

Approximation Algorithms for a Combined Facility Location Buy-at-Bulk Network Design Problem

A. Bley^{1*}, S. M. Hashemi², and M. Rezapour^{1**}

¹ Institute for Mathematics, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, {bley,rezapour}@math.tu-berlin.de

² Department of Computer Science, Amirkabir University of Technology, No. 424, Hafez Ave., Tehran, Iran, hashemi@aut.ac.ir

Preprint 2012/28

Revised version, March 22, 2013

Abstract We consider a generalization of the connected facility location problem where the clients must be connected to the open facilities via shared capacitated (tree) networks instead of independent shortest paths. This problem arises in the planning of fiber optic telecommunication access networks, for example. Given a set of clients with positive demands, a set of potential facilities with opening costs, a set of capacitated access cable types, and a core cable type of infinite capacity, one has to decide which facilities to open, how to interconnect them using a Steiner tree of infinite capacity core cables, and which access cable types to install on which potential edges such that these edges form a forest and the installed capacities suffice to simultaneously route the client demands to the open facilities via single paths. The objective is to minimize the total cost of opening facilities, building the core Steiner tree among them, and installing the access cables. In this paper, we devise a constant-factor approximation algorithm for problem instances where the access cable types obey economies of scale. In the special case where only multiples of a single cable type can be installed on the access edges, a variant of our algorithm achieves a performance guarantee of 6.72.

1 Introduction

We study a generalization of the *Connected Facility Location* (ConFL) problem where not only direct connections between clients and open facilities, but also shared access trees connecting multiple clients to an open facility are allowed. Accordingly, also more realistic capacity and cost structures with flow-dependent buy-at-bulk costs for the access edges are considered. The resulting *Connected*

* Supported by the DFG research center MATHEON ‘Mathematics for key technologies’.

** Supported by the DFG research training group ‘Methods for Discrete Structures’.

Facility Location with Buy-at-Bulk edge costs (BBConFL) problem captures the central aspects of both the buy-at-bulk network design problem and the ConFL problem. In this paper, we study the approximability of the BBConFL problem. Although both the ConFL and the buy-at-bulk network design problem have been well studied in the past, the combination of them has not been considered in the literature, to the best of our knowledge.

A typical telecommunication network consists of a backbone network with (almost) unlimited capacity on the links and several local access networks. In such a network, the traffic originating from the clients is sent through access networks to gateway or core nodes, which provide routing functionalities and access to the backbone network. The backbone then provides the connectivity among the core nodes, which is necessary to route the traffic further towards its destination. Designing such a network involves selecting the core nodes, connecting them with each other, and choosing and dimensioning the links that are used to route the traffic from the clients to the selected core nodes.

We model this planning problem as the BBConFL problem. We are given an undirected graph $G = (V, E)$ with nonnegative edge lengths $c_e \in \mathbb{Z}_{\geq 0}$, $e \in E$, obeying the triangle inequality, a set $F \subset V$ of potential facilities with opening costs $f_i \in \mathbb{Z}_{\geq 0}$, $i \in F$, and a set of clients $D \subset V$ with demands $d_j \in \mathbb{Z}_{> 0}$, $j \in D$. We are also given K types of access cables that may be used to connect clients to open facilities. A cable of type i has capacity $u_i \in \mathbb{Z}_{> 0}$ and cost (per unit length) $\sigma_i \in \mathbb{Z}_{\geq 0}$. Core cables, which are used to connect the open facilities, have a cost (per unit length) of $M \in \mathbb{Z}_{\geq 0}$ and infinite capacity. The task is to find a subset $I \subseteq F$ of facilities to open, a Steiner tree $S \subseteq E$ connecting the open facilities, and a forest $E' \subseteq E$ with a cable installation on its edges, such that E' connects each client to exactly one open facility and the installed capacities suffice to route all clients' demands to the open facilities. We are allowed to install multiple copies and types of access cables on each edge of E' . The objective is to minimize the total cost, where the cost for using edge e in the core Steiner tree is Mc_e and the cost for installing a single access cable of type i on edge e is $\sigma_i c_e$. We also consider the variant with only a single cable type, which we denote by *Single-Cable Connected facility location problem* (Single-Cable-ConFL).

The classical ConFL problem is special case of the BBConFL problem with only one cable type of unit capacity. This problem is well-studied in the literature. Gupta et al. [10] obtain a 10.66-approximation for this problem, based on LP rounding. Swamy and Kumar [15] later improved the approximation ratio to 8.55, using a primal-dual algorithm. Using sampling techniques, the guarantee was later reduced to 4 by Eisenbrand et al. [4], and to 3.19 by Grandoni et al. [7].

The (unsplittable) *Single-Sink Buy-at-Bulk problem (u-SSBB)* can be viewed as a special case of the BBConFL problem where the set of interconnected facilities are given in advance. Several approximation algorithms for u-SSBB have been proposed in the literature. Using LP rounding techniques, Garg et al. [5] developed a $O(k)$ approximation, where k is the number of cable types. Hassin et al. [11] provide a constant factor approximation for the single cable version of the problem. The first constant factor approximation for the problem with

multiple cable types is due to Guha et al. [9]. Talwar [16] showed that an IP formulation of this problem has a constant integrality gap and provided a factor 216 approximation algorithm. Using sampling techniques, the approximation was reduced to 145.6 by Jothi et al. [12], and later to 40.82 by Grandoni et al. [6].

If we omit the requirement to connect the open facilities by a core Steiner tree, then the BBConFL problem reduces to a *k-cable facility location problem*. For this problem, Ravi et al. [13] provide an $O(k)$ approximation algorithm.

The rest of the paper is structured as follows. In Section 2, we describe our constant factor approximation algorithm for BBConFL extending the algorithm of Guha et al. [9] to incorporate also the selection of facilities to open as well as the Steiner tree (of infinite capacity) interconnecting them. In Section 3, we study the single cable version of the problem and present a factor 6.72 approximation algorithm for this problem.

2 Approximating BBConFL

In this section, we present a constant factor approximation algorithm for the BBConFL, which uses the ideas of Guha's algorithm [9] for the single sink buy-at-bulk network design problem to design the access trees of the solution.

First, we define another problem similar to the BBConFL with slightly different cost function, called *modified-BBConFL*. In this problem, each access cable has a fixed cost of σ_i , a flow dependent incremental cost of $\delta_i = \frac{\sigma_i}{u_i}$, and unbounded capacity. That is, for using one copy of cable type i on edge e and transporting D flow unit on e , a cost of $(\sigma_i + D\delta_i)c_e$ is incurred.

It is not hard to see that any ρ -approximation to the modified problem gives a 2ρ -approximation to the corresponding original buy-at-bulk ConFL. Furthermore, we will show later that there exist near optimal solutions of the modified problem that have a nice tree-like structure with each cable type being installed in a corresponding layer. We will exploit this special structure in our algorithm to compute approximate solutions for the modified problem and, thereby, also approximate solutions for the original buy-at-bulk ConFL.

In the modified-BBConFL, we may assume w.l.o.g. that $\sigma_1 < \dots < \sigma_K$ and that $\delta_1 > \dots > \delta_K$. In addition, we assume that $2\sigma_K < M$. Note that in our and many other applications, it is natural to assume that $\sigma_K \ll M$.

First, we prune the set of cable types such that all cables are considerably different. As shown in [9], this can be done without increasing the cost of the optimal solution too much.

Theorem 1. *For a predefined constant $\alpha \in (0, \frac{1}{2})$, we can prune the set of cables such that, for any i , we have $\sigma_{i+1} > \frac{1}{\alpha} \cdot \sigma_i$ and $\delta_{i+1} < \alpha \cdot \delta_i$ hold and the cost of the optimal solution increases by at most $\frac{1}{\alpha}$.*

We observe that, as demand along an edge increases, there are break-points at which it becomes cheaper to use the next larger cable type. For $1 \leq i < K$, we define b_i such that $\sigma_{i+1} + b_i\delta_{i+1} = 2\alpha(\sigma_i + b_i\delta_i)$. Intuitively, b_i is the demand at which it becomes considerably cheaper to use a cable type $i + 1$ rather than a

cable type i . It has been shown in [9] that the break-points and modified cable cost functions satisfy the following properties.

Lemma 2. *For all i , we have $u_i \leq b_i \leq u_{i+1}$. For any i and $D \geq b_i$, we have $\sigma_{i+1} + D\delta_{i+1} \leq 2 \cdot \alpha(\sigma_i + D\delta_i)$.*

Let $b_K = \frac{M - \sigma_K}{\delta_K}$ be the edge flow at which the cost of using cable type K and a core link are the same. Suppose we install cable type i whenever the edge flow is in the range $[b_{i-1}, b_i]$, $1 \leq i \leq K$, where $b_0 = 0$. It can be shown that, if the edge flow is in the range $[b_{i-1}, u_i]$, then considering only the fixed cost σ_i (times the edge length) for using cable type i on the edge and ignoring the flow dependent incremental cost will underestimate the true edge cost only by a factor 2. Similarly, if the edge flow is in $[u_i, b_i]$, then considering only the flow dependent cost δ_i times the flow and ignoring the fixed cost underestimates the cost by only a factor 2. This means that any solution can be converted to a layered solution, loosing at most a factor 2 in cost, where layer i consists of (i) a Steiner forest using cable type i and carrying a flow of at least b_{i-1} on each edge, and (ii) a shortest path forest with each edge carrying a flow of at least u_i . In the following theorem, we define the structural properties of such layered solutions more formally. As in [9], for sake of simplicity, we assume that there are extra loop-edges such that property (iii) can be enforced for any solution.

Theorem 3. *Modified-BBConFL has a solution with the following properties:*

- (i) *The incoming demand of each open facility is at least b_K .*
- (ii) *Cable $i + 1$ is used on edge e only if at least b_i demand is routed across e .*
- (iii) *All demand which enters a node, except an open facility, using cable i , leaves that node using cables i or $i + 1$.*
- (iv) *The solution's cost is at most $2(\frac{1}{\alpha} + 1)$ times the optimum cost.*

Proof. Consider an optimum solution of the modified-BBConFL. Let T^* be the tree connecting the open facilities in the optimum solution. Consider those open facilities whose incoming demand is less than b_K . We can find an unsplittable flow on the edges of T^* sending the aggregated demand from these facilities to some other open facilities such that the resulting solution obeys property (i) and the total flow on any edge of the Steiner tree is at most b_K . Therefore the cost of closing these facilities and sending the corresponding demands to some other open facility using access links can be bounded by the core Steiner tree cost of the optimal solution, so we close these facilities and reroute demands. Now identify the set of remaining open facilities to a single sink, and update the edge length metric appropriately. The resulting solution is now a (possibly sub-optimal) single-sink network design solution. Results in [9] imply that there is a near-optimal solution to this single-sink instance which obeys the properties (ii) and (iii), with a factor $(\frac{2}{\alpha} + 1)$ loss in the total access cable cost. Hence, we can transform our modified-BBConFL solution to a solution which satisfies properties (ii)–(iv), too. \square

Our algorithm constructs a layered solution with the properties described in Theorem 3 in a bottom-up fashion, aggregating the client demands repeatedly

and alternating via Steiner trees and via direct assignments (or, equivalently, via shortest path trees) to values exceeding u_i and b_i . In phase i , we first aggregate the (already pre-aggregated) demands of value at least b_{i-1} to values of at least u_i using cable type i on the edges of an (approximate) Steiner tree connecting these demands. Then we further aggregate these aggregates to values of at least (a constant fraction of) b_i solving a corresponding *Lower Bounded Facility Location* (LBFL) problem [8,14,1], where all clients may serve as facilities to aggregate demand (except for the last phase, where only real facilities are eligible). The LBFL problem is a generalization of the facility location problem where each open facility is required to serve a certain minimum amount of demand.

Let D_i be the set of demand points we have at the i -th stage. Initially $D_1 = D$. Algorithm 1 describes the steps of the algorithm in more detail.

Algorithm 1

1. Guess a facility r from the optimum solution.
 2. For cable type $i = 1, 2, \dots, K - 1$ Do
 - *Steiner Trees*: Construct a ρ_{ST} -approximate Steiner tree T_i on terminals $D_i \cup \{r\}$ for edge costs σ_i per unit length. Install a cable of type i on each edge of this tree. Root this tree at r . Transport the demands from D_i upwards along the tree. Walking upwards along this tree, identify edges whose demand is larger than u_i and cut the tree at these edges.
 - *Consolidate*: For every tree in the forest created in the preceding step, transfer the total demand in the root of tree, which is at least u_i , back to one of its sources using a shortest path of cable type i . Choose this source with probability proportional to the demand at the source.
 - *Shortest Path*: Solve the LBFL problem with clients D_1 , facility opening cost 0 at all nodes, facility lower bound b_i , and edge costs δ_i per unit length. The solution is a forest of shortest path trees. Then route the current demands along these trees to their roots, installing cables of type i .
 - *Consolidate*: For every root in the forest created in the preceding step, transfer the total demand in the root of tree, which is at least b_i , back to one of its sources with probability proportional to the demand at that source using a shortest path with cables of type i . Let D_{i+1} be the resulting demand locations.
 3. For cable type K Do
 - Construct a ρ_{ST} -approximate Steiner tree T_K on terminals $D_K \cup \{r\}$ for edge costs σ_K per unit length. Install a cable of type K on each edge of this tree. Root this tree at r . Transport the demands from D_K upwards along the tree. Walking along this tree, identify edges whose demand is larger than u_K and cut the tree at these edges. For every tree in the created forest, transfer the total demand in the root of tree back to one of its sources with probability proportional to the demand at that source via a shortest path, using cables of type K .
 - Solve the LBFL problem with clients D_1 , facility set F , opening costs f_i , facility lower bound b_K , and edge costs δ_K per unit length. We obtain a forest of shortest path trees. Then route the current demands along these trees to their roots, installing cables of type K . Let F' be the set of open facilities.
 4. Compute a ρ_{ST} -approximate Steiner tree T_{core} on terminals $F' \cup \{r\}$ for edge costs M per unit length. Install the core link on the edges of T_{core} .
-

To solve LBFL, we employ the bicriteria $\mu\rho_{FL}$ -approximation algorithm devised by Guha et al. [8], which relaxes the lower bound on the minimum demand served by a facility by a factor $\beta = \frac{\mu-1}{\mu+1}$. Here ρ_{FL} is the best known approximation for the facility location problem.

It remains to show that the computed solution is an approximate solution. Let C_i^* , S^* , and O^* be the amount paid for cables of type i , for the core Steiner tree, and for opening facilities in the near-optimal solution, respectively. We define C_i to be the total cost paid for cables of type i in the returned solution. Let D_j^i be the demand of node j at stage i of the algorithm. Let T_i , P_i and N_i be cost incurred in the Steiner tree step, the shortest path step, and the consolidation steps of iteration i , respectively. Also, let T_i^I and T_i^F denote the incremental and the fixed cost components of the Steiner tree step at iteration i . Analogously, P_i^I and P_i^F denote the incremental and the fixed costs incurred in the shortest path step. Recall that the set of access cable types has been reduced depending on the constant parameter $\alpha \in (0, \frac{1}{2})$. How to choose this parameter appropriately will be discussed later.

The following Lemma carries over from the single sink buy-at-bulk problem studied in [9] to our problem in a straightforward way.

Lemma 4.

- (i) *At the end of each consolidation step, every node has $E[D_j^i] = d_j$.*
- (ii) *$E[N_i] \leq T_i + P_i$ for each i .*
- (iii) *$P_i^F \leq P_i^I$ and $T_i^I \leq T_i^F$ for each i .*

The following lemma bounds the fixed costs of the cables installed in the Steiner tree phase i of our algorithm.

Lemma 5. *$E[T_i^F] \leq \rho_{ST} \left(\sum_{j=1}^{i-1} \frac{1}{\beta} (2\alpha)^{i-j} C_j^* + \sum_{j=i}^K \alpha^{j-i} C_j^* + \frac{1}{2} \alpha^{K-i} S^* \right)$ for each i .*

Proof. We construct a feasible Steiner tree for stage i as follows. Consider the near-optimum solution, and consider only those nodes which are candidate terminals in stage i of our algorithm. We remove all the cables if the total demand flowing across it is zero. Otherwise we replace the cable with a cable of type i . Note that, being in stage i , the expected demand on each cable $j < i$ is at least βb_j . Hence, by Lemma 2, the expected cost of all replacement cables for cables of type $j < i$ is bounded by $\frac{1}{\beta} (2\alpha)^{i-j} C_j^*$.

Similarly, the expected cost of the replacement cables for the cables $j > i$ are bounded by $\alpha^{j-i} C_j^*$, using the fixed costs scale. Finally, the cost on a core link used to connect candidate terminals to r is reduced at least by $\frac{1}{2} \alpha^{K-i} S^*$. Altogether, the expected fixed cost of this Steiner tree, which is a possible solution to the Steiner tree problem in stage i , is bounded by

$$\sum_{j=1}^{i-1} \frac{1}{\beta} (2\alpha)^{i-j} C_j^* + \sum_{j=i}^K \alpha^{j-i} C_j^* + \frac{1}{2} \alpha^{K-i} S^* .$$

As we use a ρ_{ST} -approximation algorithm to solve this Steiner tree problem in our algorithm, the claim follows. \square

In a similar way, we can also bound the incremental costs of the cables installed in the shortest path phase i of our algorithm.

Lemma 6. $E[P_i^I] \leq \mu \cdot \rho_{FL} \sum_{j=1}^i \alpha^{i-j} \cdot C_j^*$ for each i .

Proof. Consider the forest defined by the edges with cable types 1 to i in the near-optimum solution and replace all cables of type less than i by cables of type i . The cost of replacing all cables of type $j < i$ is bounded by $\alpha^{i-j} \cdot C_j^*$, using the incremental costs scale. The resulting tree provides a feasible solution for the shortest path stage i . As our algorithm applies a bicriteria $\mu \cdot \rho_{FL}$ -approximation algorithm to solve the lower bounded facility location problem in this stage, the claim follows. \square

The opening costs and the incremental shortest path costs in the final stage of our algorithm can be bounded as follows.

Lemma 7. $E[P_K^I + f(F')] \leq \mu \cdot \rho_{FL} (\sum_{i=1}^K \alpha^{K-i} \cdot C_i^* + O^*)$

Proof. Now, consider the forest given by all access edges of the near-optimum solution and replace all cables (of type less than K) by cables of type K . For each $i < K$, the incremental cost of the new solution is a fraction α^{K-i} of the incremental cost of the optimal solution's cable i portion. The set of facilities opened in the solution, combined with the cables, constitutes a feasible solution for the LBFL problem solved in the final stage, and its cost is no more than $\sum_{i=1}^K \alpha^{K-i} C_i^* + O^*$. Using the bicriteria $\mu \cdot \rho_{FL}$ -approximation algorithm, the claim follows. \square

Finally, the cost of the core Steiner tree have to be bounded.

Lemma 8. $E[T_{core}] \leq \rho_{ST} (S^* + \frac{1}{\beta} \sum_{j=1}^K (C_j^* + C_j))$

Proof. Let F^* , T_{core}^* and T_{access}^* be the set of open facilities, the tree connecting them, and the forest connecting clients to open facilities in the near-optimum solution, respectively. Let T_{access} be the forest connecting clients to open facilities in the solution returned by the algorithm. We construct a feasible Steiner tree on $F' \cup \{r\}$, whose expected cost is $S^* + \frac{1}{\beta} \sum_{j=1}^K (C_j^* + C_j)$. In the algorithm's solution, each facility $l \in F'$ serves at least a total demand of βb_K . This demand is also served by the set of optimal facilities in the near-optimum solution. Therefore, at least βb_K demand can be routed between each facility $l \in F'$ and the facilities of F^* along edges of $T_{access}^* \cup T_{access}$ (using the access links). Hence, we obtain a feasible Steiner tree on $F' \cup F^*$, using core links, whose cost is at most $S^* + \frac{1}{\beta} \sum_{j=1}^K (C_j^* + C_j)$. \square

Together, Lemmas 4–8 imply our main result.

Theorem 9. *Algorithm 1 is a constant factor approximation for BBConFL.*

Proof. By Lemmas 4–6, the total expected cost of access links is bounded by

$$\begin{aligned} 4 \sum_{i=1}^K \left[\mu \rho_{FL} \sum_{j=1}^i \alpha^{i-j} C_j^* + \rho_{ST} \left(\sum_{j=i}^K \alpha^{j-i} C_j^* + \sum_{j=1}^{i-1} \frac{1}{\beta} (2\alpha)^{i-j} C_j^* + \frac{1}{2} \alpha^{K-i} S^* \right) \right] \\ \leq 4 \left(\frac{\mu \cdot \rho_{FL}}{1-\alpha} + \frac{\rho_{ST}}{1-\alpha} + \frac{\rho_{ST}}{\beta(1-2\alpha)} \right) \sum_{i=1}^K C_i^* + \frac{2 \cdot \rho_{ST}}{1-\alpha} S^* \end{aligned}$$

Additionally, using Lemmas 7 and 8, the total cost of installing core links and opening facilities is bounded by

$$\mu \rho_{FL} O^* + \rho_{ST} S^* + \frac{\rho_{ST}}{\beta} \left(\sum_{i=1}^K C_i^* + \sum_{i=1}^K C_i \right).$$

Altogether, we obtain a bound of

$$\begin{aligned} \mu \rho_{FL} O^* + \left[\frac{\rho_{ST}}{\beta} + 4 \left(1 + \frac{\rho_{ST}}{\beta} \right) \left(\frac{\mu \rho_{FL} + \rho_{ST}}{1-\alpha} + \frac{\rho_{ST}}{\beta(1-2\alpha)} \right) \right] \sum_{i=1}^K C_i^* \\ + \left[\left(1 + \frac{\rho_{ST}}{\beta} \right) \left(\frac{2\rho_{ST}}{1-\alpha} \right) + \rho_{ST} \right] S^* \end{aligned}$$

for the worst case ratio between the algorithm's solution and a near optimal solution, restricted according to Theorem 3, of the modified-BBConFL. With Theorems 1 and 3, this yields a worst case approximation guarantee of $\frac{2}{\alpha} \left(\frac{1}{\alpha} + 1 \right)$ times the above ratio against an unrestricted optimal solution of the modified-BBConFL.

Finally, we lose another factor of 2 in the approximation guarantee when evaluating the approximate solution for the modified-BBConFL with respect to the original BBConFL problem. For appropriately chosen fixed parameters α , β , and μ , we nevertheless obtain a constant factor approximation algorithm for BBConFL. \square

3 Approximating Single-Cable-ConFL

In this section, we consider a simpler version of the problem, where only multiples of a single cable type can be installed. Let $u > 0$ be the capacity of the only cable type available. We may assume that the cost of this cable is one. The algorithm presented in this paper can easily be adapted for $\sigma > 1$.

We obtain an approximation algorithm for this problem by modifying the algorithmic framework proposed in [7] as shown in Algorithm 2 on the next page. In this Algorithm, $c(v, u)$ denotes the distance between u and v , and $c(v, U) = \min_{u \in U} c(v, u)$. Again, the algorithm uses a constant parameter $\alpha \in (0, 1]$, whose setting will be discussed later.

One easily verifies that Algorithm 2 computes a feasible solution. Clearly, T' is a Steiner tree connecting the open facilities F' . The existence of (and a

Algorithm 2

1. Guess a facility r from the optimum solution.
Mark each client $j \in D$ with probability $\frac{\alpha \cdot d_j}{M \cdot u}$. Let D' be the set of marked clients.
 2. Compute a ρ_{ST} -approximate Steiner tree T_1 on terminals $D' \cup \{r\}$.
 3. Define a FL instance with clients D , facilities F , costs $c'_{ij} := \frac{d_j}{u} c(i, j)$, $j \in D$ and $i \in F$, and opening costs $f'_i := f_i + M \cdot c(i, D' \cup \{r\})$, $i \in F$.
Compute a (λ_F, λ_C) -bifactor-approximate solution $U = (F', \sigma)$ to this instance, where $\sigma(j) \in F'$ indicates the facility serving $j \in D$ in U .
 4. Augment T_1 with shortest paths from each $i \in F'$ to T_1 .
Let T' be the augmented tree.
Output F' and T' as open facilities and core Steiner tree, respectively.
 5. Compute a ρ_{ST} -approximate Steiner tree T_2 on terminals $D \cup \{r\}$.
 6. // Using the results in [11,13], we now install capacities to route the clients' demands to open facilities F' .
 - For each $j \in D$ with $d_j > u/2$, install $\lceil d_j/u \rceil$ cables from j to its closest open facility in F' .
 - Considering only clients with $d_j \leq u/2$, partition T_2 into disjoint subtrees such that the total demand of each subtree not containing r is in $[u/2, u]$ and the total demand of the subtree containing r is at most u ; see [11].
 - Install one cable on each edge contained in any subtree.
 - For each subtree not containing r , install one cable from the client closest to an open facility to this facility.
-

polynomial time algorithm to find) a partition of the tree T_2 into subtrees of total demand between $u/2$ and u each, except for the subtree containing r , has been shown in [11], given that each individual demand is at most u . From that, it follows immediately that all clients j with $d_j \leq u/2$ can be routed within their respective subtree towards the client closest to an open facility and then further on to this facility without exceeding the capacity u on these edges.

It remains to show that the computed solution is an approximate solution. Let O'_U and C'_U be the (modified) opening and connection costs of the solution U of the facility location problem solved in Step 3. Furthermore, let I^* , S^* , and F^* be the set of open facilities, the Steiner tree connecting them, and the forest connecting the clients to the open facilities in the optimal solution, respectively. Also let $\sigma^*(j) \in I^*$ be the facility serving $j \in D$ in the optimal solution. The opening costs, cable installation costs, and core Steiner tree costs of the algorithm's solution and of the optimal solution are denoted by O, C, T and O^*, C^*, T^* , respectively. Let $c(E') := \sum_{e \in E'} c_e$ for any $E' \subseteq E$,

Lemma 10. *The cable cost induced in Step 6 is at most $c(T_2) + 2 \cdot C'_U$.*

Proof. Using the result in [11], the total flow on any edge of the Steiner tree T_2 induced by grouping the demands into disjoint subtrees is at most u . Thus, one copy of the cable on all edges in T_2 is sufficient to accommodate the flow on the edges of T_2 , which contributes $c(T_2)$ to the total cable installation cost.

Let C_1, C_2, \dots, C_T be the sets of clients in each subtree and for each C_t let $j_t \in C_t$ be the client which is closest to an open facility in F' . The modified

connection costs in U are

$$\begin{aligned} C'_U &= \sum_t \sum_{j \in C_t} \frac{d_j}{u} c(j, \sigma(j)) + \sum_{j \in D: d_j > \frac{u}{2}} \frac{d_j}{u} c(j, \sigma(j)) \\ &\geq \sum_t \sum_{j \in C_t} \frac{d_j}{u} c(j, \sigma(j)) + \sum_{j \in D: d_j > \frac{u}{2}} \frac{1}{2} c(j, \sigma(j)). \end{aligned}$$

Since the algorithm sends the total demand of C_t via j_t , we have

$$C'_U \geq \sum_t \frac{\sum_{j \in C_t} d_j}{u} c(j_t, \sigma(j_t)) + \sum_{j \in D: d_j > \frac{u}{2}} \frac{1}{2} c(j, \sigma(j)) \geq \frac{1}{2} C_{AC},$$

where C_{AC} is the cost of the cables installed by the algorithm between the subtrees and the closest open facilities and between the large demand clients and the open facilities. Altogether the total cost of buying cables to route the traffic is at most $c(T_2) + 2 \cdot C'_U$. \square

Lemma 11. *The opening and core connection cost of the computed solution satisfy $O + T \leq O'_U + M \cdot c(T_1)$.*

Proof. Algorithm 2 opens the facilities chosen in the FL solution and connects these facilities by the tree T' . Since the modified opening costs f' in Step 3 include both the original cost for opening F' and the cost for augmenting T_1 to T' , the sum of the opening cost and core connection cost of the final solution are at most $O'_U + M \cdot c(T_1)$. \square

Lemma 12. *The expected cost of T_1 is at most $\frac{\rho_{ST}}{M}(T^* + \alpha C^*)$.*

Proof. We obtain a feasible Steiner tree on $D' \cup \{r\}$ by joining the optimal solution's Steiner tree S^* and the paths connecting each client in D' with its corresponding open facility in I^* in the optimal solution. The expected cost of the resulting subgraph is at most

$$\sum_{e \in S^*} c(e) + \frac{\alpha}{M} \sum_{j \in D} \frac{d_j}{u} \cdot l(j, I^*) \leq \frac{T^*}{M} + \frac{\alpha}{M} C^*,$$

where $l(j, I^*)$ denotes the length of the path connecting j to its open facility in I^* using edges of F^* . The last inequality holds since in $\sum_{j \in D} \frac{d_j}{u} \cdot l(j, I^*)$ instead of installing an integral number of cables on every edge, we install multiples of $\frac{1}{u}$ on every edge, which is a lower bound for C^* . Thus the expected cost of the ρ_{ST} -approximate Steiner tree on $D' \cup \{r\}$ is at most $\frac{\rho_{ST}}{M}(T^* + \alpha C^*)$. \square

Lemma 13. *The cost of T_2 is at most $\rho_{ST}(T^* + C^*)$.*

Proof. Clearly $S^* \cup F^*$ defines a feasible Steiner tree on $D \cup \{r\}$. \square

Lemma 14. $E[O'_U + C'_U] \leq \lambda_F(O^* + \alpha C^*) + \lambda_C(C^* + \frac{0.807}{\alpha} T^*)$.

Proof. We provide a feasible solution for the facility location problem, whose expected opening cost is $O^* + \alpha C^*$ and whose expected connection cost is $C^* + \frac{0.807}{\alpha} T^*$. Choose facilities $\sigma^*(D') \cup \{r\}$. The expected opening cost is at most

$$\sum_{i \in I^*} f_i + M \cdot \frac{\alpha}{M} \sum_{j \in D} \frac{d_j}{u} \cdot l(j, \sigma^*(j)) \leq O^* + \alpha C^* .$$

Now, replace j by several copies of co-located unit-demand clients. In order to bound the expected connection cost, we apply the core connection game described in [4] (see also Lemma 2 in [7]) for ConFL with unit-demand clients, probability $\frac{\alpha}{M \cdot u}$ (which is the same to mark each client $j \in D$ with probability $\frac{\alpha \cdot d_j}{M \cdot u}$), core S^* , mapping $\sigma = \sigma^*$, and $w(e) = \frac{c(e)}{u}$ which yields

$$\begin{aligned} E\left[\sum_{j \in D} c'(j, \sigma^*(D') \cup \{r\})\right] &\leq \sum_{j \in D} c'(j, I^*) + \frac{0.807}{M \cdot u} \cdot \frac{w(T^*)}{M} \\ &\leq \sum_{j \in D} \frac{d_j}{u} l(j, I^*) + \frac{0.807}{M \cdot u} \cdot \frac{T^*}{M \cdot u} \leq C^* + \frac{0.807}{\alpha} T^* . \end{aligned} \quad \square$$

Theorem 15. *For a proper choice of α , Algorithm 2 is an 6.72-approximation algorithm for Single-Cable-ConFL.*

Proof. By Lemmas 10–13, we have

$$E[O + T + C] \leq O'_U + 2 \cdot C'_U + \rho_{ST}(2T^* + (\alpha + 1)C^*) .$$

Applying Lemma 14, we can bound the first two terms, which yields

$$\begin{aligned} E[O + T + C] &\leq \rho_{ST}(2T^* + (\alpha + 1)C^*) + 2[\lambda_F(O^* + \alpha C^*) + \lambda_C(C^* + \frac{0.807}{\alpha} T^*)] \\ &= (2\lambda_F)O^* + 2(\lambda_C \frac{0.807}{\alpha} + \rho_{ST})T^* + (\rho_{ST}(\alpha + 1) + 2(\lambda_F \alpha + \lambda_C))C^* . \quad (1) \end{aligned}$$

Applying Byrka's $(\lambda_F, 1 + 2 \cdot e^{-\lambda_F})$ -bifactor approximation algorithm [2] for the facility location subproblem and the (currently best known) $\ln(4)$ -approximation algorithm for the Steiner tree problem [3] and setting $\alpha = 0.5043$ and $\lambda_F = 2.1488$, inequality (1) implies $E[O + T + C] \leq 6.72(O^* + T^* + C^*)$. \square

For unit demands, one can derive a stronger bound of $c(T_2) + C_U$ for the cable installation costs using the techniques proposed in [11] for the single sink network design problem. Adapting Step 6 of the algorithm and adjusting the parameters α and λ_F accordingly, one easily obtains a 4.57-approximation algorithm for the Single-Cable-ConFL problem with unit demands.

References

1. S. Ahmadian, C. Swamy: Improved Approximation Guarantees for Lower-Bounded Facility Location In Proc. of WAOA 2012.

2. J. Byrka. An Optimal Bifactor Approximation Algorithm for the Metric Uncapacitated Facility Location Problem. In Proc. of APPROX 2007, pages 29-43.
3. J. Byrka, F. Grandoni, T. Rothvoß, L. Sanità. An improved LP-based approximation for Steiner tree. In Proc. of STOC 2010, pages 583-592.
4. F. Eisenbrand, F. Grandoni, T. Rothvoß, G. Schäfer. Connected facility location via random facility sampling and core detouring. Journal of Computer and System Sciences, 76: pages 709-726, 2010.
5. N. Garg, R. Khandekar, G. Konjevod, R. Ravi, F. S. Salman and A. Sinha. On the integrality gap of a natural formulation of the single-sink buy-at-bulk network design formulation. In Proc. of IPCO 2001, pages 170-184.
6. F. Grandoni and T. Rothvoß. Network design via core detouring for problems without a core. In Proc. of ICALP 2010, pages 490-502.
7. F. Grandoni, T. Rothvoß. Approximation algorithms for single and multi-commodity connected facility location. In Proc. of IPCO 2011, pages 248-260.
8. S. Guha, A. Meyerson, K. Munagala: Hierarchical placement and network design problems. In Proc. of FOCS 2000, pages 603-612.
9. S. Guha, A. Meyerson, K. Munagala. A constant factor approximation for the single sink edge installation problems. In Proc. of STOC 2001, pages 383-388.
10. A. Gupta, J. Kleinberg, A. Kumar, R. Rastogi, and B. Yener. Provisioning a virtual private network: A network design problem for multicommodity flow. In Proc. of STOC 2001, pages 389-398.
11. R. Hassin, R. Ravi, F.S. Salman. Approximation algorithms for a capacitated network design problem. Approximation Algorithms for Combinatorial Optimization, LNCS 1913, pages 167-176, 2000.
12. R. Jothi and B. Raghavachari. Improved approximation algorithms for the single-sink buy-at-bulk network design problems. In Proc. of SWAT 2004, pages 336-348.
13. R. Ravi, A. Sinha. Integrated logistics: Approximation algorithms combining facility location and network design. In Proc. of IPCO 2002, pages 212-229.
14. Z. Svitkina. Lower-bounded facility location. Trans. on Algorithms, 6(4), 2010.
15. C. Swamy and A. Kumar: Primal-dual algorithms for connected facility location problems. Algorithmica 40, pages 245-269, 2004.
16. K. Talwar. The single-sink buy-at-bulk LP has constant integrality gap. In Proc. of IPCO 2002, pages 475-486.