Connected Facility Location Problems

Ivana Ljubić

Abstract The (uncapacitated, unrooted) Connected Facility Location Problem (ConFL) is the problem of determining a minimum-cost solution consisting of a set of open facilities used to serve a given set of clients, and to connect the open facilities using a Steiner tree. The cost of a ConFL solution consists of edge-costs for building the Steiner tree, the cost for allocating the clients to open facilities, and the cost for opening the facilities. Hence, the ConFL is a generalization of two well-known combinatorial optimization problems: the uncapacitated facility location problem, and the Steiner tree problem in graphs. In the past two decades there has been a considerable amount of research devoted to it. The aim of this article is to provide a comprehensive overview of the existing methods, theoretical results and computational studies for the ConFL and its variants.

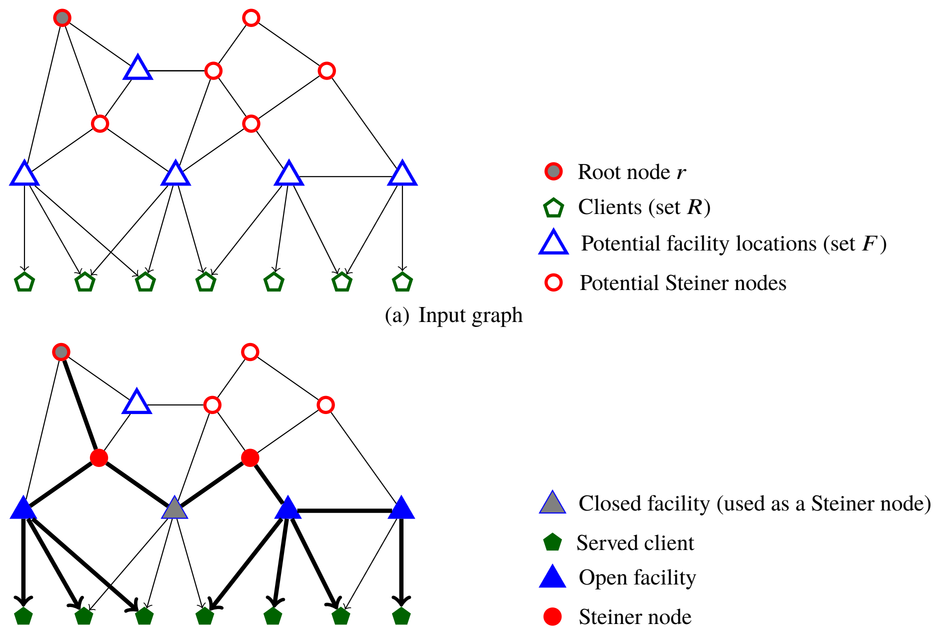
1 Introduction

Formally, the *unrooted ConFL* is defined as follows. We are given a graph $G = (V, E_S \cup A_R)$ with the set of nodes V partitioned into the set of *clients* (R), the set of *potential facility locations* (F) and the set of *potential Steiner nodes* ($V \setminus (F \cup R)$). The undirected graph $G_S = (V_S, E_S)$, with $V_S = V \setminus R$ and E_S being the set of edges between nodes from V_S, represents the *core network*. Connections between potential facilities and clients are modeled using the *assignment network*, which is a directed bipartite graph $G_R = (F \cup R, A_R)$ with $A_R \subseteq \{(i, k) : i \in F, k \in R\}$. Figure 1(a) illustrates an input graph G.

The following input parameters are associated to G:

- Facility opening cost $f_i \geq 0$, for each $i \in F$.

Ivana Ljubić
ESSEC Business School of Paris, Cergy-Pontoise, France
e-mail: ljubic@essec.edu



(a) Input graph

(b) A feasible solution

Fig. 1 Example of an input instance and its solution for the rooted, undirected ConFL.

- Edge cost $c_e \geq 0$, for each $e \in E_S$.
- Assignment cost $c_{ij} \geq 0$, for each $(i, j) \in A_R$.

The goal is to find a subnetwork of G consisting of a set of open facilities $F' \subseteq F$, a set of core edges $E'_S \subseteq E_S$ and a set of assignment arcs $A'_R \subseteq A_R$ such that:

(P1) Each client is assigned to exactly one open facility using arcs from A'_R.
(P2) The set of core edges E'_S connects all open facilities.
(P3) The sum of assignment, facility opening and core-edge costs, given by

$$\sum_{e \in E'_S} c_e + \sum_{i \in F'} f_i + \sum_{(i,j) \in A'_R} c_{ij}, \tag{1}$$

is minimized.

Since the core-edge costs are non-negative, there always exists a solution which connects the set of open facilities through a Steiner tree. When building the Steiner tree induced by the set of core edges from E'_S, nodes from V_S, together with facilities from $F \setminus F'$ can be used as interconnecting nodes, to ensure the connectivity between open facilities from F'. Such nodes are typically referred to as *Steiner nodes*. No costs are incurred for including such Steiner nodes in the solution. Figure 1(b) illustrates a feasible solution.

The ConFL is a generalization of two well-known combinatorial optimization problems: the uncapacitated facility location (UFL, see e.g., Fernández and Landete [2019]), and the Steiner tree problem in graphs (STP, see, e.g., Ljubić [2021]). The problem is clearly NP-complete as, indeed, it entails as a special case the UFL (when the core-edge costs are all equal to zero). Moreover, the problem also entails the undirected STP as a special case: indeed, if there is a unique arc connecting each client to a facility, the allocation of facilities and opening of facilities is uniquely determined, and it remains to solve the STP using the set of open facilities as terminals.

The problem has been introduced by Karger and Minkoff [2000], and since then there has been a considerable amount of research devoted to it. Several other problem variants have been studied in the literature:

- ***The rooted ConFL:*** In addition, we are given a *root node* $r \in V \setminus (F \cup R)$, which usually represents a central facility that must be included in every feasible solution. In the context of telecommunication networks, for example, the root node establishes a connection to a higher order (e.g., backbone) network.
- ***The asymmetric ConFL:*** In some applications, the core network is represented by a directed graph $G_S = (V_S, A_S)$ where the arcs from A_S connect nodes from V_S, with the costs $c_{ij} \geq 0$, $(i, j) \in A_S$. The connection between open facilities has to be established through a subset of arcs $A'_S \subseteq A_S$ such that there exists a directed path between the root r and any open facility $i \in F'$. Due to the assumption that all arc costs are non-negative, there always exists an optimal solution such that in the core-subgraph induced by A'_S, the in-degree of every node (except the root) is equal to one. Such a rooted, directed tree is referred to as a *Steiner arborescence*.
- ***The Capacitated ConFL (CapConFL):*** In this problem, we are additionally given a capacity $v_i \geq 0$ and demand $d_i \geq 0$ for each $i \in F$, and a client demand $b_k \geq 0$ for each $k \in R$. To each $e \in E_S$ we associate edge capacity $u_e \geq 0$. The goal is to find subsets E'_S, A'_R and F', as above, such that, in addition to properties (P1)-(P3), also properties (P4)-(P5) are satisfied:

  (P4) The sum of clients' demands assigned to a facility i does not exceed its capacity v_i. If the assignment of a client's demand is allowed to be split between multiple facilities, we talk of *multi-source* allocation, otherwise, a *single-source* allocation is imposed.

  (P5) In the core subnetwork induced by E'_S, we can simultaneously route the flow from the root node to satisfy the demand of all open facilities, without violating the edge capacities. The flow between the root and each open facility can be split into multiple paths.

  Hence, the CapConFL combines two other prominent combinatorial optimization problems: (1) the capacitated facility location problem [Fernández and Landete, 2019] and the single-source network loading problem [Ljubić et al., 2012].

1.1 MILP Formulation for the Asymmetric ConFL

As an illustration, we provide a MILP formulation for the asymmetric ConFL. Facility decision variables $z_i \in \{0,1\}$, $i \in F$, indicate whether facility i is opened. Arc decision variables $x_a \in \{0,1\}$, $a \in A_S$, are set to one if and only if arc a belongs to the Steiner arborescence. Finally, assignment variables $a_{ij} \in \{0,1\}$, $(i,j) \in A_R$ are equal to one if and only if facility i serves customer j. For a set $H \subset V$, we define $\delta^-(H) := \{(u,v) \in A : u \notin H, v \in H\}$ and $\delta^+(H) := \{(u,v) \in A : u \in H, v \notin H\}$, where $A = A_S \cup A_R$. Moreover, for any vector $\mu \in \{0,1\}^M$ over a ground set M, we write $\mu(M') = \sum_{m \in M'} \mu_m$, for any $M' \subseteq M$. The asymmetric ConFL is then modeled as follows:

$$\min \sum_{a \in A_S} c_a x_a + \sum_{i \in F} f_i z_i + \sum_{(i,j) \in A_R} c_{ij} a_{ij} \tag{2a}$$

$$a(\delta^-(j)) = 1 \qquad \forall j \in R \tag{2b}$$

$$a_{ij} \leq z_i \qquad \forall (i,j) \in A_R \tag{2c}$$

$$x(\delta^-(H)) \geq z_i \qquad \forall H \subseteq V_S : r \notin H, \forall i \in H \cap F \tag{2d}$$

$$x(\delta^-(s)) \leq 1 \qquad \forall s \in V_S \setminus \{r\} \tag{2e}$$

$$(x,z,a) \in \{0,1\}^{|A_S|+|F|+|A_R|} \tag{2f}$$

Constraints (2b) ensure that every client is assigned to exactly one facility, while constraints (2c) make sure that assignment arcs can only be used if the corresponding facility is opened. The directed cut-set inequalities (2d) ensure that there is a directed path from the root r to every open facility $i \in F$. If for all $a \in A_S$, $c_a > 0$, constraints (2b)-(2d) guarantee that the core part of an optimal solution will correspond to a Steiner arborescence rooted at r. In case some of the arc costs are equal to zero, in order to prevent cycles, we can impose additional constraints (2e) to ensure that the in-degree of every node in the core network is at most one. Once can replace binary constraints on a variables with $a_{ij} \leq 1$, $(i,j) \in A_R$, since the constraint matrix associated with a variables is totally unimodular.

The given MILP formulation can also be used to solve the ConFL on undirected or unrooted core graphs. Indeed, to deal with undirected core networks with a given root node, it is sufficient to replace each edge $e = \{k,l\}$ with its two bidirected counterparts, and to set $c_{kl} = c_{lk} = c_e$. To deal with unrooted problems, one has to insert an artificial root node r to the core network, add arcs (r,i), $iinF$, to A_S with $c_{ri} = 0$ and add an additional root-outdegree constranint $x(\delta^+(r)) = 1$ to the above model.

1.2 Applications

The ConFL has been used to model network design problems that arise in the design of last mile telecommunication networks. Gollowitzer and Ljubić [2011] report applications for the *fiber to the curb* (FTTC) deployment strategy in the design of local access telecommunication networks. In this setting, fiber optic cables run from a central office to a cabinet serving a neighborhood. End users connect to this cabinet using the existing copper connections. Expensive switching devices (that correspond to facilities) are installed in these cabinets. The decision where to place the switching devices represents the facility opening decision, building of the fiber optic network is modeled using a Steiner tree, and assignment between households and switching devices correspond to client-facility assignment decisions. *Fiber-to-the-building* (FTTB) and *fiber-to-the-air* (FTTA) are two other deployment architectures in the design of local access networks that can be modeled as the ConFL [Leitner et al., 2013]. Similarly, *fiber-to-the-home* (FTTH) deployment is modeled using variants of the Steiner tree problem [Leitner et al., 2014b, Rehfeldt, 2021, Ljubić et al., 2006]. All these architectures are known under a common name as FTTx deployment strategies. Indeed, after abstracting the more technical details and considering mainly topology decisions, Arulselvan et al. [2019], Grötschel et al. [2014] use the ConFL to model the deployment of FTTx architectures and report varoius case studies with real-world telecommunication networks. Similarly, the CapConFL models a more detailed planning of FTTx networks, in which capacities of the links and of multiplexor devices are taken into account (see Leitner and Raidl [2011a], Gollowitzer et al. [2013]). For further details concerning the design of local access telecommunication networks, see Grötschel et al. [2014], Fortz [2015].

Applications arising in information/content distribution networks are discussed by Karger and Minkoff [2000], Krick et al. [2003]. These applications deal with network design decisions in data distribution networks where servers (acting as facilities) need to be located. Demand nodes (clients) download the content from the closest server. When a certain content is updated on one server, it has to be updated on all other servers (and hence they need to communicate through a Steiner tree). Gupta et al. [2001] report applications for the design of Virtual Private Networks (VPNs), where a group of nodes (clients) forms a virtual subnetwork with a set of dedicated services.

Finally, Zhu et al. [2012] show the importance of the ConFL for applications in emergency management where distribution centers correspond to facilities that are used for reserving, distributing and transiting relief goods. Demand nodes (i.e., clients) are cities and open facilities need to be connected through a transportation network.

2 Exact Algorithms

Gollowitzer and Ljubić [2011] propose several (mixed-integer linear programming) MILP formulations for the ConFL with a given root node. They transform the undirected core network into its bidirected counterpart, and study two possible ways of modeling the connectivity in the solution. The first approach is based on ensuring that the root is connected to each open facility, as this is done with constraints (2d) in the model (2). The second approach imposes the connectivity between the root r and every client. In the model (2), this would correspond to replacing constraints (2d) with:

$$x(\delta^-(H)) + \sum_{i \notin H \cap F} a_{ij} \geq 1 \qquad \forall H \subset V_S : r \notin H, \forall j \in R \tag{3}$$

Inequalities (3) ensure that for each client $j \in R$ and any subset H of the core nodes (excluding r), either there is an arc entering the set H, or j has to be served by a facility outside of $H \cap F$. When $H \cap F = F$, these inequalities reduce to $x(\delta^-(H)) \geq 1$.

Gollowitzer and Ljubić [2011] show that the latter connectivity concept leads to stronger LP-relaxation bounds. To model the connectivity, besides the cut-set based models presented in this article, they also explore the flow-based models (single-commodity, multi-commodity and common-flow approaches), and a model based on Miller-Tucker-Zemlin constraints. All formulations are compared with respect to their polyhedral strength. In addition, branch-and-cut algorithms are proposed and tested on benchmark instances from the literature. The models studied by Gollowitzer and Ljubić [2011] can be also used to solve the unrooted or undirected ConFL, as explained in Section 1.1.

Leitner et al. [2018] study the polytope for the ConFL on undirected graphs. They show that generalized subtour elimination constraints (imposed on the nodes of the core network) are facet-defining. Besides, they provide several new families of valid inequalities that link core-network edge variables with facility and assignment variables. In an earlier article, Leitner et al. [2014a] study the polytope of the asymmetric ConFL. The authors give necessary and sufficient conditions under which inequalities (3) are facet-defining, and provide several new families of facet-inducing inequalities.

Leitner et al. [2017] study a larger family of problems with a tree-star topology. They show that the Steiner tree-star [Lee et al., 1993], generalized Steiner tree-star [Khuller and Zhu, 2002] and rent-or-buy problems [Swamy and Kumar, 2004] can be seen as special cases of the asymmetric ConFL. A branch-and-cut algorithm combined with sophisticated reduction and heuristic techniques is proposed. An important ingredient of this framework is a dual ascent procedure for the asymmetric ConFL. Optimal solutions are provided for many previously unsolved benchmark instances in the literature. To the best of our knowledge, the framework of Leitner et al. [2017] is currently state-of-the-art approach for the (uncapacitated) ConFL and above mentioned problems.

Gollowitzer et al. [2013] introduce the CapConFL and propose a MILP formulation, together with strong valid inequalities and a branch-and-cut algorithm. Their computational study is based on real-life FTTC and FTTB instances. To model the property (P4) stated in Section 1, they introduce constraints

$$\sum_{(i,j)\in A_R} b_j x_{ij} \le v_i z_i \qquad \forall i \in F.$$

To ensure that there is a sufficient capacity on the arcs of the core network, they impose the capacitated cut-set inequalities:

$$\sum_{(k,l)\in\delta^-(H)} u_{kl} x_{kl} \ge \sum_{i\in H\cap F} d_i z_i \qquad \forall H \subset V_S : r \notin H.$$

These cuts guarantee that the flow of d_i units can be routed from r to each open facility $i \in F$. Leitner and Raidl [2011a] study a variant of the CapConFL with capacities on facilities, but with no capacities on edges/arcs of the core network. In addition, they consider a prize-collecting objective function, in which a penalty is incurred for unserved clients. Overall, their problem combines features of the single-source capacitated facility location and the prize-collecting Steiner tree problem. Two MILP formulations are proposed: the first one is based on cut-sets between the root and served clients, together with the single-source assignment of served clients to open facilities. In the second formulation, the set of all feasible and profitable assignment patterns for each potential facility $i \in F$ is considered. Since the number of these patterns can be exponential in size, a branch-and-cut-and-price approach is implemented for finding optimal solutions. In terms of the quality of LP-relaxation bounds, the second MILP model is shown to dominate the cut-based formulation as well as the other two multi-commodity-flow-based MILP formulations proposed by the same authors in [Leitner and Raidl, 2009].

3 Approximation Algorithms

Early work on ConFL mainly includes approximation algorithms. An α-approximation algorithm for a minimization problem is a polynomial time algorithm that, for all instances of the problem, produces a solution whose value is within a factor α of the optimal solution value. Approximation algorithms developed for the ConFL work under the following assumptions. We are given an input graph $G = (V, E)$ with a set of clients $R \subset V$, edge lengths $c_e \ge 0$, $e \in E$ that form a metric, and a constant M. Every node can act as a Steiner node, and every node can act as a facility (with opening cost $f_i \ge 0$ given for each $i \in V$). If node i is assigned to node j, assignment cost c_{ij} is incurred, however if edge $\{i, j\}$ is used as part of the Steiner tree, the edge-core cost equal to $M \cdot c_{ij}$ is incurred. Under these assumptions, early approximation algorithms are given by Karger and Minkoff [2000] and Gupta et al. [2001]. The algorithm of Gupta et al. [2001] is based on an LP-rounding procedure and guarantees the approximation ratio of 10.66. Swamy and Kumar [2004] propose a primal–dual approximation algorithm whose approximation ratio is reduced to 8.55. Grandoni and Rothvoß [2011] present a new 3.19 approximation algorithm for the ConFL, which improves the previous best approximation factor of 3.92 given by Eisenbrand et al. [2010].

The directed version of the ConFL is as hard to approximate as the Steiner arborescence problem. Halperin and Krauthgamer [2003] show that the latter problem admits no $\log^{2-\varepsilon}(n)$-approximation (where n is the number of nodes in the input graph) for any constant $\varepsilon > 0$ unless NP $\subseteq$ DTIME($n^{\mathrm{polylog}(n)}$).

4 Heuristics

Heuristics have been proposed for both, capacitated and uncapacitated variants of the ConFL. Most of them rely on greedy randomized adaptive search procedure (GRASP) and various neighborhood search techniques.

For solving the rooted undirected ConFL, Ljubić [2007] uses a variable neighborhood search (VNS) heuristic that is combined with a reactive tabu search. Specifically, the solutions are encoded as binary vectors representing the set of open facilities and k-neighborhood is defined using Hamming distance between pairs of such vectors. A tabu list consists of a set of facilities that cannot be flipped during the neighborhood search, and the size of this list is dynamically adapted. A set of benchmark instances derived by combining the UFL instances from the UflLib[1], and STP instances from the OR-Library[2], is introduced.

For the unrooted, undirected ConFL, Tomazic and Ljubić [2008] propose a GRASP heuristic with the same encoding of feasible solutions as in Ljubić [2007]. GRASP is a multi-start iterative approach where in each iteration an initial solution is constructed in a greedy fashion, and later improved by a local improvement procedure. The best overall solution is reported as the final one. As a search intensification mechanism, open- and close-facility moves are applied, followed by a shortest path Steiner tree heuristic to find connections between open facilities. Furthermore, the authors show how to transform the ConFL into the minimum Steiner arborescence problem and then solve it by an exact branch-and-cut method. The empirical results show that the GRASP approach improves feasible solutions found by the branch-and-cut method and provides solutions of high-quality (with the gaps to the best found lower bound ranging between 3% and 9% on the set of considered benchmark instances).

Bardossy and Raghavan [2010] combine dual ascent with local search to derive strong lower and upper bounds for a more general variant of the ConFL. The proposed methodology can namely be applied to the four problems: the Steiner tree-star problem, the general Steiner tree-star problem, the ConFL, and the rent-or-buy

[1] https://resources.mpi-inf.mpg.de/departments/d1/projects/benchmarks/UflLib/

[2] http://people.brunel.ac.uk/ mastjjb/jeb/orlib/steininfo.html

problem. In order to derive valid lower bounds, the authors transform the input instance into an instance of the *general ConFL* in which use of a facility node necessarily incurs a facility opening cost, and there are no opening costs associated with Steiner nodes. Then, they apply the dual ascent procedure from Raghavan [1995] to the Steiner arborescence reformulation of the general ConFL. That way, a lower bound is obtained together with a heuristic solution. The latter is then improved by searching for a better tree on the existing set of nodes and by closing open facilities and possibly reassigning clients to remaining open facilities.

Leitner and Raidl [2009] study the CapConFL variant with no edge capacities in the core network, and with the prize-collecting objective function. They propose two multi-commodity-flow based formulations. In one of the two formulations (which sends one unit of flow to each served client), the authors relax flow capacity constraints in Lagrangian fashion. They show that such obtained problem decomposes into $|R|$ shortest path problems, one per each client, and $|F|$ knapsack problems, one per each facility. Largangian dual bounds are then obtained by applying the volume algorithm. Finally, a heuristic solution found by the Lagrangian decomposition approach is then improved by applying a plethora of local improvement procedures. They include: key-path improvements (that were successfully used for Steiner trees in Voß [1992]), customer-swap neighborhoods and very large scale neighborhood search. Later, Leitner and Raidl [2011b] study the same problem variant and develop a VNS and a GRASP procedure. Both, VNS and GRASP rely on several local improvement strategies, namely: key-path improvements (see Leitner and Raidl [2009]), client-swaps to achieve lower assignment costs, and facility swaps.

5 Extensions

The *hop-constrained ConFL* has been introduced by Ljubić and Gollowitzer [2013]. In this problem, the maximum number of edges between the root and any open facility cannot exceed a given parameter $H \in \mathbb{N}$. The authors examine various ways to model hop-constraints. Specifically, all formulations are derived on layered graphs. Preprocessing procedures are proposed to reduce the size of the layered graphs. A polyhedral comparison of the proposed models is given and branch-and-cut procedures are presented.

The *incremental ConFL*, a multi-period variant of the rooted ConFL, has been introduced by Arulselvan et al. [2019], see also Arulselvan et al. [2011]. The goal is to find a schedule that, for each time period, identifies a subset of facilities that can provide service, a subset of clients to be served, and assignment of served clients to open facilities, as well as the set of edges in the core network that connects open facilities. In each time period, the total demand of the served customers must satisfy the minimum coverage requirement and the chosen edges in E_S must form a network connecting the open facilities with the root node. Once a client is served, it has to be served for the remainder of the planning periods. The goal is to minimize the total costs of the network, defined as the sum of investment and maintenance costs over all periods. A branch-and-cut approach based on cover, cut-set-cover and degree-balance inequalities is proposed.

Leitner et al. [2013] introduce the *two-architecture ConFL* (2AConFL) where clients can be served using one of the two potential technologies. As in the ConFL, decisions regarding the set of open facilities and their connections using a Steiner tree need to be made. In addition, a minimum-coverage rate specifies the fraction of clients that need to be supplied by the better of the two technologies. Hence, one has to decide which one of the available technologies should be installed at facility locations, and which technology will be used for serving different clients. The major challenge from the decision maker perspective is in finding the appropriate coverage rate to properly deal with the trade-off between solution cost and the demand of clients served by the better (and thus more expensive) technology. To overcome this issue, Leitner et al. [2016] propose to look at the problem from the bi-objective perspective. They introduce the *bi-objective ConFL* where the goal is to determine a subset of clients to be served by the better of two technologies, so that on the one hand the cost of the solution is minimized (as in the standard ConFL) and on the other hand the demand of clients served by the better technology is maximized. The authors propose two novel local search methods for biobjective optimization (boundary induced neighborhood search and directional local branching heuristics) and a new exact method called adaptive search in the objective space. Benchmark instances representing telecommunication access networks from Germany are used in the empirical study. The obtained results show that with the proposed methodology the Pareto front of instances of realistic size can be efficiently explored.

The ConFL with uncertain input data has also been studied in the recent literature. Bardossy and Raghavan [2016] assume that assignment costs are subject to interval uncertainty, i.e., for each $(i, j) \in A_R$ the assignment cost belongs to the interval $[c_{ij}, c_{ij} + \Delta_{ij}]$ where $\Delta_{ij} > 0$ represents the deviation from the nominal assignment cost c_{ij}. This deviation can represent the uncertainty with respect to the location of client j or to their demand. Following the Γ-robustness concept introduced by Bertsimas and Sim [2003], the goal is to find a ConFL solution that minimizes the maximum value of the solution when up to Γ assignment costs are at their worst-case value $(c_{ij} + \Delta_{ij})$ and the remaining assignment costs are at their nominal value. In other words, the objective function (1) of the deterministic ConFL is replaced by

$$\min \sum_{e \in E'_S} c_e + \sum_{i \in F'} f_i + \sum_{(i,j) \in A'_R} c_{ij} + \max_{U \subseteq A'_R, |U| \leq \Gamma} \sum_{(i,j) \in U} \Delta_{ij},$$

Due to the result of Bertsimas and Sim [2003], see also Álvarez Miranda et al. [2013], the optimal solution of the Γ-robust ConFL problem can be found by solving a series of deterministic ConFL problems with modified assignment costs. Bardossy and Raghavan [2016] exploit this result and introduce the Approximate Robust Optimization (ARO) method which provides valid lower and upper bounds for Γ-robust optimization problems, in general. The ARO method uses lower and upper bounds obtained from a heuristic procedure (e.g., a dual ascent approach) applied to the series of deterministic counterparts. In the context of the robust ConFL, the authors apply the dual ascent heuristic proposed earlier in Bardossy and Raghavan [2010] and show that the ARO method finds high-quality solutions to the robust ConFL very rapidly.

Bardossy and Raghavan [2017] introduce the *stochastic ConFL* where the facility opening costs and the connection costs between them are assumed to be known a priori, while the assignment costs are subject to uncertainty and dependent upon a set of random scenarios. The here-and-now decisions are concerned with finding a set of open facilities and a Steiner tree that connects them. After that, the uncertainty regarding the assignment costs is revealed (i.e., one of the scenarios is realized) and customers are assigned to closest open facilities. If we denote by $\mathcal{S}$ the set of possible discrete scenario realizations, and by $c^s_{ij} \geq 0$ the assignment cost in a scenario $s \in \mathcal{S}$, $(i, j) \in A_R$, the objective function of the deterministic ConFL (1) is replaced by its two-stage stochastic counterpart

$$\min \sum_{e \in E'_S} c_e + \sum_{i \in F'} f_i + \sum_{s \in \mathcal{S}} \pi_s \sum_{(i,j) \in A^s_R} c^s_{ij},$$

where $\pi_s > 0$ refers to the probability of a scenario s and $A^s_R \subset A_R$ represents the optimal assignment of clients under s, $s \in \mathcal{S}$. In contrast to the Sample Average Approximation (SAA) method (where the sample average problems are solved exactly), the authors solve the sample average problems with a heuristic coupled with a lower bound mechanism. This new approach, called the the Inexact SAA method, is then applied to the stochastic ConFL where the dual ascent procedure from Bardossy and Raghavan [2010] is used as the underlying lower- and upper-bounding heuristic.

We close this section by mentioning some other problems closely related to the ConFL. Cherkesly et al. [2019] introduce the Median Problem with Interconnected Facilities and the Covering Problem with Interconnected Facilities, see also [Kuzbakov and Ljubić, 2023]. For the former one, all customers have to be served by open facilities, and for the latter one, if it is not profitable to cover all customers, a penalty is imposed for each customer that does not receive the service. Similar as for the ConFL, all facilities have to be connected, however no connection costs are imposed, i.e., $c_e = 0$ for all $e \in E_S$, and in addition, opening costs have to be paid for all facilities, even if they are used as Steiner nodes. Interconnectivity between the facilities is modeling communications in wireless sensor networks or radio networks, which explains why no connection costs are imposed on the core network. Other problems related to the ConFL are the p-cable-trench problem [Calik et al., 2017, Marianov et al., 2012] and p-arborescence star problem [Morais et al., 2019]. We refer to Contreras and Fernández [2012] for a more detailed survey on network design problems with facility location decisions.

6 Conclusions

In this paper we have provided a survey on existing methods for solving the ConFL and its variants. Recent research has focused on extensions of the ConFL, notably in introducing more general network design concepts and combining them with facility location. For example, capacities on the edges of the core network, or on the open facilities were imposed, and covering variants (in which non-profitable customers are left unserved) are studied as well. Also, the problem variants under customer demand uncertainty have drawn the attention of the researchers in the past.

When it comes to future developments, we highlight the work of Grandoni and Rothvoß [2011] who introduce a multi-commodity generalization of the ConFL. In this problem we are given source-sink pairs (i.e., *commodities*) that wish to communicate. A feasible solution consists of a subset of open facilities, and a Steiner forest (rather than a Steiner tree) spanning them. While a constant approximation algorithm has been proposed for this problem variant, to the best of our knowledge, exact methods and meta/matheuristics are still to be developed. Moreover, the studies on the impact of data uncertainty have been very limited so far (focusing on customer demand only), however there are other sources of uncertainty that need to be considered in real-life applications, such as potential failures of connections in the core network, potential failures of open facilities, as well as uncertainty in the cost/capacity of core-edges or open facilities.

References

A. Arulselvan, A. Bley, S. Gollowitzer, I. Ljubić, and O. Maurer. MIP modeling of incremental connected facility location. In *Proceedings of INOC 2011*, volume 6701 of *Lecture Notes in Computer Science*, pages 490–502. 2011.

A. Arulselvan, A. Bley, and I. Ljubić. The incremental connected facility location problem. *Computers & Operations Research*, 112:104763, 2019.

M. Bardossy and S. Raghavan. Dual-Based Local Search for the Connected Facility Location and Related Problems. *INFORMS Journal on Computing*, 22(4):584–602, 2010.

M. G. Bardossy and S. Raghavan. Approximate robust optimization for the connected facility location problem. *Discretete Applied Mathematics*, 210:246–260, 2016.

M. G. Bardossy and S. Raghavan. An inexact sample average approximation approach for the stochastic connected facility location problem. *Networks*, 70(1): 19–33, 2017.

D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical Programming*, 98:49–71, 2003.

H. Calik, M. Leitner, and M. Luipersbeck. A Benders decomposition based framework for solving cable trench problems. *Computers & Operations Research*, 81: 128–140, 2017.

M. Cherkesly, M. Landete, and G. Laporte. Median and covering location problems with interconnected facilities. *Computers & Operations Research*, 107:1–18, 2019.

I. Contreras and E. Fernández. General network design: A unified view of combined location and network design problems. *European Journal of Operational Research*, 219(3):680–697, 2012. Feature Clusters.

F. Eisenbrand, F. Grandoni, T. Rothvoß, and G. Schäfer. Connected facility location via random facility sampling and core detouring. *Journal of Computer and System Sciences*, 76:709–726, 2010.

E. Fernández and M. Landete. *Fixed-Charge Facility Location Problems*, pages 67–98. Springer International Publishing, Cham, 2019.

B. Fortz. *Location Problems in Telecommunications*, pages 537–554. Springer International Publishing, Cham, 2015.

S. Gollowitzer and I. Ljubić. MIP models for connected facility location: A theoretical and computational study. *Computers and Operations Research*, 38(2): 435–449, 2011.

S. Gollowitzer, B. Gendron, and I. Ljubić. A cutting plane algorithm for the capacitated connected facility location problem. *Computational Optimization and Applications*, 55(3):647–674, 2013.

F. Grandoni and T. Rothvoß. Approximation algorithms for single and multi-commodity connected facility location. In O. Günlük and G. J. Woeginger, editors, *Integer Programming and Combinatoral Optimization - 15th International Conference, IPCO 2011, New York, NY, USA, June 15-17, 2011. Proceedings*, volume 6655 of *Lecture Notes in Computer Science*, pages 248–260. Springer, 2011.

M. Grötschel, C. Raack, and A. Werner. Towards optimizing the deployment of optical access networks. *EURO Journal on Computational Optimization*, 2(1): 17–53, 2014.

A. Gupta, J. Kleinberg, A. Kumar, R. Rastogi, and B. Yener. Provisioning a virtual private network: A network design problem for multicommodity flow. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, pages 389–398, New York, NY, USA, 2001. Association for Computing Machinery. ISBN 1581133499.

E. Halperin and R. Krauthgamer. Polylogarithmic inapproximability. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 585–594, 2003.

D. R. Karger and M. Minkoff. Building Steiner Trees with Incomplete Global Knowledge. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 613–623. IEEE Computer Society, 2000.

S. Khuller and A. Zhu. The General Steiner Tree-Star problem. *Information Processing Letters*, 84(4):215–220, 2002.

C. Krick, H. Räcke, and M. Westermann. Approximation algorithms for data management in networks. *Theory of Computing Systems*, 36(5):497–519, 2003.

Y. Kuzbakov and I. Ljubić. New formulations for two location problems with interconnected facilities. 2023.

Y. Lee, L. Lu, Y. Qiu, and F. Glover. Strong formulations and cutting planes for designing digital data service networks. *Telecommunication Systems*, 2:261–274, 1993.

M. Leitner and G. Raidl. Branch-and-cut-and-price for capacitated connected facility location. *Journal of Mathematical Modelling and Algorithms*, 10(3):245–267, 2011a.

M. Leitner and G. Raidl. Variable neighborhood and greedy randomized adaptive search for capacitated connected facility location. In R. Moreno-Díaz, F. Pichler, and A. Quesada-Arencibia, editors, *EUROCAST (1)*, volume 6927 of *Lecture Notes in Computer Science*, pages 295–302. Springer, 2011b.

M. Leitner and G. R. Raidl. A Lagrangian decomposition based heuristic for capacitated connected facility location. In *Proceedings of the 8th Metaheuristic International Conference (MIC 2009), Hamburg, Germany*, 2009.

M. Leitner, I. Ljubić, M. Sinnl, and A. Werner. On the two-architecture connected facility location problem. *Electronic Notes in Discrete Mathematics*, 41:359–366, 2013.

M. Leitner, I. Ljubić, J. J. S. González, and M. Sinnl. On the asymmetric connected facility location polytope. In P. Fouilhoux, L. E. N. Gouveia, A. R. Mahjoub, and V. T. Paschos, editors, *Combinatorial Optimization - Third International Symposium, ISCO 2014, Lisbon, Portugal, March 5-7, 2014, Revised Selected Papers*, volume 8596 of *Lecture Notes in Computer Science*, pages 371–383. Springer, 2014a.

M. Leitner, I. Ljubić, M. Luipersbeck, M. Prossegger, and M. Resch. New Real-world Instances for the Steiner Tree Problem in Graphs. Technical report, ISOR, University of Vienna, 2014b.

M. Leitner, I. Ljubić, M. Sinnl, and A. Werner. ILP heuristics and a new exact method for bi-objective 0/1 ILPs: Application to FTTx -network design. *Computers & Operations Research*, 72:128 – 146, 2016.

M. Leitner, I. Ljubić, and M. S. J.J. Salazar-González and. An algorithmic framework for the exact solution of tree-star problems. *European Journal of Operational Research*, 261(1):54–66, 2017.

M. Leitner, I. Ljubić, J. Salazar-González, and M. Sinnl. The connected facility location polytope. *Discrete Applied Mathematics*, 234:151–167, 2018.

I. Ljubić. A hybrid VNS for connected facility location. In T. Bartz-Beielstein, M. J. B. Aguilera, C. Blum, B. Naujoks, A. Roli, G. Rudolph, and M. Sampels, editors, *Hybrid Metaheuristics, 4th International Workshop, HM 2007, Dortmund, Germany, October 8-9, 2007, Proceedings*, volume 4771 of *Lecture Notes in Computer Science*, pages 157–169. Springer, 2007.

I. Ljubić. Solving Steiner trees: Recent advances, challenges, and perspectives. *Networks*, 77(2):177–204, 2021.

I. Ljubić and S. Gollowitzer. Layered graph approaches to the hop constrained connected facility location problem. *INFORMS Journal on Computing*, 25(2): 256–270, 2013.

I. Ljubić, R. Weiskircher, U. Pferschy, G. Klau, P. Mutzel, and M. Fischetti. An algorithmic framework for the exact solution of the prize-collecting Steiner tree problem. *Mathematical Progamming, Series B*, 105(2-3):427–449, 2006.

I. Ljubić, P. Putz, and J. J. S. González. Exact approaches to the single-source network loading problem. *Networks*, 59(1):89–106, 2012.

V. Marianov, G. Gutiérrez-Jarpa, C. Obreque, and O. Cornejo. Lagrangean relaxation heuristics for the p-cable-trench problem. *Computers & Operations Research*, 39 (3):620–628, 2012.

V. Morais, B. Gendron, and G. R. Mateus. The p-arborescence star problem: Formulations and exact solution approaches. *Computers & Operations Research*, 102:91–101, 2019.

S. Raghavan. *Formulations and algorithms for network design problems with connectivity requirements*. PhD thesis, Massachusetts Institute of Technology, Cambridge, 1995.

D. Rehfeldt. *Faster algorithms for Steiner tree and related problems*. PhD thesis, Technische Universität Berlin, 2021.

C. Swamy and A. Kumar. Primal-dual algorithms for connected facility location problems. *Algorithmica*, 40(4):245–269, 2004.

A. Tomazic and I. Ljubić. A GRASP algorithm for the connected facility location problem. In *Proceedings of the 2008 International Symposium on Applications and the Internet, SAINT 2008, 28 July - 1 August 2008, Turku, Finland*, pages 257–260. IEEE Computer Society, 2008.

S. Voß. Steiner's problem in graphs: Heuristic methods. *Discrete Applied Mathematics*, 40(1):45–72, 1992.

J. Zhu, J. Huang, D. Liu, and J. Han. Connected distribution center location problem under traffic network in emergency management. In *2012 Fifth International Joint Conference on Computational Sciences and Optimization*, pages 160–163, 2012.

E. Álvarez Miranda, I. Ljubić, and P. Toth. A note on the Bertsimas & Sim algorithm for robust combinatorial optimization problems. *4OR—A Quarterly Journal of Operations Research*, 11:349–360, 2013.