Cover-based inequalities for the single-source capacitated facility location problem with customer preferences

Christina Büsing[1] Markus Leitner[2] Sophia Wrede[1]

[1]Teaching and Research Area Combinatorial Optimization, RWTH Aachen University, Ahornstraße 55, 52074 Aachen, Germany. {buesing,wrede}@combi.rwth-aachen.de
[2]Department of Operations Analytics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. m.leitner@vu.nl

27. August 2024

Abstract

The *single-source capacitated facility location problem with customer preferences (SSCFLPCP)* is known to be strongly NP-hard. Computational tests imply that state-of-the-art solvers struggle with computing exact solutions. In this paper, we contribute two novel preprocessing methods which reduce the size of the considered integer programming formulation, and introduce sets of valid inequalities which decrease the integrality gap. Each of the introduced results utilises structural synergies between capacity constraints and customer preferences in the SSCFLPCP. First, we derive two preprocessing methods where the first method fixes location variables and the second method fixes allocation variables. Afterwards, we study cover-based inequalities. Here, we first strengthen the well-known cover inequalities: when determining covers, we also consider demands of customers not in the cover that must be assigned to the covered facility if a customer in the cover is assigned to it. We further strengthen these inequalities by including information on the assignments of customers in a cover if they are not assigned to the covered facility. Afterwards, we derive a new family of valid inequalities, which expresses the relation of open facilities based on sets of customers covering a facility. We then discuss solution methods for the corresponding separation problems and, finally, test our results for two preference types in a computational study. Our results show a clear positive impact of the preprocessing methods and inequalities, in particular when preferences are defined by assignment costs.

Keywords: capacitated facility location, customer preferences, valid inequalities, preprocessing

1 Introduction

Facility location problems (FLPs) play an essential role in operations research literature (Laporte et al., 2020, Celik Turkoglu and Erol Genevois, 2020). In their basic version, facilities providing some sort of service for customers need to be opened at potential sites and customers are assigned to these facilities. The aim is to minimise the total cost consisting of opening costs for the facilities and costs for assigning the customers. We refer to this problem as the *uncapacitated facility location problem* (*UFLP*). Naturally, each customer is assigned to the open facility with smallest assignment cost. The class of FLPs is widely applicable to many real-world problems such as locating health care institutions (Ahmadi-Javid et al., 2017), charging stations for electrical cars (Ahmad et al., 2022), and many more (Celik Turkoglu and Erol Genevois, 2020).

There are no limitations on the service a facility can provide in UFLPs. However, such limitations have to be taken into account for many real-world problems, for example, when locating hospitals (Mestre et al., 2015). Often, customers incur a certain demand and facilities have a certain capacity for serving customers' demands. Such problems belong to the class of *capacitated facility location problems* (*CFLP*s). It is no longer guaranteed that each customer will be assigned to the open facility with lowest assignment cost when considering capacities. The problem is called *single-source* CFLP (SSCFLP) if each customer has to be assigned to exactly one open facility.

If customers have an individual agenda regarding the facilities they want to be served at, then their assignment in the solution to the UFLP or CFLP might be in conflict with their preferences. In reality, however, customers then deviate to one of their most preferred open facilities. While this yields a feasible, although possibly worse, solution in the UFLP, it might turn a feasible solution of the (SS)CFLP infeasible. A real-world application for this class of location problems lies in the location of health care emergency centers. Here, facilities offer medical services, which are limited due to time and space. Customers' demands have to be served at these facilities without violating the facilities' capacities. In an emergency, it can be assumed that customers will seek the service of their closest open facility.

We refer to the problem in which each customer has to be served at one of their most preferred open facilities as the *(single-source capacitated) facility location problem with customer preferences* ((SSC)FLPCP). In this work, we study the SSCFLPCP, for which determining a feasible solution is already strongly NP-complete; cf., e.g., Büsing et al. (2022). Computational tests show that state-of-the-art solvers struggle with computing exact solutions. We contribute two novel preprocessing methods, which reduce the size of the considered integer programming formulation, and introduce sets of valid inequalities, which decrease the integrality gaps. More specifically:

- We propose two preprocessing methods, which fix location and allocation decisions, respectively.
- We propose the concept of *implied-demand covers*, which generalise normal covers by simultaneously considering preferences of customers and capacities of facilities.
- We propose and analyse inequalities for the SSCFLPCP based on implied-demand covers, which utilise information on assignments of customers in the cover who are not assigned to the covered facility.
- We perform a computational study in which our results are tested for the case that preferences are defined by assignment costs as well as *perturbed assignment costs*. Our results indicate that our approaches work well for the former case and show potential for the latter one.

With these contributions, we aim to offer results for improving the performance of exact solution methods. The remainder of this article is organized as follows. Section 2 provides an overview over related literature and adresses the research gap we aim to close with this paper. A formal problem definition is given in Section 3, in which we also provide an integer linear programming formulation and summarize used notation. Section 4 introduces the novel preprocessing routines. Section 5 introduces several sets of new valid inequalities that utilise information from customers covering a facility. In Section 6, we briefly discuss algorithms for solving the separation problems of our newly developed valid inequalities. Finally, we discuss the results of our computational study in Section 7.

2 Related work

Facility location problems are thoroughly studied in the literature, see, e.g., Laporte et al. (2020) or Celik Turkoglu and Erol Genevois (2020) for recent surveys. From the complexity-theoretical point of view, feasible solutions for UFLP instances can be computed in polynomial time while the computation of a cost-minimising solution is strongly NP-hard (Mirchandani and Francis, 1990). Conversely to the UFLP, finding a feasible solution for the single-source capacitated facility location problem is already strongly NP-complete. Even if the set of open facilities is already known, it might be hard to find a feasible assignment that meets the capacity constraints. This can be seen via a reduction from 3-*partition* (Garey and Johnson, 1979). The relevance and complexity of FLPs has triggered a large number of scientific articles related to different aspects such as exact solution methods (see, e.g., Avella and Boccia (2009), Görtz and Klose (2012) and Fischetti et al. (2016)), the investigation of their polyhedral structure (see, e.g., Leung and Magnanti (1989), Aardal et al. (1995), Avella and Boccia (2009) and Avella et al. (2021)), and heuristic solution methods (see, e.g., Mirchandani and Francis (1990) and Korte and Vygen (2018)).

Most articles considering facility location problems with customer preferences focus on the uncapacitated case. The work of Hanjoul and Peeters (1987) is considered to be the first occurrence of preference constraints in the context of facility location problems. The authors propose an exact algorithm, which utilises a branch-and-bound procedure, as well as two heuristics for solving the UFLP with customer preferences. Several articles focus on preprocessing strategies and valid inequalities, see, e.g., Cánovas et al. (2007), Vasilyev et al. (2010) and Vasilyev et al. (2013). A semi-Lagrangian relaxation heuristic approach is proposed by Cabezas and García (2022).

Rojeski and ReVelle (1970), Wagner and Falkson (1975) and Gerrard and Church (1996) focus on the special case in which preferences are defined by distances, i.e., each customer prefers to be served at their closest open facility. Rojeski and ReVelle (1970) additionally consider capacities, allow to split customer demands and to expand capacities. Wagner and Falkson (1975) present models for the location of public facilities which maximise social welfare and propose a set of closest assignment constraints. Gerrard and Church (1996) review constraints for integer linear programming formulations for enforcing closest assignments and identify applications for facility location problems with closest assignment constraints. Espejo et al. (2012) theoretically compare all closest assignment constraints in the literature of discrete location theory up until 2011 and contribute a new set of constraints.

Adding closest assignment constraints to the classical single-source capacitated facility location problem yields surprising complexity results. If assignment costs correspond to distances in an underlying graph and the graph is a path or a cycle, then an optimal solution can be computed in polynomial time (Büsing et al., 2022). In contrast, the single-source CFLP is strongly NP-hard independently of the structure of assignment costs, as discussed at the beginning of this section.

When studying CFLPs with customer preferences, there are two main approaches to dealing with overloaded facilities. In the first approach, the overloaded facility may be opened and some customers will be served at a facility they like less. In the second approach, each customer has to be served at their most preferred open facility and facilities that would be overloaded can never be opened. Note that whether a facility will be overloaded depends on which other facilities are opened. Most research revolves around the first approach (Casas-Ramírez et al., 2018, Calvete et al., 2020, Polino et al., 2023). Here, the authors consider variations of a bilevel setting where the leader opens facilities with the aim to minimise the total sum of opening and assignment costs; the follower assigns each customer, for which a ranking of all potential facilities is known, to the open facilities with the objective to optimise the sum of achieved preference rankings of the customers. Calvete et al. (2020) compare computational results of the first and second approach to dealing with overloaded facilities.

To the best of our knowledge, no research has been conducted on the second approach for dealing with overloaded facilities besides the work by Büsing et al. (2022). Kang et al. (2023) study a variation of the CFLPCP in which each customer has to be served at their most preferred open facility and the capacities installed at facilities are part of the decision process. In contrast to the setting considered in this article, the authors allow for customers to not be served at any of the open facilities.

The consideration of capacities in facility location problems induces a knapsack-like structure in the CFLP. For knapsack problems, cover inequalities are well studied (Crowder et al., 1983, Gu et al., 1999), and they can be strengthened to so-called extended and lifted cover inequalities (Balas, 1975). Klabjan et al. (1998) show that the separation of cover inequalities is weakly NP-hard. Kaparis and Letchford (2010) present exact and heuristic separation algorithms. Due to the knapsack-like structure occurring in capacitated facility location problems, cover inequalities are useful in these problems as well (Aardal et al., 1995).

To the best of our knowledge, almost no research has been conducted on the SSCFLPCP in which facilities that would be overloaded cannot be opened. Polynomial time algorithms for special cases suggest further potential arising from the combination of capacities and closest assignments. With this work, we aim to shed light on cover-based inequalities in the context of single-source capacitated facility location problems with customer preferences when customers have to be served at their most preferred open facility.

3 Problem definition, notation, and an integer linear programming formulation

In the single-source capacitated facility location problem with customer preferences (SSCFLPCP), we are given a set of customers I with demands $d_i \in \mathbb{Z}_{\geq 0}$ for each $i \in I$ and a set of potential facilities J. A capacity $Q_j \in \mathbb{Z}_{\geq 0}$ and opening costs $f_j \geq 0$ are associated with each potential facility $j \in J$. Assigning a customer $i \in I$ to a facility $j \in J$ yields assignment costs $c_{ij} \geq 0$. Each customer has ranked all potential facilities, inducing a complete, weak ordering of all facilities for every customer. To that end, $j <_i k$ indicates that customer $i \in I$ strictly prefers potential facility $j \in J$ over potential facility $k \in J$, while $j =_i k$ indicates that customer $i \in I$ is indifferent about facilities $j, k \in J$, i.e., neither $j <_i k$ nor $k <_i j$ holds. Similarly, $j \leq_i k$ indicates that customer $i \in I$ either strictly prefers facility $j \in J$ over facility $k \in J$ or is indifferent about them.

A solution to the SSCFLPCP consists of a subset $F \subseteq J$ of facilities that are opened and an assignment $\Lambda : I \to F$ of customers to these facilities, which respects the capacity limits of all open facilities. That is, $\sum_{i \in \Lambda^{-1}(j)} d_i \leq Q_j$ must hold for each $j \in F$ with $\Lambda^{-1}(j)$ the set of customers assigned to facility j. Furthermore, each customer must be assigned to a facility they prefer most among all open facilities, i.e., $\Lambda(i) \leq_i j$ holds for all $i \in I$ and $j \in F$. Note that this property implies that a facility must be closed if the total demand of all customers preferring it over all other open facilities exceeds its capacity. The cost of a solution (F, Λ) equals the total opening and assignment costs, i.e., $\sum_{j \in F} f_j + \sum_{i \in I} c_{i\Lambda(i)}$. The objective of the SSCFLPCP is to find a solution with minimum cost.

The SSCFLPCP is strongly NP-hard as it contains the single-source capacitated facility location problem as a special case when each customer is indifferent about all facilities. For ease of readability, we introduce further notation for preference sets.

Notation In the remainder of this article, we use notations $J_{ij}^{<} = \{k \in J : k <_i j\}$, $J_{ij}^{=} = \{k \in J : k =_i j\}$, $J_{ij}^{\leq} = J_{ij}^{<} \cup J_{ij}^{=}$, $J_{ij}^{\geq} = J \setminus J_{ij}^{<}$, and $J_{ij}^{>} = J \setminus J_{ij}^{\leq}$ for each customer $i \in I$ and facility $j \in J$ to indicate the sets of facilities a customer strictly prefers over j, is indifferent compared to j, prefers at least as much as j, does not prefer more than j, and prefers less than j, respectively.

We use formulation (1) as a basis for our study of preprocessing rules and valid inequalities. The formulation uses decision variables $y_j \in \{0, 1\}$, which indicate whether facility $j \in J$ is opened ($y_j = 1$), and variables $x_{ij} \in \{0, 1\}$, which indicate whether customer $i \in I$ is assigned to facility $j \in J$ ($x_{ij} = 1$).

$$\min \quad \sum_{j \in J} f_j y_j + \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \tag{1a}$$

$$\text{s.t.} \quad \sum_{j \in J} x_{ij} = 1 \qquad \forall i \in I \tag{1b}$$

$$\sum_{i \in I} d_i x_{ij} \leq Q_j y_j \qquad \forall j \in J \tag{1c}$$

$$\sum_{k \in J_{ij}^{>}} x_{ik} + y_j \leq 1 \qquad \forall i \in I, j \in J \tag{1d}$$

$$x_{ij} \leq y_j \qquad \forall i \in I, j \in J \tag{1e}$$

$$x_{ij}, y_j \in \{0, 1\} \qquad \forall i \in I, j \in J \tag{1f}$$

Objective function (1a) minimises the sum of total opening costs and the cost of assigning each customer to their serving facility. Constraints (1b) guarantee that each customer is assigned to a facility. Constraints (1c) ensure that the total demand assigned to an open facility respects its capacity and that customers can only be assigned to open facilities. Constraints (1d), which have first been introduced by Wagner and Falkson (1975), guarantee that customer preferences are respected. More specifically, these constraints ensure that if facility j is open, then customer $i \in I$ cannot be assigned to a facility $k \in J$ which they prefer less than j. Linking constraints (1e) are redundant but well known to significantly strengthen the linear programming relaxation of capacitated facility location problems.

4 Preprocessing

In this section, we introduce preprocessing methods that allow to fix or eliminate variables from formulation (1). Combining arguments related to capacities of facilities and customer preferences, they are specific for the SSCFLPCP. Similar ideas could, however, be used for related problems with capacity constraints in which assignment decisions cannot be (fully) controlled by a central decision maker. The first procedure was observed by Cánovas et al. (2007) for the uncapacitated version with strict preferences and is based on the fact that any open facility $j \in J$ must always serve all customers who prefer j over all other facilities.

Proposition 1 (Cánovas et al. (2007)). *Consider an arbitrary facility $j \in J$ and a customer $i \in I$. If there is no facility customer i prefers over j or is indifferent to, customer i must be assigned to facility j if the facility is open, i.e., inequality $y_j \leq x_{ij}$ holds.*

This preprocessing is also valid for the capacitated problem. When considering capacities, we can further strengthen this result. Proposition 2 states that a facility that is unable to accommodate the demand of all customers who prefer it most has to be kept closed in any feasible solution.

Proposition 2 (*CloseViolatedFacs*). *Consider an arbitrary facility $j \in J$ and the set of customers $I(j) = \{i \in I : J_{ij}^{\leq} = \{j\}\}$ that prefer j over all other facilities. If the total demand of these customers exceeds the capacity of facility j, i.e., $\sum_{i \in I(j)} d_i > Q_j$, then facility j can not be opened. Consequently, variables y_j and x_{ij} are equal to zero for all $i \in I$ in any feasible solution.*

Notice that closing one or more facilities according to this preprocessing, i.e., forcing the associated variables to zero, may increase the demand to be served at other facilities: since the most preferred facilities of some customers are no longer available, such customers will always prefer their second most-preferred facility most among all facilities that may occur in a feasible solution. Consequently, preprocessing based on Proposition 2 should be applied iteratively until no new facilities are eliminated.

Cánovas et al. (2007) also observe a property for the uncapacitated problem that connects allocation variables. Their result works for the case that customers are not indifferent between multiple facilities. In the following, we strengthen their result so that it also allows indifferent customers.

Proposition 3 (Cánovas et al. (2007)). *Consider an arbitrary facility $j \in J$ and two customers $i, \ell \in I$ such that the set of facilities customer ℓ prefers over j or is indifferent to, i.e., $J_{\ell j}^{\leq}$, is subset of the union of facility j with the set of facilities customer i strictly prefers over j, i.e., $J_{\ell j}^{\leq} \subseteq J_{ij}^{<} \cup \{j\}$. Then, customer ℓ must be assigned to facility j if customer i is assigned to it, i.e., inequality $x_{ij} \leq x_{\ell j}$ holds.*

The observations stated in the latter two propositions reveal that assigning one or more customers to a specific facility may imply the assignment of other customers to the same facility. For the capacitated problem, such a set of implied customers also implies a certain demand at this facility. These concepts of *implied customers* and *implied demand* are crucial for the preprocessing rule stated in Proposition 5 and several of the valid inequalities proposed in Section 5. We formally introduce them in Definitions 1 and 2.

Definition 1 (Implied customers). *The set of customers $\mathcal{I}(i, j) \subseteq I$ implied by customer $i \in I$ at facility $j \in J$ is the set of customers that must be assigned to facility $j \in J$ if customer $i \in I$ is assigned to it, i.e., $\mathcal{I}(i, j) = \{i\} \cup \{\ell \in I : J_{\ell j}^{\leq} \subseteq J_{ij}^{<} \cup \{j\}\}$. Similarly, the set of customers $\mathcal{I}(I', j) \subseteq I$ implied by customer set $I' \subseteq I$ at facility $j \in J$ is the set of customers that must be assigned to facility $j \in J$ if all customer $i \in I'$ are assigned to it, i.e., $\mathcal{I}(I', j) = \cup_{i \in I'} \mathcal{I}(i, j)$.*

Definition 2 (Implied demand)**.** *The demand implied by customer $i \in I$ at facility $j \in J$ is the total demand of all customers implied by customer $i \in I$ at facility j, i.e., $\mathcal{D}(i,j) = \sum_{\ell \in \mathcal{I}(i,j)} d_\ell$. Similarly, the demand implied by customer set $I' \subseteq I$ at facility $j \in J$ is the total demand of all customers implied by customer set I' at facility j, i.e., $\mathcal{D}(I',j) = \sum_{\ell \in \mathcal{I}(I',j)} d_\ell$.*

The function defining implied demands is submodular.

Proposition 4. *Consider a facility $j \in J$. Function $\mathcal{D}(\cdot, j) : 2^{|I|} \to \mathbb{Z}_{\geq 0}, X \mapsto \mathcal{D}(X,j)$ is submodular.*

Proof. The function defining the implied demand is submodular if and only if $\mathcal{D}(X \cup \{x\}, j) - \mathcal{D}(X,j) \geq \mathcal{D}(Y \cup \{x\}, j) - \mathcal{D}(Y,j)$ holds for any two subsets $X, Y \subseteq I$ with $X \subseteq Y$ and a customer $x \in I \setminus Y$. (McCormick, 2005). The value of $\mathcal{D}(X \cup \{x\}, j) - \mathcal{D}(X,j)$ is defined by the customers in set $\mathcal{I}(X \cup \{x\}, j) \setminus \mathcal{I}(X,j)$. Writing out this set yields

$$\begin{aligned}
\mathcal{I}(X \cup \{x\}, j) \setminus \mathcal{I}(X,j) &= \cup_{b \in X \cup \{x\}} \left(\{b\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{bj} \cup \{j\}\}\right) \setminus \cup_{b \in X}(\{b\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{bj} \cup \{j\}\}) \\
&=(\{x\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{xj} \cup \{j\}\}) \setminus \cup_{b \in X}(\{b\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{bj} \cup \{j\}\}) \\
&\supseteq(\{x\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{xj} \cup \{j\}\}) \setminus \cup_{b \in Y}(\{b\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{bj} \cup \{j\}\}) \\
&= \cup_{b \in Y \cup \{x\}} \left(\{b\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{bj} \cup \{j\}\}\right) \setminus \cup_{b \in Y}(\{b\} \cup \{\ell \in I : J^{\leq}_{\ell j} \subseteq J^{<}_{bj} \cup \{j\}\}) \\
&=\mathcal{I}(Y \cup \{x\}, j) \setminus \mathcal{I}(Y,j).
\end{aligned}$$

Therefore, inequality $\mathcal{D}(X \cup \{x\}, j) - \mathcal{D}(X,j) \geq \mathcal{D}(Y \cup \{x\}, j) - \mathcal{D}(Y,j)$ is true and the claim follows. □

Proposition 5, which can be seen as an extension of Proposition 3, uses the concept of implied demands to derive conditions under which a customer can never be assigned to a facility.

Proposition 5 (*ImplDem*)**.** *Consider an arbitrary facility $j \in J$ and a customer $i \in I$. If the total demand implied by customer i at facility j exceeds the capacity of j, i.e., $\mathcal{D}(i,j) > Q_j$, then i can not be assigned to j. Consequently, variable x_{ij} is equal to zero in any solution and can be removed from formulation* (1).

The preprocessing rule following from Proposition 6 has been proposed by Cánovas et al. (2007). It considers the case when the sets of facilities two customers i and ℓ prefer over a given facility j are identical. If there are no facilities to which i and ℓ are indifferent to compared to j, then either both of these customers are assigned to j or none of them.

Proposition 6 (Cánovas et al. (2007))**.** *Consider an arbitrary facility $j \in J$ and two customers $i, \ell \in I$. If both customers prefer the same facilities over j, i.e., $J^{<}_{ij} = J^{<}_{\ell j}$, and there does not exist a facility $k \in J$, $k \neq j$, that either i or ℓ is indifferent to with respect to j, i.e., if $J^{=}_{ij} = J^{=}_{\ell j} = \{j\}$, then either both customers i and ℓ are assigned to j or none of them. Consequently, the equation $x_{ij} = x_{\ell j}$ holds and can be added to formulation* (1).

Note that this result is a direct implication of Proposition 3. We study the performance of the various preprocessing methods in Section 7. We will see that method *ImplDem* (Proposition 5) either on its own or together with the methods proposed by Cánovas et al. (2007) performs best if preferences are defined by assignment costs. If preferences are defined by perturbed assignment costs, all introduced methods fix variables or add constraints - yet they do not improve the optimality gap.

5 Cover-based inequalities

In this section, we propose several sets of valid inequalities for the SSCFLPCP that are based on utilising the combination of customer preferences and capacity constraints of facilities. One basic concept used in these inequalities are customer sets whose (implied) demand cannot be served by a particular facility, i.e., (implied-demand) covers, that are formally introduced in Definition 3.

Definition 3 ((Implied-demand) cover)**.** *A* cover *for a facility $j \in J$ is a set of customers $I' \subseteq I$ whose demand exceeds the capacity of j, i.e., $\sum_{i \in I'} d_i > Q_j$. An* implied-demand cover *for a facility $j \in J$ is a set of customers $\tilde{I} \subseteq I$ whose implied demand exceeds the capacity of j, i.e., $\mathcal{D}(\tilde{I}, j) > Q_j$.*

From this definition, we can immediately observe a relation between covers and implied-demand covers.

Proposition 7. *Consider an arbitrary facility $j \in J$. Then, for every cover $I' \subseteq I$ for j there exists an implied-demand cover $\tilde{I}$ for j such that $\tilde{I} \subseteq I'$ holds. An implied-demand cover consists of all customers in set I' whose assignment to j is not implied by the assignment of any other customer in I' to j.*

Proof. Consider a facility $j \in J$ and let set $I' \subseteq I$ be a cover for j. Suppose set I' consists of customers whose assignment to j does not imply the assignment of any other customer in I' to j. Then, it is $\tilde{I} = I'$. Suppose there are two customers $i, \ell \in I'$ and the assignment of i to j implies the assignment of ℓ to j, i.e., $\ell \in \mathcal{I}(i, j)$. Then, set $\tilde{I} = I' \setminus \{\ell\} \subseteq I'$ is an implied-demand cover of j. In general, set $\tilde{I}$ consists of all customers in set I' whose assignment to j is not implied by another customer in I'. □

An (implied-demand) cover $I' \subseteq I$ is called *minimal* if for any customer $i \in I'$ set $I' \setminus \{i\}$ is no longer an (implied-demand) cover. A classical minimal cover can be found in polynomial time: start by considering a cover consisting of all customers and iteratively remove a customer with smallest demand; stop when removing the next customer would yield a set of customers whose demands meet the capacity at the considered facility. A *minimum cover*, i.e., a minimal cover with a minimum number of elements, can be identified in polynomial time by ordering the customers by their demand, and adding the customers starting from the one with most to least demand until there is no capacity left. As Theorem 1 shows, finding a *minimum implied-demand cover* is, on the contrary, strongly NP-hard.

Theorem 1. *It is strongly NP-hard to find a minimum implied-demand cover $\tilde{I} \subseteq I$ for a facility $j \in J$.*

We prove this theorem via a reduction from the strongly NP-hard *minimum set cover problem* (Garey and Johnson, 1979) in the Appendix.

5.1 Cover inequalities

Cover inequalities are known to be valid for the capacitated facility location problem due to the knapsack structure of the capacity constraints (1c). As the considered customer preferences do not impact their validity, classic cover inequalities

$$\sum_{i \in I'} x_{ij} \leq |I'| - 1 \qquad \forall j \in J, I' \subseteq I : \sum_{i \in I'} d_i > Q_j \tag{2}$$

are also valid for the SSCFLPCP. We denote this set of inequalities with **C**. It is not difficult to see that the same arguments as for inequalities (2) can be used to show validity of inequalities (3) that instead consider an implied-demand cover $I' \subseteq I$ for facility $j \in J$:

$$\sum_{i \in I'} x_{ij} \leq |I'| - 1 \qquad \forall j \in J, I' \subseteq I : \mathcal{D}(I', j) > Q_j. \tag{3}$$

We denote this set of inequalities with **IC**. Proposition 7 implies that for each inequality (2) there exists at least one inequality (3) whose set of left-hand side variables is a (potentially identical) subset of the set of left-hand side variables of the former inequality. Thus, Corollary 1 holds.

Corollary 1. *Inequalities* (3) *imply inequalities* (2)*.*

Note that this result implies that the strongest implied-demand cover inequalities derived from a traditional cover I' are those in which no two customers in the implied-demand cover imply each other. In the following, we focus on implied-demand cover inequalities based on such implied-demand covers.

It is well known, that classical cover inequalities can be strengthened to so-called *extended* or *lifted* cover inequalities. Next, we use analogous ideas to lift implied-demand cover inequalities (3). To ease notation and arguments in the proofs of the next two theorems, we first define the concepts of a customer's *individual contribution* and the *maximum individual contribution* to the total demand implied by a set of customers.

Definition 4. *For each set of customers $I' \subseteq I$, every customer $i \in I'$ and facility $j \in J$, we denote by $r_i(I', j) = \mathcal{D}(I', j) - \mathcal{D}(I' \setminus \{i\}, j)$ the* individual contribution *of customer i to the demand implied by I' at facility j. We denote by $r_{\max}(I', j) = \max_{i \in I'} r_i(I', j)$ the* maximum individual contribution *among all customers in set I'.*

The function of individual contributions of a customer $i \in I$ at a facility $j \in J$ is monotonically decreasing.

Proposition 8. *Function $r_i(\cdot, j) : 2^{|I|} \to \mathbb{Z}_{\geq 0}, X \mapsto r_i(X, j)$ defining the individual contribution of a customer $i \in X \subseteq J$ at a facility $j \in J$ is monotonically decreasing.*

Proof. Function $r_i(\cdot, j)$ is monotonically decreasing if and only if $r_i(X, j) \geq r_i(Y, j)$ holds for any $X, Y \subseteq I$ with $i \in X \subseteq Y$ and $j \in J$. We can rewrite inequality $r_i(X, j) \geq r_i(Y, j)$ as

$$r_i(X, j) \geq r_i(Y, j) \Leftrightarrow \mathcal{D}(X, j) - \mathcal{D}(X \setminus \{i\}, j) \geq \mathcal{D}(Y, j) - \mathcal{D}(Y \setminus \{i\}, j)$$

The second inequality holds due to the submodularity of the implied demand function (Proposition 4). □

We define extended implied-demand cover inequalities as follows.

Theorem 2. *Let $I' \subseteq I$ be an implied-demand cover for facility $j \in J$ and let $E \neq \emptyset$ be a set of customers not implied by I', i.e., $E \subseteq I \setminus \mathcal{I}(I', j)$. Let $(s_1, \ldots, s_{|E|})$ be a sequence of customers in E such that $r_{s_\ell}(I' \cup \{s_1, \ldots s_\ell\}, j) \geq \max_{i \in I'} \mathcal{D}(\{i\}, j)$ holds for all $\ell \in \{1, \ldots, |E|\}$. Then, the* extended implied-demand cover inequality (**EIC**)

$$\sum_{i \in I'} x_{ij} + \sum_{i \in E} x_{ij} \leq |I'| - 1 \tag{4}$$

is valid and dominates implied-demand cover inequality (3) *defined on I' and j.*

Proof. Let $I' \subseteq I$ be an implied-demand cover for facility $j \in J$ and let $\emptyset \neq E \subseteq I \setminus \mathcal{I}(I', j)$ be the set of customers not implied by I'. Let $(s_1, \ldots, s_{|E|})$ be a sequence of customers not implied by I' such that $r_{s_\ell}(I' \cup \{s_1, \ldots s_\ell\}, j) \geq \max_{i \in I'} \mathcal{D}(\{i\}, j)$ holds for all $\ell \in \{1, \ldots, |E|\}$. Note that $\ell \geq 1$ since $E \neq \emptyset$.

We prove this theorem by contradiction. Let $\ell \in \{1, \ldots, |E|\}$ be the first index for which the theorem is violated. Then, there exists a set $I'' \subset I' \cup \{s_1, s_2, \ldots, s_\ell\}$ of cardinality $|I'|$ whose implied demand respects the capacity at facility j, i.e., $\mathcal{D}(I'', j) \leq Q_j$. Clearly, $s_\ell \in I''$. Thus, at least one customer $i \in I'$ is not element of set I''. Then, due to Proposition 8 and assumption $r_{s_\ell}(I' \cup \{s_1, \ldots s_\ell\}, j) \geq \max_{i \in I'} \mathcal{D}(\{i\}, j)$,

$$\begin{aligned}
Q_j \geq \mathcal{D}(I'', j) &= \mathcal{D}(I'' \setminus \{s_\ell\}, j) + r_{s_\ell}(I'', j) \\
&= \mathcal{D}(I'' \setminus \{s_\ell\}, j) + r_{s_\ell}(I'', j) + r_i((I'' \cup \{i\}) \setminus \{s_\ell\}, j) - r_i((I'' \cup \{i\}) \setminus \{s_\ell\}, j) \\
&= \mathcal{D}((I'' \cup \{i\}) \setminus \{s_\ell\}, j) + r_{s_\ell}(I'', j) - r_i((I'' \cup \{i\}) \setminus \{s_\ell\}, j) \\
&\geq \mathcal{D}((I'' \cup \{i\}) \setminus \{s_\ell\}, j) + r_{s_\ell}(I' \cup \{s_1, \ldots, s_\ell\}, j) - r_i((I'' \cup \{i\}) \setminus \{s_\ell\}, j) \\
&\geq \mathcal{D}((I'' \cup \{i\}) \setminus \{s_\ell\}, j) + \mathcal{D}(\{i\}, j) - r_i((I'' \cup \{i\}) \setminus \{s_\ell\}, j) \geq \mathcal{D}((I'' \cup \{i\}) \setminus \{s_\ell\}, j)
\end{aligned}$$

holds. Since $|(I'' \cup \{i\}) \setminus \{s_\ell\}| = |I'|$, this contradicts the original assumption that the first $\ell - 1$ customers together with I' form an extended implied-demand cover. □

Next, we introduce a set of lifted implied-demand cover inequalities that are conceptually similar to those known for standard covers (Conforti et al., 2014) and dominate extended implied-demand cover inequalities (4).

Theorem 3. *Consider an implied-demand cover $I' \subseteq I$ for a $j \in J$ and an ordering $(s_1, \ldots, s_{|I \setminus \mathcal{I}(I',j)|})$ of all customers not implied by I'. Then, the* lifted implied-demand cover inequality (**LIC**)

$$\sum_{i \in I'} x_{ij} + \sum_{i=1}^{|I \setminus \mathcal{I}(I',j)|} \alpha_{s_i j} x_{s_i j} \leq |I'| - 1 \tag{5}$$

with

$$\alpha_{s_i j} = |I'| - 1 - \max_{C \subseteq I' \cup \{s_1, \ldots, s_{i-1}\}} \{\sum_{k \in C} \alpha_{kj} : \mathcal{D}(C,j) \leq Q_j - r_{s_i}(C \cup \{s_i\}, j),\ (I' \cup \{s_1, \ldots, s_{i-1}\}) \cap \mathcal{I}(s_i, j) \subseteq C\}$$

for all $i \in \{1, \ldots, |I \setminus \mathcal{I}(I',j)|\}$ is valid and dominates inequality (4) *obtained for the same sequence of customers not in I'.*

Proof. Consider an implied-demand cover $I' \subseteq I$ for a facility $j \in J$ and an ordering $(s_1, \ldots, s_{|I \setminus \mathcal{I}(I',j)|})$, with $|I \setminus \mathcal{I}(I',j)| \geq 1$, of all customers not implied by I'. We prove our claim by contradiction.

Let $\ell \in \{1, \ldots, |I \setminus \mathcal{I}(I',j)|\}$ be the first index for which the theorem is violated, and denote with $I'' \subset I' \cup \{s_1, \ldots, s_\ell\}$ the set of customers for who the lifted implied-demand cover inequality is violated. Clearly, $s_\ell \in I''$. Since I'' violates the lifted implied-demand cover inequality, we have $Q_j \geq \mathcal{D}(I'',j) = \mathcal{D}(I'' \setminus \{s_\ell\}, j) + r_{s_\ell}(I'', j)$ as well as $|I'| \leq \alpha_{s_\ell j} + \sum_{i \in I'' \setminus \{s_\ell\}} \alpha_{ij}$ with $\alpha_{kj} = 1$ for $k \in I'$ and $\alpha_{s_k j}$ with $k \in \{1, \ldots, |I \setminus \mathcal{I}(I',j)|\}$ defined as

$$\alpha_{s_k j} = |I'| - 1 - \max_{C \subseteq I' \cup \{s_1, \ldots, s_{k-1}\}} \big\{ \sum_{k \in C} \alpha_{kj} : \mathcal{D}(C,j) \leq Q_j - r_{s_k}(C \cup \{s_k\}, j),\ (I' \cup \{s_1, \ldots, s_{k-1}\}) \cap \mathcal{I}(s_k, j) \subseteq C \big\}.$$

This is equivalent to

$$\alpha_{s_k j} + \max_{C \subseteq I' \cup \{s_1, \ldots, s_{k-1}\}} \big\{ \sum_{k \in C} \alpha_{kj} : \mathcal{D}(C,j) \leq Q_j - r_{s_k}(C \cup \{s_k\}, j),\ (I' \cup \{s_1, \ldots, s_{k-1}\}) \cap \mathcal{I}(s_k, j) \subseteq C \big\} = |I'| - 1.$$

Since I'' meets the capacity at j, the relation

$$\sum_{i \in I'' \setminus \{s_\ell\}} \alpha_{ij} \leq \max_{C \subseteq I' \cup \{s_1, \ldots, s_{\ell-1}\}} \big\{ \sum_{k \in C} \alpha_{kj} : \mathcal{D}(C,j) \leq Q_j - r_{s_\ell}(C \cup \{s_\ell\}, j),\ (I' \cup \{s_1, \ldots, s_{\ell-1}\}) \cap \mathcal{I}(s_\ell, j) \subseteq C \big\}$$

holds and inequality

$$\begin{aligned} |I'| \leq \alpha_{s_\ell j} + \sum_{i \in I'' \setminus \{s_\ell\}} \alpha_{ij} \leq \alpha_{s_\ell j} + \max_{C \subseteq I' \cup \{s_1, \ldots, s_{\ell-1}\}} \big\{ \sum_{k \in C} \alpha_{kj} : \mathcal{D}(C,j) \leq Q_j - r_{s_\ell}(C \cup \{s_\ell\}, j), \\ (I' \cup \{s_1, \ldots, s_{\ell-1}\}) \cap \mathcal{I}(s_\ell, j) \subseteq C \big\} = |I'| - 1 \end{aligned}$$

holds - a contradiction.

We also observe that $\alpha_{s_\ell j} \geq 1$ holds if $r_{s_\ell}(I' \cup \{s_1, \ldots, s_\ell\}, j) \geq \max_{i \in I'} \mathcal{D}(\{i\}, j)$ holds. Therefore, the resulting lifted implied-demand cover inequality dominates the extended one. □

Value $\alpha_{s_k j}$ corresponds to the number of elements in $I' \cup \{s_1, \ldots, s_{k-1}\}$ that can be replaced in the cover by adding element s_k with $k \in \{1, \ldots, |I \setminus \mathcal{I}(I',j)|\}$ due to s_k's (implied) demand. We discuss a lifting-sequence in Section 6.

5.2 Strengthened implied-demand cover inequalities

We next introduce two additional sets of inequalities (6) and (7) that are based on further observations concerning customer preferences. Notice that we introduce and discuss these inequalities based on implied-demand covers. Similar to above, weaker counterparts based on traditional covers are valid too. The first set of inequalities is defined in Theorem 4.

Theorem 4. *Consider a facility $j \in J$, an implied-demand cover $I' \subseteq I$ for this facility consisting of at least two customers, i.e., $|I'| \geq 2$, and let $\ell \in I'$. Then, implied-demand cover inequality* (3) *can be strengthened to valid inequality*

$$\sum_{i \in I' \setminus \{\ell\}} x_{ij} \leq \sum_{k \in J_{\ell j}^{\leq} \setminus (\cup_{i \in I' \setminus \{\ell\}} J_{ij}^{<} \cup \{j\})} x_{\ell k} + (|I'| - 2) \tag{6}$$

and we denote this set of inequalities with $\mathbf{RemElem}_{IC}$.

Proof. Note that any considered implied-demand cover I' must consist of at least two customers. Suppose cover $I' = \{k\}$ consists of one customer. If facility j is closed and customer k does not prefer any facility over j, inequality (6) turns into $0 \leq 0 + 1 - 2 = -1$, which is infeasible; if facility j is open and i assigned to j, we have $1 \leq 0 + 1 - 2 = -1$, which is infeasible.

In the following, we assume that each implied-demand cover consists of at least two customers. Consider an implied-demand cover $I' \subseteq I$ for facility $j \in J$ and the associated implied-demand cover inequality $\sum_{i \in I'} x_{ij} \leq |I'| - 1$. Let $\ell \in I'$ be an arbitrary customer from I' and recall that every customer must be assigned to some facility, i.e., $\sum_{j \in J} x_{\ell j} = 1$. Using the equation $J = (J_{\ell j}^{\leq} \setminus \{j\}) \cup \{j\} \cup J_{\ell j}^{>}$, we can rewrite the left-hand side of the initial cover inequality as follows:

$$\sum_{i \in I'} x_{ij} = x_{\ell j} + \sum_{i \in I' \setminus \{\ell\}} x_{ij} = 1 - (\sum_{k \in J_{\ell k}^{\leq} \setminus \{j\}} x_{\ell k} + \sum_{k \in J_{\ell k}^{>}} x_{\ell k}) + \sum_{i \in I' \setminus \{\ell\}} x_{ij}.$$

Thus, the initial cover inequality is equivalent to

$$\sum_{i \in I' \setminus \{\ell\}} x_{ij} \leq |I'| - 2 + \sum_{k \in J_{\ell k}^{\leq} \setminus \{j\}} x_{\ell k} + \sum_{k \in J_{\ell k}^{>}} x_{\ell k}$$

which states that all customers in $I' \setminus \{\ell\}$ can be assigned to facility j only if customer ℓ is assigned to another facility.

Observe that customer ℓ cannot be assigned to a facility $k \in J_{\ell j}^{>}$ if j is open and that the cover inequality is redundant if j is closed - in which case the left-hand side is equal to zero. Thus, all assignment variables considering facilities in $J_{\ell j}^{>}$ can be removed from the right-hand side leading to the stronger inequality

$$\sum_{i \in I' \setminus \{\ell\}} x_{ij} \leq |I'| - 2 + \sum_{k \in J_{\ell k}^{\leq} \setminus \{j\}} x_{\ell k}.$$

Further notice that the variable term on the right-hand side is only relevant for the validity of the inequality if all customers in $I' \setminus \{\ell\}$ are assigned to facility j. Otherwise the left-hand side value is at most $|I'| - 2$. In this case, no facility that is preferred over j by any of these customers may be open and the previous inequality can therefore be further strengthened to

$$\sum_{i \in I' \setminus \{\ell\}} x_{ij} \leq |I'| - 2 + \sum_{k \in J_{\ell j}^{\leq} \setminus (\cup_{i \in I' \setminus \{\ell\}} J_{ij}^{<} \cup \{j\})} x_{\ell k}.$$

□

The following corollary can be shown by repeating the steps in the proof of Theorem 4 using a standard cover instead of an implied-demand cover as starting point.

Corollary 2. *A set of valid inequalities is obtained from inequalities* (6) *by considering covers instead of implied-demand covers. The resulting set of inequalities dominates cover inequalities* (2).

Further utilising the idea of removing customers included in an (implied-demand) cover from the left-hand side of a strengthened cover inequality (6), we can also derive the next result.

Theorem 5. *Consider a facility $j \in J$, an implied-demand cover $I' \subseteq I$ for this facility and let $i \in I'$. Then, implied-demand cover inequality* (3) *can be strengthened to valid inequality*

$$x_{ij} \leq \sum_{\ell \in I' \setminus \{i\}} \sum_{a \in J^{\leq}_{\ell j} \setminus (J^{<}_{ij} \cup \{j\})} x_{\ell a} \tag{7}$$

and we denote this set of inequalities with $\mathbf{RemAll}_{IC}$.

Proof. Note that inequality (7) is valid for implied-demand covers of size one: in this case, the inequality reduces to $x_{ij} \leq 0$, which is valid since i can not be assigned to j as its implied demand is too high.

Let $\sum_{i \in I'} x_{ij} \leq |I'| - 1$ be the implied-demand cover inequality associated to implied-demand cover I'. We substitute all left-hand side variables $x_{\ell j}$ corresponding to $\ell \in I' \setminus \{i\}$ for some $i \in I'$ by $1 - (\sum_{k \in J^{\leq}_{\ell j} \setminus \{j\}} x_{\ell k} + \sum_{k \in J^{>}_{\ell j}} x_{\ell k})$ and obtain

$$x_{ij} \leq \sum_{\ell \in I' \setminus \{i\}} \sum_{a \in J^{\leq}_{\ell j} \setminus (\{j\})} x_{\ell a} + \sum_{\ell \in I' \setminus \{i\}} \sum_{a \in J^{>}_{\ell j}} x_{\ell a}.$$

Like in the previous proof, no customer $\ell \in I' \setminus \{i\}$ can be assigned to a facility $k \in J^{>}_{\ell j}$ if j is open and the cover inequality is redundant if j is closed. Thus, we can, again, remove the facilities in $J^{>}_{\ell j}$, for $\ell \in I' \setminus \{i\}$, from the right hand side. Furthermore, the term on the right-hand side is only relevant for the validity of the inequality if customer i is assigned to facility j. In this case, no facility that i strictly prefers over j may be open. The combination of these two arguments leads to the stronger inequality

$$x_{ij} \leq \sum_{\ell \in I' \setminus \{i\}} \sum_{a \in J^{\leq}_{\ell j} \setminus (J^{<}_{ij} \cup \{j\})} x_{\ell a}.$$

□

Theorem 6 analyses the impact of removing customers included in an (implied-demand) cover one by one from the left-hand side of a cover-like inequality.

Theorem 6. *Consider a facility $j \in J$, an implied-demand cover $I' \subseteq I$ and a customer $i \in I'$. Let $\nu = (\pi(I' \setminus \{i\}), i) \in I^{|I|}$ be an ordering of customers in I' and $\pi : (i)_{i \in I} \mapsto \pi((i)_{i \in I})$ a permutation. Then,*

$$x_{ij} \leq \sum_{m=2}^{|I'|} \sum_{\substack{a \in J^{\leq}_{\nu(m-1)j} \setminus \\ (\cup_{i'=\nu(m)}^{\nu(|I'|)} J^{<}_{i'j} \cup \{j\})}} x_{\nu(m-1)a} \tag{8}$$

is valid and dominates inequalities (6) and (7). We denote this set of inequalities with $\boldsymbol{RemOBO}_{IC}$.

Proof. We first observe, that inequalities (8) are valid for implied-demand covers of size one in which case they reduce to $x_{ij} \leq 0$, which is clearly valid since customer i cannot be assigned to this facility as its implied demand is too high.

Thus, we restrict our attention to implied-demand covers $I' \subseteq I$ at facilities $j \in J$ such that $|I'| \geq 2$. In this case inequality (6)

$$\sum_{i \in I' \setminus \{\ell\}} x_{ij} \leq \sum_{k \in J^{\leq}_{\ell j} \setminus (\cup_{i \in I' \setminus \{\ell\}} J^{<}_{ij} \cup \{j\})} x_{\ell k} + (|I'| - 2)$$

is valid for any $\ell \in I'$, cf. Theorem 4. Next, we use use relation $x_{kj} = 1 - \sum_{a \in J^{\leq}_{kj} \setminus \{j\}} x_{ka} - \sum_{a \in J^{>}_{kj}} x_{ka}$ to substitute x_{kj} for a $k \in I' \setminus \{\ell\}$. Rearranging the terms, we obtain

$$\sum_{i \in I' \setminus \{k, \ell\}} x_{ij} \leq \sum_{a \in J^{\leq}_{\ell j} \setminus (\cup_{i' \in I' \setminus \{\ell\}} J^{<}_{i'j} \cup \{j\})} x_{\ell k} + (|I'| - 2) - (1 - \sum_{a \in J^{\leq}_{kj} \setminus \{j\}} x_{ka} - \sum_{a \in J^{>}_{kj}} x_{ka})$$

$$= \sum_{a \in J^{\leq}_{\ell j} \setminus (\cup_{i' \in I' \setminus \{\ell\}} J^{<}_{i'j} \cup \{j\})} x_{\ell a} + (|I'| - 3) + \sum_{a \in J^{\leq}_{kj} \setminus \{j\}} x_{ka} + \sum_{a \in J^{>}_{kj}} x_{ka}.$$

We observe that the rightmost term $\sum_{a \in J^{>}_{kj}} x_{ka}$ can only be non-zero if facility j is closed as these assignment variables all relate to assigning some customer to a facility customer k prefers less than j. Thus, we can drop this term from the inequality as the left-hand side $\sum_{i \in I' \setminus \{k,\ell\}} x_{ij}$ must be equal to zero if this rightmost expression is greater than zero. The inequality above is only of interest if all customers $i \in I' \setminus \{k, \ell\}$ are assigned to j. Thus, all facilities such customers strictly prefer over j have to be closed and we may exclude all assignments of k to such facilities in the term $\sum_{a \in J^{\leq}_{kj} \setminus \{j\}} x_{ka}$. Then, we can strengthen this term to $\sum_{a \in J^{\leq}_{kj} \setminus (\cup_{i' \in I' \setminus \{k,\ell\}} J^{<}_{i'j} \cup \{j\})} x_{ka}$, leading to

$$\sum_{i \in I' \setminus \{k,\ell\}} x_{ij} \leq (|I'| - 3) + \sum_{a \in J^{\leq}_{\ell j} \setminus (\cup_{i' \in I' \setminus \{\ell\}} J^{<}_{i'j} \cup \{j\})} x_{\ell a} + \sum_{a \in J^{\leq}_{kj} \setminus (\cup_{i' \in I' \setminus \{k,\ell\}} J^{<}_{i'j} \cup \{j\})} x_{ka}$$
$$= (|I'| - 3) + \sum_{m=2}^{3} \sum_{a \in J^{\leq}_{\nu(m-1)j} \setminus (\cup_{i'=\nu(m)}^{\nu(|I'|)} J^{<}_{i'j} \cup \{j\})} x_{\nu(m-1)a}$$

for $\nu = (\ell, k, \pi(I' \setminus \{k, \ell\}))$ an ordering of the customers with customers ℓ, k at the first and second position. Repeating this procedure for $|I'| - 3$ further elements and defining ordering ν according to the order the respective elements are moved to the right-hand side yields the claim.

It follows immediately that inequality (8) dominates inequalities (6) and (7). □

We can also apply this procedure to extended implied-demand cover inequalities. Consider an extended implied-demand cover I', which is an extension of an implied-demand cover $\tilde{I}$ with $\tilde{I} \subsetneq I'$. Then, we can move at most $|\tilde{I}| - 1$ elements from the left-hand side to the right-hand side in order to strengthen implied-demand cover inequalities (3). Otherwise, assigning no customer in extended implied-demand cover I' to facility j would result in inequality $0 \leq -1$.

Finally, we consider another approach to utilise information provided by covers.

5.3 Location-centered inequalities

In the following, we study a different approach to derive valid inequalities from (implied-demand) covers. First, consider a cover $I' \subseteq I$ for some facility $j \in J$. From the definition of a cover it is immediate that facility j can only be opened if at least one customer from I' is assigned to a different facility. The latter is, however, only possible if at least one facility that is preferred by some customer from I' over j or to which such a customer is indifferent is opened too. Thus, the set of inequalities

$$y_j \leq \sum_{k \in (\cup_{i \in I'} J^{\leq}_{ij}) \setminus \{j\}} y_k \qquad \forall j \in J,\ I' \subseteq I : \sum_{i \in I'} d_i > Q_j \tag{9}$$

is valid. We denote this set of inequalities with ***ImplFac***. It is not difficult to see, that the same arguments as for inequalities (9) can be used to show validity of inequalities (10) that instead consider an implied-demand cover $I' \subseteq I$ for facility $j \in J$:

$$y_j \leq \sum_{k \in (\cup_{i \in I'} J^{\leq}_{ij}) \setminus \{j\}} y_k \qquad \forall j \in J,\ I' \subseteq I : \mathcal{D}(I', j) > Q_j. \tag{10}$$

For this family of valid inequalities, the consideration of implied-demands does not provide an advantage.

Proposition 9. *Inequalities* (9) *and inequalities* (10) *are equivalent.*

Proof. Consider a valid inequality (9) based on cover C. Define a set C' which consists of all non-implied customers in C. Then, it is $C \subseteq \mathcal{I}(C', j)$ and set C' is an implied-demand cover. The right-hand side of inequalities (9) and (10) is defined by the preference sets of non-implied customers. Hence, the right-hand sides of both inequalities coincide.

Conversely, consider an inequality (10) based on implied-demand cover C'. Build a solution C for inequalities (9) as follows: set $C = \mathcal{I}(C', j)$. Obviously, set C is still a cover. Due to the definition of implied customers, it is $\cup_{i \in C} J_{ij}^{\leq} = \cup_{k \in C'} J_{kj}^{\leq}$ and the value on the right-hand side coincides for implied-demand cover C' and cover C. □

We only consider inequalities (9) instead of inequalities (10) since the former inequalities do not demand the computation of implied-demand covers - which is likely to be more complex than the computation of traditional covers. We study the performance of all valid inequalities derived above in Section 7.

6 Separation

For evaluating the impact of the valid inequalities above, we have to solve the corresponding separation problems. Since inequalities **C**, **EC** and **LC** are already well studied in the literature, we mainly focus on the remaining inequalities. Throughout this section, we assume that $(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \in [0,1]^{|I||J|+|J|}$ is the current solution to the LP relaxation of (1).

6.1 Implied-demand cover inequalities

In the following, we first provide an integer programming formulation that solves the separation problem for implied-demand cover inequalities. Then, we introduce heuristics for extending an implied demand-cover inequality to an extended and a lifted one, respectively.

Formulation (11) solves the separation problem for implied-demand cover inequalities (3) for each facility $j \in J$, i.e., it identifies an implied-demand cover in which no two customers imply each other. We consider binary decision variables $z_i, u_i \in \{0,1\}$. Variable z_i is equal to one if customer i is either in the implied-demand cover or implied by an element of it ($z_i = 0$ otherwise); variable u_i is equal to one if i is implied by another customer in the implied-demand cover ($u_i = 0$ otherwise). Then, a customer $i \in I$ is in an implied-demand cover where no two customers imply one another if $z_i = 1$ and $u_i = 0$.

$$\min \quad \sum_{i \in I} (1 - \bar{x}_{ij}) \cdot (z_i - u_i) \tag{11a}$$

$$\text{s.t.} \quad \sum_{i \in I} d_i z_i \geq Q_j + 1 \tag{11b}$$

$$z_i \leq z_k \qquad \forall i, k \in I : k \in \mathcal{I}(i, j) \tag{11c}$$

$$u_k \leq \sum_{i \in I : k \in \mathcal{I}(i,j) \setminus \{i\}} (z_i - u_i) \qquad \forall k \in I \tag{11d}$$

$$z_i, u_i \in \{0,1\} \qquad \forall i \in I \tag{11e}$$

Objective function (11a) maximises the violation of inequality $\sum_{i \in I'} \bar{x}_{ij} \leq |I'| - 1 \Leftrightarrow 1 \leq \sum_{i \in I'} (1 - \bar{x}_{ij})$ at a facility $j \in J$. In the separation problem, our goal is to determine an implied-demand cover which maximises the total violation of this inequality. We achieve this by determining an implied-demand cover which minimises the right-hand side of the reformulated implied-demand cover inequality, i.e., $\sum_{i \in I'} (1 - \bar{x}_{ij})$. If the objective value is strictly lower than one, we have detected a violated inequality. Inequality (11b) guarantees that the implied-demand of the found solution exceeds the capacity at j. Inequalities (11c) ensure that implied demands are considered, i.e., if a customer is considered in the cover, we also consider the demands of all customers implied by said customer in inequality (11b). Constraints (11d) ensure that no two customers in the computed implied-demand cover imply one another. In order to better understand this inequality, consider two cases. First, suppose that it is already decided that customer $i \in I$ occurs in the implied-demand cover. Let the assignment of a customer $k \in I \setminus \{i\}$ to j be implied by i's assignment to j. Then, it is $k \in \mathcal{I}(i,j)$ as well as $z_i = 1$ and $u_i = 0$. Hence, the right-hand side of inequality (11d) is greater than or equal to one with u_k on the left-hand side. Assigning variable u_k the value of one decreases the objective value and is therefore advantageous. Second, suppose that customer $k \in I$ is not implied by another customer in the implied-demand cover. In this case, there is no customer $i \in I \setminus \{k\}$ with $k \in \mathcal{I}(i,j)$ and $z_i = 1 - u_i = 1$. Therefore, the right-hand side of inequality (11d) is equal to zero for customer k and variable u_k is also equal to zero. In conclusion, inequalities (11d) ensure that variable u_k is equal to one if customer k is not an element of the implied-demand cover but implied by at least one customer occuring in it; variable u_k is equal to zero either if k's demand is not considered in the implied-demand cover at all or if k occurs in the implied-demand cover.

In order to compute a classical cover inequality, remove inequalities (11d) and set $u_i = 0$ for all $i \in I$.

The following result holds for the computational complexity of finding a maximum violated inequality (3).

Theorem 7. *It is strongly NP-hard to find an implied-demand cover $I' \subseteq I$ for a facility $j \in J$ that maximises the violation $\sum_{i \in I'} x_{ij} - |I'| + 1$ of inequality* (3) *for optimal solutions $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in [0,1]^{|I||J|+|J|}$ to the linear relaxation of* (1)*.*

We prove this theorem in the Appendix. Note that this result implies that it is also strongly NP-hard to separate inequalities $\mathbf{EIC}_{IC}$, $\mathbf{LIC}_{IC}$, $\mathbf{RemElem}_{IC}$, $\mathbf{RemAll}_{IC}$ and **RemOBO**.

Extended implied-demand cover inequalities We separate extended implied-demand cover inequalities heuristically by a two-phase approach. In the first phase, we identify a maximum violated implied-demand cover I' for each facility $j \in J$ by solving the corresponding separation problem with formulation (11). Among all customers in implied-demand cover I', we determine the maximum implied-demand to j. In the second phase, we determine the customers in the extended implied-demand cover in a greedy manner. Denote with I' the implied-demand cover we aim to extend, and denote the set of customers which may occur in the extended implied-demand cover as *customers of interest*; from Theorem 2, the set of customers of interests consists of all customers in $I \setminus \mathcal{I}(I', j)$ with an individual contribution greater than or equal to the maximum implied-demand among the customers in I'. The next customer added to the extended implied-demand cover is the customer with smallest individual contribution among all remaining customers of interest.

This approach is similar to the standard procedure used to determine traditional extended covers; they coincide if there is no customer implying the assignment of another customer to the same facility. This rule is, however, just an heuristic. Suppose the implied demand of a customer $i \in I$ at facility j is the lowest possible so that i is still a customer of interest; let i's allocation variable value $\bar{x}_{ij}$ be also very small. Suppose that customer i is added to the extended implied-demand cover. Now, considering i in the extended implied-demand cover might decrease the individual contribution to the extended implied-demand cover of another customer of interest $k \in I \setminus I'$ with a greater allocation variable value to j than $\bar{x}_{ij}$. This might result in a solution where considering i does not violate extended implied-demand cover inequality (4) while considering k would. This wrong choice potentially weakens inequalities **EIC**.

Once the next customer to be added to the extended implied-demand cover is chosen, we update the set of customers of interest by excluding demands of customers implied by customers in the current extended implied-demand cover and repeat this procedure. We stop when no new customers are found.

Lifted implied-demand cover inequalities We proceed with a heuristic for solving the separation problem of inequalities **LIC**. First, we compute an implied-demand cover I' at a facility $j \in J$ which violates implied-demand cover inequalities (3) most by solving separation problem (11). Next, consider all customers who neither occur in I' nor are implied by a customer in I', i.e., all customers in $I \setminus \mathcal{I}(I', j)$. Compute a lifting coefficient for each of these customers according to the formula in Theorem 3. For that, we have to define a lifting-sequence H and then compute value

$$\alpha_{pj} = \max_{C \subseteq I' \cup \{1,\dots,p-1\}} \{\sum_{k \in C} \alpha_{kj} : \mathcal{D}(C,j) \leq Q_j - r_p(C \cup \{p\}, j),\, (I' \cup \{1,\dots,p-1\}) \cap \mathcal{I}(p,j) \subseteq C\}$$

for each $p \in H$. We derive sequence H based on the number of implied customers at facility j for each customer $i \in I \setminus \mathcal{I}(I',j)$. The customer whose assignment to j implies most customers is considered as the last element in the lifting sequence; the customer implying the least number of customers is considered as the first element in the sequence. The motivation for this procedure raises from the observation that the greater the number of customers implied by a customer in $I \setminus \mathcal{I}(I',j)$ is, the greater might be set $(I' \cup \{1,\dots,p-1\}) \cap \mathcal{I}(p,j)$. Per definition of the lifting coefficients, we must consider all customers in $(I' \cup \{1,\dots,p-1\}) \cap \mathcal{I}(p,j)$ in the corresponding maximisation problem. This forced consideration of customers might block better solutions in the maximisation problem.

We compute the desired lifting coefficients α_{pj} for a customer $p \in I \setminus \mathcal{I}(I',j)$ in sequence H and a facility $j \in J$ via integer linear program (12). Binary decision variables $z_i \in \{0,1\}$ indicate whether customer $i \in I' \cup \{1,2,\dots,p-1\}$ is assigned to facility j ($z_i = 1$) or not ($z_i = 0$). Binary decision variables $u_k \in \{0,1\}$ indicate whether customer $k \in \mathcal{I}(I' \cup \{1,2,\dots,p-1\},j)$ is implied by a customer $i \in I' \cup \{1,2,\dots,p-1\}$ considered in the solution ($u_k = 1$) or not ($u_k = 0$). Based on lifting-sequence H, we solve the following integer linear program for each element $p \in H$.

$$\max \quad \sum_{i \in I' \cup \{1,2,\dots,p-1\}} \alpha_{ij} z_i \tag{12a}$$

$$\text{s.t.} \quad \sum_{\substack{i \in \mathcal{I}(I' \cup \{1,2,\dots,p-1\},j) \setminus \\ \mathcal{I}(p,j)}} d_i u_i \leq Q_j - \mathcal{D}(p,j) \tag{12b}$$

$$z_i \leq u_k \quad \forall i \in I' \cup \{1,2,\dots,p-1\}, k \in \mathcal{I}(i,j) \tag{12c}$$

$$z_i = 1 \quad \forall i \in (I' \cup \{1,2,\dots,p-1\}) \cap \mathcal{I}(p,j) \tag{12d}$$

$$z_i, u_k \in \{0,1\} \quad \forall i \in I' \cup \{1,2,\dots,p-1\}, k \in \mathcal{I}(I' \cup \{1,2,\dots,p-1\},j). \tag{12e}$$

Objective function (12a) sums up the weighted value of all decision variables in sequence H that are considered before element p; the goal is to find a subset of customers in $I' \cup \{1,2,\dots,p-1\}$ that maximises this sum. Constraint (12b) combined with constraints (12c) limit the number of variables z_i that can be set to one. That is, they ensure that the joint contribution of customers from set $(I' \cup \{1,2,\dots,p-1\}) \setminus \mathcal{I}(p,j)$ that are considered in the solution does not exceed the remaining capacity of facility j after assigning customer p to j. Constraints (12d) ensure that each customer in set $I' \cup \{1,2,\dots,p-1\}$ that is implied by customer p is considered in the solution as well. Given a solution (z^*, u^*) with objective value $c(z^*, u^*)$ to this integer linear program, set coefficient $\alpha_{pj} = |I'| - 1 - c(z^*, u^*)$.

6.2 Strengthened implied-demand cover inequalities

In the following, we consider methods to solve the separation problems for inequalities $\mathbf{RemElem}_{IC}$, $\mathbf{RemAll}_{IC}$ as well as $\mathbf{RemOBO}_{IC}$ and their versions without implied demands. We start with inequalities $\mathbf{RemElem}_{IC}$.

Instead of inequalities (6) for $\mathbf{RemElem}_{IC}$, we consider reformulation $1 \leq \sum_{i \in I' \setminus \{l\}} (1 - x_{ij}) + \sum_{k \in J^{\leq}_{lj} \setminus (\cup_{i \in I' \setminus \{l\}} J^{\leq}_{ij} \cup \{j\}} x_{lk}$. The separation problem for finding a (maximum) violated reformulated inequality $\mathbf{RemElem}_{IC}$ for a facility $j \in J$ is stated in (13). Binary decision variables $z_i \in \{0,1\}$ indicate whether the demand of customer $i \in I$ is considered in the implied-demand cover ($z_i = 1$) or not ($z_i = 0$). Binary decision variables $u_i \in \{0,1\}$ state whether customer i is implied by a customer in the cover ($u_i = 1$) or not ($u_i = 0$). We model the decision regarding which customer is to be considered on the right-hand side of inequalities (6) through binary decision variables $\hat{w}_i, w_{ik}, \tilde{w}_k \in \{0,1\}$ for $i \in I$ and $k \in J$. Variable $\hat{w}_i$ indicates whether customer $i \in I$ is considered on the right-hand side ($\hat{w}_i = 1$) or not ($\hat{w}_i = 0$). Variable w_{ik} indicates whether customer i is considered on the right-hand side and prefers a facility k, that is not strictly preferred by a customer remaining on the left-hand side, over j ($w_{ik} = 1$) or not ($w_{ik} = 0$). Variable $\tilde{w}_k$ indicates whether there is another customer in the cover besides the customer on the right-hand side who strictly prefers facility k over j ($\tilde{w}_k = 1$) or not ($\tilde{w}_k = 0$).

$$\min \quad \sum_{i\in I}(1-\bar{x}_{ij})\cdot(z_i - u_i - \hat{w}_i) \quad + \quad \sum_{i\in I}\sum_{k\in J}\bar{x}_{ik}w_{ik} \tag{13a}$$

$$\text{s.t.} \quad \sum_{i\in I} d_i z_i \geq Q_j + 1 \tag{13b}$$

$$z_i \leq z_k \qquad \forall i \in I, k \in \mathcal{I}(i,j) \tag{13c}$$

$$u_k \leq \sum_{i\in I: k\in\mathcal{I}(i,j)\setminus\{i\}} (z_i - u_i) \qquad \forall k \in I \tag{13d}$$

$$\sum_{i\in I}\hat{w}_i = 1 \tag{13e}$$

$$\hat{w}_i + u_i \leq z_i \qquad \forall i \in I \tag{13f}$$

$$\hat{w}_i \leq w_{ik} + \tilde{w}_k \qquad \forall i \in I, k \in J^{\leq}_{ij} \setminus \{j\} \tag{13g}$$

$$\tilde{w}_k \leq \sum_{i\in I: k\in J^{<}_{ij}} (z_i - u_i - \hat{w}_i) \qquad \forall k \in J \tag{13h}$$

$$2 \leq \sum_{i\in I}(z_i - u_i) \tag{13i}$$

$$z_i, u_i, \hat{w}_i, w_{ik}, \tilde{w}_k \in \{0,1\} \qquad \forall i \in I, k \in J. \tag{13j}$$

In the objective function (13a), we add $1 - \bar{x}_{ij}$ to the objective value for each non-implied customer $i \in I$ in computed cover that is on the left-hand side of inequality (6). Furthermore, we add value $\bar{x}_{\ell k}$ for customer $\ell \in I$ that is considered on the right-hand side of the inequality and all facilities only customer ℓ prefers over j, i.e., facilities $k \in J^{\leq}_{\ell j} \setminus (\cup_{i\in I'\setminus\{\ell\}} J^{<}_{ij} \cup \{j\})$. Constraint (13b) combined with constraints (13c) ensure that the implied-demand of the computed cover violates the capacity of facility j. Constraints (13d) indicate whether a customer is a non-implied customer in the cover or an implied customer. Constraint (13e) ensures that exactly one customer is considered on the right-hand side. Constraints (13f) guarantee that a customer can either be implied or be considered on the right-hand side or none of both. In the latter case, the corresponding customer is considered either in the cover or not at all. Inequalities (13g) and (13h) enforce that the correct allocation variable values are considered in the objective function in order to model the sum of the right-hand side in inequality (6). That is, given that customer ℓ is considered on the right-hand side, we add $\bar{x}_{\ell k}$ to the objective function value if no other customer who is considered in the solution strictly prefers k over j. Last but not least, constraint (13i) ensure that the implied-demand cover consists of at least two customers.

There is a violated inequality if the objective value is strictly lower than one. When solving the separation problem of inequalities **RemElem** without implied-demands, remove constraints (13c) and (13d) and set $u_i = 0$ for all $i \in I$.

We derive the separation problem for finding a maximum violated inequality $\mathbf{RemAll}_{IC}$ by reformulating inequalities (7) in a similar manner to what we did with inequalities (6). We aim to solve integer linear program (14) for each facility $j \in J$. We consider binary decision variables $z_i, u_i, \hat{w}_i, w_{ik}, \tilde{w}_k \in \{0,1\}$ for $i \in I$ and $k \in J$. Like in formulation (13), binary decision variable $z_i \in \{0,1\}$ indicates whether the demand of customer $i \in I$ is considered in the implied demand cover ($z_i = 1$) or not ($z_i = 0$); binary decision variable $u_i \in \{0,1\}$ indicates whether customer $i \in I$ is implied by a customer in the cover ($u_i = 1$) or not ($u_i = 0$). Binary decision variable $\hat{w}_i \in \{0,1\}$ indicates whether customer $i \in I$ is considered on the left-hand side ($\hat{w}_i = 1$) or not ($w_i = 0$). Decision variable w_{ik} indicates whether customer $i \in I$ is considered on the right-hand side in inequality (7) and prefers facility $k \in J$ over j while the customer remaining on the left-hand side does not strictly prefer k over j ($w_{ik} = 1$) or not ($w_{ik} = 0$). Decision variable $\tilde{w}_k \in \{0, 1\}$ indicates whether facility $k \in J$ is strictly preferred by the customer remaining on the left-hand side of inequality (7) ($\tilde{w}_k = 1$) or not ($\tilde{w}_k = 0$).

$$
\begin{aligned}
\min \quad & \sum_{i\in I}(1-\bar{x}_{ij})\cdot\hat{w}_i & + & \sum_{i\in I}\sum_{k\in J^{\leq}_{ij}\setminus\{j\}} \bar{x}_{ik}w_{ik} & & \text{(14a)}\\
\text{s.t.} \quad & \sum_{i\in I} d_i z_i & \geq & \; Q_j + 1 & & \text{(14b)}\\
& z_i & \leq & \; z_k & \forall i\in I, k\in\mathcal{I}(i,j) & \quad \text{(14c)}\\
& u_k & \leq & \sum_{i\in I: k\in\mathcal{I}(i,j)\setminus\{i\}} (z_i - u_i) & \forall k\in I & \quad \text{(14d)}\\
& \sum_{i\in I}\hat{w}_i & = & \; 1 & & \text{(14e)}\\
& \hat{w}_i + u_i + w_{ik} & \leq & \; z_i & \forall i\in I, k\in J^{\leq}_{ij}\setminus\{j\} & \quad \text{(14f)}\\
& z_i & \leq & \; \hat{w}_i + u_i + w_{ik} + \tilde{w}_k & \forall i\in I, k\in J^{\leq}_{ij}\setminus\{j\} & \quad \text{(14g)}\\
& \tilde{w}_k & \leq & \sum_{i\in I: k\in J^{<}_{ij}} \hat{w}_i & \forall k\in J & \quad \text{(14h)}\\
& z_i, u_i, \hat{w}_i, w_{ik}, \tilde{w}_k & \in & \; \{0,1\} & \forall i\in I, k\in J. & \quad \text{(14i)}
\end{aligned}
$$

Constraints (14b) – (14e) coincide with constraints (13b) – (13e) in formulation (13). Constraint (14e), however, indicates here that a customer stays on the left-hand side in inequality (7).

In the objective function (14a), we consider value $1 - \bar{x}_{ij}$ if customer $i \in I$ represents a customer on the left-hand side of valid inequality (7); we add value $\bar{x}_{ik}$ for facilities $k \in J^{\leq}_{ij} \setminus \{j\}$ to the objective function value if i represents a customer on the right-hand side of inequality (7) and the customer on the left-hand side does not strictly prefer k over j. Constraints (14f) ensure that a customer $i \in I$ can be considered either on the left-hand side of inequality (7), or as an implied customer, or their assignment to facility $k \in J$ is of interest if their demand is considered in the implied-demand cover, or they are not considered in the cover at all. Constraints (14g) enforce that a customer $i \in I$ whose demand is considered in the implied-demand cover is either considered on the left-hand side of the valid inequality, or is an implied customer, or their assignment to facility $k \in J^{\leq}_{ij}$ is considered, or facility k is strictly preferred by the customer on the left-hand side. Constraints (14h) allow the assignment of value one to variable $\tilde{w}_k$ for $k \in J$ if facility k is strictly preferred by the customer considered on the left-hand side of valid inequality (7).

There is a violated inequality if the objective value is strictly lower than one. When considering the separation problem of inequalities **RemAll** without implied demands, remove constraints (14c) and (14d) and set $u_i = 0$ for all $i \in I$.

The complexity of separating inequality (8) increases compared to separating inequalities (6) and (7): given the decision of which customers are considered on the right-hand side in inequality (8), we are still in need of an ordering of these customers - conversely to inequality (7). Due to this complexity, we solve the separation problem for inequalities $\mathbf{RemOBO}_{IC}$ heuristically. The heuristic proceeds as follows.

1. Compute an implied-demand cover $I' \subseteq I$ with maximum violation of inequality $\mathbf{RemAll}_{IC}$ for solution $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ by solving formulation (14). Let $\ell \in I'$ be the customer on the left-hand side of the inequality.

2. Compute for each customer $i \in I' \setminus \{\ell\}$, i.e., all customers on the right-hand side in inequality $\mathbf{RemAll}_{IC}$, the sum of the allocation variable values of i to facilities they prefer over j or are

indifferent to according to solution $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$.

3. Consider the customers in decreasing order according to the values computed in the previous step. This decision enforces that customers with most of their demand assigned to facilities they prefer over j or are indifferent to are considered early in the ordering. If a customer $i \in I'$ is considered early in the ordering, the set of facilities for which the allocation variable value of i is ignored on the right-hand side of inequality (8) is greater than or equal to the set of facilities to be ignored for customers at a later point in the ordering.

4. Check whether this newly constructed valid inequality is violated.

If we do not consider implied demands, we solve the separation problem of inequalities **RemAll** in step 1. The remaining steps coincide with the heuristic approach for separating inequalities **RemOBO**$_{IC}$.

6.3 Location-centered inequalities

Finally, we study the separation problem of inequalities **ImplFac**. To identify a violated inequality (9), we are interested in finding the most violated one, i.e., the inequality that minimises $\sum_{k\in(\cup_{i\in I'} J^{\leq}_{ij})\setminus\{j\}} \bar{y}_k - \bar{y}_j$ with I' a cover. If this minimum is non-negative, we know that there are no violated inequalities.

In order to compute a most violated inequality, we solve integer linear program (15) for each facility $j \in J$. Binary decision variable z_i indicates whether customer $i \in I$ is either considered in the cover or the assignment of i to j is implied by a customer in the cover ($z_i = 1$) or not ($z_i = 0$). Binary decision variable v_k indicates whether at least one customer $i \in I$ in the computed cover prefers $k \in J$ over j ($v_k = 1$) or not ($v_k = 0$).

$$
\begin{aligned}
\min \quad & \sum_{k\in J\setminus\{j\}} \bar{y}_k v_k && && \text{(15a)}\\
\text{s.t.} \quad & \sum_{i\in I} d_i z_i && \geq \quad Q_j + 1 && && \text{(15b)}\\
& z_i && \leq \quad v_k && \forall i \in I, k \in J^{\leq}_{ij} \setminus \{j\} && \text{(15c)}\\
& v_k, z_i && \in \quad \{0,1\} && \forall k \in J, i \in I. && \text{(15d)}
\end{aligned}
$$

Inequality (15b) enforces that the computed set is indeed a cover. Inequalities (15c) ensure that the allocation value of a facility is counted in the objective function if the facility is preferred by at least one customer in the computed cover. Objective function (15a) minimises the total sum of all values assigned to location variables in a solution to the relaxation of integer linear program (1) which are preferred by customers in the computed cover.

Separating inequalities (9) is strongly NP-hard.

Theorem 8. *It is strongly NP-hard to find an implied-demand cover $I' \subseteq I$ for a facility $j \in J$ that maximises the violation $y_j - \sum_{k\in(\cup_{i\in I'} J^{\leq}_{ij})\setminus\{j\}} y_k$ of inequality (9) for optimal solutions $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in [0,1]^{|I||J|+|J|}$ to the linear relaxation of (1).*

We prove this result in the Appendix.

7 Computational study

In this section, we report on the results of our computational study.

7.1 Set-up

In our tests, we consider 44 instances developed for the classical capacitated facility location problem with metric assignment costs (Avella and Boccia, 2009, Beasley, 1990), which can be currently accessed at Università degli Studi di Brescia. The considered instances consist of three different sizes: small (75 customers, 50 facilities), medium (100 customers, 75 facilities), and large (300 customers, 300 facilities). The small and medium sized instances are modified variants of instances *capa, capb* and *capc* from Beasley (1990). Analogously to Cánovas et al. (2007), we consider the first 75 (100) customers and 50 (75) facilities for small (medium) sized instances and multiply the opening costs with 0.0375 (0.075). In a similar manner to Calvete et al. (2020), we set the capacity of each facility $j \in J$ to $\lceil \bar{Q}_j/(\xi \cdot \bar{d}) \rceil$, where $\bar{Q}_j$ is the original capacity, $\bar{d}$ is the average of the customer demands in the full instance, and $\xi = 2, 3, 4$ for *capa, capb* and *capc*, respectively. As opposed to Calvete et al. (2020) who set $\xi = 1$, we choose higher values as several instances would be infeasible otherwise if preferences are defined by assignment costs. This allows us to better compare the impact of our methods depending on the preference type.

We consider two preference types. In the first type, preferences are defined by assignment costs: customer $i \in I$ prefers facility $j \in J$ over facility $k \in J$ if the assignment costs of i to j are lower than the costs of assigning i to k. In the second preference type, we slightly perturb the first preference type. Like before, facilities with small assignment costs are generally preferred over facilities with great assignment costs. Yet, a customer's preference ordering regarding facilities with similar assignment costs might deviate from the preference ordering according to lowest assignment costs. In order to determine the second preference ordering, we follow the procedure introduced by Cánovas et al. (2007).

In the former preference type, instances occur in which customers are indifferent between several potential facilities; in the latter preference type, each customer has a strict preference ordering of the potential facilities. Feasible instances for the capacitated facility location problem might turn infeasible when considering customer preferences due to more restrictive assignment rules. Among the instances studied here, instances $i300-5$ as well as *capa-1* and *capc-1* of medium size are infeasible if preferences are defined by assignment costs. We exclude these instances from our study. Thus, we study 41 instances for the case that preferences are defined by assignment costs and 44 otherwise.

All algorithms have been implemented in python 3.10.4 and all experiments have been performed on a Linux machine (Rocky Linux 8.9 Green Obsidian) with CPU clock 2.1 GHz and 5 GB RAM. We use Gurobi version 10.0.0 with default settings to solve integer linear programs if not stated otherwise.

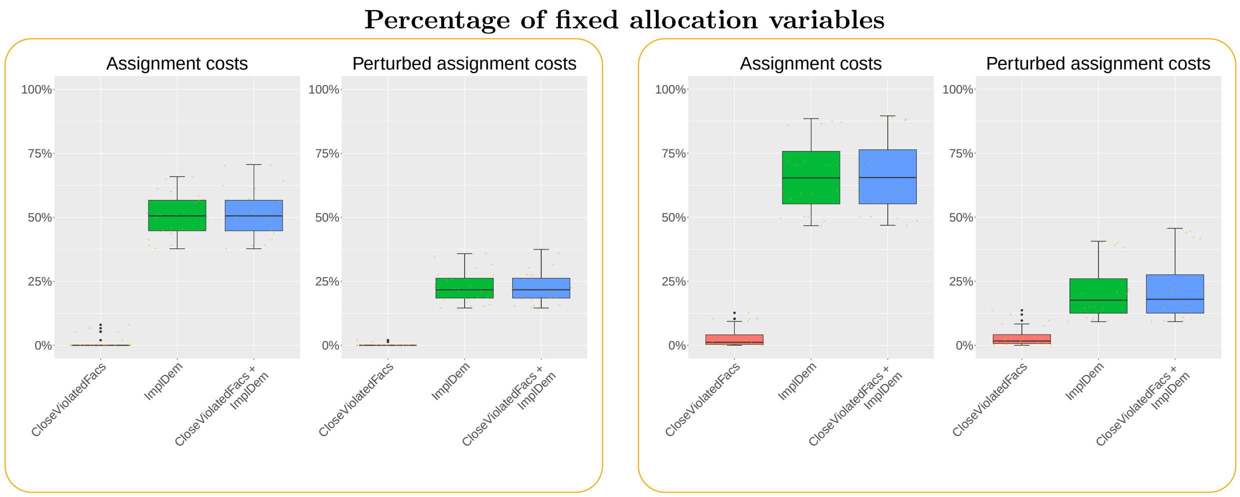
7.2 Preprocessing

We first analyse the impact of the preprocessing methods introduced in Section 4 on the computational performance.

7.2.1 Decision variables

Figure 1 shows the percentage of allocation variables fixed to zero by methods *CloseViolatedFacs* and *ImplDem* as well as their combination. The percentage of fixed location variables by *CloseViolatedFacs* corresponds to the percentage of fixed allocation variables; therefore, we only focus on allocation variables. We depict the results for the large instances separately from the remaining instances as they belong to a different family of instances. Note, however, that the preprocessing methods behave similarly across all instances. Therefore, we focus on a general description of the performance.

Method *CloseViolatedFacs* performs weakest for both preference types. The percentage of fixed variables by this method is not strongly affected by the preference type. Method *ImplDem* on its own as well as its combination with method *CloseViolatedFacs* perform similar within the same preference type. If preferences are defined by assignment costs, more than 40% of all allocation variables can be fixed for nearly all small and median sized instances; for large instances, more than 50% of the decision variables are fixed for nearly all instances. If preferences are defined by perturbed assignment costs, the performance decreases visibly.



(a) Small and medium instances.

(b) Large instances.

Figure 1: Percentages of allocation variables fixed to zero by methods *CloseViolatedFacs*, *ImplDem* and their combination.

The similarity between the computational results of method *ImplDem* and its combination with method *CloseViolatedFacs* is due to the following. Suppose a facility j has to be closed since it can not serve the demands of all customers that prefer it most. Then, the implied demand of any customer at j includes the demand of customers that prefer j most. Hence, no customer can be assigned to j even if j's location variable is not explicitly set to zero. This similarity only holds for decisions made in the first iteration of applying preprocessing *CloseViolatedFacs*, though. Method *CloseViolatedFacs* is reiterated multiple times since closing one facility changes the set of customers who prefer a facility most for the remaining facilities. This additional impact, however, is not strong in the considered instances.

The drop in the performance of method *ImplDem* and its combination with *CloseViolatedFacs* between the two preference types is likely due to a decrease in the size of implied customer sets when considering preferences defined by perturbed assignment costs. Method *ImplDem* depends highly on these sets, which occur naturally when preferences are defined by metric assignment costs.

In conclusion, method *ImplDem* is affected by the choice of preferences, and its impact is much stronger when considering preferences defined by assignment costs. Method *CloseViolatedFacs* is not affected by the choice of preferences but does not perform strongly in the first place. In the remainder of this study, we therefore drop the consideration of method *CloseViolatedFacs*.

7.2.2 Computational performance

In the following, we discuss the impact of preprocessing method *ImplDem* with and without preprocessing methods by Cánovas et al. (2007), which were developed for the uncapacitated facility location problem with customer preferences, on the computational performance. First, we revisit the preprocessing developed by Cánovas et al. (2007). Then, we analyse the time needed to build the model in Gurobi and to perform the considered preprocessing methods. We refer to this time needed until the solver starts computing a solution as the *build-up time*. Afterwards, we compare the optimality gaps of the various approaches achieved by Gurobi within one hour minus the build-up time, and determine the best-performing preprocessing methods depending on the considered preference type.

According to Cánovas et al. (2007), further constraints are added to the integer programming formulation if (a) a customer prefers a facility most, cf. Proposition 1; (b) the preference set of a customer $i \in I$

Build-up times in seconds

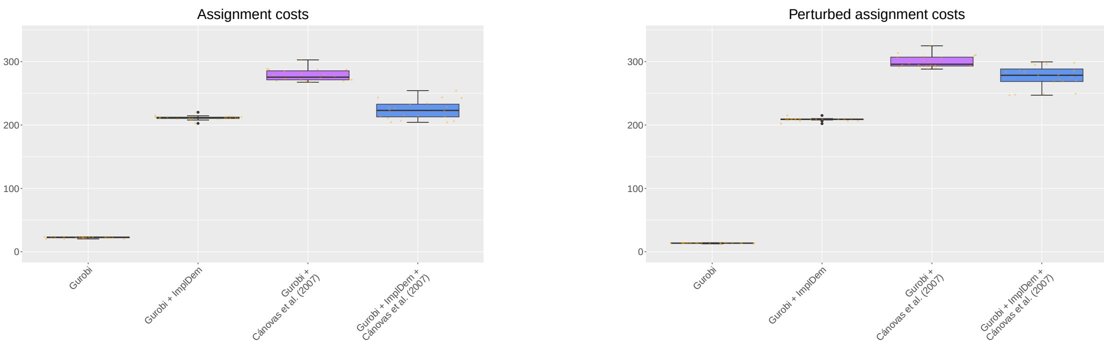


Figure 2: Build-up times for different preprocessing methods and preference types for large instances.

regarding a facility $j \in J$ is *maximally* contained in the preference set of another customer $k \in I \setminus \{i\}$, i.e., there is no third customer $\ell \in I \setminus \{i, k\}$ with $J_{ij}^{\leq} \setminus \{j\} \subset J_{\ell j}^{\leq} \setminus \{j\} \subset J_{kj}^{\leq} \setminus \{j\}$, cf. Proposition 3; or (c) if two customers have coinciding preference sets, cf. Proposition 6. In the following, we perform the methods introduced by Cánovas et al. (2007) after executing method *ImplDem*.

Figure 2 shows the build-up times for the considered preprocessing combinations and both preference types for large instances. The build-up times for small and medium sized instances are smaller than 1.25 and 4 seconds for nearly all instances.

First, note that methods proposed by Cánovas et al. (2007) on their own take longest, followed by the combination of *ImplDem* and Cánovas et al. (2007), only *ImplDem* and then Gurobi. The drop in the build-up times when considering both *ImplDem* and Cánovas et al. (2007) is likely due to a decrease in the needed comparisons due to already fixed assignment variables after executing *ImplDem*.

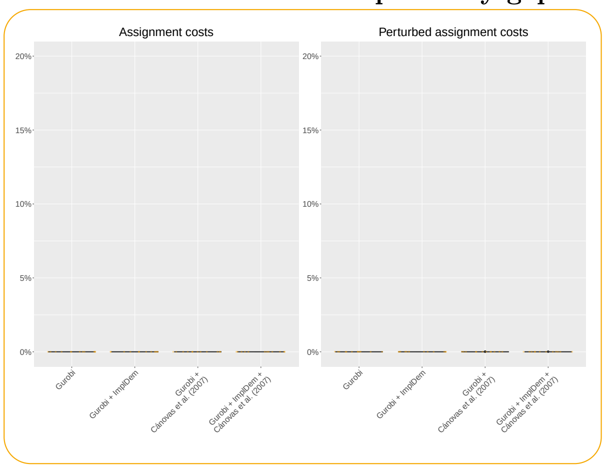
Figure 3 depicts the optimality gaps in percentage achieved by Gurobi after one hour for different preprocessing methods and preference types. Within the given time limit, the preprocessing has to be performed as well. We set the number of threads used in Gurobi to one in order to prevent parallelisation as this allows us to get a better understanding of the computational impact of our methods. Besides fixing the number of threads, we do not modify any other default parameters.

First, we analyse the results for the large instances. We observe that these instances are not solved optimally within the given time limit. The optimality gaps are smaller if preferences are defined by perturbed assignment costs in case Gurobi is used without additional preprocessing methods. If preferences are defined by assignment costs, applying method *ImplDem* either on its own or together with the methods developed by Cánovas et al. (2007) yields an improvement. Since the performance of both approaches is so similar, it is impossible to deem one approach better. If preferences are defined by perturbed assignment costs, none of the considered preprocessing methods yield a clear improvement.

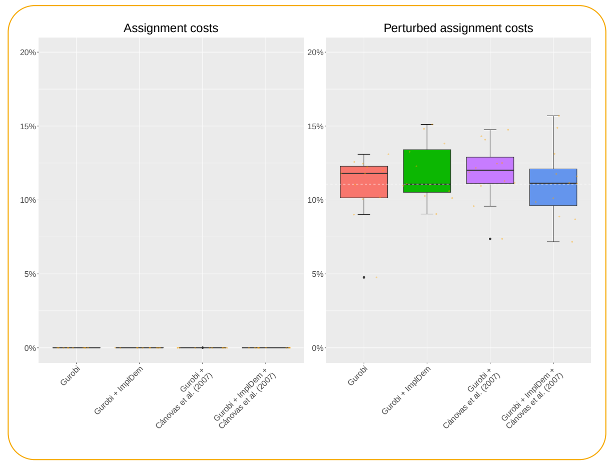
On the contrary to large instances, we observe that medium sizded instances are easier to solve when preferences are defined by assignment costs. In this case, all instances are solved to optimality, while none of the instances for the second preference type are solved to optimality. Figure 4 (right) shows the time needed to solve instances in which preferences are defined by assignment costs. The general impact of the tested preprocessing methods behaves similarly to large instances: if preferences are defined by assignment costs, *ImplDem* on its own as well as combined with the methods by Cánovas et al. (2007) show the best performance, if preferences are defined by perturbed assignment costs, the performance of all approaches behaves similarly.

Lastly, let us consider small instances. Here, all instances are solved to optimality. If we look at the time it takes to solve the instances, cf. Figure 4 (left), instances with preferences defined by assignment costs are

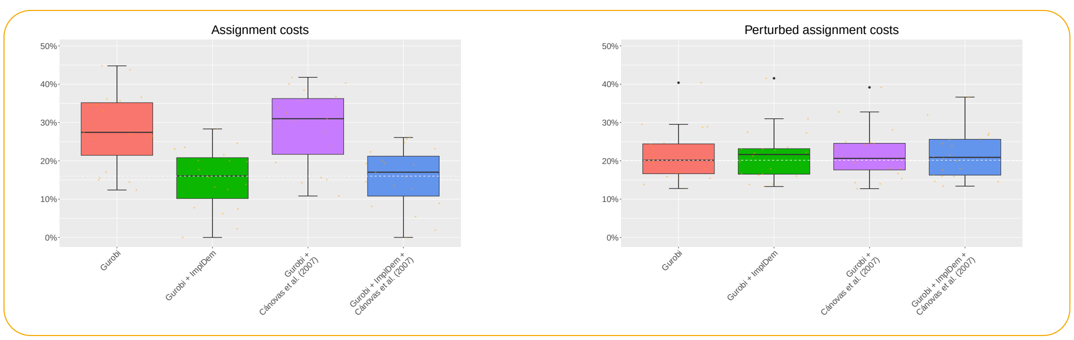
Optimality gap after one hour in Gurobi



(a) Small instances.

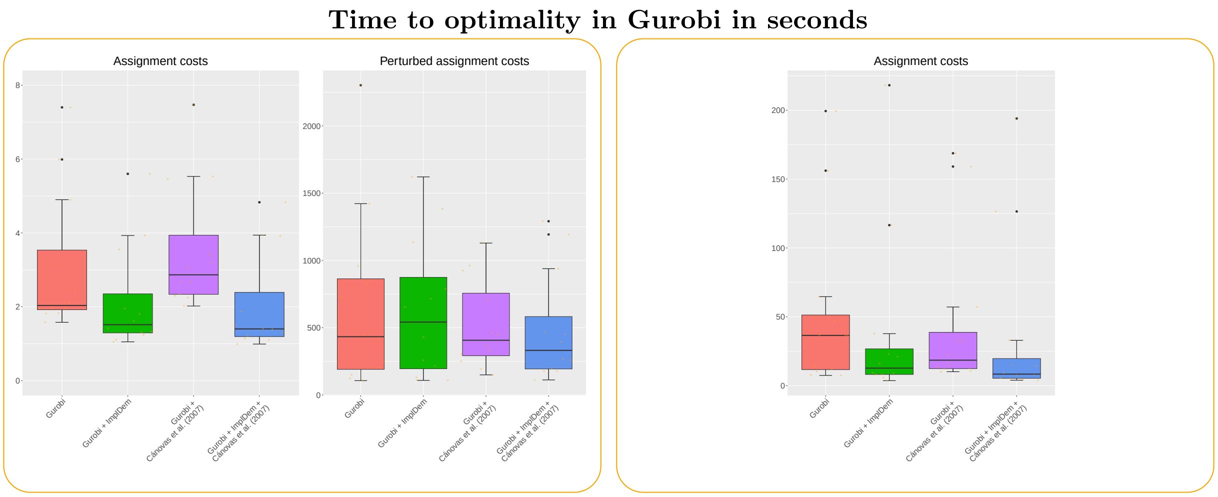


(b) Medium instances.



(c) Large instances.

Figure 3: Optimality gaps in percentage achieved for different preprocessing methods and preference types. The white dotted line depicts the level of the lowest median among all approaches.



(a) Small instances.

(b) Medium instances.

Figure 4: Time to optimality in Gurobi in seconds for different preprocessing methods and preference types. Note the different time scales on the y-axes in the three plots.

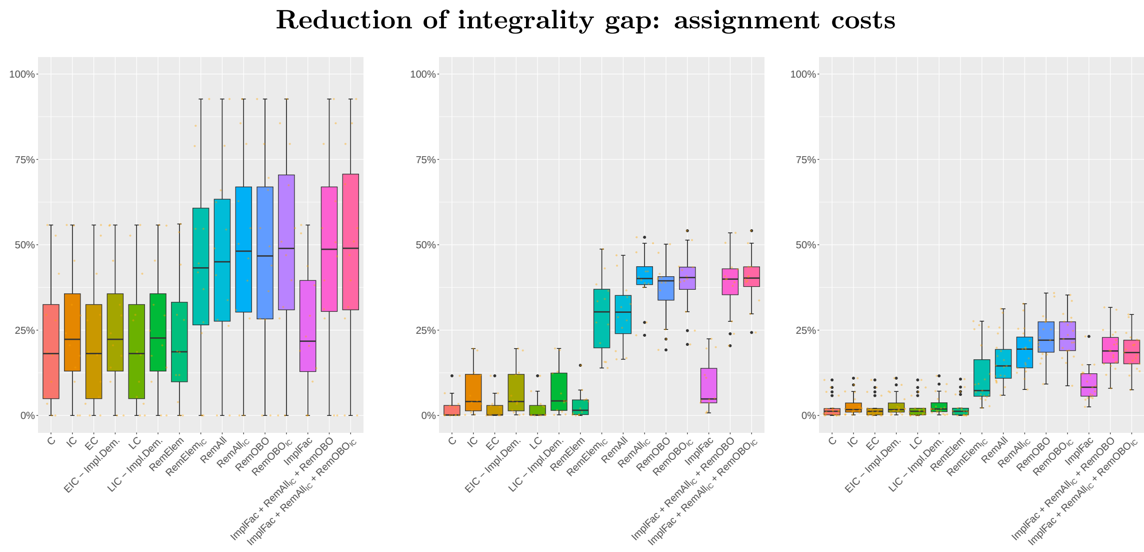
solved in less than 10 seconds. Solving the instances for the second preference type is more difficult. Note the different time scales on the y-axes in the three plots.

We conclude that the behaviour of the preprocessing methods is consistent throughout the different instance sizes. If preferences are defined by assignment costs, method *ImplDem* either on its own or together with the methods developed by Cánovas et al. (2007) yields an improvement. In the following, we decide to only consider preprocessing method *ImplDem* as this approach reduces the model size. If preferences are defined by perturbed assignment costs, none of the additional approaches yields a clear improvement. The computations for small and medium instances indicate that this lack of impact is due to the problem's structure: the build-up times of the models are low, yet, the fixed decisions do not yield an advantage. Future work might revolve around new preprocessing ideas for the case that preferences are not defined by assignment costs.

7.3 Reduction of the integrality gap

In the following, we evaluate the relevance of valid inequalities **(I)C**, **E(I)C**, **L(I)C** as well as inequalities **RemElem**, **RemAll**, **RemOBO** with and without implied-demands, and **ImplFac**. We consider the performance of each of these inequalities on their own and, for selected inequalities, their combinations. We utilise the findings from the analysis in Section 7.2 and consider the best-performing preprocessing approaches for each preference type. Hence, if preferences are defined by assignment costs, we consider preprocessing *ImplDem*. Otherwise, we do not consider any preprocessing at all. We measure the theoretical performance of the considered inequalities through the *reduction of the integrality gap*.

In order to compute the reduction of the integrality gap, we start by considering a solution for the relaxation of integer linear program (1), which provides a lower bound on the optimal integer solution. We gradually add violated inequalities to integer program (1) as cuts; these inequalities are found by solving the separation problem for the considered inequality-family. We stop once no further inequalities are found. We denote the final lower bound with $x^*_{LP'}$. Denote the best known IP-solution computed in the previous subsection with x^*_{IP} and the lower bound corresponding to the solution of the relaxation of formulation (1) with x^*_{LP}. Then, we measure the reduction of the integrality gap through $r = (1 - (x^*_{IP} - x^*_{LP'})/(x^*_{IP} - x^*_{LP})) \cdot 100$.



(a) Small instances. (b) Medium instances. (c) Large instances.

Figure 5: Reduction of the integrality gaps in percentage for the valid inequalities discussed in Section 5 with preferences defined by assignment costs.

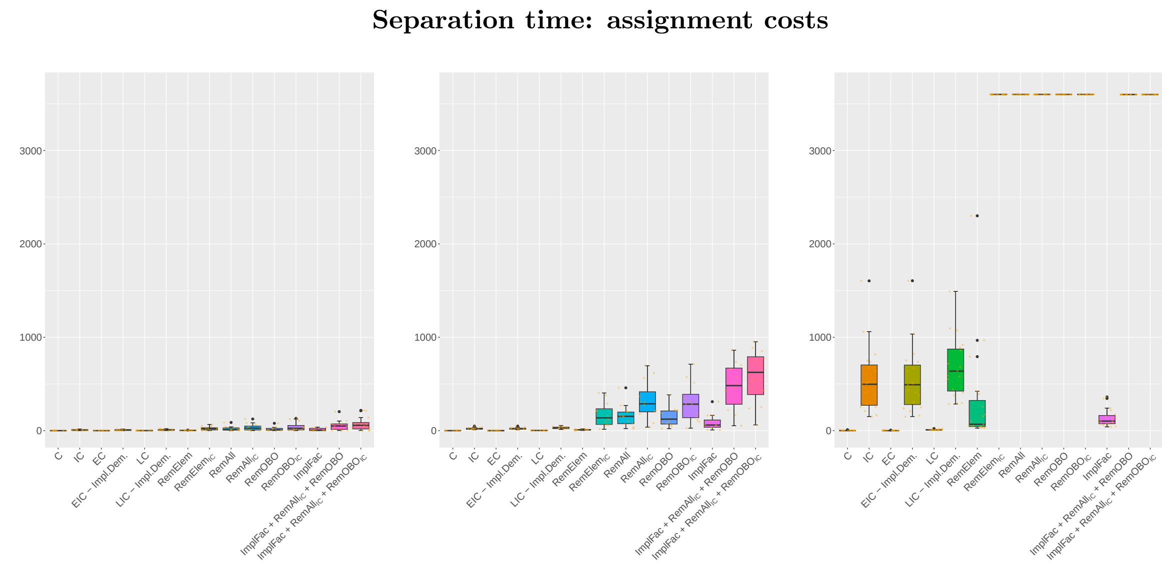
To assess the impact of our inequalities, we provide one hour for the process of computing the reduction of the integrality gap. Compare Section 6 for further information on solution approaches for the separation problems to be solved in this section. We set x^*_{LP} as the solution of the relaxation of formulation (1), i.e., no preprocessing methods are considered. If we combine inequalities, the ordering of the combination is reflected in the name. Furthermore, we distribute the given time equally among all considered inequalities. If no more violated inequalities are detected within the given time interval, we distribute the remaining time equally among the remaining inequalities. As an example, suppose four inequalities are combined. Then, we determine violated inequalities for the first inequality-family and add them as cuts for 15 minutes. Based on the formulation after 15 minutes, we detect and add violated inequalities for the second inequality-family for 15 minutes. If no further violated inequalities are found after 10 minutes, each of the remaining two inequality-families will be considered for 17.5 minutes.

Figure 5 depicts the reduction of the integrality gap for the inequalities discussed in Section 5 in case that preferences are defined by assignment costs.

First, we observe that the impact of the inequalities decreases with increasing instance size. Yet, especially for the small instances, we observe a significant reduction of the integrality gap. Furthermore, the behaviour of the inequalities stays consistent across all instance families. That is, the impact of the traditional cover inequalities, of **RemAll** and **ImplFac** is notably lower than the impact of the strengthened cover-based inequalities. Second, we observe that the concept of implied-demand covers yields an improvement for all inequalities. Third, we observe that inequality family $\mathbf{RemOBO}_{IC}$ belongs to the best-performing inequalities across all instance sizes. Lastly, combining the best-performing valid inequalities does not yield a (noteworthy) additional improvement.

Figure 6 depicts the separation time. For all inequality families besides the strengthened cover-based inequalities, all violated inequalities are found within the considered time window throughout all instance sizes. For the remaining inequalities, some violated inequalities are not found within the time limit for large instances. Hence, these inequalities have further potential in case of the large instances.

The graphics depicted in Figure 7 show the reduction of the integrality gap for the inequalities discussed



(a) Small instances. (b) Medium instances. (c) Large instances.

Figure 6: Separation time in seconds if preferences are defined by assignment costs.

in Section 5 if preferences are defined by perturbed assignment costs.

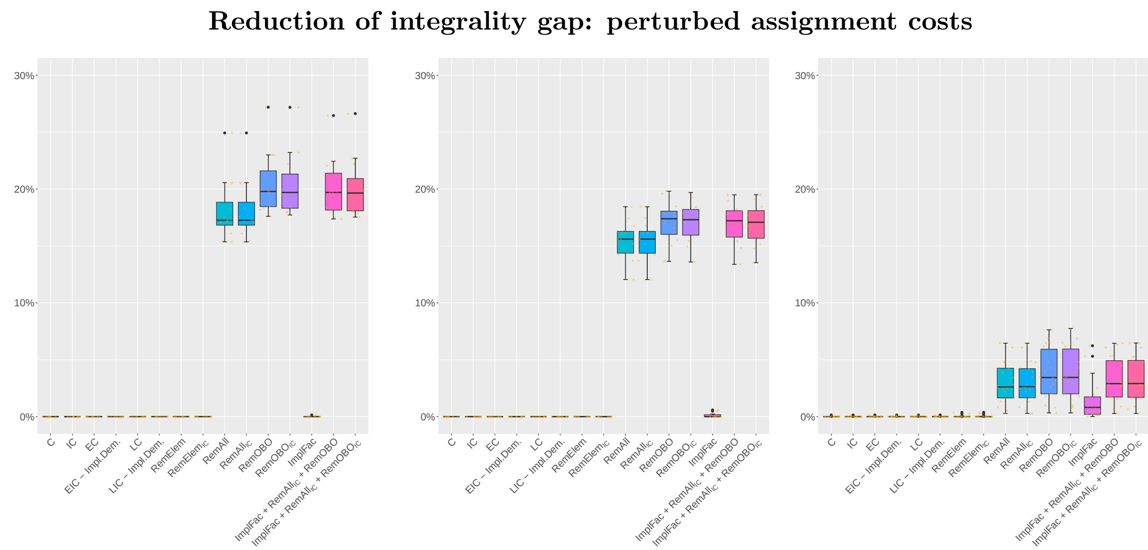
Similar to the previous preference type, we observe a decrease in the reduction of integrality gaps with increasing instance size for nearly all considered inequalitiy families. However, inequality family **ImplFac** profits from increasing instance sizes; recall that this instance family utilises a different structure of the SSCFLPCP than the remaining inequality families. Conversely to the previous preference type, traditional cover inequalities as well as inequality-families **RemElem**, **RemElem**$_{IC}$ and **ImplFac** have nearly no impact on the integrality gap. Inequality-families **RemOBO** and **RemOBO**$_{IC}$ perform best throughout all instance sizes; especially for small and medium sized instances, they significantly reduce the integrality gap. Considering inequalities based on implied-demand covers instead of traditional covers does, however, not offer a noteworthy improvement. This behaviour is similar across all instance sizes. Lastly, we notice a strong drop in the performance of the considered inequalities if preferences are defined by perturbed assignment costs instead of assignment costs. We therefore derive that the definition of preferences has a strong impact on the performance of our inequalities.

Figure 8 depicts the separation time. For most strengthened cover-based inequality families, we run into the time limit for large instances. Hence, these inequalities are likely to have more potential in case of the large instances.

In conclusion, the computations indicate that the behaviour of our inequality families is consistent across all instance sizes. That is, traditional (implied-demand) cover inequalities and their extensions have a lower impact than the strengthened (implied-demand) cover inequalities. The inequalities perform better if preferences are defined by assignment costs, and the theoretical impact of the inequalities decreases with the instance size. Finally, note that some of our inequality families significantly reduce the integrality gap, especially for small and medium sized instances for both preference types.

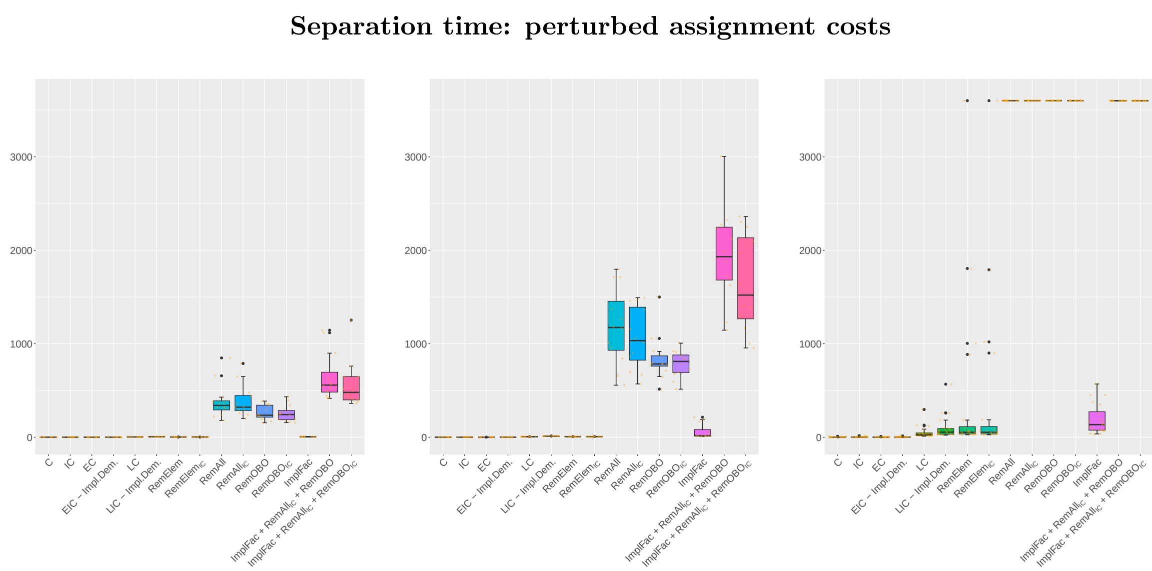
7.4 Computational performance of a branch-and-cut approach

In the following, we test the impact of our inequalities in a *branch-and-cut* approach in Gurobi. If preferences are defined by assignment costs, we only consider large instances as nearly all small and medium



(a) Small instances. (b) Medium instances. (c) Large instances.

Figure 7: Reduction of the integrality gaps in percentage for the valid inequalities discussed in Section 5 with preferences defined by perturbed assignment costs.



(a) Small instances. (b) Medium instances. (c) Large instances.

Figure 8: Separation time in seconds if preferences are defined by perturbed assignment costs.

sized instances are solved to optimality within 8 and 50 seconds, respectively, cf. Figure 4. Due to the strong NP-hardness of the separation problems for the inequalities, it does not seem promising to test the inequalities in a branch-and-cut approach for small and medium sized instances that are solved relatively quickly already withouth the inequalities. If preferences are defined by perturbed assignment costs, we only consider small and medium instances since the theoretical impact of the inequalities for large instances is relatively small, cf. Figure 7. We therefore do not expect a noteworthy impact of the inequalities on large instances when using the exact separation routines considered in this paper. We focus on strengthened cover-based inequalities, except for families **RemElem** and **ImplFac**, as they did lead to significant reductions of the integrality gap, cf. Figures 5 and 7.

For large and medium instances, we study the remaining optimality gap after one hour. We do this by considering the ratio between the gap computed by the branch-and-cut approach and the approach without cuts. This ratio allows us to compare the impact of the branch-and-cut approach for each instance; we immediately see which instances improve and which worsen. We define the ratio for an instance i as $r_G(i) = \frac{(Gap_{C\&B}(i) - Gap(i))}{\max\{Gap_{C\&B}(i), Gap(i)\}} \in [-1, 1]$, where $Gap_{C\&B}(i)$ denotes the optimality gap of i when considering our cuts and $Gap(i)$ denotes the optimality gap of i without our cuts. If $r_G(i)$ takes a negative (positive) value, the gap computed by the branch-and-cut approach is better (worse). Lastly, if $r_G(i)$ is equal to 1 (-1), the approach without (with) cuts solved the instance to optimality while the other approach did not.

Since small instances are solved to optimality within the given time horizon, we compare the ratio between the solution times of the branch-and-cut approach and the time it takes Gurobi to solve the instances to optimality. If the ratio is greater (lower) than one, Gurobi performs better (worse).

We consider the following setting. If preferences are defined by assignment costs, we consider preprocessing method *ImplDem*. Otherwise, we do not consider any preprocessing at all. In order to reserve enough time for the actual optimisation, we allow the use of our inequalities within the first 450 seconds of the solution process only and up to a depth of 50 in the branching tree; furthermore, we only solve the separation problem for a facility $j \in J$ if its location variable value in the solution to the relaxed problem is strictly greater than 0.5. Lastly, we set the *PreCrush* parameter in Gurobi to 1 when we integrate our inequalities as cuts, as demanded by the Gurobi documentation; if we do not consider our inequalities, we keep the *PreCrush* parameter at its default value, i.e., zero.

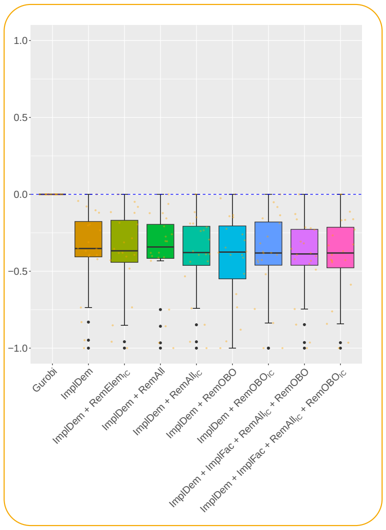
Figure 9 shows the ratio between the gap computed by the branch-and-cut approach and Gurobi for large (left) and medium (center) instances; the graphic in Figure 9 on the right shows the ratio between the solution times of the branch-and-cut approach and Gurobi for small instances.

If preferences are defined by assignment costs, the inequalities combined with preprocessing method *ImplDem* outperform Gurobi in each instance. By considering inequality family **RemOBO** in a branch-and-cut approach, we reach a notable improvement compared to the performance of Gurobi together with preprocessing method *ImplDem*.

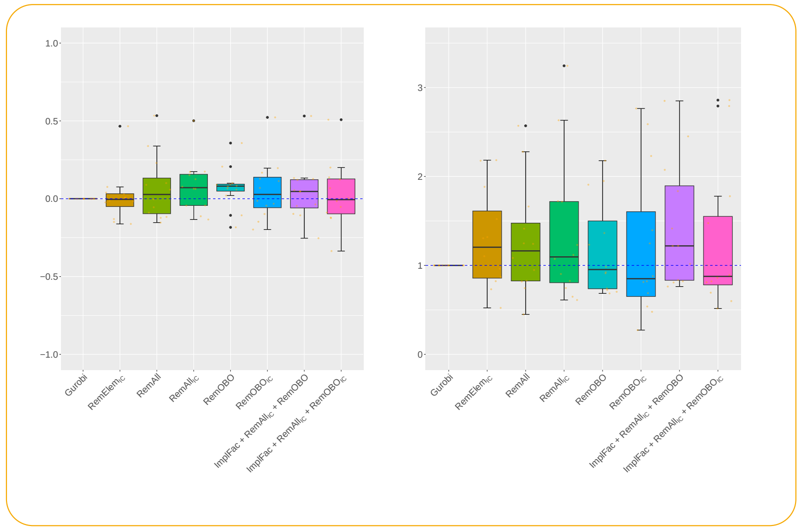
If preferences are defined by perturbed assignment costs, none of the inequalities yields a clear improvement. Inequality families **RemOBO**, $\mathbf{RemOBO}_{IC}$ as well as the combination of **ImplFac**, $\mathbf{RemAll}_{IC}$ and $\mathbf{RemOBO}_{IC}$ show, however, potential on small instances.

Our computations indicate that the inequalities introduced in this paper have a positive impact on solving instances of the SSCFLPCP if preferences are defined by assignment costs - even though we focused on exact separation routines and did not yet develop heuristic ones that may provide a better trade-off between the time needed and the resulting improvements of lower bounds. We observe that simple time management in the branch-and-cut approach already uncovers potential of our inequalities. If preferences are defined by perturbed assignment costs, we do not observe a clear impact. This is in contrast to the theoretical potential of the inequalities we observed before; cf. Figure 7. Future work might include research on heuristic approaches for solving the arising separation problems for both preference types.

Ratio of the branch-and-cut approach VS Gurobi



(a) Assignment costs. Ratio of the optimality gaps for large instances.



(b) Perturbed assignment costs. Ratio of the optimality gaps for medium instances (left) and ratio of the time to optimality for small instances (right).

Figure 9: Performance of the valid inequalities in a branch-and-cut approach measured by the ratio of the optimality gaps for large and medium instances; measured by the ratio of the solution times for small instances.

8 Conclusion

In this paper, we study preprocessig methods and valid inequalities for the single-source capacitated facility location problem with customer preferences (SSCFLPCP). While (capacitated) facility location problems are already well studied, less research has been conducted on the SSCFLPCP. Considering customer preferences, however, is important in settings where customers choose a facility which serves their demand.

The main contributions in this work are (a) two preprocessing methods which reduce the size of the considered integer programming formulation, (b) the introduction of the concept of implied-demand covers which generalise normal covers, (c) a study of the complexity of finding implied-demand covers as well as (d) new, cover-based valid inequalities which decrease the integrality gaps.

More specifically, the first preprocessing method fixes location variables if the corresponding facility is not capable of serving the customers who prefer it most. The second preprocessing method fixes allocation variables if a facility is not able to serve the demand of a customer assigned to it as well as the demands of customers who strictly prefer the same facilities over the considered facility as the assigned customer. We propose and analyse inequalities for the SSCFLPCP based on implied-demand covers, which incorporate information on assignments of customers in the cover that are not assigned to the covered facility. We discuss solution approaches for the separation problems belonging to our inequalities. All of our inequalities are also valid, yet weaker, for traditional covers. Finally, we test the computational potential of our findings in a computational study. We discover that the choice of preferences has a great impact on the performance of our preprocessing and valid inequalities. While our approaches work well if preferences are defined by assignment costs, their impact is smaller if preferences are defined by perturbed assignment costs.

This work reveals that there are a lot of open research questions waiting. First, the computational study shows that there is potential in the strengthened cover-based inequalities. This potential is held back by slow methods for solving the separation problems. Here, it is important to develop faster, well-performing heuristics for solving these problems. Second, we observe that our preprocessing methods and inequalities perform worse when preferences are defined by perturbed assignment costs. Future work might also include the study of preprocessing methods and inequalities which perform better for this preference type.

Acknowledgements

We would like to warmly thank Timo Gersing for fruitful discussions and helpful comments. We would like to further thank the anonymous referees for their helpful comments. This research has been partially supported by the German Federal Ministry of Education and Research (grant no. 05M16PAA) within the project "HealthFaCT - Health: Facility Location, Covering and Transport", by the Freigeist-Fellowship of the Volkswagen Stiftung and by the German research council (DFG) Research Training Group 2236 UnRAVeL.

References

K. Aardal, Y. Pochet, and L. A. Wolsey. Capacitated facility location: Valid inequalities and facets. *Mathematics of Operations Research*, 20(3):562–582, 1995.

F. Ahmad, A. Iqbal, I. Ashraf, M. Marzband, and I. Khan. Optimal location of electric vehicle charging station and its impact on distribution network: A review. *Energy Reports*, 8:2314–2333, 2022.

A. Ahmadi-Javid, P. Seyedi, and S. S. Syam. A survey of healthcare facility location. *Computers & Operations Research*, 79:223–263, 2017.

P. Avella and M. Boccia. A cutting plane algorithm for the capacitated facility location problem. *Computational Optimization and Applications*, 43:39–65, 2009.

P. Avella, M. Boccia, S. Mattia, and F. Rossi. Weak flow cover inequalities for the capacitated facility location problem. *European Journal of Operational Research*, 289(2):485–494, 2021.

E. Balas. Facets of the knapsack polytope. *Mathematical Programming*, 8(1):146–164, 1975.

J. E. Beasley. OR-Library: capacitated warehouse location. `http://people.brunel.ac.uk/~mastjjb/jeb/orlib/capinfo.html`, 1990. Accessed: 2024-12-05.

C. Büsing, T. Gersing, and S. Wrede. Analysing the complexity of facility location problems with capacities, revenues, and closest assignments. *Proceedings of the 10th International Network Optimization Conference (INOC), Aachen, Germany, June 7–10, 2022*, pages 81–86, 2022.

X. Cabezas and S. García. A semi-lagrangian relaxation heuristic algorithm for the simple plant location problem with order. *Journal of the Operational Research Society*, 0(0):1–12, 2022.

H. Calvete, C. Galé, J. Iranzo, J. Camacho-Vallejo, and M. Casas-Ramírez. A matheuristic for solving the bilevel approach of the facility location problem with cardinality constraints and preferences. *Computers & Operations Research*, 124:105066, 2020.

L. Cánovas, S. García, M. Labbé, and A. Marín. A strengthened formulation for the simple plant location problem with order. *Operations Research Letters*, 35(2):141–150, 2007.

M. Casas-Ramírez, J. Camacho-Vallejo, and I. Martínez-Salazar. Approximating solutions to a bilevel capacitated facility location problem with customer's patronization toward a list of preferences. *Applied Mathematics and Computation*, 319:369–386, 2018.

D. Celik Turkoglu and M. Erol Genevois. A comparative survey of service facility location problems. *Annals of Operations Research*, 292:399–468, 2020.

M. Conforti, G. Cornuejols, and G. Zambelli. *Integer Programming.* Springer Publishing Company, 2014.

H. Crowder, E. L. Johnson, and M. Padberg. Solving large-scale zero-one linear programming problems. *Operations Research*, 31:803–834, 1983.

I. Espejo, A. Marín, and A. M. Rodríguez-Chía. Closest assignment constraints in discrete location problems. *European Journal of Operational Research*, 219(1):49–58, 2012.

M. Fischetti, I. Ljubić, and M. Sinnl. Benders decomposition without separability: A computational study for capacitated facility location problems. *European Journal of Operational Research*, 253(3):557–569, 2016.

M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of $\mathcal{NP}$-completeness*. W. H. Freeman & Co., 1979.

R. A. Gerrard and R. L. Church. Closest assignment constraints and location models: Properties and structure. *Location Science*, 4(4):251–270, 1996.

S. Görtz and A. Klose. A simple but usually fast branch-and-bound algorithm for the capacitated facility location problem. *INFORMS Journal on Computing*, 24(4):597–610, 2012.

Z. Gu, G. L. Nemhauser, and M. W. P. Savelsbergh. Lifted cover inequalities for 0-1 integer programs: Complexity. *INFORMS Journal on Computing*, 11(1):117–123, 1999.

Gurobi. `https://www.gurobi.com/`. Accessed: 2024-12-05.

P. Hanjoul and D. Peeters. A facility location problem with clients' preference orderings. *Regional Science and Urban Economics*, 17(3):451–473, 1987.

C.-N. Kang, L.-C. Kung, P.-H. Chiang, and J.-Y. Yu. A service facility location problem considering customer preference and facility capacity. *Computers & Industrial Engineering*, 177:109070, 2023.

K. Kaparis and A. N. Letchford. Separation algorithms for 0-1 knapsack polytopes. *Mathematical Programming*, 124:69–91, 2010.

D. Klabjan, G. Nemhauser, and C. Tovey. The complexity of cover inequality separation. *Operations Research Letters*, 23:35–40, 08 1998.

B. H. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*. Springer-Verlag, 2018.

G. Laporte, S. Nickel, and F. Saldanha da Gama, editors. *Location Science*. Springer, 2 edition, 2020.

J. M. Y. Leung and T. L. Magnanti. Valid inequalities and facets of the capacitated plant location problem. *Mathematical Programming*, 44:271–291, 1989.

S. T. McCormick. Submodular function minimization. In K. Aardal, G. Nemhauser, and R. Weismantel, editors, *Discrete Optimization*, volume 12 of *Handbooks in Operations Research and Management Science*. Elsevier, 2005.

A. M. Mestre, M. D. Oliveira, and A. P. Barbosa-Póvoa. Location–allocation approaches for hospital network planning under uncertainty. *European Journal of Operational Research*, 240(3):791–806, 2015.

P. Mirchandani and R. Francis, editors. *Discrete Location Theory*. Wiley Interscience, 1990.

S. Polino, J.-F. Camacho-Vallejo, and J. G. Villegas. A facility location problem for extracurricular workshop planning: bi-level model and metaheuristics. *International Transactions in Operational Research*, 0:1–43, 2023.

P. Rojeski and C. ReVelle. Central facilities location under an investment constraint. *Geographical Analysis*, 2(4):343–360, 1970.

M. Sipser. *Introduction to the Theory of Computation: 2nd Edition*. International Thompson Publishing, 1996.

Università degli Studi di Brescia. `https://or-brescia.unibs.it/instances/instances_sscflp`. Accessed: 2024-12-05.

I. Vasilyev, X. Klimentova, and Y. A. Kochetov. New lower bounds for the facility location problem with clients' preferences. *Computational Mathematics and Mathematical Physics*, 49:2010–2020, 2010.

I. Vasilyev, X. Klimentova, and M. Boccia. Polyhedral study of simple plant location problem with order. *Operations Research Letters*, 41:153–158, 2013.

J. L. Wagner and L. M. Falkson. The optimal nodal location of public facilities with price-sensitive demand. *Geographical Analysis*, 7(1):69–83, 1975.

Appendix A: Computational Complexity

In the following, we provide the proof for Theorem 1.

Theorem 1. *It is strongly NP-hard to find a minimum implied-demand cover $\tilde{I} \subseteq I$ for a facility $j \in J$.*

Proof. In order to prove the first statement of the theorem, we show that the decision version of the minimum implied-demand cover problem is strongly NP-complete. That is, we show that it is strongly NP-complete to determine whether there is an implied-demand cover $\tilde{I} \subseteq I$ for a facility $j \in J$ with size K for some integer $K \in \mathbb{Z}_{\geq 0}$. We prove this via a reduction from the *minimum set cover problem*, which is known to be strongly NP-complete (Garey and Johnson, 1979).

In the minimum set cover problem, we are given an instance $\mathcal{I}$ consisting of a set of elements $U = \{1, 2, \ldots, n\}$, also known as the *universe*, and a collection S of m subsets of universe U. The goal is to decide whether there is a sub-collection $S^K \subseteq S$ consisting of $K \in \mathbb{Z}_{\geq 0}$ subsets such that the union of the subsets in S^K equals the universe, i.e., $\cup_{s \in S^K} s = U$. Without loss of generality, we assume that collection S does not contain subsets s_1, s_2 with $s_1 \subseteq s_2$ and that each element of the collection S contains at least two elements.

We start by presenting a transformation of any minimum set cover instance $\mathcal{I}$ into an instance $\mathcal{I}'$ of the minimum implied-demand cover problem. Introduce one customer i^u for each element $u \in U$ in the universe and one customer i^s for each element $s \in S$. Denote the set of customers with $I = I^U \cup I^S$, where I^U (I^S) contains all customers corresponding to elements in U (S) in the corresponding minimum set cover instance. Set the demand d_i of each customer $i \in I$ to one, i.e., $d_i = 1$. Define the set of facilities as $J = \cup_{i \in I}\{j_i\} \cup \{j\}$. That is, there are $|I| + 1$ facilities. We set the assignment costs of a customer $i \in I$ to a facility $p \in J$ to one, i.e., $c_{ip} = 1$, and set the costs for opening facility $p \in J$ to one as well, i.e., $f_p = 1$. Our goal is to determine whether an implied-demand cover of size K for facility j exists. Therefore, it suffices to explicitly define sets $J^<_{ij}$ and $J^=_{ij}$ for facility j and all customers $i \in I$. The set of facilities a customer $i^u \in I^U$ prefers over j or is indifferent to is defined as $J^<_{i^u j} = \{j_{i^u}\}$ and $J^=_{i^u j} = \{j\}$. Hence, the assignment of any customer in I^U to j does not imply the assignment of another customer in I to j. The set of facilities a customer $i^s \in I^S$ prefers over j or is indifferent to is defined as $J^<_{i^s j} = \bigcup_{\ell \in s} J^<_{i^\ell j} \cup \{j_{i^s}\}$ and $J^=_{i^s j} = \{j\}$. These preference sets ensure that assigning customer i^s to j implies the assignment of any customer in I^U corresponding to a universe element $\ell \in s$ to j. Finally, we set the capacity of facility j to $Q_j = K + |U| - 1$. The goal is to decide whether there is an implied-demand cover $\tilde{I} \subseteq I$ of facility j with size K.

Instance $\mathcal{I}'$ can be constructed in polynomial time from a minimum set cover instance $\mathcal{I}$ of polynomial size as we see next. The set of customers in instance $\mathcal{I}'$ corresponds to the elements in universe U and collection S. Since instance $\mathcal{I}$ has polynomial size, introducing the customers in $\mathcal{I}'$ takes polynomial time. Introducing the set of facilities takes also polynomial time: the number of facilities exceeds the number of customers by one. Lastly, we introduce the preference sets, customer demands, assignment and opening costs as well as a capacity for the considered facility. Given that the number of customers and facilities is of polynomial size, introducing these parameters takes also polynomial time.

Given the transformation from any minimum set cover instance $\mathcal{I}$ into a minimum implied-demand cover instance $\mathcal{I}'$, it remains to prove that there is a set cover with size K for instance $\mathcal{I}$ *if and only if* there is an implied-demand cover $\tilde{I} \subseteq I$ of facility j with size K in instance $\mathcal{I}'$.

Suppose there is a set cover with size K for instance $\mathcal{I}$. Denote the sub-collection of size K covering all elements in U with S^K. We construct a solution for instance $\mathcal{I}'$ of the implied-demand cover problem as follows. Define set $\tilde{I}$ as the set of all customers $i \in I^S$ corresponding to elements in set S^K. Set $\tilde{I}$ is of size K. Due to the definition of the preference sets and since S^K covers universe U, assigning the customers in $\tilde{I}$ to j implies the assignment of all customers $i \in I^U$ to j. Thus, the implied demand by the customers in set $\tilde{I}$ is $K + |U|$. This violates the capacity at j, and set $\tilde{I}$ is an implied-demand cover of j with size K.

Conversely, suppose that $\tilde{I}$ corresponds to an implied-demand cover of facility j with size K in instance $\mathcal{I}'$. First, we show that the assignment of $\tilde{I}$ to j implies the assignment of all customers in I^U to j. Suppose the assignment of $\tilde{I}$ to j does not imply the assignment of at least one customer in I^U to j. Then, the demand from customers in set I^U served at j is at most $|U| - 1$. In order to still construct an implied-demand cover of facility j, we have to assign at least $K + 1$ customers from set I^S to j. Then, set $\tilde{I}$ consists of $K + 1$ customers since the assignment of customers in I^S to j is not implied by the assignment of any other customer. This is a contradiction to the assumption that $\tilde{I}$ is of size K. Second, we show that set $\tilde{I}$ does not contain any customer from set I^U. Suppose set $\tilde{I}$ contains at least one customer from set I^U. In this case, there are at least $K + 1$ customers in set $\tilde{I}$: otherwise, the implied demand to be served at j is at most $K + |U| - 1$, which respects the capacity at j. This is a contradiction to the assumption that set $\tilde{I}$ is of size K. Given these results, we construct a solution for instance $\mathcal{I}$ as follows. Define a cover consisting of all elements in collection S that correspond to customers in set $\tilde{I}$. Due to the second property of set $\tilde{I}$, this cover is of size K. Due to the first property of set $\tilde{I}$, all elements in U are covered by at least one element in the cover. Hence, the constructed cover is indeed a set cover of size K for instance $\mathcal{I}$ and our claim follows.

Finally, note that the problem of determining whether a set of customers is an implied-demand cover of facility j with size K is in NP, i.e., it can be checked in polynomial time whether the given set is indeed an implied-demand cover of size K: determining the size of the set and comparing it to K can be done in polynomial time; the same holds for computing the implied demand of the customers in the considered set and comparing it to the capacity of facility j. □

Computational complexity of the separation problems

Next, we prove the open claims at the end of Section 6. Note that all results proving strong NP-hardness depend on solutions for the relaxation of (1). However, these solutions are no extreme points. For optimal solutions that are extreme points, the complexity is still open. An example for an algorithm returning extreme points is the simplex algorithm.

We start with analysing the complexity of the separation problem of inequalities **IC**.

Separating implied-demand cover inequalities is weakly NP-hard if each customer only implies their own assignment to any facility. In this case implied-demand cover inequalities coincide with classical cover inequalities, and the separation of classical cover inequalities is known to be weakly NP-hard (Klabjan et al., 1998). This relation provides a lower bound on the complexity of separating implied-demand cover inequalities. An upper bound on the complexity is provided by Theorem 1, which categorises the problem of computing a minimum implied-demand cover to be strongly NP-hard. This result does not clarify the complexity of separating implied-demand cover inequalities based on a solution for the relaxation of (1): due to the constraints in formulation (1), any solution follows a certain structure, making these instances special cases of the minimum implied-demand cover problem. These special cases are not necessarily strongly NP-hard. The next result gives insights into the complexity of separating inequalities (3).

Theorem 7. *It is strongly NP-hard to find an implied-demand cover $I' \subseteq I$ for a facility $j \in J$ that maximises the violation $\sum_{i \in I'} x_{ij} - |I'| + 1$ of inequality* (3) *for optimal solutions $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in [0,1]^{|I||J|+|J|}$ to the linear relaxation of* (1).

Proof. We prove the claim via a reduction from the strongly NP-hard *minimum set cover* problem (Garey and Johnson, 1979). In the minimum set cover problem, we are given an instance $\mathcal{I}$ consisting of a set of elements $U = \{1, 2, \ldots, n\}$, also known as the *universe*, and a collection S of m subsets of universe U. The goal is to decide whether there is a sub-collection $S^K \subseteq S$ of size $K \in \mathbb{Z}_{\geq 0}$ whose union equals the universe. Without loss of generality, we assume that collection S does not contain subsets s_1, s_2 with $s_1 \subseteq s_2$ and that each element of the collection S contains at least two elements.

We start by presenting a transformation of any minimum set cover instance $\mathcal{I}$ into an SSCFLPCP-instance $\mathcal{I}'$ in a similar manner to the proof of Theorem 1. That is, we introduce one customer for each element $u \in U$ in the universe and one customer for each element $s \in S$. We denote the set of customers with $I = I^U \cup I^S$ with $I^U (I^S)$ denoting the set of customers corresponding to elements in $U(S)$. We set the demand of each customer $i \in I$ to one, i.e., $d_i = 1$. Define the set of facilities as $J = \cup_{i \in I}\{j_i\} \cup \{j\}$. The set of facilities customer $i^u \in I^U$ prefers over j is defined as $J^{<}_{i^u j} = \{j_{i^u}\}$ and $J^{=}_{i^u j} = \{j\}$. The set of facilities customer $i^s \in I^S$ prefers over j is defined as $J^{<}_{i^s j} = \cup_{e \in s} J^{<}_{i^e j} \cup \{j_{i^s}\}$ and $J^{=}_{i^s j} = \{j\}$. Set the opening cost of facility j to $f_j = 0$ and its capacity to $Q_j = K + |U|$. Furthermore, set the capacities of facilities $a \in J \setminus \{j\}$ to $Q_a = 1$ and their opening costs to $f_a = (|U| + |S|)^2$. We set the allocation cost of customer $i \in I$ to any facility $a \in J^{\leq}_{ij}$ to $c_{ia} = 1$ and the remaining allocation costs to two. Instance $\mathcal{I}'$ is an instance of SSCFLPCP.

We define a vector $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ with

$$\bar{y}_k = \begin{cases} 1 & \text{if } k = j \\ (|S| - K)/(|S| + |U|) & \text{otherwise} \end{cases} \qquad \text{and} \qquad \bar{x}_{ik} = \begin{cases} (K + |U|)/(|S| + |U|) & \text{if } k = j \\ (|S| - K)/(|S| + |U|) & \text{if } k = j_i \\ 0 & \text{otherwise} \end{cases}$$

for each customer $i \in I$ and facility $k \in J$. This is a feasible solution for the LP-relaxation of the considered SSCFLPCP-instance $\mathcal{I}'$ as we see next.

- Each customer $i \in I$ is completely assigned among all facilities as $\sum_{k \in J} \bar{x}_{ik} = (K + |U|)/(|S| + |U|) + (|S| - K)/(|S| + |U|) = 1$ holds. Thus, constraints (1b) are met.
- Each facility meets their capacity: for facility j it is $\sum_{i \in I} \bar{x}_{ij} = \sum_{i=1}^{|S|+|U|} (K + |U|)/(|S| + |U|) = K + |U|$; for facility $a \in J \setminus \{j\}$ it is $\sum_{i \in I} \bar{x}_{ia} = (|S| - K)/(|S| + |U|) + \sum_{i \in I \setminus \{a\}} 0 = (|S| - K)/(|S| + |U|) \leq 1 = Q_j$. Thus, constraints (1c) are met.
- Preference constraints (1d) are met. Since facility j is fully open, no customer may be assigned to a facility they like less than j; this is true in the constructed solution. Consider a facility $a \in J \setminus \{j\}$ with location variable value $\bar{y}_a = (|S| - K)/(|S| + |U|)$. Consider a customer $i^u \in I^U$. They prefer facility j_{i^u} over j, and at most $1 - \bar{y}_{j_{i^u}} = 1 - (|S| - K)/(|S| + |U|) = (K + |U|)/(|S| + |U|) = \bar{x}_{i^u j}$ units of i^u's demand may be assigned to a facility customer i^u likes less than j_{i^u}. Hence, the preference constraints are met for all customers in set I^U. We can argue analogously that the preference constraint is met for all customers in I^S.
- It is easy to see that linking constraints (1e) are also met.

In conclusion, vector $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ is a feasible solution for the LP-relaxation of the considered SSCFLPCP-instance $\mathcal{I}'$. Note that there is no solution with strictly lower cost. Since opening facility j is cheap, it will be open in any feasible solution. However, facility j can only serve a demand of $K + |U|$. Thus, the remaining demand to be served at different facilities is $(|S| + |U|) - (|U| + K) = |S| - K$. Opening a few facilities completely or all of them partially does not effect the objective value as all remaining facilities have the same opening costs. Moreover, all customers have the same travel-distance to facility j or the facilities they prefer over j. Hence, the allocation costs do not have an impact on the objective value. In conclusion, this solution is of minimum cost.

Finally, we show that any instance $\mathcal{I}'$ is equivalent to the constructed minimum implied-demand cover instance in the proof of Theorem 1, which we denote with $\mathcal{I}''$. Multiply the cost-parameter of each customer $i \in I$ in instance $\mathcal{I}''$ with factor $1-(K+|U|)/(|S|+|U|)$; modify the goal so that we are asking for a solution with cost of at most $K \cdot (1-(K+|U|)/(|S|+|U|))$. Then, computing a feasible solution for this modified minimum implied-demand cover instance is equivalent to computing a maximum violated implied-demand cover inequality in instance $\mathcal{I}'$. In conclusion, Theorem 7 holds. □

Location-centered inequalities

Finally, we show that determining a maximum violated inequality (9) is strongly NP-hard. The proof is similar to the reduction from 3-Sat to the maximum clique problem in Sipser (1996).

Theorem 8. *It is strongly NP-hard to find an implied-demand cover $I' \subseteq I$ for a facility $j \in J$ that maximises the violation $y_j - \sum_{k\in(\cup_{i\in I'} J_{ij}^{\leq})\setminus\{j\}} y_k$ of inequality (9) for optimal solutions $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in [0,1]^{|I||J|+|J|}$ to the linear relaxation of (1).*

Proof. In order to prove the claim, we show that the decision version of the separation problem is strongly NP-complete. That is, we show that it is strongly NP-complete to determine whether there is a cover $I' \subseteq I$ for a facility $j \in J$ for which the sum of the location variable values of all facilities preferred by at least one customer in cover I' over j is lower than or equal to a value $K \cdot \frac{|I|-(K\cdot(K-1)-2)/2}{2\cdot|I|}$ with $K \in \mathbb{Z}_{\geq 0}$ and $|I| \geq K$. Note that the left-hand side in inequality (9) is constant for any fixed $j \in J$; hence, finding a cover I' with maximum violation of inequality (9) for a fixed j is equivalent to finding a cover I' which minimises the inequality's right-hand side for a fixed j. We prove Theorem 8 via a reduction from the strongly NP-complete *3-Sat* problem (Garey and Johnson, 1979).

In the 3-Sat problem, we are given an instance $\mathcal{I}$ consisting of a set U of boolean variables and a collection C of clauses over U such that each clause $c \in C$ consists of three literals. The goal is to decide whether there is a satisfying truth assignment to the variables such that in every clause $c \in C$ at least one literal is true. We assume that set C consists of $K > 2$ clauses

We start by presenting a transformation of any 3-Sat instance $\mathcal{I}$ into an instance $\mathcal{I}'$ of the separation problem for inequalities (9). Introduce one customer i for any two clause-literal-pairs $\{(c,u),(d,v)\}$ with clauses $c,d \in C$, $c \neq d$, and literals $u \in c, v \in d$ with u, v are not each other's negations. Denote the set of customers with I. The number of customers lies in $\mathcal{O}((3\cdot|C|)^2)$. We set the demand of each customer $i \in I$ to one, i.e., $d_i = 1$. We define the set of facilities J as follows. Introduce one facility j. Furthermore, introduce one facility j_{cu} for each clause-literal-pair (c,u) with $c \in C, u \in c$; this yields $\mathcal{O}(3 \cdot |C|)$ facilities. We define the preference set of a customer $i \in I$ corresponding to clause-literal pair $\{(c,u),(d,v)\}$ as $J_{ij}^{=} = \{j\}$ and $J_{ij}^{<} = \{j_{cu}, j_{dv}\}$. Note that no customer is implied by another customer regarding facility j in this instance; hence, this proof also holds for the case where implied-demand covers are considered. Set the capacity of facility j to $Q_j = (K\cdot(K-1)-2)/2$, and the capacity of the remaining facilities $a \in J \setminus \{j\}$ to $Q_a = |\{i \in I:\ a \in J_{ij}^{<}\}|$. We set the opening costs of facility j to $f_j = 0$ and the opening costs of all remaining facilities to 10. We set the assignment costs of each customer $i \in I$ to each facility $j \in J$ to $c_{ij} = 0$.

Last but not least we construct a fractional solution for the LP-relaxation of (1). Define vector $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ as

$$\bar{y}_a = \begin{cases} 1 & \text{if } a = j, \\ \frac{|I|-Q_j}{2\cdot|I|} & \text{if } a \in J \setminus (\{j\}) \end{cases} \qquad \text{and} \qquad \bar{x}_{ia} = \begin{cases} Q_j/|I| & \text{if } a = j, \\ \frac{|I|-Q_j}{2\cdot|I|} & \text{if } a \in J_{ij}^{<}. \end{cases}$$

for each $i \in I$. This solution corresponds to a feasible solution with optimum value, as we discuss next.

First, observe that the values of all decision variables lie in interval $[0, 1]$. Note that inequality $Q_j = (K \cdot (K-1) - 2)/2 < |I|$ holds, as we see next. Given a feasible 3-Sat instance $\mathcal{I}$ consisting of K clauses, there needs to be at least one pair $\{(c, u), (d, v)\}$ for any two clauses $c, d \in C$, $c \neq d$, with at least one literal $u \in c, v \in d$ in each clause such that u, v are not each other's negations. Otherwise, we can determine in polynomial time that the instance is infeasible. Therefore, I consists of at least $Q_j + 1 = K \cdot (K-1)/2$ customers. Hence, it is $Q_j/|I| < (Q_j + 1)/|I| \leq 1$ as well as $(|I| - Q_j)/(2|I|) \leq 1$ and all values assigned to the decision variables lie in $[0, 1]$.

Second, we show that the constructed LP solution meets all constraints in the relaxation of formulation (1). Each customer is assigned completely to open facilities since inequality

$$\sum_{a \in J} \bar{x}_{ia} = \bar{x}_{ij} + \sum_{a \in J_{ij}^{<}} \bar{x}_{ia} = \frac{Q_j}{|I|} + 2 \cdot \frac{|I| - Q_j}{2 \cdot |I|} = 1$$

holds. Per construction, the capacity of each facility is met: facility j is completely full while the remaining facilities are evenly filled; moreover, each facility in $J \setminus \{j\}$ provides enough capacity to completely serve all customers who prefer it over j. Furthermore, the preference constraints hold. In order to see this, consider a customer $i \in I$. The location variable values of facilities i prefers over j are set to $(|I| - Q_j)/(2 \cdot |I|)$. Due to constraints (1d), the values assigned to the allocation variables may be at most $1 - (|I| - Q_j)/(2 \cdot |I|) = (|I| + Q_j)/(2 \cdot |I|)$. This is greater than $Q_j/|I|$ due to property $Q_j < |I|$. Hence, solution $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ meets the preference constraints. Lastly, note that the linking constraints, i.e., $\bar{x}_{ij} \leq \bar{y}_j$ for $i \in I, j \in J$, which can be seen easily by comparing the values assigned to the location and allocation variables. In conclusion, the considered solution is feasible.

Finally, observe that fractional solution $(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \in [0, 1]^{|I| \cdot |J| + |J|}$ is a cost-minimising solution for instance $\mathcal{I}'$ per construction, as the assignment costs of each customer to each facility are equal to zero and opening facility j incurs the cheapest opening costs.

It remains to prove that there is a satisfying truth assignment for collection C in 3-Sat instance $\mathcal{I}$ *if and only if* there is a subset $I' \subseteq I$ of customers with weight of at least $Q_j + 1$ and value $\sum_{a \in \cup_{i \in I'}} y_a$ lower than or equal to $K \cdot (|I| - Q_j)/(2 \cdot |I|)$ for instance $\mathcal{I}'$.

Suppose C has a satisfying assignment in instance $\mathcal{I}$, i.e., at least one literal is true in each clause. For each clause $c \in C$ consider exactly one arbitrary literal $u \in c$ with value TRUE, and denote the set consisting of these clause-literal-pairs with $\mathcal{T}$. Set $\mathcal{T}$ consists of K elements: one positive clause-literal-pair for each clause. Define a set I' as the set of all customers in I corresponding to any two clause-literal-pairs $\{(c, u), (d, v)\}$ with $(c, u), (d, v) \in \mathcal{T}$, and define the set of facilities preferred by at least one customer in I' over j as $J^{<} = \{a \in J : a \text{ corresponds to a pair } (c, u) \in \mathcal{T}\}$. First, we show that set I' is a cover, i.e., $\sum_{i \in I'} d_i = |I'| \geq Q_j + 1 = K \cdot (K-1)/2$ holds. We add a customer corresponding to each pair of clause-literal-pairs in set $\mathcal{T}$ to set I'. Each of these customers occurs in set I because none of the clause-literal-pairs in set $\mathcal{T}$ satisfies one of the exceptions described before: we only add one clause-literal-pair per clause to set $\mathcal{T}$ and the literals in the corresponding clause-literal-pairs do not have contradictory truth-values - otherwise, at least one boolean variable were to be simultaneously TRUE and FALSE. Since $\mathcal{T}$ consists of K elements, set I' consists of $K \cdot (K-1)/2$ customers, and the sum of the demands of customers in I' exceeds the capacity at facility j. Next, we show that the sum of the location variable values of the facilities preferred by at least one customer in I' adds up to at most $K \cdot (|I| - Q_j)/(2 \cdot |I|)$. Since set $\mathcal{T}$ consists of K elements, set $J^{<}$ consists of K distinct facilities. The sum of the location variable values for those facilities in $J^{<}$ adds up to $\sum_{a \in J^{<}} \bar{y}_a = K \cdot (|I| - Q_j)/(2 \cdot |I|)$. Hence, set I' corresponds to a solution of instance $\mathcal{I}'$.

Conversely, consider a feasible solution for instance $\mathcal{I}'$ consisting of a subset of customers I' with demand of at least $K \cdot (K-1)/2$, and $J^{<} = \cup_{i \in I'} J_{ij}^{<}$ the set of facilities at least one customer in I' strictly prefers over j such that the sum of the location variable values of facilities in $J^{<}$ adds up to at most $K \cdot (|I| - Q_j)/(2 \cdot |I|)$. Due to the choice of the customer demands, set I' consists of at least $K \cdot (K-1)/2$ customers; due to the choice of the location variable values in solution $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$, set $J^{<}$ consists of at most K facilities. We construct a solution for 3-Sat instance $\mathcal{I}$ next. Assign TRUE to literal u in clause c if there is a customer in set I' who strictly prefers the facility corresponding to clause-literal-pair (c, u) over facility j. Observe that there is one literal with value TRUE in each clause. Suppose there is a clause containing two literals with value TRUE. This implies that I' contains a customer corresponding to a pair of clause-literal pairs with the same clause. However, per construction, such a customer does not exist in the first place in set I. Thus, this case does not occur and each of the K clauses has exactly one literal set to TRUE. Suppose two literals which are each other's negations are both set to TRUE. Again, this case does not occur since customers corresponding to such a pair do not occur in I and, therefore, not in I'. Next, we assign the truth-values to the variables in U so that each clause-literal-pair (c, u) corresponding to a facility considered in set $J^<$ is set to TRUE. This assignment to the variables satisfies a truth-assignment because, for each clause, there is one clause-literal-pair corresponding to a facility in $J^<$ and, hence, each clause contains a literal that is assigned a TRUE-value. Thus, our claim follows.

Since we can verify in polynomial time whether the sum of the location variable values of facilities customers in a proposed cover I' prefer over j is lower than or equal to $K \cdot \frac{|I|-(K\cdot(K-1)-2)/2}{2\cdot|I|}$, the decision problem lies in NP. □