Citation for published version:
Arslan, O, Jabali, O & Laporte, G 2018, 'Exact solution of the evasive flow capturing problem', *Operations Research*, vol. 66, no. 6, pp. 1625-1640. https://doi.org/10.1287/opre.2018.1756

DOI:
10.1287/opre.2018.1756

Publication date:
2018

Document Version
Peer reviewed version

Link to publication

University of Bath

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 08. Jul. 2024

Submitted to *Operations Research*
manuscript OPRE-2017-08-434

Exact Solution of the Evasive Flow Capturing Problem

Okan Arslan
CIRRELT and HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7, Canada, okan.arslan@hec.ca

Ola Jabali
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy, ola.jabali@polimi.it

Gilbert Laporte
CIRRELT and HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7, Canada, gilbert.laporte@cirrelt.ca

The Evasive Flow Capturing Problem (EFCP) is defined as the problem of locating a set of law enforcement facilities on the arcs of a road network to intercept unlawful vehicle flows traveling between origin-destination pairs, who in turn deviate from their route to avoid any encounter with such facilities. Such deviations are bounded by a given tolerance. We first propose a bilevel program which, in contrast to previous studies, does not require a priori route generation. We then transform this bilevel model into a single-stage equivalent model using duality theory to yield a compact formulation. We finally reformulate the problem by describing the extreme rays of the polyhedral cone of the compact formulation and by projecting out the auxiliary variables, which leads to facet-defining inequalities and a cut formulation with an exponential number of constraints. We develop a branch-and-cut algorithm for the resulting model, as well as two separation algorithms to solve the cut formulation. Through extensive experiments on real and randomly generated networks, we demonstrate that our best model and algorithm accelerate the solution process by at least two orders of magnitude compared with the best published algorithm. Furthermore, our best model significantly increases the size of the instances that can be solved optimally.

Key words: location, routing, bilevel programming, projection, branch-and-cut, evasive flow capturing
Subject classifications: Transportation:Location; Programming:Integer:Cutting plane/facet; Networks/graphs:Applications.
Area of review: Transportation
History: Received: Aug 14, 2017; Revised: Nov 21, 2018, Feb 20, 2018; Accepted: Mar 2, 2018

1. Introduction

The purpose of this paper is to present an exact algorithm for the Evasive Flow Capturing Problem (EFCP), defined as follows. In a transportation network, the regulatory authority locates law enforcement facilities to intercept unlawful vehicles, while these vehicles may deviate from their routes to avoid encountering these facilities. Taking this non-cooperative behavior into account, the objective of the EFCP is to determine the location of such law enforcement facilities. The problem has applications in transportation, revenue management and security management.

The EFCP belongs to a wider class of flow capturing problems (FCPs), which are rooted in the work of Hodgson (1981). The main difference between FCPs and classical facility location problem lies in the definition of demands. In classical facility location problems, demand is usually concentrated at nodes of a network. Applications in which the demand is non-stationary and the movement of vehicles is to be modeled require a different demand definition. To this end, Hodgson (1990) and Berman, Larson, and Fouska (1992) defined the demand as vehicle flows between origin-destination (OD) pairs. This definition is particularly suitable for transportation applications in which the vehicle trips or path choices need to be taken into account. There exist several motives for intercepting vehicle flows: to provide the drivers with an opportunity to visit discretionary service facilities such as day-care centers, automated teller machines, convenience stores or gasoline stations (Berman, Larson, and Fouska 1992), to intercept law breakers such as overweight trucks or drunk drivers (Hodgson, Rosing, and Zhang 1996, Gendreau, Laporte, and Parent 2000), or to prevent transportation of unlawful or dangerous materials such as explosives or hazardous materials (Mirchandani, Rebello, and Agnetis 1995). The objective of FCPs is to locate facilities such that the intercepted vehicle flow is maximized or that the reduction in risk is maximized, the former objective being a special case of the latter (Gendreau, Laporte, and Parent 2000).

There exist three main categories of flow capturing problems characterized by different driver behaviors. In the first category, the drivers are neutral to facility locations and use predetermined fixed paths. They are assumed to receive service if there exists a facility on their fixed path and the demand is covered. There are no routing decisions in this first category. Examples include locating billboards (Averbakh and Berman 1996) or traffic counting stations (Yang and Zhou 1998). In the second category, the fixed path assumption is relaxed and the drivers are willing to multiply the length of their path by a *deviation tolerance* $\lambda \geq 1.0$ to visit a station. Note that the problems of the first category can be interpreted as having a deviation tolerance $\lambda = 1.0$. An example arises in the context of alternative fuel vehicles (AFV), where the facilities to be located are alternative refueling stations (AFS). The objective is to capture as many AFV travelers as possible in order to encourage AFV usage. To this end, Kuby and Lim (2005) introduced the flow refueling location problem (FRLP) as an extension of FCPs in which the coverage of a vehicle flow possibly requires multiple stops at an AFS on its path. The deviation flow capturing problem introduced by Kim and Kuby (2012) is an extension of FRLP in which the vehicles change their routes to get serviced at refueling facilities. Under this behavior, the problem is effectively transformed into a location-routing problem. A heuristic (Kim and Kuby 2013), a branch-and-price algorithm (Yıldız, Arslan, and Karaşan 2016) and a branch-and-cut algorithm (Arslan et al. 2017) have been developed for this problem. In the third category, which is the topic of this paper, the drivers non-cooperatively modify their route within a deviation tolerance λ. They exhibit an evasive behavior and try to avoid stations in order not to be intercepted. Examples include location of tollbooths or security checkpoints. In transportation settings, one particular application, which motivated our work, is the location of weigh-in-motion (WIM) systems that are used to enforce weight limits. In transportation networks, the regulatory authority imposes strict truck weight rules in order to prevent excessive damage on roads. To enforce these rules, WIM facilities are located to intercept overweight trucks. The location optimization of such facilities has typically been studied by assuming that the truck drivers travel on fixed paths (Hodgson, Rosing, and Zhang 1996). However, there is empirical evidence that the drivers of overweight trucks learn the location of WIM facilities and change their path between their origin and destination accordingly (Cottrell 1992, Cunagin, Mickler, and Wright 1997). Therefore, the fixed path assumption was relaxed by Marković, Ryzhov, and Schonfeld (2015), who presented the Evasive Flow Capturing Problem. The problem is defined as locating a set of law enforcement facilities on a road network in order to capture the vehicle flows trying to avoid the facilities in such a way that the total cost of facility location and the cost of damage on the roads due to overweight transportation by non-intercepted vehicles is minimized. Note that the road damage component does not appear in all EFCP applications but is specific to the WIM station location problem, which is therefore more general.

To capture a flow between an OD pair, all paths within the deviation tolerance of the drivers must be intercepted. The underlying logic is that if the stations are located so that all paths within the drivers' deviation tolerance are covered, then the law-breaking truck drivers are deterred from overweight transportation and they abide by the law because overloading trucks and taking excessive detours would represent a costlier alternative.

There also exists a stream of literature on fare evasion of travelers in transit networks. Borndörfer et al. (2012) consider a problem, in which the regulatory authority determines probabilities for inspecting passengers at different locations. The fare-evading travelers prefer either to travel without buying a ticket, or to buy a ticket to avoid being intercepted and paying excessive fines. The travelers in their application are non-adaptive and follow fixed routes. Correa et al. (2017) relax the fixed path assumption and consider a variant with adaptive followers while accounting for the traveled distance as a cost component for the travelers. The authors further consider a fixed-fee policy in which the fees are exogenously determined, and a flexible-fee policy in which setting the fees as well as the inspection probabilities is an instrument for the regulatory authority to maximize revenues. The non-intercepted evaders do not contribute to the objective function, and the number of agents are assumed to be limited by a given budget. In the EFCP, on the other hand, the regulatory authority minimizes the cost of damage caused by non-intercepted evaders. There is also no fine associated with interception, the global objective being to deter the travelers from illegal weights transportation. Furthermore, there is a cost associated with opening a station in the EFCP.

In order to solve the EFCP, Marković, Ryzhov, and Schonfeld (2015) pregenerated routes up to the drivers' deviation tolerance, and presented a binary integer program as the basis for an exact solution methodology. This model is further elaborated in the following section. The technical details of WIM systems can be found in the same paper. More recently, Marković, Ryzhov, and Schonfeld (2017) considered a multi-period EFCP, and Lu et al. (2017) developed a heuristic for this problem.

Our aim is to propose new models for the EFCP, study their theoretical properties, and ultimately devise efficient solution algorithms. Our scientific contributions are as follows:

- We present a novel bilevel model that does not require the pregeneration of routes and that is applicable in a more generic cost setting.
- We linearize the model to obtain an equivalent single-state formulation by using a constraint transformation and the duality theory. This compact formulation accelerates the solution process and extends the size of the instances that can be solved to optimality.
- Furthermore, we describe the extreme rays of the polyhedral cone of the compact formulation and obtain a formulation with an exponential number of constraints by projecting out the auxiliary variables from the compact formulation.
- We further investigate the obtained inequality to arrive at another facet-defining cut inequality.
- We devise a branch-and-cut algorithm based on the cut inequality. The separation problem is a shortest path problem in the case of integer solutions and a resource constrained shortest path problem in the case of fractional solutions.
- Through extensive computational experiments, we demonstrate the effectiveness of our methodology and its marked superiority over an existing algorithm for the same problem.

In the following section, we formally present the problem. The methodology and computational study are presented in Sections 3 and 4, respectively. We conclude in Section 5.

2. Formal Problem Definition

Consider a network $G = (N, A)$ where N and A are the sets of nodes and arcs, respectively. A vehicle flow f is defined as quadruple $\langle s_f, t_f, q_f, \lambda_f \rangle$, where s_f and t_f are the origin and destination nodes, respectively, q_f is the vehicle flow, and λ_f is the value of deviation tolerance. Let F be the set of all vehicle flows in the network. The drivers minimize their travel distance and attempt to avoid any encounter with stations on their paths. If the stations are located such that all paths between an OD pair within the deviation tolerance are intercepted, then this flow is assumed to be *covered*. The objective of the EFCP is to determine the location of stations in such a way that the total cost of opening stations and the cost of damage due to non-intercepted vehicle flows is minimized.

For completeness, we start our discussion by presenting the model by Marković, Ryzhov, and Schonfeld (2015), which we refer to as the MRS model. This model requires the enumeration of all paths within the deviation tolerance value of λ_f and requires successively solving k shortest path problems, where the $(k+1)^{st}$ shortest path is the first to exceed the tolerance. Let P_f be the set of all paths from s_f to t_f within the deviation tolerance λ_f. We refer to the set of arcs in path p as A^p. The parameter c_f^p represents the damage produced by vehicle flow $f \in F$ on path $p \in P_f$. The damage changes linearly with the distance traveled. Since the paths are assumed to be known before solving the model, c_f^p can be calculated as $c_f^p = \sum_{(i,j)\in A^p} d_{ij} c_f$, where d_{ij} and c_f are the length of arc $(i,j) \in A$ and the damage coefficient of vehicles of flow $f \in F$, respectively. Lastly, w_{ij} is the cost of opening a station on arc $(i,j) \in A$. We need the following decision variables:

$$x_{ij} = \begin{cases} 1 & \text{if a station is located on arc } (i,j) \in A \\ 0 & \text{otherwise} \end{cases}$$

$$y_f^p = \begin{cases} 1 & \text{if at least one station is located on path } p \in P_f \text{ of flow } f \in F \\ 0 & \text{otherwise} \end{cases}$$

$$y_f = \begin{cases} 1 & \text{if at least one station is located on } \textit{all} \text{ paths } p \in P_f \text{ of flow } f \in F \\ 0 & \text{otherwise} \end{cases}$$

$$z_f^p = \begin{cases} 1 & \text{if flow } f \in F \text{ travels non-intercepted on path } p \in P_f \\ 0 & \text{otherwise.} \end{cases}$$

The MRS model is as follows:

$$\text{(MRS) minimize} \sum_{(i,j)\in A} x_{ij} w_{ij} + \sum_{f\in F} \sum_{p\in P_f} z_f^p c_f^p \tag{1}$$

subject to

$$y_f^p \leq \sum_{(i,j)\in A^p} x_{ij} \qquad f \in F, p \in P_f \tag{2}$$

$$x_{ij} \leq y_f^p \qquad f \in F, p \in P_f, (i,j) \in A^p \tag{3}$$

$$y_f \leq y_f^p \qquad f \in F, p \in P_f \tag{4}$$

$$z_f^p \le 1 - y_f^p \qquad f \in F, p \in P_f \tag{5}$$

$$1 - y_f \le \sum_{p \in P_f} z_f^p \qquad f \in F \tag{6}$$

$$x_{ij} \in \{0,1\} \qquad (i,j) \in A \tag{7}$$

$$y_f, y_f^p, z_f^p \in \{0,1\} \qquad f \in F, p \in P_f. \tag{8}$$

The objective function minimizes the total station setup cost and the cost of damage on roads due to unintercepted vehicle flows. Constraints (2) and (3) enforce $y_f^p = 1$ if at least one station is located on path $p \in P_f$ of flow $f \in F$. Due to Constraints (4), $y_f = 1$ if at least one station is located on all paths $p \in P_f$. Constraints (5) ensure that vehicle flows only travel on non-intercepted paths. If not all paths of a flow $f \in F$ are intercepted, then the vehicle flow travels on the shortest non-intercepted path due to Constraints (6). Finally (7)–(8) are the integrality constraints. Note that the variables y_f, y_f^p and z_f^p assume binary values when their integrality constraints are relaxed (Marković, Ryzhov, and Schonfeld 2015).

Even though the EFCP is similar to the flow capturing problems (FCP) in a broad sense, it differs substantially from it since the behavior of the drivers is to avoid visiting the facilities. The main challenge arising in the EFCP is to determine the non-intercepted vehicle flows. Marković, Ryzhov, and Schonfeld (2015) tackle this challenge in their model by introducing binary variables y_f^p and z_f^p for each possible path, which introduces additional computational complexities, compared to the FCP. Furthermore, the way the deviation tolerance is handled is not applicable to problem instances of moderate sizes, since it is almost impossible to a priori enumerate all paths up to even very small deviation tolerance values for networks of practical sizes.

The following assumptions were made by Marković, Ryzhov, and Schonfeld (2015), which enabled them to model EFCP as a single-level mathematical program:

(a) the damage by overweight trucks increases linearly with the distance traveled;

(b) truckers seek to minimize their travel distance on non-intercepted paths.

Therefore, the minimization of the travel distance coincides with the minimization of the damage costs. Relaxing (b) yields a more general cost and damage setting and requires the problem to be tackled by a bilevel program. In our study, we also adopt the same assumptions, but our bilevel model, proposed in the following section, can be extended to more general cost settings.

3. Methodology

In this section, we present alternative models for the EFCP. We start by formulating a bilevel program in Section 3.1 to solve EFCP which does not require a priori route generation. In Section 3.2, we transform it into a single-stage equivalent formulation in compact form using constraint transformation and duality theory. We then reformulate the problem by projection in Section 3.3 and obtain a model with an exponential number of constraints in Section 3.4. We analyze the polyhedral properties of the model in Section 3.5 to show that the cut inequality is facet defining. Finally, in Section 3.6, we devise an exact branch-and-cut algorithm and present separation problems for integer and fractional solutions.

3.1. A Bilevel Program

There exist various network optimization problems for which bilevel programs have been developed, such as design (Kara and Verter 2004, Brotcorne et al. 2008), pricing (Labbé, Marcotte, and Savard 1998, Brotcorne et al. 2001, Kuiteing, Marcotte, and Savard 2017), and fare evasion (Correa et al. 2017) problems. The EFCP is also a bilevel problem in which two sets of players have different objectives: the law enforcement authority (the leader) locates the stations, and the unlawful drivers (the followers) attempt to drive by avoiding stations on their paths. The objective of the leader is to minimize the cost of setting up new stations and the total cost due to damage on roads caused by non-intercepted vehicle flows. The followers' objective, on the other hand, is to travel on their OD pairs by means of shortest non-intercepted paths.

Let u_f be an indicator variable equal to one if flow $f \in F$ is *not* intercepted, and zero otherwise. A flow f is intercepted if all paths between s_f and t_f within the deviation tolerance are intercepted. Hence, $u_f + y_f = 1$. Let ξ_{ij} be the shortest path distance between nodes $i, j \in N$ and $\bar{\lambda}_f = \lambda_f \xi_{s_f t_f}$. For a given arc $(i,j) \in A$ and flow $f \in F$, if $\xi_{s_f,i} + d_{ij} +$

$\xi_{j,t_f} > \bar{\lambda}_f$, then arc (i,j) is dominated for flow f. In other words, it is not possible for flow f to traverse arc $(i,j) \in A$ even when the vehicle travels on the shortest of all paths between nodes s_f and t_f using arc (i,j). For $f \in F$, let $A_f \subset A$ be the set of non-dominated arcs, $N_f \in N$ be the set of nodes induced by A_f and graph $G_f = (N_f, A_f)$. We define flow variables r^f_{ij} to equal one if arc $(i,j) \in A_f$ is traveled by vehicle flow $f \in F$. When the context is clear, we use $\mathbf{r^f}$ to refer to set of r^f_{ij} variables for all $(i,j) \in A_f$, and $\mathbf{r}$ to refer to all r^f_{ij} variables in the model. The other variables and parameters used in our model are as previously introduced. We refer to x, u and r variables as location, flow and arc variables, respectively. The new formulation, which we refer to as the 'Bilevel Model' (BM), is as follows:

$$\text{(BM)} \quad \underset{\mathbf{x},\mathbf{r}}{\text{minimize}} \sum_{(i,j)\in A} w_{ij} x_{ij} + \sum_{f\in F} \sum_{(i,j)\in A_f} c_f d_{ij} r^f_{ij} \tag{9}$$

subject to

$$x_{ij} \in \{0,1\} \qquad (i,j) \in A \tag{10}$$

where $\mathbf{r^f}$ solves the model of follower f, which we refer to as FM$_f$, for all $f \in F$:

$$\underset{\mathbf{u},\mathbf{r}}{\text{maximize}} \quad (\bar{\lambda}_f + \varepsilon) u_f - \sum_{(i,j)\in A_f} d_{ij} r^f_{ij} \tag{11}$$

$$\text{subject to} \quad \sum_{(i,j)\in A_f} r^f_{ij} - \sum_{(j,i)\in A_f} r^f_{ji} = \begin{cases} u_f & \text{if } i = s_f \\ -u_f & \text{if } i = t_f \\ 0 & \text{otherwise} \end{cases} \qquad i \in N_f \tag{12}$$

$$r^f_{ij} \le 1 - x_{ij} \qquad (i,j) \in A_f \tag{13}$$

$$u_f, r^f_{ij} \in \{0,1\} \qquad (i,j) \in A_f, \tag{14}$$

where ε is a very small value which can be taken as half the length of the shortest arc in G. The leader's objective is to minimize the total cost of setting up new stations and the cost of damage due to non-intercepted flows. There is a separate model for each follower $f \in F$ and the arc variables solve the followers' problems. The followers' objective in FM_f is to drive on a shortest non-intercepted path. Therefore, the objective is twofold: the primary objective is to drive without being intercepted between s_f and t_f on a path of length at most $\bar{\lambda}_f$. If this is achieved, then the secondary objective is to minimize the total traveled distance. In (11), we combine these two objectives into one. Constraints (12) ensure flow conservation, and Constraints (13) restrict the arcs with a station. Note that there is no additional constraint to enforce the deviation tolerance. The objective function ensures that if the shortest non-intercepted distance is longer than the tolerance value, then the flow variable and all arc variables will equal zero.

In the BM model, the non-intercepted flow $f \in F$ is indicated by the binary variable u_f. An alternative way of modeling this is to add a dummy arc (s_f, t_f) of length $\lambda_f + \varepsilon$ to graph G_f for all $f \in F$, in which case the corresponding arc variable $r_{s_f,t_f} = 1 - u_f$.

3.2. Linearization of the Bilevel Program

The general idea of using strong duality conditions for the followers' problem to convert a bilevel problem to a single level one can be traced back to Labbé, Marcotte, and Savard (1998). Building on the same idea, we first show that FM_f can be solved as a linear program (LP). Considering the fact that x_{ij} is a parameter in the followers' models, $1 - x_{ij}$ can be considered as the capacity of arc $(i,j) \in A_f$. It follows that, the constraint matrix of the FM_f formulation is the same as that of the classical maximum flow formulation (Ahuja, Magnanti, and Orlin 1993). Therefore the matrix defined by (12)–(14) is totally unimodular, and the followers' problems can be solved as LPs. Furthermore, for reasons that will become clearer in the following parts of the text, we replace (13) with an equivalent inequality:

$$x_{ij} r_{ij}^f \leq 0 \qquad (i,j) \in A_f.$$

This substitution preserves the linearity of the model and the integrality property of the LP solutions. Without loss of generality, we also change the equalities in (12) into inequalities. The final version of FM_f is as follows:

$$(\mathrm{FM}_f) \quad \underset{\mathbf{u},\mathbf{r}}{\text{maximize}} \quad (\bar{\lambda}_f + \varepsilon) u_f - \sum_{(i,j) \in A_f} d_{ij} r_{ij}^f \tag{15}$$

$$\text{subject to} \quad \sum_{(i,j) \in A_f} r_{ij}^f - \sum_{(j,i) \in A_f} r_{ji}^f \leq \begin{cases} u_f & \text{if } i = s_f \\ -u_f & \text{if } i = t_f \\ 0 & \text{otherwise} \end{cases} \qquad i \in N_f \tag{16}$$

$$x_{ij} r_{ij}^f \leq 0 \qquad (i,j) \in A_f \tag{17}$$

$$u_f \leq 1 \tag{18}$$

$$u_f, r_{ij}^f \geq 0 \qquad (i,j) \in A_f. \tag{19}$$

Let π_i^f, μ_{ij}^f and η^f be dual variables associated with Constraints (16), (17) and (18), respectively. Then the dual problem is

$$\text{minimize } \eta^f \tag{20}$$

$$\text{subject to } \pi_i^f - \pi_j^f + \mu_{ij}^f x_{ij} + d_{ij} \geq 0 \qquad (i,j) \in A_f \tag{21}$$

$$\pi_{t_f}^f - \pi_{s_f}^f + \eta^f \geq \bar{\lambda}_f + \varepsilon \tag{22}$$

$$\eta^f, \pi_k^f, \mu_{ij}^f \geq 0 \qquad k \in N_f, (i,j) \in A_f. \tag{23}$$

Remark 1. For $f \in F$, we define the 'non-intercepted shortest path' to be the shortest path on the subgraph of G_f induced by the arcs on which no station exists. For a given solution (u^*, r^*) of the FM_f formulation, if $u_f^* = 1$, then a non-intercepted shortest path can be extracted from the r^* variables. Therefore, the π_i variable in the dual formulation has a very natural interpretation: it is a label for the non-intercepted shortest path from node s_f to node $i \in N_f$. □

At this point, the objective is to bring the primal and the dual formulations together to form a single-stage equivalent formulation. Note that due to multiplication of the μ and x variables in Constraints (21), the single-stage formulation contains non-linearities. However, these can be eliminated without introducing additional variables or constraints as a result of Constraints (13) being replaced with Constraints (17):

Proposition 1. *For $f \in F$, there exists an optimal solution of formulation (20)–(23) with $\mu_{ij}^f = \bar{\lambda}_f + \varepsilon$, for all $(i,j) \in A_f$.*

Proof At any optimal solution, we have $\eta^f = \max\{\bar{\lambda}_f + \varepsilon - \pi_{t_f}^f + \pi_{s_f}^f, 0\}$ due to Constraint (22). We can therefore reformulate the problem as

$$\begin{gathered} \text{maximize } \pi_{t_f}^f - \pi_{s_f}^f \\ \text{subject to } (21), (23). \end{gathered}$$

Without loss of generality, set $\pi^f_{s_f}=0$. The x_{ij} and μ^f_{ij} variables do not directly contribute to the objective function. However their product $\mu^f_{ij}x_{ij}$ can be set arbitrarily large to relax Constraints (21) in order to possibly increase the value of $\pi^f_{t_f}$ in the objective function. It is therefore straightforward to see that the variable μ^f_{ij} enforces a constraint when $x_{ij}=0$ and relaxes it when $x_{ij}=1$. This line of logic is in accordance with the interpretation of the π variables as labels to keep the non-intercepted shortest path from node s_f to node i (Remark 1). Since the path between s_f and t_f can at most be $\bar{\lambda}_f$, we have the desired result. □

A similar transformation was also applied by Amaldi, Bruglieri, and Fortz (2011). We now transform the BM into a single-stage model. Instead of using the standard Karush-Kuhn-Tucker conditions, we reformulate the model by appending the primal constraints, dual constraints and the so-called reverse weak duality inequality, which states that the objective function value of the primal maximization problem is at least equal to the objective function value of the dual minimization problem. Observe that, due to strong duality in LPs, only an optimal solution satisfies these conditions. The same conditions were also used by Fontaine and Minner (2014), Amaldi, Bruglieri, and Fortz (2011), Cao and Chen (2006) to transform bilevel programs into single stage in different applications. We further eliminate the $u_f \leq 1$ constraints for all $f \in F$ since they are implied. We have now converted the BM into a single-stage mixed integer linear programming model, which we refer to as the single-stage model (SM):

$$\text{(SM) minimize} \sum_{(i,j)\in A} w_{ij}x_{ij} + \sum_{f\in F}\sum_{(i,j)\in A_f} c_f d_{ij} r^f_{ij} \tag{24}$$

$$\text{subject to } \pi^f_j \leq \pi^f_i + d_{ij} + (\bar{\lambda}_f+\varepsilon)x_{ij} \qquad f\in F, (i,j)\in A_f \tag{25}$$

$$(\bar{\lambda}_f+\varepsilon)(1-u_f) + \sum_{(i,j)\in A_f} d_{ij}r^f_{ij} \leq \pi^f_{t_f} - \pi^f_{s_f} \qquad f\in F \tag{26}$$

$$\sum_{(i,j)\in A_f} r^f_{ij} - \sum_{(j,i)\in A_f} r^f_{ji} \leq \begin{cases} u_f & \text{if } i=s_f \\ -u_f & \text{if } i=t_f \\ 0 & \text{otherwise} \end{cases} \qquad f\in F, i\in N_f \tag{27}$$

$$r^f_{ij} \leq 1-x_{ij} \qquad f\in F, (i,j)\in A_f \tag{28}$$

$$x_{ij}\in\{0,1\} \qquad (i,j)\in A \tag{29}$$

$$u_f, \pi^f_k, r^f_{ij} \geq 0 \qquad f\in F, k\in N_f, (i,j)\in A_f. \tag{30}$$

Constraints (25) are imposed for dual feasibility. Constraints (26) are the reverse weak duality inequality, enforcing $u_f = 1$ if the non-intercepted shortest path length from the origin to the destination does not exceed the tolerance $\bar{\lambda}_f$, for all $f \in F$. Constraints (27)–(28) are imposed for primal feasibility. Finally, Constraints (29) and (30) are the integrality requirements and the non-negativity constraints, respectively. Note that when $u_f = 1$, the existence of a path with a length at most $\bar{\lambda}_f$ is ensured by (25)–(26). Therefore, we do not explicitly force a path expressed by arc variables to have a length of at most $\bar{\lambda}_f$.

3.3. Reformulation by Projection

We now project SM onto a subspace in order to obtain an alternative formulation. Let SMR be the linear relaxation of SM. For $f \in F$, we define $\mathcal{Q}_f^{SM}$ to be the projection of the SMR polytope onto the space spanned by the variables x, u and r. More formally, $\mathcal{Q}_f^{SM} = \{(u,x,r) \in \mathbb{R}^{|F|+|A|+\sum_{f\in F}|A_f|} |$ there exists $\pi \in \mathbb{R}^{|F|}$ such that $(\pi,u,x,r) \geq 0$ satisfies (25) and (26)$\}$. To describe this set by means of linear functions, we first investigate the projection cone and its extreme rays. Let α_{ij}^f and β_f be the dual variables associated with Constraints (25) and (26), respectively. When the context is clear, we refer to the vector of α_{ij}^f for all $(i,j) \in A_f$ as α^f, to the vector of α^f for all $f \in F$ as α and to the vector of β_f for all $f \in F$ as β.

By the Farkas lemma (see, e.g., Conforti, Cornuéjols, and Zambelli 2010), we have the following result: Given a solution (x,u,r), there exists a vector π satisfying (25) and (26) if and only if

$$0 \leq \sum_{f\in F}\Big[\sum_{(i,j)\in A_f}(d_{ij} + (\bar{\lambda}_f + \varepsilon)x_{ij})\alpha_{ij}^f + (\bar{\lambda}_f + \varepsilon)(u_f - 1)\beta_f - \sum_{(i,j)\in A_f} d_{ij} r_{ij}^f \beta_f\Big] \tag{31}$$

for all $(\alpha,\beta) \geq 0$ such that

$$\sum_{j:(i,j)\in A_f}\alpha_{ij}^f - \sum_{j:(j,i)\in A_f}\alpha_{ji}^f = \begin{cases}\beta_f & \text{if } i = s_f\\ -\beta_f & \text{if } i = t_f\\ 0 & \text{otherwise}\end{cases} \qquad f \in F, i \in N_f. \tag{32}$$

Let $\mathcal{C}_f$ be the cone of $(\alpha^f,\beta_f) \geq 0$ satisfying inequalities (32) for $f \in F$ and $\mathcal{C} = \cup_{f\in F}\mathcal{C}_f$. Note that $\mathcal{C}$ is a pointed cone with apex at $\mathbf{0}$. Non-dominated projection inequalities of type (31) are defined by the extreme rays of $\mathcal{C}$, which we characterize in the following proposition:

Proposition 2. *For $f \in F$, a vector $(\alpha^f, \beta_f) \in \mathcal{C}_f$ of size $|A_f|+1$ defines an extreme ray if and only if $\beta_f = \epsilon$ and*

$$\alpha^f_{ij} = \begin{cases} \epsilon & \textit{if } (i,j) \in A^p \\ 0 & \textit{otherwise} \end{cases} \qquad (i,j) \in A_f, \tag{33}$$

where $\epsilon \geq 0$ and p is a directed path in G_f.

Proof Let $(\alpha^f, \beta_f) \in \mathcal{C}_f$ be a vector where α^f is defined as in (33) and $\beta_f = \epsilon$. Now, suppose that $\alpha^f = \nabla \upsilon^1 + (1-\nabla)\upsilon^2$ for some $(\upsilon^1, \psi^1), (\upsilon^2, \psi^2) \in \mathcal{C}_f$ and $0 < \nabla < 1$. Then $\nabla \upsilon^1_{ij} + (1-\nabla)\upsilon^2_{ij} = 0$ for $(i,j) \in A_f \setminus A^p$. This implies that $\upsilon^1_{ij} = \upsilon^2_{ij} = 0$ for all $(i,j) \in A_f \setminus A^p$. If there exists $(k,l) \in A^p$ such that $\upsilon^1_{kl} = 0$, we have $\upsilon^1_{ij} = 0$, for all $(i,j) \in A^p$ due to (32). It then follows that υ^2 is a positive multiple of α^f, which is a contradiction. Therefore, suppose $\upsilon^1_{ij} > 0$ and $\upsilon^2_{ij} > 0$ for all $(i,j) \in A^p$. But then, due to (32), υ^1 is a positive multiple of υ^2. Hence (α^f, β_f) defines an extreme ray.

Let $(\upsilon, \psi) \in \mathcal{C}_f$ be an extreme ray. Then υ satisfies (32). If $\psi = 0$, then $\upsilon = 0$ due to (32) and we get the apex of the cone. Let $\epsilon > 0$ and $\psi = \epsilon$. Then, (32) enforces υ to define a path q from s_f to t_f and the value of $\upsilon_{ij} = \psi = \epsilon$, for all $(i,j) \in A^q$. Thus any extreme ray is a path in G_f and the proposition follows. $\square$

Let $\overline{P_f}$ be the set of all paths from s_f to t_f in G_f. Consider a vehicle flow $f \in F$, a path $p \in \overline{P_f}$ and the (α^f, β_f) setting in Proposition 2 for $\epsilon = \dfrac{1}{(\bar{\lambda}_f + \varepsilon)} > 0$. Then, (31) is transformed into the following inequality:

$$1 - \sum_{(i,j)\in A^p} x_{ij} \leq u_f + \sum_{(i,j)\in A^p} \frac{d_{ij}}{(\bar{\lambda}_f + \varepsilon)} - \sum_{(i,j)\in A_f} \frac{d_{ij} r^f_{ij}}{(\bar{\lambda}_f + \varepsilon)}. \tag{34}$$

We refer to length of path p as $\mathcal{L}(p) = \sum_{(i,j)\in A^p} d_{ij}$. Rewriting (34) for all $f \in F$ and $p \in \overline{P_f}$, we have the following *projection inequality*:

$$1 - \sum_{(i,j)\in A^p} x_{ij} \leq u_f + \frac{\mathcal{L}(p) - \sum_{(i,j)\in A_f} (d_{ij} r^f_{ij})}{(\bar{\lambda}_f + \varepsilon)} \qquad f \in F, p \in \overline{P_f}. \tag{35}$$

With this projection, we now have a linear description of the $\mathcal{Q}^{SM}_f$. We can remove Constraints (25) and (26) from the SM formulation and append (35) to it to yield an equivalent formulation. We refer to this formulation as the Projection Model (ProjM).

At this point we further investigate the projection inequality. A few remarks are in order.

Remark 2. Preliminary computational experiments with the Porta software (Christof, Löbel, and Stoer 1997) have shown that the extreme points of the ProjM formulation are generally fractional, due to the fractional term in the projection inequality.

Remark 3. The projection inequality (35) is valid for all paths in G_f. However, for both intercepted and non-intercepted paths with $\mathcal{L}(p) > \bar{\lambda}_f$, the inequality can be satisfied by setting the flow and arc variables trivially equal to zero. Hence, we consider paths of length $\mathcal{L}(p) \leq \bar{\lambda}_f$, depending on the following two cases:

(a) Consider an intercepted path p of length $\mathcal{L}(p) \leq \bar{\lambda}_f$. Since p is intercepted, we have $\sum_{(i,j)\in A^p} x_{ij} \geq 1$. Therefore the projection inequality can be satisfied by setting the flow and arc variables trivially equal to zero.

(b) Consider the only non-trivial case when a path p of length $\mathcal{L}(p) \leq \bar{\lambda}_f$ is non-intercepted. Then, we have $\sum_{(i,j)\in A^p} x_{ij} = 0$ and the projection inequality can be rewritten as

$$1 + \frac{\sum_{(i,j)\in A_f} (d_{ij} r_{ij}^f)}{(\bar{\lambda}_f + \varepsilon)} \leq u_f + \frac{\mathcal{L}(p)}{(\bar{\lambda}_f + \varepsilon)}.$$

The inequality $\mathcal{L}(p) \leq \bar{\lambda}_f$ implies that $u_f > 0$. Since this is a binary variable, we have $u_f = 1$. Furthermore, u_f cancels out with 1 on the left-hand side, and the inequality states that the arc variables cannot exceed the length of path p:

$$\sum_{(i,j)\in A_f} (d_{ij} r_{ij}^f) \leq \mathcal{L}(p).$$

This is true for all unintercepted paths $p \in P_f$. Therefore the length of the unintercepted path that trucks travel (identified by the arc variables) is the shortest of all unintercepted paths. This is valuable information, but it is not required for the correctness of the model since the path identified by arc variables is the shortest of all unintercepted paths due to Constraints (27), (28) and the objective function.

Remark 4. For all the cases we considered, the fractional term in the projection inequality lies in the half-closed $[0, 1)$ interval. Having the location and flow variables as binary allows us to omit the fractional term, and we obtain the following inequality, which we refer to as cut inequality:

$$1 - \sum_{(i,j)\in A^p} x_{ij} \leq u_f \qquad\qquad f\in F, p\in P_f.$$

The following observations are due to the cut inequality:

(a) For $f \in F$ and a path $p \in P_f$ with $\mathcal{L}(p) \leq \bar{\lambda}_f$, having no station on p (i.e. $\sum_{(i,j)\in A^p} x_{ij} = 0$) implies that p is non-intercepted.

(b) The existence of at least one such non-intercepted path with $\mathcal{L}(p) \leq \bar{\lambda}_f$ implies that the overweight trucks can travel non-intercepted (i.e. $u_f = 1$).

3.4. Pathcut Formulation

In this section, we present an alternative formulation based on the cut inequality. We refer to the following as the pathcut model (PM):

$$\text{(PM) minimize} \sum_{(i,j)\in A} w_{ij}x_{ij} + \sum_{f\in F}\sum_{(i,j)\in A_f} c_f d_{ij} r_{ij}^f \tag{36}$$

subject to

$$1 - \sum_{(i,j)\in A^p} x_{ij} \leq u_f \qquad f\in F, p\in P_f \tag{37}$$

$$\sum_{(i,j)\in A_f} r_{ij}^f - \sum_{(j,i)\in A_f} r_{ji}^f = \begin{cases} u_f & \text{if } i = s_f \\ -u_f & \text{if } i = t_f \\ 0 & \text{otherwise} \end{cases} \qquad f\in F, i\in N_f \tag{38}$$

$$r_{ij}^f \leq 1 - x_{ij} \qquad f\in F, (i,j)\in A_f \tag{39}$$

$$x_{ij} \in \{0,1\} \qquad (i,j)\in A \tag{40}$$

$$u_f, r_{ij}^f \geq 0 \qquad f\in F, (i,j)\in A_f. \tag{41}$$

The objective function is the same as that of the SM formulation. Constraints (37) ensure that u_f equals one if there exists at least one non-intercepted path between s_f and t_f in G_f. Constraints (38)–(41) are the same in SM. Even though there is an exponential number of constraints in the PM formulation, we eliminate the need to use big-M type constraints. Furthermore, observe that there is no constraint explicitly handling the λ_f parameter. The deviation tolerance is implicitly taken into account by the definition of path $p \in P_f$. Also note that the cut inequality can also be used to model the multi-period EFCP (Marković, Ryzhov, and Schonfeld 2017). In the following section, we investigate the strength of the cut inequality and then discuss effective separation algorithms for Constraints (37).

3.5. Polyhedral Analysis

Let X be the set of feasible solutions in formulation PM, and $\mathscr{P} = conv(X)$ be the convex hull of X. In this section, we prove that the cut inequality defines a facet of $\mathscr{P}$. For $f \in F$, we assume that graph G_f is connected (if not, then without loss of generality, f can be removed from the set F).

Let e_{ij} be the unit vector of size $|A|$ corresponding to arc $(i,j) \in A$, h_f be the unit vector of size $|F|$ corresponding to flow $f \in F$, p_f be a path in G_f, $v_{p_f} = \sum_{(i,j) \in A^{p_f}} e_{ij}$ be the vector corresponding to path p_f (in the domain of location variables), $w_{p_f} \in \{0,1\}^{\sum_{g \in F} |A_g|}$ be the vector corresponding to $f \in F$ and path p_f (in the domain of arc variables) and $F(S) \subset F$ be the set of vehicle flows, that can travel from their origins to their destinations in the graph induced by the arcs in S. In the following, we refer to a solution in $\mathscr{P}$ as $(\mathbf{x}, \mathbf{u}, \mathbf{r})$. In order to determine the dimension of $\mathscr{P}$, we start by presenting a graph theoretical notion, the number of linearly independent paths, which is usually used as a complexity measure in software source codes (McCabe 1976).

Definition 1. For $f \in F$, a set of paths P is referred to as linearly independent if the set of corresponding vectors $v_p,\ p \in P$ is linearly independent.

Lemma 1. *(McCabe 1976) In a connected graph $G = (N, A)$, the number of linearly independent paths between a source $s \in N$ and a destination $t \in N$ is equal to $|A| - |N| + 2$.*

The proof is by induction, see McCabe (1976) and Berge and Minieka (1973). Let $\mathcal{Y}_f$ be a set of linearly independent paths in G_f.

To avoid considering single-arc paths, we introduce, without loss of generality, a dummy node on all such arcs. We are now ready to discuss the dimension of $\mathscr{P}$.

Proposition 3. $dim(\mathscr{P}) = |A| + \sum_{f \in F}(|A_f| - |N_f| + 2)$.

Proof The number of variables in PM is $|A|+|F|+\sum_{f\in F}|A_f|$. In all solutions of PM, Constraints (38) are satisfied at equality. Since they express node balance equations in graphs, the rank of the corresponding matrix is $\sum_{f\in F}(|N_f|-1)$. Therefore $|A|+\sum_{f\in F}(|A_f|-|N_f|+2)$ is an upper bound on the dimension. Now, consider the following solutions:

$$\begin{pmatrix}\mathbf{1}\\0\\0\end{pmatrix},\begin{pmatrix}\mathbf{1}-e_{ij}\\0\\0\end{pmatrix}\quad (i,j)\in A,\begin{pmatrix}\kappa_{p_f}\\\phi_{p_f}\\\pi_{p_f}\end{pmatrix}\quad p_f\in\mathcal{Y}_f, f\in F,$$

where $\kappa_{p_f}=\mathbf{1}-v_{p_f}, \phi_{p_f}=\sum_{g\in F(A^{p_f})}h_g$ and $\pi_{p_f}=\sum_{\substack{p_g:g\in F(A^{p_f}),\\A^{p_g}\subset A^{p_f}}}w_{p_g}$. Having these $|A|+\sum_{f\in F}(|A_f|-|N_f|+2)+1$ number of affinely independent solutions gives a lower bound on the dimension and the proposition follows. $\square$

PROPOSITION 4. *For $f\in F$ and $p\in P_f$, the inequality $1-\sum_{(i,j)\in A^p}x_{ij}\le u_f$ is facet-defining for $\mathscr{P}$ if and only if no subpath of p constitutes a complete path between $s_{\hat f}$ and $t_{\hat f}$ for a vehicle flow $\hat f\in F\setminus\{f\}$.*

Proof In a solution with $x_{ij}=1$ for all $(i,j)\in A^p$, the cut inequality strictly holds. Since $\mathscr{P}$ is not full dimensional, this establishes that it is not an implicit equation. Let $\mathscr{R}=\{(x,u,r)\in\mathscr{P}:\sum_{(i,j)\in A^p}x_{ij}+u_f=1\}$. Suppose that all solutions $(x,u,r)\in\mathscr{R}$ also satisfy

$$ax+bu+cr=\varphi, \tag{42}$$

where a,b and c are coefficient vectors of appropriate sizes. Consider an arc $(i,j)\in A^p$ and a solution $(\sum_{(k,l)\in S_{ij}}e_{kl},0,0)\in\mathscr{R}$, where S_{ij} is a minimal cut in G with $S_{ij}\cap A^p=\{(i,j)\}$, disconnecting $(s_g,\,t_g)$ OD pairs for all $g\in F$. Since no subpath of p is a path for another OD pair, the existence of S_{ij} is well defined. Then, (42) implies $\sum_{(k,l)\in S_{ij}}a_{kl}=\varphi$. Now, consider another solution $(\sum_{(k,l)\in A\setminus A^p}e_{kl}+e_{ij},0,0)\in\mathscr{R}$, which implies $\sum_{(k,l)\in A\setminus A^p}a_{kl}+a_{ij}=\varphi$. Therefore, $a_{kl}=0$ for all $(k,l)\in A\setminus(A^p\cup S_{ij})$. If $S_{ij}=\{(i,j)\}$, then we have $a_{ij}=\varphi$. Else, consider another arc $(m,n) \in S_{ij} \setminus \{(i,j)\}$. Observe that there exists another minimal cut $\bar{S}_{ij}$ in G, disconnecting all OD pairs:

$$\bar{S}_{ij} = \begin{cases} S_{ij} \setminus \{(m,n)\} \cup \{(n,l) \in \delta^+(n)\} & \text{if } m = s_f, \\ S_{ij} \setminus \{(m,n)\} \cup \{(k,m) \in \delta^-(m)\} & \text{otherwise} \end{cases}$$

where $\delta^+(n)$ and $\delta^-(m)$ are the emanating arcs from node n and entering arcs into node m, respectively. Similar to the previous argument with arc (i,j), considering solutions $(\sum_{(k,l)\in\bar{S}_{ij}} e_{kl}, 0, 0)$ and $(\sum_{(k,l)\in A\setminus A^p} e_{kl} + e_{ij}, 0, 0) \in \mathscr{R}$ yields $a_{kl} = 0$ for all $(k,l) \in A \setminus (A^p \cup \bar{S}_{ij})$, which implies that $a_{mn} = 0$. Iterating over all arcs $(m,n) \in S_{ij} \setminus \{(i,j)\}$, we obtain the following result: $a_{ij} = \varphi$ and $a_{kl} = 0$ for all $(k,l) \in A \setminus A^p$. Iterating over all arcs $(i,j) \in A^p$ yields

$$a_{ij} = \begin{cases} \varphi & \text{if } (i,j) \in A^p, \\ 0 & \text{otherwise.} \end{cases}$$

Then inequality (42) becomes $\sum_{(i,j)\in A^p} \varphi x_{ij} + \sum_{g\in F} b_g u_g + \sum_{\substack{g\in F \\ (i,j)\in A_g}} c_{ij} r_{ij}^g = \varphi$. Next, consider a solution $(\sum_{(k,l)\in A\setminus A^p} e_{kl}, h_f, w_p) \in \mathscr{R}$, which gives $b_f + \sum_{(i,j)\in A^p} c_{ij} = \varphi$. Let Q be a set of paths with a single path from each vehicle flow's set of linearly independent paths, except flow f. Observe that there exist $\prod_{g\in F\setminus\{f\}} |\mathcal{Y}_f| = \prod_{g\in F\setminus\{f\}} (|A_g| - |N_g| + 2)$ alternative ways of forming the set Q. Consider all such sets and the corresponding solutions $(0, \sum_{g\in F} h_g, \sum_{q\in Q} w_q + w_p) \in \mathscr{R}$. We then obtain $\sum_{g\in F\setminus\{f\}} b_g + \sum_{\substack{g\in F\setminus\{f\} \\ (i,j)\in A_g}} c_{ij} = 0$. This implies $b_g = 0$ for all $g \in F \setminus \{f\}$, $c_{ij} = 0$ for all $g \in F \setminus \{f\}, (i,j) \in A_g$. Then (42) becomes

$$\sum_{(i,j)\in A^p} \varphi x_{ij} + b_f u_f + \sum_{(i,j)\in A_f} c_{ij} r_{ij}^f = \varphi. \tag{43}$$

Now consider path $q \in \mathcal{Y}_f$ and solution $(\sum_{(k,l)\in A\setminus(A^p\cup A^q)} e_{kl}, h_f, w_q) \in \mathscr{R}$. This gives the following equations:

$$b_f + \sum_{(i,j)\in A^q} c_{ij} = \varphi \qquad\qquad q \in \mathcal{Y}_f. \tag{44}$$

For the considered solution corresponding to path q, we have $r^q_{ij} = u_f$, for all $(i,j) \in A^q$ due to node balance constraints in PM. Therefore, (43) can be rewritten as

$$\sum_{(i,j)\in A^p} \varphi x_{ij} + b_f u_f + \sum_{(i,j)\in A^q} c_{ij} u_f = \sum_{(i,j)\in A^p} \varphi x_{ij} + (b_f + \sum_{(i,j)\in A^q} c_{ij}) u_f = \varphi.$$

From (44), we have $\sum_{(i,j)\in A^p} \varphi x_{ij} + \varphi u_f = \varphi$, which is a positive multiple of the cut inequality.

Now assume that the condition does not hold, that is, there exists a vehicle flow $\bar{f}$ such that path p has a subpath sp from $s_{\bar{f}}$ to $t_{\bar{f}}$. Then, $1 - \sum_{(i,j)\in A^{sp}} x_{ij} \le u_{\bar{f}}$ and $u_{\bar{f}} \le u_f + \sum_{(i,j)\in A^p \setminus A^{sp}} x_{ij}$ are valid, they imply the cut inequality and therefore the cut inequality cannot be facet defining. □

Remark 5. For any $0 \le \chi \le 1$, equation (44) can be satisfied by:

$$b_f = \chi\varphi \text{ and } c_{ij} = \begin{cases} (1-\chi)\varphi & \text{if } (i,j) \in S \\ 0 & \text{otherwise,} \end{cases}$$

where S is an s_f-t_f minimum cut in G_f. This implies that the family of inequalities $1 - \sum_{(i,j)\in A^p} x_{ij} \le \chi u_f + (1-\chi) \sum_{(i,j)\in S} r^f_{ij}$ are valid for $\mathscr{P}$ for $0 \le \chi \le 1$. However, all these inequalities are equivalent since $u_f = \sum_{(i,j)\in S} r^f_{ij}$ due to the flow balance equations in PM.

3.6. Separation Problem

For a given solution (x^*, u^*, r^*) and a vehicle flow f^* with $u^*_f < 1$, the separation problem is to identify a path $p \in P_f$ such that $1 - \sum_{(i,j)\in A^p} x^*_{ij} > u^*_f$, or to conclude that none exists. Note that by definition, we have $\mathcal{L}(p) \le \bar{\lambda}_f$. In the following, we present an exact separation algorithm for integer solutions by solving a shortest path problem, and an exact separation algorithm for fractional solutions by solving a resource constrained shortest path problem.

3.6.1. Exact separation algorithm for integer solutions. When the solution (x^*, u^*, r^*) is integer, we construct graph G^*_{Int} induced by arcs with $x^*_{ij} = 0$. Any path in this graph is a non-intercepted path. For a vehicle flow $f \in F$ with $u^*_f = 0$, if the length of the shortest path between s_f and t_f is at most $\bar{\lambda}_f$, then we have identified a path p^* with $\sum_{(i,j)\in A^{p^*}} x^*_{ij} = 0$, separating the current solution. If the length is greater than $\bar{\lambda}_f$ or nodes s_f and t_f are disconnected, then the solution does not need to be separated.

3.6.2. Exact separation algorithm for fractional solutions. When the solution (x^*, u^*, r^*) is fractional, we construct the graph $G^*_{Frac} = (N, \hat{A}_{Frac})$ where $\hat{A}_{Frac} = \{(i,j) \in A\}$ with the weight of arc $(i,j) \in \hat{A}_{Frac}$ equals x^*_{ij}. We consider actual path lengths as resource in the transformed graph. For a vehicle flow with $u^*_f < 1$, we compute a resource-constrained elementary shortest path p^* between s_f and t_f with a total resource of $\bar{\lambda}_f$. If $1 - \sum_{(i,j)\in A^{p^*}} x^*_{ij} > u^*_f$, we separate the solution. If not, we conclude that no such path exists.

3.6.3. Branching. The branching rules are applied with the default settings of CPLEX.

3.6.4. Initial set of rows. For each OD pair $f \in F$, we add Constraint of type (37) for the shortest path between s_f and t_f.

3.6.5. Feasibility heuristic. Consider an integral solution (x^*, u^*, r^*) and let $F^* = \{f \in F : u^*_f = 0 \text{ and } \mathcal{L}(p^*_f) \leq \bar{\lambda}_f\}$ where p^*_f is the shortest path in G_f between s_f and t_f with $\sum_{(i,j)\in A^{p^*_f}} x^*_{ij} = 0$. Note that the infeasibility stems from the vehicle flows $f \in F^*$, and we identify set F^* when separating integer solutions as in Section 3.6.1. Using this information, a feasible solution can readily be constructed by setting $u_f = 1$ and the arc variables corresponding to path p^*_f equal to 1, $r^f_{ij} = 1$, for all $(i,j) \in A^{p^*_f}$ and leaving all remaining variables unaltered. Note that such a feasible solution can be obtained at no cost at every integer separation iteration.

4. Computational Study

We now present the experimental setup and the computations to assess the efficiency of the proposed solution methodologies: the Marković, Ryzhov, and Schonfeld model (MRS), our single-stage model (SM) and our pathcut model (PM). We have implemented the algorithms using Java under Linux and CPLEX 12.6.1, and all experiments were conducted on a cluster of 27 machines each having two Intel(R) Xeon(R) X5675 3.07 GHz processors with 96 GB of RAM running on Linux. Each machine has 12 cores and each experiment was run using a single thread. The time limit for all experiments is 3600 seconds. For graph representations, we use JGraphT Java Library (Sichi and Contributors 2017). For the shortest path problem, we use the Dijkstra's algorithm with Fibonacci heap implementation of JGraphT. To solve the resource constrained shortest path problem, we use the pulse algorithm introduced by Lozano and Medaglia (2013). The Java code of the algorithms is available as an online supplement. In the following, we discuss the data needed for our experiments, the implementation details of the algorithms, and the results. In all experiments, we use a common λ value, independent of the vehicle flow.

4.1. Data

To test our algorithms, we use WIM station location instances of the EFCP. The pavement damage caused by trucks depends on several parameters such as axle weights, pavement structure and climate and the environmental damage includes accidents, emissions and noise (Marković, Ryzhov, and Schonfeld 2015). For a given truck class, the regulations impose an upper bound on the total transportation weight. Provided with the class of a truck and the excessive weight, there are methods to estimate the total damage. This estimation provides a cents per kilometer value for a given truck. There are several types of trucks traveling between a given OD pair with varying overweights. Marković et. al assume that the percent of total flow by a specific truck type can be estimated. Therefore, all calculations boil down to a single parameter c_f, which is the excessive damage cost per mile for vehicles traveling between OD pair $f \in F$. This parameter is calculated using OD pair specific data for various truck classes and overweight loads. Once an estimate is made, it is multiplied by the number of vehicles traveling yearly between an OD pair and the distance of the road to find the annual cost of damage due to transportation of overweight trucks between the OD pairs. For further details on the comprehensive calculation of these values, we refer the reader to Marković, Ryzhov, and Schonfeld (2015). As in Marković et. al, we conducted experiments for various realistic values of $c_f \in \{0.025, 0.05, 0.10, 0.20\}$ in dollars per kilometer and $w_i \in \{10, 60, 110, 160, 260, 360\}$ in thousands of dollars. As in Hodgson (1990), we calculate the vehicle flows between OD pairs based on a gravity model. According to this model, the flow is directly proportional to the total population of the origin and destination nodes, and inversely proportional to the shortest distance of the OD pair. We additionally assume that for every 1000 people at origin and destination nodes, one truck travels between this OD pair.

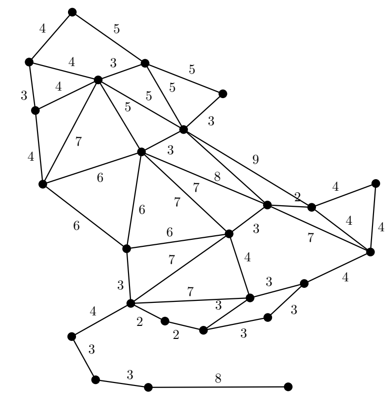


Figure 1 25-node network with 25 nodes and 86 arcs

In their experiments, Marković, Ryzhov, and Schonfeld (2015) solved a 205-node instance with an average node degree of 1.078, which corresponds to the road network of Nevada with state-designated links for large commercial vehicles. This network is used to show that the FCP model performs very poorly even when there are very few alternative paths. Instances with better connected networks are also solved. In our experiments, we consider three network topologies (Table 1): a 25-node network (Figure 1) used by Hodgson (1990), Kuby and Lim (2005), Lim and Kuby (2010), MirHassani and Ebrazi (2013) and Capar et al. (2013), the 339-node California road network (CA) (Figure 2) used by Arslan, Yıldız, and Karaşan (2014), Yıldız, Arslan, and Karaşan (2016), Arslan and Karaşan (2016), and three 500-node randomly generated road networks (see Section 4.4). The average node degrees of these networks vary between 3.44 and 3.64. We consider three deviation tolerances: 1.0, 1.5 and 2.0.

Table 1 Network characteristics

			Node degree			OD Pairs			
Network	# nodes	# arcs	min	avg	max	# OD pairs	min dist	avg dist	max dist
25-node	25	86	1	3.44	6	300	20.00	142.33	380.00
CA	339	1234	1	3.64	7	1167	30.06	153.37	463.50
N500-50	500	1820	1	3.64	11	500	30.6	61.89	108.65
N500-75	500	1820	1	3.64	11	1000	30.18	60.95	106.82
N500-100	500	1820	1	3.64	11	2000	30.02	59.37	123.38

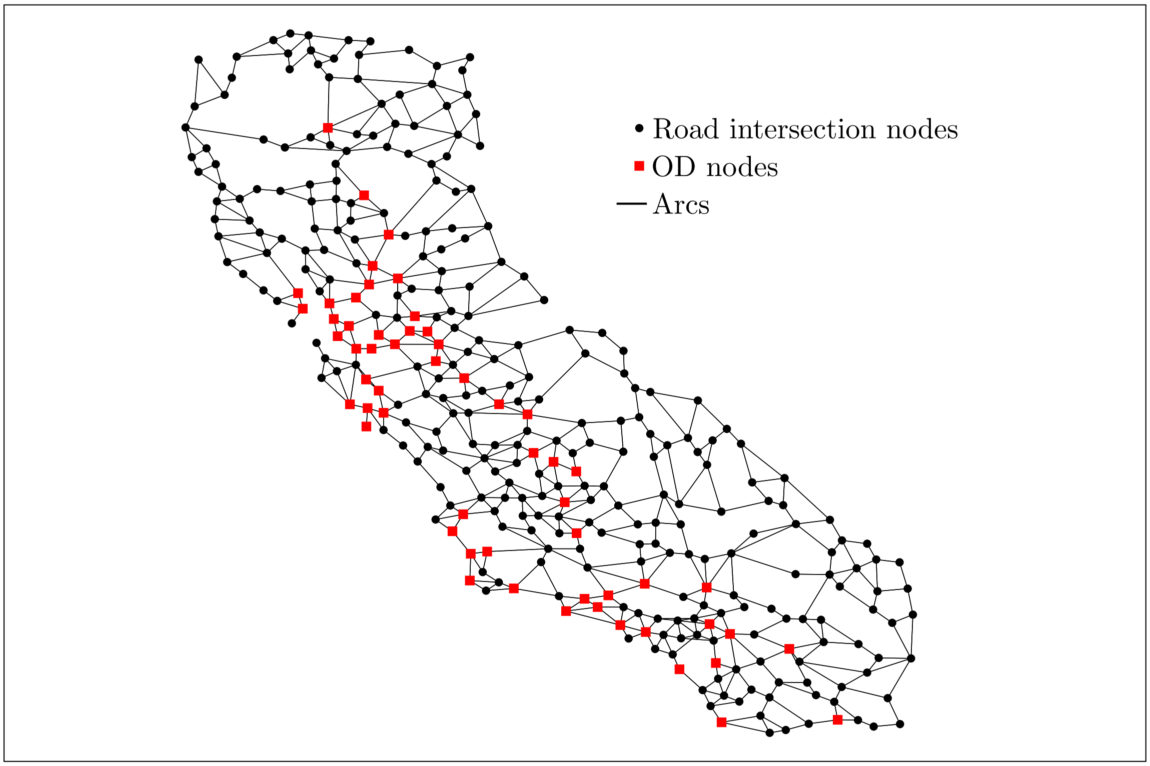


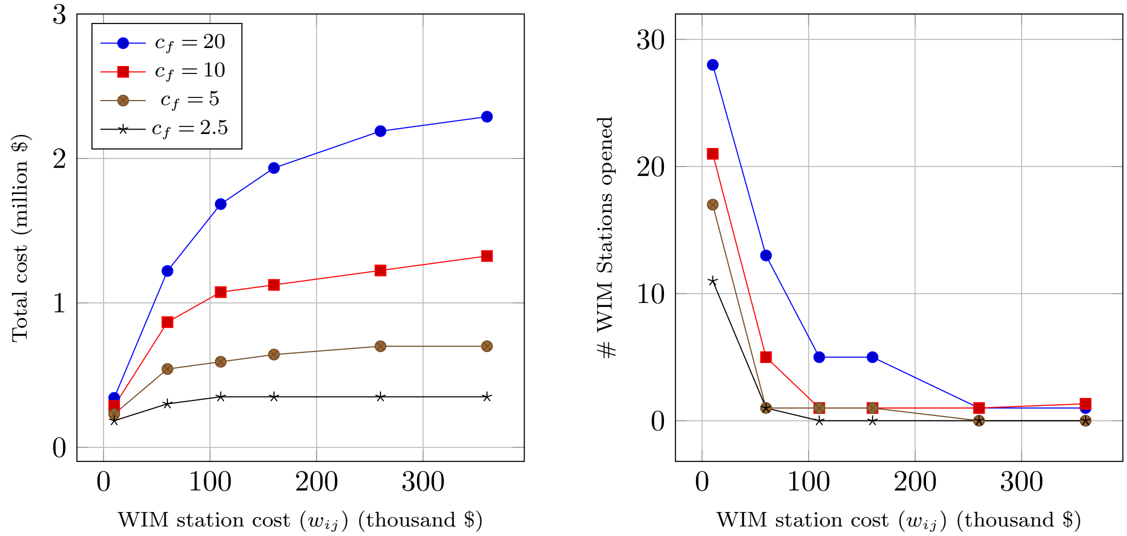
Figure 2 The California road network with 339 nodes and 1234 arcs

4.2. The 25-node Network

Figure 3(a) displays the total cost as a function of the WIM station cost (w_{ij}) and of the damage coefficient of vehicles of flows (c_f) in the 25-node network for $\lambda = 2.0$. Figure 3(b) plots the corresponding number of opened WIM stations as a function of w_{ij} and c_f. We only present results for $\lambda = 2.0$ since other deviation tolerance values exhibit similar trends, but the change is more pronounced for higher tolerances.

The costs gradually increase with larger station costs. The concave shape is due to the fact that increasing costs imply opening fewer stations to intercept the flows. Conversely, Figure 3(b) shows that the number of open stations is quite high when the station costs are low at \$10K, but the numbers show a steep decent for increasing station costs.

Table 2 displays the solution times in seconds for MRS, SM and PM. All instances are solved to optimality by all three solution techniques. When generating paths as an input for



(a) Total cost

(b) Number of WIM stations

Figure 3 (a) Objective function change and (b) number of WIM stations opened as a function of w_{ij} and c_f for $\lambda = 2.0$ in the 25-node network

the MRS, we used the Carlyle and Wood (2005) path enumeration algorithm for loopless paths rather than solve k shortest path problems successively, since the ranking of paths is not required in the problem. All path enumerations in the 25-node network for MRS take less than one second. We also tested the MRS version in which only location variables are binary and the integrality conditions on all remaining variables are relaxed, since it provides a tighter LP-relaxation. However, as reported in Marković, Ryzhov, and Schonfeld (2015), the model with fewer constraints and some binary flow-based variables provides shorter run times with respect to the version tested here. Regarding the PM, we only implemented the integer separation in the 25-node network. The performance comparison of alternative separation procedures is carried out in the experiments on larger networks.

One of the most important results that can be observed from this table is the exponential rate of increase in MRS solution times for increasing deviation tolerance values. The average runtimes when $\lambda = 1.0$ are 0.46, 0.58 and 1.17 for MRS, SM and PM, respectively. However when $\lambda = 2.0$, the solution times of MRS become prohibitive whereas the SM and PM can still solve the instances within an average of 3.41 and 1.96 seconds, respectively.

Table 2 Solution times with three models for the 25-node network

		MRS CPU time (s)			SM CPU time (s)			PM CPU time (s)		
w_e(\$)	c_f(\$)	$\lambda = 1.0$	$\lambda = 1.5$	$\lambda = 2.0$	$\lambda = 1.0$	$\lambda = 1.5$	$\lambda = 2.0$	$\lambda = 1.0$	$\lambda = 1.5$	$\lambda = 2.0$
10K	0.025	0.4	12.6	854.9	0.6	4.2	7.1	0.5	2.7	3.1
	0.05	0.3	7.6	392.4	0.6	1.7	2.7	1.5	1.7	1.4
	0.10	0.4	7.1	466.7	0.5	1.8	2.4	1.0	1.7	1.5
	0.20	0.3	6.6	531.8	0.5	1.3	3.2	0.9	1.6	1.3
60K	0.025	0.5	9.6	298.5	0.6	1.4	2.4	1.1	2.0	2.1
	0.05	0.5	10.9	308.9	0.6	2.9	4.1	1.3	1.7	2.2
	0.10	0.4	10.3	1066.5	0.6	2.3	6.0	0.8	1.3	2.1
	0.20	0.4	7.7	515.1	0.6	2.3	3.3	0.7	1.3	1.7
110K	0.025	0.4	13.7	373.1	0.6	1.7	5.3	1.7	1.9	1.7
	0.05	0.4	8.8	284.1	0.5	1.6	2.4	0.9	1.6	1.8
	0.10	0.5	11.6	332.1	0.6	3.3	5.7	1.4	1.3	1.8
	0.20	0.2	7.9	617.3	0.6	2.6	4.5	1.4	1.0	2.3
160K	0.025	0.5	11.8	285.8	0.6	1.6	2.6	1.7	2.0	1.9
	0.05	0.6	10.1	307.4	0.5	1.7	2.2	1.3	1.8	1.8
	0.10	0.4	11.8	370.7	0.6	1.4	3.1	1.3	1.8	2.0
	0.20	0.5	16.7	894.7	0.6	4.1	4.6	1.3	1.9	2.2
260K	0.025	0.5	12.8	303.1	0.5	1.3	1.8	1.0	1.7	2.0
	0.05	0.5	9.0	319.7	0.6	1.7	2.4	1.3	1.4	1.9
	0.10	0.8	6.9	235.3	0.5	2.0	2.4	1.2	1.2	2.2
	0.20	0.3	8.4	321.3	0.6	2.5	4.0	1.0	1.0	2.4
360K	0.025	0.8	12.9	315.4	0.5	1.3	1.9	1.0	1.4	2.1
	0.05	0.5	13.8	296.2	0.5	1.5	2.5	1.2	1.7	2.1
	0.10	0.4	11.8	239.8	0.5	1.6	2.8	1.3	1.0	2.0
	0.20	0.4	9.7	318.4	0.5	1.5	2.3	1.4	1.5	1.3
	Average	0.5	10.4	427.1	0.6	2.1	3.4	1.2	1.6	2.0

This shows that even though all paths can be generated and used as an input to MRS, the performance of MRS is still inferior to those of SM and PM.

4.3. The CA Network

We now present the results of the SM and PM models. Note that the MRS model does not scale up to the size of the CA network. This can be understood from Figure 4 which plots the number of paths for deviation tolerances of 1.00 to 1.08 of 0.01 increments in the CA network. For a deviation of 1.08, there exist almost 100 million paths within the deviation tolerance, which makes it impossible to solve the corresponding MRS model.

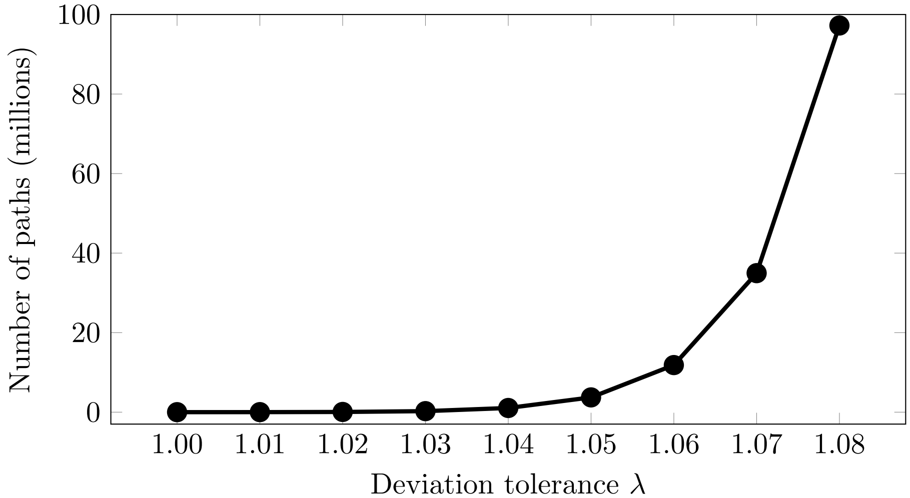


Figure 4 Number of paths in the CA network for different deviation tolerances λ

Figure 5 plots the total costs and number of open stations in the CA network against varying w_{ij}, c_f and λ values. These plots exhibit patterns similar to those of Figures 3(a) and 3(b) for the 25-node network.

An important result is the actual savings yielded by the location of WIM stations. In the CA network, when there are no stations and all overweight trucks drive on their shortest paths, the total annual cost is \$5.84 million for each cent of damage coefficient. In other words, if one kilometer of damaged road can be repaired by only spending one cent, then the cost is still excessive. When a realistic value of five to ten cents is used, then the total annual cost of not intercepting overweight trucks varies between \$29.2M and \$58.4M, which is three to six times higher than the largest cost appearing in Figure 5.

A final remark regarding the figure is related to the number of open stations in the optimal solutions. Observe that some of the curves are not convex. This is basically due to the indivisibility of the number of stations. The marginal contribution of the last station opened depends on the vehicle flow structure in the road network.

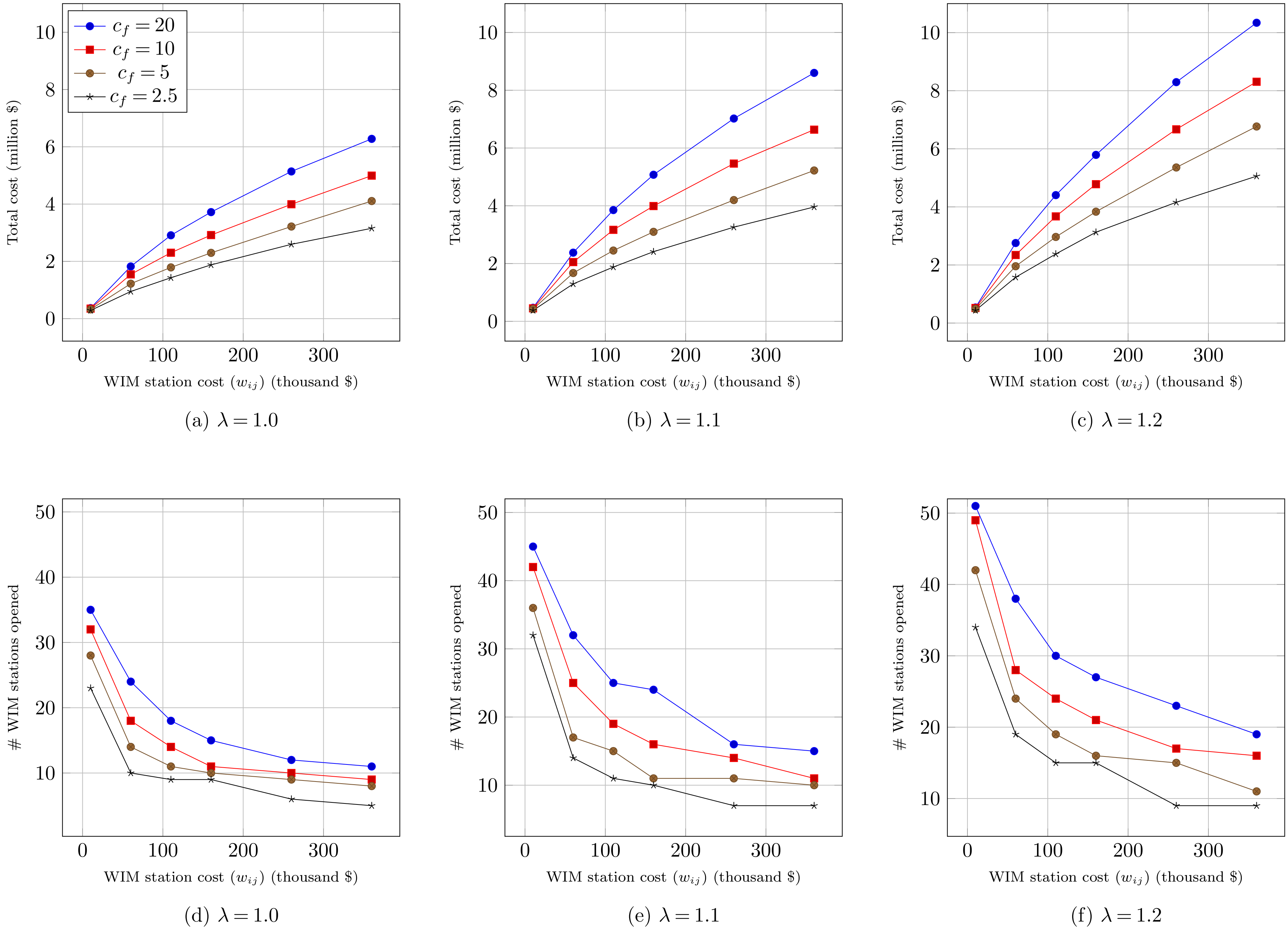


Figure 5 Objective function change (Figures 5(a), 5(b) and 5(c)) and number of WIM stations opened (Figures 5(d), 5(e) and 5(f)) for varying w_{ij}, c_f and λ values for the California network

Table 3 Computational results for the CA network

λ	Method	# opt.sol.	avg.gap (%)	CPU time(s)	nNodes	nIntCuts	nFracCuts
1.0	SM	24	0	72.0	1037.88	N/A	N/A
	PM-00	24	0	12.7	0	0	0
	PM-01	24	0	14.5	0	0	0
	PM-10	24	0	13.9	0	0	0
	PM-11	24	0	16.0	0	0	0
1.1	SM	1	23.13	3610.9	856.71	N/A	N/A
	PM-00	24	0	151.2	47.38	1436.13	0
	PM-01	24	0	158.0	45.79	1419.50	0
	PM-10	24	0	130.0	13.25	746.33	597.88
	PM-11	24	0	144.0	12.13	698.25	647.88
1.2	SM	0	84.53	3614.0	N/A	N/A	N/A
	PM-00	20	3.02	1886.7	366.21	5005.08	0
	PM-01	20	10.99	1996.2	336.75	5305.29	0
	PM-10	24	0	703.6	39.29	1625.75	2418.17
	PM-11	24	0	785.9	64.58	1613.54	2431.33

Computational results for the CA network are presented in Table 3. The first column provide the deviation tolerance (λ), the second column indicates which model is used. The two-digit in the PM version identifies the settings: if the first digit is 1, then the exact fractional separation is implemented; if the second digit is 1, then the feasibility heuristic is used. The exact fractional separation is only implemented at the root node, since this is already another NP-hard problem. The third column, '# opt.sol.' is the number of instances solved to optimality. Each line corresponds to an average of a total of 24 runs, corresponding to the combinations of c_f and w_i. The fourth column, 'avg.gap' displays the percent average gap. The CPU times in seconds are presented in the fifth column. Next, 'nNodes', average number of nodes in the branch-and-cut tree is presented. The last two columns, 'nIntCuts' and 'nFracCuts', are the number of cuts added by integer and fractional separation algorithms, respectively.

The results show that SM does not scale up well for large-size networks. All instances with $\lambda = 1.0$ are solved, but only one instance can be solved for the larger deviation tolerance values. Furthermore, the average number of nodes explored with SM is very large compared with the other methods. Note that the average number of nodes is not reported when $\lambda = 1.2$. This is due to SM not being able to solve the root node relaxation within the time limit. The feasibility heuristic provides several good upper bounds, but significantly increased the runtimes on average. This is mainly related to the inner dynamics of CPLEX. Our feasibility heuristic usually provided several feasible solutions before CPLEX did, but CPLEX's own heuristics for the generation of feasible solutions provided better solutions than ours.

Out of the four PM implementations, the exact fractional separation algorithm proved to be highly effective in providing better bounds. It performed the best among all implementations with an average solution time of 703.6 seconds and an average of 39.29 nodes explored. Without fractional separation, both the PM-00 and PM-01 implementations were unable to solve four of the 24 instances for $\lambda = 1.2$. Even though there is no significant reduction in the number of total cuts added in the implementations with and without fractional separation, the average number of nodes explored and the runtimes are significantly lower.

One final important observation concerning SM and PM regards the node domination in the transformed graphs. When we generate the graphs for each OD pair, we only consider non-dominated nodes as presented in the Section 3.1. This turns out to be highly effective when solving the models. In the CA network, there are in total of 1,440,078 dominated and undominated arc variables; however, the proportions of undominated nodes are only 1.25%, 8.5% and 16.1% for $\lambda = 1.0, 1.1$ and 1.2, respectively.

4.4. The Three 500-node Networks

We generated 500 random nodes on a 1000 $\times$ 1000 km^2 region and assigned a population number according to Pareto distribution with scale parameter of 5000 and shape parameter of 0.9. Each node is connected to other nodes in the network within 100 kilometers and the arc lengths are Euclidian distances. We thus obtained 23 disconnected node clusters. Randomly selecting a node from each cluster and connecting it to the cluster having the largest number of nodes, we ensured strong connectivity of the graph. We then selected 50, 75 and 100 most densely populated nodes as OD pair nodes. We refer to these networks as N500-50, N500-75 and N500-100. Using the shortest path distances and population

Table 4 Computational results for 500-node networks

Network	λ	Method	CPU time (s)	nNodes	nIntCuts	nFracCuts
N500-50	1.0	PM-00	9.8	0	0	0
		PM-01	11.6	0	0	0
		PM-10	10.6	0	0	0
		PM-11	11.7	0	0	0
	1.1	PM-00	107.9	7.04	460.00	0
		PM-01	128.4	8.42	448.08	0
		PM-10	82.5	0.75	306.21	91.00
		PM-11	86.1	0.25	299.21	96.92
	1.2	PM-00	215.4	11.79	527.63	0
		PM-01	234.5	13.58	549.04	0
		PM-10	177.2	5.00	396.88	92.83
		PM-11	180.1	3.79	397.71	121.83
N500-75	1.0	PM-00	17.8	0	0	0
		PM-01	20.2	0	0	0
		PM-10	18.6	0	0	0
		PM-11	21.8	0	0	0
	1.1	PM-00	231.1	4.63	897.79	0
		PM-01	304.6	8.50	1042.38	0
		PM-10	246.1	2.46	767.42	201.96
		PM-11	248.9	1.08	693.96	244.04
	1.2	PM-00	586.1	20.88	1494.13	0
		PM-01	593.7	16.79	1403.92	0
		PM-10	481.1	5.50	978.54	427.04
		PM-11	562.1	8.08	1025.88	429.96
N500-100	1.0	PM-00	34.5	0	0	0
		PM-01	38.8	0	0	0
		PM-10	35.7	0	0	0
		PM-11	42.2	0	0	0
	1.1	PM-00	965.3	25.04	2535.54	0
		PM-01	1155.9	27.88	2735.79	0
		PM-10	750.2	6.13	1029.17	1130.42
		PM-11	794.7	5.92	1078.67	1072.25
	1.2	PM-00	2241.1	39.22	3338.22	0
		PM-01	2299.2	35.00	3235.67	0
		PM-10	1515.7	6.25	1276.96	1452.25
		PM-11	1550.3	6.50	1365.83	1452.46

numbers, we generated vehicle flows between the nodes. We selected 500, 1000 and 2000 most dense vehicle flows for N500-50, N500-75 and N500-100, respectively. Further details of these graphs are presented in Table 1.

Table 4 reports the results of experiments for various w_{ij}, c_f and λ settings. All instances are solved to optimality. We observe very similar results to the CA network. Again, PM-10 performs the best among four implementations for all λ values. For all experiments with non-zero deviations from the shortest paths, it always outperforms the other implementations in terms of solution time.

5. Conclusions

We have modeled and solved the Evasive Flow Capturing Problem (EFCP) (Marković, Ryzhov, and Schonfeld 2015), which consists of optimally locating law enforcement facilities to intercept non-cooperative vehicle flows in a transportation network. The EFCP adds an important dimension to the flow capturing location problem literature by accounting for the fact that the unlawful vehicle drivers may change their routes in order not to be intercepted. The previous available models for the EFCP worked with an exhaustive pregeneration of all routes, but this soon becomes prohibitive when the instance size increases, especially when non-cooperative drivers are willing to significantly deviate from their shortest paths. One of the major challenges results from the presence of the conflicting aims of the law enforcement body and of the unlawful drivers, which we have tackled by developing a bilevel model. Another challenge stems from the transformation of the two-stage model into a single-stage model, which we have handled by transforming a constraint and by using duality theory. Furthermore, we have applied a polyhedral study on this compact single-stage model and we have projected out auxiliary variables to arrive at a formulation with an exponential number of constraints, associated with the paths in the network. We have proved that the projection inequality yields a facet-defining cut inequality. This final model was solved by branch-and-cut. The solutions were separated by solving resource constrained shortest path problem. When the solutions are integral, the separation problem further reduces to a shortest path problem. Extensive numerical experiments on networks containing up to 500 nodes and a deviation tolerance $\lambda = 1.2$ have confirmed the efficacy of our methodology. In particular, we have shown that solving our best model by branch-and-cut yielded a CPU time reduction of at least two orders of magnitude on the 25-node instances with respect to the best published algorithm. We have significantly increased the size of the instances that can be solved to optimality.

The Java code

A compressed file is attached as an online supplement that includes the Java code, data and a readme file.

Acknowledgments

The authors gratefully acknowledge funding provided by the Canadian Natural Sciences and Engineering Research Council under grants 436014-2013 and 2015-06189 and by the Fonds de recherche du Québec-Nature et technologies under grant NC-198837. The authors thank the associate editor and two referees whose comments helped improve the quality of this paper.

References

Ahuja RK, Magnanti TL, Orlin JB, 1993 *Network Flows: Theory, Algorithms, and Applications* (Prentice Hall, Upper Saddle River, New Jersey).

Amaldi E, Bruglieri M, Fortz B, 2011 *On the hazmat transport network design problem.* Pahl J, Reiners T, Voß S, eds., *Network Optimization*, 327–338 (Springer, Berlin Heidelberg).

Arslan O, Karaşan OE, Majhoub AR, Yaman H, 2017 *A branch-and-cut approach for the alternative fuel refueling station location problem with routing. Manuscript under review.*

Arslan O, Karaşan OE, 2016 *A Benders decomposition approach for the charging station location problem with plug-in hybrid electric vehicles. Transportation Research Part B* 93, Part A:670–695.

Arslan O, Yıldız B, Karaşan OE, 2014 *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(0):168–178.

Averbakh I, Berman O, 1996 *Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. European Journal of Operational Research* 91(3):495–506.

Berge C, Minieka E, 1973 *Graphs and Hypergraphs* (North-Holland, Amsterdam).

Berman O, Larson RC, Fouska N, 1992 *Optimal location of discretionary service facilities. Transportation Science* 26(3):201–211.

Borndörfer R, Omont B, Sagnol G, Swarat E, 2012 *A Stackelberg game to optimize the distribution of controls in transportation networks*, 224–235 (Springer, Berlin Heidelberg).

Brotcorne L, Labbé M, Marcotte P, Savard G, 2001 *A bilevel model for toll optimization on a multicommodity transportation network. Transportation Science* 35(4):345–358.

Brotcorne L, Labbé M, Marcotte P, Savard G, 2008 *Joint design and pricing on a network. Operations Research* 56(5):1104–1115.

Cao D, Chen M, 2006 *Capacitated plant selection in a decentralized manufacturing environment: a bilevel optimization approach. European Journal of Operational Research* 169(1):97–110.

Capar I, Kuby M, Leon VJ, Tsai YJ, 2013 *An arc-cover path-cover formulation and strategic analysis of alternative-fuel station locations. European Journal of Operational Research* 227(1):142–151.

Carlyle MW, Wood KR, 2005 *Near-shortest and k-shortest simple paths. Networks* 46(2):98–109.

Christof T, Löbel A, Stoer M, 1997 *Porta-polyhedron representation transformation algorithm. Software package, available at http://www.zib.de/Optimization/Software/Porta* .

Conforti M, Cornuéjols G, Zambelli G, 2010 *Polyhedral approaches to mixed integer linear programming.* Jünger M, Liebling TM, Naddef D, Nemhauser GL, Pulleyblank WR, Reinelt G, Rinaldi G, Wolsey LA, eds., *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*, 343–385 (Springer, Berlin Heidelberg).

Correa J, Harks T, Kreuzen VJ, Matuschke J, 2017 *Fare evasion in transit networks. Operations Research* 65(1):165–183.

Cottrell J BH, 1992 *The avoidance of weigh stations in Virginia by overweight trucks.* Technical report, Virginia Transportation Research Council.

Cunagin W, Mickler W, Wright C, 1997 *Evasion of weight-enforcement stations by trucks. Transportation Research Record: Journal of the Transportation Research Board* 1570:181–190.

Fontaine P, Minner S, 2014 *Benders decomposition for discrete–continuous linear bilevel problems with application to traffic network design. Transportation Research Part B* 70:163–172.

Gendreau M, Laporte G, Parent I, 2000 *Heuristics for the location of inspection stations on a network. Naval Research Logistics* 47(4):287–303.

Hodgson MJ, 1981 *The location of public facilities intermediate to the journey to work. European Journal of Operational Research* 6(2):199–204.

Hodgson MJ, 1990 *A flow-capturing location-allocation model. Geographical Analysis* 22(3):270–279.

Hodgson MJ, Rosing KE, Zhang J, 1996 *Locating vehicle inspection stations to protect a transportation network. Geographical Analysis* 28(4):299–314.

Kara BY, Verter V, 2004 *Designing a road network for hazardous materials transportation. Transportation Science* 38(2):188–196.

Kim JG, Kuby M, 2012 *The deviation-flow refueling location model for optimizing a network of refueling stations. International Journal of Hydrogen Energy* 37(6):5406–5420.

Kim JG, Kuby M, 2013 *A network transformation heuristic approach for the deviation flow refueling location model. Computers & Operations Research* 40(4):1122–1131.

Kuby M, Lim S, 2005 *The flow-refueling location problem for alternative-fuel vehicles. Socio-Economic Planning Sciences* 39(2):125–145.

Kuiteing AK, Marcotte P, Savard G, 2017 *Network pricing of congestion-free networks: The elastic and linear demand case. Transportation Science* 51(3):791–806.

Labbé M, Marcotte P, Savard G, 1998 *A bilevel model of taxation and its application to optimal highway pricing. Management science* 44(12-part-1):1608–1622.

Lim S, Kuby M, 2010 *Heuristic algorithms for siting alternative-fuel stations using the flow-refueling location model. European Journal of Operational Research* 204(1):51–61.

Lozano L, Medaglia AL, 2013 *On an exact method for the constrained shortest path problem. Computers & Operations Research* 40(1):378–384.

Lu CC, Yan S, Ko HC, Chen HJ, 2017 *A bilevel model with a solution algorithm for locating weigh-in-motion stations. IEEE Transactions on Intelligent Transportation Systems* PP(99):1–10.

Marković N, Ryzhov IO, Schonfeld P, 2015 *Evasive flow capture: Optimal location of weigh-in-motion systems, tollbooths, and security checkpoints. Networks* 65(1):22–42.

Marković N, Ryzhov IO, Schonfeld P, 2017 *Evasive flow capture: A multi-period stochastic facility location problem with independent demand. European Journal of Operational Research* 257(2):687–703.

McCabe TJ, 1976 *A complexity measure. IEEE Transactions on Software Engineering* (4):308–320.

Mirchandani PB, Rebello R, Agnetis A, 1995 *The inspection station location problem in hazardous material transportation: some heuristics and bounds. INFOR: Information Systems and Operational Research* 33(2):100–113.

MirHassani SA, Ebrazi R, 2013 *A flexible reformulation of the refueling station location problem. Transportation Science* 47(4):617–628.

Sichi J, Contributors, 2017 *JGraphT a free Java graph library.* Jgrapht.org.

Yang H, Zhou J, 1998 *Optimal traffic counting locations for origin-destination matrix estimation. Transportation Research Part B: Methodological* 32(2):109–126.

Yıldız B, Arslan O, Karaşan OE, 2016 *A branch and price approach for routing and refueling station location model. European Journal of Operational Research* 248(3):815–826.

Okan Arslan is a postdoctoral research fellow at HEC Montréal. He received his Ph.D. in industrial engineering at Bilkent University. His research is mainly centered around investigating the impacts of new technologies in transportation and logistics for businesses, households, and governments.

Ola Jabali is an assistant professor in the operations research and discrete optimization group of DEIB (Dipartimento di Elettronica, Informazione e Bioingegneria) at the Politecnico di Milano. She received her Ph.D. in Industrial Engineering from the group of Operations, Planning, Accounting and Control at Eindhoven University of Technology. She was a faculty member of the Department of Logistics and Operations Management at HEC Montréal. Her research focuses on optimizing goods transportation, in particular problems related to pollution, uncertainty, fleet composition and customer service.

Gilbert Laporte is a professor of operations research at HEC Montréal and Canada Research Chair in Distribution Management. He is the former editor-in-chief of Transportation Science and Computers & Operations Research. He is a member of the Royal Society of Canada, from which he received the Innis-Gérin Medal. In 2009 he received the Robert M. Herman Lifetime Achievement Award in Transportation Science from the Transportation Science and Logistics Society of INFORMS, and in 2014 he was awarded the INFORMS Lifetime Achievement in Location Analysis.