ABSTRACT

Title of dissertation: EVASIVE FLOW CAPTURE

Nikola Marković, Doctor of Philosophy, 2013

Dissertation directed by: Professor Paul Schonfeld
Department of Civil and
Environmental Engineering

The flow-capturing location-allocation problem (FCLAP) consists of locating facilities in order to maximize the number of flow-based customers that encounter at least one of these facilities along their predetermined travel paths. In FCLAP, it is assumed that if a facility is located along (or "close enough" to) a predetermined path of a flow, the flow of customers is considered captured. However, existing models for FCLAP do not consider the likelihood that targeted users may exhibit non-cooperative behavior by changing their travel paths to avoid fixed facilities. Examples of facilities that targeted subjects may have an incentive to avoid include weigh-in-motion stations used to detect and fine overweight trucks, tollbooths, and security and safety checkpoints. The location of these facilities cannot be adequately determined with the existing flow-capturing models.

This dissertation contributes to the literature on facility location by introducing a new type of flow capturing framework, called the "Evasive Flow Capturing Problem" (EFCP), in which targeted flows exhibit non-cooperative behavior by trying to avoid the facilities. The EFCP proposed herein generalizes the FCLAP and has relevant applications in transportation, revenue management, and security and safety management. This work formulates several variants of EFCP. In particular, three optimization models, deterministic, two-stage stochastic, and multi-stage stochastic, are developed to allocate facilities given different availability of information and planning policies. Several properties are proved and exploited to make the models computationally tractable. These results are crucial for solving optimally the instances of EFCP that include real-world road networks, which is demonstrated on case studies of Nevada and Vermont.

EVASIVE FLOW CAPTURE

by

Nikola Marković

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Paul Schonfeld, Chair/Advisor
Professor Ilya O. Ryzhov
Professor Bruce Golden
Professor Elise Miller-Hooks
Professor Cinzia Cirillo

Acknowledgments

First and foremost, I am grateful to my advisor, Professor Paul Schonfled, for his guidance and support during my graduate studies at University of Maryland. I was fortunate to have an advisor with an extensive knowledge in transportation engineering and great enthusiasm to help his students, both professionally and personally. It has been an honor and pleasure to work under his mentorship.

I would like to thank Professors Bruce Golden, Cinzia Cirillo, and Elise Miller-Hooks for participating in my dissertation committee and for contributing with helpful suggestions. A special thank is due to Professor Ilya O. Ryzhov for numerous discussions regarding this thesis and for encouraging me to dive deeper into the world of operations research. His help and advice are much appreciated.

I am grateful to Professor Dušan Teodorović at University of Belgrade, Serbia, who encouraged me to pursue my graduate studies in US and ensured that I obtain a research assistantship. Moreover, I thank Željko Drobnjak, Nikola Bešinović, Rahul Nair, and Sanjin Milinković, for fruitful collaboration on various projects during my studies. A big thanks is also due to Stevan Turajlić for inspiring me to pursue an engineering degree in the first place.

Finally, I would like to thank my beloved Jelena for her unconditional love and support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Motivation and Contributions

Facility location is a fundamental problem in operations research, due to its importance in strategic planning and efficient resource allocation. Traditional network facility location models assume that customers are concentrated at nodes of a transportation network and travel to nearby facilities to obtain services. One such model is the maximal covering location problem, which locates a given number of facilities in order to maximize total node-based demand within a specified radius from at least one facility [2].

An important generalization of the maximal covering location problem is the flow-capturing location-allocation problem (FCLAP), in which demand is defined in terms of flows of customers traveling between their origin and destination nodes [3]. The objective of the FCLAP is to locate a given number of facilities in order to maximize the number of flow-based customers who encounter at least one facility on their preplanned travel paths. FCLAP was independently introduced in [4] and [5], and has been extensively studied within operations research, various areas of engineering, economics, and geography. Some of the applications of the original FCLAP and its variants included the optimal location of bank ATMs [5], vehicle inspection stations [6, 7], traffic counting points [8], rail park-and-ride facilities [9], and alternative-fuel stations [10, 11].

A common theme for FCLAP is the assumption that, if a facility is located along (or "close enough" to) a predetermined path of a flow of customers, then that flow is considered captured. The literature on flow capturing does acknowledge that implementation of certain fixed facilities could encourage the targeted users to avoid them by changing their travel paths. For example, Mirchandani et al. [12] argue that truckers transporting hazardous materials may find out or guess the locations of inspection stations and try to avoid them by changing their routes. However, existing models for FCLAP do not consider the possibility that targeted users would exhibit such non-cooperative behavior. As a result, the existing models for FCLAP cannot find adequate locations of flow-capturing facilities, which the targeted flows may wish to avoid.

This work addresses the problem of locating facilities that targeted flows may have an incentive to evade by changing their travel paths. Examples of such facilities include the weigh-in-motion stations that are used to detect and fine overweight trucks, tollbooths, and security and safety checkpoints. This dissertation introduces a new type of flow-capturing problem, called the "Evasive Flow Capturing Problem" (EFCP), which generalizes FCLAP by assuming that a flow can travel along multiple paths as long as the detour is not too large, and that a targeted flow chooses to travel along the shortest path not covered by a facility. The dissertation presents three models and a realistic case study, whose main findings and contributions are summarized below.

Chapter 2 documents work on the deterministic EFCP [13], which makes the following contributions:

1. It introduces and mathematically formulates the EFCP, which has broad applications in transportation, revenue management, security and safety management. It shows that the EFCP generalizes the FCLAP, and it establishes relations between the two problems and their optimal solutions. One consequence of non-cooperative behavior is that any solution always incurs higher (or equal) costs under the EFCP objective than the FCLAP objective.

2. It studies the mathematical properties of the EFCP (e.g. submodularity and computational complexity) and concludes that this problem is structurally different from FCLAP. Specifically, existing performance guarantees on the performance of a greedy heuristic for FCLAP do not hold for EFCP. In fact, the greedy approximation for EFCP can perform arbitrarily poorly. However, a partial linear relaxation will always yield an optimal solution at a reduced computational cost.

3. It presents numerical examples including real-world transportation networks. These case studies are used to show the applicability of the proposed flow-capturing framework to realistic problem instances. In addition, the real-world road networks are used to numerically contrast EFCP and FCLAP. This comparison demonstrates that solutions optimal for FCLAP do poorly when targeted subjects try to avoid the facilities, thus showing that EFCP adds considerable value.

Chapter 3 presents work on the two-stage stochastic EFCP [13], which makes the following contributions:

1. It introduces a stochastic extension of EFCP that accounts for flows whose intensities and willingness to avoid facilities are not known with certainty. This extension adds considerably to the applicability of the EFCP methodology since, in the real-world applications, intensities of flows and their willingness to avoid facilities could be estimated through data collection or expert opinion. Both approaches yield scenarios which could be used as inputs for the two-stage stochastic program proposed herein.

2. The structural properties of the stochastic EFCP are studied and exploited in order to make the problem computationally tractable. First, it is shown that, under certain independence assumptions, random intensities of flows can be replaced with their expected values without affecting the solution. Then, the second-stage is reformulated recursively. These two results significantly reduce the number of variables and constraints, making the two-stage stochastic EFCP only slightly more difficult to solve in the extensive form than the deterministic EFCP. In addition, it is argued that many scenario-dependent binary variables can be linearly relaxed, which further reduces the solution times.

3. It presents numerical experiments on actual transportation networks of two US states. Numerical tests show that exploiting the specific structure of the problem is crucial for efficiently solving the real-world-size stochastic EFCP. To emphasize this point, it is shown that a standard solution approach, the integer

L-shaped method, fails to find the optimal solution in a reasonable amount of time. In addition, we numerically compute the value of the stochastic solution, i.e. the benefit from solving the stochastic EFCP over solving deterministic EFCP in which all random parameters are replaced with their expected values. Numerical tests show significant value of the stochastic solution, which implies the relevance of the proposed stochastic EFCP.

Chapter 4 presents the work on the multi-stage stochastic EFCP, which makes the following contributions:

1. It introduces a multi-stage extension of the stochastic EFCP in which decisions about the implementation of facilities are made at different time points (e.g. annually or every few years) given probabilistic information about the flows and their willingness to avoid facilities. This resembles a realistic long-term investment planning problem and is particularly suitable for the case when intensities of the flows change over time (e.g. the expected number of heavy trucks increases by 2% annually). Thus, it considerably adds to the applicability of the EFCP methodology.

2. It further exploits the structural properties of the two-stage stochastic model to make the multi-stage stochastic EFCP tractable. It proposes an exact mathematical programming formulation which can be solved optimally even for real-world transportation networks. In addition, it develops a dynamic programming formulation of the problem and proposes an approximate dynamic programming approach which statistically estimates the downstream values of

the objective function.

3. It presents numerical examples including the real-world road networks and solves them optimally. The considered case studies assume a 30-year planning horizon during which the decisions about the implementation of facilities are made every 2 or 3 years. These case studies indicate that the proposed multi-stage stochastic EFCP is applicable in designing realistic long-term investment plans.

Chapter 5 applies the proposed EFCP methodology to a real-world case study and makes the following contributions:

1. It presents a case study including the allocation of weigh-in-motion sttions in the road network of Nevada. The case study includes the actual road network designated for large commercial vehicles, truck flows simulated based on data available in the literature, and realistically estimated damage produced by overweight trucks. This analysis demonstrates applicability of the proposed models and solution techniques to a real-world problem.

2. It contrasts the proposed EFCP with the existing FCLAP in order to estimate the value that the EFCP framework adds in allocating facilities which targeted truck flows wish to avoid. The numerical comparison indicates that results optimal for FCLAP perform poorly in the setting where targeted flows try to evade the facilities. Moreover, the FCLAP-based allocations often incur greater damage than no weigh-in-motion implementation. The reason for this is the so called weigh-in-motion paradox which is discussed in greater detail.

3. It considers the real-world implementation of static weigh scales in Nevada and explores whether the current allocation could be improved through application of the EFCP. The conducted analysis implies that the current allocation can be significantly improved given the available information about the flows in Nevada. This comparison suggests that the proposed EFCP can serve as a useful decision support tool with the potential to improve solutions based on human judgment and intuition.

1.2 Background on Flow Capture

Many flow-capturing problems were proposed since FCLAP was first introduced. The characteristics of various flow-capturing problems found in the literature are summarized below. Different aspects of these problems include:

1. Deviations from preplanned trips where a flow is considered captured not only if a facility is located along the predetermined path of a flow, but also in its relative proximity [14, 15]. This extension was considered in the context of locating gas stations and restaurants and it was approached by preparing inputs for FCLAP differently (i.e. by enlarging the set of potential facility locations from which a flow could be intercepted).

2. Limited capacity of the facilities [16, 17], as well as decisions about the size of facilities [18].

3. Temporal aspects such as time spent in a facility [16], determining service

start times [3], and multi-period planning where decisions about the facility locations are made over several years [19].

4. Multiple counting of consumers in which the level of consumption depends on the number of facilities (e.g. billboards) that customers encounter [20], and consumers' preference for obtaining a service at the beginning, middle, or end of their trips [21].

5. Probabilistic information about the travel origins, turning movements to visit facilities, and customer arrival and service rates [16, 22–24].

6. Competition between the facilities that may be within the same or different chain [25, 26].

7. Synthesis with demand coverage, where flow capture (e.g. intercepting customers along their trips) is addressed jointly with covering fixed customers residing at nodes. [27, 28].

The introduction of FCLAP and its variants also initiated work seeking more efficient problem formulations [15, 29, 30], as well as developing exact and approximate solution techniques [7, 31, 32] for efficiently solving realistic problem instances. Probably the main reason for such fruitful research on FCLAPs is the applicability of this class of facility location problems to various areas of human endeavor. More information about such applications can be found in [33] and a review of over 30 different FCLAPs is provided in [29].

All the previous studies found about flow-capturing problems assume that if a facility is located along (or "close enough" to) a predetermined path of a flow, the flow of customers is considered captured. This assumption raises a serious issue in applications where targeted flows have an incentive to avoid the facilities. For example, consider the placement of weigh-in-motion systems, tollbooths, or security and safety checkpoints. The EFCP model, introduced in this dissertation, generalizes FCLAP by assuming that a flow can travel along multiple paths and that a targeted flow chooses to travel along the shortest path not covered by a facility, as long as the detour is not too large. Optimal solutions of this problem behave very differently from those of FCLAP.

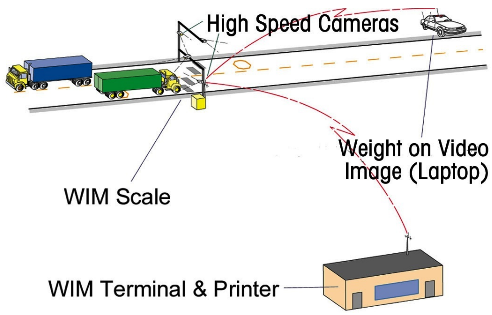
Like previously described FCLAPs, EFCPs encompass many variants that may include different objectives (e.g. cost minimization in WIM allocation, profit maximization in tollbooth allocation, or risk minimization in locating safety and security checkpoints), constraints (depending on the application), temporal aspects (single-stage vs. multi-stage location of facilities), and treatment of information (deterministic vs. probabilistic inputs). Thus, different variants of EFCP require different modeling features and solution techniques due to potentially different structural properties of the problem. This dissertation proposes three distinct variants of EFCP: deterministic, two-stage stochastic, and multi-stage stochastic. The three models can be used in different applications, whose relevance is explained in the following sections.

1.3 Weigh-in-Motion Allocation

Truckers have an incentive to overload their vehicles because it increases their productivity and thus their profits. However, these "extra profits" for the truckers come at the expense of severe pavement and environmental damages, whose costs are passed to the taxpayers. In particular, the taxpayers have to pay hundreds of millions of dollars annually for the damages that are due to overweight trucks. For example, only the pavement damage attributed to these trucks in California was roughly estimated at $23 to $35 million per year (adjusted for inflation from [34]). When extrapolated to the entire US, this damage exceeds $200 millions/year. However, the total damage is much higher because it also includes external costs associated with the extra loads, such as emissions, noise, and accidents [35].

An efficient way of reducing this damage is to implement weigh-in-motion (WIM) systems that are designed to detect overweight trucks (Figure 1.1). As a truck drives over a WIM scale, the category of truck, axle weights, velocity, and other data are recorded and stored by the WIM system. The information gathered by a WIM system can be associated with the truck license plate and registration number through the use of high speed cameras. These data can then be transmitted to the weight-enforcing authorities and trucks violating weight restrictions can be cited [34]. Note that the WIM stations are uncapacitated and collect data at all time, which makes them much more efficient than static weigh stations that may have limited hours of operations and where considerable queuing delays may occur.

WIM technology is expensive and hence it cannot be implemented on every



(a) Concept of WIM [36]

(b) WIM Implementation [37]

Figure 1.1: WIM systems: Real-time image data are monitored on a computer in a fixed facility or a vehicle. When a suspect truck is identified, an enforcement unit can intercept and weigh the truck to confirm the violation.

road link. Recent implementations of WIM checkpoints reveal that their location in a road network is determined by prioritizing the most damaged road links. Such an approach was used in Montana, where officials reported an estimated reduction of annual pavement damage by $700,000 [38]. This intuitive approach towards allocating WIM systems can be improved by developing operations research models that optimize the number and location of WIM checkpoints. Several such models are found in the literature [39,40], but they are built on the assumption that trucks travel along the shortest paths from their origins to their destinations and that locating WIM checkpoints along the trucks' shortest paths suffices to enforce weight control. However, this simplifying assumption misrepresents the real world, where truck drivers quickly learn the location of checkpoints, communicate with other truckers, and start avoiding the checkpoints by taking detours (see [41] for a discussion of the empirical evidence). If this fact is ignored in allocating WIM checkpoints, then the

(a) Before WIM (b) WIM Implementation (c) After WIM

Figure 1.2: WIM Paradox: If there is a reasonably short detour, trucks traveling from A to B will bypass the WIM checkpoint and produce greater damage due to the longer distance traveled.

implementation of WIM technology can potentially result in greater damage due to additional vehicle-miles traveled. This phenomenon is called WIM paradox [42] and an example is shown in Figure 1.2.

In allocating WIM checkpoints there is a tradeoff between investing in WIM technology and the excessive damage that overweight trucks produce (e.g. damage associated with loads that exceed legal limits). Therefore, among the important inputs in optimizing the allocation of WIM systems are the estimated origins, destinations, and intensities of truck flows. As in most FCLAPs, it is assumed that this information about the flows is known (e.g. can be reliably estimated). In addition, six assumptions are outlined below in order to clarify relations incorporated in the mathematical formulations, which are presented the following chapters.

1. An agency allocating WIM checkpoints wishes to minimize total cost that includes investment in WIM systems and excessive damage due to overweight

trucks.

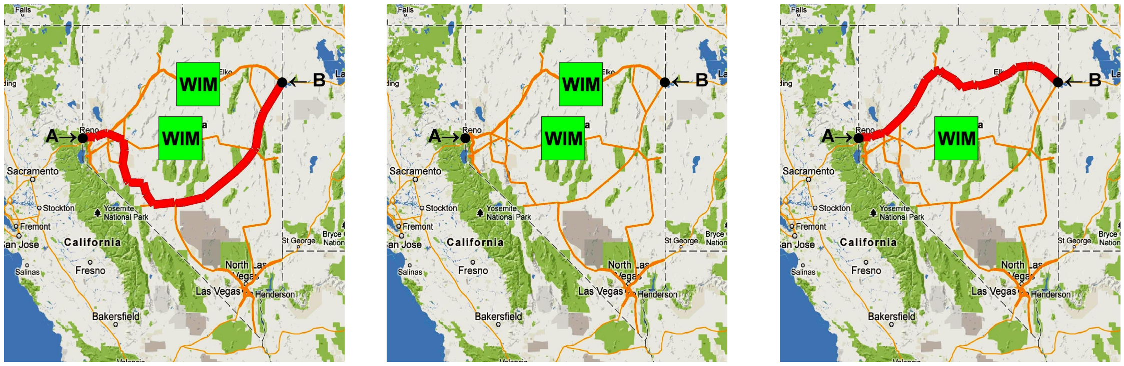
2. WIM checkpoints are located on road links. The cost of implementing checkpoints depends on the number of lanes.

3. The damage produced by a truck flow (i.e. a group of trucks with the same origin and destination) increases linearly with the distance traveled. This is clearly the case for pavement and environmental damage. (This should not be confused with the nonlinear relation between the weight of a vehicle and the per mile damage it produces.)

4. A truck flow can travel along k_f shortest paths from its origin to destination. The number k_f can be determined so that the $(k_f + 1)$-th shortest path would represent an excessive detour for truckers (i.e. that the cost of taking such a long detour would exceed the benefit from overloading the truck). For example, k_f can be determined so that the $(k_f + 1)$-th shortest path is 30% longer than the shortest path. (Figure 1.3 provides an example of an excessive detour.)

5. A truck flow is considered captured if at least one WIM checkpoint is located along each of the k_f paths. There is no excessive damage associated with captured flows.

6. An uncaptured flow travels along the shortest of its k_f paths that have not been covered by checkpoints because that minimizes the truckers' cost (see Figure 1.3).



(a) Excessive Detour (b) Captured Flow (c) Uncaptured Flow

Figure 1.3: Flow from A to B: Example of an excessive detour and WIM allocations that do (not) capture this flow.

1.4 Allocation of Vehicle Inspection Stations

About 500,000 shipments containing hazardous materials are made in the US every day [43]. The vast majority of these shipments are moved by trucks, whereas about 50% of all shipments include corrosive or flammable petroleum products. Since accidents that occur in transportation of hazardous materials may result in death, severe injuries, and destruction of environment or property, the transportation of hazardous materials is heavily regulated. The principal federal law regulating the transportation of hazardous materials is the Hazardous Materials Transportation Act (HMTA), which was introduced in 1975 and further enhanced through several major amendments in the 90s. The regulations include four aspects: 1) Procedures and Policies, 2) Material Designations and Labeling, 3) Packaging Requirements, and 4) Operational Rules. Violations of these rules may result in civil or criminal penalties.

To enforce the rules defined by HMTA, the regulating agencies need to determine where to inspect the trucks in the underlying transportation network [12]. The problem of locating inspections stations for trucks transporting hazardous materials is studied by Mirchandani et al. [12] who seek to locate these stations in order to maximize the number of inspected trucks. The authors state that their model is not applicable when there is game-playing behavior between the truckers and inspectors. They argue that truckers may know or guess where the inspectors are located and modify their routes to avoid inspection. Thus, their model is applicable when the truckers are required to take given routes or when the inspectors are mobile and can relocate accordingly.

The EFCP framework proposed in this dissertation is applicable to the location of safety checkpoints and it fills the gap discussed by Mirhcandani et al. [12]. Namely, the EFCP accounts for the non-cooperative behavior which one would expect from those truckers who violate regulations for the transportation of hazardous materials. The EFCP for this particular application is based on the assumptions similar to those in the allocation of WIMs. The main difference will arise in estimating parameters and defining the objective which would consist of minimizing risk.

1.5 Allocation of Tollbooths

The toll roads are usually designed with tollbooths located at each entry/exit point, which prevents drivers from avoiding them. However, tollbooths are expensive (in their delays to users, operating costs and capital costs) and their use is thereby restricted to limited-access freeways and main bottlenecks in road networks (such as major bridges or tunnels). Transportation economists and planners [44] argue convincingly that most congested road networks, including urban street networks, could be operated far more efficiently and beneficially if appropriate congestion prices could be charged without incurring excessive collection costs or user delays. For many potential applications where road pricing may be desirable, the density of conventional tollbooths needed to prevent evasion would be quite unaffordable, i.e. the delays and other costs would greatly exceed the revenues and other benefits. Instead of conventional tollbooths we envision much cheaper and less obtrusive tolling systems that detect vehicles, charge appropriate tolls and duly inform the vehicle operators about those tolls. (Please note that motorists should know the locations of those tolling locations, both for ethical reasons and because prices should be known to motorists in order to appropriately influence their travel decisions.) The proposed EFCP could be used to locate such "virtual tollbooths to maximize system-wide net benefits while considering user routing behavior and various benefits and costs, including the costs of the virtual tollbooths.

1.6 Dissertation Outline

Chapter 2 presents the deterministic EFCP which assumes that information about the flows (i.e. origins/destinations, intensities) and their willingness to avoid facilities can be estimated with certainty. Chapter 3 extends the deterministic model into a two-stage stochastic program, which assumes probabilistic information about the intensities of flows and their willingness to avoid facilities. Chapter 4 formulates the multi-stage stochastic EFCP in which decisions about the facility locations are made over multiple time periods (e.g. years), given the probabilistic information about the intensities of flows and their willingness to avoid facilities. Chapter 5 applies the deterministic EFCP to a realistic case study including the optimal location of WIM systems in Nevada, and contrasts the EFCP-based solution with the actual implementation of weigh stations. Chapter 6 summarizes contributions, emphasizes the potential benefits of this work to society, and discusses further extensions.

Chapter 2: Deterministic EFCP

Deterministic EFCP that assumes perfect information about the origins, destinations, and intensities of flows is introduced herein. This chapter is organized in six sections. First, a non-linear and a linear formulation of the problem are provided. Second, the relation between EFCP and FCLAP is established. Third, the structural properties of EFCP are analyzed and contrasted with those of FCLAP. Fourth, the exact and approximate solution methods are proposed. Fifth, the deterministic EFCP is tested on case studies involving real-world road networks of Nevada and Vermont. Finally, the conclusions are drawn.

2.1 Problem Formulation

Let $G(N, A)$ be a bidirectional road transportation network, where N is a set of nodes and A is a set of arcs (i, j). Denote by F a set of flows and define P_f as a set of paths which contains k_f shortest paths of the flow $f \in F$. Let A_f^p be the set of arcs along path $p \in P_f$ of flow $f \in F$. Additionally, let w_{ij} denote the cost of implementing and maintaining a facility at arc (i, j), and let c_f^p be the excessive damage cost (or risk) incurred if flow $f \in F$ passes unintercepted along path $p \in P_f$. Let x_{ij} be a binary variable equal to 1 if a facility is located at arc (i, j)

and 0 otherwise. Moreover, define $\boldsymbol{x} = \{x_{ij} \mid (i,j) \in A\}$ and $\boldsymbol{w} = \{w_{ij} \mid (i,j) \in A\}$ as vectors of $|A|$ elements.

The deterministic EFCP can now be formulated as minimization problem

$$\mathbf{P1}: \quad \min_{\boldsymbol{x} \in \{0,1\}^{|A|}} \quad \boldsymbol{w}^T \boldsymbol{x} + Q(\boldsymbol{x}),$$

where $Q(\boldsymbol{x})$ is an oracle that, given an allocation of checkpoints $\boldsymbol{x}$, computes the cost of excessive damage (or risk) associated with flows. If a flow is captured, then the corresponding damage is 0. Otherwise, the flow seeks to minimize its travel distance, and produces the damage by traveling along its shortest unmonitored path. More formally, if we let $P_f^2 = \left\{p \in P_f \mid \sum_{x_{ij} \in A_f^p} x_{ij} = 0\right\}$ be the set of paths of flow $f \in F$ not covered by facilities, then $Q(\boldsymbol{x}) = \sum_{f \in F} Q_f(\boldsymbol{x})$, where

$$Q_f(\boldsymbol{x}) = \begin{cases} \min\limits_{p \in P_f^2} \left\{c_f^p\right\}, & P_f^2 \neq \emptyset; \\ 0, & P_f^2 = \emptyset. \end{cases}$$

The nonlinear Problem **P1** can be linearized by introducing three sets of auxiliary binary variables. These variables are used to check whether a flow is captured, and direct the uncaptured flows along their shortest unmonitored paths.

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

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

$$z_f^p = \begin{cases} 1, & \text{if flow } f \in F \text{ travels unintercepted along path } p \in P_f \\ 0, & \text{otherwise} \end{cases}$$

The EFCP can now be formulated as a linear binary integer program:

$$\mathbf{P2}: \quad \min_{x_{ij}, y_f^p, y_f, z_f^p \in \{0,1\}} \sum_{(i,j)\in A} x_{ij} w_{ij} + \sum_{f\in F} \sum_{p\in P_f} z_f^p c_f^p \tag{2.1}$$

$$\text{s.t.} \sum_{(i,j)\in A_f^p} x_{ij} \geq y_f^p \qquad \forall p \in P_f \quad \forall f \in F \tag{2.2}$$

$$z_f^p \leq 1 - y_f^p \qquad \forall p \in P_f \quad \forall f \in F \tag{2.3}$$

$$\sum_{(i,j)\in A_f^p} x_{ij} \leq \left|A_f^p\right| \cdot y_f^p \qquad \forall p \in P_f \quad \forall f \in F \tag{2.4}$$

$$y_f \leq y_f^p \qquad \forall p \in P_f \quad \forall f \in F \tag{2.5}$$

$$\sum_{p\in P_f} z_f^p \geq 1 - y_f \qquad \forall f \in F \tag{2.6}$$

The objective (2.1) minimizes the investment cost and excessive damage (or risk) due to flows whose paths are not all covered by at least one checkpoint. Constraints (2.2)-(2.4) ensure that if at least one facility is allocated along a path of a flow ($y_f^p = 1$), the flow cannot pass unintercepted along that path ($z_f^p = 0$). Constraints (2.5) tie the variables guaranteeing that y_f can take a value of 1, if all the corresponding paths are covered by at least one facility. Constraints (2.6) require the unintercepted flows to count towards the objective function by producing the excessive damage along the shortest unmonitored path.

The above linearization includes three sets of auxiliary binary variables and five additional sets of constraints. The following result demonstrates that the two formulations are indeed equivalent. The full proof is presented below; the technique, which is based on the separability of the second-stage objective, will also be used in later proofs.

Proposition 1. *Problems **P1** and **P2** are equivalent.*

Proof. To prove this we need to show that $Q(\boldsymbol{x}) = \bar{Q}(\boldsymbol{x})$, where

$$\bar{Q}(\boldsymbol{x}) = \min_{y_f^p, y_f, z_f^p \in \{0,1\}} \left\{ \sum_{f \in F} \sum_{p \in P_f} z_f^p c_f^p \quad \text{s.t. constraints (2)-(6)} \right\}$$

Note that the above summation is separable in f, whence

$$\bar{Q}(\boldsymbol{x}) = \sum_{f \in F} \bar{Q}_f(\boldsymbol{x})$$

where

$$\bar{Q}_f(\boldsymbol{x}) = \min_{y_f^p, y_f, z_f^p \in \{0,1\}} \left\{ \sum_{p \in P_f} z_f^p c_f^p \quad \text{s.t. constraints (2)-(6) for fixed } f \right\}$$

Next, partition each set P_f into sets P_f^1 such that $\sum_{(i,j) \in A_f^p} x_{ij} \geq 1$ for $p \in P_f^1$, and P_f^2 such that $\sum_{(i,j) \in A_f^p} x_{ij} = 0$ for $p \in P_f^2$. It follows that:

1. For $p \in P_f^1$, constraints (2.4) and (2.3) imply $y_f^p = 1$ and $z_f^p = 0$, respectively;

2. For $p \in P_f^2$, constraints (2.2) and (2.3) imply $y_f^p = 0$ and $z_f^p \leq 1$, respectively.

Now, note that constraint (2.5) is defined over $p \in P_f$, and so is the summation in constraint (2.6). We can determine $\bar{Q}_f(\boldsymbol{x})$ depending on whether set P_f^2 is empty:

1. If $P_f^2 \neq \emptyset$, then constraint (2.5) implies $y_f = 0$ and constraint (2.6) is equivalent to $\sum_{p \in P_f^2} z_f^p \geq 1$. In this case, we have

$$\bar{Q}_f(\boldsymbol{x}) = \min_{z_f^p \in \{0,1\}} \left\{ \sum_{p \in P_f^2} z_f^p c_f^p \quad \text{s.t.} \sum_{p \in P_f^2} z_f^p \geq 1 \right\} = \min_{p \in P_f^2} \left\{ c_f^p \right\}.$$

2. If $P_f^2 = \emptyset$, then constraint (2.5) implies $y_f \leq 1$ and constraint (2.6) is equivalent to $\sum_{p \in P_f^1} z_f^p \geq 1 - y_f$. Since $z_f^p = 0$ for all $p \in P_f^1$, (2.6) implies that $y_f \geq 1$. Hence, $y_f = 1$ and $\bar{Q}_f(\boldsymbol{x}) = 0$.

The two cases can be summarized as

$$\bar{Q}_f(\boldsymbol{x}) = \begin{cases} \min\limits_{p \in P_f^2} \left\{ c_f^p \right\}, & P_f^2 \neq \emptyset; \\ \\ 0, & P_f^2 = \emptyset, \end{cases}$$

which is precisely the definition of $Q_f(\boldsymbol{x})$ given in Problem **P1**. Thus, we have $Q(\boldsymbol{x}) = \bar{Q}(\boldsymbol{x})$, and the result follows. □

2.2 Relation to FCLAP

Recall that FCLAP locates facilities in order to maximize the number of flow-based customers that encounter these facilities along their predetermined travel paths. Here we consider a case with a variable number of facilities and note that maximizing a weighted sum of captured flows is equivalent to minimizing the weighted sum of uncaptured flows. Using our notation, this variant of FCLAP is formulated as

$$\text{FCLAP'}: \min_{x_{ij}, y_f^p \in \{0,1\}} \sum_{(i,j)\in A} x_{ij} w_{ij} + \sum_{f \in F} \left(1 - y_f^p\right) c_f^p$$

$$\text{s.t.} \sum_{(i,j)\in A_f^p} x_{ij} \geq y_f^p \qquad \forall f \in F$$

where p denotes the predetermined path of a flow.

The following part provides two propositions that 1) argue that **P2** encompasses FCLAP', and 2) establish relations between solutions to EFCP and FCLAP'. The first result will be further used to analyze the computational complexity of EFCP. The second result will be later illustrated through numerical examples, which show that allocations suggested by FCLAP' do poorly in a setting where flows try to evade facilities and thus motivate the application of the proposed EFCP.

Proposition 2. *For $k_f = 1$, Problem **P2** reduces to FCLAP'.*

Proof. For $k_f = 1$, we have $|P_f| = 1$ and thus:

1. We can omit condition $\forall p \in P_f$ from constraints (2.2)-(2.5) and (2.6).

2. Variables y_f and y_f^p are equivalent by construction and thus constraints (2.5) can be omitted.

3. Constraints (2.6), $\sum_{p \in P_f} z_f^p \geq 1 - y_f$, are equivalent to $z_f^p \geq 1 - y_f$, which is same as $z_f^p \geq 1 - y_f^p$.

4. Constraints (2.3) and (2.6), $z_f^p \leq 1 - y_f^p$ and $z_f^p \geq 1 - y_f^p$, imply $z_f^p = 1 - y_f^p$.

Now we can replace z_f^p from (2.1) with $1 - y_f^p$ and omit (2.3) and (2.6). This reduces (2.1)-(2.6) to the following mathematical program:

$$
\begin{aligned}
\min_{x_{ij}, y_f^p \in \{0,1\}} \; & \sum_{(i,j) \in A} x_{ij} w_{ij} + \sum_{f \in F} \left(1 - y_f^p\right) c_f^p \\
\text{s.t.} \; & \sum_{(i,j) \in A_f^p} x_{ij} \geq y_f^p & \forall f \in F \\
& \sum_{(i,j) \in A_f^p} x_{ij} \leq \left|A_f^p\right| \cdot y_f^p & \forall f \in F
\end{aligned}
$$

Note the following relations defined with the two above inequalities:

1. If $\sum_{(i,j) \in A_f^p} x_{ij} = 0$ then the first inequality implies $y_f^p = 0$;

2. If $\sum_{(i,j) \in A_f^p} x_{ij} \geq 1$ then the second inequality implies $y_f^p = 1$.

Considering that the objective function would force y_f^p to take the value of 1 whenever $\sum_{(i,j) \in A_f^p} x_{ij} \neq 0$, the second inequality can be omitted. □

Proposition 3. *Let* $\boldsymbol{x^*_{FCLAP'}}$ *denote the optimal solution for FCLAP' in which p is defined as the shortest path of a flow. Similarly, let* $\boldsymbol{x^*_{EFCP}}$ *be the optimal solution for EFCP (Problem* ***P2****). If we let* $FCLAP'(\boldsymbol{x})$ *and* $EFCP(\boldsymbol{x})$ *denote values of the facility allocation* $\boldsymbol{x}$ *in these two problems, then*

1. $EFCP(\boldsymbol{x}) \geq FCLAP(\boldsymbol{x})$

2. $EFCP(\boldsymbol{x^*_{EFCP}}) \geq FCLAP(\boldsymbol{x^*_{FCLAP}})$

3. $EFCP(\boldsymbol{x^*_{EFCP}}) \leq EFCP(\boldsymbol{x^*_{FCLAP}})$

Proof. The third inequality obviously holds because $\boldsymbol{x^*_{EFCP}}$ is the optimal solution for EFCP, which is a minimization problem.

To show that the first inequality holds, note that constraints (2.5) hold for all $p \in P_f$ and thus $y_f \leq y_f^{p^*(f)}$, where $p^*(f)$ denotes the shortest path of a flow. This relation and constraint (2.6) imply

$$\left(\sum_{p \in P_f} z_f^p \geq 1 - y_f\right) \wedge \left(y_f \leq y_f^{p^*(f)}\right) \quad \Rightarrow \quad \sum_{p \in P_f} z_f^p \geq 1 - y_f^{p^*(f)}$$

First, we include c_f^p in the summation on the left hand side and multiply the right hand side with $c_f^{p^*(f)}$. Note that we are allowed to do this because $c_f^p \geq c_f^{p^*(f)}$ for all $p \in P_f$ (based on Assumption 3 as well as the definition of $p^*(f)$ as the shortest path of a flow). Second, we sum the obtained inequality for all the flows:

$$\sum_{p \in P_f} z_f^p c_f^p \geq \left(1 - y_f^{p^*(f)}\right) c_f^{p^*(f)} \quad \Rightarrow \quad \sum_{f \in F} \sum_{p \in P_f} z_f^p c_f^p \geq \sum_{f \in F} \left(1 - y_f^{p^*(f)}\right) c_f^{p^*(f)}$$

Finally, we add the facility cost to both sides of the last inequality and note that

$$\sum_{(i,j) \in A} x_{ij} w_{ij} + \sum_{f \in F} \sum_{p \in P_f} z_f^p c_f^p \geq \sum_{(i,j) \in A} x_{ij} w_{ij} + \sum_{f \in F} \left(1 - y_f^{p^*(f)}\right) c_f^{p^*(f)}$$

whence $EFCP\left(\boldsymbol{x}\right) \geq FCLAP'\left(\boldsymbol{x}\right)$ as required.

The second inequality follows immediately from the first, because $\boldsymbol{x}^*_{\boldsymbol{FCLAP'}}$ is optimal for FCLAP' and thus $EFCP\left(\boldsymbol{x}^*_{\boldsymbol{EFCP}}\right) \geq FCLAP'\left(\boldsymbol{x}^*_{\boldsymbol{FCLAP'}}\right)$. □

2.3 Structural Properties

Since Problem **P1** represents minimization of a set function, we would be interested in checking whether this set function is submodular or supermodular. On the one hand, submodular set functions can be minimized in strongly polynomial time [45, 46]. On the other hand, a simple greedy heuristic is guaranteed to perform well when applied to minimization of supermodular functions. The bound on this greedy approximation was extensively used in the literature on FCLAP and is stated below for completeness. Before we proceed, recall that a set function is nondecreasing, submodular, and supermodular if for all $S \subset T \subset N$ and $k \notin T$ the following holds:

1. nondecreasing: $h\left(S\right) \leq h\left(T\right)$
2. submodular: $h\left(T \cup \{k\}\right) - h\left(T\right) \leq h\left(S \cup \{k\}\right) - h\left(S\right)$
3. supermodular: $h\left(T \cup \{k\}\right) - h\left(T\right) \geq h\left(S \cup \{k\}\right) - h\left(S\right)$ (i.e. $-h$ is submodular)

Theorem 1 (Nemhauser, Wolsey, and Fisher 1978 [47])**.** *Consider the optimization problem*

$$Z^* = \max_{S \subset N,\, |S| \leq m} h(S).$$

Let Z^G be a value returned by the greedy heuristic that sequentially selects elements in N that myopically maximize the objective function. If $h(S)$ is submodular and nondecreasing, then

$$\frac{Z^G}{Z^*} \geq 1 - \left(1 - \frac{1}{m}\right)^m \geq 1 - \frac{1}{e} \approx 0.63.$$

Numerous papers on FCLAP show that the problem of locating m facilities to maximize the weighted sum of captured flows, can be expressed using the framework of Theorem 1 [7, 14, 20, 23, 25]. This result guarantees that a greedy heuristic will quickly provide solutions for FCLAP that are within 37% of the optimum. Numerical comparison with exact solution techniques, e.g. branch and bound, shows that a greedy algorithm performs exceptionally well yielding optimal or near optimal solutions [14, Table 2].

However, although EFCP is closely related to FCLAP, it is a substantially more complex problem. Our next result shows that EFCP is neither submodular (and thus existing results on polynomial complexity do not apply) nor supermodular (and thus a greedy heuristic is not guaranteed to perform well). In fact, as we show later on, a greedy heuristic can perform arbitrarily poorly in EFCP.

Proposition 4. *The objective function of Problem **P1** is non-submodular, non-supermodular, and non-monotonic.*

Proof. We prove this by contradiction. Let $\boldsymbol{w}$ be a vector of zeros. Consider a case of a single flow f and its k_f shortest paths indexed $p = 1, ..., k_f$. Let S denote an allocation of facilities such that only the first r shortest paths are covered by checkpoints and $r < k_f$. Let T denote the allocation of facilities covering the first

r shortest paths like in S, as well as shortest paths indexed $p = r+2, ..., k_f - 1$. Moreover, let $h(S)$ denote the objective function value of **P1** given allocation S. Clearly, we have $S \subset T \subset A$ and $h(S) = h(T) = c_f^{r+1}$ because the flow travels along the shortest unmonitored path, which is path $r+1$ in both cases. Now let $k \notin T$ be the location of a checkpoint such that only the $(r+1)$-shortest path is intercepted and observe the following:

$$h(T \cup \{k\}) - h(T) = c_f^{k_f} - c_f^{r+1}$$

$$h(S \cup \{k\}) - h(S) = c_f^{r+2} - c_f^{r+1}$$

The above equalities imply $h(T \cup \{k\}) - h(T) \geq h(S \cup \{k\}) - h(S)$ because $c_f^{k_f} \geq c_f^{r+2}$. Thus, submodularity does not hold for all $S \subset T \subset A$ and $k \notin T$.

To show that the function is neither supermodular nor monotonic, let T denote the allocation of facilities that cover the first r shortest paths like in S, as well as shortest paths indexed $p = r+2, ..., k_f$. Then,

$$h(T \cup \{k\}) - h(T) = 0 - c_f^{r+1} \leq 0$$

$$h(S \cup \{k\}) - h(S) = c_f^{r+2} - c_f^{r+1} \geq 0$$

The above expressions imply that the objective function of **P1** is not monotonic. Moreover, since this time we have $h(T \cup \{k\}) - h(T) \leq h(S \cup \{k\}) - h(S)$, supermodularity does not hold for all $S \subset T \subset A$ and $k \notin T$. □

Proposition 4 indicates that standard solution approaches for FCLAP are not guaranteed to work well in EFCP. We now address the computational complexity of EFCP.

Proposition 5. *Problem **P2** is NP-hard.*

Proof. To prove that Problem **P2** is NP-hard, we reduce a known NP-hard problem, namely the problem of "Locating Uncapacitated Inspection Stations" (LUIS), studied by [12], to an instance of Problem **P2**. The goal of this problem is to place the smallest possible number of inspection stations needed to cover all truck flows (thus ensuring that all trucks are inspected). Using our notation, it can be written as

$$\text{LUIS:} \quad \min_{x_{ij} \in \{0,1\}} \sum_{(i,j) \in A} x_{ij}$$

$$\text{s.t.} \quad \sum_{(i,j) \in A_f^p} x_{ij} \geq 1$$

where (i, j) are edges in a graph, f denotes a truck flow, A_f^p is the set of edges along the single predetermined path of a flow, and A and x_{ij} are as defined earlier. Given an arbitrary instance of LUIS, we construct an instance of **P2** whose optimal solution yields an optimal solution to LUIS.

First, let $w_{ij} = 1$ and let $c_f^p = \sum_{(i,j) \in A} w_{ij}$ for all f. Then the problem

$$\text{LUIS':} \quad \min_{x_{ij}, y_f^p \in \{0,1\}} \sum_{(i,j) \in A} x_{ij} w_{ij} + \sum_{f \in F} (1 - y_f^p) \cdot c_f^p$$

$$\text{s.t.} \quad \sum_{(i,j) \in A_f^p} x_{ij} \geq y_f^p$$

is an instance of FCLAP'. In this formulation, the variable y_f^p equals 1 if flow f is captured and 0 otherwise. However, if we do not capture flow f, we incur a penalty c_f^p that exceeds the cost of implementing a station on each edge. Therefore, the optimal solution to LUIS' never leaves any flows uncaptured, and will remain unchanged if we require $y_f^p = 1$, in which case LUIS and LUIS' are identical. Since FCLAP' is an instance of **P2** with $k_f = 1$, we can conclude that problem P2 is NP-hard. □

Note that Problem **P1** minimizes the total investment in facilities and excessive damage associated with unintercepted flows. While this is a reasonable economic objective, most work on FCLAP considers a fixed number of facilities, and focuses on placing them to maximize the number of captured customers (which is equivalent to minimizing excessive damage associated with uncaptured flows). Thus, we also consider a variant of **P1** whose objective function only includes excessive damage, not the cost of implementing the facilities. This problem is denoted by

$$\textbf{P1'}: \quad \min_{\boldsymbol{x} \in \{0,1\}^{|A|}, \sum_{(i,j) \in A} x_{ij} \leq m} \quad Q(\boldsymbol{x}).$$

It is straightforward to show that the structural properties of **P1** obtained in Proposition 4 also hold for **P1'**. Additionally, for $k_f = 1$, Problem **P1'** transforms into a classic FCLAP, which is known to be NP-hard [5].

2.4 Solution Techniques

Formulation **P2** represents a binary integer program which can be tackled in any mathematical programming software using branch-and-bound-based algorithms. This section shows that the binary variables y_f and z_f^p can be linearly relaxed without altering the optimal solution or the value of the objective function. As it will be illustrated numerically, this partial linear relaxation typically reduces solution time for **P2** by about 30%. In addition, a tighter formulation of **P2** is proposed, which enables linear relaxation of all the variables except x_{ij}. These results are summarized in the following two theorems.

Theorem 2. *Let $EFCP^1_{LR}$ denote a partial linear relaxation of EFCP (Problem **P2**), such that $y_f, z^p_f \geq 0$ and $x_{ij}, y^p_f \in \{0,1\}$. Then, an optimal facility allocation for $EFCP^1_{LR}$ is also optimal for $EFCP$, and the two problems have the same optimal objective values.*

Proof. To prove this, it suffices to show that for any fixed binary $\boldsymbol{x}$ and $\boldsymbol{y} = \{y^p_f \mid p \in P_f, f \in F\}$ which satisfy (2.2) and (2.4), the two problems have the same optimal second-stage value. We can show this by noting that for fixed binary $\boldsymbol{x}$ and $\boldsymbol{y}$ which satisfy (2.2) and (2.4), the objective function of problem $EFCP^1_{LR}$ corresponds to

$$\sum_{(i,j)\in A} x_{ij} w_{ij} + \sum_{f\in F} B_f(\boldsymbol{y}),$$

where

$$\begin{aligned} B_f(\boldsymbol{y}) = \min_{y_f, z^p_f \geq 0} & \sum_{p\in P_f} z^p_f c^p_f \\ \text{s.t. } & z^p_f \leq 1 - y^p_f & \forall p \in P_f \\ & y_f \leq y^p_f & \forall p \in P_f \\ & \sum_{p\in P_f} z^p_f \geq 1 - y_f \end{aligned}$$

We proceed by partitioning each set P_f into P^1_f and P^2_f, such that $y^p_f = 1$ for $p \in P^1_f$ and $y^p_f = 0$ for $p \in P^2_f$. Now note that $B_f(\boldsymbol{y})$ can be determined based on whether set P^2_f is empty. Using arguments similar to those in the proof of Proposition 1, we conclude that, for fixed $\boldsymbol{x}$ and $\boldsymbol{y}$, the objective function of

$EFCP^1_{LR}$ is given by

$$\sum_{(i,j)\in A} x_{ij}w_{ij} + \sum_{f\in F} B_f(\boldsymbol{y}),$$

where

$$B_f(\boldsymbol{y}) = \begin{cases} \min\limits_{p\in P^2_f} \left\{c^p_f\right\}, & P^2_f \neq \emptyset; \\ 0, & P^2_f = \emptyset. \end{cases}$$

Finally, we observe that $\sum_{f\in F} B_f(\boldsymbol{y})$ corresponds to $Q(\boldsymbol{x})$ from Problem **P1**. This implies that for fixed $\boldsymbol{x}$ and $\boldsymbol{y}$, problems $EFCP^1_{LR}$ and $EFCP$ have the same objective function values. Thus, they also have the same objective function values for the optimal $\boldsymbol{x}$ and $\boldsymbol{y}$. □

Remark 1. *The partial linear relaxation stated in Theorem 2 reduces the number of binary integer variables from* $\underbrace{|A|}_{x_{ij}} + \underbrace{|F|}_{y_f} + \underbrace{2\cdot\sum_{f\in F}|P_f|}_{y^p_f \,\&\, z^p_f}$ *to* $|A| + \sum_{f\in F}|P_f|$.

In Theorem 2, it was shown that y_f and z^p_f can be linearly relaxed without altering the optimal solution. The following result shows that we can additionally relax y^p_f, provided that we tighten formulation **P2**. In this case, however, the relaxation comes at the cost of additional constraints. Whether it will run faster than $EFCP^1_{LR}$ is problem-dependent.

Theorem 3. *Let $EFCP^2_{LR}$ denote a partial linear relaxation of EFCP (Problem **P2**), such that:*

1. *Constraints* (2.4) *are replaced with constraints*

$$x_{ij} \leq y^p_f \qquad \forall (i,j)\in A^p_f \quad \forall p\in P_f \quad \forall f\in F \tag{2.7}$$

   *which tighten the formulation **P2**;*

2. *All auxiliary variables are linearly relaxed* $y_f^p, y_f, z_f^p \geq 0$, *whereas the facility location variables are kept binary* $x_{ij} \in \{0,1\}$.

Let $\boldsymbol{x^*_{EFCP^2_{LR}}}$ *denote its optimal solution with objective* $EFCP^2_{LR}(\boldsymbol{x^*_{EFCP^2_{LR}}})$. *Then,* $\boldsymbol{x^*_{EFCP^2_{LR}}} = \boldsymbol{x^*_{EFCP}}$ *and* $EFCP^2_{LR}(\boldsymbol{x^*_{EFCP^2_{LR}}}) = EFCP(\boldsymbol{x^*_{EFCP}})$.

Proof. We first prove that for any allocation of checkpoints $\boldsymbol{x}$, two problems have the same objective functions, $EFCP^2_{LR}(\boldsymbol{x}) = EFCP(\boldsymbol{x})$. We can show this by working through the constraints of $EFCP^2_{LR}$ similarly to Proposition 1. We begin by noting that objective is separable in f, whence

$$EFCP^2_{LR}(\boldsymbol{x}) = \boldsymbol{x^T w} + \sum_{f \in F} \bar{Q}_f(\boldsymbol{x})$$

where $\bar{Q}_f(\boldsymbol{x})$ equals

$$\min_{y_f^p, y_f, z_f^p \geq 0} \left\{ \sum_{p \in P_f} z_f^p c_f^p \quad \text{s.t. (2.2)-(2.3), (2.5)-(2.7) for fixed } f \right\}.$$

We partition each set P_f into sets P_f^1 such that $\sum_{(i,j) \in A_f^p} x_{ij} \geq 1$ for $p \in P_f^1$, and P_f^2 such that $\sum_{(i,j) \in A_f^p} x_{ij} = 0$ for $p \in P_f^2$, and observe the following:

1. For $p \in P_f^1$, constraints (2.7) and (2.3) imply $y_f^p \geq 1$ and $z_f^p = 0$;

2. For $p \in P_f^2$, constraints (2.2) and (2.3) imply $y_f^p = 0$ and $z_f^p \leq 1$, respectively.

Now we can compute $\bar{Q}_f(\boldsymbol{x})$ similarly to Proposition 1. Thus, we omit the corresponding steps and conclude that the objective function of $EFCP^2_{LR}(\boldsymbol{x})$ can be given as

$$EFCP^2_{LR}(\boldsymbol{x}) = \boldsymbol{x^T w} + \sum_{f \in F} \bar{Q}_f(\boldsymbol{x})$$

where

$$\bar{Q}_f(\boldsymbol{x}) = \begin{cases} \min\limits_{p \in P_f^2} \left\{ c_f^p \right\}, & P_f^2 \neq \emptyset; \\ 0, & P_f^2 = \emptyset. \end{cases}$$

The above expression for $EFCP_{LR}^2(\boldsymbol{x})$ matches the objective function of **P1** and is thus equivalent to the objective of **P2**. Since $EFCP_{LR}^2(\boldsymbol{x}) = EFCP(\boldsymbol{x})$ and two problems have the same feasible regions for $\boldsymbol{x}$, then $\boldsymbol{x}^*_{\boldsymbol{EFCP_{LR}^2}} = \boldsymbol{x}^*_{\boldsymbol{EFCP}}$ and $EFCP_{LR}^2(\boldsymbol{x}^*_{\boldsymbol{EFCP_{LR}^2}}) = EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})$. □

Remark 2. *The partial linear relaxation stated in Theorem 3 reduces the number of binary integer variables from* $|A| + |F| + 2 \cdot \sum_{f \in F} |P_f|$ *to* $|A|$*. Moreover, this partial linear relaxation includes reformulation of constraints (4), which increases the total number of constraints from* $\underbrace{|F|}_{(2.6)} + \underbrace{4 \cdot \sum_{f \in F} |P_f|}_{(2.2)-(2.5)}$ *to* $|F| + 3 \cdot \sum_{f \in F} |P_f| + \sum_{f \in F} \sum_{p \in P_f} |A_f^p|$ *constraints.*

We also consider the performance of a greedy heuristic that introduces checkpoints at the best current locations as long as the facility implementation improves the objective function (Algorithm 1). Recall that such heuristics are often used in FCLAP, where they can be guaranteed to perform within 37% of optimality. However, in EFCP, the greedy heuristic cannot be guaranteed to perform within any fraction of the optimal value. Our numerical experiments include cases where the heuristic performs very poorly.

Proposition 6. *For any* $0 < \varepsilon < 1$*, there exists an instance of EFCP (Problem* ***P1****) for which* $EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}}) \leq \varepsilon \cdot EFCP(\boldsymbol{x}^{\boldsymbol{G}})$*, where* $\boldsymbol{x}^{\boldsymbol{G}}$ *represents the allocation of checkpoints found by the greedy heuristic.*

Algorithm 1 Greedy heuristic for **P1**

Initialize $R^0 \leftarrow \emptyset$

For $t = 0, ..., |A|$

$(i,j)^* = \arg \min_{(i,j) \in A \setminus R^t} h\left(R^t \cup (i,j)\right)$

if $h\left(R^t \cup (i,j)^*\right) \leq h\left(R^t\right)$ then

$R^{t+1} \leftarrow R^t \cup (i,j)^*$

else

break

end-if

end-for

return $R^G = R^t$, $Z^G = h\left(R^t\right)$

Proof. Let $0 < \varepsilon < 1$, and suppose that there is a single flow f that can travel along at least two edge-disjoint paths. In this case, the optimal value can be expressed as $EFCP(\boldsymbol{x^*_{EFCP}}) = \min\left(c_f^1, \sum_{(i,j) \in S} w_{ij}\right)$, where S is the least expensive allocation of facilities that covers all the paths of flow f. Moreover, the greedy heuristic is initialized with a solution that includes no facility implementation and the corresponding damage c_f^1. Since facility implementation only exacerbates the objective function in the first iteration (i.e. a flow diverts and/or facility cost is incurred), the greedy heuristic stops after the first pass and returns the solution $EFCP(\boldsymbol{x^G}) = c_f^1$.

Recall that c_f^1 represents the excessive damage produced if flow f travels along the shortest path and note that c_f^1 can be arbitrarily high depending on the intensity of the flow and length of the path. Suppose that $c_f^1 = \frac{2}{\varepsilon} \sum_{(i,j) \in S} w_{ij}$ and observe the following:

$$\frac{EFCP(\boldsymbol{x^*_{EFCP}})}{EFCP(\boldsymbol{x^G})} = \frac{\min\left(c_f^1, \sum_{(i,j)\in S} w_{ij}\right)}{c_f^1} = \frac{\sum_{(i,j)\in S} w_{ij}}{\frac{2}{\varepsilon}\sum_{(i,j)\in S} w_{ij}} = \frac{\varepsilon}{2}$$

The above equality shows that for any $0 < \varepsilon < 1$, there exists an instance of **P1** for which $EFCP(\boldsymbol{x^*_{EFCP}}) \leq \varepsilon \cdot EFCP(\boldsymbol{x^G})$. □

Moreover, we consider Problem **P1'** and a greedy heuristic that places a given number of facilities (e.g. m facilities) in the best current position, as in [14]. We show that a bound cannot be determined for this greedy algorithm either. Our numerical experiments also include instances where it performs poorly.

Proposition 7. *For any $\varepsilon > 0$, there exists an instance of EFCP (Problem **P1'**) for which $EFCP(\boldsymbol{x^*_{EFCP}}) \leq \varepsilon \cdot EFCP(\boldsymbol{x^G})$, where $\boldsymbol{x^G}$ represents the allocation of checkpoints found by the greedy heuristic.*

Proof. Let $\varepsilon > 0$, and consider a completely connected network with $m + 1$ nodes. Suppose that there are m flows with the same origin node O and m distinct destination nodes. Furthermore, suppose flows can travel from O to their destinations through all m remaining nodes (i.e. these nodes are not "too far" apart and thus all possible paths are acceptable). Clearly, the optimal solution consists of locating m facilities along m links adjacent to node O and thus $EFCP(\boldsymbol{x^*_{EFCP}}) = 0$.

On the other hand, the greedy heuristic is initialized with a solution that includes no facility implementation and all flows traveling freely from O to their m destination nodes. In the first step, the greedy heuristic tries implementing a facility on all the links. However, placing a facility on any of the links adjacent to node O yields an increased excessive damage since the corresponding flow diverts. The greedy heuristic proceeds by implementing all m facilities on links connecting m destination nodes, without intercepting any flows. Thus, we have $EFCP(\boldsymbol{x^G}) = \sum_{f=1}^{m} c_f^1$.

In the described case, we have $EFCP(\boldsymbol{x^*_{EFCP}})/EFCP(\boldsymbol{x^G}) = 0 < \varepsilon$. Thus, for any $\varepsilon > 0$, there exists an instance of **P1'** for which $EFCP(\boldsymbol{x^*_{EFCP}}) \leq \varepsilon \cdot EFCP(\boldsymbol{x^G})$. □

2.5 Numerical Experiments

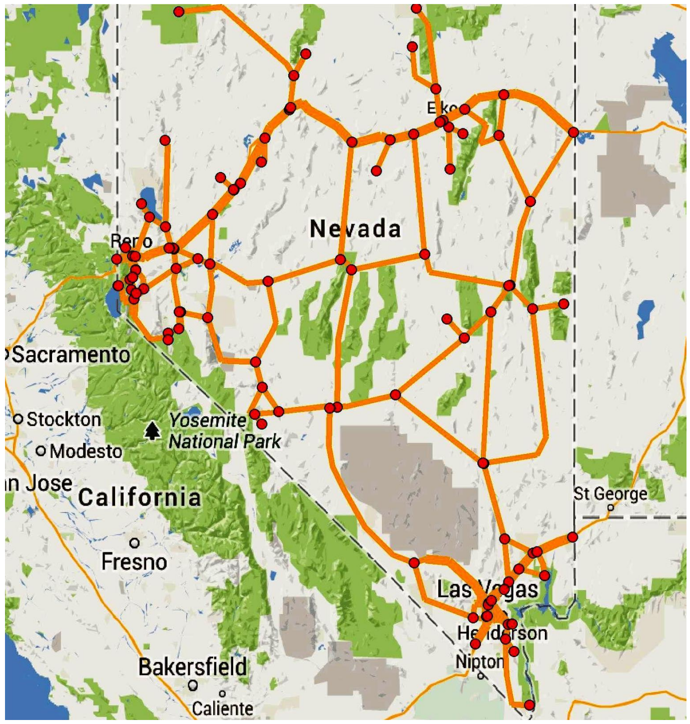
A set of simulated problems are solved to obtain insights on 1) the benefits of the proposed partial linear relaxations, 2) the performance of the greedy heuristics, and 3) performance of FCLAP-based facility allocations in a setting where flows evade facilities. The random instances are based on real-world road networks of Nevada and Vermont. They include 400 and 200 randomly simulated flows, respectively, as well as differently specified:

1. Willingness of flows to avoid facilities (i.e. k_f is defined so that the $k_f + 1$ shortest path is 1.1 or 1.2 times longer than the shortest path);

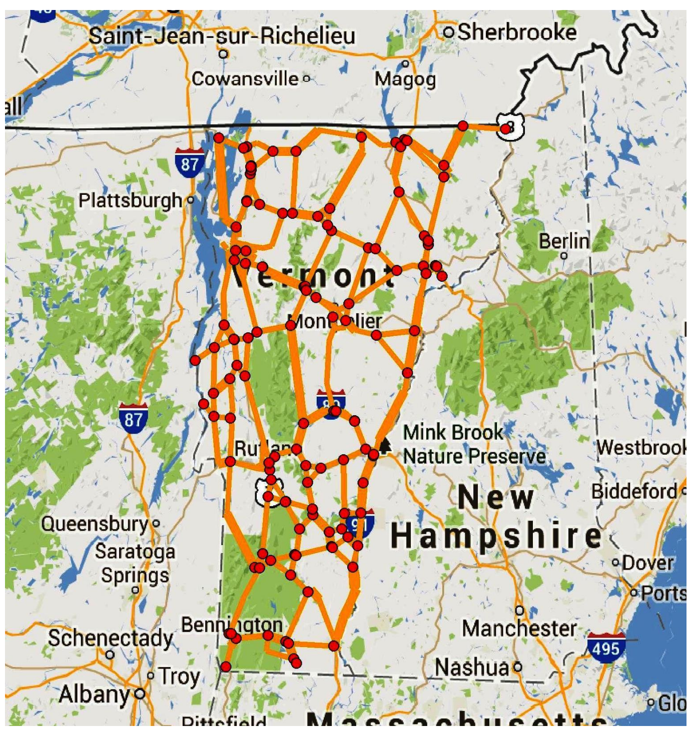
2. Cost of facilities for Problem **P1**, or number of facilities for Problem **P1'**.

The mathematical programming formulations were implemented in GAMS 23.5 and solved using GAMS/CPLEX solver for mixed integer programs on a PC with an AMD Athlon 3300 GHz processor with 4 GB of RAM. Tables 2.1 and 2.2 report computation times for different formulations. The $EFCP$ refers to solving

Problem (2.1)-(2.6) as a binary integer program, whereas $EFCP^1_{LR}$ refers to solving the partial linear relaxation proposed in Theorem 2. The greedy heuristics were implemented in C++ in a Linux environment. Run times of greedy heuristics where below one second in almost all instances, so they were omitted from Tables 2.1 and 2.2.



(a) Road network of Nevada includes 130 links



(b) Road network of Vermont includes 178 links

Figure 2.1: Real-world road networks used in numerical experiments

Table 2.1: Summary of Results for Nevada

Threshold	Problem **P2**					Problem **P2'**				
	WIM (k$)	$EFCP$ (sec)	$EFCP^1_{LR}$ (sec)	$\frac{EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})}{EFCP(\boldsymbol{x}^{\boldsymbol{greedy}})}$	$\frac{EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})}{EFCP(\boldsymbol{x}^*_{\boldsymbol{FCLAP'}})}$	m	$EFCP$ (sec)	$EFCP^1_{LR}$ (sec)	$\frac{EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})}{EFCP(\boldsymbol{x}^{\boldsymbol{greedy}})}$	$\frac{EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})}{EFCP(\boldsymbol{x}^*_{\boldsymbol{FCLAP'}})}$
1.1	10	34	24	0.86	0.75	2	65	34	0.95	0.79
	60	9	8	0.89	0.94	4	101	28	0.65	0.49
	110	15	9	0.87	0.99	5	21	16	0.59	0.35
	160	13	13	0.86	0.99	7	52	35	0.71	0.61
	210	17	14	0.84	0.99	8	45	43	0.64	0.51
	260	20	23	0.91	0.84	10	20	20	0.59	0.35
	310	32	34	0.97	0.88	11	28	28	0.72	0.74
	360	7	7	1.00	0.93	13	22	19	0.58	0.19
1.2	10	507	410	0.77	0.60	2	481	388	0.95	0.49
	60	779	517	0.68	0.46	4	762	421	0.49	0.17
	110	353	297	0.73	0.57	5	505	514	0.44	0.12
	160	772	475	0.76	0.66	7	553	670	0.39	0.52
	210	537	405	0.83	0.72	8	719	371	0.36	0.33
	260	336	298	0.89	0.61	10	542	577	0.37	0.24
	310	340	301	0.95	0.61	11	445	437	0.43	0.64
	360	273	259	1.00	0.83	13	462	390	0.52	0.20

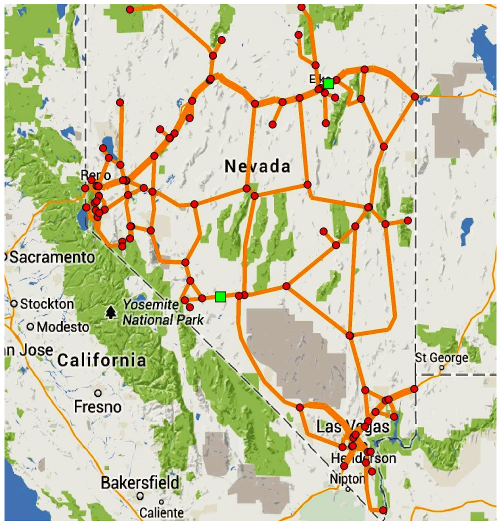
Table 2.2: Summary of Results for Vermont

Threshold	Problem **P2**					Problem **P2'**				
	WIM (k$)	$EFCP$ (sec)	$EFCP^1_{LR}$ (sec)	$\frac{EFCP(\boldsymbol{x^*_{EFCP}})}{EFCP(\boldsymbol{x^{greedy}})}$	$\frac{EFCP(\boldsymbol{x^*_{EFCP}})}{EFCP(\boldsymbol{x^*_{FCLAP'}})}$	m	$EFCP$ (sec)	$EFCP^1_{LR}$ (sec)	$\frac{EFCP(\boldsymbol{x^*_{EFCP}})}{EFCP(\boldsymbol{x^{greedy}})}$	$\frac{EFCP(\boldsymbol{x^*_{EFCP}})}{EFCP(\boldsymbol{x^*_{FCLAP'}})}$
1.1	5	57	67	0.83	0.28	2	229	80	0.92	0.77
	7.5	82	88	0.88	0.38	5	13	11	0.49	0.27
	10	77	82	0.88	0.42	6	15	13	0.54	0.26
	20	22	21	0.72	0.48	8	37	42	0.58	0.22
	30	23	22	0.72	0.54	11	37	41	0.49	0.16
	40	25	23	0.81	0.67	14	37	31	0.35	0.06
	50	24	22	0.92	0.81	17	58	57	0.22	0.01
	60	42	28	1.00	0.85	20	57	12	0.00	0.00
1.2	5	3117	2071	0.69	0.24	2	3457	3312	0.99	0.81
	7.5	3542	3710	0.92	0.28	5	586	490	0.38	0.23
	10	1439	1279	0.78	0.31	6	994	818	0.40	0.22
	20	907	804	0.57	0.41	8	1079	772	0.43	0.18
	30	1089	984	0.68	0.50	11	987	848	0.28	0.11
	40	1574	1472	0.80	0.62	14	814	950	0.11	0.04
	50	1414	1272	0.92	0.71	17	1140	962	0.05	0.01
	60	1011	1393	1.00	0.76	20	593	492	0.00	0.00

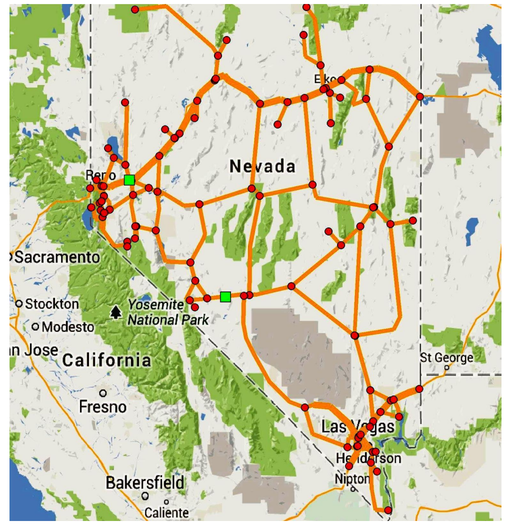
The numerical results summarized in Tables 2.1 and 2.2 indicate that the partial linear relaxation $EFCP^1_{LR}$ proposed in Theorem 2 on average reduced the computation time by about 29%. In particular, in 47/64 cases it reduced the computation time by 42% on average, in 4/64 cases it made no difference, whereas in 13/64 cases it increased the computation time by 11% on average. Tables 2.1 and 2.2 also illustrate the results in Propositions 6 and 7: although the greedy heuristic often performs well, and typically runs in less than a second, there are problem instances where it performs extremely poorly. Moreover, the performance of the greedy heuristic is much worse in Problem **P1'**. We also omit the running times of the tighter formulation proposed in Theorem 3, as the increased number of constraints led to slower computation times for these problem instances.

We also apply FCLAP' to find $\boldsymbol{x}^*_{\boldsymbol{FCLAP'}}$ and then evaluate these allocations in a setting where flows try to evade facilities. We compare the obtained $EFCP(\boldsymbol{x}^*_{\boldsymbol{FCLAP'}})$ with the optimal values $EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})$ and show their ratios in Tables 2.1 and 2.2. The FCLAP' produced solutions that were on average only within 52% of the optimal. This clearly shows the additional value that proposed EFCP adds in allocating facilities that targeted flows try to avoid.

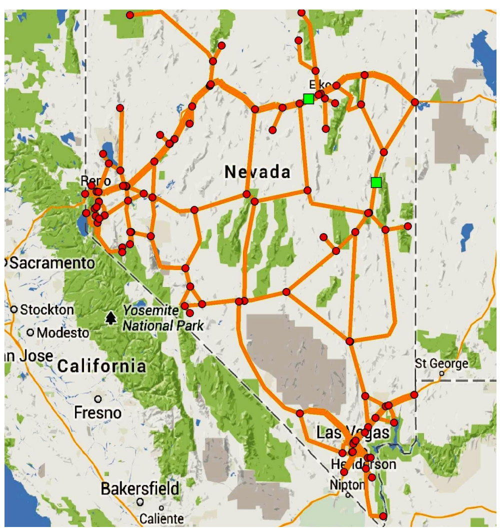
In the remainder of this section we illustrate some of the optimal allocations from Tables 2.1 and 2.2, and show how they change when we vary m, w_{ij}, or the threshold for determining k_f. These illustrations are provided in Figures 2.2-2.5.



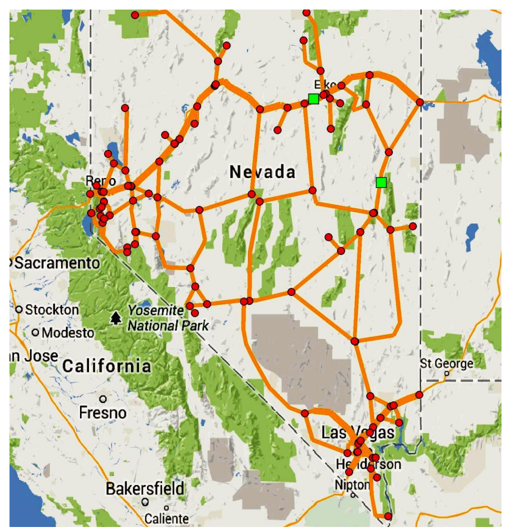
(a) $\boldsymbol{x}^*_{\boldsymbol{FCLAP}}$ for $m = 2$



(b) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 2$ and threshold 1.1

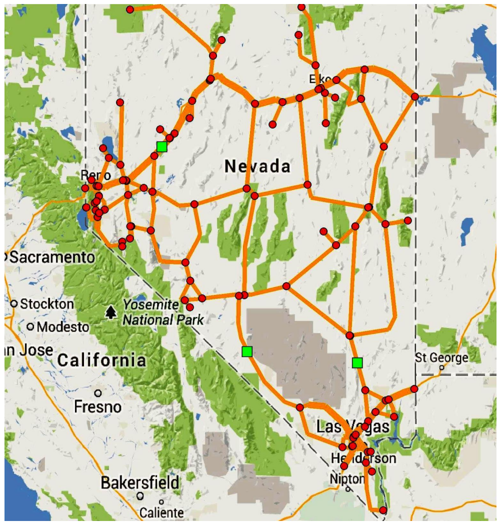


(c) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 2$ and threshold 1.2

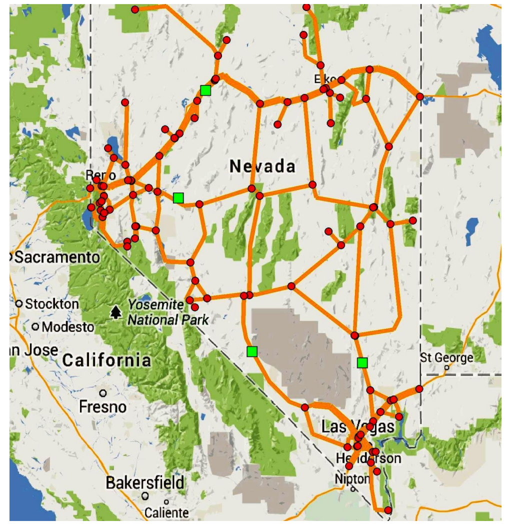


(d) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 2$ and threshold 1.3

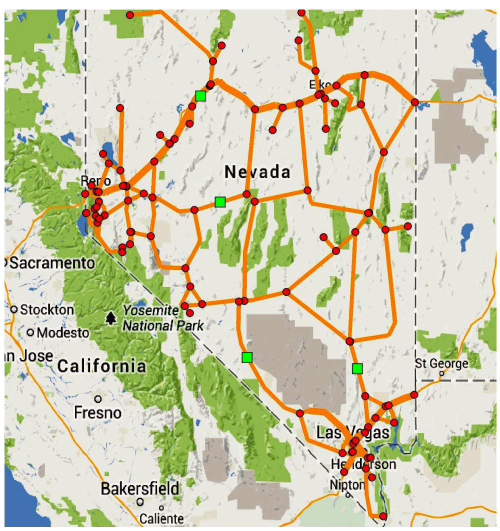
Figure 2.2: Optimal solutions for Nevada based on differently specified threshold (i.e. willingness of flows to evade facilities) and $m = 2$



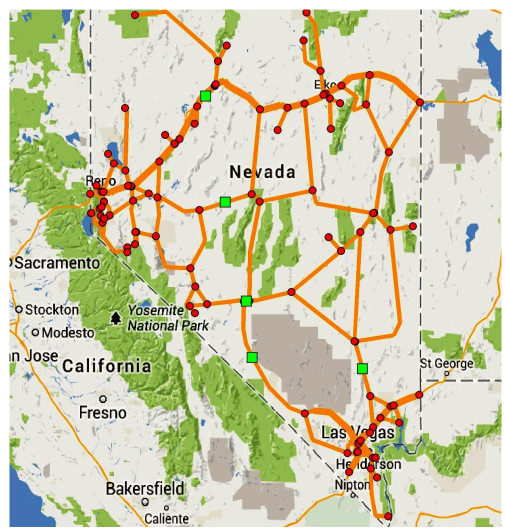
(a) $\boldsymbol{x^*_{FCLAP}}$



(b) $\boldsymbol{x^*_{EFCP}}$ for threshold 1.1

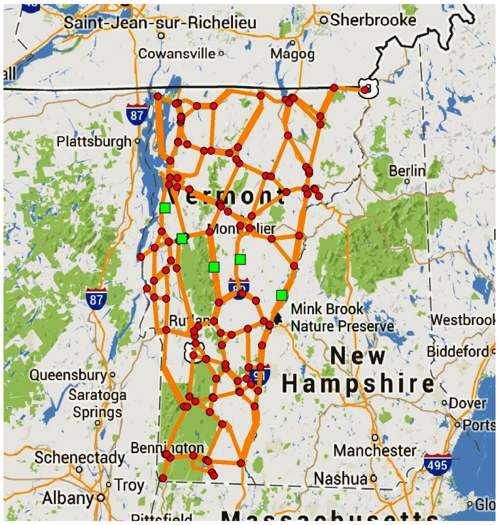


(c) $\boldsymbol{x^*_{EFCP}}$ for threshold 1.2

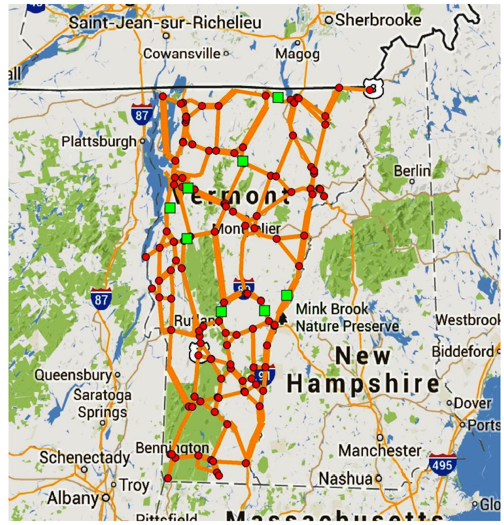


(d) $\boldsymbol{x^*_{EFCP}}$ for threshold 1.3

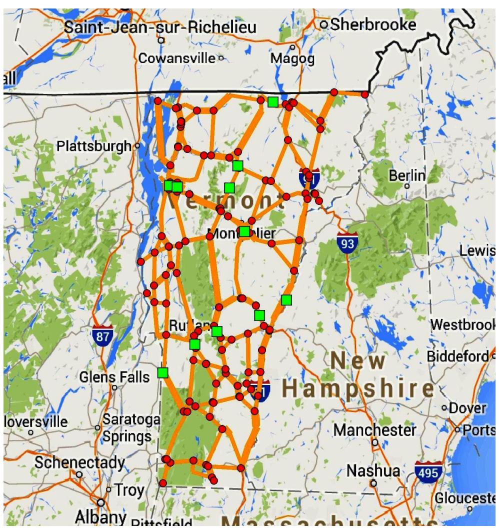
Figure 2.3: Optimal solutions for Nevada based on differently specified threshold (i.e. willingness of flows to evade facilities) and facility cost of $60,000/lane-year



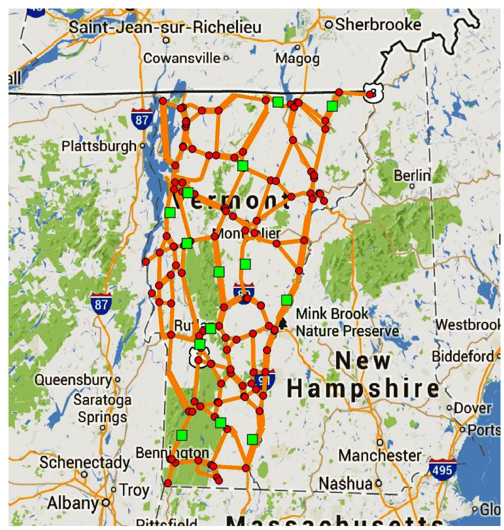
(a) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 5$



(b) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 8$

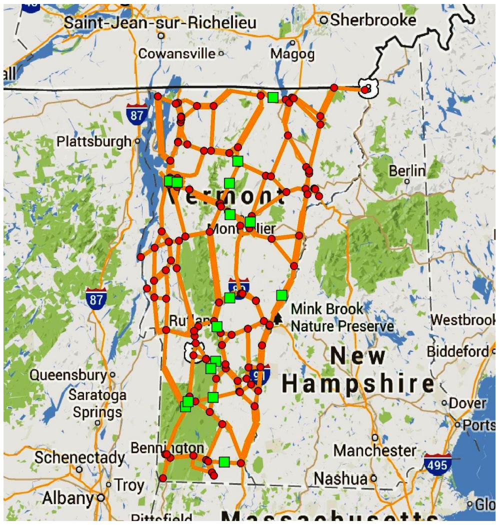


(c) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 11$

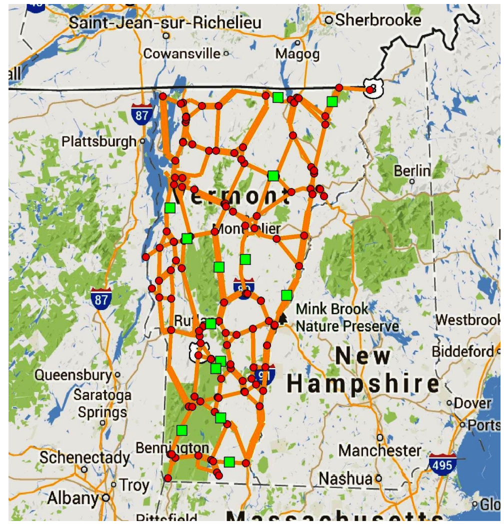


(d) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for $m = 14$

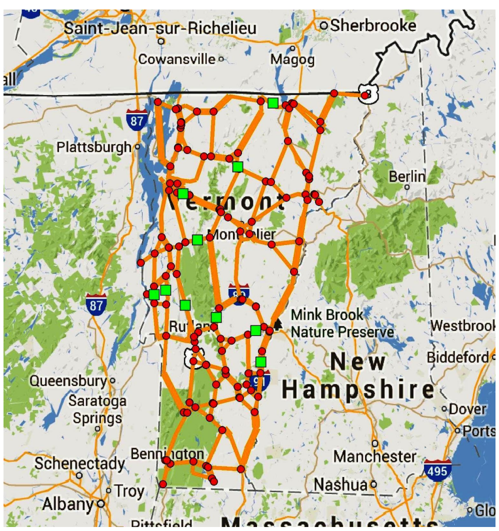
Figure 2.4: Optimal solutions for Vermont based on differently specified m and threshold 1.2 (i.e. willingness the avoid facilities)



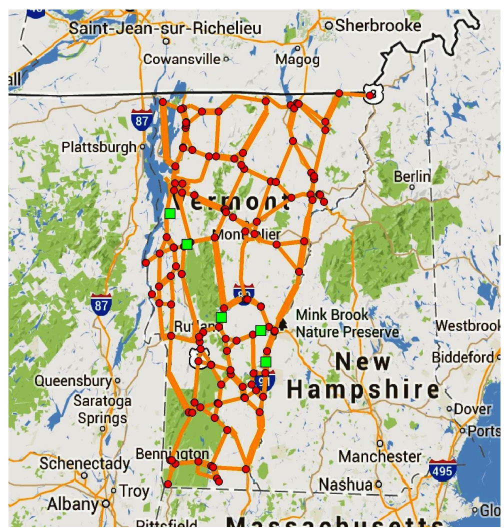
(a) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for 5,000/lane-year



(b) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for 7,500/lane-year



(c) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for 10,000/lane-year



(d) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for 20,000/lane-year

Figure 2.5: Optimal solutions for Vermont based on differently specified cost of facilities and threshold 1.2 (i.e. willingness the avoid facilities)

2.6 Conclusions

This chapter introduced a deterministic flow-capturing model in which targeted flows try to evade facilities that are being located. The proposed EFCP generalizes the previously studied FCLAP, but includes structurally different properties that, for example, can cause a greedy heuristic to perform arbitrarily poorly. It was shown that many binary variables can be linearly relaxed without altering the optimal solution or the value of the objective function. This result proved to be very useful in solving realistic problem instances involving the road networks of Nevada and Vermont, as it considerably reduced the computation time needed to find the optimal solutions. In addition, the numerical comparison of EFCP and FCLAP indicates that solutions optimal to FCLAP do very poorly in a setting where targeted flows try to evade facilities. These results, as well as wide applicability of EFCP in transportation, revenue management, and security and safety management, show the relevance of the proposed type of flow-capturing problem and encourage further extensions of EFCP, which are addressed in the following chapters of this thesis.

Chapter 3: Two-Stage Stochastic EFCP

The EFCP proposed in Chapter 2 represents an optimization problem in which all the parameters are assumed to be known with certainty. For example, the damage that a flow produces (i.e. parameter c_f^p) and its willingness to avoid facilities (i.e. the size of set P_f containing shortest paths) are assumed to be known. However, in real-world applications, this information could be obtained through expert opinion or data collection, which result in different estimations or realizations of these parameters. To address the case when c_f^p and P_f are not known with certainty, this chapter proposes a stochastic extension of EFCP and develops theoretical results that are crucial for solving this problem optimally.

It should be noted that *per mile* pavement and environmental damages that a flow produces vary with the number and types of vehicles within a flow, excessive loads, climate, and weather. On the other hand, the willingness to avoid facilities may depend on both physiological and economic factors (e.g. price of gasoline, driver's hourly pay, trucker's ability to overload the truck which depends on demand, the age of the truck, tires, types of loads, and road conditions). In this analyses, it is assumed that these two parameters are independent.

3.1 Problem Formulation

Let $\boldsymbol{\xi} = \{\xi_f \mid f \in F\}$ be a vector of random variables denoting unit intensities of flows $f \in F$ (i.e. damage or risk produced per unit of distance traveled). Similarly, let $\boldsymbol{\zeta} = \{\zeta_f \mid f \in F\}$ be a vector of discrete random variables denoting the willingness of flows to evade facilities. This quantity could be defined as a percentage by which drivers are willing to increase the distance traveled (e.g. 20% of the shortest path). A particular realization of these random parameters will be denoted by $\omega \in \Omega$. As a result, in the stochastic extension, we will have $P_f(\omega)$, $y_f^p(\omega)$, $y_f(\omega)$, $z_f^p(\omega)$, and $c_f^p(\omega)$, associated with each realization. Recall that, by Assumption 3 from Section 1.3, we can write

$$c_f^p(\omega) = l_f^p(\omega)\xi_f(\omega),$$

where $l_f^p(\omega)$ is the length of path $p \in P_f(\omega)$. The set $P_f(\omega)$ itself is determined by realization of $\zeta(\omega)$.

The two-stage stochastic EFCP, with a fixed number of facilities, can now be formulated as a minimization problem

$$\textbf{SP1:} \qquad \min_{\boldsymbol{x} \in \{0,1\}^{|A|}} \quad \mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}} Q(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{\zeta})$$
$$\text{s.t.} \sum_{(i,j) \in A} x_{ij} \le m$$

where $Q(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega))$ is an oracle that, given an allocation $\boldsymbol{x}$ of checkpoints, computes the excessive damage (or risk) associated with flows and a particular realization of $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$. If a flow is captured, then the corresponding damage is 0. Otherwise, the flow seeks to minimize its travel distance, and produces the damage by traveling along its shortest unmonitored path. More formally, if we let $P_f^2(\omega) = \left\{p \in P_f(\omega) \mid \sum_{x_{ij} \in A_f^p(\omega)} x_{ij} = 0\right\}$ be the set of paths of flow $f \in F$ not covered by facilities, then $Q(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega)) = \sum_{f \in F} Q_f(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega))$, where

$$Q_f(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega)) = \begin{cases} \min\limits_{p \in P_f^2(\omega)} \left\{c_f^p(\omega)\right\}, & P_f^2(\omega) \neq \emptyset; \\ 0, & P_f^2(\omega) = \emptyset. \end{cases}$$

The nonlinear Problem **SP1** can be linearized by introducing three sets of auxiliary binary variables. These variables are used to check whether a flow is captured, and direct the uncaptured flows along their shortest unmonitored paths. Moreover, these variables depend on the realization of random intensity of flows and their willingness to evade facilities.

$$y_f^p(\omega) = \begin{cases} 1, & \text{if at least one facility is located along path } p \in P_f(\omega) \text{ of flow } f \in F \\ 0, & \text{otherwise} \end{cases}$$

$$y_f(\omega) = \begin{cases} 1, & \text{if at least one facility is located along all } p \in P_f(\omega) \text{ of flow } f \in F \\ 0, & \text{otherwise} \end{cases}$$

$$z_f^p(\omega) = \begin{cases} 1, & \text{if flow } f \in F \text{ travels unintercepted along path } p \in P_f(\omega) \\ 0, & \text{otherwise} \end{cases}$$

The two-stage stochastic EFCP can now be formulated as a linear binary integer program, which we denote **SP2**.

First Stage:

$$\min_{\boldsymbol{x} \in \{0,1\}^{|A|}} \quad \mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}} Q(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{\zeta}) \tag{3.1}$$

$$\text{s.t.} \sum_{(i,j) \in A} x_{ij} \leq m \tag{3.2}$$

Second Stage:

$$Q(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega)) = \min_{\substack{y_f^p(\omega) \in \{0,1\} \\ y_f(\omega) \in \{0,1\} \\ z_f^p(\omega) \in \{0,1\}}} \sum_{f \in F} \sum_{p \in P_f(\omega)} z_f^p(\omega) c_f^p(\omega) \tag{3.3}$$

$$\text{s.t.} \sum_{(i,j) \in A_f^p(\omega)} x_{ij} \geq y_f^p(\omega) \qquad \forall p \in P_f(\omega) \quad \forall f \in F \tag{3.4}$$

$$z_f^p(\omega) \leq 1 - y_f^p(\omega) \qquad \forall p \in P_f(\omega) \quad \forall f \in F \tag{3.5}$$

$$\sum_{(i,j) \in A_f^p(\omega)} x_{ij} \leq \left| A_f^p(\omega) \right| \cdot y_f^p(\omega) \qquad \forall p \in P_f(\omega) \quad \forall f \in F \tag{3.6}$$

$$y_f(\omega) \leq y_f^p(\omega) \qquad \forall p \in P_f(\omega) \quad \forall f \in F \tag{3.7}$$

$$\sum_{p \in P_f(\omega)} z_f^p(\omega) \geq 1 - y_f(\omega) \qquad \forall f \in F \tag{3.8}$$

The first-stage objective (3.1) minimizes the expected excessive damage given the maximum implementation of m facilities (3.2). Second stage objective function (3.3) computes the excessive damage associated with unintercepted flows that travel along the shortest unmonitored paths. Constraints (3.4)-(3.6) ensure that if at least one facility is allocated along a path of a flow ($y_f^p(\omega) = 1$), the flow cannot pass unintercepted along that path ($z_f^p(\omega) = 0$). Constraints (3.7) tie the variables guaranteeing that $y_f(\omega)$ can take a value of 1, if all the corresponding paths are covered by at least one facility. Constraints (3.8) require the unintercepted flows to count towards the objective function by producing the excessive damage along the shortest unmonitored path.

Since **SP1** and **SP2** are stochastic extensions of **P1** and **P2** from Chapter 2, it follows that:

1. Problems **SP1** and **SP2** are equivalent;

2. The objective function of Problem **SP1** is non-submodular, non-supermodular, and non-monotonic;

3. Problem **SP2** is NP-hard;

4. A bound cannot be established on the greedy approximation of **SP1**.

3.2 Reducing the Noise

The following theorem argues that some of the randomness inherent to **SP1** and **SP2** can be reduced without altering the problem. In particular, it shows that, under independence assumptions, stochastic flow intensities (i.e. per mile damage or risk) can be replaced with their means while preserving the randomness associated with the willingness of targeted subjects to evade the facilities. This result considerably reduces the noise, which enables us to consider fewer scenarios and thus solve the problem much more efficiently.

Theorem 4. *Suppose that* $\boldsymbol{\xi}$ *and* $\boldsymbol{\zeta}$ *are independent, and let* $\bar{\boldsymbol{\xi}} = \mathbb{E}(\boldsymbol{\xi})$ *denote the expected intensity of flows (i.e. damage or risk per distance traveled). Then the following holds:*

$$\min_{\substack{\boldsymbol{x}\in\{0,1\}^{|A|} \\ \sum_{(i,j)\in A} x_{ij}\leq m}} \mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}} Q(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{\zeta}) = \min_{\substack{\boldsymbol{x}\in\{0,1\}^{|A|} \\ \sum_{(i,j)\in A} x_{ij}\leq m}} \mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x},\bar{\boldsymbol{\xi}},\boldsymbol{\zeta})$$

Proof. For a fixed feasible allocation of facilities $\boldsymbol{x}$ and realization ω, the damage produced by a particular flow $f \in F$, is given by

$$Q_f(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega)) = \begin{cases} \min\limits_{p \in P_f^2(\omega)} \left\{ c_f^p(\omega) \right\}, & P_f^2(\omega) \neq \emptyset; \\ 0, & P_f^2(\omega) = \emptyset. \end{cases}$$

where $P_f^2(\omega)$ is a set of paths such that $\sum\limits_{(i,j) \in A_f^p(\omega)} x_{ij} = 0$ (i.e. a set of paths not covered by a facility). Furthermore, let $s_f(\omega) = \arg \min\limits_{p \in P_f^2(\omega)} \{c_f^p(\omega)\}$ be the shortest unmonitored path of flow f. Let us now define a random variable

$$d_f(\omega) = \begin{cases} l_f^{s_f(\omega)}, & P_f^2(\omega) \neq \emptyset; \\ 0, & P_f^2(\omega) = \emptyset, \end{cases}$$

where $l_f^{s_f(\omega)}$ is the length of the shortest unmonitored path which depends on allocation $\boldsymbol{x}$ and realization of ζ_f.

The total damage produced by all the flows can now be computed as

$$Q(\boldsymbol{x}, \boldsymbol{\xi}(\omega), \boldsymbol{\zeta}(\omega) = \sum_{f \in F} d_f(\omega) \xi_f(\omega),$$

where d_f is a function of the allocation $\boldsymbol{x}$ as well as random variable ζ_f. On the other hand, $\xi_f(\omega)$ represents the intensity of flow f (i.e. per mile damage or risk). Based on the assumed independence of $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$, we have

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}} Q(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{\zeta}) &= \mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}} \left[\sum_{f \in F} d_f \xi_f \right] \\ &= \sum_{f \in F} \mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}} \left[d_f \xi_f \right] \\ &= \sum_{f \in F} \mathbb{E}_{\boldsymbol{\zeta}} \left[d_f \right] \mathbb{E}_{\boldsymbol{\xi}} \left[\xi_f \right] \\ &= \mathbb{E}_{\zeta} \left[\sum_{f \in F} d_f \bar{\xi}_f \right] \end{aligned}$$

and the result follows. □

3.3 Reformulating the Second-Stage

The number of second-stage binary variables and constraints in **SP2** can be further reduced by exploiting the special structure of the problem. Namely, as the willingness of a flow to evade facilities increases, so does the number of its paths. However, some of the paths remain the same for different realizations of ζ_f, so we can use this to reduce the number of path-based constraints.

Let $\zeta_f(\omega^r)$ denote the r-th realization of the random willingness of a flow to avoid facilities, where $r = 1, ..., R$. Moreover, assume that realizations are ordered so that $\zeta_f(\omega^r) \geq \zeta_f(\omega^{r-1})$, and thus $P_f(\omega^{r-1}) \subseteq P_f(\omega^r)$. We can use this to reduce the size of the sets over which constraints (3.5)-(3.7) are defined, while including only one additional constraint. We formulate the scenario-based constraints recursively, while assuming for notational convenience that $P_f(\omega^0) = \emptyset$ and $y_f(\omega^0) = 1$.

First Stage:

$$\min_{x_{ij} \in \{0,1\}} \mathbb{E}_\zeta \sum_{f \in F} \tilde{Q}_f(\boldsymbol{x}, \bar{\xi}, \zeta) \tag{3.9}$$

$$\text{s.t.} \sum_{(i,j) \in A} x_{ij} \leq m \tag{3.10}$$

Second Stage:

$$\tilde{Q}_f(\boldsymbol{x}, \bar{\xi}_f, \zeta_f(\omega^r)) = \min_{\substack{y_f^p(\omega^r) \in \{0,1\} \\ y_f(\omega^r) \in \{0,1\} \\ z_f^p(\omega^r) \in \{0,1\}}} \sum_{p \in P_f(\omega^r)} z_f^p(\omega^r) c_f^p(\omega^r) \tag{3.11}$$

$$\text{s.t.} \sum_{(i,j) \in A_f^p(\omega^r)} x_{ij} \geq y_f^p(\omega^r) \qquad \forall p \in P_f(\omega^r) \setminus P_f(\omega^{r-1}) \tag{3.12}$$

$$z_f^p(\omega^r) \leq 1 - y_f^p(\omega^r) \qquad \forall p \in P_f(\omega^r) \setminus P_f(\omega^{r-1}) \tag{3.13}$$

$$\sum_{(i,j) \in A_f^p(\omega^r)} x_{ij} \leq \left|A_f^p(\omega^r)\right| \cdot y_f^p(\omega^r) \qquad \forall p \in P_f(\omega^r) \setminus P_f(\omega^{r-1}) \tag{3.14}$$

$$y_f(\omega^r) \leq y_f^p(\omega^r) \qquad \forall p \in P_f(\omega^r) \setminus P_f(\omega^{r-1}) \tag{3.15}$$

$$y_f(\omega^r) \leq y_f(\omega^{r-1}) \tag{3.16}$$

$$\sum_{p \in \bigcup_{r'=1}^{r} P_f(\omega^{r'})} z_f^p(\omega^r) \geq 1 - y_f(\omega^r) \tag{3.17}$$

Program (3.9)-(3.17) describes the same relations as (3.1)-(3.8), but includes recursively defined path-based constraints. In this regard, the newly introduced constraint (3.16) ensures that each flow f can be captured in the r-th realization only if it is also captured in realization $r-1$, which includes fewer paths.

The following two remarks imply that, after we apply Theorem 4 and reformulate the problem as in (3.1)-(3.8), the two-stage stochastic EFCP becomes only slightly more difficult than the deterministic EFCP with the largest realizations of ζ_f.

Remark 3. *Let $\zeta_f(\omega^R)$ denote the largest realization of the willingness of a flow to avoid facilities, and let M_f denote the number of realizations of ζ_f. Then the two-stage stochastic EFCP defined with* (3.9)-(3.17) *includes:*

- $\underbrace{|A|}_{x_{ij}} + \underbrace{\sum_{f \in F} M_f}_{y_f(\omega)} + \underbrace{2 \cdot \sum_{f \in F} \left|P_f(\omega^R)\right|}_{y_f^p(\omega)\ \&\ z_f^p(\omega)}$ *binary variables;*
- $\underbrace{1}_{(3.10)} + \underbrace{4 \cdot \sum_{f \in F} \left|P_f(\omega^R)\right|}_{(3.12)-(3.15)} + \underbrace{2 \cdot \sum_{f \in F} M_f}_{(3.16)-(3.17)}$ *constraints.*

Remark 4. *Consider* (3.9)-(3.17) *given a single realization of flow's willingness to avoid facilities,* $\zeta_f(\omega^R)$*. This case represents deterministic EFCP. In such setting, constraint* (3.16) *becomes redundant, so the deterministic EFCP includes:*

- $\underbrace{|A|}_{x_{ij}} + \underbrace{|F|}_{y_f(\omega)} + \underbrace{2 \cdot \sum_{f \in F} \left|P_f(\omega^R)\right|}_{y_f^p(\omega) \ \& \ z_f^p(\omega)}$ *binary variables;*
- $\underbrace{1}_{(3.10)} + \underbrace{4 \cdot \sum_{f \in F} \left|P_f(\omega^R)\right|}_{(3.12)-(3.15)} + \underbrace{|F|}_{(3.17)}$ *constraints.*

The above remarks imply that two-stage stochastic EFCP includes more flow-based variables and constraints (i.e. $y_f(\omega)$ and (3.16)-(3.17)). However, the number of most numerous path-based variables and constraints is the same in both problems. This makes the two-stage stochastic EFCP only slightly more difficult than deterministic problem, provided that Theorem 4 and reformulation (3.9)-(3.17) are applied.

3.4 The Value of the Stochastic Solution

Let VSS denote the value of the stochastic solution, which represents the benefit from solving the two-stage stochastic EFCP over solving its deterministic counterpart in which random parameters are replaced with their expected values [48]. While recalling formulation **SP1** and result from Theorem 4, we can formally define VSS for the two-stage stochastic EFCP as

$$VSS = \mathbb{E}_\zeta Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta) - \mathbb{E}_\zeta Q(\boldsymbol{x}^*, \bar{\xi}, \zeta),$$

where

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}\in\{0,1\}^{|A|},\ \sum_{(i,j)\in A} x_{ij}\leq m}{\arg\min} \quad \mathbb{E}_\zeta Q(\boldsymbol{x}, \bar{\xi}, \zeta)$$

$$\bar{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\{0,1\}^{|A|},\ \sum_{(i,j)\in A} x_{ij}\leq m}{\arg\min} \quad Q(\boldsymbol{x}, \bar{\xi}, \bar{\zeta}).$$

The following proposition argues that one can design an instance of EFCP with an arbitrarily large VSS. Then, a sufficient condition for which $VSS = 0$ is provided. In Section 3.6, we numerically compute the VSS for case studies involving two real-world transportation networks and contrast $\bar{\boldsymbol{x}}$ with $\boldsymbol{x}^*$.

Proposition 8. *For any finite $\varepsilon > 0$, there exists an instance of the two-stage stochastic EFCP for which $VSS > \varepsilon$.*

Proof. Let $\varepsilon > 0$, and assume that $m = 1$. Now suppose there is a single flow that can travel along two edge-disjoint paths. Let l denote the length of the shorter, and $\gamma \cdot l$ be the length of the longer path ($\gamma > 1$). Furthermore, let ζ denote the maximum distance that a flow is willing to travel to avoid facilities. Assume that ζ has two possible realizations, $\mathbb{P}(\zeta = \gamma \cdot l) = \delta$ and $\mathbb{P}(\zeta = l) = 1 - \delta$, where $\delta < 1$.

Since $\bar{\zeta} < \gamma \cdot l$, in the deterministic counterpart of EFCP, the flow f can travel only along the shorter path. Thus, $\bar{\boldsymbol{x}}$ implies implementation of a facility anywhere along this path. The corresponding expected cost is computed as $\mathbb{E}_\zeta Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta) = (1-\delta) \cdot 0 + \delta \cdot (\gamma \cdot l \cdot \bar{\xi}) = \delta \cdot \gamma \cdot l \cdot \bar{\xi}$, where $\bar{\xi}$ is the expected unit damage (i.e. per mile damage). On the other hand, x^* implies implementation of the facility along the shorter path if $\delta \cdot \gamma \leq 1$, or along the longer path if $\delta \cdot \gamma > 1$. Assume that $\delta \cdot \gamma > 1$ and note that the expected cost for the corresponding optimal solution is

$\mathbb{E}_{\zeta} Q(\boldsymbol{x}^*, \bar{\xi}, \zeta) = l \cdot \bar{\xi}$.

In the afore-described case, the value of the stochastic solution is given as

$$VSS = \mathbb{E}_{\zeta} Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta) - \mathbb{E}_{\zeta} Q(\boldsymbol{x}^*, \bar{\xi}, \zeta)$$
$$= l \cdot \bar{\xi} \cdot (\delta \cdot \gamma - 1).$$

Finally, note that we can define parameters l, $\bar{\xi}$, δ and γ (such that $\delta < 1$ and $\delta \cdot \gamma > 1$) to make VSS in the above example arbitrarily large. Thus, for any finite $\varepsilon > 0$, we can design an instance of the two-stage stochastic EFCP such that $VSS > \varepsilon$. □

Remark 5. *Given the network topology and willingness of flows to evade facilities, $VSS = 0$ if realizations of $\boldsymbol{\zeta}$ are such that $|P_f(\omega)| = 1$ for all $\omega \in \Omega$. This result follows from the definition of VSS, as well as observation that $\mathbb{E}_{\zeta} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}) = Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\zeta}})$ when $|P_f(\omega)| = 1$ for all $\omega \in \Omega$.*

3.5 Solution Techniques

Formulation (3.9)-(3.17) is only slightly more difficult than the deterministic EFCP which was efficiently solved with a mathematical programming software for the real-world transportation networks in Chapter 2. Thus, formulation (3.9)-(3.17) can also be tackled in the extensive form (i.e. as a binary integer program defined over all the scenarios) using similar software packages. The following section argues that many binary variables in (3.9)-(3.17) can be linearly relaxed, which considerably reduces the solution time. These results are based on relaxations proposed in Chapter 2 and their proofs are thus omitted for brevity. Section 3.5.2 discusses application of the integer L-shaped method to Problem **SP1**.

3.5.1 Partial Linear Relaxations

This section presents two partial linear relaxations that do not alter the optimal solution for **SP2** and the corresponding objective function value. The computational benefits of these partial linear relaxations are explored in Section 3.6.

Remark 6. *The partial linear relaxation stated in Theorem 2 is applicable to* ***SP2****, where it reduces the number of binary integer variables from* $|A| + \sum_{f\in F} M_f + 2\cdot\sum_{f\in F}\left|P_f(\omega^R)\right|$ *to* $|A| + \sum_{f\in F}\left|P_f(\omega^R)\right|$.

Remark 7. *The partial linear relaxation stated in Theorem 3 is applicable to* ***SP2****, where it reduces the number of binary integer variables from* $|A| + \sum_{f\in F} M_f + 2\cdot\sum_{f\in F}\left|P_f(\omega^R)\right|$ *to* $|A|$*. However, the total number of constraints is increased from* $1 + 4\cdot\sum_{f\in F}\left|P_f(\omega^R)\right| + 2\cdot\sum_{f\in F} M_f$ *to* $1 + 3\cdot\sum_{f\in F}\left|P_f(\omega^R)\right| + 2\cdot\sum_{f\in F} M_f + \sum_{f\in F}\sum_{p\in P_f(\omega^R)}\left|A_f^p(\omega^R)\right|$.

3.5.2 Integer L-shaped Method

The integer L-shaped method [49] is a standard procedure for solving two-stage stochastic programs with binary first-stage variables. It represents a branch-and-cut algorithm that can be readily applied to Problem **SP1**. This solution method approximates $\mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}}Q(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{\zeta})$ with the variable θ and a set of cuts. At a given stage of the method, we consider the so called current problem (CP) which is defined as a linear program:

$$\min_{0\leq x_{ij}\leq 1,\, \theta\geq 0} \quad \theta \tag{3.18}$$

$$\text{s.t.} \quad \sum_{(i,j)\in A} x_{ij} \leq m \tag{3.19}$$

$$\theta \geq \theta_k \left(\sum_{(i,j)\in S_k} x_{ij} - \sum_{(i,j)\notin S_k} x_{ij} - |S_k| + 1 \right), \quad k = 1, ..., t \tag{3.20}$$

where (3.20) represent optimality cuts that are iteratively added to the current problem when feasible solutions are found. Set S_k is defined so that for the k-th feasible solution, $x_{ij} = 1$ for $(i,j) \in S_k$ and $x_{ij} = 0$ for $(i,j) \notin S_k$. Moreover, θ_k represents the expected recourse for the k-th feasible solution. The outline of the procedure is given as Algorithm 2.

It should be noted that the optimality cut is obtained from [49] for $L = 0$, which is a lower bound on $\mathbb{E}_{\boldsymbol{\xi}\boldsymbol{\zeta}}Q(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{\zeta})$, as shown later on in Lemma 2. Additionally, [49] have shown that their procedure finds an optimal solution in a finite number of steps. In the following theorem we state this result together with the necessary conditions.

Theorem 5 (Laporte and Louveaux, 1993)**.** *The integer L-shaped method finds an optimal solution in a finite number of steps if the following conditions are satisfied:*

1. *For fixed* $\boldsymbol{x}$, $\mathbb{E}_{\boldsymbol{\zeta}}Q(\boldsymbol{x}, \boldsymbol{\zeta})$ *is computable in a finite number of steps.*

2. *There exists a finite lower bound* L*, such that* $\min_{\boldsymbol{x}}\{\mathbb{E}_{\boldsymbol{\zeta}}Q(\boldsymbol{x}, \boldsymbol{\zeta}) \quad s.t. \quad Ax = b, x \in X\} \geq L$

3. *Two-stage stochastic program has complete recourse*

Algorithm 2 Integer L-Shaped Method for **SP1**

Step 0: Set $t = \nu = 0$, $\bar{z} = +\infty, \theta = 0$. The only pendant node corresponds to the initial CP.

Step 1: Select the pendant node from the list; if none exists, stop.

Step 2: Set $\nu = \nu + 1$; solve the CP. If the CP has no feasible solution, fathom the current node; go to Step 1. Otherwise, let $(\boldsymbol{x}^{\nu}, \theta^{\nu})$ be an optimal solution.

Step 3: Check for integrality restrictions in $\boldsymbol{x}^{\nu}$. If one is violated, create two new branches; append the new nodes to the list of pendant nodes; return to Step 1.

Step 4: Compute $\mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}^{\nu}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$ by running a simple algorithm for $\forall \omega \in \Omega$ and let $z^{\nu} = \mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}^{\nu}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$. If $z^{\nu} < \bar{z}$, then update $\bar{z} = z^{\nu}$.

Step 5: If $\theta^{\nu} \geq \mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}^{\nu}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$, then fathom the current node and return to Step 1. Otherwise, impose an optimality cut (3.20) where $\theta_k = \mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}^{\nu}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$, set $t = t + 1$ and return to Step 2.

Corollary 1. *Integer L-shaped method finds an optimal solution for **SP1** in a finite number of steps. This follows from Lemma 1, 2, and 3.*

Lemma 1. *For fixed $\boldsymbol{x}$, $\mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$ is computable in polynomial time.*

Proof. For fixed $\boldsymbol{x}$, a simple algorithm (which assigns unintercepted flows to their shortest unmonitored paths) finds $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}(\omega))$ in a number of steps that is bounded from above by $O(|F| \cdot |P_f(\omega)| \cdot \left|A_f^p(\omega)\right|)$. Furthermore, $\mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$ is computed by finding $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}(\omega))$ for all $\omega \in \Omega$ and taking the average, which is done in polynomial time. □

Lemma 2. *There exists a finite lower bound L, such that*

$$\min_{\boldsymbol{x} \in \{0,1\}^{|A|}} \left\{ \mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}) \right\} \geq L.$$

Proof. Recall that $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}(\omega)) = \min\{\sum_{f \in F} \sum_{p \in P_f} z_f^p(\omega) c_f^p(\omega) \text{ s.t. } (3.4) - (3.8)\}$ and note that $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}(\omega)) \geq 0$ for any $\boldsymbol{x}$ because $c_f^p(\omega) \geq 0$ and $z_f^p(\omega) \in \{0, 1\}$. Taking the expectation we get $\mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}) \geq 0$ for any $\boldsymbol{x}$. Since this inequality holds for any $\boldsymbol{x}$, it holds for $\boldsymbol{x}$ that minimizes $\mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$, i.e. $\min_{\boldsymbol{x} \in \{0,1\}^{|A|}} \left\{ \mathbb{E}_{\boldsymbol{\zeta}} Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}) \right\} \geq 0$. This implies that for $L = 0$, the previous inequality holds. □

Lemma 3. *Problem **SP2** has complete recourse.*

Proof. For any fixed $\boldsymbol{x}$, a simple algorithm (which assigns unintercepted flows to their shortest unmonitored paths) finds $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}(\omega))$. This implies that $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta}(\omega))$ is feasible for any $\boldsymbol{x}$. Moreover, $Q(\boldsymbol{x}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$ is bounded from above by

$$\max_{\omega \in \Omega} \left\{ \sum_{f \in F} \max_{p \in P_f(\omega)} \left\{ c_f^p(\omega) \right\} \right\} < +\infty.$$

□

3.6 Numerical Examples

The road networks of Nevada and Vermont are used to explore the performance of the proposed solution techniques, numerically compute VSS, and contrast $\boldsymbol{x}^*$ with $\bar{\boldsymbol{x}}$. The relevant data are extracted from Matlog [50], which contains the Oak Ridge National Highway Network [51]. Since many of the observed road links are non-separated, it is assumed that $x_{ij} = x_{ji}$ as in an undirected graph. Hence, the observed road networks include 130 edges for Nevada, and 178 edges for Vermont.

Two hundred flows are randomly simulated, all with the same expected intensity of $\bar{\xi}_f = 200$ units/mile. Moreover, ζ_f is assumed to have three equally likely realizations, $\zeta_f(\omega) \in \{1, 1.1, 1.2\}$. The first realization, $\zeta_f(\omega) = 1$, corresponds to the case when the flow is willing to travel only along its shortest path. Second, $\zeta_f(\omega) = 1.1$, implies that the flow would be willing to travel an additional 10% of its shortest path to bypass the facilities. Similarly, when $\zeta_f(\omega) = 1.2$, the flow would travel an extra 20% to bypass the facilities. The k-shortest path algorithm [52] is used to find the necessary number of paths for each flow.

The following two problem are solved: 1) the deterministic counterpart of the stochastic EFCP (i.e. $\bar{\zeta}_f = 1.1$) and 2) the two-stage stochastic EFCP. Tables 3.1 and 3.2 contrast $\bar{\boldsymbol{x}}$ with $\boldsymbol{x}^*$ for the different number of facilities. Moreover, $\bar{\boldsymbol{x}}$ is evaluated over the three scenarios, $\mathbb{E}_{\boldsymbol{\zeta}} Q(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \boldsymbol{\zeta})$, and the VSS is computed. The last column of Tables 3.1 and 3.2 indicates that cost reductions achieved by solving the stochastic EFCP ranges between 0% and 100%, with the average reduction of 15.5%. Moreover, the $VSS > 0$ in 62% of the considered instances, which certainly motivates the application of the two-stage EFCP as opposed to just solving its deterministic counterpart. To show the differences that arise between the two solutions, we graphically contrast some of the $\bar{\boldsymbol{x}}$ and $\boldsymbol{x}^*$ in Figures 3.1-3.4.

Table 3.3 compares the performance of solution techniques. The partial linear relaxation proposed in Theorem 2 reduced the computation time in 35/37 instances by a median of 19% (SP vs. SP^1_{LR}). In Section 3.3 it was argued that the stochastic EFCP is only slightly more difficult than the deterministic EFCP for $\zeta_f(\omega^R)$, provided that Theorem 4 and reformulation (3.9)-(3.17) are applied. Table 3.3 indicates that computation times for the two cases are fairly similar (SP^1_{LR} vs. $EFCP^1_{LR}$ for $\zeta_f(\omega^R)$). In 16 instances the two-stage EFCP took more time, with a median overhead of 37%. In 13 instances the two-stage EFCP was solved more efficiently, with a median of 20% less computation time. Finally, in 8 instances the difference in solution times was within 1 second.

The integer L-shaped method was implemented in C++ and applied to same problem instances. Since it did not find an optimal solution in the vast majority of instances within the 4 hour time limit, the integer L-shaped was omitted from Table 3.3. Moreover, the partial linear relaxation proposed in Theorem 3 includes an increased number of constraints, which led to longer computation times. Thus, the corresponding computation times were omitted from Table 3.3.

Table 3.1: Computing the VSS for the Road Network of Nevada

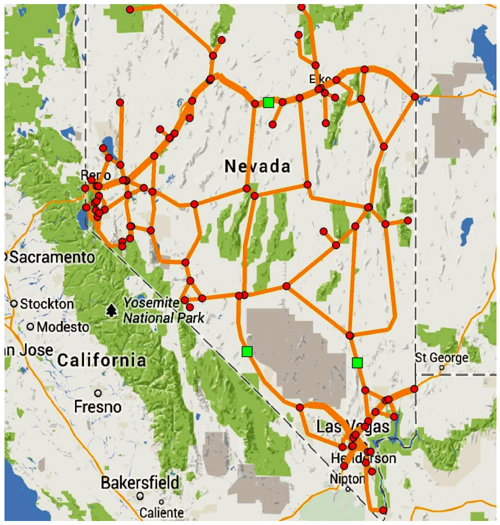
m	$\bar{\boldsymbol{x}}$	$\mathbb{E}_\zeta Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta)$	$\boldsymbol{x}^*$	$\mathbb{E}_\zeta Q(\boldsymbol{x}^*, \bar{\xi}, \zeta)$	$\frac{100 \cdot VSS}{\mathbb{E}_\zeta Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta)}$
1	79	8,453,873	79	8,453,873	0.00
2	21, 33	4,755,020	21, 33	4,755,020	0.00
3	21, 33, 66	3,617,553	21, 33, 88	3,439,407	4.92
4	33, 62, 67, 79	2,118,247	33, 69, 77, 79	2,090,827	1.29
5	33, 62, 67, 79, 88	1,445,200	33, 68, 69, 79, 83	1,297,647	10.21
6	28, 33, 62, 67, 79, 88	1,224,200	28, 33, 68, 69, 79, 83	899,860	26.49
7	21, 28, 33, 62, 67, 84, 88	689,373	21, 28, 33, 66, 67, 84, 88	671,767	2.55
8	21, 28, 33, 62, 67, 84, 88, 103	509,300	21, 28, 33, 62, 67, 84, 88, 103	509,300	0.00
9	21, 29, 33, 42, 63, 67, 84, 88, 103	381,720	24, 28, 33, 63, 68, 79, 88, 93, 97	355,680	6.82
10	24, 29, 33, 38, 63, 68, 79, 88, 93, 97	256,540	24, 28, 33, 38, 63, 68, 79, 88, 93, 97	256,540	0.00
11	16, 21, 29, 33, 63, 64, 70, 81, 84, 88, 103	210,193	16, 21, 29, 33, 63, 64, 70, 81, 84, 88, 103	210,193	0.00
12	16, 20, 21, 29, 33, 63, 64, 77, 81, 84, 88, 103	175,127	16, 20, 22, 29, 33, 63, 64, 77, 81, 84, 88, 103	153,273	12.48
13	16, 20, 22, 29, 33, 63, 64, 70. 81, 84, 88, 103, 112	102,633	16, 20, 22, 29, 33, 63, 64, 70, 81, 84, 88, 103, 112	102,633	0.00
14	16, 20, 22, 29, 33, 38, 63, 64, 77, 81, 84, 88, 103, 112	74,233	16, 20, 22, 29, 33, 63, 64, 70, 81, 84, 88, 101, 104, 112	71,280	3.98
15	16, 20, 21, 29, 33, 38, 63, 64, 77, 81, 84, 88, 95, 103, 112	60,153	16, 20, 22, 29, 33, 38, 63, 64, 70, 81, 84, 88, 101, 104, 112	42,880	28.72
16	14, 20, 22, 29, 33, 38, 52, 63, 70, 79, 88, 93, 97, 101, 102, 112	83,767	16, 21, 27, 29, 31, 33, 53, 63, 64, 70, 81, 84, 88, 101, 104, 112	25,500	69.56
17	10, 16, 20, 29, 33, 35, 38, 62, 70, 71, 79, 88, 93, 97, 101, 102, 112	164,827	16, 21, 27, 29, 31, 33, 35, 63, 64, 70, 81, 84, 88, 95, 101, 104, 112	11,720	92.89
18	14, 16, 21, 27, 29, 31, 33, 35, 37, 62, 70, 79, 88,	140,993	16, 21, 29, 30, 33, 37, 38, 53, 63, 64, 70, 81, 84,	0	100.0
	93, 97, 101, 102, 112		88, 95, 101, 104, 112		

Note: The last column indicates a percentage reduction achieved by solving the stochastic EFCP.

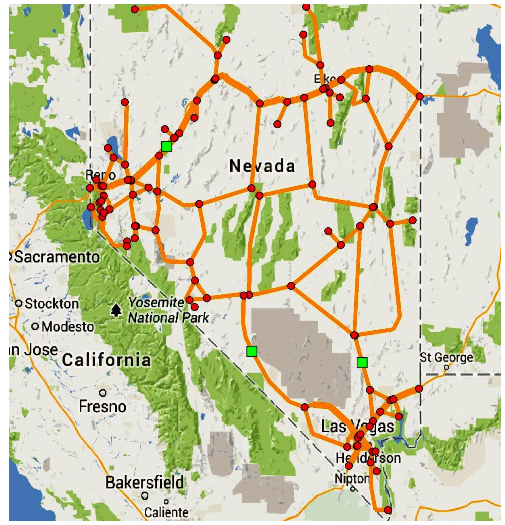
Table 3.2: Computing the VSS for the Road Network of Vermont

m	$\bar{\boldsymbol{x}}$	$\mathbb{E}_\zeta Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta)$	$\boldsymbol{x}^*$	$\mathbb{E}_\zeta Q(\boldsymbol{x}^*, \bar{\xi}, \zeta)$	$\frac{100 \cdot VSS}{\mathbb{E}_\zeta Q(\bar{\boldsymbol{x}}, \bar{\xi}, \zeta)}$
1	107	3,553,473	45	3,434,947	3.34
2	26, 88	3,069,940	8, 45	3,051,600	0.60
3	36, 82, 115	2,394,960	36, 82, 115	2,394,960	0.00
4	36, 82, 88, 115	1,935,313	36, 82, 88, 115	1,935,313	0.00
5	26, 37, 82, 88, 115	1,071,427	26, 37, 82, 88, 115	1,071,427	0.00
6	26, 37, 82, 88, 107, 115	938,660	26, 37, 45, 82, 88, 115	931,473	0.77
7	26, 37, 82, 88, 95, 115, 146	780,660	26, 37, 82, 88, 95, 115, 146	780,660	0.00
8	34, 45, 58, 87, 88, 95, 97, 146	657,773	34, 43, 45, 58, 87, 88, 95, 97	633,800	3.64
9	34, 45, 58, 77, 87, 88, 95, 97, 146	532,087	34, 43, 45, 58, 77, 87, 88, 95, 97	508,113	4.51
10	34, 42, 45, 58, 63, 77, 87, 88, 97, 111	371,613	34, 42, 45, 58, 63, 77, 87, 88, 97, 111	371,613	0.00
11	34, 42, 45, 58, 80, 88, 95, 97, 107, 114, 146	277,067	34, 42, 45, 58, 80, 87, 88, 95, 97, 114, 146	271,200	2.12
12	34, 42, 45, 58, 80, 88, 95, 97, 106, 107, 114, 146	197,247	34, 42, 45, 58, 80, 88, 95, 97, 107, 114, 131, 146	194,427	1.43
13	27, 38, 42, 45, 58, 80, 88, 95, 97, 107, 114, 131, 146	123,320	27, 38, 42, 45, 58, 80, 88, 95, 97, 107, 114, 131, 146	123,320	0.00
14	27, 38, 42, 45, 58, 80, 87, 88, 95, 97, 108, 114, 127, 146	85,440	27, 38, 42, 45, 58, 80, 87, 88, 95, 97, 108, 114, 127, 146	85,440	0.00
15	27, 38, 42, 45, 58, 80, 87, 88, 95, 97, 108, 114, 127, 146, 164	52,000	8, 27, 38, 42, 45, 58, 80, 87, 88, 95, 97, 108, 114, 127, 146	52,000	0.00
16	27, 38, 42, 44, 45, 58, 80, 87, 88, 95, 97, 108, 114, 127, 146, 164	41,080	8, 26, 37, 38, 45, 55, 59, 68, 82, 87, 88, 95, 108, 115, 127, 147	30,847	24.91
17	8, 27, 38, 41, 44, 45, 58, 80, 87, 88, 95, 97, 108, 114, 129, 130, 146	25,220	8, 26, 37, 38, 45, 55, 59, 68, 82, 87, 88, 95, 108, 115, 129, 130, 147	20,107	20.27
18	27, 38, 42, 44, 45, 58, 80, 87, 88, 95, 97, 108, 114,	19,820	8, 26, 37, 38, 45, 55, 59, 68, 82, 87, 88, 95, 108,	9,587	51.63
	129, 138, 146, 164, 175		115, 129, 138, 147, 170		
19	26, 35, 36, 38, 41, 44, 45, 82, 85, 88, 93, 95, 96	32,127	26, 37, 38, 41, 45, 55, 59, 68, 82, 87, 88, 104, 108,	0	100.0
	105, 114, 115, 129, 130, 164		112, 115, 129, 130, 138, 147		

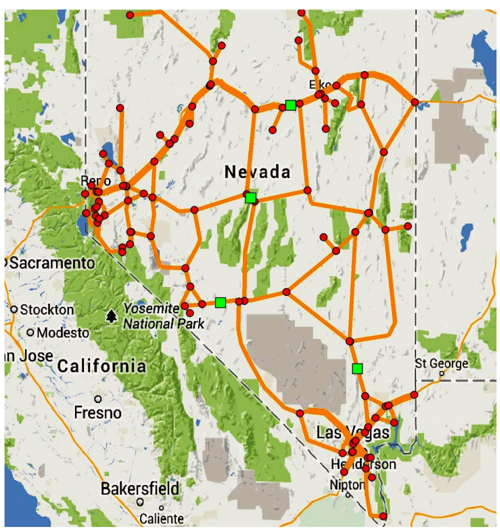
Note: The last column indicates a percentage reduction achieved by solving the stochastic EFCP.



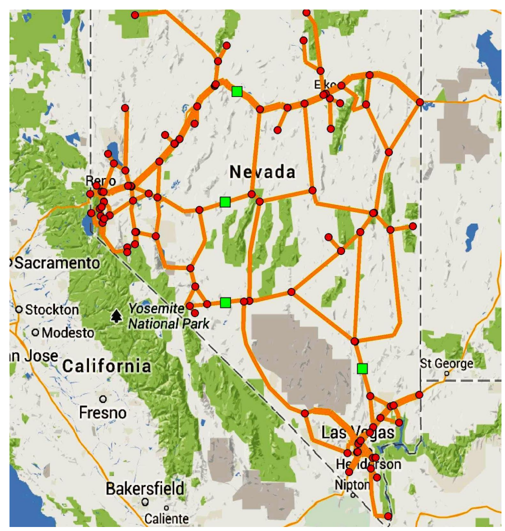
(a) $\bar{\boldsymbol{x}}$ for $m = 3$



(b) $\boldsymbol{x}^*$ for $m = 3$

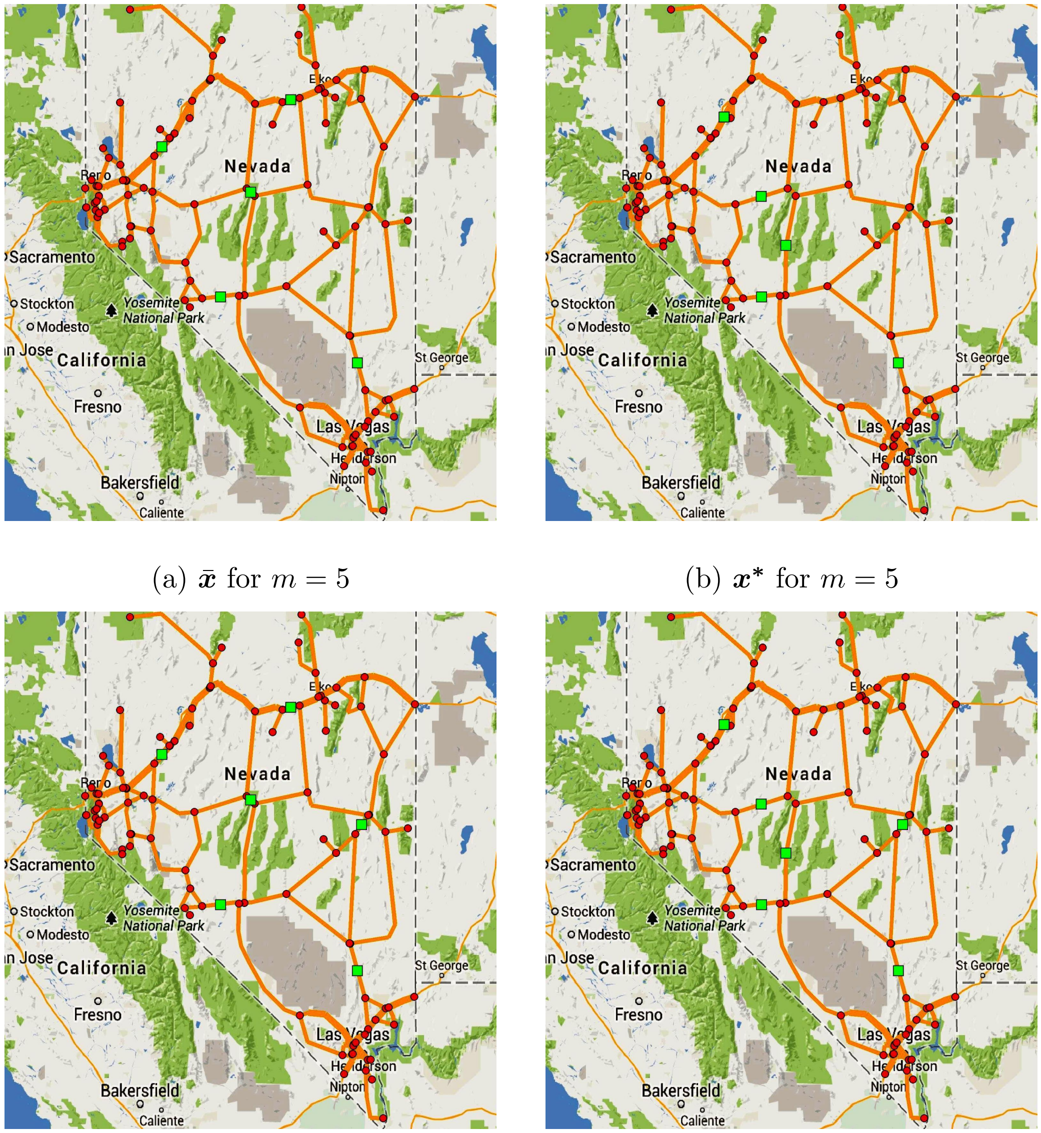


(c) $\bar{\boldsymbol{x}}$ for $m = 4$



(d) $\boldsymbol{x}^*$ for $m = 4$

Figure 3.1: Nevada: comparison of stochastic and deterministic solutions for $m = 3, 4$.



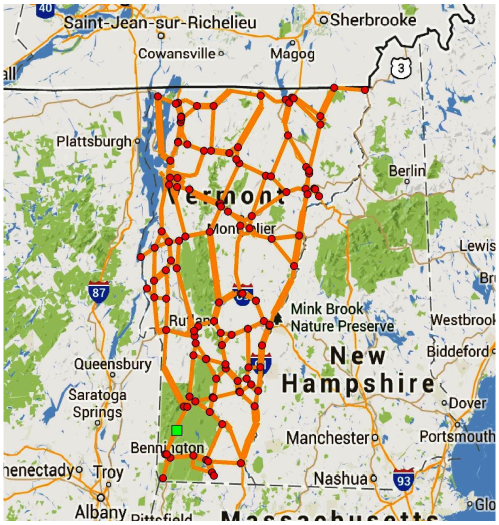
(a) $\bar{\boldsymbol{x}}$ for $m = 5$

(b) $\boldsymbol{x}^*$ for $m = 5$

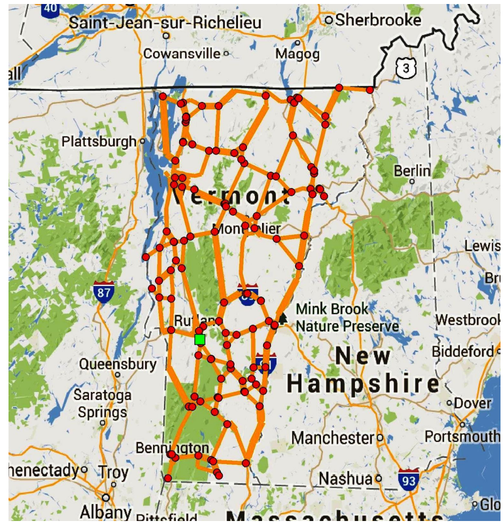
(c) $\bar{\boldsymbol{x}}$ for $m = 6$

(d) $\boldsymbol{x}^*$ for $m = 6$

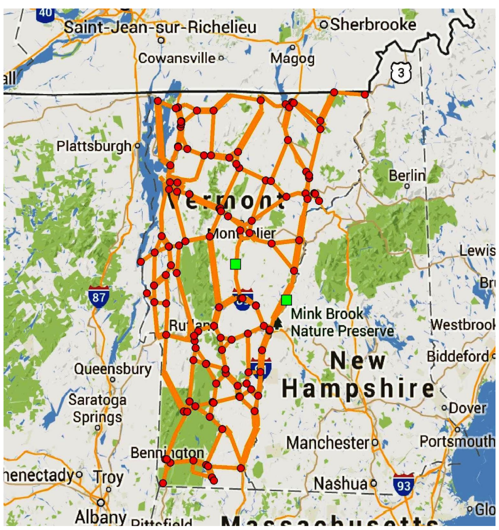
Figure 3.2: Nevada: comparison of stochastic and deterministic solutions for $m = 5, 6$.



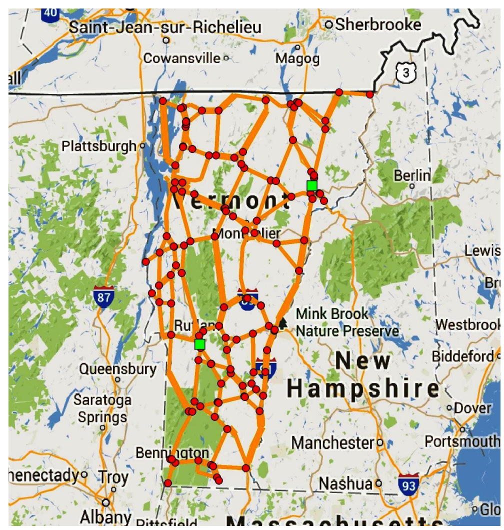
(a) $\bar{\boldsymbol{x}}$ for $m = 1$



(b) $\boldsymbol{x}^*$ for $m = 1$

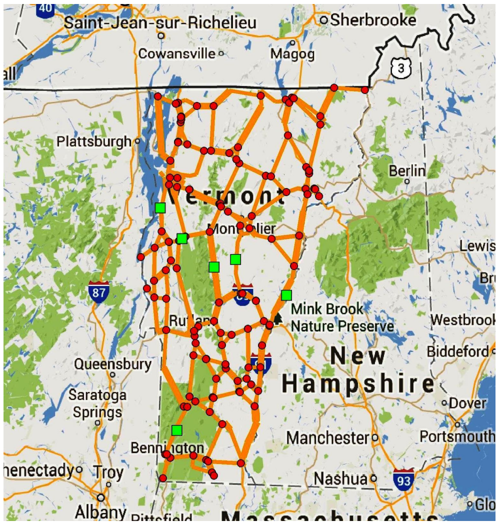


(c) $\bar{\boldsymbol{x}}$ for $m = 2$

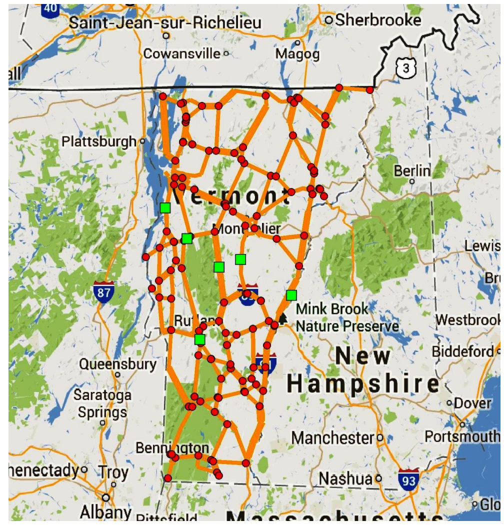


(d) $\boldsymbol{x}^*$ for $m = 2$

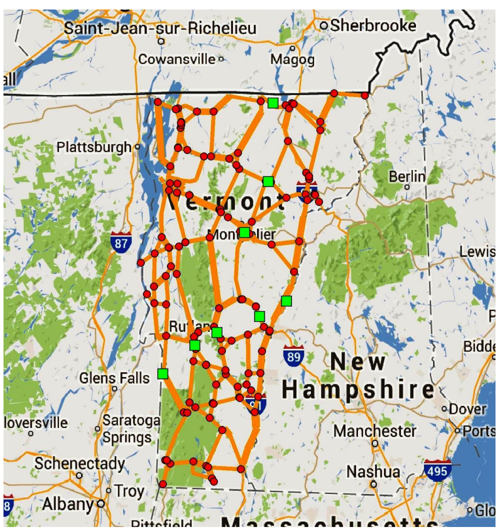
Figure 3.3: Vermont: comparison of stochastic and deterministic solutions for $m = 1, 2$



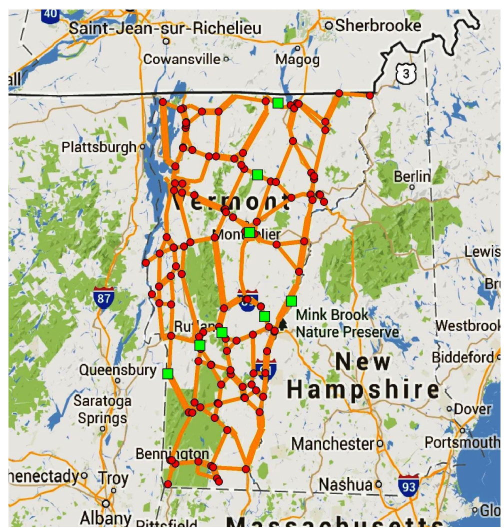
(a) $\bar{\boldsymbol{x}}$ for $m = 6$



(b) $\boldsymbol{x}^*$ for $m = 6$



(c) $\bar{\boldsymbol{x}}$ for $m = 8$



(d) $\boldsymbol{x}^*$ for $m = 8$

Figure 3.4: Vermont: comparison of stochastic and deterministic solutions for $m = 6, 8$

Table 3.3: Computation Times: Comparison with the Deterministic EFCP

m	Nevada			Vermont		
	SP (sec)	SP^1_{LR} (sec)	$EFCP^1_{LR}$ for $\zeta_f(\omega^R)$ (sec)	SP (sec)	SP^1_{LR} (sec)	$EFCP^1_{LR}$ for $\zeta_f(\omega^R)$ (sec)
1	84	59	60	760	231	121
2	74	63	62	4983	13134	3738
3	96	144	92	7073	3949	2643
4	80	61	247	18945	7926	2454
5	74	60	56	835	709	662
6	72	59	54	1125	899	1079
7	65	62	56	1197	827	744
8	114	88	53	1513	940	619
9	63	51	54	3273	1160	915
10	60	50	48	1716	1352	1239
11	90	77	77	981	868	593
12	66	52	86	1551	1038	1306
13	60	49	59	614	564	609
14	59	48	49	628	558	605
15	60	49	49	706	521	651
16	58	48	48	713	615	961
17	58	47	47	799	1206	1226
18	59	48	47	1175	1034	1096
19				605	472	545

3.7 Conclusions

This chapter presented a stochastic extension of EFCP where intensities of flows and their willingness to avoid facilities are characterized with random distributions. The two-stage stochastic EFCP is made computationally tractable by exploiting the structural properties of the problem. This is achieved by 1) reducing the noise associated with the intensities of the flows, 2) reformulating the second stage recursively, and 3) linearly relaxing many scenario-based binary variables. The proposed approach yields an instance which is only slightly more difficult than the deterministic EFCP and is thus crucial in efficiently solving the stochastic EFCP. This point is emphasized through application of the standard solution method, the integer L-shaped, which is not capable of finding an optimal solutions to real-world problems in a reasonable amount of time.

The proposed stochastic EFCP is tested on case studies involving real-world transportation networks, which shows the applicability of the model and solution methods. Moreover, the stochastic EFCP is contrasted with its deterministic counterpart in which all random parameters are replaced with their expected values. This comparison showed that solving the stochastic model added considerable value, as it reduced the cost of the deterministic solution by more than 15% on average. These results show the relevance of the proposed two-stage stochastic EFCP and motivate its application.

Chapter 4: Multi-Stage Stochastic EFCP

Suppose that decisions about the implementation of facilities are made at different time points (e.g. biannually) given probabilistic information about the flows which varies over time. This resembles a realistic long-term investment planning during which intensities of the flows typically increase over time (e.g. the expected number of heavy trucks increases 2% annually). The resulting model for optimal decision making is then a multi-stage stochastic optimization model [53]. Let $\boldsymbol{\xi}^t = \{\xi_f^t \mid f \in F\}$ be a vector of random variables denoting intensities of flows $f \in F$ in stage $t \in T$ (i.e. damage or risk produced per distance traveled). Similarly, let $\boldsymbol{\zeta}^t = \{\zeta_f^t \mid f \in F\}$ be a vector of discrete random variables denoting the willingness of flows to evade facilities in stage $t \in T$.

Random vectors $\boldsymbol{\xi}^t$ and $\boldsymbol{\zeta}^t$ are assumed to be independent from one another. Moreover, they are assumed to be independent from $\boldsymbol{\xi}^{t-1}$ and $\boldsymbol{\zeta}^{t-1}$. In Chapter 3, it was shown that random vector $\boldsymbol{\xi}$ denoting intensities of flows can be replaced with $\bar{\boldsymbol{\xi}}$ without altering the two-stage stochastic EFCP (Theorem 4). This result can be extended to the multi-stage setting and vector $\boldsymbol{\xi}^t$, given the aforementioned independence assumption.

The remainder of this section is organized as follows. Section 4.1 presents a mathematical programming formulation of the multi-stage stochastic EFCP which can be solved optimally for the real-world transportation networks using the partial linear relaxations given in Section 4.1.1. Section 4.2 reformulates the problem as a dynamic program and proposes an approximate dynamic programming approach which can be used to tackle the multi-stage EFCP more efficiently, but without guaranteeing an optimal solution. This approximate solution technique can also be used for problems that are intractable with mathematical programming techniques. The proposed solution techniques are tested on the real-world road networks in Section 4.3. Finally, Section 4.4 draws the conclusions.

4.1 Mathematical Programming

Let $G(N, A)$ be a bidirectional road transportation network, where N is a set of nodes and A is a set of arcs (i, j). Define $P_f^t(\omega)$ as a set of paths which contains $k_f^t(\omega)$ shortest paths of the flow $f \in F$. Let $A_f^{pt}(\omega)$ be the set of links along path $p \in P_f^t(\omega)$ of flow $f \in F$. Additionally, let w_{ij}^t denote the cost of implementing a facility at arc (i, j) in period $t \in T$ and maintaining it during its life duration. Let $c_f^{pt}(\omega)$ be the excessive damage cost if flow $f \in F$ travels freely along path $p \in P_f^t(\omega)$ in period $t \in T$. Let x_{ij}^t be a binary variable equal to 1 if a facility is located at arc (i, j) in period $t \in T$ and 0 otherwise. The life expectancy of a facility is L time periods. Moreover, we define $\boldsymbol{x}^t = \{x_{ij}^t \mid (i, j) \in A\}$ and $\boldsymbol{w}^t = \{w_{ij}^t \mid (i, j) \in A\}$ as vectors of $|A|$ elements. Let b^t be the budget allocated for stage $t \in T$, which can be either spent or carried over to the next stage. Denote by u^t the total investment budget available at stage $t \in T$.

The multi-stage stochastic EFCP is defined as a minimization problem

$$\textbf{MSP1}: \quad \min_{\substack{\boldsymbol{x}^t \in \{0,1\}^{|A|} \\ u^t \geq 0}} \quad \sum_{t \in T} \left(\sum_{(i,j) \in A} w_{ij}^t x_{ij}^t + \mathbb{E}_{\boldsymbol{\zeta}^t} Q^t \left(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ..., \boldsymbol{x}^{t-L+1}, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t \right) \right)$$

$$\text{s.t.} \sum_{(i,j) \in A} w_{ij}^t x_{ij}^t \leq u^t \qquad \forall t \in T$$

$$u^t = u^{t-1} - \sum_{(i,j) \in A} w_{ij}^{t-1} x_{ij}^{t-1} + b^t \qquad \forall t \in T$$

where $Q^t \left(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ..., \boldsymbol{x}^{t-L+1}, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t(\omega) \right)$ is an oracle that, given an allocation of checkpoints in the current as well as previous $L - 1$ stages, computes the excessive damage (or risk) associated with flows and a particular realization of $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$. If a flow is captured, then the corresponding damage is 0. Otherwise, the flow seeks to minimize its travel distance, and produces the damage by traveling along its shortest unmonitored path. More formally, if we let

$$P_f^{2,t}(\omega) = \left\{ p \in P_f^t(\omega) \mid \sum_{t'=t-L+1}^{t} \sum_{(i,j) \in A_f^{pt}(\omega)} x_{ij}^{t'} = 0 \right\}$$

be the set of paths of flow $f \in F$ not covered by facilities, then

$$Q^t \left(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ..., \boldsymbol{x}^{t-L+1}, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t(\omega) \right) = \sum_{f \in F} Q_f^t \left(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ..., \boldsymbol{x}^{t-L+1}, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t(\omega) \right)$$

where

$$Q_f^t \left(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ..., \boldsymbol{x}^{t-L+1}, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t(\omega) \right) = \begin{cases} \min_{p \in P_f^2(\omega)} \left\{ c_f^p(\omega) \right\}, & P_f^2(\omega) \neq \emptyset; \\ 0, & P_f^2(\omega) = \emptyset. \end{cases}$$

Problem **MSP1** can be linearized. First, four sets of auxiliary binary variables are introduced:

$$s_{ij}^t = \begin{cases} 1 \text{ if link } (i,j) \in A \text{ is equipped with facility in time period } t \in T \\ 0 \;\; \text{otherwise} \end{cases}$$

$$y_f^{pt}(\omega) = \begin{cases} 1 \text{ if a facility is located along path } p \in P_f^t(\omega) \text{ of } f \in F \text{ in } t \in T \\ 0 \;\; \text{otherwise} \end{cases}$$

$$y_f^t(\omega) = \begin{cases} 1 \text{ if a facility is located along all paths } p \in P_f^t(\omega) \text{ of } f \in F \text{ in } t \in T \\ 0 \text{ otherwise} \end{cases}$$

$$z_f^{pt}(\omega) = \begin{cases} 1 \text{ if flow } f \in F \text{ travels freely along path } p \in P_f^t(\omega) \text{ in period } t \in T \\ 0 \;\; \text{otherwise} \end{cases}$$

Second, some additional notation is introduced to formulate the path-based constraints recursively, as in Section 3.3. Let $\zeta_f^t(\omega^{rt})$ denote the r-th realization of the random willingness of flow $f \in F$ to avoid facilities in stage $t \in T$. Moreover, assume that realizations are ordered so that $\zeta_f^t(\omega^{rt}) \geq \zeta_f^t(\omega^{r-1,t})$, and thus $P_f^t(\omega^{r-1,t}) \subseteq P_f^t(\omega^{rt})$. We now formulate Problem **MSP2**, while assuming that $v^0 = 0$, $P_f^t(\omega_f^{0,t}) = \emptyset$ and $y_f(\omega_f^{0,t}) = 1$.

$$\min_{\substack{x_{ij}^t \in \{0,1\} \\ u^t \geq 0}} \sum_{t \in T} \sum_{(i,j) \in A} x_{ij}^t w_{ij}^t + \sum_{t \in T} \sum_{f \in F} \mathbb{E}_{\zeta_f^t} \tilde{Q}_f^t \left(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ..., \boldsymbol{x}^{t-L+1}, \bar{\xi}_f^t, \zeta_f^t \right) \tag{4.1}$$

$$\text{s.t.} \sum_{(i,j) \in A} w_{ij}^t x_{ij}^t \leq u^t \qquad \forall t \in T \tag{4.2}$$

$$u^t = u^{t-1} - \sum_{(i,j) \in A} w_{ij}^{t-1} x_{ij}^{t-1} + b^t \qquad \forall t \in T \tag{4.3}$$

$$\tilde{Q}_f^t(\boldsymbol{x}^t, \boldsymbol{x}^{t-1}, ., \boldsymbol{x}^{t-L+1}, \bar{\xi}_f^t, \zeta_f^t(\omega^{rt})) = \min_{\substack{s_{ij}^t y_f^{pt}(\omega^{rt}) \in \{0,1\} \\ y_f^t(\omega^{rt}), z_f^{pt}(\omega^{rt}) \in \{0,1\}}} \sum_{p \in P_f^t(\omega^{rt})} z_f^{pt}(\omega^{rt}) c_f^{pt}(\omega^{rt}) \quad (4.4)$$

$$\text{s.t. } s_{ij}^t \leq \sum_{t'=t-L+1}^{t} x_{ij}^{t'} \qquad \forall (i,j) \in A \quad (4.5)$$

$$\sum_{(i,j) \in A_f^{pt}(\omega^{rt})} s_{ij}^t \geq y_f^{pt}(\omega^{rt}) \qquad \forall p \in P_f^t(\omega^{rt}) \setminus P_f^t(\omega^{r-1,t}) \quad (4.6)$$

$$z_f^{pt}(\omega^{rt}) \leq 1 - y_f^{pt}(\omega^{rt}) \qquad \forall p \in P_f^t(\omega^{rt}) \setminus P_f^t(\omega^{r-1,t}) \quad (4.7)$$

$$\sum_{(i,j) \in A_f^{pt}(\omega^{rt})} s_{ij}^t \leq \left| A_f^{pt}(\omega^{rt}) \right| \cdot y_f^{pt}(\omega^{rt}) \qquad \forall p \in P_f^t(\omega^{rt}) \setminus P_f^t(\omega^{r-1,t}) \quad (4.8)$$

$$y_f^t(\omega^{rt}) \leq y_f^{pt}(\omega^{rt}) \qquad \forall p \in P_f^t(\omega^{kt}) \setminus P_f^t(\omega^{r-1,t}) \quad (4.9)$$

$$y_f^t(\omega^{rt}) \leq y_f(\omega^{r-1,t}) \quad (4.10)$$

$$\sum_{p \in \bigcup_{r'=1}^{r} P_f^t(\omega^{r',t})} z_f^{pt}(\omega^{rt}) \geq 1 - y_f^t(\omega^{rt}) \quad (4.11)$$

The objective (4.1) minimizes the investment cost and the expected excessive damage subject to budget constraints (4.2)-(4.3). Constraint (4.5) checks whether the allocation at stage $t \in T$ includes a facility implementation. All together, (4.4)-(4.11), model the same relations as in Section 3.3.

Since **MSP1** and **MSP2** are extensions of **SP1** and **SP2** from Chapter 3, it follows that:

1. Problems **MSP1** and **MSP2** are equivalent;
2. The objective of Problem **MSP1** is non-submodular, non-supermodular, and non-monotonic;
3. Problem **MSP2** is NP-hard;

4. A bound cannot be established on the greedy approximation of **MSP1**.

4.1.1 Partial Linear Relaxations

In Chapter 3 it was shown that, after applying Theorem 4 and reformulating the second stage, the two-stage stochastic EFCP is as difficult as the deterministic EFCP. Since Problem **MSP2** is a multi-stage extension of **SP2**, the multi-stage stochastic EFCP is as difficult as a multi-stage deterministic program. Thus, Problem **MSP2** can also be tackled in the extensive form with a mathematical programming software using branch-and-bound-based methods. The following section argues that many binary variables in (4.1)-(4.11) can be linearly relaxed, which considerably reduces the solution time. These results are based on relaxations proposed in Chapter 2 and their proofs are thus omitted for brevity.

Remark 8. *The partial linear relaxation stated in Theorem 2 is applicable to **MSP2**, where it reduces the number of binary integer variables from* $\underbrace{2 \cdot |A| \cdot |T|}_{x_{ij}^t \ \& \ s_{ij}^t} + \underbrace{\sum_{t \in T} \sum_{f \in F} M_f^t}_{y_f^t(\omega)}$ $+ \underbrace{2 \cdot \sum_{t \in T} \sum_{f \in F} \left| P_f^t(\omega^{Rt}) \right|}_{y_f^{pt}(\omega) \ \& \ z_f^{pt}(\omega)}$ *to* $2 \cdot |A| \cdot |T| + \sum_{t \in T} \sum_{f \in F} \left| P_f^t(\omega^{Rt}) \right|$.

Remark 9. *The partial linear relaxation stated in Theorem 3 is applicable to **MSP2**, where it reduces the number of binary integer variables from* $2 \cdot |A| \cdot |T| + \sum_{t \in T} \sum_{f \in F} M_f^t$ $+ 2 \cdot \sum_{t \in T} \sum_{f \in F} \left| P_f^t(\omega^{Rt}) \right|$ *to* $2 \cdot |A| \cdot |T|$. *However, the total number of constraints is increased from* $2 \cdot |T| + |T| \cdot |A| + 4 \cdot \sum_{t \in T} \sum_{f \in F} \left| P_f^t(\omega^{Rt}) \right| + 2 \cdot \sum_{t \in T} \sum_{f \in F} M_f^t$ *to* $2 \cdot |T| + |T| \cdot |A| + 3 \cdot \sum_{t \in T} \sum_{f \in F} \left| P_f^t(\omega^{Rt}) \right| + 2 \cdot \sum_{t \in T} \sum_{f \in F} M_f^t + \sum_{t \in T} \sum_{f \in F} \sum_{p \in P_f^t(\omega^{Rt})} \left| A_f^{pt}(\omega^{Rt}) \right|$.

4.2 Approximate Dynamic Programming

To reformulate the multi-stage stochastic EFCP as a dynamic program, some additional notation is introduced to characterize the state of the system. Note that a state of the system at time $t \in T$ is determined with the current allocation of facilities, their age, and available budget. Let S_{ij}^t denote the remaining lifespan of a facility on link $(i, j) \in A$ at the beginning of stage $t \in T$ (i.e. the number of stages until it expires). It is assumed that $S_{ij}^t = 0$ implies no facility implementation. Moreover, let $S^t = \{S_{ij}^t \mid (i, j) \in A\} || \{u^t\}$ denote the vector of $|A| + 1$ elements, which completely defines the state of the system. The system dynamics (i.e. transition from stage t to $t + 1$) can now be given as

$$S_{ij}^{t+1} = \begin{cases} \max\left(S_{ij}^t - 1, 0\right), & \text{if } x_{ij}^t = 0; \\ L - 1, & \text{if } x_{ij}^t = 1; \end{cases} \tag{4.12}$$

$$u^{t+1} = u^t - \sum_{(i,j)\in A} w_{ij}^t x_{ij}^t + b^{t+1}, \tag{4.13}$$

where L is the deterministic life span of a facility. Given S^t, the set of feasible decisions at stage t is

$$\mathcal{X}(S^t) = \left\{ x_{ij}^t \ \text{ s.t. } \sum_{(i,j)\in A} x_{ij}^t w_{ij}^t \leq u^t \right\} \tag{4.14}$$

The multi-stage stochastic EFCP is given as

$$\min_{\boldsymbol{x}^t} \mathbb{E}\left[\sum_{t\in T} \left((\boldsymbol{w^t})^T \boldsymbol{x^t} + Q^t(S^t, \boldsymbol{x}^t, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t) \right) \mid S^0 \right] \tag{4.15}$$

where Q^t is again an oracle that returns the damage associated with the uncaptured flows. This problem can be reformulated using Bellman's principle of optimality [54], as

$$V^t(S^t) = \min_{\boldsymbol{x}^t \in \mathcal{X}(S^t)} \left\{ (\boldsymbol{w^t})^T \boldsymbol{x^t} + \mathbb{E}_{\boldsymbol{\zeta}^t} Q^t(S^t, \boldsymbol{x}^t, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t) + \mathbb{E}\left[V^{t+1}(S^{t+1})\right] \right\}, \tag{4.16}$$

where the transition from one stage into another (e.g. ageing of checkpoints, budgets carried over to subsequent years) and action space are defined in (4.12), (4.13), and (4.14).

An issue that arises in solving Bellman's equation is the so called curse of dimensionality. Actually, in solving (4.16), three curses of dimensionality are typically encountered: state space, outcome space (expectation is over a vector of random variables), and action space [55]. Since an approach to efficiently deal with the outcome space was already proposed (Theorem 4 and reformulating the scenarios recursively), the approximate dynamic programming (ADP) [56] is used to overcome the issues of large state and action spaces.

In approximate dynamic programming, we replace the expected value function $\mathbb{E}\left[V^{t+1}(S^{t+1})\right]$ with an approximation, denoted $\hat{V}^{t+1}$, and solve the following problem

$$\tilde{V}^t(S^t) = \min_{\boldsymbol{x}^t \in \mathcal{X}(S^t)} \left\{ (\boldsymbol{w}^t)^T \boldsymbol{x}^t + \mathbb{E}_{\boldsymbol{\zeta}^t} Q^t(S^t, \boldsymbol{x}^t, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t) + \hat{V}^{t+1}(S^{t+1}) \right\} \tag{4.17}$$

Problem (4.17) is referred to as the subproblem. Starting with a set of value-function approximations $\hat{V}^{t+1}$ and an initial state vector S^t, we sequentially solve (4.17) for each $t \in T$ while moving forward in time. The information obtained while solving (4.17) is used to update and improve the value-function approximation $\hat{V}^t$. After the updating procedure, a new set of value-function approximations is obtained. Then, the subproblems are solved again using the new value-function approximations [57].

The following section discusses in greater detail the approximation and updating of the value function.

4.2.1 Approximating the Value Function

Linear regression is a very efficient way to approximate the downstream values of the objective function. In this approach, the objective function approximation is given as

$$\hat{V}^{t+1}(S^{t+1}) = \sum_{k=1}^{K} \theta_k^{t+1} \phi_k(S^{t+1}), \tag{4.18}$$

where ϕ denotes features of state S^{t+1} and θ^{t+1} are regression parameters. Every time the subproblem

$$\tilde{V}^t(S^t) = \min_{\boldsymbol{x}^t \in \mathcal{X}(S^t)} \left\{ (\boldsymbol{w}^t)^T \boldsymbol{x}^t + \mathbb{E}_{\boldsymbol{\zeta}^t} Q^t(S^t, \boldsymbol{x}^t, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t) + \sum_{k=1}^{K} \theta_k^{t+1} \phi_k(S^{t+1}) \right\} \tag{4.19}$$

is solved, an observation $\hat{v}^t$ is obtained and used together with $\phi(S^t)$ to recursively update the estimate of θ^t. This procedure is repeated iteratively, while adding more observations which improve the estimate of θ^t. In updating θ^t, more weight is put on recent observations. This dynamics is determined with the step size α^n. The outline of the procedure is given in Algorithm 3. It should be noted that Step 1.2 in Algorithm 3 applies the recursive least squares to update parameter θ^t. This procedure is described in [56].

The feature function $\phi(S^{t+1})$ in specified in three different ways. Let $D_{ij}^{t+1,a}$ be a "dummy" variable which equals 1 if $S_{ij}^{t+1} = a$ (i.e. the facility on link (i, j) is of age a in stage $t + 1$), and 0 otherwise. The following specifications of $\phi(S^{t+1})$ are considered:

1. Let $\phi(S^{t+1}) = \{1\} \parallel \left\{S^t_{ij}\right\}_{(i,j)\in A} \parallel \{u^{t+1}\}$ be the vector of $2 + |A|$ elements;

2. Let $\phi(S^{t+1}) = \{1\} \parallel \left\{D^{t+1,a}_{ij}\right\}_{a=0,\ldots,L-1;\ (i,j)\in A} \parallel \{u^{t+1}\}$ be the vector of $2+L\cdot|A|$ elements;

3. Let $\phi(S^{t+1}) = \{1\} \parallel \left\{D^{t+1,a}_{ij} \cdot D^{t+1,a'}_{(ij)'}\right\}_{\substack{a=0,\ldots,L-1;\ (i,j)\in A;\\ a'=0,\ldots,L-1;\ (i,j)'\in A}} \parallel \{u^{t+1}\}$ be the vector of $2 + (L \cdot |A|)^2$ elements.

It should be noted that the last option for defining $\phi(S^{t+1})$ would be reasonable (i.e. computationally tractable) only for small transportation networks.

4.3 Numerical Examples

The proposed solution techniques are applied to real-world road networks of Nevada and Vermont. The relevant data are extracted from Matlog [50], which contains the Oak Ridge National Highway Network [51]. Since many of the observed road links are non-separated, it is assumed that $x_{ij} = x_{ji}$ as in an undirected graph. Hence, the observed road networks of Nevada and Vermont include 130 and 178 edges, respectively.

Case studies with 10 stages

The first set of numerical examples assumes that $|T| = 10$ stages, while the lifespan of a facility is $L = 3$ stages. In the context of WIM technology whose average lifespan is 8-12 years, the designed case studies would correspond to a 30-year planning horizon where the investments in WIM systems can be made every 3 years.

Algorithm 3 Approximate Dynamic Programming for the Multi-Stage Stochastic EFCP: Approximating the Value Function with Linear Regression

Step 0: Initialization: start with some initial parameters $\bar{\theta}^{t,0}$ and $B^{t,0}$ for all $t \in T$. Let n = 1 and choose an initial state $S^{0,1}$ (no facility implementation) and step size α^n.

Step 1: For $t \in T$ do the following:

- **Step 1.1:** Solve the problem:

$$\hat{v}^{t,n} = \min_{x \in \mathcal{X}(S^{t,n})} \left\{ (c^t)^T x + \mathbb{E}_{\boldsymbol{\zeta}^t} Q^t(S^{t,n}, \boldsymbol{x}, \bar{\boldsymbol{\xi}}^t, \boldsymbol{\zeta}^t) + \sum_{k=1}^{K} \bar{\theta}_k^{t+1,n-1} \phi_k(S^{t+1,n}) \right\},$$

  where

$$\mathcal{X}(S^{t,n}) = \left\{ x_{ij} \ \text{ s.t. } \sum_{(i,j)\in A} x_{ij} c_{ij}^t \leq u^{t,n} \right\}$$

$$S_{ij}^{t+1,n} = \begin{cases} \max\left(S_{ij}^{t,n} - 1, 0\right), & \text{if } x_{ij} = 0; \\ L - 1, & \text{if } x_{ij} = 1; \end{cases}$$

$$u^{t+1,n} = u^{t,n} - \sum_{(i,j)\in A} c_{ij}^t x_{ij} + b^{t+1}$$

  Let $x^{t,n}$ be the x that minimizes above expression.

- **Step 1.2:** Update $\bar{\theta}^{t,n-1}$ and $B^{t,n-1}$ with observation $\hat{v}^{t,n}$ and vector of features $\phi(S^{t,n})$ by setting

$$\lambda^n = \alpha^{n-1}\left(\frac{1-\alpha^n}{\alpha^n}\right)$$

$$\bar{\theta}^{t,n} = \bar{\theta}^{t,n-1} + \frac{\hat{v}^{t,n} - (\phi(S^{t,n}))^T \bar{\theta}^{t,n-1}}{\lambda^n + (\phi(S^{t,n}))^T B^{t,n-1} \phi(S^{t,n})} B^{t,n-1} \phi(S^{t,n})$$

$$B^{t,n} = \frac{1}{\lambda^n} \left(B^{t,n-1} - \frac{B^{t,n-1}\phi(S^{t,n})(\phi(S^{t,n}))^T B^{t,n-1}}{\lambda^n + (\phi(S^{t,n}))^T B^{t,n-1} \phi(S^{t,n})} \right)$$

- **Step 1.3:** Apply the system dynamics by setting

$$S_{ij}^{t+1,n} = \begin{cases} \max\left(S_{ij}^{t,n} - 1, 0\right), & \text{if } x_{ij}^{t,n} = 0; \\ L - 1, & \text{if } x_{ij}^{t,n} = 1; \end{cases}$$

$$u^{t+1,n} = b^{t+1} + u^{t,n} - \sum_{(i,j)\in A} c_{ij}^t x_{ij}^{t,n}$$

Step 2: Set n = n + 1. Go to step 1.

The flows are generated randomly, 100 flows for Nevada and 50 flows for Vermont, all with the same expected intensity of $\bar{\xi}_f^1 = 200$ units/mile. The expected increase of their intensities is set to be 10% per stage. Moreover, it is assumed that ζ_f^1 has three equally probably realizations, $\zeta_f^1(\omega) \in \{1, 1.1, 1.2\}$. The first realization, $\zeta_f^1(\omega) = 1$, corresponds to the case when the flow is willing to travel only along its shortest path. Second, $\zeta_f^1(\omega) = 1.1$, implies that the flow would be willing to travel an additional 10% of its shortest path to bypass the facilities. Similarly, when $\zeta_f^1(\omega) = 1.2$, the flow would travel an extra 20% distance to bypass the facilities. We assume ζ_f^t is the same for all $t \in T$, and use the k-shortest path algorithm [52] to find the necessary number of paths for each flow.

The objective is to minimize the excessive damage associated with uncaptured flows over the 10 stages. At each stage, we are given a budget that suffices for implementing m facilities. The budget can be either spent or carried on to the next stage. Table 4.1 provides optimal 10-stage investment plans for different values of m, as well as the corresponding excessive damages and solution times. Some of the investment plans are illustrated in Figures 4.1-4.4.

The optimal results reveal some interesting patterns. First, some facilities are renewed after the end of their 3-stage lifespan. This appears to be the case with those facilities that are located at "good strategic" locations. Second, the budget is often saved and accumulated for future stages. This happens for two reasons:

1. Due to evasive nature of flows, it is often better to implement several facilities at the same time, because sequential implementation could allow the flows

to bypass the facilities and produce greater damage until all the facilities are implemented. This is most notable in cases when $m = 1$ or $m = 2$.

2. Due to increasing intensities of flows, it is better to save facilities for later stages which include flows that produce greater damage. As a result, a larger number of facilities is typically implemented in stage 8, in order to cover flows during the last three stages of the planning horizon. Again, this is particularly notable in cases when $m = 1$ or $m = 2$.

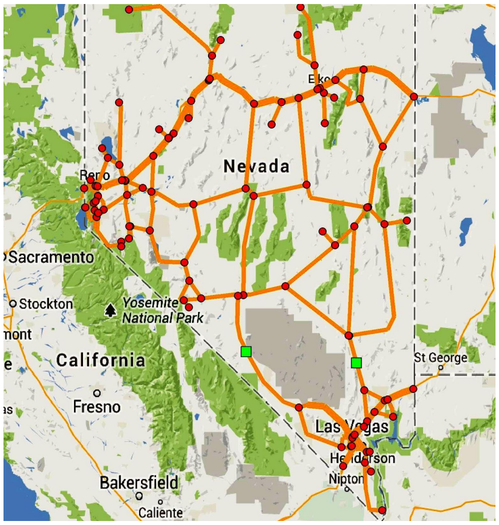
Case studies with 15 stages

The second set of numerical examples assumes that $|T| = 15$, while the lifespan of a facility is $L = 5$ stages. In the context of WIM technology, these case studies correspond to a 30-year planning horizon where the investments in WIM systems are made every 2 years. We use flows from the experiments including 10 stages, this time with the expected increase of 6% per stage. The optimal results for different budget availability are given in Table 4.2.

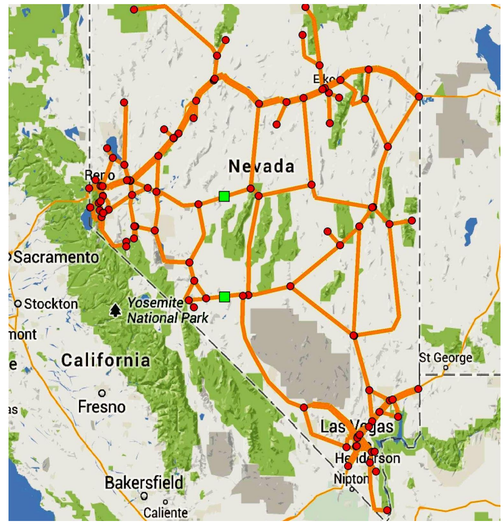
The optimal results follow similar patterns as in the case with 10 stages. Renewal of the facilities again takes place at several important links. In addition, the accumulation of budget is observed. This time, we observe that more facilities are implemented in stage 11, in order to cover the flows during the last 5 stages of the planning horizon.

Table 4.1: Optimal 10-Stage Investment Plans for Different Budgets

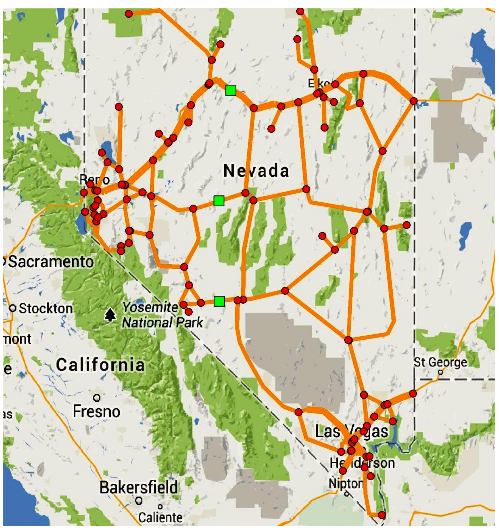
	Nevada					Vermont	
	1 facility/stage	2 facilities/stage	3 facilities/stage	4 facilities/stage	5 facilities/stage	1 facility/stage	2 facilities/stage
t=1	No Implement.	69, 79	69, 77, 79	33, 66, 67, 79	21, 33, 69, 84, 88	No Implement.	26, 88
t=2	21, 33	33, 77	21, 33, 88	52, 60, 88, 103	16, 30, 63, 77, 93	No Implement.	82, 115
t=3	No Implement.	21	52, 63, 84	38, 85, 101	38, 73, 101, 103, 127	88	37
t=4	No Implement.	69, 79	69, 77	33, 66, 67, 79	9, 33, 67, 79, 88	26	26, 88
t=5	69, 79	33, 88	21, 33, 88	52, 60, 88, 103	16, 19, 60, 66, 98	157	82, 115
t=6	77	21, 77	52, 63, 84	38, 73, 75, 101	38, 73, 99, 101, 103	88	37
t=7	33	69, 84	77, 81, 85	33, 66, 67, 79	33, 67, 79, 88, 92	26	26, 88
t=8	69, 79	33, 58, 88	21, 31, 33, 88	31, 52, 63, 88, 103	16, 19, 60, 66, 98	37, 82, 115	18, 82, 115, 132
t=9	77	21, 77	16, 63, 79	38, 73, 85, 101	38, 73, 99, 101, 103	88	36, 41
t=10	33	69, 84	77, 81, 85	33, 67, 77, 79	9, 33, 67, 78, 88	26	26, 88
Damage	28,672,418	9,241,450	3,823,222	1,411,430	678,628	8,225,243	4,129,683
CPU	15,826 sec	2,389 sec	2,431 sec	656 sec	2,142 sec	8,205 sec	7,881 sec



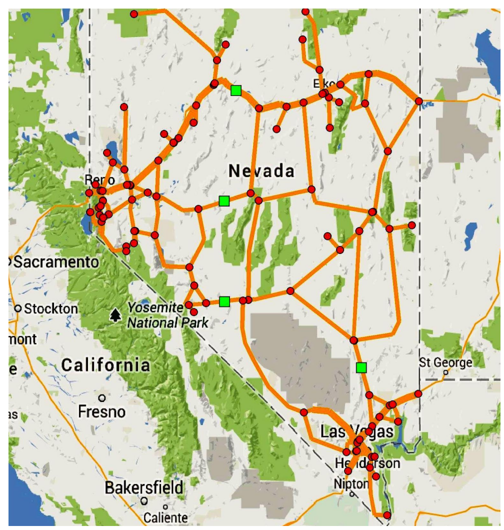
(a) Allocation for $t = 2, 3, 4$.



(b) Allocation for $t = 5$.

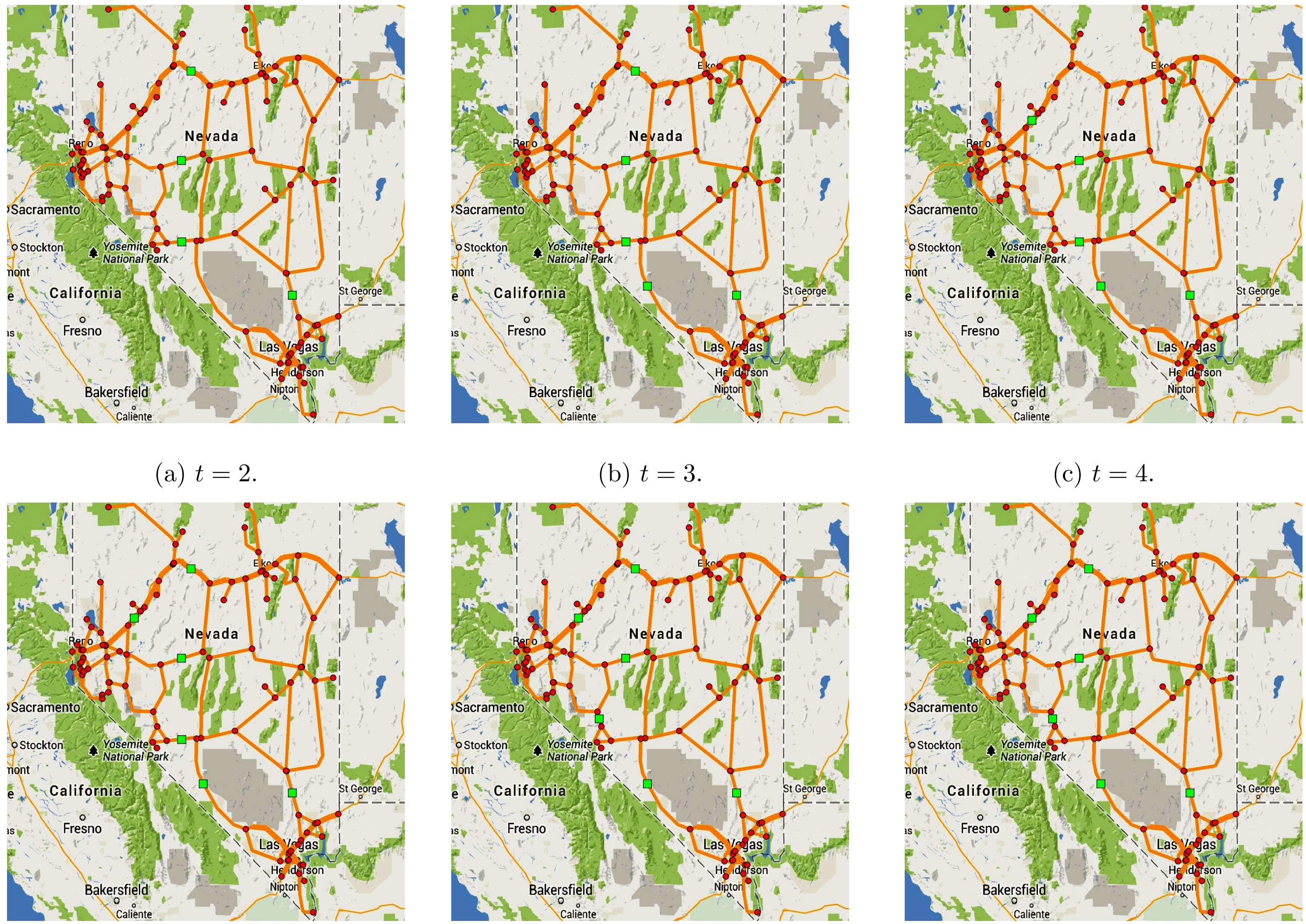


(c) Allocation for $t = 6$.



(d) Allocation for $t = 7, 8, 9, 10$.

Figure 4.1: Nevada: optimal implementation of facilities over the course of 10 stages with budget allowing one facility per stage



(a) $t = 2$.

(b) $t = 3$.

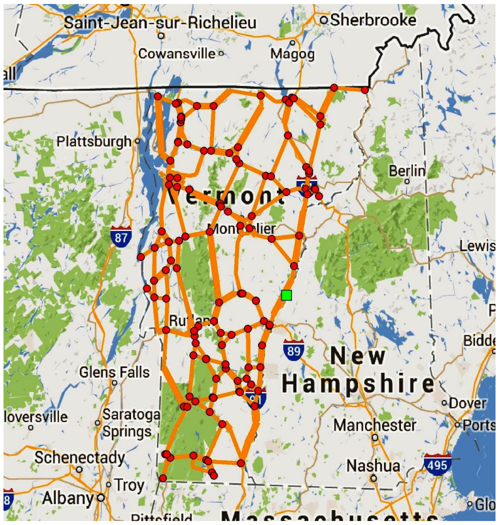
(c) $t = 4$.

(d) $t = 5$.

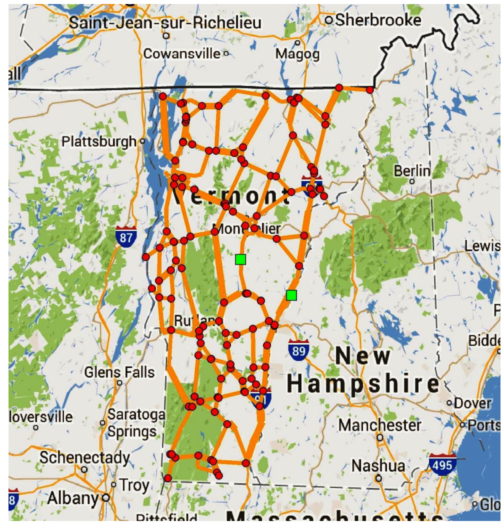
(e) $t = 6$.

(f) $t = 7, 8, 9, 10$.

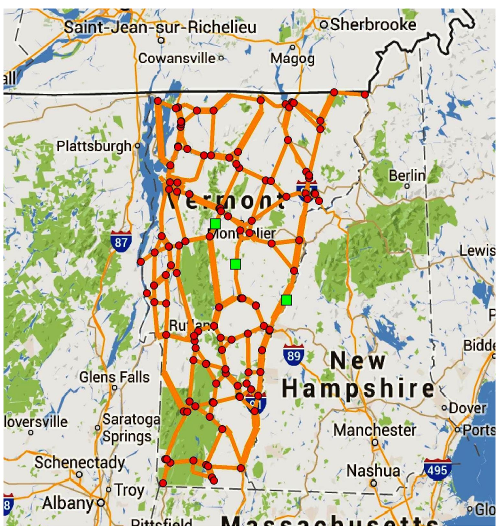
Figure 4.2: Nevada: optimal implementation of facilities over the course of 10 stages with budget allowing two facilities per stage



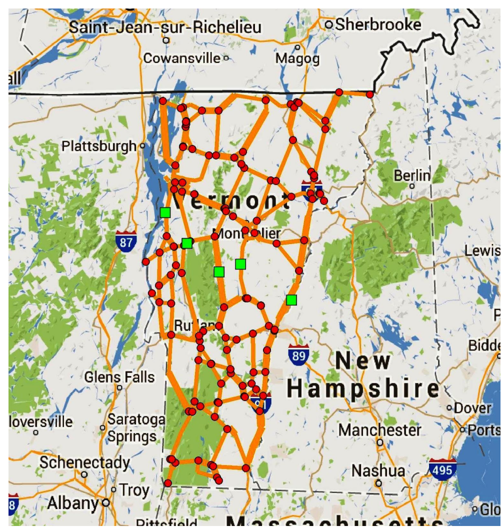
(a) Allocation for $t = 3$.



(b) Allocation for $t = 4$.

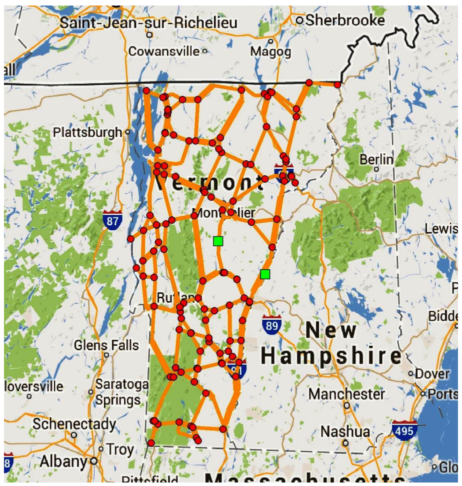


(c) Allocation for $t = 5, 6, 7$.

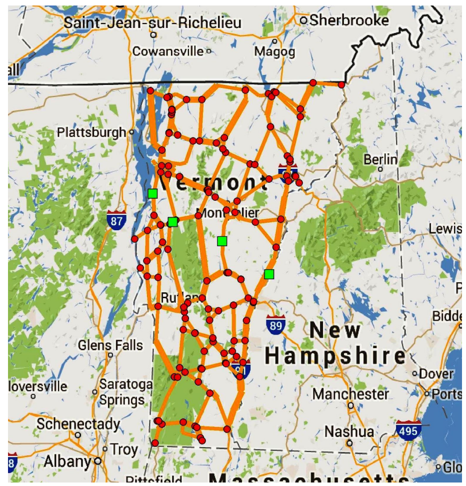


(d) Allocation for $t = 8, 9, 10$.

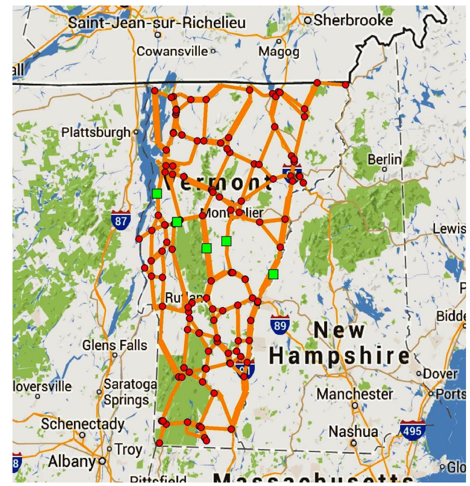
Figure 4.3: Vermont: optimal implementation of facilities over the course of 10 stages with budget allowing one facility per stage



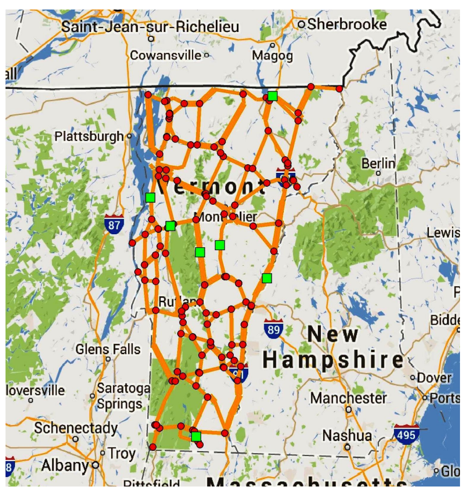
(a) $t = 1$.



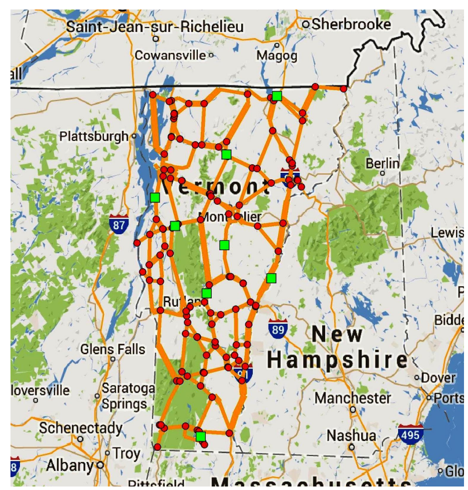
(b) $t = 2$.



(c) $t = 3, 4, 5, 6, 7$.



(d) $t = 8$.



(e) $t = 9, 10$.

Figure 4.4: Vermont: optimal implementation of facilities over the course of 10 stages with budget allowing two facilities per stage

Table 4.2: Optimal 15-Stage Investment Plans for Nevada and Different Budgets

	1 facility/stage	2 facilities/stage	3 facilities/stage	4 facilities/stage	5 facilities/stage
t=1	79	69, 79	69, 77, 79	33, 66, 67, 79	21, 33, 69, 84, 88
t=2	69	33, 77	21, 33, 88	16, 60, 88, 103	30, 52, 63, 77, 93
t=3	77	21, 88	16, 60, 103	38, 73, 85, 101	5, 43, 73, 101, 127
t=4	33	52, 63	30, 85, 101	19, 92, 98, 99	1, 15, 20, 26, 41
t=5	No implement.	No implement.	38, 73, 114	21, 33	94, 98
t=6	21, 79	69, 84	69, 77, 79	52, 58, 69, 71, 77, 100	36, 44, 49, 58, 69, 80, 90, 128
t=7	69	33, 77	33, 63, 88	56, 74, 103, 115	31, 51, 77, 99, 101
t=8	88	21, 88	16, 103	25, 30, 50, 73,	5, 20, 41, 73, 94
t=9	33	16, 31, 63	30, 85, 101, 112	41, 92, 98, 99	15, 30, 74, 100, 115
t=10	No implement.	85,	38, 73	15, 21, 36, 49	98
t=11	21, 79	67, 79, 101, 103	66, 67, 79, 81	58, 69, 77, 100	36, 44, 49, 58, 69, 77, 80, 82, 90
t=12	69	33, 66	33, 60, 88	30, 74, 103, 115	26, 41, 51, 99, 103
t=13	88	38, 88	19, 52, 103	15, 25, 50, 73	19, 21, 50, 73, 92
t=14	33	52, 60	85, 101, 114	41, 92, 98, 99	15, 94, 114, 117, 124
t=15	77	73, 85	38, 41, 73	21, 36, 49, 109	20, 30, 93, 110, 121
Damage	20,817,490	6,135,153	2,357,105	1,120,433	675,057
CPU	9,380 sec	15,983 sec	57,663 sec	9,652 sec	1,236 sec

The proposed ADP algorithm was implemented in Matlab with feature functions specified in three different ways, as described in Section 4.2.1. Moreover, three different step sizes were considered: harmonic, polynomial, and McClain's [56]. The parameters of the linear regression diverged even for the small problem instances including as little as 12 links. This issue is often encountered and discussed in the literature on ADP [56]. In future work, this issue may be overcome through the use of artificial neural networks which could also capture the nonlinear relationship between the states of the system and the downstream values of the objective function.

4.4 Conclusions

An extension of the stochastic EFCP is proposed in which the decisions about the implementation of facilities are made over multiple stages. The structural properties of the problem are exploited to make instances involving real-world transportation networks tractable with the exact solution technique. This is achieved by 1) reducing the noise associated with the intensities of the flows, 2) formulating the scenario-dependent constraints recursively in each stage, and 3) linearly relaxing many scenario-dependent binary variables in each stage. The proposed methodology is tested on the road networks of Nevada and Vermont, which shows the applicability of the proposed model and solution technique.

In addition to the exact solution approach, we study approximate solution techniques. The problem is formulated as a dynamic program and an approximate dynamic programming algorithm is proposed, in which the objective function is estimated through the use of linear regression. This approach turns out to be unsuccessful even for small problem instances. Future work may include application of neuro-dynamic programming which could account for the nonlinear relation between the states of the system and the downstream values of the objective function.

Chapter 5: Optimal Location of WIM in Nevada

An application of the deterministic EFCP is shown in a realistic case study including the allocation of WIM in the road network of Nevada. This case study considers the road network designated for large commercial vehicles, truck flows simulated based on data available in the literature, and realistically estimated damage produced by overweight trucks. It also contrasts EFCP with FCLAP in order to estimate the value that proposed EFCP framework adds in allocating facilities which targeted flows wish to avoid. In addition, it contrasts EFCP with the real-world implementation of static weigh scales in Nevada in order to explore whether current allocations could be improved through application of the EFCP. This comparison is conducted given the *limited* available information about the truck flows in Nevada. In sum, this chapter:

1. Demonstrates applicability of the proposed work to a realistic case study and discusses input preparation;

2. Numerically estimates the benefits of applying EFCP and thereby explores the concrete contributions of this dissertation.

The remainder of this chapter is organized in seven sections. Sections 5.1 and 5.2 estimate the inputs and explain the design of the case study. Section 5.3 provides

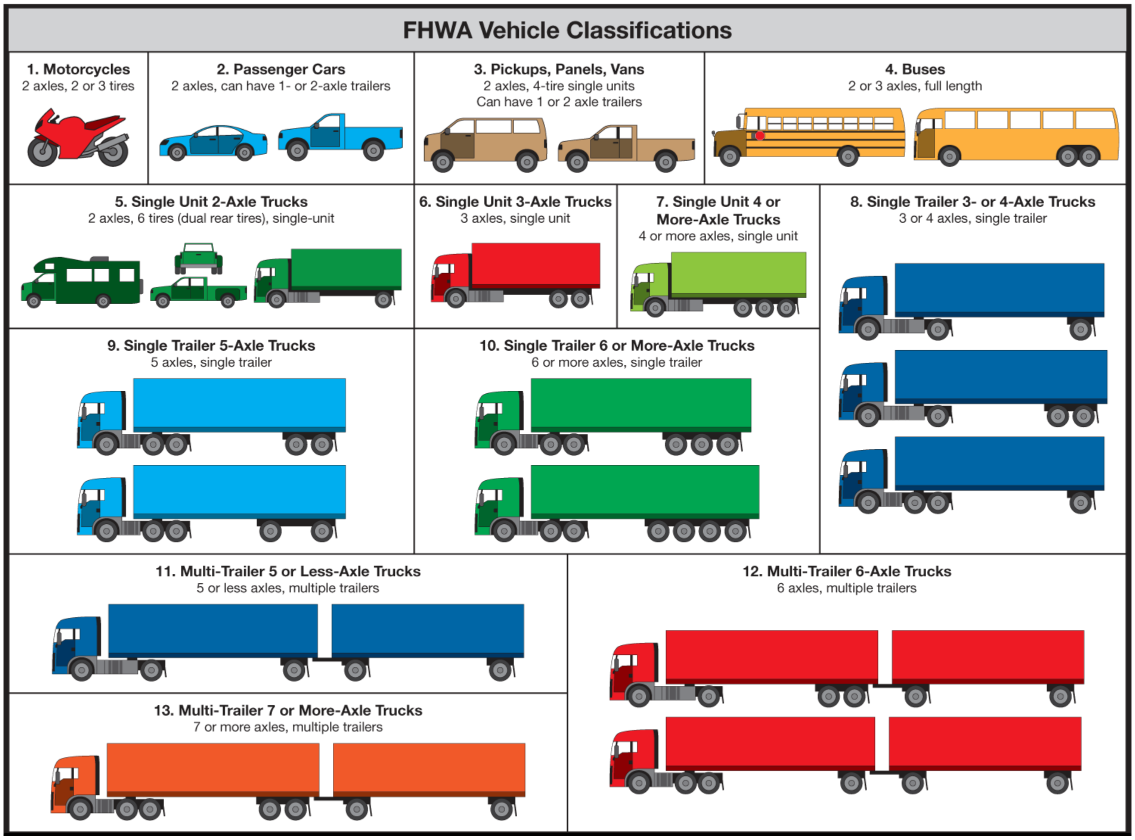


Figure 5.1: Federal Highway Classification of Vehicles [1]

optimal allocations of WIM stations given different costs per lane of WIM technology. Section 5.4 contrasts EFCP with FCLAP, whereas 5.5 explores the application of approximate solution techniques. Section 5.6 explores whether the current allocation of facilities can be improved. Finally, Sections 5.7 draws conclusions.

5.1 Excessive Damage Estimation

This section discusses estimation of parameter c_f^p which denotes the excessive damage cost if flow $f \in F$ passes unintercepted along path $p \in P_f$. It has been argued earlier that overweight trucks damage the pavement and environment. Thus,

c_f^p is estimated by roughly computing the aforementioned damage costs associated with loads that exceed legal limits.

Pavement damage depends on many factors including axle weights, axle configuration, pavement structure, and climate. Since detailed information about the pavement structure and climate may not be available for the entire transportation network, the pavement damage can be estimated based on the equivalent single axle load (ESAL). This method allows different axle types (single, tandem, and tridem) to be summed together and is widely used in pavement design since it provides a reasonably accurate indicator of the pavement damage [34]. ESALs may be estimated with the formula

$$ESAL = \alpha \left[\frac{(W/\alpha)}{80} \right]^{4.2} \tag{5.1}$$

where α is the number of individual axles in an axle group (for steering and single $\alpha = 1$; for tandem $\alpha = 2$; for tridem $\alpha = 3$) and W is weight of an axle [kN]. In computing the excessive pavement damage, the following axle loads [34] are used as legal limits for each axle group shown:

1. Steering: 55 kN, which corresponds to 0.21 ESALs;
2. Single: 88 kN, which corresponds to 1.49 ESALs;
3. Tandem: 151 kN, which corresponds to 1.57 ESALs;
4. Tridem: 233 kN, which corresponds to 2.65 ESALs.

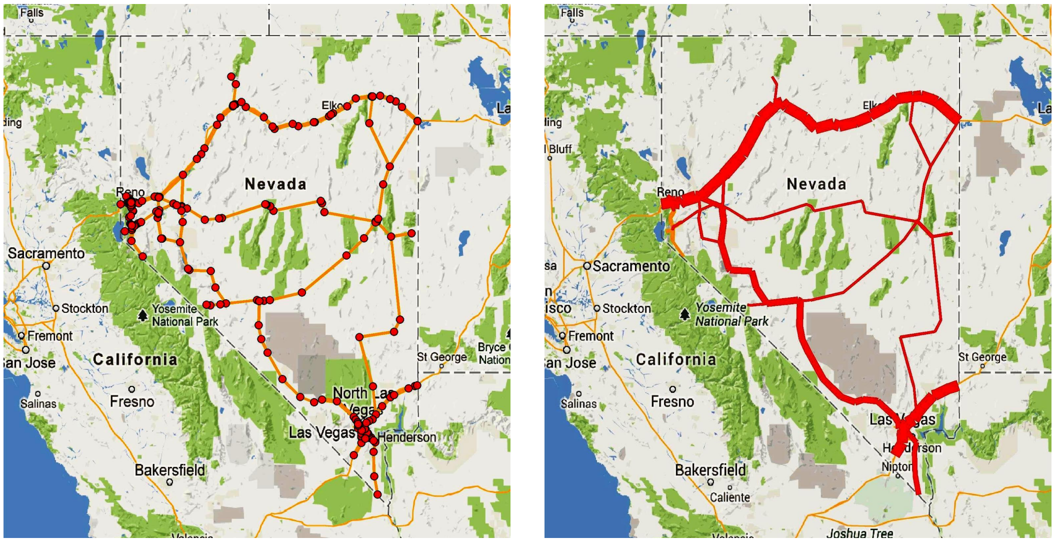
Table 5.1 provides an example of how the excessive pavement damage is computed for a 17 ton truck that has front steering and a single rear axle. The assumed

Table 5.1: Computing Excessive Pavement Damage: An Example

	Weight [kN]	ESALs	ESAL Limits	Excess ESALs	Total Excess ESALs
Steering	64.12	0.39	0.21	0.18	0.18+1.35=1.53
Single	102.59	2.84	1.49	1.35	

gross truck weight distribution is 38:62 between the front steering and single rear axle, the same as the maximum axle load ratio in kN (e.g. 55:88). In particular, Table 5.1 provides the axle weights in kN and corresponding ESALs computed with equation (5.1). The obtained ESALs are compared with the limits to obtain a total of 1.53 excessive ESALs. Finally, assuming the fee of 4 cents per ESAL-mile (adjusted for inflation from [58]), the excessive pavement damage of 6.12 cents per mile for this particular truck is computed.

Environmental damage includes accidents (fatalities, injuries, and property damage), emissions (air pollution and greenhouse gases), noise, and unrecovered costs associated with the provision, operation, and maintenance of public facilities [35]. The average environmental damage cost is assumed to be 1.53 cents per ton-mile (adjusted for inflation from [35]). Thus, assuming that the truck from Table 5.1 is overloaded by 2.7 tons, the corresponding excessive environmental damage is 4.13 cents per mile.



(a) Road Network (b) Truck flows

Figure 5.2: Nevada's Road Network and Major Truck Flows

5.2 Road Network, Flows, and Other Inputs

The proposed model is tested on the road network of Nevada, considering only road links that are state designated for Surface Transportation Assistance Act (STAA) vehicles. Most of the observed road links have either 2 or 4 lanes. Since many of these road links are non-separated, it is assumed that $x_{ij} = x_{ji}$ like in undirected graphs. The relevant data are extracted from Matlog [50], which contains the Oak Ridge National Highway Network [51] with approximately 500,000 miles of roadway in the US, Canada, and Mexico, including all rural arterials and urban principal arterials in the US.

The truck flows along three major transit routes are specified based on data from [59] and [60]. They include 5,000 trucks/day on I-15 (southwest of Reno - Salt Lake City) and I-80 (passing by Las Vegas), as well as 2,000 trucks/day along the route stretching from northwest of Reno to south of Las Vegas (Figure 5.2b). In addition, 59 local truck flows are randomly generated with their origins and destinations at least 50 miles apart. Moreover, the number of trucks within the flow is assumed to be Poisson distributed with a mean 50 trucks/day. Ten different types of trucks with different numbers and combinations of axles are considered. Table 5.2 provides truck weights, weight limits, and the assumed percentages within the total flow for each truck type. Since truck weights are typically bimodally distributed [38] due to imbalanced flows, the trucks are simulated so that 60% are traveling with heavy loads and 40% are traveling with light loads (e.g. empty or nearly empty trucks returning to their origins). Discrete distributions of load weights in tons are provided in Table 5.2. The expected number of overweight trucks generated based on the assumed inputs from Table 5.2 is 4.5% of the total number of trucks. It should be noted that this percentage is within the range reported in the literature, such as 2.6% for California [34] and 8.8% for Montana [38].

Yen's k-shortest path algorithm [52] is used to find k_f shortest loopless paths, such that the $(k_f + 1)$-th shortest path is at least 20% longer than the shortest path. Thus, k_f varies considerably with flows. For example, $k_f = 5$ for transit flow passing by Las Vegas, whereas $k_f = 910$ for flow traversing Nevada east-west. It should be noted that the 59 local truck flows are randomly generated so that their origins and destinations are at least 50 miles apart and that the maximum number of paths that must be considered is 30 (e.g. the $k_f = 30$ shortest path is more than 20% longer than the shortest path).

A single set of flows, trucks, and truck loads is generated using Monte Carlo

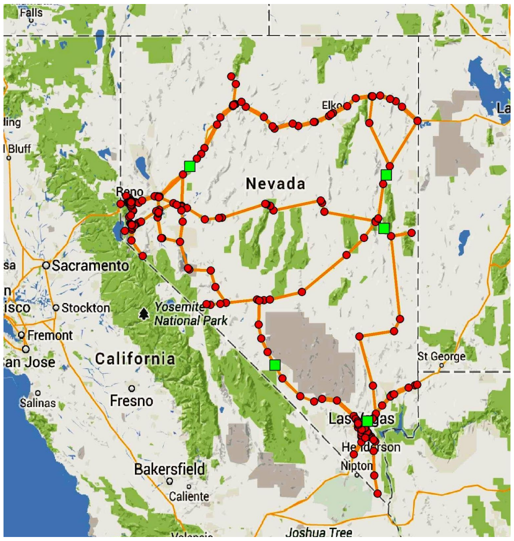
Table 5.2: Simulating Truck Flows

Type	Number of Axles	Empty Truck Weight (ton)	Loads Light (ton)	Weight Heavy (ton)	Weight Limit (ton)	Percent in Total Flow
S	2	6	B(3, 0.45)	B(15, 0.40)	14.3	9
S	3	8	B(4, 0.50)	B(22, 0.45)	20.6	17
T	3	10	B(5, 0.45)	B(25, 0.40)	23.1	3
T	4	13	B(6, 0.45)	B(31, 0.40)	29.4	4
T	5	15	B(7, 0.45)	B(39, 0.40)	35.7	46
T	6	16	B(9, 0.50)	B(50, 0.45)	43.9	3
T	5	15	B(8, 0.50)	B(46, 0.45)	40.7	7
MT	6	18	B(9, 0.50)	B(53, 0.45)	47.0	3
MT	8	21	B(12, 0.50)	B(68, 0.45)	59.0	4
MT	7	20	B(11, 0.50)	B(59, 0.45)	52.7	4

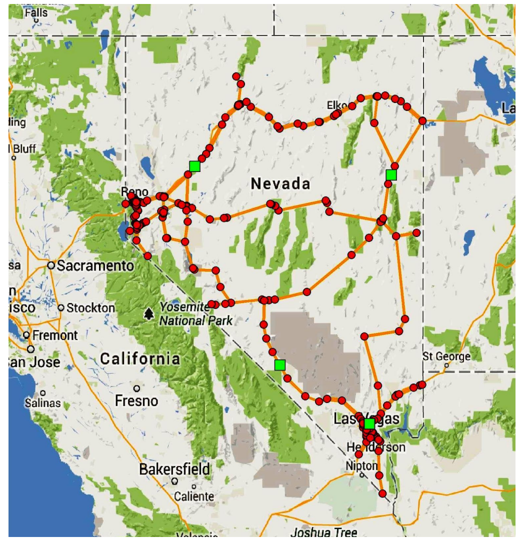
S - single unit truck, T/MT - single/multi trailer truck, B(n,k) - Binomial distribution simulation. The corresponding excessive damage c_f^p is computed for all the flows and their paths, as described in Section 5.1. Finally, WIM cost includes the cost of hardware and software, implementation, maintenance, re-calibration, office and personnel. Available references indicate that total cost can vary considerably depending on the technology (e.g. sensors, cameras) and location (e.g. state within the same country). Numerical results are provided for WIM cost ranging between \$10 thousand and \$360 thousand per lane per year. For this analysis, \$60 thousand per lane-year is the most realistic cost, since the cost of only WIM inroad equipment ranges between \$7 thousand and \$12 thousand per lane-year depending on the technology (adjusted for inflation from [61]).

5.3 Optimal Results for EFCP

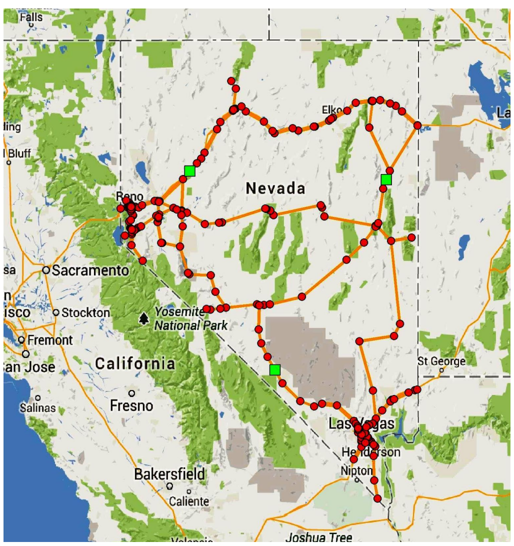
The binary program (2.1)-(2.6) is implemented in GAMS 23.5 and solved using GAMS/CPLEX solver for mixed integer programs on a PC with an AMD Athlon 3300 GHz processor with 4 GB of RAM. The optimal results for different WIM costs are provided in Table 5.3 and the corresponding allocations of checkpoints are shown in Figure 5.3. To simplify the comparison, the links in Table 5.3 and throughout this section are denoted with tags (e.g. 1-221 for 221 road links), rather than with their origin and destination nodes. Finally, it should be noted that all the results are obtained within 3 to 4 seconds of computation time.



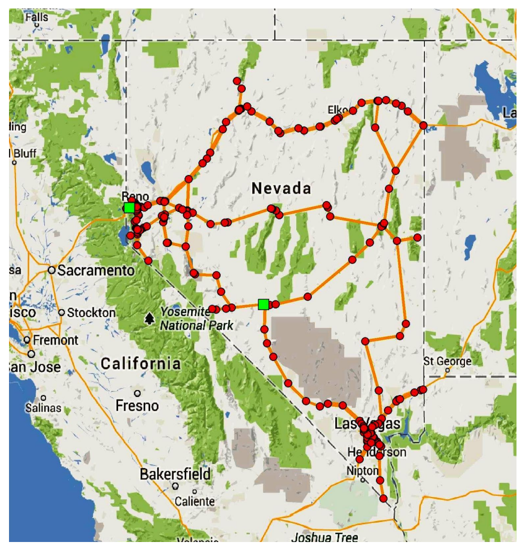
(a) $\boldsymbol{x^*_{EFCP}}$ for \$10,000/lane-year



(b) $\boldsymbol{x^*_{EFCP}}$ for \$60,000/lane-year



(c) $\boldsymbol{x^*_{EFCP}}$ for \$110,000/lane-year



(d) $\boldsymbol{x^*_{EFCP}}$ for \$160,000/lane-year

Figure 5.3: Optimal Allocations of WIM Checkpoints for Different Costs of the WIM Technology

Table 5.3: Optimal Results for Different WIM Costs

WIM Costs ($/lane-year)	$\boldsymbol{x^*_{EFCP}}$ (links covered)	WIM Systems ($/year)	Excessive Damage ($/year)	Total Cost ($/year)
10,000	30, 32, 93, 130, 216	140,000	26,947	166,947
60,000	32, 62, 93, 130	720,000	56,370	776,370
110,000	32, 93, 130	880,000	403,640	1,283,640
160,000	105, 164	960,000	681,633	1,641,633
210,000	105, 164	1,260,000	681,633	1,941,633
260,000	105, 164	1,560,000	681,633	2,241,633
310,000	105	620,000	1,723,607	2,343,607
360,000	no WIMs	0	2,349,907	2,349,907

5.4 Numerical Comparison of EFCP and FCLAP

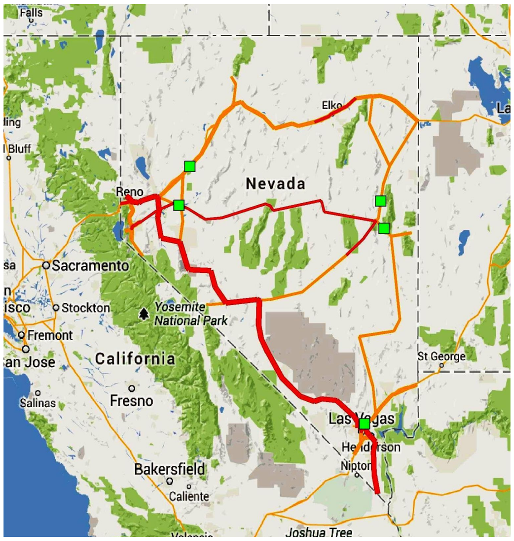
Now let us observe what would happen if the FCLAP was applied to determine the optimal allocation of WIM checkpoints. Recall from Proposition 2 that EFCP reduces to FCLAP when $k_f = 1$. First, the EFCP for $k_f = 1$ is applied to find the optimal WIM allocation $\boldsymbol{x^*_{FCLAP}}$ and the corresponding objective function $FCLAP(\boldsymbol{x^*_{FCLAP}})$. Second, this solution is evaluated for the EFCP where k_f is determined so that the $(k_f + 1)$-th shortest path is at least 20% longer than the shortest path. This value is denoted by $EFCP(\boldsymbol{x^*_{FCLAP}})$ and contrasted with the optimal solution $EFCP(\boldsymbol{x^*_{EFCP}})$.

The last column in Table 5.4 indicates that the solution obtained from FCLAP performs poorly in the setting where trucks try to avoid WIM systems by taking reasonably long detours. The graphical comparison and dispersion of the *uncaptured* flows for the two solutions is provided in Figures 5.4 and 5.5. This comparison indicates that truck flows simply bypass the facilities allocated with FCLAP. For example, Figures 5.5a and 5.5c show that the flow traversing Nevada east-west bypasses the implemented facility at a small increase in travel distance. A similar situation occurs in Figure 5.4c, but at a higher increase in driving distance that also includes greater excessive damage associated with the same transit flow.

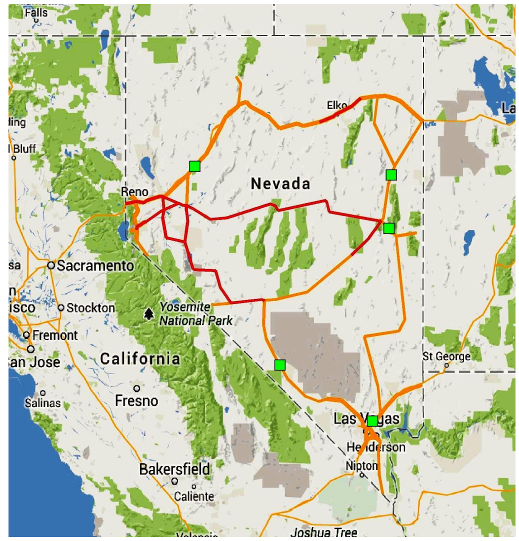
Table 5.4 also clearly illustrates the WIM paradox, in which inefficient use of WIM technology actually causes excessive damage (and total system cost) to increase. In particular, the allocation $\boldsymbol{x}^*_{\boldsymbol{FCLAP}}$ based on the FCLAP incurs a cost of approximately \$2.9-3.9M/year for the WIM technology cost of \$110-360k/lane-year. On the other hand, Table 5.3 indicates a total cost of roughly \$2.4M/year when no WIM technology is implemented. Hence, the FCLAP allocation is counterproductive, and actually incurs greater total cost than a solution that includes no WIMs at all. This clearly demonstrates the potential pitfalls of using FCLAP in settings where users behave non-cooperatively.

5.5 Heuristic Results for EFCP

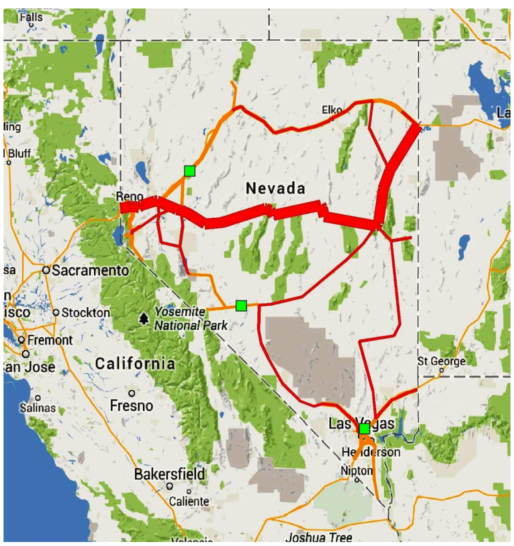
This section discusses performance of the greedy heuristic implemented in Matlab, as well as a binary genetic algorithm available in Matlab 2013a. Table 5.5 indicates good performance of the greedy heuristic, which took few seconds of computation time. In this particular instance, the greedy heuristic performs within



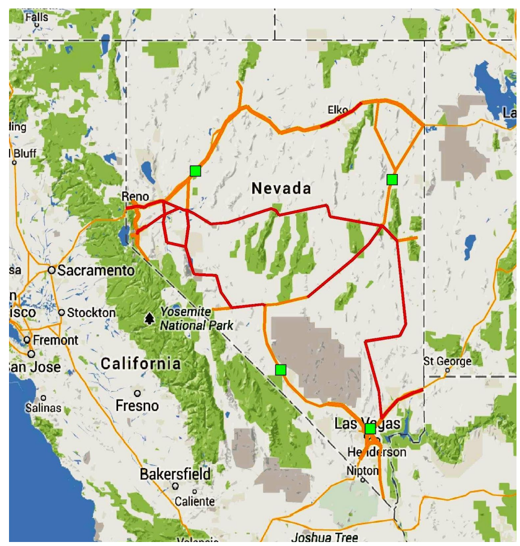
(a) $\boldsymbol{x}^*_{\boldsymbol{FCLAP}}$ for \$10,000/lane-year



(b) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for \$10,000/lane-year

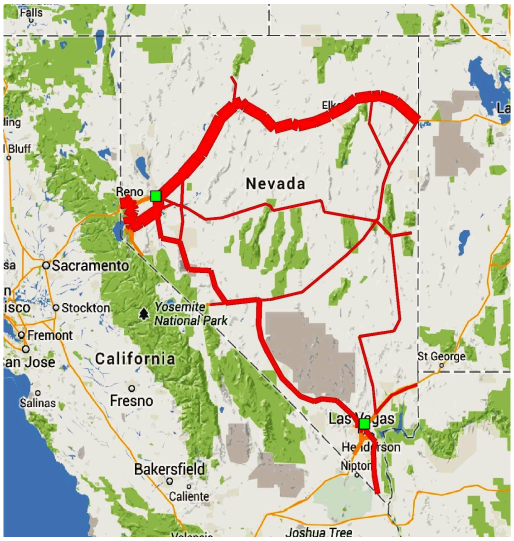


(c) $\boldsymbol{x}^*_{\boldsymbol{FCLAP}}$ for \$60,000/lane-year

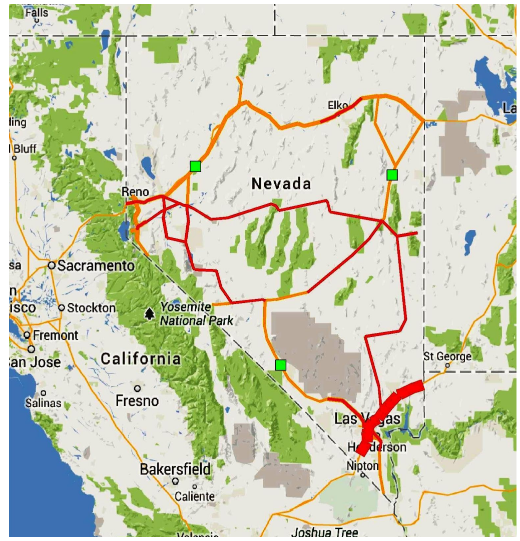


(d) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for \$60,000/lane-year

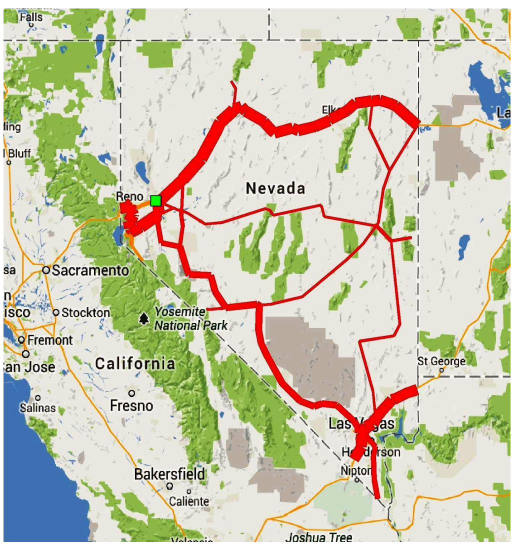
Figure 5.4: Comparison of FCLAP and EFCP for WIM Cost of \$10,000 and \$60,000 per lane-year



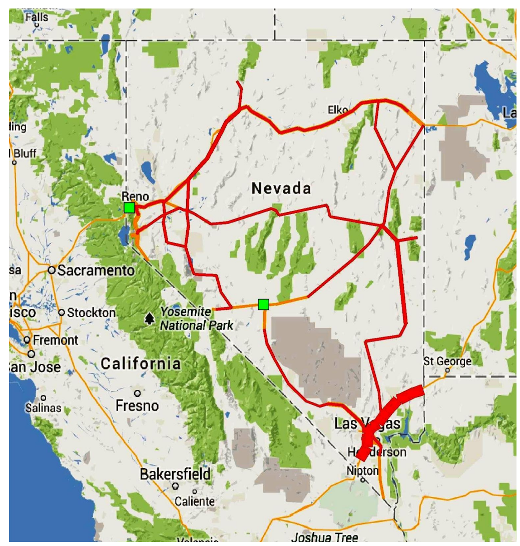
(a) $\boldsymbol{x^*_{FCLAP}}$ for \$110,000/lane-year



(b) $\boldsymbol{x^*_{EFCP}}$ for \$110,000/lane-year



(c) $\boldsymbol{x^*_{FCLAP}}$ for \$160,000/lane-year



(d) $\boldsymbol{x^*_{EFCP}}$ for \$160,000/lane-year

Figure 5.5: Comparison of FCLAP and EFCP for WIM Cost of \$110,000 and \$160,000 per lane-year

Table 5.4: Comparison of EFCP and FCLAP for Different WIM Costs

WIM Costs ($/lane-year)	$\boldsymbol{x}^*_{\boldsymbol{FCLAP}}$ (links covered)	$FCLAP(\boldsymbol{x}^*_{\boldsymbol{FCLAP}})$ ($/year)	$EFCP(\boldsymbol{x}^*_{\boldsymbol{FCLAP}})$ ($/year)	$\frac{EFCP(\boldsymbol{x}^*_{\boldsymbol{EFCP}})}{EFCP(\boldsymbol{x}^*_{\boldsymbol{FCLAP}})}$
10,000	30, 37, 62, 130, 138	152,791	667,245	0.250
60,000	62, 107, 130	699,362	1,911,587	0.406
110,000	62, 154	1,102,999	2,874,007	0.447
160,000	154	1,367,532	3,094,203	0.530
210,000	154	1,567,532	3,294,203	0.589
260,000	154	1,767,532	3,494,203	0.641
310,000	154	1,967,532	3,694,203	0.634
360,000	154	2,167,532	3,894,203	0.603

20% of optimality; recall, however, that a bound on its worst-case performance cannot be established by Proposition 6. As one would expect, the genetic algorithm outperforms the greedy heuristic at the cost of a considerably increased computation time. The initial population of 500 individuals is generated randomly, after which the first 221 individuals are assigned a WIM checkpoint at 221 possible locations. The algorithm is run 5 times through 2000 generations and the best solutions are presented in Table 5.5. All the reported solutions are within 2.5% of the optimum and the average computation time is about 2.5 hours.

Table 5.5: Performance of Greedy Heuristic and Genetic Algorithm for EFCP

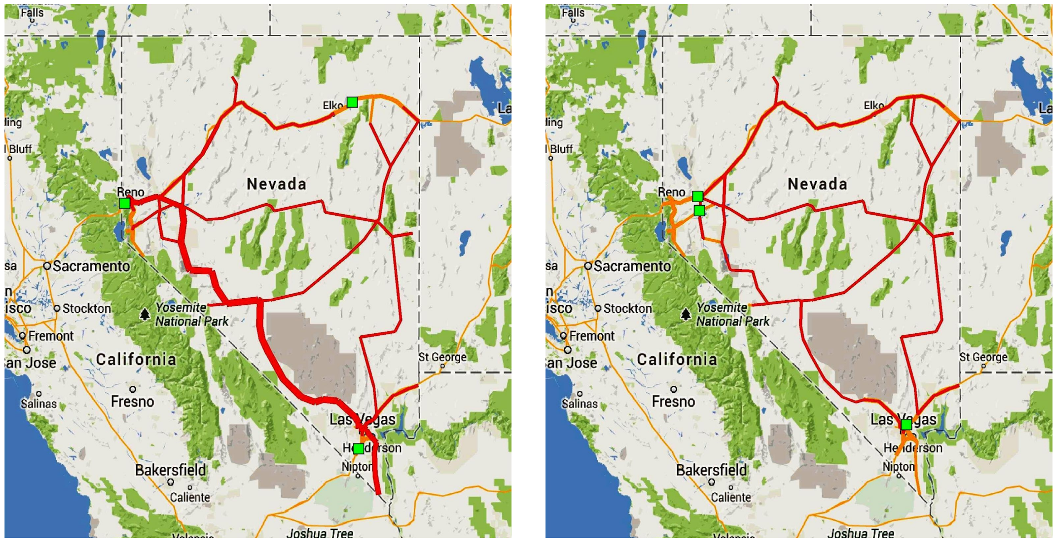
WIM cost ($/lane-year)	$\boldsymbol{x}^{greedy}$ (links covered)	$\frac{EFCP(\boldsymbol{x}^*_{EFCP})}{EFCP(\boldsymbol{x}^{greedy})}$	$\boldsymbol{x}^{genetic}$ (links covered)	$\frac{EFCP(\boldsymbol{x}^*_{EFCP})}{EFCP(\boldsymbol{x}^{genetic})}$
10,000	30, 32, 62, 105, 130, 164	0.801	37, 46, 62, 93, 130	0.975
60,000	62, 105, 164	0.857	37, 38, 66, 130	1
110,000	105, 164	0.957	42, 93, 130	0.979
160,000	105, 164	1	107, 170	1
210,000	105, 164	1	107, 164	1
260,000	105, 164	1	107, 170	1
310,000	105	1	107	1
360,000	no WIMs	1	no WIMs	1

5.6 Comparison of EFCP to the Real-World Solution

The real-world implementation of *static* weigh systems in Nevada is contrasted with the solution suggested by the EFCP for WIM allocation. Several grounds for caution should be noted in interpreting this comparison. First, locations of static weigh scales are more restricted than those of WIMs because static scales require considerable land for their ramps and truck queues. Thus, the authorities may have considered only a subset of the links considered in the model (e.g. only links that are further away from towns). The reasons for this could be the land ownership and price, or space availability. Second, the model focuses on road links that are either state designated for STAA vehicles or federally designated for large commercial vehicles. On the other hand, in allocating static scales the authorities may have considered additional roads (i.e. not only roads designated for STAA or large commercial vehicles) as potential bypasses. Third, the assumed intensities of truck flows are based on recent references, but the flows may have been different when the static weigh stations were originally implemented. Furthermore, our experiment includes some randomly simulated local truck flows.

Since Nevada currently has three static stations [62], Problem **P1'** for $m = 3$ is applied to minimize the excessive damage. The real-world implementation and optimal solution for EFCP are given in Figure 5.6 together with graphical representation of the uncaptured flows and the corresponding excessive damage. The main difference between the two solutions arises in (not) capturing the transit flow between northwest of Reno and south of Las Vegas (note that the two checkpoints in Figure 5.6b are grouped together to capture this flow). Thus, under the assumptions of the model, the optimal solution for EFCP outperforms the real-world implementation by about \$670,000/year. While the exact dollar amount reflects the assumptions made in our experiments, it suggests that there is significant economic potential in modeling evasive transportation flows.

The real-world solution suggests that practitioners, unlike the FCLAPs, have considered that overloaded trucks would try to evade the checkpoints, as they have placed them at locations that cannot be avoided at a small increase in driving distance. These locations include links close to border crossings and other areas where road network is not well connected. As it happens, however, the optimal allocation for three stations is somewhat counter-intuitive, as it is better to implement 2 of 3 checkpoints very close together, instead of spreading them out across the net-



(a) Real-world locations and the corresponding damage \$925,640/year

(b) $\boldsymbol{x}^*_{\boldsymbol{EFCP}}$ for and $m = 3$ and the corresponding damage \$248,941/year

Figure 5.6: Comparison of real-world locations of weigh stations with those suggested by EFCP

work. This case suggests that EFCP can be a useful decision support tool with the potential to improve solutions based on human judgment and intuition.

5.7 Conclusions

This chapter provides a case study where deterministic EFCP is applied to optimally allocate WIM systems in Nevada. The EFCP and FCLAP are also contrasted in this realistic case study and the numerical comparison indicates that results optimal for FCLAP perform poorly in the setting where targeted flows try to avoid the facilities. Moreover, the EFCP-based facility locations are contrasted the actual implementation of static weigh scales in Nevada. This comparison showed that current allocation of static weigh scales could be considerably improved through application of EFCP. These results show the relevance of the proposed EFCP and indicate that it adds a considerable value in the allocation of facilities which targeted flows have an incentive to avoid.

Chapter 6: Conclusions

This dissertation contributes to the literature on facility location by introducing a new type of flow-capturing framework in which targeted flows exhibit non-cooperative behavior by changing their routes in order to avoid the facilities. This work develops three models to allocate facilities given different availability of information and planning policies. Several case studies including real-world transportation networks are conducted to demonstrate applicability and efficiency of the proposed models and solution techniques. The EFCP solutions are also compared with those suggested by the FCLAP. This comparison demonstrates that solutions optimal for FCLAP do poorly when targeted subjects try to avoid the facilities, showing that proposed EFCP adds considerable value.

The EFCP-based allocation is also contrasted with the actual implementation of weigh stations in Nevada, given the available information about the truck flows. This comparison shows that EFCP-based allocation significantly outperforms the actual implementation, which indicates that application of EFCP could yield great economic benefits. These results, as well as wide applicability of EFCP in transportation, revenue management, and security and safety management, show the relevance of the proposed type of flow-capturing problem and encourage further research on EFCP.

6.1 Benefits to Society

The EFCP has many important applications pertaining to preservation of transportation infrastructure and environment (weigh-in-motion systems), safety (inspection stations for transportation of hazardous material) and profit maximization (tollbooths). Thus, the line of research proposed in this dissertation could:

1. Improve the current practice of transportation agencies in locating WIM systems that consists of simply prioritizing the most damaged road links. The proposed EFCP for WIM allocation could both speed up the decision making process of highway agencies and provide more cost effective solutions that 1) reduce government expenditures for road maintenance and 2) decrease environmental damage due to overweight commercial vehicles.

2. Improve toll collection for transportation agencies through optimal allocation of tollbooths. The EFCP for WIM allocation can be readily applied to allocation of tollbooths in a road transportation network. This application would only require different estimation of the parameter c_f^p, which would represent the lost revenue and road deterioration associated with those flows that bypass the tollbooths.

3. Improve safety management through optimal allocation of security checkpoints (e.g. inspection stations for vehicles transporting hazardous material). The

EFCP for WIM allocation can be directly applied in allocating fixed security and safety checkpoints to manage risk. This application may include allocation of a fixed number of facilities (i.e. problem **P1'**) in order to minimize the risk associated with unintercepted flows. In such a setting, c_f^p would represent an estimated risk.

6.2 Extensions

One limitation of the proposed EFCP framework is that k_f can be determined so that the $(k_f + 1)$-th shortest path would represent an excessive detour. This approach is appropriate for highway road networks, as it was shown in case studies involving real-world networks of Nevada and Vermont. It would be more difficult to apply this approach to well-connected road networks (e.g. urban areas like Manhattan) due to a very large number of possible paths. For these cases, an alternative cut-based formulation could circumvent the issue of the large number of path-based variables that would currently arise in instances involving well-connected networks. Another way to cope with the well-connected networks would be to apply the network aggregation techniques to reduce the size of the network and hence the number of shortest paths to be considered within the EFCP.

The proposed EFCP framework assumes that flows seek to minimize their travel distance by choosing to travel along their shortest unmonitored paths. Thus, the three formulations introduced herein are flow-separable. A possible extension would be to assume that flows seek to minimize their travel times. This extension would include equilibrium constraints, which would imply different structural properties of such an evasive flow-capturing framework.

Bibliography

[1] Federal Highway Administration. Classification of vehicles, 2013. `http://onlinemanuals.txdot.gov/txdotmanuals/tri/images/FHWA_Classification_Chart_FINAL.png`.

[2] R. Church and C. ReVelle. The maximal covering location problem. *Papers in Regional Science*, 32(1):101–118, 1974.

[3] K. Tanaka. Maximum flow-covering location and service start time problem and its application to Tokyo metropolitan railway network. *Journal of the Operations Research Society of Japan-Keiei Kagaku*, 54(4):237, 2011.

[4] M. J. Hodgson. A flow-capturing location-allocation model. *Geographical Analysis*, 22(3):270–279, 1990.

[5] O. Berman, R. C. Larson, and N. Fouska. Optimal location of discretionary service facilities. *Transportation Science*, 26(3):201–211, 1992.

[6] M. J. Hodgson, K. E. Rosing, and J. Zhang. Locating vehicle inspection stations to protect a transportation network. *Geographical Analysis*, 28(4):299–314, 1996.

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

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

[9] M. W. Horner and S. Groves. Network flow-based strategies for identifying rail park-and-ride facility locations. *Socio-Economic Planning Sciences*, 41(3):255–268, 2007.

[10] M. Kuby and S. Lim. 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.

[11] M. Kuby, L. Lines, R. Schultz, Z. Xie, J. G. Kim, and S. Lim. Optimization of hydrogen stations in Florida using the flow-refueling location model. *International Journal of Hydrogen Energy*, 34(15):6045–6064, 2009.

[12] P. B. Mirchandani, R. Rebello, and A. Agnetis. The inspection station location problem in hazardous material transportation: some heuristics and bounds. *Infor-Information Systems and Operational Research*, 33(2):100–113, 1995.

[13] N. Marković, I. O. Ryzhov, and P. Schonfeld. Evasive flow capture: Optimal location of weigh-in-motion systems, tollbooths, and safety checkpoints. Submitted for publication.

[14] O. Berman, D. Bertsimas, and R. C. Larson. Locating discretionary service facilities, II: maximizing market size, minimizing inconvenience. *Operations Research*, 43(4):623–632, 1995.

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

[16] O. Berman. The maximizing market size discretionary facility location problem with congestion. *Socio-Economic Planning Sciences*, 29(1):39–46, 1995.

[17] C. Upchurch, M. Kuby, and S. Lim. A model for location of capacitated alternative-fuel stations. *Geographical Analysis*, 41(1):85–106, 2009.

[18] K. Tanaka and T. Furuta. A hierarchical flow capturing location problem with demand attraction based on facility size, and its Lagrangian relaxation solution method. *Geographical Analysis*, 44(1):15–28, 2012.

[19] S. H. Chung and C. Kwon. Multi-period planning for electric-car charging station locations: a case of Korean expressways. `http://infohost.nmt.edu/~schung/chung2012.pdf`, 2012.

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

[21] W. Zeng, M. J. Hodgson, and I. Castillo. The pickup problem: consumers' locational preferences in flow interception. *Geographical Analysis*, 41(1):107–126, 2009.

[22] O. Berman, D. Krass, and C. W. Xu. Locating discretionary service facilities based on probabilistic customer flows. *Transportation Science*, 29(3):276–290, 1995.

[23] O. Berman, D. Krass, and C. W. Xu. Locating flow-intercepting facilities: New approaches and results. *Annals of Operations Research*, 60(1):121–143, 1995.

[24] S. H. R. Pasandideh and S. T. A. Niaki. Genetic application in a facility location problem with random demand within queuing framework. *Journal of Intelligent Manufacturing*, 23(3):651–659, 2012.

[25] O. Berman and D. Krass. Flow intercepting spatial interaction model: a new approach to optimal location of competitive facilities. *Location Science*, 6(1):41–65, 1998.

[26] T. H. Wu and J. N. Lin. Solving the competitive discretionary service facility location problem. *European Journal of Operational Research*, 144(2):366–378, 2003.

[27] M. J. Hodgson and K. E Rosing. A network location-allocation model trading off flow capturing and p-median objectives. *Annals of Operations Research*, 40(1):247–260, 1992.

[28] O. Berman. Deterministic flow-demand location problems. *Journal of the Operational Research Society*, 48(1):75–81, 1997.

[29] W. Zeng, I. Castillo, and M. J. Hodgson. A generalized model for locating facilities on a network with flow-based demand. *Networks and Spatial Economics*, 10(4):579–611, 2010.

[30] S. A. MirHassani and R. Ebrazi. A flexible reformulation of the refueling station location problem. *Transportation Science*, 2012.

[31] F. Gzara and E. Erkut. A lagrangian relaxation approach to large-scale flow interception problems. *European Journal of Operational Research*, 198(2):405–411, 2009.

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

[33] H. K. Smith, G. Laporte, and P. R. Harper. Locational analysis: highlights of growth to maturity. *Journal of the Operational Research Society*, pages S140–S148, 2009.

[34] N. J. Santero, W. Nokes, and J. T. Harvey. Virtual weigh stations: The business case. 2005. `http://caltrans.ca.gov/research/researchreports/2002-2006/2005/ucprc-tm-2005-03.pdf`.

[35] D. J. Forkenbrock. External costs of intercity truck freight transportation. *Transportation Research Part A: Policy and Practice*, 33(7):505–526, 1999.

[36] Mettler-Toledo. Virtual weigh station, 2013. http://www.mtwim.com/VirtualWeighStation.htm.

[37] Cardinal. Weigh-in-motion implementation, 2013. http://www.cardinalscale.com/wp-content/uploads/2012/11/QWIM-1-1024x835.jpg.

[38] J. Stephens, J. L. Carson, D. Reagor, and M. Harrington. *An Evaluation of Montana's State Truck Activities Reporting System*. Montana Department of Transportation, 2003. http://www.mdt.mt.gov/publications/docs/brochures/mcs_stars_report.pdf.

[39] S. A. AlGadhi. Optimizing truck weigh stations' locations on the highway network of Saudi Arabia. *Journal of King Saud University*, 22(1):01–19, 2002.

[40] M. Šelmić, N. Bešinović, and D. Teodorović. Locating weigh-in-motion checkpoints in traffic networks using genetic algorithm. *E-Society Journal*, 2(1):55–66, 2011.

[41] B. H. Cottrell Jr. The avoidance of weigh stations in Virginia by overweight trucks. Technical report, Virginia Transportation Research Council, 1992.

[42] N. Bešinović, N. Marković, and P. Schonfeld. Optimal allocation of truck inspection stations based on k-shortest paths. In *Transportation Research Board 92nd Annual Meeting*, number 13-3123, Compendium of Papers DVD, 2013.

[43] L. N. Moses and D. Lindstrom. *Transportation of Hazardous Materials: Issues in Law, Social Science, and Engineering Ser.* Kluwer Academic Pub, 1993.

[44] Martin Wohl and Chris T Hendrickson. *Transportation investment and pricing principles: an introduction for engineers, planners and economists.* Number 1-102. Chris Hendrickson, 1984.

[45] A. Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *Journal of Combinatorial Theory, Series B*, 80(2):346–355, 2000.

[46] S. Iwata, L. Fleischer, and S. Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *Journal of the ACM (JACM)*, 48(4):761–777, 2001.

[47] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions-I. *Mathematical Programming*, 14(1):265–294, 1978.

[48] John R Birge. The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical programming*, 24(1):314–325, 1982.

[49] G. Laporte and F. Louveaux. The integer l-shaped method for stochastic integer programs with complete recourse. *Operations research letters*, 13(3):133–142, 1993.

[50] M. G. Kay. Matlog: logistics engineering matlab toolbox, 2006. http://www.ise.ncsu.edu/kay/matlog/.

[51] B. Peterson. Description of the Oak Ridge national highway network, version 1.3. *Center for Transportation Analysis, Oak Ridge National Laboratory, Oak Ridge, Tenn. http://www-cta. ornl. gov*, 1997.

[52] J. Y. Yen. Finding the k shortest loopless paths in a network. *Management Science*, 17(11):712–716, 1971.

[53] J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Verlag, 1997.

[54] R. Bellman. Dynamic programming, princeton univ. *Prese Princeton, 19S7*, 1957.

[55] H. P. Simão, J. Day, A. P. George, T. Gifford, J. Nienow, and W. B. Powell. An approximate dynamic programming algorithm for large-scale fleet management: A case application. *Transportation Science*, 43(2):178–197, 2009.

[56] W. B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. Wiley-Interscience, 2007.

[57] H. Topaloglu and W. B. Powell. Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems. *INFORMS Journal on Computing*, 18(1):31–42, 2006.

[58] G. Allen, A. Moruza, and B. Diefenderfer. Oversize and overweight vehicle studies, August 2010. http://dls.virginia.gov/GROUPS/transaccount/meetings/080410/oversize.pdf.

[59] Federal-Highway-Administration. Major flows by truck to, from, and within Nevada: 2007 and 2040. 2011. http://ops.fhwa.dot.gov/freight/freight_analysis/state_info/nevada/truckflow.htm.

[60] D. Andersen. I-15 corridor system master plan. 2011. http://www.i15alliance.org/pdfs/tech_memos/freight_goods_movement/I-15CSMP_Freight_Ports_FINAL.pdf.

[61] R. Bushman and A. J. Pratt. Weigh in motion technology–economics and performance. In *Presentation on the North American Travel Monitoring Exhibition and Conference (NATMEC). Charlotte, North Carolina*, 1998.

[62] NevSta. Locations of weigh stations in Nevada, 2013. http://www.allstays.com/c/weigh-scales-nevada-locations-map.htm.