if $\{1, \ldots, r\} \setminus \{t\} = \emptyset$,

    Let $C_r^0 = 0, \bar{f}_t = f_t, \bar{c}_t = c_t, W = 1$.

    end if

    if $\{1, \ldots, r\} \setminus \{t\} \neq \emptyset$

    Let $W = \frac{\epsilon C_r^0}{k}$, where $C_r^0 = \max\{C_i \mid i \in \{1, \ldots, r\} \setminus \{t\}\}$.

    For each facility $i \in \{1, \ldots, r\} \setminus \{t\}$, define $\bar{C}_i = \lfloor \frac{C_i}{W} \rfloor$.

    Let $\bar{f}_t = \frac{f_t}{W}, \bar{c}_t = \frac{c_t}{W}$.

    end if

    Consider the subproblem P_{rt} involving items $\{1, \ldots, r\} \cup \{t\}$, in which only t can be not fully used, that is, $x_i \in \{0, s_i\}, i \in \{1, \ldots, r\} \setminus \{t\}$; $0 \leq x_t \leq s_t$. With the above scaled costs, compute $S_g(\bar{r}, p)$ for each $0 \leq g \leq k-1, 0 \leq p \leq (k-1) \lfloor \frac{C_r^0}{W} \rfloor$, where $\bar{r} = r$ if $r \neq t, \bar{r} = r - 1$ otherwise. Then, find a solution with total scaled cost:

    $\min\{p + (d - S_g(\bar{r}, p))\bar{c}_t + \bar{f}_t \mid 0 \leq d - S_g(\bar{r}, p) \leq s_t, 0 \leq$

    $g \leq$

    $k - 1, \ 0 \leq p \leq (k-1)\lfloor \frac{C_r^0}{W} \rfloor \}$, if a feasible solution exists.

    end for

    end for
3. Output the solution with the minimum total original cost.

Recall that $W = \frac{\epsilon C_r^0}{k}$. We have

$$\text{Cost}(x, y) = \left(\sum_{i \in F_1} C_i \right) + x_{t^0} c_{t^0} + f_{t^0}$$

$$\leq \left(\sum_{i \in F_1} (W\bar{C}_i + W) \right) + W x_{t^0} \bar{c}_{t^0} + W\bar{f}_{t^0}$$

$$\leq W\left(\left(\sum_{i \in F_1} \bar{C}_i \right) + x_{t^0}\bar{c}_{t^0} + \bar{f}_{t^0} \right) + kW$$

$$\leq W \cdot \text{Scaled\_cost}(x, y) + kW,$$

where the second inequality holds as $|F_1| \leq k - 1$.

The scaled total cost of solution (x^*, y^*) in this iteration is $\left(\sum_{i \in F_1^*} \bar{C}_i\right) + x_{t^0}^* \bar{c}_{t^0} + \bar{f}_{t^0}$. Clearly,

$$\text{Scaled\_cost}(x, y) \leq \left(\sum_{i \in F_1^*} \bar{C}_i \right) + x_{t^0}^* \bar{c}_{t^0} + \bar{f}_{t^0},$$

since (x, y) is optimal in this iteration. That is,

$$\text{Scaled\_cost}(x, y) \leq \left(\sum_{i \in F_1^*} \left\lfloor \frac{C_i}{W} \right\rfloor \right) + x_{t^0}^* \frac{c_{t^0}}{W} + \frac{f_{t^0}}{W}.$$

Then, we have

$$W \cdot \text{Scaled\_cost}(x, y) \leq W\left(\sum_{i \in F_1^*} \left\lfloor \frac{C_i}{W} \right\rfloor \right) + W x_{t^0}^* \frac{c_{t^0}}{W} + W\frac{f_{t^0}}{W}$$

$$\Rightarrow W \cdot \text{Scaled\_cost}(x, y) + kW \leq \left(\sum_{i \in F_1^*} C_i \right) + x_{t^0}^* c_{t^0} + f_{t^0} + kW.$$

Therefore, we get

$$\text{Cost}(x, y) \leq \left(\sum_{i \in F_1^*} C_i \right) + x_{t^0}^* c_{t^0} + f_{t^0} + kW$$

$$= Z_{\text{MIP}} + \epsilon C_i^0 \leq (1 + \epsilon) Z_{\text{MIP}},$$

where the equality holds by the definition of W and the last inequality holds as $C_i^0 \leq Z_{\text{MIP}}$.

For fixed t, the running time of the subproblem $P_{rt}, r = 1, \ldots, n$ is $O(n^3 \lfloor \frac{C_r^0}{W} \rfloor)$. That is, $O(n^3 \frac{k}{\epsilon})$. Thus, the total running time of our algorithm is $O(\frac{n^6}{\epsilon})$ as we have $O(n^2)$ subproblems. □

3. The capacitated k-facility location problem with uniform capacities

In this section, we aim to show the following result for the capacitated k-facility location problem with uniform capacities (CKFLU). Let $m = |D|, n = |F|$ and $s_i = s, i \in F$.

Theorem 2. *For fixed m, the capacitated k-facility location problem with uniform capacities can be solved in polynomial time $O(n^m \cdot n^3)$.*

Note that if $n \leq m$, then Theorem 2 holds since n is also fixed now. Thus, from now on we consider the case $n > m$.

We need new notation to describe our idea. We consider an optimal solution (x, y) for CKFLU as a weighted bipartite graph $G = (V, E)$, where $V = \{i \in F \mid y_i = 1\} \cup D$ and $E = \{\{i, j\} \mid x_{ij} > 0, i \in F, j \in D\}$. To be more precise, if $x_{ij} > 0$, we add an edge $\{i, j\}$ between facility i and client j with weight x_{ij}. Moreover, let $\bar{E} = \{\{i, j\} \in E \mid 0 < x_{ij} < s\}$ be the set of edges in E with positive weight strictly less than s, and $\bar{V} = \bigcup_{e \in \bar{E}} e \cup D$ be the set of nodes consisted of the clients and the nodes incident to an edge in $\bar{E}$, as shown in Fig. 2. We call $(\bar{V}, \bar{E})$ the *untight* weighted subgraph of G.

Define $r_j := d_j / s$ for all $j \in D$. If all r_j are integral, we say that the CKFLU is *divisible*.

Lemma 2. *The divisible capacitated k-facility location problem with uniform capacities can be solved in $O(n^3)$ time.*

Proof. We transform the divisible CKFLU to a balanced transportation problem, in which the total capacity is equal to total demand. Then, to get an integer solution to this transportation problem, we can consider this problem as a minimum weight perfect matching problem that can be solved in $O(n^3)$ time (Gabow, 1976), by splitting the demands. Since the problem is infeasible if $k < \sum_{j \in D} r_j$, we only consider the case: $|F| \geq k \geq \sum_{j \in D} r_j$.

By dividing the capacity and demand constraints by s, we can get an equivalent formulation for the divisible CKFLU, in which the new capacity of each facility is 1 and the new demand of each client j is r_j.

First, we show that there exists an optimal integral solution for this equivalent formulation. We add a dummy client j' to D with demand $r_{j'} = n - \sum_{j \in D} r_j$. Take the cost of shipping one unit from $i \in F$ to $j \in D \setminus \{j'\}$ to be $s c_{ij} + f_i$, from $i \in F$ to j' to be 0. Now the divisible
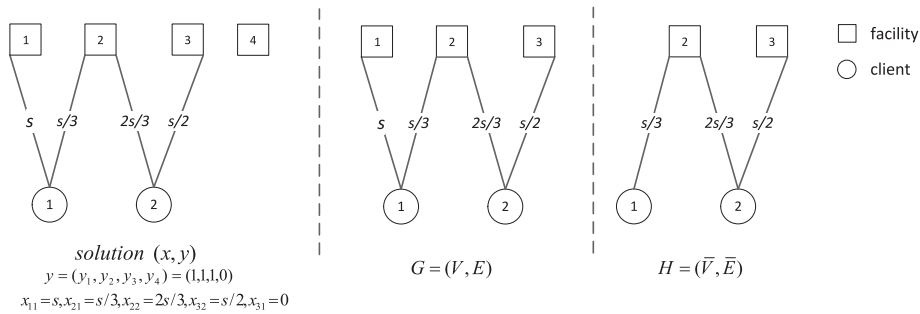
Fig. 2. An illustration of a solution with its corresponding graphs.

CKFLU can be considered as a balanced transportation problem with total demand n. Since $r_j, j \in D$ are integers, there is an integer optimal solution for this transportation problem (see for instance Hoffman & Kruskal, 1956, or Theorem 21.14 in Schrijver, 2003). Note that based on the optimal integer solution for this transportation problem, we can easily construct an optimal solution for our original problem.

Then, to get an optimal integer solution for the constructed transportation problem, we can split each $j \in D$ to r_j copies each with demand 1. Now we can consider the balanced transportation problem as a minimum weight perfect matching problem that can be solved in $O(n^3)$ time (Gabow, 1976).   □

Note that if we know the exact structure of $(\bar{V}, \bar{E})$, then according to the definition of G the remaining part $(V, E \setminus \bar{E})$ can be generated by an optimal integer solution to an instance of the divisible CKFLU problem. Thus, the high-level idea is that we reduce our original problem to a collection of divisible CKFLU problems by checking all the possible structures of $(\bar{V}, \bar{E})$. To prove that we can examine all the structures in polynomial time, we show some useful properties of the untight weighted subgraph of G first.

Lemma 3. *Let $G = (V, E)$ be the graph corresponding to a vertex (x, y) of the convex hull of feasible solutions of the MIP to CKFLU, and $H = (\bar{V}, \bar{E})$ be its corresponding untight subgraph. Then,*

(a) G is acyclic;
(b) in each connected component of H, there is at most one $i \in F \cap \bar{V}$ with $0 < \sum_{j \in D} x_{ij} < s$;
(c) H contains at most m facilities and $2m - 1$ edges.

Proof.

(a) Suppose that there is a cycle $O = (e_1, e_2, \ldots, e_{2p-1}, e_{2p})$ in G. Note that O must have even number of edges as G is bipartite. Let $\chi^O \in \mathbb{R}^E$ be the signed incidence vector of this path:

$$\chi^O(e_i) = (-1)^i \text{ for } i = 1, 2, \ldots, 2p;$$
$$\chi^O(e) = 0 \text{ for } e \in E \setminus \{e_1, \ldots, e_{2p}\}.$$

For sufficiently small $\epsilon > 0$ both $(x + \epsilon \chi^O, y)$ and $(x - \epsilon \chi^O, y)$ are feasible solutions, contradicting the fact that (x, y) is a vertex.

(b) The idea is similar to (a). Consider any connected component B of H. Suppose for contradiction that we have two facilities i_1, i_2 in B with $0 < \sum_{j \in D} x_{i_1 j} < s, 0 < \sum_{j \in D} x_{i_2 j} < s$. Since B is connected, there is a path $P = (e_1, e_2, \ldots, e_{2p-1}, e_{2p})$ from i_1 to i_2. Again, we can construct two feasible solutions $(x + \epsilon \chi^P, y)$ and $(x - \epsilon \chi^P, y)$, contradicting the fact that (x, y) is a vertex.

(c) Consider any connected component of H with at least one edge. Note that each component is a tree with $0 < x_{ij} < s$ for each edge $\{i, j\}$. If there is a facility i^* in this component with $0 < \sum_{j \in D} x_{i^* j} < s$, then take i^* as the root. Otherwise, take an arbitrary facility as the root. Since $0 < x_{ij} < s$ for each edge $\{i, j\}$

and $\sum_{j \in D} x_{ij} = s$ for each facility $i \neq i^*$, each facility except i^* has at least two neighbors. Then, each facility in this connected component has at least one child (client) as each facility has at most one parent. Moreover, no two facilities have a common child (by the definition of a rooted tree). Therefore, the number of facilities in each connected component is at most the number of clients. Thus, we have at most m facilities in H as there are at most m clients. Clearly, the number of edges is at most $2m - 1$ since H is a forest.   □

Lemma 4. *Let $G = (V, E)$ be the graph corresponding to a vertex (x, y) of the convex hull of feasible solutions of the MIP to CKFLU, and $H = (\bar{V}, \bar{E})$ be its corresponding untight subgraph. Given the set $I = \{i \in F \cap \bar{V} \mid 0 < \sum_{j \in D} x_{ij} < s\}$ and $\bar{E}$, we can get the unique weight x_{ij} for each edge $\{i, j\} \in \bar{E}$ in $O(m)$ time.*

Proof. Consider any connected component of H. Note that each connected component must be in the form of a tree. If there is a facility $i^* \in I$ in this component, then take i^* as the root. Otherwise, take an arbitrary facility i^* in this component as the root. Then, all leaves are clients since $\sum_{j \in D} x_{ij} = s$ for each facility $i \neq i^*$ in the considered connected component (Lemma 3(b)) and $0 < x_{ij} < s$ for each edge $\{i, j\}$.

We will show that in each connected component, if node (client) j is a leaf, we can obtain the exact value of x_{ij}, where i is the father of j; and for each other node in this tree, we can compute the value of the edge between this node and its father based on the values of edges between this node and its children. Then, we can obtain the values of all edges in the tree by induction.

Consider a client j. Let $f(j)$ be the father node (facility) of j in the tree and $c(j)$ be the set of children (facilities) of j. If j is a leaf, that is $c(j) = \emptyset$, then we know $|\{i \in F \mid 0 < x_{ij} < s\}| = 1$. Otherwise, j cannot be a leaf. Thus, we can get the exact value for $x_{f(j),j} = d_j - \lfloor \frac{d_j}{s} \rfloor \cdot s$ since j has exactly one father. If j is not a leaf, the value $x_{f(j),j} = (d_j - \sum_{i \in c(j)} x_{ij}) - \lfloor \frac{d_j - \sum_{i \in c(j)} x_{ij}}{s} \rfloor \cdot s$ as $x_{tj} \in \{0, s\}, \forall t \in V \setminus \bar{V}$.

Consider a facility $i \neq i^*$. Let $f(i)$ be the father node (client) of i in the tree and $c(i)$ be the set of children (clients) of i. We can obtain the value of $x_{i,f(i)}$ as long as all values of $x_{ij}, j \in c(i)$ are known, since i must be fully used by Lemma 3. That is, $x_{i,f(i)} = s - \sum_{j \in c(i)} x_{ij}$. Note that if $i = i^*$, we can stop since $f(i^*) = \emptyset$.

Moreover, the computing time is $O(m)$ since each edge is examined at most twice.   □

Consider an optimal vertex (x, y) of the convex hull of feasible solutions for CKFLU whose corresponding graph $G = (V, E)$ is a forest. The graph $H = (\bar{V}, \bar{E})$ (the untight subgraph of G) can be viewed as a subgraph of some spanning tree of the complete bipartite graph $K_{\bar{F}, D}$, where $\bar{F} = F \cap \bar{V}$. Consequently, checking all the possible structures of H means checking all the subgraphs of these spanning trees. Note that H and $K_{\bar{F}, D}$ have the same vertices. Then, in order to prove Theorem 2, it now suffices to answer the following questions:

1. How many different choices do we have for $K_{\bar{F},D}$?
2. How to list all the spanning trees for a complete bipartite graph?
3. How many subgraphs, that have the same vertices as the considered spanning tree, does a spanning tree have?
4. For a fixed structure of H, how to compute the corresponding total cost?

If all the above questions can be solved in polynomial time, we can get all the possibilities of H in polynomial time. Consequently, Theorem 2 can be proved by Lemmas 2 and 4.

Lemma 5 (Kapoor & Ramesh, 1995). *For an undirected graph without weights $G = (V, E)$, all spanning trees can be correctly generated in $O(N + |V| + |E|)$ time, where N is the number of spanning trees.*

Lemma 6 (Scoins, 1962). *The number of spanning trees of a complete bipartite graph is $m^{n-1}n^{m-1}$, where m and n are respectively the cardinalities of two disjoint sets in this bipartite graph.*

Proof of Theorem 2. Since $H = (\bar{V}, \bar{E})$ contains at most m facilities by Lemma 3, the number of all the possible cases for $K_{\bar{F},D}$ can be bounded by $\sum_{t=1}^{m} \binom{n}{t} \leq m \cdot n^m$. So, we can answer question 1.

Lemmas 5 and 6 answer question 2. The time to list all the spanning trees for the complete bipartite graph is $O(m^{2m-2} + 2m + m^2)$ since we have at most m facilities and m clients in $K_{\bar{F},D}$ by Lemma 3. Note that at this stage, we do not need to consider the weight x_{ij} of edge $\{i, j\}$.

By Lemma 3, we know that the number of edges is at most $2m - 1$ in a spanning tree. Thus, each spanning tree has at most 2^{2m-1} subgraphs that have the same vertices as the spanning tree. This answers question 3.

Then, the total time to list all the possible untight subgraphs is $O(m \cdot n^m \cdot (m^{2m-2} + 2m + m^2) \cdot 2^{2m-1})$.

By Lemma 4, we can get the cost for any untight subgraph in polynomial time $O(m)$ as long as $I = \{i \in F \cap \bar{V} \mid 0 < \sum_{j \in D} x_{ij} < s\}$ is fixed. Note that the opening costs for facilities are easy to get if we know the structure of H. Indeed, it is $\sum_{i \in F \cap \bar{V}} f_i$. The remaining part $(V, E \setminus \bar{E})$ can be considered as an optimal integer solution to a divisible CKFLU, which means we can get the total cost in polynomial time $O(n^3) + O(m)$ by Lemma 2. This answers question 4. Moreover, the number of all the choices for I is bounded by 2^m since there are at most m facilities in each spanning tree by Lemma 3.

Combining all the pieces together, we can get all the possibilities of solutions in computing time $O(m \cdot n^m \cdot (m^{2m-2} + 2m + m^2) \cdot 2^{2m-1} \cdot 2^m \cdot (m + n^3)) = O(n^m \cdot (m^{2m-1} + 2m^2 + m^3) \cdot 2^{3m-1} \cdot (m + n^3))$, that is, $O(n^m \cdot n^3)$ since m is fixed. Finally, we output the solution with at most k open facilities and the smallest total cost. □

4. The hard capacitated k-facility location problem with non-uniform capacities

In this section, we show how to bound the number of fractionally open facilities by a simple rank-counting argument on an extreme point solution. Then, together with an algorithm to group clients, we give a pseudo-approximation algorithm for the hard CKFL (with uniform opening costs) with approximation ratio $7 + \epsilon$ by using at most $2k$ facilities. Finally, we extend this algorithm to improve the previous best approximation ratio for the capacitated facility location problem with uniform opening costs.

As a simple illustration of the techniques used, we first give a 2-approximation algorithm for the single-sink hard capacitated k-facility location problem (SCKFL). Note that this ratio is worse than that of the FPTAS in Section 2. Here we aim to show that this upper bound is helpful to design approximation algorithms. And the approach is totally different from the FPTAS.

4.1. A simple illustration of using the structure of extreme point solutions

4.1.1. The structure of extreme point solutions to SCKFL

Definition 1. Let $Ax \leq a, Bx \geq b, Cx = c$ be a system of linear (in)equalities. For a feasible solution z we define *the rank at z of the* system to be the (row)rank of $[A_z^\top \ B_z^\top \ C^\top]^\top$, where $A_z x \leq a_z, B_z x \geq b_z, Cx = c$ is the subsystem consisting of the (in)equalities that are satisfied with equality by z.

Note that for two subsystems, the sum of the ranks at z of those two subsystems is at least the rank at z of their union.

Let P be the set of feasible solutions to the system SCKFL-LP consisting of (7), (9), (11) and $\sum_{i \in F} y_i = k$ (Note that in this section we consider constraint $\sum_{i \in F} y_i = k$ instead of the corresponding inequality (10)). That is,

$$P := \{(x, y) : \text{SCKFL-LP}\},$$

where SCKFL-LP is a system of constraints given below:

$$\sum_{i \in F} x_i = d, \quad \sum_{i \in F} y_i = k, \tag{15}$$
$$0 \leq x_i \leq s_i y_i, \quad \forall i \in F,$$
$$0 \leq y_i \leq 1, \quad \forall i \in F.$$

Lemma 7. *Let (x, y) be a vertex of P. Then either y is integer, or y has exactly two noninteger components and for every $i \in F$ we have $x_i = 0$ or $x_i = s_i y_i$.*

Proof. Let $F' := \{i \in F \mid 0 < y_i < 1\}$. If $|F'| = 0$ we are done. As $|F'| = 1$ is ruled out because the sum of the y_i is k, $k \in \mathbb{Z}$, we may assume that $|F'| \geq 2$.

The rank of system SCKFL-LP at (x, y) is equal to $2n$, $n = |F|$ (Theorem 5.7 in Schrijver, 2003). We partition the (in)equalities in this system and bound the rank at (x, y) for each subsystem:

- The rank at (x, y) of the subsystem $\sum_{i \in F} x_i = d$, $\sum_{i \in F} y_i = k$ is at most 2.
- For every $i \in F'$, the rank at (x, y) of the subsystem $0 \leq x_i, x_i \leq s_i y_i, 0 \leq y_i, y_i \leq 1$ is at most 1 and equality holds if and only if $x_i = 0$ or $x_i = s_i y_i$.
- For every $i \in F \setminus F'$, the rank at (x, y) of the subsystem $0 \leq x_i, x_i \leq s_i y_i, 0 \leq y_i, y_i \leq 1$ is at most 2 and equality holds if and only if $x_i = 0$ or $x_i = s_i y_i$.

Since the rank is subadditive, we find that the rank at (x, y) of SCKFL-LP is at most

$$2 + |F'| + 2|F \setminus F'| = 2n + 2 - |F'| \leq 2n,$$

where the inequality holds as $|F'| \geq 2$, with equality only if $|F'| = 2$ and for each i we have $x_i = 0$ or $x_i = s_i y_i$. □

We give a 2-approximation algorithms for SCKFL to show the potential power of this nice structure.

4.1.2. 2-Approximation algorithm for SCKFL

We give an alternative approach to get an approximate solution for SCKFL, compared to the FPTAS in Section 2. This approach can be viewed as incomplete implement of a branch and bound technique, branching on the 0-1 variables y_i. To obtain a 2-approximation algorithm that runs in polynomial time, we use two key ideas. First, by Lemma 7, we know that in any vertex of the feasible region of the LP-relaxation either 0 or 2 components of y are fractional. We exploit this to guide the branching. Secondly, we show that for a branch $y_i = 1$ either there is no 2-approximation solution, or we can find a 2-approximation solution in polynomial time by again exploiting the structure of the vertices of the feasible region to the LP-relaxation. A precise description of this algorithm is given in Algorithm 2.

Algorithm 2 A 2-approximation algorithm for the single-sink hard capacitated k-facility location problem

Input Finite set F of facilities, costs $c \in \mathbb{Z}_{\geq 0}^F$, costs $f \in \mathbb{Z}_{\geq 0}^F$, capacities $s \in \mathbb{Z}_{\geq 1}^F$, demand $d \in \mathbb{Z}_{\geq 1}$, integer $k \in \mathbb{Z}_{\geq 1}$.

Output A feasible solution (x, y) to MIP: (8), (11), (12), and (15), that is within a factor 2 of optimum, if a feasible solution exists.

Description

1. Find an optimal vertex (x, y) of the feasible region of the LP-relaxation.
   If no solution exists then stop. If y is integer then return (x, y) and stop.

2. Let $i_1 \neq i_2$ in F with $y_{i_1}, y_{i_2} \in (0, 1)$ and $s_{i_1} \geq s_{i_2}$.

3. Define x^1 by $x_{i_1}^1 := x_{i_1} + x_{i_2}$, $x_{i_2}^1 := 0$ and $x_i^1 := x_i$ for $i \neq i_1, i_2$.
   Define y^1 by $y_{i_1}^1 := 1$, $y_{i_2}^1 := 0$, $y_i^1 := y_i$ for $i \neq i_1, i_2$.

4. Recursively compute a 2-approximation solution (x^0, y^0) for the restriction to $F \setminus \{i_1\}$ and extend it by setting $x_{i_1}^0 := 0$ and $y_{i_1}^0 := 0$.

5. Set $F_0 := \emptyset$. While $|F_0| \leq |F| - k$ do:

   a. Find an optimal vertex (x', y') of the feasible region of the LP-relaxation intersected with $\{(x, y) \mid y_{i_1} = 1, y_i = 0 \, \forall i \in F_0\}$.

   b. If y' is integer, return the best solution among (x', y'), (x^0, y^0) and (x^1, y^1) and stop.

   c. If $x_{i_1}' = s_{i_1}$, return the best solution among (x^0, y^0) and (x^1, y^1) and stop.

   d. Let $i_3 \neq i_4$ in F with $y_{i_3}', y_{i_4}' \in (0, 1)$ and $f_{i_3} \leq f_{i_4}$.

   e. Define y'' by $y_{i_1}'' := 0$, $y_{i_3}'' := y_{i_4}'' := 1$ and $y_i'' := y_i'$ for $i \neq i_1, i_3, i_4$.
      If (x', y'') has smaller value than (x^0, y^0), set $(x^0, y^0) \leftarrow (x', y'')$.

   f. Set $F_0 \leftarrow F_0 \cup \{i_4\}$.

Theorem 3. *For the single-sink hard capacitated k-facility location problem, Algorithm 2 finds a solution that is within a factor 2 of optimum, or it concludes correctly that there is no feasible solution. The running time is polynomially bounded in the number $|F|$ of facilities.*

Proof. Notice that an optimal vertex of the feasible region of the LP-relaxation can be found in polynomial time (see for instance Grötschel, Lovász, & Schrijver, 1988). Furthermore, since the number of recursive calls is no more than $|F| - 1$, the polynomial running time is evident. It now suffices to show that when the MIP: (8), (11), (12), and (15) is feasible, the solution given by Algorithm 2 is within a factor 2 of optimum.

Clearly, if y is integer in Step 1 of Algorithm 2, then the output (x, y) is an optimal feasible solution. Hence, by Lemma 7, we may assume that y has exactly two fractional components y_{i_1} and y_{i_2}. Then, we know $y_{i_1} + y_{i_2} = 1$ since $\sum_{i \in F} y_i = k$, and all y_i, $i \in F$ are integer except y_{i_1} and y_{i_2}. Without loss of generality we can assume that $s_{i_1} \geq s_{i_2}$.

To see that (x^1, y^1) defined in Step 3 of Algorithm 2 is indeed a feasible solution, it suffices to show that $x_{i_1}^1 \leq s_{i_1}$. This follows directly from the fact that $s_{i_1} \geq s_{i_2}$, since

$$x_{i_1}^1 = x_{i_1} + x_{i_2} \leq y_{i_1} s_{i_1} + y_{i_2} s_{i_2} \leq y_{i_1} s_{i_1} + y_{i_2} s_{i_1} = s_{i_1}.$$

Further, we find an upper bound for the value of (x^1, y^1),

$$c^\mathsf{T} x^1 + f^\mathsf{T} y^1 \leq (c^\mathsf{T} x + f^\mathsf{T} y) + (c_{i_1} s_{i_1} + f_{i_1}), \quad (16)$$

which is at most the optimum plus $c_{i_1} s_{i_1} + f_{i_1}$.

To conclude the proof, we analyze Step 5 of Algorithm 2. Observe that the initial solution (x^0, y^0) may be replaced, but only by a better solution. Also observe, that the solution that is returned is always at least as good as (x^0, y^0) and (x^1, y^1). Hence, we may assume that (x^0, y^0) (at the end of the algorithm) and (x^1, y^1) are not 2-approximations. Let (x^*, y^*) be an optimal solution. We have $y_{i_1}^* = 1$, since otherwise (x^0, y^0) would be a 2-approximation already at Step 4. It suffices to show that (x^*, y^*) remains feasible throughout the iterations of Step 5, until a solution of the same value is returned in Step 5b. For this, we observe that while (x^*, y^*) is feasible, the situation $x_{i_1}' = s_{i_1}$ as in Step 5c cannot occur, because otherwise, by (16), we would have

$$c^\mathsf{T} x^1 + f^\mathsf{T} y^1 \leq c^\mathsf{T} x + f^\mathsf{T} y + (c_{i_1} s_{i_1} + f_{i_1})$$
$$\leq c^\mathsf{T} x + f^\mathsf{T} y + c^\mathsf{T} x' + f^\mathsf{T} y' \leq 2(c^\mathsf{T} x^* + f^\mathsf{T} y^*),$$

contradicting the fact that (x^1, y^1) is not a 2-approximation.

In Step 5d, the fact that y' has exactly two fractional components follows from Lemma 7 as y' is a vertex of a face of the feasible region of SCKFL-LP, and hence of that region itself. Observe that this implies that $y_{i_3}' + y_{i_4}' = 1$. Moreover, $x_{i_1}' = 0$ again by Lemma 7 since $x_{i_1}' \neq s_{i_1}$ in this step. Hence (x', y'') defined in Step 5e is a feasible solution.

In Step 5f, we have $y_{i_4}^* = 0$. Indeed, for the cost of (x', y'') we find:

$$c^\mathsf{T} x' + f^\mathsf{T} y'' = (c^\mathsf{T} x' + f^\mathsf{T} y') - f_{i_1} + (1 - y_{i_3}') f_{i_3} + (1 - y_{i_4}') f_{i_4}$$
$$\leq (c^\mathsf{T} x' + f^\mathsf{T} y') - f_{i_1} + (1 - y_{i_3}' + 1 - y_{i_4}') f_{i_4}$$
$$\leq (c^\mathsf{T} x' + f^\mathsf{T} y') + f_{i_4} \leq (c^\mathsf{T} x^* + f^\mathsf{T} y^*) + f_{i_4}.$$

Since (x^0, y^0) and hence (x', y'') is not a 2-approximation, we find that $f_{i_4} > c^\mathsf{T} x^* + f^\mathsf{T} y^*$ and hence $y_{i_4}^* = 0$. This shows that (x^*, y^*) remains feasible after adding i_4 to F_0. □

4.2. A pseudo-approximation algorithm for CKFL with uniform opening costs

In this section, we consider the capacitated k-facility location problem with uniform opening costs, i.e., $f_i = f$, $i \in F$. Since we have an upper bound on the number of fractionally open facilities based on Lemma 8 below, a natural idea is to design a constant factor approximation algorithm for CKFL by relaxing the cardinality constraint with a constant factor. We give a simple algorithmic framework that can extend any α-approximation algorithm for the (uncapacitated) k-median problem (UKM) to a $(1 + 2\alpha)$-approximation algorithm for CKFL using at most $2k$ facilities ($2k - 1$ for uniform capacities).

The (uncapacitated) k-median problem (UKM) can be formulated as follows, where variable x_{ij} indicates the fraction of the demand of client j that is served by facility i, and y_i indicates whether facility i is open.

$$\min \sum_{i \in F} \sum_{j \in D} d_j c_{ij} x_{ij}$$

subject to:
$$\sum_{i \in F} x_{ij} = 1, \quad \forall j \in D,$$
$$x_{ij} \leq y_i, \quad \forall i \in F, \forall j \in D,$$
$$\sum_{i \in F} y_i \leq k,$$
$$x_{ij}, y_i \in \{0, 1\}, \forall i \in F, \forall j \in D.$$

4.2.1. The structure of extreme point solutions to CKFL

Let Q be the set of feasible solutions (x, y) to the system CKFL-LP consisting of (2), (3), (4), (5) and (7). That is,

$$Q := \{(x, y) : \text{CKFL-LP}\},$$

where CKFL-LP is a system of constraints given below:

$$\sum_{i \in F} x_{ij} = d_j, \quad \forall j \in D; \quad \sum_{i \in F} y_i \leq k,$$

$$\sum_{j \in D} x_{ij} \leq s_i y_i, \quad \forall i \in F,$$

$$x_{ij} \geq 0, \quad \forall i \in F, \forall j \in D,$$

$$0 \leq y_i \leq 1, \quad \forall i \in F.$$

Lemma 8. *Let (x, y) be a vertex of Q. Then y has at most $m + 1$ non-integer components, where $m = |D|$. And if the number of noninteger components is exactly $m + 1$, then $x_{ij} = 0$ or $x_{ij} = s_i y_i$ for each x_{ij}.*

Proof. The proof is similar to the proof of Lemma 7. Let $F' = \{i \in F \mid 0 < y_i < 1\}$. The rank of system CKFL-LP at (x, y) is equal to $(m + 1)n, n = |F|, m = |D|$ (Theorem 5.7 in Schrijver, 2003). We partition the (in)equalities in this system and bound the rank at (x, y) for each subsystem:

- The rank at (x, y) of the subsystem $\sum_{i \in F} x_{ij} = d_j, \forall j \in D; \sum_{i \in F} y_i \leq k$ is at most $m + 1$.
- For every $i \in F'$, the rank at (x, y) of the subsystem $\sum_{j \in D} x_{ij} \leq s_i y_i; 0 \leq x_{ij}, j \in D; 0 \leq y_i; y_i \leq 1$ is at most m and equality holds if and only if $x_{ij} = 0$ or $x_{ij} = s_i y_i$ for each x_{ij}.
- For every $i \in F \setminus F'$, the rank at (x, y) of the subsystem $\sum_{j \in D} x_{ij} \leq s_i y_i; 0 \leq x_{ij}, j \in D; 0 \leq y_i; y_i \leq 1$ is at most $m + 1$ and equality holds if and only if $x_{ij} = 0$ or $x_{ij} = s_i y_i$ for each x_{ij}.

Since the rank is subadditive, we find that the rank of CKFL-LP is at most

$$m + 1 + m|F'| + (m + 1)|F \setminus F'| = m + 1 + (m + 1)n - |F'|.$$

So, we have $|F'| \leq m + 1$ as $m + 1 + (m + 1)n - |F'| \geq (m + 1)n$. $\square$

For the uniform capacities case ($s_i = s > 0, \forall i \in F$), we will show a stronger property that there is an optimal solution (x, y) to the LP-relaxation with at most m noninteger components in y. Indeed, consider an optimal solution (x, y) with $|\{i \mid 0 < y_i < 1\}|$ minimal. Suppose for contradiction that y has more than m fractional components. Then there exist a client j and two facilities i_1, i_2 such that y_{i_1} and y_{i_2} are fractional and $x_{i_1 j}, x_{i_2 j} > 0$, and $x_{i_1 j} = s y_{i_1}, x_{i_2 j} = s y_{i_2}$ by Lemma 8. Without loss of generality, we assume that $c_{i_1 j} \leq c_{i_2 j}$. Let $\epsilon := \min\{s y_{i_2}, s(1 - y_{i_1})\}$. Now modify (x, y) by setting

$$x_{i_1 j} := x_{i_1 j} + \epsilon \quad y_{i_1} := y_{i_1} + \epsilon/s$$

$$x_{i_2 j} := x_{i_2 j} - \epsilon \quad y_{i_2} := y_{i_2} - \epsilon/s,$$

to obtain a new optimal solution, while $|\{i \mid 0 < y_i < 1\}|$ decreases, a contradiction. Thus, we can find an optimal solution (x, y) to the LP-relaxation for which y has at most m noninteger components.

4.2.2. The algorithm

The high level idea is as follows. We convert our original instance to a new instance with at most k clients while incurring some bounded extra costs. Then, at most $2k$ facilities are (fractionally or fully) opened for the new instance according to Lemma 8. Finally, we simply open all the fractionally and fully opened facilities to get an integral solution. A detailed description of this algorithm is given in Algorithm 3.

Theorem 4. *By Algorithm 3, each α-approximation algorithm for UKM can be extended to get a $(1 + 2\alpha)$-approximation algorithm for CKFL with uniform opening costs using at most $2k$ facilities.*

Proof. Without loss of generality, we suppose that exactly k facilities are opened in an optimal solution to our original problem (as we check all the cases in our algorithm).

Let $OPT(*)$ denote the optimal cost of the instance $*$. Let $COST(\cdot, \cdot)$ be the total cost of the solution $(\cdot, \cdot)$. We consider the following instances and solutions.

Algorithm 3 A $(1 + 2\alpha)$-approximation algorithm for CKFL with uniform opening costs using at most $2k$ facilities

---

Input Finite set F of facilities, D of clients, costs $c \in \mathbb{Q}_{\geq 0}^{F \times D}$, opening cost $f \in \mathbb{Q}_{\geq 0}$, capacities $s \in \mathbb{Q}_{\geq 0}^{F}$, demands $d \in \mathbb{Q}_{\geq 0}^{D}$, integer $k \in \mathbb{Z}_{\geq 1}$, an α-approximation algorithm for UKM.

Output A solution (x, y) to MIP (1)–(6) using at most $2k$ facilities that is within a factor $1 + 2\alpha$ of optimum, if a feasible solution exists.

Description

Suppose exactly l facilities are opened in an optimal solution. That is, we can consider a stronger constraint $\sum_{i \in F} y_i \leq l$.

Step 1. Reduce the input instance I_0 of CKFL to an instance I_1 of UKM as follows.

Let F and D be the set of facilities and clients of our input instance I_0 respectively. Let $F' = F$ (located at the same sites) be the set of facilities of UKM while with infinite capacities and without opening costs. Let $D' = D$ be the set of clients of UKM. Solve this constructed instance (denoted by I_1) by the existing α-approximation algorithm for UKM. Let (x', y') be the obtained solution. Note that for UKM, each feasible solution is in a form of so-called stars. That is, each client is served by exactly one open facility. Without loss of generality, suppose $y'_1 = \cdots = y'_l = 1$. Then, we can consider (x', y') as l stars $\{T_1, \ldots, T_l\}$, where $T_r = \{j \in D' \mid x'_{rj} = 1\}$ and the center of T_r is the facility r.

Step 2. Consolidate clients and construct a new instance I_2 of CKFL with at most l clients as follows.

For each star T_r in (x', y'), we set a client t_r at the location of facility r with the total demand of clients in T_r, i.e., $d_{t_r} = \sum_{j \in T_r} d_j$. Let $\bar{D} = \{t_1, \ldots, t_l\}$ be the set of our new clients. Now we get a new instance of CKFL, denoted by I_2, with facilities F and clients $\bar{D}$.

Step 3. Find an optimal vertex (x, y) of the feasible region of the LP-relaxation to the constructed instance I_2 in Step 2 with $\sum_{i \in F} y_i = l$.

Step 4. We simply open all the facilities with $y_i > 0$ in our original instance I_0 and then solve a transportation problem to get an integer solution (x^*, y^*).

Since we do not know how many facilities are opened in an optimal solution in advance, we repeat the above 4 steps for $l := 1, \ldots, k$. Then, output the solution with smallest total cost.

I_0	The original instance.
(x^*, y^*)	The obtained solution of our algorithm for instance I_0 while using at most $2k$ facilities.
I_1	The constructed instance in Step 1, that is a (uncapacitated) k-median problem.
(x', y')	The obtained integral solution by α-approx. alg. for instance I_1.
I_2	The constructed instance in Step 2 in which we have at most k clients.
(x, y)	An optimal fractional solution of instance I_2.

Clearly, we have $COST(x, y) \leq OPT(I_2)$, and $COST(x', y') \leq \alpha OPT(I_1)$.

By construction of instance I_2, we have $OPT(I_0) + COST(x', y') \geq OPT(I_2)$. To see this, we will show that we can construct a feasible solution for I_2 with total cost at most $OPT(I_0) + COST(x', y')$.

Consider any client $j \in D$. Let $F_j = \{i \in F \mid x_{ij}^0 > 0\}$ be the set of facilities which serve some positive amount of demand of j in an optimal solution (x^0, y^0) of instance I_0. Let t be the facility that serves j in solution (x', y') of I_1. Note that $\sum_{i \in F_j} x_{ij}^0 = x'_{tj} = d_j$. So, we can serve d_j
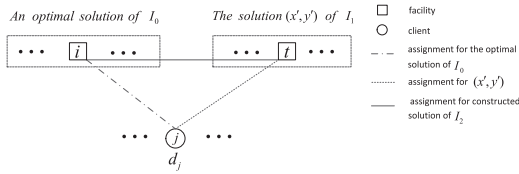
Fig. 3. An illustration for bounding the serve cost of the constructed solution for I_2.

units of demand of the new client t in instance I_2 by facilities F_j. That is, we can construct a feasible solution for I_2 from (x^0, y^0) and (x', y'), since the above operation works for any client $j \in D$.

Further, it is easy to bound the total cost of this constructed solution for I_2. By triangle inequalities, we know that $c_{ij} + c_{tj} \geq c_{it}$, where $i \in F_j$, as shown in Fig. 3. Since we open exactly k facilities in both the constructed solution for I_2 and the solution (x^0, y^0) for I_0, we get that the total cost of this constructed solution for I_2 is at most $OPT(I_0) + COST(x', y')$. Then, we obtain

$$OPT(I_0) + COST(x', y') \geq OPT(I_2)$$

since this constructed solution is impossible better than the optimal solution.

Moreover, we know that $OPT(I_0) \geq OPT(I_1) + kf$.

We will prove

$$COST(x^*, y^*) \leq COST(x', y') + COST(x, y) + kf.$$

We first show that we can obtain an integer solution for I_2 with the total cost at most $COST(x, y) + kf$ in Step 4 of Algorithm 3. We have $|\{i \mid 0 < y_i < 1\}| \leq k + 1$, since Lemma 8 still holds when $\sum_{i \in F} y_i = k$. Moreover, if $|\{i \mid 0 < y_i < 1\}| > 0$, then $|\{i \mid y_i = 1\}| \leq k - 1$. So, we open at most $2k$ facilities. Thus, the total cost of the obtained solution for I_2 is at most $COST(x, y) + kf$.

Then, based on the above solution for I_2 we can construct an integer solution for I_0 by moving the demand of t_r, which is located at the same position with facility r, back to all clients in $T_r = \{j \in D' \mid x'_{rj} = 1\}$ with increasing at most $COST(x', y')$ cost as $COST(x', y') = \sum_{r=1}^{k} \sum_{j \in D'} d_j c_{r,j} x'_{r,j}$. Therefore, the solution obtained by Step 4 has $COST(x^*, y^*) \leq COST(x', y') + COST(x, y) + kf$.

Then, we have

$$COST(x^*, y^*) \leq COST(x', y') + OPT(I_2) + kf$$
$$= (OPT(I_2) - COST(x', y')) + 2COST(x', y') + kf$$
$$\leq OPT(I_0) + 2COST(x', y') + kf$$
$$\leq OPT(I_0) + 2\alpha OPT(I_1) + kf$$
$$\leq OPT(I_0) + 2\alpha (OPT(I_1) + kf) \leq (1 + 2\alpha)OPT(I_0).$$

That is, the approximation ratio is $1 + 2\alpha$. □

We can obtain the following result as there is a $(3 + \epsilon)$-approximation algorithm for the (uncapacitated) k-median problem in Arya et al. (2001), and we can make sure that at most $2k - 1$ facilities are opened in Step 4 if all capacities are equal.

Corollary 1. *Algorithm 3 can get an integer solution within $7 + \epsilon$ times of the optimal cost by using at most $2k$ facilities $(2k - 1$ facilities) for the hard capacitated k-facility location problem with uniform opening costs (with uniform opening costs and uniform capacities).*

4.3. Extension

We show how to combine Algorithm 3 with the algorithm of Charikar and Guha (1999) for the uncapacitated facility location problem (UFL) to improve the approximation ratio for the capacitated facility location problem (CFL) with uniform opening costs. Recall that the capacitated facility location problem is a special case of CKFL with $k = n = |F|$. The uncapacitated facility location problem is a special

case of CFL with $s_i = +\infty$, $\forall i \in F$. In this section, we only consider uniform opening costs. To simplify the description, sometimes we omit "with uniform opening costs" when we refer to the problems.

A (β, δ)-approximation algorithm for the (uncapacitated) k-median problem (UKM) outputs an integer solution by using at most δk facilities, with service cost at most β times the optimal total cost.

Theorem 5. *Each (β, δ)-approximation algorithm for the k-median problem gives rise to a $\max\{2\beta + 1, \delta + 1\}$-approximation algorithm for the CFL with uniform opening costs.*

Proof. A crucial observation is that if exactly k facilities are opened in the optimal solution for an instance I of CFL, then this solution is also an optimal solution to the corresponding instance of CKFL, where the input is the same as that in I but with an extra constraint that at most k facilities can be opened. Thus, if for each $k = 1, \ldots, n, n = |F|$ we can obtain the optimal solution for CKFL, then the solution with smallest total cost is the optimal solution for CFL.

Our algorithm for CFL is as follows: Repeat the 4 steps in Algorithm 3 for $l := 1, \ldots, n, n = |F|$. In each iteration, we consider the constraint $\sum_{i \in F} y_i \leq l$. Then, output the solution with smallest total cost.

To get a better approximation ratio for CFL, we replace the α-approximation algorithm for the k-median problem in Step 1 of Algorithm 3 by a (β, δ)-approximation algorithm. Then, for each iteration l, we obtain an integer solution for the instance I_1 with at most δl open facilities. Thus, we have at most δl clients in instance I_2.

Again, without loss of generality, we suppose that exactly k facilities are opened in an optimal solution for the original instance I_0.

We maintain all the notations in the proof of Theorem 4. Notice that $COST(x, y) \leq OPT(I_2)$, $OPT(I_0) + COST(x', y') \geq OPT(I_2)$ and $OPT(I_0) \geq OPT(I_1) + kf$ still hold. Moreover, $COST(x', y') \leq \beta OPT(I_1)$,

We will prove

$$COST(x^*, y^*) \leq COST(x', y') + COST(x, y) + \delta kf.$$

The proof is similar to that in proof of Theorem 4. We first show that we can obtain an integer solution for I_2 with the total cost at most $COST(x, y) + \delta kf$ in Step 4. Note that $|\{i \mid 0 < y_i < 1\}| \leq \delta k + 1$, since Lemma 8 still holds when $\sum_{i \in F} y_i = k$. So, at most $(\delta + 1)k$ facilities are opened at the end, since if $|\{i \mid 0 < y_i < 1\}| > 0$, then $|\{i \mid y_i = 1\}| \leq k - 1$. Thus, the total cost of the obtained solution for I_2 is at most $COST(x, y) + \delta kf$.

Then, by moving the demand of t_r back to all clients in $T_r = \{j \in D' \mid x'_{rj} = 1\}$, we can construct an integer solution for I_0. This operation increases at most $COST(x', y')$ cost as $COST(x', y') = \sum_{r=1}^{\delta k} \sum_{j \in D'} d_j c_{r,j} x'_{r,j}$. Therefore, the solution obtained by Step 4 has $COST(x^*, y^*) \leq COST(x', y') + COST(x, y) + \delta kf$.

Then, we have

$$COST(x^*, y^*) \leq COST(x', y') + OPT(I_2) + \delta kf$$
$$\leq COST(x', y') + OPT(I_0) + COST(x', y') + \delta kf$$
$$= OPT(I_0) + 2COST(x', y') + \delta kf$$
$$\leq OPT(I_0) + 2\beta OPT(I_1) + \delta kf$$
$$\leq OPT(I_0) + 2\beta (OPT(I_0) - kf) + \delta kf$$
$$\leq (1 + 2\beta)(OPT(I_0) - kf) + (\delta + 1)kf.$$

Recall that we assume that exactly k facilities are opened in the optimal solution to I_0. So, the total service cost of the optimal solution to I_0 is $OPT(I_0) - kf$. Then, $COST(x^*, y^*) \leq \max\{2\beta + 1, \delta + 1\}OPT(I_0)$. □

Theorem 6 (Charikar & Guha, 1999). *Let SOL be any solution to the uncapacitated facility location problem (possibly fractional), with facility cost F_{SOL} and service cost C_{SOL}. For any $\gamma > 0$, the local search heuristic proposed (together with scaling) gives a solution with facility cost at most $(1 + \frac{2}{\gamma})F_{SOL}$ and service cost at most $(1 + \gamma)C_{SOL}$. The approximation is up to multiplicative factors of $(1 + \epsilon)$ for arbitrarily small $\epsilon > 0$.*

Based on Theorem 6, we can obtain the following corollary.

Corollary 2. *For any $\gamma > 0$, there exists a $((1+\epsilon)(1+\gamma), (1+\epsilon)(1+\frac{2}{\gamma}))$-approximation algorithm for the k-median problem, where $\epsilon > 0$ can be arbitrarily small.*

Proof. Let (x', y') be an optimal solution with total cost T to the LP relaxation of UKM. A crucial observation is that (x', y') is also a feasible fractional solution to the corresponding UFL with uniform opening cost $f > 0$. Let $SOL = (x', y')$ with total facility cost F_{SOL} and total service cost C_{SOL}. Note that $C_{SOL} = T$ and $F_{SOL} \leq kf$.

Now, it is easy to see that, based on the Charikar and Guha algorithm (Charikar & Guha, 1999), we could get an integer solution with at most $(1+\gamma)(1+\epsilon)$ times the optimal cost while using at most $(1+\frac{2}{\gamma})(1+\epsilon)k$ facilities for the k-median problem. That is, there exists a $((1+\epsilon)(1+\gamma), (1+\epsilon)(1+\frac{2}{\gamma}))$-approximation algorithm for the k-median problem, where $\epsilon > 0$ can be arbitrarily small. $\square$

The following theorem can be obtained by combining Corollary 2 with Theorem 5 and setting $\gamma = 0.78078$.

Theorem 7. *There is a $(4.562 + \epsilon)$-approximation algorithm for the capacitated facility location problem with uniform opening costs, where $\epsilon > 0$ can be arbitrarily small.*

Acknowledgment

References

Aggarwal, A., Louis, A., Bansal, M., Garg, N., Gupta, N., Gupta, S., et al. (2013). A 3-approximation algorithm for the facility location problem with uniform capacities. *Mathematical Programming, 141*(1-2), 527–547.

Archer, A., Rajagopalan, R., & Shmoys, D. B. (2003). Lagrangian relaxation for the *k*-median problem: New insights and continuity properties. In G. D. Battista, & U. Zwick (Eds.), *Algorithms—ESA 2003, Proceedings of the 11th annual European symposium, Budapest, Hungary, September, 2003*. Lecture Notes in Computer Science, Vol. 2832 (pp. 31–42). Springer.

Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., & Pandit, V. (2001). Local search heuristic for *k*-median and facility location problems. In: *Proceedings of 33th annual ACM symposium on theory of computing* (pp. 11–20). Heraklion, Crete, Greece: ACM.

Bansal, M., Garg, N., & Gupta, N. (2012). A 5-approximation for capacitated facility location. In L. Epstein, & P. Ferragina (Eds.), *Algorithms—ESA 2012, Proceedings of the 20th annual European symposium, Ljubljana, Slovenia, September, 2012*. Lecture Notes in Computer Science, Vol. 7501 (pp. 133–144). Springer.

Bartal, Y., Charikar, M., & Raz, D. (2001). Approximating min-sum *k*-clustering in metric spaces. In: *Proceedings of 33th annual ACM symposium on theory of computing* (pp. 11–20). Heraklion, Crete, Greece: ACM.

Bradley, P. S., Fayyad, U. M., & Mangasarian, O. L. (1999). Mathematical programming for data mining: Formulations and challenges. *INFORMS Journal on Computing, 11*(3), 217–238.

Byrka, J., Fleszar, K., Rybicki, B., & Spoerhase, J. (2013). Bi-factor approximation algorithms for hard capacitated *k*-median problems. *CoRR, abs/1312.6550*.

Byrka, J., Pensyl, T., Rybicki, B., Srinivasan, A., & Trinh, K. (2014). An improved approximation for *k*-median, and positive correlation in budgeted optimization. *CoRR, abs/1406.2951*.

Caprara, A., Kellerer, H., Pferschy, U., & Pisinger, D. (2000). Approximation algorithms for knapsack problems with cardinality constraints. *European Journal of Operational Research, 123*(2), 333–345.

Charikar, M., & Guha, S. (1999). Improved combinatorial algorithms for the facility location and *k*-median problems. In: *Proceedings of 40th annual symposium on foundations of computer science* (pp. 378–388). New York, NY: IEEE Computer Society.

Charikar, M., Guha, S., Tardos, É., & Shmoys, D. B. (1999). A constant-factor approximation algorithm for the *k*-median problem (extended abstract). In: *Proceedings of the 31th annual ACM symposium on theory of computing* (pp. 1–10). Atlanta, GA: ACM.

Chudak, F. A., & Williamson, D. P. (2005). Improved approximation algorithms for capacitated facility location problems. *Mathematical Programming, 102*(2), 207–222.

Chuzhoy, J., & Rabani, Y. (2005). Approximating *k*-median with non-uniform capacities. In: *Proceedings of the 16th annual ACM-SIAM symposium on discrete algorithms* (pp. 952–958). Vancouver, British Columbia, Canada: SIAM.

Cornuéjols, G., Nemhauser, G. L., & Wolsey, L. A. (1990). *Discrete location theory*. chapter The uncapacitated facility location problem. (pp. 119–171). New York: Wiley.

Edwards, N. J. (2001). Approximation algorithms for the multi-level facility location problem (Ph.D. thesis). Cornell University.

Gabow, H. N. (1976). An efficient implementation of Edmonds' algorithm for maximum matching on graphs. *Journal of the ACM, 23*(2), 221–234.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. New York, NY: W. H. Freeman & Co.

Görtz, S., & Klose, A. (2009). Analysis of some greedy algorithms for the single-sink fixed-charge transportation problem. *Journal of Heuristics, 15*(4), 331–349.

Grötschel, M., Lovász, L., & Schrijver, A. (1988). *Geometric algorithms and combinatorial optimization*. Berlin: Springer.

Herer, Y. T., Rosenblatt, M. J., & Hefter, I. (1996). Fast algorithms for single-sink fixed charge transportation problems with applications to manufacturing and transportation. *Transportation Science, 30*(4), 276–290.

Hoffman, A. J., & Kruskal, J. B. (1956). *Linear inequalities and related systems*. chapter Integral boundary points of convex polyhedra. pp. 233–246. Princeton, NJ: Princeton Univ. Press.

Hsu, V. N., Lowe, T. J., & Tamir, A. (1997). Structured *p*-facility location problems on the line solvable in polynomial time. *Operations Research Letters, 21*(4), 159–164.

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for clustering data*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Jain, K., Mahdian, M., Markakis, E., Saberi, A., & Vazirani, V. V. (2003). Greedy facility location algorithms analyzed using dual fitting with factor-revealing LP. *Journal of the ACM, 50*(6), 795–824.

Jain, K., Mahdian, M., & Saberi, A. (2002). A new greedy approach for facility location problems. In: *Proceedings of the 34th annual ACM symposium on theory of computing* (pp. 731–740). Montréal, Québec, Canada: ACM.

Jain, K., & Vazirani, V. V. (2001). 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.

Kapoor, S., & Ramesh, H. (1995). Algorithms for enumerating all spanning trees of undirected and weighted graphs. *SIAM Journal on Computing, 24*(2), 247–265.

Korupolu, M. R., Plaxton, C. G., & Rajaraman, R. (2000). Analysis of a local search heuristic for facility location problems. *Journal of Algorithms, 37*(1), 146–188.

Kuehn, A. A., & Hamburger, M. J. (1963). A heuristic program for locating warehouses. *Management Science, 9*(4), 643–666.

Lawler, E. L. (1979). Fast approximation algorithms for knapsack problems. *Mathematical Methods of Operations Research, 4*(4), 339–356.

Levi, R., Shmoys, D. B., & Swamy, C. (2012). LP-based approximation algorithms for capacitated facility location. *Mathematical Programming, 131*(1-2), 365–379.

Li, S. (2014). An improved approximation algorithm for the hard uniform capacitated *k*-median problem. In K. Jansen, J. D. P. Rolim, N. R. Devanur, & C. Moore (Eds.), *Approximation, randomization, and combinatorial optimization. algorithms and techniques (approx/random 2014)*. Leibniz International Proceedings in Informatics (LIPIcs), Vol. 28 (pp. 325–338). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Li, S., & Svensson, O. (2013). Approximating *k*-median via pseudo-approximation. In *Proceedings of the 2013 ACM symposium on theory of computing* (pp. 901–910). Palo Alto, CA: ACM.

Mahdian, M., & Pál, M. (2003). Universal facility location. In G. D. Battista, & U. Zwick (Eds.), *Algorithms—ESA 2003, Proceedings of the 11th annual European symposium, Budapest, Hungary, September 2003* Lecture Notes in Computer Science, Vol. 2832. (pp. 409–421). Springer.

Mahdian, M., Ye, Y., & Zhang, J. (2003). A 2-approximation algorithm for the soft-capacitated facility location problem. In S. Arora, K. Jansen, J. D. P. Rolim, & A. Sahai (Eds.), *Approximation, randomization, and combinatorial optimization: Algorithms and techniques, Proceedings of the 6th international workshop on approximation algorithms for combinatorial optimization problems, APPROX 2003 and 7th international workshop on randomization and approximation techniques in computer science, RANDOM 2003, Princeton, NJ, USA, August 24–26, 2003* Lecture Notes in Computer Science, Vol. 2764. (pp. 129–140). Springer.

Pál, M., Tardos, É., & Wexler, T. (2001). Facility location with nonuniform hard capacities. In: *Proceedings of 42th annual symposium on foundations of computer science* (pp. 329–338). Las Vegas, NV: IEEE Computer Society.

Schrijver, A. (2003). *Combinatorial optimization: Polyhedra and efficiency*. Berlin: Springer-Verlag.

Scoins, H. I. (1962). The number of trees with nodes of alternate parity. *Mathematical Proceedings of the Cambridge Philosophical Society, 58*(1), 12–16.

Shmoys, D. B., Tardos, É., & Aardal, K. (1997). Approximation algorithms for facility location problems (extended abstract). In *Proceedings of the 29th annual ACM symposium on the theory of computing* (pp. 265–274). El Paso, TX: ACM.

Zhang, J., Chen, B., & Ye, Y. (2005). A multiexchange local search algorithm for the capacitated facility location problem. *Mathematical Methods of Operations Research, 30*(2), 389–403.

Zhang, P. (2007). A new approximation algorithm for the *k*-facility location problem. *Theoretical Computer Science, 384*(1), 126–135.