

The Strength of Flow Refueling Location Problem Formulations and an Extension to Cyclic Routing*

Nagisa Sugishita^{1*}, Margarida Carvalho² and Ribal Atallah³

^{1*}CIRRELT and Département de sciences de la décision, HEC Montréal,
3000 Chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7, QC,
Canada.

^{2*}CIRRELT and Département d’informatique et de recherche
opérationnel, Université de Montréal, 2920 Chemin de la Tour,
Montréal, H3T 1N8, QC, Canada.

³Institut de Recherche d’Hydro-Quebec, 1800 Boulevard Lionel-Boulet,
Varenes, J3X 1S1, QC, Canada.

*Corresponding author(s). E-mail(s): nagisa.sugishita@hec.ca;

Contributing authors: carvalho@iro.umontreal.ca;

atallah.ribal@hydroquebec.com;

Abstract

The Flow Refueling Location Problem (FRLP) is a stylized model for determining the optimal placement of refueling stations for vehicles with limited travel ranges, such as hydrogen fuel cell vehicles and electric vehicles. A notable extension, the deviation FRLP, accounts for the possibility that drivers may deviate from their preferred routes to refuel or recharge. While solution techniques based on various mathematical programming formulations have been thoroughly explored for this extension, there is a lack of theoretical insights into the relationships and strengths of these formulations.

In this work, for the deviation extension, we study two prominent FRLP formulations from the literature and compare their strengths in terms of linear programming (LP) relaxations. We show that the LP relaxation of one formulation yields a bound that is at least as tight as that of the other, which may explain

*This version of the article has been accepted for publication, after peer review and is subject to Springer Nature’s AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <https://doi.org/10.1007/s10288-026-00617-2>.

its observed superior performance. Building on these insights, we address a common modeling assumption in the FRLP that requires drivers to use the same paths for their outbound and inbound trips. Specifically, we relax this assumption and introduce the cyclic FRLP, where drivers may use different paths in each direction. We show how existing formulations can be naturally extended to accommodate this setting and describe a branch-and-cut algorithm to solve the problem. We provide numerical experiments highlighting the benefits of such asymmetric routing. For example, in an instance based on the Californian network, the original and cyclic FRLPs yield identical solutions when the deviation tolerance is small. However, differences emerge and become more pronounced as the deviation tolerance increases. When a 50% deviation is permitted, the cyclic FRLP satisfies all demand using 30% fewer facilities than the original FRLP.

Keywords: Branch-and-Cut, Integer Programming, Flow Refueling Location Problems, Electric Vehicles

1 Introduction

The Flow Refueling Location Problem (FRLP), introduced by Kuby and Lim [13], extends flow-capturing location models by Hodgson [8] to address the optimal placement of refueling facilities for vehicles with limited range. In the model by Kuby and Lim [13], each demand is associated with an origin, a destination and a flow volume, and is considered served if there are enough refueling facilities to ensure that the drivers can complete their journey with limited-range vehicles. The goal of the FRLP is to maximize the demands that are served.

On the practical level, the growing popularity of hydrogen-powered vehicles and electric vehicles (EVs), as well as governmental incentives in some countries to support their adoption, underscores the relevance of the FRLP (see Arslan and Karaşan [1] and Kadri et al. [15]). Indeed, given the current insufficiency of refueling and charging infrastructure, it is critical to deploy these facilities effectively. In what follows, we use the EVs application to unify our terminology around the FRLP, referring to charging stations for refueling facilities and EVs for limited-range vehicles.

On the methodological side, the literature on the FRLP has mostly focused on integer programming-based approaches. Kuby and Lim [13] introduced a mixed-integer programming (MIP) formulation for the FRLP. However, its size grows rapidly with the problem scale. To address this, Capar and Kuby [4] proposed a more scalable formulation, which was further improved by Capar et al. [5]. Building on this, Arslan and Karaşan [1] applied Benders decomposition to accelerate the solution process. Kadri et al. [15] extended the FRLP to a multi-stage stochastic setting, using a decomposition approach closely related to that of Arslan and Karaşan [1], with an identical structure for the separation subproblem.

The previously discussed works focus on the classical FRLP, where drivers are assumed to strictly follow predetermined (e.g., shortest) paths without deviation. However, in practice, drivers may deviate slightly from their usual routes to access

charging stations. To capture this behavior, Kim and Kuby [12] introduced the deviation FRLP, which allows for multiple feasible paths per origin-destination (OD) pair. Their formulation requires enumerating all paths and, for each path, all combinations of station placements that ensure traversability. This leads to rapid growth in model size as the instance scale increases. Yıldız et al. [17] proposed an improved formulation of the deviation FRLP that eliminates the need to enumerate combinations of station placements. Leveraging techniques from Capar et al. [5], they express path traversability using set-covering constraints. Additionally, they introduced a path-segment-based formulation, solved via a branch-and-price algorithm. While this formulation, together with a branch-and-price approach, yields better performance, the set-covering formulation offers greater flexibility and modeling capabilities for incorporating additional constraints, as it directly uses binary variables indicating station presence at each node. Independently, Arslan et al. [2] and Göpfert and Bock [7] proposed another MIP formulation, using set-covering constraints to express the condition that each demand is served, rather than expressing the condition that each path is traversable, as in Yıldız et al. [17]. This approach eliminates the need for path enumeration and, when solved with a branch-and-cut algorithm, significantly outperforms previous methods, successfully handling instances of practical size.

On the theoretical front, the comparative strengths of the formulations for the deviation FRLP remain underexplored. We aim to address that gap. To be more specific, we compare the MIP formulation developed independently by Arslan et al. [2] and Göpfert and Bock [7] with the formulation by Yıldız et al. [17] based on the set-covering constraints. By analyzing these formulations, we aim to understand the reasons behind the superior computational performance of the state-of-the-art formulation. We note that although Yıldız et al. [17] proposed two formulations, the set-covering formulation and path-segment formulation, we concentrate on the set-covering one, excluding the path-segment formulation due to its limited flexibility for integrating additional constraints.

In addition, we propose an extension of the deviation FRLP, which we call the cyclic FRLP, to support more flexible routing. All prior work mentioned above assumes route symmetry, i.e., drivers must use the same path in both directions. While simplifying the problem, this assumption may lead to suboptimal outcomes in more realistic settings. Furthermore, this assumption requires the underlying transportation network to be undirected. In the cyclic FRLP, drivers are allowed to choose different outbound and inbound routes, enabling more efficient utilization of transportation infrastructure. Notably, the formulation of Arslan et al. [2] and Göpfert and Bock [7] can be naturally extended to accommodate this generalization. We develop a solution method for the cyclic FRLP by introducing a labeling algorithm that searches for a cycle (round-trip route) that can be repeatedly traversed by vehicles with limited range.

Using the Californian network developed by Arslan et al. [3], we demonstrate the difference between the deviation and cyclic FRLP. Furthermore, to complement our experiments and highlight the value of the cyclic FRLP, we consider a variant that minimizes the number of charging stations while covering a specified portion or all of the demand. This variant has been studied in the literature under the symmetric routing assumption (e.g., Capar et al. [5] and Kınay et al. [10]) and can be readily

addressed by adapting formulations such as those by Arslan et al. [2] and Göpfert and Bock [7]. Including this variant allows us to directly compare the added benefit of incorporating cyclic routes against the classical FRLP with symmetric assumptions. For example, in the instances based on the Californian network of Arslan et al. [3], we observe no difference between the two models when the deviation tolerance is small. However, differences emerge as the deviation tolerance increases. When the tolerance reaches 50%, the cyclic FRLP is able to serve all demand with 30% fewer charging stations than the deviation FRLP.

We note that our approach and techniques are flexible and can be incorporated into other models. For example, Kinay et al. [11] recently extended the deviation FRLP to incorporate both sizing decisions and service availability constraints. The resulting model is formulated as a bilevel programming problem, and a heuristic based on decomposition algorithms was proposed. In their model, each user is assumed to follow the same route for both outbound and inbound trips. Using the techniques developed in this paper, it is straightforward to extend their model and solution approach to account for cyclic routing.

The paper is organized as follows. In Section 2, we introduce the necessary notation and present two formulations of the FRLP. Section 3 compares the strengths of these formulations via their LP relaxations. Section 4 introduces the cyclic FRLP, which accommodates more flexible routing. Numerical experiments are also provided to illustrate the differences between the original and cyclic FRLP. Finally, Section 5 concludes the paper.

2 Formulations of the Deviation FRLP

In this section, we present the arguments used to define the deviation FRLP and recap two key mathematical formulations of it from the literature. The notation, namely the parameters and variables used in the formulations in this section, is summarized in Table 1.

Table 1: Symbols and their description

Symbol	Description
N	Set of nodes
A	Set of edges
G	Graph (N, A)
Q	Set of demands
R_q	Set of paths used by drivers of demand $q \in Q$
R	Union of R_q for $q \in Q$
X	Set of feasible placement of charging stations
$\bar{X}$	LP relaxation of X
C_q^{agg}	Family of node sets to express the condition of $q \in Q$ to be served
C_r^{disagg}	Family of node sets to express the condition of $r \in R$ to be traversable
f_q	Demand volume of $q \in Q$
x_j	Binary variable that is 1 if and only if there is a charging station at node j
y_q	Binary variable that is 1 if and only if $q \in Q$ is served
z_{qr}	Binary variable that is 1 if and only if $r \in R_q$ is used by $q \in Q$

Let $G = (N, A)$ be an undirected graph, where N is the set of nodes and A is the set of edges. We consider the installation of charging stations in nodes of N . For each $j \in N$, we use the binary variable x_j to indicate the installation of a charging station at node j . We define the set of feasible charging station placements as $X = \bar{X} \cap \{0, 1\}^{|N|}$, where $\bar{X}$ is a polytope. For example, in our numerical experiments in Section 4.3, X represents charging station locations that satisfy a budget constraint, with

$$\bar{X} = \left\{ x : \sum_{j \in N} x_j \leq a \right\},$$

where a denotes the budget. Our discussion, however, is independent of the particular choice of X .

We denote a set of demands by Q . Each demand $q \in Q$ is associated with an origin node, a destination node, a flow volume f_q , and a set R_q of paths from the origin to the destination. The set R_q is the set of paths that drivers of q are willing to use. Demand $q \in Q$ is considered served if there is at least one path in R_q that is traversable without running out of battery. Since the travel range of a vehicle is limited, it may require charging the EV multiple times to complete the travel. We assume that the graph does not contain edges longer than the travel range of a vehicle and $R_q \neq \emptyset$ for each $q \in Q$. Note that we can remove excessively long edges or demands q with empty R_q without modifying the set of feasible solutions, which is defined later.

In the classical FRLP proposed by Kuby and Lim [13], each demand is associated with a single shortest path. Thus, for each $q \in Q$, the set R_q is a singleton containing only the shortest path between the origin and destination. Kim and Kuby [12] introduced the deviation FRLP, an extension of the classical FRLP in which multiple paths are considered for each origin-destination (OD) pair. For example, in their work, Kim and Kuby [12] defined R_q to include all paths whose lengths are below a specified deviation threshold, for each $q \in Q$. Figure 1 illustrates the problem definition.

As discussed by Capar et al. [5], given a path and the travel range of a vehicle, one can construct a family $C \subseteq \mathcal{P}(N)$ of node sets such that the path is traversable repeatedly with the vehicle if and only if $\sum_{j \in S} x_j \geq 1$ for all $S \in C$, where $\mathcal{P}(N)$ is the power set of N . In words, the path is traversable if and only if each $S \in C$ contains at least one node with a charging station. Thus, for each $q \in Q$ and $r \in R_q$, we can find a family $C_r^{\text{disagg}} \subseteq \mathcal{P}(N)$ of node sets such that r is traversable if and only if $\sum_{j \in S} x_j \geq 1$ for all $S \in C_r^{\text{disagg}}$. Demand $q \in Q$ is served if and only if there exists at least one $r \in R_q$ such that $\sum_{j \in S} x_j \geq 1$ for all $S \in C_r^{\text{disagg}}$. Thus, we can obtain the following formulation of the deviation FRLP:

$$\begin{aligned} \max_{x \in X, z \in \{0, 1\}^{|Q| \times |R|}} \quad & \sum_{q \in Q} \sum_{r \in R_q} f_q z_{qr} \\ \text{s.t.} \quad & \sum_{r \in R_q} z_{qr} \leq 1, \quad \forall q \in Q, \\ & \sum_{j \in S} x_j \geq z_{qr}, \quad \forall q \in Q, r \in R_q, S \in C_r^{\text{disagg}}, \end{aligned} \tag{1}$$

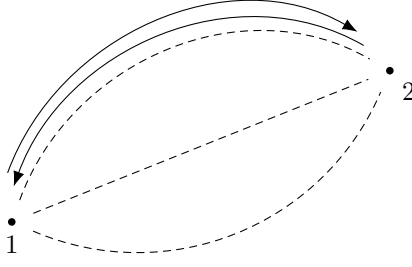


Fig. 1: OD pair with origin and destination at nodes 1 and 2, respectively; dashed lines indicate possible OD paths, and solid directed lines show a feasible route. In the classical and deviation FRLP, a demand is served if there is a path that is traversable repeatedly without a vehicle running out of battery. Drivers are required to use the same path for both outbound and inbound trips.

where $R = \cup_{q \in Q} R_q$ and, for each $q \in Q$, the auxiliary variable z_{qr} is 1 if path r is used and 0 otherwise. Formulation (1) was proposed by Yıldız et al. [17] as formulation (14) in their paper. As discussed in the introduction, Yıldız et al. [17] also proposed a path-segment formulation that can be solved with a branch-and-price approach. We do not consider the path-segment formulation since it is less flexible than those considered in this section.

Example 1. Consider the graph G shown in Figure 2. Suppose the travel range is d , and there is a single OD pair q_1 , with origin and destination at nodes 1 and 4, respectively. Assume $R_{q_1} = \{(1, 2, 4), (1, 2, 3, 4)\}$. Let us denote $r_1 = (1, 2, 4)$ and $r_2 = (1, 2, 3, 4)$.

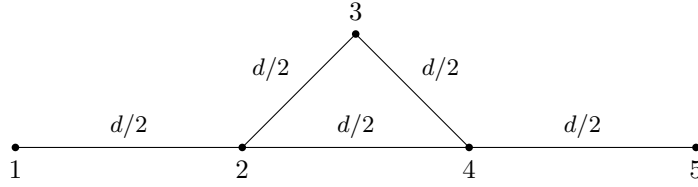


Fig. 2: Graph G , where d is the travel range

It can be shown that path r_1 is traversable if and only if there is at least one charging station at each of the following sets of nodes:

- nodes 1 or 2,
- nodes 2 or 4,
- nodes 4 or 5.

See Capar et al. [5] for the derivation. Thus, $C_{r_1}^{disagg} = \{\{1, 2\}, \{2, 4\}, \{4, 5\}\}$. Similarly, we have $C_{r_2}^{disagg} = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$. Therefore, formulation (1)

becomes

$$\begin{aligned}
f'_1 = & \max_{x \in X, z_{q_1} \in \{0,1\}^2} f_{q_1} z_{q_1 r_1} + f_{q_1} z_{q_1 r_2} \\
& \text{s.t. } z_{q_1 r_1} + z_{q_1 r_2} \leq 1, \\
& x_1 + x_2 \geq z_{q_1 r_1}, x_2 + x_4 \geq z_{q_1 r_1}, x_4 + x_5 \geq z_{q_1 r_1}, \\
& x_1 + x_2 \geq z_{q_1 r_2}, x_2 + x_3 \geq z_{q_1 r_2}, x_3 + x_4 \geq z_{q_1 r_2}, x_4 + x_5 \geq z_{q_1 r_2}.
\end{aligned}$$

Next, we derive an alternative formulation. While this formulation has been previously introduced, we provide a new and straightforward derivation here to keep the exposition self-contained. To this end, let us look at another way to formulate the condition of demand $q \in Q$ being served. For $q \in Q$, let $C_q^{\text{agg}} = \{\cup_{r \in R_q} S_r : S_r \in C_r^{\text{disagg}}, r \in R_q\}$, i.e., each element of C_q^{agg} is the union of one S_r for each $r \in R$. The set C_q^{agg} has the following property.

Lemma 2. *Demand $q \in Q$ is served if and only if $\sum_{j \in S} x_j \geq 1$ for all $S \in C_q^{\text{agg}}$.*

Proof. Suppose that $q \in Q$ is served. Then, there exists $r' \in R_q$ such that

$$\sum_{j \in S'} x_j \geq 1, \quad \forall S' \in C_{r'}^{\text{disagg}}.$$

Let $S \in C_q^{\text{agg}}$. By definition, we can write S as

$$S = \bigcup_{r \in R_q} S_r,$$

where $S_r \in C_r^{\text{disagg}}$ for each $r \in R_q$. It follows that

$$\sum_{j \in S} x_j \geq \sum_{j \in S_{r'}} x_j \geq 1.$$

Next, suppose that $q \in Q$ is not served. For any $r \in R_q$, there exists $S'_r \in C_r^{\text{disagg}}$ such that

$$\sum_{j \in S'_r} x_j = 0.$$

Let

$$S' = \bigcup_{r \in R_q} S'_r.$$

Then, $S' \in C_q^{\text{agg}}$ and

$$\sum_{j \in S'} x_j \leq \sum_{r \in R_q} \sum_{j \in S_r} x_j = 0.$$

□

Therefore, we obtain the second formulation of the deviation FRLP for maximizing the amount of served demands as

$$\max_{x \in X, y \in \{0,1\}^{|Q|}} \sum_{q \in Q} f_q y_q \quad (2a)$$

$$\text{s.t.} \quad \sum_{j \in S} x_j \geq y_q, \quad \forall q \in Q, S \in C_q^{\text{agg}}, \quad (2b)$$

where y_q is 1 if demand $q \in Q$ is served, and 0 otherwise. This is the formulation used by Arslan et al. [2] and Göpfert and Bock [7]. Note that when R_q is a singleton for each $q \in Q$, formulations (1) and (2) coincide and are equivalent to the one used by Capar et al. [5].

Example 3. Let us consider the graph G from Example 1 again. We have

$$C_{q_1}^{\text{agg}} = \left\{ \{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4\}, \{1, 2, 4, 5\}, \{1, 2, 4\}, \{2, 3, 4\}, \right. \\ \left. \{2, 4, 5\}, \{2, 3, 4, 5\}, \{3, 4, 5\}, \{4, 5\} \right\}. \quad (3)$$

Therefore, formulation (2) is

$$f'_2 = \max_{x \in X, y \in \{0,1\}} f_{q_1} y_{q_1} \quad (4)$$

$$\begin{aligned} \text{s.t.} \quad & x_1 + x_2 \geq y_{q_1}, x_1 + x_2 + x_3 \geq y_{q_1}, x_1 + x_2 + x_3 + x_4 \geq y_{q_1}, \\ & x_1 + x_2 + x_4 + x_5 \geq y_{q_1}, x_1 + x_2 + x_4 \geq y_{q_1}, x_2 + x_3 + x_4 \geq y_{q_1}, \\ & x_2 + x_4 + x_5 \geq y_{q_1}, x_2 + x_3 + x_4 + x_5 \geq y_{q_1}, x_3 + x_4 + x_5 \geq y_{q_1}, \\ & x_4 + x_5 \geq y_{q_1}. \end{aligned}$$

Formulation (2) has fewer variables than formulation (1), but many constraints since

$$|C_q^{\text{agg}}| = \prod_{r \in R_q} |C_r^{\text{disagg}}|,$$

which grows rapidly as the number of paths in R_q increases. However, some of the constraints are redundant. For instance, the following three constraints of (4) imply the others:

$$x_1 + x_2 \geq y_{q_1}, x_2 + x_3 + x_4 \geq y_{q_1}, x_4 + x_5 \geq y_{q_1}.$$

Note that these constraints correspond to the minimal node sets $\{\{1, 2\}, \{2, 3, 4\}, \{4, 5\}\}$ in $C_{q_1}^{\text{agg}}$. We can therefore formalize this reasoning. To that end, let us define

$$\bar{P}(C) = \left\{ (x, y) \in [0, 1]^{|N|+1} : \sum_{j \in S} x_j \geq y, \quad \forall S \in C \right\},$$

where $C \subseteq \mathcal{P}(N)$ is a family of node sets.

Proposition 4. Let $C \subseteq \mathcal{P}(N)$ be a family of node sets and $C' \subseteq C$ denote the subfamily of minimal ones: $C' = \{S' \in C : \nexists S \in C \text{ such that } S \subset S'\}$.

- (i) It holds $\bar{P}(C) = \bar{P}(C')$.
- (ii) Let $S \in C'$, then there exists $(x, y) \in \{0, 1\}^{|N|+1}$ such that $(x, y) \in \bar{P}(C \setminus \{S\}) \setminus \bar{P}(C)$.
- (iii) Let $S \in C'$, then $\sum_{j \in S} x_j \geq y_q$ is facet-defining for $\text{conv}(\bar{P}(C) \cap \{0, 1\}^{|N|+1})$.

Proof. (i) It is clear that $\bar{P}(C) \subseteq \bar{P}(C')$ since $C' \subseteq C$, so we will show $\bar{P}(C) \supseteq \bar{P}(C')$. To derive a contradiction, suppose that $\bar{P}(C) \not\supseteq \bar{P}(C')$ and let $(x, y) \in \bar{P}(C') \setminus \bar{P}(C)$. Since $(x, y) \in \bar{P}(C')$, $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Furthermore, since $(x, y) \notin \bar{P}(C)$, there exists $S \in C$ such that $\sum_{j \in S} x_j < y$. However, we can find $S' \in C'$ such that $S' \subseteq S$, for which we have $\sum_{j \in S'} x_j \leq \sum_{j \in S} x_j < y$. This contradicts the assumption that $(x, y) \in \bar{P}(C')$. Thus, it follows that $\bar{P}(C) \supseteq \bar{P}(C')$.

(ii) Let S be any node set in C' . Define

$$x_j = \begin{cases} 0 & \text{if } j \in S, \\ 1 & \text{otherwise.} \end{cases}$$

Then, for any $S' \in C' \setminus \{S\}$, we have $S' \setminus S \neq \emptyset$ (otherwise S is not minimal and $S \notin C'$) and thus

$$\sum_{j \in S'} x_j \geq 1.$$

Therefore, $(x, 1) \in \bar{P}(C \setminus \{S\})$. However, since

$$\sum_{j \in S} x_j = 0,$$

we have $(x, 1) \notin \bar{P}(C)$.

- (iii) Observe that $\text{conv}(\bar{P}(C) \cap \{0, 1\}^{|N|+1})$ is full dimensional and that the inequality is valid for $\text{conv}(\bar{P}(C) \cap \{0, 1\}^{|N|+1})$. The following set contains $|N| + 1$ affinely independent points that satisfy the inequality with equality:

$$\left\{ (0, 0) \right\} \cup \left\{ (e_j, 0) : j \in N \setminus S \right\} \cup \left\{ \left(e_j + \sum_{j' \in N \setminus S} e_{j'}, 1 \right) : j \in S \right\},$$

where $e_j \in \mathbb{R}^{|N|}$ is the j -th unit vector. Therefore, the given inequality is facet-defining. □

In particular, Proposition 4 implies that, for $q \in Q$, we can drop the node sets from C_q^{agg} that are not minimal, reducing the number of constraints (2b). However, dropping a minimal node set from C_q^{agg} may introduce extra integer points. We also note that Assertion (iii) in Proposition 4 concerns the polytope $\bar{P}(C_q^{\text{agg}})$, which is associated with a single OD pair $q \in Q$. Proposition 6 by Arslan et al. [2] gives

the condition under which constraint (2b) is facet-defining for the convex hull of the feasible region of formulation (2).

For $q \in Q$, when R_q consists of a single, simple path, the polytope $\bar{P}(C_q^{\text{agg}})$ has the following property.

Proposition 5. *Let $q \in Q$. Suppose R_q consists of a single, simple path. Then, $\bar{P}(C_q^{\text{agg}})$ is integral (i.e. all the vertexes of $\bar{P}(C_q^{\text{agg}})$ are integral).*

To show Proposition 5, we first present two lemmata.

Lemma 6. *Let $C \subseteq \mathcal{P}(N)$ be a family of node sets and (x^*, y^*) be a vertex of $\bar{P}(C)$. Then, we have $y^* \in \{0, 1\}$.*

Proof. Let us rewrite the definition of polyhedron $\bar{P}(C)$ as

$$\bar{P}(C) = \left\{ (x, y) \in \mathbb{R}^{|N|+1} : \begin{array}{ll} \sum_{j \in S} x_j \geq y, & S \in C, \\ x_j \geq 0, & j \in N, \\ x_j \leq 1, & j \in N, \\ a_j y \geq b_j, & j \in \{1, 2\} \end{array} \right\},$$

where $a = (1, -1)$ and $b = (0, -1)$. There are $|N| + 1$ inequalities that are active at (x^*, y^*) and whose coefficients are linearly independent. That is, we can find $I_1 \subseteq C$, $I_2, I_3 \subseteq N$ and $I_4 \subseteq \{1, 2\}$ such that $|I_1| + |I_2| + |I_3| + |I_4| = |N| + 1$ and (x^*, y^*) is a unique solution to

$$\begin{cases} \sum_{j \in S} x_j = y, & S \in I_1, \\ x_j = 0, & j \in I_2, \\ x_j = 1, & j \in I_3, \\ a_j y = b_j, & j \in I_4. \end{cases} \quad (5)$$

If $1 \in I_4$, $y \geq 0$ is active at (x^*, y^*) and $y^* = 0$. Similarly, if $2 \in I_4$, $y \leq 1$ is active at (x^*, y^*) and $y^* = 1$. In the following, suppose that $I_4 = \emptyset$.

If $I_3 \cap S = \emptyset$ for any $S \in I_1$, then

$$x_j = \begin{cases} 1 & \text{if } j \in I_3, \\ 0 & \text{otherwise,} \end{cases} \quad y = 0$$

is a solution to (5). Since (x^*, y^*) is the unique solution to (5), we have $y^* = 0$.

If there exists $S \in I_1$ such that $I_3 \cap S \neq \emptyset$, then

$$y^* = \sum_{j \in S} x_j^* \geq \sum_{j \in S \cap I_3} x_j^* = |S \cap I_3| \geq 1,$$

where we used the assumption that $\sum_{j \in S} x_j \geq y$ and $x_j \leq 1$ for $j \in I_3$ are active at (x^*, y^*) . Since $(x^*, y^*) \in \bar{P}(C)$, we have $y^* \leq 1$. Therefore, $y^* = 1$. $\square$

Lemma 7. *Let $q \in Q$ and define*

$$U(C_q^{\text{agg}}) = \left\{ x \in [0, 1]^{|N|} : \sum_{j \in S} x_j \geq 1, \quad \forall S \in C_q^{\text{agg}} \right\}.$$

If R_q contains a single, simple path, $U(C_q^{\text{agg}})$ is integral.

Proof. By relabeling the nodes N if necessary, we can assume that $N = \{1, 2, \dots, |N|\}$ and the path in R_q is $(1, 2, \dots, t)$. The argument of Capar et al. [5] ensures that any $S \in C_q^{\text{agg}}$ is a set of consecutive integers. That is, if we write

$$U(C_q^{\text{agg}}) = \left\{ x \in [0, 1]^{|N|} : Bx \geq 1 \right\},$$

where $B \in \{0, 1\}^{|C_q^{\text{agg}}| \times |N|}$, 1s in each row of B occur in consecutive positions. It is known that B is totally unimodular (see Fulkerson and Gross [6]). It is straightforward to see that all vertexes of $U(C_q^{\text{agg}})$ are integral. $\square$

Proof of Proposition 5. Let us denote $\{0, 1\}^{|N|} = \{\beta_i : i = 1, 2, \dots, 2^{|N|}\}$ and the set of vertexes of a polyhedron P by $V(P)$. Then, Lemma 6 implies that

$$V(\bar{P}(C_q^{\text{agg}})) = \{(\beta_i, 0) : i = 1, 2, \dots, 2^{|N|}\} \cup \{(x, 1) : x \in V(U(C_q^{\text{agg}}))\}.$$

Now invoke Lemma 7 to show that $V(U(C_q^{\text{agg}}))$ only contains integral points. $\square$

3 Strength of Formulations of the Deviation FRLP

In Section 2, two formulations of the deviation FRLP were presented. In this section, we compare their strengths in terms of the tightness of the bounds given by their LP relaxations.

To facilitate the discussion, we rewrite formulations (1) and (2). Following the arguments of Kuby and Lim [13], it can be shown that the optimal objective value of (1) remains unchanged if the integrality constraints on z are relaxed (see Appendix A). Thus, we can rewrite formulation (1) as

$$\begin{aligned} & \max_{x \in X, z \in [0, 1]^{|Q| \times |R|}} \left\{ \sum_{q \in Q} \sum_{r \in R_q} f_q z_{qr} : \sum_{r \in R_q} z_{qr} \leq 1, \quad \forall q \in Q, \right. \\ & \quad \left. \sum_{j \in S} x_j \geq z_{qr}, \quad \forall q \in Q, r \in R_q, S \in C_r^{\text{disagg}} \right\} \\ &= \max_{x \in X} \sum_{q \in Q} \max_{z_q \in [0, 1]^{|R_q|}} \left\{ \sum_{r \in R_q} f_q z_{qr} : \sum_{r \in R_q} z_{qr} \leq 1, \quad \forall r \in R_q, S \in C_r^{\text{disagg}} \right\} \\ &= \max_{x \in X} \sum_{q \in Q} v_q^{\text{disagg}}(x), \end{aligned} \tag{6}$$

where for each $q \in Q$ and $x \in [0, 1]^{|N|}$, we have

$$v_q^{\text{disagg}}(x) = \max_{z_q \in [0, 1]^{|R_q|}} \left\{ \sum_{r \in R_q} f_q z_{qr} : \sum_{r \in R_q} z_{qr} \leq 1, \sum_{j \in S} x_j \geq z_{qr}, \quad \forall r \in R_q, S \in C_r^{\text{disagg}} \right\}. \tag{7}$$

We have $x \in X$ is optimal for (1) if and only if it is optimal for (6). Note that being the value function of a maximization LP, v_q^{disagg} is concave for each $q \in Q$.

Similarly, it can be shown that the optimal objective value does not change if we relax the integrality constraint of y_q for $q \in Q$ in formulation (2). Therefore, we can rewrite formulation (2) as

$$\max_{x \in X} \sum_{q \in Q} v_q^{\text{agg}}(x), \quad (8)$$

where, for each $q \in Q$ and $x \in [0, 1]^{|N|}$, we have

$$v_q^{\text{agg}}(x) = \max_{y_q \in [0, 1]} \left\{ f_q y_q : \sum_{j \in S} x_j \geq y_q, \quad \forall S \in C_q^{\text{agg}} \right\}. \quad (9)$$

Again, $x \in X$ is optimal for (2) if and only if it is optimal for (8). Note that v_q^{agg} is a concave function for each $q \in Q$.

For any $x \in \{0, 1\}^{|N|}$ and $q \in Q$, both $v_q^{\text{disagg}}(x)$ and $v_q^{\text{agg}}(x)$ model the amount of served demand from q . Thus, we have

$$v_q^{\text{disagg}}(x) = v_q^{\text{agg}}(x) = \begin{cases} f_q & \text{if demand } q \text{ is served,} \\ 0 & \text{otherwise,} \end{cases}$$

and thus

$$\max_{x \in X} \sum_{q \in Q} v_q^{\text{disagg}}(x) = \max_{x \in X} \sum_{q \in Q} v_q^{\text{agg}}(x).$$

However, functions v_q^{disagg} and v_q^{agg} may not agree on $x \in \bar{X}$ (i.e. if x is fractional), and in particular, the LP relaxations of (1) and (2) may give different upper bounds.

In general, for each $q \in Q$, if we have a concave function v_q on $[0, 1]^{|N|}$ that agrees with v_q^{agg} on $\{0, 1\}^{|N|}$, we can replace v_q^{agg} with v_q without affecting the optimal objective value nor the structure of the problem:

$$\max_{x \in X} \sum_{q \in Q} v_q^{\text{agg}}(x) = \max_{x \in X} \sum_{q \in Q} v_q(x). \quad (10)$$

Let us define

$$K_q = \left\{ v : v \text{ is a concave function on } [0, 1]^{|N|} \text{ and } v_q^{\text{agg}}(x') = v(x'), \quad \forall x' \in \{0, 1\}^{|N|} \right\}. \quad (11)$$

We have $K_q \neq \emptyset$ since $v_q^{\text{agg}} \in K_q$. Now, let

$$v_q^{\text{tight}}(x) = \inf_{v \in K_q} \{v(x)\}.$$

Being the point-wise infimum of concave functions, v_q^{tight} is concave and thus $v_q^{\text{tight}} \in K_q$. The following lemma gives an alternative definition of v_q^{tight} .

Lemma 8. Let us denote $\{0, 1\}^{|N|} = \{\beta_i : i = 1, 2, \dots, 2^{|N|}\}$, that is, $\beta_i \in \{0, 1\}^{|N|}$ for each i and $\beta_i \neq \beta_j$ for any $i \neq j$. Then, we have

$$v_q^{tight}(x) = \max_{\alpha} \left\{ \sum_{i=1}^{2^{|N|}} \alpha_i v_q^{agg}(\beta_i) : \sum_{i=1}^{2^{|N|}} \alpha_i \beta_i = x, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\}.$$

Proof. Let us write

$$v_q(x) = \max_{\alpha} \left\{ \sum_{i=1}^{2^{|N|}} \alpha_i v_q^{agg}(\beta_i) : \sum_{i=1}^{2^{|N|}} \alpha_i \beta_i = x, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\},$$

where v_q is polyhedral and concave since it is the value function of an LP. Furthermore, $v_q(x) = v_q^{agg}(x)$ for any $x \in \{0, 1\}^{|N|}$. We show that for any $v'_q \in K_q$, we have $v_q \leq v'_q$, which implies that $v_q^{tight} = v_q$. Fix any $v'_q \in K_q$. To derive a contradiction, suppose there exists $x' \in [0, 1]^{|N|}$ such that $v_q(x') > v'_q(x')$. Let α' be the solution of the LP in the definition in v_q given $x = x'$. In particular, it follows that

$$\sum_{i=1}^{2^{|N|}} \alpha'_i \beta_i = x'.$$

We have

$$v'_q(x') < v_q(x') = \sum_{i=1}^{2^{|N|}} \alpha'_i v_q^{agg}(\beta_i) = \sum_{i=1}^{2^{|N|}} \alpha'_i v'_q(\beta_i),$$

contradicting the assumption that v'_q is concave. Therefore, we have $v_q \leq v'_q$. $\square$

In particular, Lemma 8 implies that v_q^{tight} is polyhedral for any $q \in Q$.

The set K_q consists of concave functions that model the amount of served demand corresponding to $q \in Q$. The tightest function in K_q is v_q^{tight} in the sense that it gives the best continuous (LP) relaxations among K_q :

$$\max_{x \in \bar{X}} \sum_{q \in Q} v_q^{tight}(x) \leq \max_{x \in \bar{X}} \sum_{q \in Q} v_q(x), \quad \forall v_q \in K_q, q \in Q.$$

Now we are ready to state our first comparison result.

Proposition 9. Let $q \in Q$ and $x \in [0, 1]^{|N|}$.

(i) It holds

$$v_q^{tight}(x) \leq v_q^{agg}(x) \leq v_q^{disagg}(x).$$

(ii) If R_q consists of a single path, we have

$$v_q^{agg}(x) = v_q^{disagg}(x).$$

(iii) If R_q consists of a single, simple path, we have

$$v_q^{\text{tight}}(x) = v_q^{\text{agg}}(x) = v_q^{\text{disagg}}(x). \quad (12)$$

Proof. (i) The inequality $v_q^{\text{tight}}(x) \leq v_q^{\text{agg}}(x)$ follows from the definition, thus we will show $v_q^{\text{agg}}(x) \leq v_q^{\text{disagg}}(x)$. If $v_q^{\text{disagg}}(x) = f_q$, the inequality follows from $v_q^{\text{agg}}(x) \leq f_q$. Thus, assume $v_q^{\text{disagg}}(x) < f_q$. Let $z \in [0, 1]^{|R|}$ and $y_q \in [0, 1]$ be the solution of the LPs in (7) and (9). Since $v_q^{\text{disagg}}(x) < f_q$, we have

$$\sum_{r \in R_q} z_{qr} < 1.$$

For each $r \in R_q$, there exists $S_{qr} \in C_q^{\text{disagg}}$ such that

$$\sum_{j \in S_{qr}} x_j = z_{qr},$$

since otherwise we can increase z_{qr} and improve the objective value. Now, let

$$S_q = \bigcup_{r \in R_q} S_{qr}.$$

We have

$$v_q^{\text{agg}}(x) = f_q y_q \leq f_q \sum_{j \in S_q} x_j \leq f_q \sum_{r \in R_q} \sum_{j \in S_{qr}} x_j = f_q \sum_{r \in R_q} z_{qr} = v_q^{\text{disagg}}(x).$$

- (ii) The equality follows from the definitions.
 (iii) We have seen that $v_q^{\text{agg}}(x) = v_q^{\text{disagg}}(x)$, so we will show $v_q^{\text{tight}}(x) = v_q^{\text{agg}}(x)$. Observe that

$$\text{subgraph}(v_q^{\text{agg}}) = \left\{ (x, \theta) \in [0, 1]^{|N|} \times \mathbb{R} : f_q \sum_{j \in S} x_j \geq \theta, S \in C_q^{\text{agg}} \right\}.$$

Similar to Proposition 5, we can show that if R_q consists of a single, simple path, all the vertices of $\text{subgraph}(v_q^{\text{agg}})$ are $\{(x, v_q^{\text{agg}}(x)) : x \in \{0, 1\}^{|N|}\}$. Thus,

$$\text{subgraph}(v_q^{\text{agg}}) = \text{conv}\{(x, v_q^{\text{agg}}(x)) : x \in \{0, 1\}^{|N|}\} + \{(0, r) : r \leq 0\}.$$

Lemma 8 implies that $\text{subgraph}(v_q^{\text{tight}})$ has the same set of vertexes and

$$\text{subgraph}(v_q^{\text{tight}}) = \text{conv}\{(x, v_q^{\text{agg}}(x)) : x \in \{0, 1\}^{|N|}\} + \{(0, r) : r \leq 0\}.$$

Thus, $v_q^{\text{agg}}(x) = v_q^{\text{tight}}(x)$ for any $x \in [0, 1]^{|N|}$. □

Proposition 9 implies that the LP relaxation of (1) may yield a weaker (i.e., looser) upper bound than that of (2). A natural question, then, is how large the difference between these two bounds can be. The following proposition shows that the relative gap between the bounds can be arbitrarily large.

Proposition 10. (i) For any $\gamma > 1$, there exists a deviation FRLP instance such that

$$\underbrace{\gamma \cdot \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{agg}(x)}_{\text{UB from LP relaxation of (2)}} \leq \underbrace{\max_{x \in \bar{X}} \sum_{q \in Q} v_q^{disagg}(x)}_{\text{UB from LP relaxation of (1)}}$$

(ii) For any $\gamma > 1$, there exists a deviation FRLP instance such that

$$\underbrace{\gamma \cdot \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{tight}(x)}_{\text{UB from LP relaxation with } v^{tight}} \leq \underbrace{\max_{x \in \bar{X}} \sum_{q \in Q} v_q^{agg}(x)}_{\text{UB from LP relaxation of (2)}}$$

Proof. (i) Fix $\gamma > 1$ and let n be a positive integer such that

$$\gamma \leq \frac{n+1}{2}.$$

Consider the following graph $G = (N, A)$, where $N = \{1, 2, \dots, 2n\}$ and

$$A = \{(1, j), (j, n+2) : j = 2, 3, \dots, n+1\} \cup \{(j, j+1) : j = n+2, n+3, \dots, 2n-1\}.$$

Suppose each edge is of length d , where d is the travel range. The graph is shown in Figure 3.

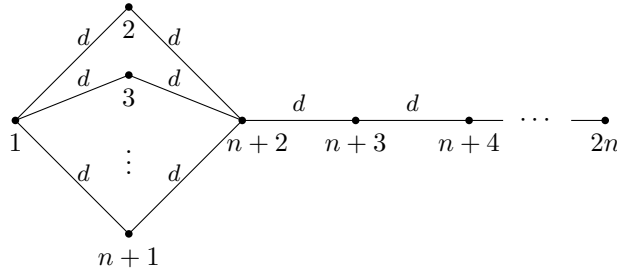


Fig. 3: Graph G , where d is the travel range

Let $Q = \{1\}$, where the origin and the destination of the demand are 1 and $2n$, respectively, and R_1 contains all the $(1, 2n)$ -path of length nd :

$$R_1 = \{r_j := (1, j+1, n+2, n+3, \dots, 2n) : j = 1, 2, \dots, n\}.$$

Finally, suppose that

$$\bar{X} = \left\{ x \in \mathbb{R}^{2n} : \sum_{j=1}^{2n} x_j \leq 2 \right\}.$$

On this instance, following the argument by Capar et al. [5], for $j = 1, 2, \dots, n$, we have

$$C_{r_j}^{\text{disagg}} = \left\{ \{1\}, \{j+1\}, \{n+2\}, \{n+3\}, \dots, \{2n\} \right\}.$$

That is, each path $r \in R_1$ is traversable if and only if all nodes in r have charging stations. Let

$$x'_j = \frac{1}{n}, \quad j = 1, 2, \dots, 2n.$$

Note that $x' \in \bar{X}$. Since

$$z'_{1r} = \frac{1}{n}, \quad r \in R_1,$$

is feasible for the LP in the RHS of (7) for $x = x'$,

$$\max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{disagg}}(x) \geq v_1^{\text{disagg}}(x') \geq \sum_{r \in R_1} f_1 z'_{1r} = f_1.$$

Next, we will obtain a bound on v_1^{agg} . It is clear that

$$\{1\}, \{2, 3, \dots, n+1\}, \{n+2\}, \{n+3\}, \dots, \{2n\} \in C_1^{\text{agg}}.$$

Therefore, for any $x \in \bar{X}$, we have

$$\begin{aligned} v_1^{\text{agg}}(x) &= \max_{y_1 \in [0,1]} \left\{ f_1 y_1 : \sum_{j \in S} x_j \geq y_1, \quad \forall S \in C_1^{\text{agg}} \right\} \\ &= \max_{y_1 \in [0,1]} \left\{ f_1 y_1 : x_1 \geq y_1, \sum_{j=2}^{n+1} x_j \geq y_1, x_{n+2} \geq y_1, x_{n+3} \geq y_1, \dots, x_{2n} \geq y_1 \right\}. \end{aligned}$$

Sum up the constraints to get

$$\begin{aligned} &\max_{y_1 \in [0,1]} \left\{ f_1 y_1 : x_1 \geq y_1, \sum_{j=2}^{n+1} x_j \geq y_1, x_{n+2} \geq y_1, x_{n+3} \geq y_1, \dots, x_{2n} \geq y_1 \right\} \\ &\leq \max_{y_1 \in [0,1]} \left\{ f_1 y_1 : \sum_{j=1}^{2n} x_j \geq (n+1)y_1 \right\}. \end{aligned}$$

For $x \in \bar{X}$, we have $\sum_{j=1}^{2n} x_j \leq 2$. Therefore,

$$\max_{y_1 \in [0,1]} \left\{ f_1 y_1 : \sum_{j=1}^{2n} x_j \geq (n+1)y_1 \right\} \leq \max_{y_1 \in [0,1]} \{ f_1 y_1 : 2 \geq (n+1)y_1 \} \leq \frac{2}{n+1} f_1.$$

By combining the above inequalities, we get

$$\max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{agg}}(x) \leq \frac{2}{n+1} f_1.$$

Therefore, we have

$$\gamma \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{agg}}(x) \leq \frac{n+1}{2} \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{agg}}(x) \leq f_1 \leq \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{disagg}}(x).$$

- (ii) Fix $\gamma > 1$ and let $n \geq 6\gamma$ be an integer. Consider a graph $G = (N, A)$, where $N = \{1, 2, \dots, 2n+4\}$ and

$$A = \{(1, 2), (2, 3), (3, 2n+3), (2n+3, 2n+4)\} \cup \{(i, j) : 3 \leq i, j \leq 2n+2, i \neq j\}.$$

The edges 1-2 and $(2n+3)$ -($2n+4$) are of length $d-\delta$, edges 2-3 and 3-($2n+3$) are of length 2δ and all the other edges are of length $(d-\delta)/n$, where $0 < \delta < 1/n^2$ and d is the travel range. The graph is illustrated in Figure 4.

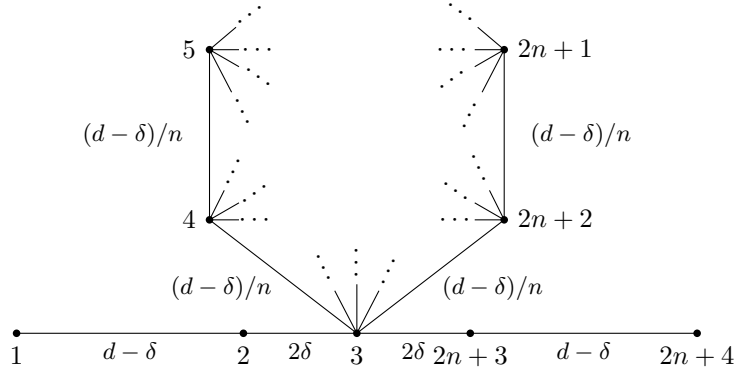


Fig. 4: Graph G , where $0 < \delta < 1/n^2$ and d is the travel range

Suppose that $Q = \{1\}$, where the origin and the destination of the demand are 1 and $2n+4$, respectively. Furthermore, assume that R_1 contains a single path $r = (1, 2, r', 2n+3, 2n+4)$, where r' is a path that starts and ends at 3 and

contains all the permutations of $\{3, 4, \dots, 2n+2\}$. For example, in case of $n = 3$, we may have

$$r = (1, 2, \quad 3, 4, 5, 6, 7, 8, \quad 3, 4, 5, 6, 8, 7, \quad \dots, \quad 8, 7, 6, 5, 4, 3, \quad 9, 10).$$

Since δ is sufficiently small, after leaving node 2 with a full battery, one can reach the next n nodes without recharging, but not the $(n+1)$ -th node. Therefore, in order to traverse the path, there must be at least one node with a charging station among the n nodes following node 2. By the same reasoning, every set of n nodes from $\{3, 4, \dots, 2n+2\}$ must contain at least one node with a charging station, and

$$C_1^{\text{agg}} = \left\{ \{1\}, \{2\}, \{2n+3\}, \{2n+4\} \right\} \cup \left\{ S : S \subseteq \{3, 4, \dots, 2n+2\}, |S| = n \right\}.$$

Note that the demand is served if and only if

$$x_1 = x_2 = x_{n+3} = x_{n+4} = 1 \text{ and } \sum_{j=3}^{2n+2} x_j \geq n+1. \quad (13)$$

Now, let us consider $\bar{X}$ given by a simple budget constraint

$$\bar{X} = \left\{ x \in \mathbb{R}^{2n} : \sum_{j=1}^{2n+4} x_j \leq 6 \right\}.$$

It is straightforward to see that $v_1^{\text{agg}}(x') = f_1$ for

$$x'_j = \begin{cases} 1, & j \in \{1, 2, 2n+3, 2n+4\}, \\ 1/n, & \text{otherwise.} \end{cases}$$

Since $\|x'\|_1 = 6$, we have $x' \in \bar{X}$, and the bound given by v_1^{agg} is loose. We will obtain a bound on $v_1^{\text{tight}}(x)$ for $x \in \bar{X}$. In the proof of Lemma 8, we showed that

$$v_1^{\text{tight}}(x) = \max_{\alpha} \left\{ \sum_{i=1}^{2^{|N|}} \alpha_i v_1^{\text{agg}}(\beta_i) : \sum_{i=1}^{2^{|N|}} \alpha_i \beta_i = x, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\},$$

where $\{0, 1\}^{|N|} = \{\beta_i : i = 1, 2, \dots, 2^{|N|}\}$. Let I_1, I_2 be a partition of $\{1, 2, \dots, 2^{|N|}\}$ such that

$$\|\beta_i\|_1 \leq n-1, \quad \forall i \in I_1, \quad \|\beta_i\|_1 \geq n, \quad \forall i \in I_2. \quad (14)$$

It follows from (13) that

$$v_1^{\text{agg}}(\beta_i) = \begin{cases} 0, & i \in I_1, \\ f_1, & i \in I_2, \end{cases}$$

and thus, for $x \in \bar{X}$,

$$v_1^{\text{tight}}(x) = \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \sum_{i=1}^{2^{|N|}} \alpha_i \beta_i = x, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\}.$$

Now, we will obtain a bound by constructing relaxations. First, observe that

$$\begin{aligned} v_1^{\text{tight}}(x) &= \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \sum_{i=1}^{2^{|N|}} \alpha_i \beta_i = x, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\} \\ &\leq \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \left\| \sum_{i=1}^{2^{|N|}} \alpha_i \beta_i \right\|_1 = \|x\|_1, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\} \\ &= \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \sum_{i=1}^{2^{|N|}} \alpha_i \|\beta_i\|_1 = \|x\|_1, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\} \\ &= \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \sum_{i \in I_1} \alpha_i \|\beta_i\|_1 + \sum_{i \in I_2} \alpha_i \|\beta_i\|_1 = \|x\|_1, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\} \\ &\leq \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \sum_{i \in I_2} \alpha_i \|\beta_i\|_1 \leq \|x\|_1, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\}. \end{aligned}$$

Since $\|\beta_i\|_1 \geq n$ for $i \in I_2$ and $\|x\|_1 \leq 6$ for $x \in \bar{X}$, we have

$$\begin{aligned} &\max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : \sum_{i \in I_2} \alpha_i \|\beta_i\|_1 \leq \|x\|_1, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\} \\ &\leq \max_{\alpha} \left\{ f_1 \sum_{i \in I_2} \alpha_i : n \sum_{i \in I_2} \alpha_i \leq 6, \sum_{i=1}^{2^{|N|}} \alpha_i = 1, \alpha \geq 0 \right\}. \end{aligned}$$

Therefore, it follows that for $x \in \bar{X}$

$$v_1^{\text{tight}}(x) \leq \frac{6}{n} f_1.$$

Thus,

$$\gamma \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{tight}}(x) \leq \frac{n}{6} \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{tight}}(x) \leq f_1 = \max_{x \in \bar{X}} \sum_{q \in Q} v_q^{\text{agg}}(x).$$

□

In Appendix B, we observe the difference between the two bounds numerically.

4 Cyclic FRLP

4.1 Formulations

In Sections 2 and 3, we studied the FRLP under the standard setting considered in the literature (e.g., Arslan et al. [2] and Göpfert and Bock [7]), where it is assumed that drivers use the same path in both directions, from the origin to the destination and back. An OD pair is thus considered served if drivers can traverse the same path repeatedly. Furthermore, this symmetric routing requires all roads to be undirected. However, in many practical situations, it may be natural to consider the possibility that drivers use different paths for their outbound and return trips. For example, consider a case where a driver deviates from the shortest path on the outbound trip to visit a charging station. In the original FRLP, it is assumed that the driver follows the same path on the return trip and visits the same charging station again. However, if the battery level is sufficiently high at the destination, the driver may choose a shorter path for the return trip, deferring recharging to the next trip. Thus, we consider the setting in which an OD path is considered served if one can repeat a cycle from the origin to the destination and back, without running out of battery. We call this extension of the FRLP, which allows such flexibility, as the *cyclic FRLP*; see Figure 5 for an illustration. In contrast, we refer to the deviation FRLP (including the classic FRLP), where paths are symmetric, as the *original FRLP*. In this section, we develop a formulation and solution method for the cyclic FRLP.

The following example highlights the difference between the original and cyclic FRLP.

Example 11. Consider the undirected graph shown in Figure 6, with a single demand q_1 from node 1 to node 2 (i.e., $Q = \{q_1\}$). Assume that the travel range is d .

Table 2: Paths from node 1 to node 2 within a 50% deviation from the shortest path

Path	Length	Deviation
(1, 2)	$(1/3)d$	0%

Table 3: Cycles from node 1 to 2 and back, within a 50% deviation from the shortest cycle

Cycle	Length	Deviation
(1, 2, 1)	$(2/3)d$	0%
(1, 2, 3, 1)	d	50%

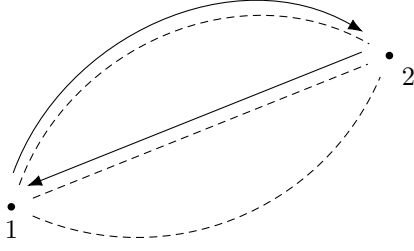


Fig. 5: OD pair with origin and destination at nodes 1 and 2, respectively; dashed lines indicate possible OD paths, and solid directed lines show a feasible route for the cyclic FRLP. In the cyclic FRLP, drivers may choose different paths for their outbound and inbound trips. An OD pair is served if there is a cycle that can be repeated without running out of battery (recall Figure 1 for comparison).

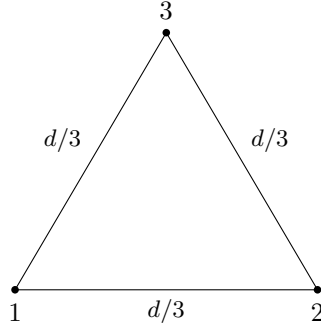


Fig. 6: Graph G , where d is the travel range

In the original FRLP, if drivers are willing to take any path that is at most 50% longer than the shortest one, there is only one possible path, as shown in Table 2. Note that path $(1, 3, 2)$ has length $2d/3$, which exceeds the 50% deviation limit.

Now consider the cyclic FRLP. If drivers are willing to take any cycle that is at most 50% longer than the shortest round-trip cycle, there are two possible cycles, as listed in Table 3 (excluding symmetric duplicates). The first cycle corresponds to the path considered in the original FRLP.

Suppose that a charging station is available only at node 3. Under the original FRLP, demand q_1 is considered unserved, as no feasible path passes through node 3. In contrast, under the cyclic FRLP, demand q_1 is considered served, since the cycle $(1, 2, 3, 1)$ is feasible and can be repeated without violating the travel range constraint.

In the original FRLP, Capar et al. [5] provided an argument that for a given path, one can construct a family of node sets $C \subseteq \mathcal{P}(N)$ such that the path is traversable repeatedly if and only if $\sum_{j \in S} x_j \geq 1$ for all $S \in C$. Thus, for each $q \in Q$ and $r \in R_q$, one can construct either C_r^{agg} or C_q^{disagg} to encode the condition for r to be

traversable, or for q to be considered served, where R_q denotes the set of allowable paths for demand q .

It is straightforward to extend the argument to the cyclic FRLP. Given a cycle, one can construct a corresponding set $C \subseteq \mathcal{P}(N)$ such that the cycle is traversable repeatedly if and only if $\sum_{j \in S} x_j \geq 1$ for all $S \in C$. In the cyclic FRLP, each demand $q \in Q$ is associated with a set of cycles that drivers of q are willing to use. We also use R_q to denote this set of cycles. Then, for each $q \in Q$ and $r \in R_q$, one can construct C_r^{disagg} , or alternatively C_q^{agg} , and use them to formulate (1) or (2).

In the case of Example 11, there are two cycles as listed in Table 3, $r_1 = (1, 2, 1)$ and $r_2 = (1, 2, 3, 1)$. Each cycle is usable repeatedly if there is at least one node in the cycle with a charging station. Thus, $C_{r_1}^{\text{disagg}} = \{\{1, 2\}\}$ and $C_{r_2}^{\text{disagg}} = \{\{1, 2, 3\}\}$. Hence, $C_{q_1}^{\text{agg}} = \{S_{r_1} \cup S_{r_2} : S_{r_1} \in C_{r_1}^{\text{disagg}}, S_{r_2} \in C_{r_2}^{\text{disagg}}\} = \{\{1, 2, 3\}\}$.

Observe that the same formulations (1) and (2) can be used for both the original and cyclic FRLP. Furthermore, the original FRLP is a special case of the cyclic FRLP in which all cycles are “symmetric” (drivers use the same path in both directions). Thus, many of the results established in Section 3 extend naturally to the cyclic FRLP. To be more specific, the bound given by the LP relaxation of formulation (2) is tighter than that of formulation (1) (Proposition 9 (i)), and the difference between them can be arbitrarily large (Proposition 10).¹ Therefore, we focus on formulation (2) in the remainder of this section.

4.2 Solution Method

Arslan et al. [2] and Göpfert and Bock [7] solve formulation (2) for the original FRLP using branch-and-cut algorithms. In this section, we demonstrate how to adapt the method proposed by Arslan et al. [2] to solve the cyclic FRLP using the same formulation. Among the variants proposed in their work, we focus on the one that yields the best performance in their experiments.

In a branch-and-cut algorithm, formulation (2) is initially considered without constraint (2b), and a standard branch-and-bound algorithm is applied. Each time the branch-and-bound algorithm finds a candidate integer solution (x', y') , the algorithm checks whether any constraint (2b) is violated. Specifically, it searches for $q \in Q$ and $S \in C_q^{\text{agg}}$ such that the corresponding constraint is not satisfied at (x', y') . If a violated constraint is found, it is added to the model, and the current solution is discarded. Otherwise, the candidate solution is feasible for the full problem.

While the scheme of the branch-and-cut algorithm outlined above is quite generic, the task of finding violated constraints for a given integer solution is typically problem-specific, as it must be done efficiently. We now describe this separation procedure. Let (x', y') be a candidate integer solution for which we aim to identify a violated constraint (2b) or its feasibility. Fix $q \in Q$ such that $y'_q = 1$ (if $y'_q = 0$, constraint (2b) is satisfied for all $S \in C_q^{\text{agg}}$). Define $S' = \{j : x'_j = 0\}$, the set of nodes where no charging station is installed in the candidate solution. We first check if there exists a cycle in R_q traversable repeatedly, assuming we only install charging stations at nodes

¹We note that assertions (ii) and (iii) in Proposition 9 require drivers to use a single, simple path between the origin and the destination, which does not generalize to a single, simple cycle in case of the cyclic FRLP.

in $N \setminus S'$. If such a cycle exists, then the constraint is satisfied for any $S \in C_q^{\text{agg}}$. Otherwise, the constraint with $S = S'$ is violated.

In light of Proposition 4, it is desirable to find a minimal node set that violates constraint (2b). To this end, we iteratively attempt to remove nodes from S' . For each $j \in S'$, we check whether there exists a traversable cycle in R_q using the modified node set $S' \setminus \{j\}$ (i.e., we search for a traversable cycle assuming we install charging stations at nodes in $N \setminus (S' \setminus \{j\})$). If no such cycle exists, we can replace S' with $S' \setminus \{j\}$. The final (reduced) S' is then added to the set of violated constraints. The pseudocode is provided in Algorithm 1.

Algorithm 1 Separation of violated constraints

Input: Set of demands Q , set of cycles R , candidate integer solution (x', y')

Ensure: Set $\mathcal{O}$ such that for each $(q, S) \in \mathcal{O}$ constraint (2b) is violated

```

1:  $\mathcal{O} \leftarrow \emptyset$ 
2: for  $q \in Q$  such that  $y_q = 1$  do
3:    $S' \leftarrow \{j : x'_j = 0\}$ 
4:   if there exists a cycle in  $R_q$  traversable using nodes in  $N \setminus S'$  then
5:     continue // Demand  $q$  is indeed served, so check the next demand
6:   end if
7:   for  $j \in S'$  do
8:     if No cycle in  $R_q$  is traversable using nodes in  $N \setminus (S' \setminus \{j\})$  then
9:        $S' \leftarrow S' \setminus \{j\}$ 
10:    end if
11:  end for
12:   $\mathcal{O} \leftarrow \mathcal{O} \cup \{(q, S')\}$ 
13: end for

```

The branch-and-cut approach and the separation procedure described above follow those of Arslan et al. [2]. However, in the deviation FRLP, identifying a violated constraint corresponds to finding a traversable path, whereas in the cyclic FRLP, it involves finding a traversable cycle. The symmetric path assumption in the deviation FRLP (i.e., that drivers must use the same path for both outbound and inbound trips) considerably simplifies the search for a traversable path. It can be shown that a path is repeatedly traversable if and only if a driver can leave the origin with a half-charged battery and reach the destination with at least 50% of the battery remaining. This condition reduces the search for a feasible path to a shortest path problem on a transformed network (see Arslan et al. [3] for details). In contrast, it is more involved to find a traversable cycle from the origin to the destination and back. For instance, it is not known in advance how much energy can be assumed at the origin or how much must remain upon arrival at the destination. Hence, we propose a new procedure to find a traversable cycle.

In the cyclic FRLP, the existence of a traversable cycle in R_q can be tested using a labeling algorithm (see Irnich and Desaulniers [9]). For example, suppose R_q includes all cycles of length at most τ that start at the origin, visit the destination, and return

to the origin. The graph is first modified by adding a sink node s connected from the origin via a zero-length edge. Each label has the form $(\delta_{\text{charge}}, \delta_{\text{dest}}, l_{\text{start}}, l_{\text{charge}}, \gamma_{\text{end}})$, where

- δ_{charge} is 1 if and only if we have visited at least one charging station and 0 otherwise,
- δ_{dest} is 1 if and only if we have visited the destination and 0 otherwise,
- l_{start} is the total distance traveled from the origin,
- l_{charge} is the distance since the last charging station (equal to l_{start} if $\delta_{\text{charge}} = 0$),
- γ_{end} is the amount of charge (in distance units) needed at the origin before completing the cycle (equal to ∞ if $\delta_{\text{charge}} = 0$).

We initialize the origin with a label $(0, 0, 0, 0, \infty)$ if there are no charging stations at the origin or otherwise $(1, 0, 0, 0, 0)$. A label $(\delta_{\text{charge}}, \delta_{\text{dest}}, l_{\text{start}}, l_{\text{charge}}, \gamma_{\text{end}})$ at node j_1 can be extended to node $j_2 \neq s$ if and only if $l_{\text{start}} + l_{j_1 j_2} \leq \tau$ and $l_{\text{charge}} + l_{j_1 j_2} \leq d$ where $l_{j_1 j_2}$ is the length of the edge from node j_1 to node j_2 and d is the vehicle's travel range. Similarly, a label $(\delta_{\text{charge}}, \delta_{\text{dest}}, l_{\text{start}}, l_{\text{charge}}, \gamma_{\text{end}})$ at the origin can be extended to the sink s if and only if $l_{\text{charge}} + \gamma_{\text{end}} \leq d$. The label at node j_2 resulting from an extension is given by the resource extension function f . It is defined by $f : (\delta_{\text{charge}}, \delta_{\text{dest}}, l_{\text{start}}, l_{\text{charge}}, \gamma_{\text{end}}) \mapsto (\delta'_{\text{charge}}, \delta'_{\text{dest}}, l'_{\text{start}}, l'_{\text{charge}}, \gamma'_{\text{end}})$, where

$$\begin{aligned} \delta'_{\text{charge}} &= \begin{cases} 0, & \delta_{\text{charge}} = 0 \text{ and } x_{j_2} = 0 \text{ (i.e., there is no charging stations at } j_2), \\ 1, & \text{otherwise,} \end{cases} \\ \delta'_{\text{dest}} &= \begin{cases} 0, & \delta_{\text{dest}} = 0 \text{ and } j_2 \text{ is not the destination,} \\ 1, & \text{otherwise,} \end{cases} \\ l'_{\text{start}} &= l_{\text{start}} + l_{j_1 j_2}, \\ l'_{\text{charge}} &= \begin{cases} l_{\text{charge}} + l_{j_1 j_2}, & x_{j_2} = 0, \\ 0, & \text{otherwise,} \end{cases} \\ \gamma'_{\text{end}} &= \begin{cases} l_{\text{start}} + l_{j_1 j_2}, & \delta_{\text{charge}} = 0 \text{ and } x_{j_2} = 1, \\ \gamma_{\text{end}}, & \text{otherwise.} \end{cases} \end{aligned}$$

To determine the order in which labels are extended, we use a scoring rule. Let $\text{dist}(j_1, j_2)$ denote the shortest distance from node j_1 to node j_2 , and let o and t represent the origin and destination, respectively. Given a label $(\delta_{\text{charge}}, \delta_{\text{dest}}, l_{\text{start}}, l_{\text{charge}}, \gamma_{\text{end}})$ attached to a node j , its score is computed as $\text{dist}(j, t) + \text{dist}(t, o)$ if $\delta_{\text{dest}} = 0$, and as $\text{dist}(j, o)$ otherwise. In other words, the score represents the length of the shortest cycle that could be formed by extending the label, ignoring travel range or distance limits. The label with the smallest score is selected for extension.

A traversable cycle in R_q exists if and only if the sink node s receives a label during the labeling process. When s is labeled, the condition $l_{\text{charge}} + \gamma_{\text{end}} \leq d$ ensures that the sum of the distance from the first charger after departure and the distance from the last charger before returning is at most the vehicle's travel range, thereby guaranteeing that the cycle can be repeated. Conversely, if a traversable cycle exists,

there must be a corresponding sequence of labels that will be identified during the execution of the labeling algorithm.

Figure 7 illustrates an execution of the labeling algorithm on a simple 3-node transportation network. In this example, the sink s is labeled after five steps, confirming the existence of a cycle that satisfies both the travel range and total length constraints.

4.3 Numerical Experiments

In this section, numerical experiments are presented to demonstrate the effect of the flexibility introduced by the cyclic FRLP. In Section 4.3.1, we run experiments to see how much demand the solution of the original and cyclic FRLP can cover. In Section 4.3.2, we use the original and cyclic FRLP to determine the minimum number of charging stations required to serve all demands.

4.3.1 FRLP for Maximizing Served Demand

In this section, we compare the original FRLP and the cyclic FRLP for the problem of placing charging stations so as to serve as much demand as possible. In both models, we assume that drivers are willing to deviate from their shortest cycle by at most a factor of α , which is varied across the experiments.

First setup (original FRLP with reevaluation). We first solve the original FRLP, which assumes that drivers must use the same path for their outbound and return trips, with a budget constraint imposing that at most five charging stations may be located

$$X = \left\{ x \in \{0, 1\}^{|N|} : \sum_{j \in N} x_j \leq 5 \right\}.$$

This yields a location decision x . Keeping this charging station placement fixed, we then reevaluate the amount of served demand under the more general cyclic routing assumption, in which drivers are willing to use different paths for their outbound and inbound trips. The reevaluation uses the same deviation tolerance parameter α . Consequently, any demand served in the original solution remains served, while additional demand that was unserved in the original solution may become served in the reevaluation. Hence, the reevaluated served demand is greater than or equal to the objective value obtained from the original FRLP.

Second setup (cyclic FRLP). In the second setup, we directly solve the cyclic FRLP, which assumes drivers may use different paths in the two directions, and determine the placement of at most five charging stations that maximizes the served demand.

The original FRLP is solved with the method of Arslan et al. [2] with formulation (2), while the cyclic FRLP is solved with the method described in Section 4.2. We use two transportation networks introduced by Simchi-Levi and Berman [16] (see also Kuby and Lim [14]) and by Arslan et al. [3] (which is also used by Arslan et al. [2] and Göpfert and Bock [7]). See Table 4 for the statistics. All methods are implemented

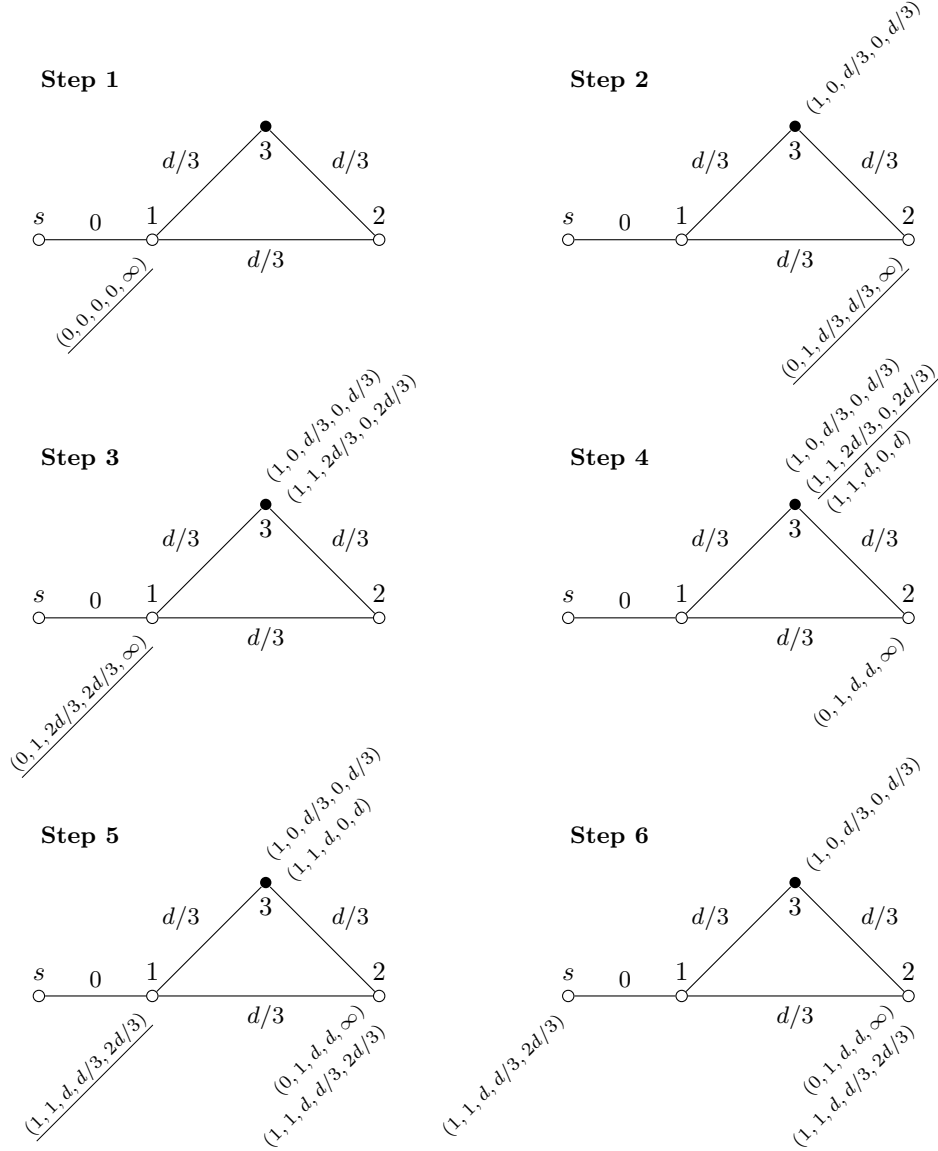


Fig. 7: An illustrative example of a 3-node transportation network and a corresponding execution sequence of the labeling algorithm. Node 1 is the origin, and node 2 is the destination. Suppose both the travel range and the maximum travel distance are d (corresponding to a 50% deviation tolerance). Open nodes indicate locations without charging stations, whereas closed nodes indicate locations with a charging station. Labels are shown at their respective nodes, and underlined labels denote those selected for extension in subsequent iterations.

in C++ with CPLEX 22.1 (executed on a single thread) and executed on a machine with 2 “Intel(R) Xeon(R) Gold 6226 CPU” (12 cores each) and 384GB of RAM.²

Table 4: Characteristics of the transportation networks used in the experiments.

Name	Reference	Number of nodes	Number of edges	Number of OD pairs
NETWORK25	[16]	25	42	300
CALIFORNIA	[3]	339	1234	1167

Table 5 shows the results. For the original FRLP, it reports the objective value based on the formulation (2), and the actual demand served by the solution output from the original FRLP, reevaluated using the assumption of the cyclic FRLP. For the cyclic FRLP, the objective value equals the amount of served demand, since it correctly calculates the objective value under our assumption.

On the smaller network NETWORK25, the difference in served demand between the two methods is small, except when the deviation tolerance is $\alpha = 1.5$. We observe larger discrepancies on the larger network CALIFORNIA. However, the time required by the cyclic FRLP tends to be larger. For both methods, the solution time tends to increase as the deviation tolerance α gets larger. We note that the increase is more rapid for the cyclic FRLP compared to the original FRLP. In fact, the majority of the computational time is spent on the separation procedure, indicating that the labeling algorithm takes longer to execute as the deviation tolerance increases. When solving the original FRLP, most of the computational time is also devoted to separation. Since, as discussed in Section 4.1, in the original FRLP the separation reduces to a shortest path problem, whereas for the cyclic FRLP the same procedure is not applicable, the higher solving times for the cyclic FRLP are to be expected. More detailed results are provided in Appendix C.

4.3.2 FRLP for Serving All Demands

Although we have focused on the FRLP for maximizing the amount of served demand, formulation (2) can be modified to solve the FRLP with other objectives. For example, as discussed by Arslan et al. [2], the FRLP that finds the minimum number of charging stations to serve all demands can be formulated as

$$\begin{aligned}
& \min_{x \in \{0,1\}^{|N|}} \sum_{j \in N} x_j \\
& \text{s.t.} \quad \sum_{j \in S} x_j \geq 1, \quad \forall q \in Q, S \in C_q^{\text{agg}}.
\end{aligned} \tag{15}$$

The adaptation of the solution method to solve this variant is straightforward, and hence omitted.

²The computer program is available at <https://github.com/nsugishita/frlp>.

Table 5: Comparison of original and cyclic FRLP models for maximizing served demand. Columns indicate deviation tolerance (α), original FRLP objective value (Objective), served demand assuming drivers may use different paths outbound and inbound (Served demand), and computation times in seconds (Time).

	α	Original FRLP			Cyclic FRLP	
		Objective	Served demand	Time (s)	Served demand	Time (s)
NETWORK25	1.0	56.16	56.16	0.10	56.16	0.04
	1.1	58.38	58.38	0.03	58.38	0.05
	1.2	59.20	59.20	0.03	59.20	0.09
	1.3	60.12	60.77	0.04	60.77	0.23
	1.4	61.70	63.14	0.05	63.17	0.27
	1.5	63.17	65.41	0.06	68.21	0.35
CALIFORNIA	1.0	79.94	79.94	3.32	79.94	0.82
	1.1	85.91	86.71	19.66	87.89	160.43
	1.2	89.08	91.80	24.31	92.42	356.69
	1.3	91.61	94.37	11.21	95.31	467.08
	1.4	92.95	95.85	29.58	97.04	695.10
	1.5	94.51	96.70	29.18	97.96	865.44

In the next experiment, we run the original FRLP and the cyclic FRLP to find the minimum number of charging stations required to serve all demands by solving (15). Table 6 shows the results. Similar to the FRLP for maximizing served demands, it takes more time to solve the cyclic FRLP than the original FRLP. However, the cyclic FRLP tends to lead to solutions using fewer charging stations than the original FRLP, especially when the deviation tolerance α is large. For both the original FRLP and cyclic FRLP, the solution time grows as the deviation tolerance α increases. Therefore, it is important to use the correct variant of the FRLP depending on driver behavior. Additional results are again provided in Appendix C.

Table 6: Comparison of original and cyclic FRLP models for serving all demands with the minimum number of charging stations.

	α	Original FRLP		Cyclic FRLP	
		Num. of charging stations	Time (s)	Num. of charging stations	Time (s)
NETWORK25	1.0	17	0.02	17	0.02
	1.1	17	0.02	17	0.04
	1.2	17	0.02	17	0.08
	1.3	17	0.02	16	0.12
	1.4	17	0.03	16	0.16
	1.5	16	0.03	14	0.19
CALIFORNIA	1.0	24	6.14	24	1.40
	1.1	22	10.66	19	43.12
	1.2	19	14.57	16	106.23
	1.3	18	17.55	12	160.47
	1.4	16	18.86	9	212.63
	1.5	14	19.79	9	279.44

5 Conclusions

We studied the FRLP, an optimization problem for determining the optimal placement of refueling stations for vehicles with limited travel ranges, such as EV charging stations. Specifically, we examined two formulations from the literature, one developed independently by Arslan et al. [2] and Göpfert and Bock [7], and another one proposed by Yıldız et al. [17]. The former formulation is used in the state-of-the-art algorithm, which is based on a branch-and-cut approach. We showed that this formulation is tighter than the latter and that the relative difference in bounds obtained from their LP relaxation can be arbitrarily large. This tightness may partially explain the superior performance of the current state-of-the-art solution method.

We also addressed a common routing assumption used in the literature of the FRLP, which requires drivers to use the same path for their outbound and inbound trips. To relax this, we introduced a variant called the cyclic FRLP, where drivers may take different paths on their outbound and return trips. We demonstrated how existing formulations for the original FRLP can be naturally extended to model this variant and presented its solution method. Furthermore, we presented a labeling algorithm for finding a traversable cycle, enabling the use of a branch-and-cut approach to solve the cyclic FRLP. Through a small example and numerical experiments on the Californian network from Arslan et al. [3], we highlighted the significant impact of correctly modeling driver behavior on solution quality.

The cyclic FRLP is more flexible than the original FRLP, which requires drivers to use the same path for both outbound and return trips. It relaxes this assumption by allowing different paths in each direction, but still assumes that the same cycle is used for every round trip. A natural extension is to allow drivers to repeat an n -sequence of cycles with some predetermined number n , which could model medium-distance travel where refueling or recharging is not needed after every trip. The solution method, including the separation procedure based on the labeling algorithm, can be readily adapted to handle this extension. However, evaluating its practical relevance is left for future work.

A Integrality of Auxiliary Variables

Here, we show that dropping the integrality constraints in (1) does not change the optimal objective value. Consider

$$\begin{aligned}
 \max_{x \in X, z \in [0,1]^{|Q| \times |R|}} \quad & \sum_{q \in Q} \sum_{r \in R_q} f_q z_{qr} \\
 \text{s.t.} \quad & \sum_{r \in R_q} z_{qr} \leq 1, \quad \forall q \in Q, \\
 & \sum_{j \in S} x_j \geq z_{qr}, \quad \forall q \in Q, r \in R_q, S \in C_r^{\text{disagg}},
 \end{aligned} \tag{16}$$

which is obtained by relaxing the integrality constraints on z in (1). It suffices to show that there exists an optimal solution to (16) that is also feasible and optimal for (1).

Let (x, z) be an optimal solution to (16). For each $q \in Q$, let $<_q$ be a well ordering on R_q . We construct a solution (x, z') as follows: for each $q \in Q$ and $r \in R_q$,

$$z'_{qr} = 1 \text{ if and only if } z_{qr} > 0 \text{ and } z_{qr'} = 0 \text{ for all } r' \in R_q \text{ with } r' <_q r,$$

and $z'_{qr} = 0$ otherwise. It is straightforward to observe that (x, z') is feasible for (7) and (16), and has the same objective value as (x, z) . Thus, (1) and (16) have the same optimal objective value.

B Difference between Two Relaxations

In this section, we numerically examine the difference in the strength of the two formulations, (1) and (2). As discussed in Section 3, particularly in Proposition 10, the gap between the two formulations can be arbitrarily large in terms of the tightness of the bounds provided by their LP relaxations. In this section, we solve the LP relaxations for both formulations and numerically observe the differences in the resulting bounds. To perform this analysis, we use the NETWORK25 instance, as described in Section 4.3. For each deviation tolerance α and OD pair $q \in Q$, we enumerate all paths R_q whose lengths are below the deviation tolerance. For each $q \in Q$ and $r \in R_q$, we then construct the set C_q^{disagg} .

Formulation (1). The LP relaxation of this formulation is directly constructed and solved.

Formulation (2). Since formulation (2) involves a large number of constraints, they are generated lazily. The LP relaxation of this formulation is first solved without the constraints, and any violated constraints are identified. If violations are found, the constraints are added, and the problem is resolved. Note that the search for violated constraints itself is an integer programming problem.

Table 7 summarizes the bounds obtained from the LP relaxations, the gap between these bounds and the optimal objective values (in percentage), and the solution times (in seconds), which include the time spent searching for violated constraints in formulation (2). Interestingly, the gap of the bound obtained from the LP relaxation of formulation (2) remains below 2% for all values of α between 1 and 1.5. However, the bound from the LP relaxation of formulation (1) tends to loosen as α increases. Additionally, the LP relaxation of formulation (2) generally takes longer to solve than that of formulation (1). Both formulations exhibit increasing solution times as α increases.

It is important to note that the decomposition method used for formulation (2) in Section 4.3 does not solve the LP relaxation, as constraints are generated lazily. Consequently, the solution time for the LP relaxation of formulation (2) reported in Table 7 may exceed the solution time of the decomposition method used to solve formulation (2). Nonetheless, the tightness of the formulation together with the efficient decomposition method makes formulation (2) a suitable choice.

Table 7: Comparison of the LP relaxations of formulations (1) and (2). Columns indicate deviation tolerance (α), the bounds obtained from the LP relaxations (Bound), the gap between the bounds and the optimal objective value (Gap), and computation times in seconds (Time).

α	Formulation (1)			Formulation (2)		
	Bound	Gap (%)	Time (s)	Bound	Gap (%)	Time (s)
1.0	58.12	3.50	0.02	57.03	1.55	0.98
1.1	62.07	6.32	0.03	58.60	0.38	0.71
1.2	68.69	16.02	0.07	59.30	0.17	2.04
1.3	73.16	21.69	0.24	60.47	0.58	2.99
1.4	79.32	28.54	0.81	62.41	1.15	9.84
1.5	82.29	30.27	2.82	63.47	0.47	15.74

C Detailed Results

This section contains detailed results. Table 8 corresponds to the experiment in Section 4.3.1 (the FRLP for maximizing the served demand). Table 9 is the result of the experiment in Section 4.3.2 (the FRLP for covering all demands).

Data Availability

Data available on request from the authors.

Acknowledgements

This research was supported by Hydro-Québec, NSERC Collaborative Research and Development Grant CRDPJ 536757 - 19.

The authors are grateful to Professor Okan Arslan, who kindly shared the Californian network used in our numerical experiments, and to the two anonymous reviewers, whose comments greatly improved the quality of the paper.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- [1] Arslan, O., Karaşan, O.E.: A Benders decomposition approach for the charging station location problem with plug-in hybrid electric vehicles. *Transportation Research. Part B: Methodological* **93**, 670–695 (2016) <https://doi.org/10.1016/j.trb.2016.09.001>
- [2] Arslan, O., Karaşan, O.E., Mahjoub, A.R., Yaman, H.: A branch-and-cut algorithm for the alternative fuel refueling station location problem with routing.

Table 8: Comparison of original and cyclic FRLP models for maximizing served demand. Columns indicate deviation tolerance (α), total computation times in seconds (Time), computation times spent in separation (Separation time), the number of branch-and-bound nodes (# B&B nodes), and the number of violated constraints added (# cuts).

Instance	Routing	α	Time (s)	Separation time (s)	Num. of B&B nodes	Num. of cuts
NETWORK25	original	1.0	0.10	0.01	7	721
		1.1	0.03	0.01	3	586
		1.2	0.03	0.02	8	636
		1.3	0.04	0.02	7	673
		1.4	0.05	0.03	5	727
		1.5	0.06	0.04	3	834
	cyclic	1.0	0.04	0.01	15	761
		1.1	0.05	0.03	11	666
		1.2	0.09	0.07	5	722
		1.3	0.23	0.21	17	955
		1.4	0.27	0.24	10	848
		1.5	0.35	0.33	12	858
CALIFORNIA	original	1.0	3.32	3.28	0	1448
		1.1	19.66	19.59	7	2178
		1.2	24.31	24.25	3	2116
		1.3	11.21	11.18	0	1390
		1.4	29.58	29.51	5	1906
		1.5	29.18	29.13	2	2026
	cyclic	1.0	0.82	0.77	0	2134
		1.1	160.43	160.26	32	3185
		1.2	356.69	356.44	45	2710
		1.3	467.08	466.91	39	3025
		1.4	695.10	694.90	130	2918
		1.5	865.44	865.33	0	2592

Transportation Science **53**(4), 1107–1125 (2019) <https://doi.org/10.1287/trsc.2018.0869>

- [3] Arslan, O., Yıldız, B., Karaşan, O.E.: Impacts of battery characteristics, driver preferences and road network features on travel costs of a plug-in hybrid electric vehicle (PHEV) for long-distance trips. Energy Policy **74**, 168–178 (2014) <https://doi.org/10.1016/j.enpol.2014.08.015>
- [4] Capar, I., Kuby, M.: An efficient formulation of the flow refueling location model for alternative-fuel stations. IIE Transactions **44**(8), 622–636 (2012) <https://doi.org/10.1080/0740817X.2011.635175>
- [5] Capar, I., Kuby, M., Leon, V.J., Tsai, Y.: An arc cover–path-cover formulation and strategic analysis of alternative-fuel station locations. European Journal of Operational Research **227**(1), 142–151 (2013) <https://doi.org/10.1016/j.ejor.2012.11.033>

Table 9: Comparison of original and cyclic FRLP models for serving all demands with the minimum number of charging stations.

Instance	Routing	α	Time (s)	Callback time (s)	Num. of B&B nodes	Num. of cuts
NETWORK25	original	1.0	0.02	0.01	3	945
		1.1	0.02	0.01	3	934
		1.2	0.02	0.02	2	929
		1.3	0.02	0.02	2	926
		1.4	0.03	0.02	4	928
		1.5	0.03	0.02	3	924
	cyclic	1.0	0.02	0.02	0	901
		1.1	0.04	0.04	0	902
		1.2	0.08	0.08	3	901
		1.3	0.12	0.12	3	907
		1.4	0.16	0.16	3	906
		1.5	0.19	0.18	2	900
CALIFORNIA	original	1.0	6.14	6.12	24	3712
		1.1	10.66	10.64	14	3535
		1.2	14.57	14.56	16	3501
		1.3	17.55	17.54	8	3554
		1.4	18.86	18.84	13	3513
		1.5	19.79	19.77	1	3517
	cyclic	1.0	1.40	1.38	73	3903
		1.1	43.12	43.10	7	3509
		1.2	106.23	106.20	28	3517
		1.3	160.47	160.46	0	3554
		1.4	212.63	212.62	0	3534
		1.5	279.44	279.41	6	3576

- [6] Fulkerson, D., Gross, O.: Incidence matrices and interval graphs. *Pacific Journal of Mathematics* **15**(3), 835–855 (1965) <https://doi.org/10.2140/pjm.1965.15.835>
- [7] Göpfert, P., Bock, S.: A branch&cut approach to recharging and refueling infrastructure planning. *European Journal of Operational Research* **279**(3), 808–823 (2019) <https://doi.org/10.1016/j.ejor.2019.06.031>
- [8] Hodgson, M.J.: A flow-capturing location-allocation model. *Geographical Analysis* **22**(3), 270–279 (1990) <https://doi.org/10.1111/j.1538-4632.1990.tb00210.x>
- [9] Irnich, S., Desaulniers, G.: Shortest path problems with resource constraints. In: *Column Generation*, pp. 33–65. Springer, New York (2005). https://doi.org/10.1007/0-387-25486-2_2
- [10] Kınay, Ö.B., Gzara, F., Alumur, S.A.: Full cover charging station location problem with routing. *Transportation Research Part B: Methodological* **144**, 1–22 (2021) <https://doi.org/10.1016/j.trb.2020.12.001>

- [11] Kınay, Ö.B., Gzara, F., Alumur, S.A.: Charging station location and sizing for electric vehicles under congestion. *Transportation Science* **57**(6), 1433–1451 (2023) <https://doi.org/10.1287/trsc.2021.0494>
- [12] Kim, J., Kuby, M.: The deviation-flow refueling location model for optimizing a network of refueling stations. *International Journal of Hydrogen Energy* **37**(6), 5406–5420 (2012) <https://doi.org/10.1016/j.ijhydene.2011.08.108>
- [13] Kuby, M., Lim, S.: The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences* **39**(2), 125–145 (2005) <https://doi.org/10.1016/j.seps.2004.03.001>
- [14] Kuby, M., Lim, S.: Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. *Networks and Spatial Economics* **7**(2), 129–152 (2007) <https://doi.org/10.1007/s11067-006-9003-6>
- [15] Kadri, A.A., Perrouault, R., Boujelben, M.K., Gicquel, C.: A multi-stage stochastic integer programming approach for locating electric vehicle charging stations. *Computers & Operations Research* **117**, 104888 (2020) <https://doi.org/10.1016/j.cor.2020.104888>
- [16] Simchi-Levi, D., Berman, O.: A heuristic algorithm for the traveling salesman location problem on networks. *Operations research* **36**(3), 478–484 (1988) <https://doi.org/10.1287/opre.36.3.478>
- [17] Yıldız, B., Arslan, O., Karaşan, O.E.: A branch and price approach for routing and refueling station location model. *European Journal of Operational Research* **248**(3), 815–826 (2016) <https://doi.org/10.1016/j.ejor.2015.05.021>