OPTIMIZATION APPROACHES TO PROTECT TRANSPORTATION INFRASTRUCTURE AGAINST STRATEGIC AND RANDOM DISRUPTIONS

June 2016

A thesis submitted to
The University of Kent
In the subject of Management Science
For the degree
Of Doctor of Philosophy

by
Stefano Starita

To my family

Contents

List of Figures

List of Tables

Abstract

Past and recent events have proved that critical infrastructure are vulnerable to natural catastrophes, unintentional accidents and terrorist attacks. Protecting these systems is critical to avoid loss of life and to guard against economical upheaval. A systematic approach to plan security investments is paramount to guarantee that limited protection resources are utilized in the most efficient manner. This thesis provides a detailed review of the optimization models that have been introduced in the past to identify vulnerabilities and protection plans for critical infrastructure. The main objective of this thesis is to study new and more realistic models to protect transportation infrastructure such as railway and road systems against man made and natural disruptions. Solution algorithms are devised to efficiently solve the complex formulations proposed. Finally, several illustrative case studies are analysed to demonstrate how solving these models can be used to support efficient protection decisions.

Acknowledgement

First and foremost, I would like to say a very big thank you to my supervisors Dr Maria Paola Scaparra and Dr Jesse O'Hanley. They have been supportive from the very first day, providing the guidance and feedback I needed. It has been an honour working with you. Many thanks also to Dr Claudio Sterle and Prof Antonio Sforza for convincing me to undertake this Ph.D.

I would also like to thank my examiners Dr Nagy and Prof Bektas for their insightful feedbacks which have significantly improved this thesis.

I am grateful to all the colleagues and friends I have met in the past three years. We had a lot of fun!

Finally, a special thanks to my family for their unconditional support and encouragement.

1 Introduction

The European Program for Critical Infrastructure Protection (EPCIP), defines critical infrastructure systems as that *"which are of the highest importance for the Community and which if disrupted or destroyed would affect two or more member states"* (EU COM, 2006). Nowadays, the well-being of society relies heavily on the proper functioning of infrastructure such as transportation, telecommunication, energy supply, and information. Planning and protecting infrastructure systems is a complex task, especially because of their large sizes and interdependencies. Even small, random disruptions can severely affect the normal functioning of one or more infrastructure systems. Intelligent attacks or large natural catastrophes can have even more dramatic consequences in terms of economic damage and loss of life. Examples of such events include the 1995 Paris metro bombing, the 2004 Madrid train bombing, the 2005 London underground suicide attacks, the 2010 Moscow bombing, and the 2016 Brussels attacks. Most recently, severe floods hit some western regions of the UK and forced Network Rail to pay £12.5M for disrupted services and a further £15M to repair the rail network (Wintour and Topham, 2014). It is therefore paramount to protect infrastructure systems in the most efficient way so as to guarantee continuity in service provision and safety for users in the event of disruption.

A critical aspect in planning infrastructure protection is the scarce availability of protection resources. Protecting all the components of an infrastructure system to desired safety levels is often cost prohibitive. For example, the Kent (UK) railway system serves 179 stations and has 1094 miles of tracks. Protecting every station and all the tracks is economically impossible.

1.1 Research contributions

In this thesis my aim is to contribute to research on transportation systems protection planning. Novel optimization models are formulated and solved that incorporate characteristics and issues that have been considered or neglected in the literature. Specifically, my primary research contribution includes:

- A model is introduced to protect a railway network system. The aim is to identify the set of tracks and stations to protect so that the impact of a worst case disruption is minimised. Impact is estimated in terms of post-disruption unserved demand. The model is tested on an Italian railroad network instance.

- A second model is studied that extends the previous one by adding a temporal component. Considering the possibility that budget is allocated over time renders the model more realistic but more complex as well. Two different solution approaches are proposed and tested on the Kent railway network.

- A further extension of the first model is also presented which tries to capture user-behaviour in a post-disruption state. Specifically, it is assumed that the demand for railway travel changes with the extent of post-disruption delays. A heuristic is developed to solve the problem efficiently and tested on the London tube network.

- The final model proposed switches from focusing on worst-case to random disruption scenarios and deals with protecting a road network against flooding. A heuristic is devised to efficiently solve the problem. Key insights from a case study of the Hertfordshire A-road network are made.

1.2 Outline

The remainder of this thesis is organized as follows. Section 2 provides a detailed review of the literature related to this current work. In Section 3, I discuss the various methodologies utilized in model development and solution methods. Sections 4, 5, and 6 describe three novel models for protecting railway infrastructure against worst-case disruptions. In Section 7, I introduce a model for the protection of road networks against flooding. Finally, in Section 8, I provide some conclusions and recommendations for future research.

2 Literature review

The issue of protecting critical infrastructure systems can be investigated from different angles. Generally, the identification of sound protection strategies requires a mixed approach which includes both qualitative and quantitative aspects. In this thesis, the main focus is on quantitative research and, more specifically, on optimization models. Figures 1 and 2 show the number of documents found by Scopus©database using *interdiction models* and *fortification models* as keywords, respectively. The graphs indicate that the focus on disruption and protection models grew significantly over the last decades.

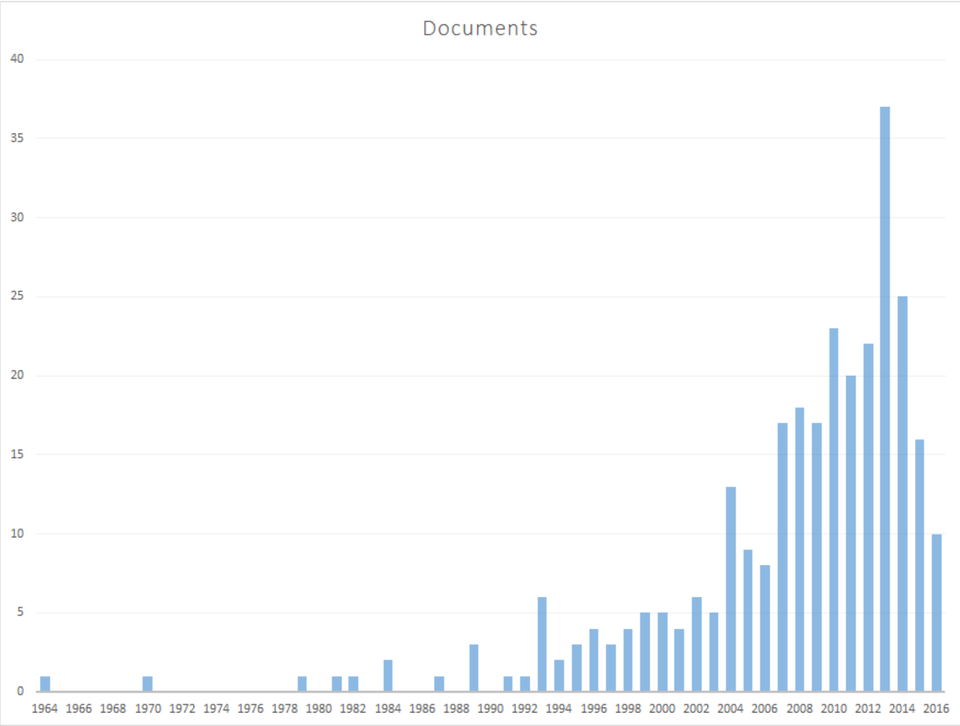


Figure 1: Number of papers on interdictions

In the optimization area, papers can be classified in different ways, for example by grouping them according to the methodology, underlying model, or type

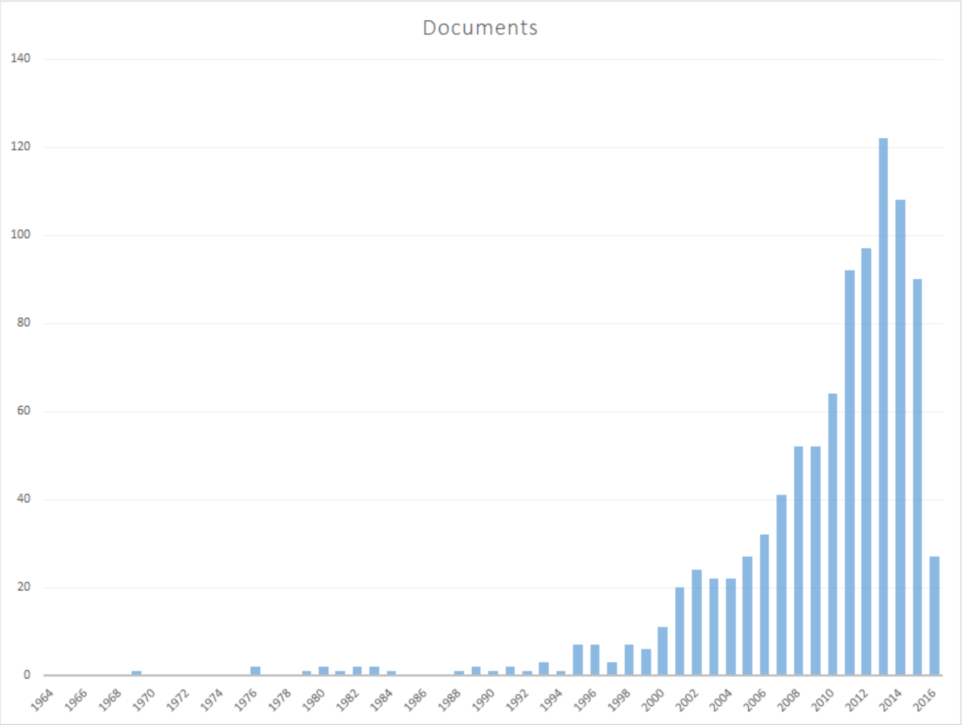


Figure 2: Number of papers on fortifications

of infrastructure.

The remainder of this chapter presents a review of the literature on optimization models for protection that has been relevant to this thesis. The main categorization has been done in terms of the methodology used to model the problem. Figure 3 outlines how research papers have been classified for the literature review.

2.1 Protection

In this category we consider all the papers that, given a system already in place, study the problem of identifying its vulnerabilities and how a protection budget should be invested to minimize the impact of disruptions.

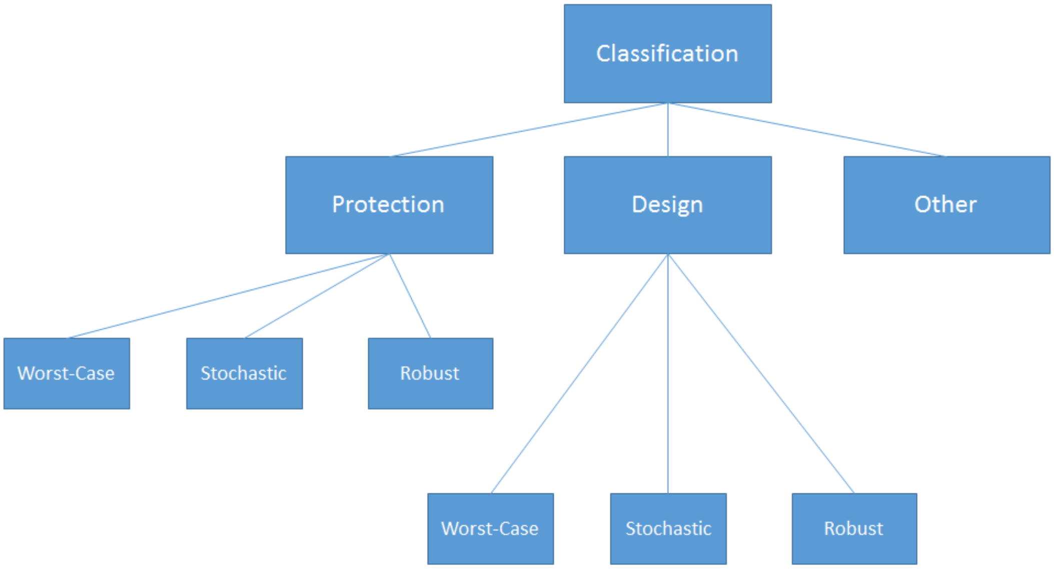


Figure 3: Classification used in the literature review

2.1.1 Worst-case scenario models

Worst-case scenario models assume that the disruption is perpetrated by an intelligent actor who knows the actions to take so as to inflict the worst damage possible to the system. These models are particularly suitable to plan against terrorist and deliberate attacks. Nonetheless, decision makers often need to have a risk-averse approach against natural events as well to safeguard assets and system's users. As a consequence, these models can be used to drive decisions also when planning protection against random natural disasters.

Interdiction models Interdiction models study the problem from an intelligent attacker point of view. These models simulate the game between two actors: the attacker and the system user. The attacker aims at inflicting the highest possible damage, the system user aims at minimizing the operational cost or maximizing the system's value. Interdiction models are generally bi-level (each level represents an actor). By solving an interdiction model one can identify the critical assets of a system, i.e. the assets that, when disrupted have the greatest impact on the ability of the system to perform its functions. Several papers have dealt with interdiction problems on flow-based networks. The seminal work by Wollmer (1964) proposed an algorithm to identify the n arcs that, when removed, reduce the max-flow the most. Later on, Wood (1993) proposed an integer programming formulation of the problem, where the interdictor is subject to a budget constraint. They formulated further extensions with partial arc interdiction, multiple sources and sinks, undirected networks, multiple resources and multiple commodities. Finally, they introduced valid inequalities to tighten the formulation. Different solution approaches to this problem were studied by Royset and Wood (2007). They solved the problem with a Lagrangian relaxation-based algorithm and with a customized branch and bound with partial path enumeration algorithm. Cormican et al. (1998) further extended the interdiction problem on max-flow network by considering the outcome of an interdiction uncertain. They proposed lower and upper bounds used in a sequential approximation solution algorithm. They also studied other cases where arc capacities are uncertain or can assume a discrete number of realizations, and multiple interdictions are attempted. Lim and Smith (2007) applied the interdiction model to multi-commodity flow networks. They considered both discrete interdiction (i.e., an arc is either operational or disrupted) and continuous (i.e., an arc capacity is reduced according to the amount of resources used to disrupt the arc). They introduced an exact partitioning algorithm and an approximate heuristic to solve the continuous case. Myung and Kim (2004), Murray et al. (2007), and Matisziw and Murray (2009) also studied the same problem, providing a different formulation, different solution approaches and insights obtained from different case studies. Bell et al. (2008) applied the interdiction model to study a road network vulnerability problem. They proposed a solution algorithm based on the successive averages method and a case study on central London road network. Altner et al. (2010) studied the model introduced by Wood (1993) and proposed two classes of valid inequalities. Zenklusen (2010) introduced an efficient algorithm to solve the interdiction problem in planar graphs. They studied the multi-source version of the problem and proposed an extension for single-source networks. Rad and Kakhki (2013) added a temporal component to the problem. They assumed that each arc is characterized by its traversal time and the disruption occurs in a time interval. They proposed a Benders' decomposition algorithm and discussed a second algorithm based on the concept of most vital arcs.

Shifting the focus to distance based networks, Fulkerson and Harding (1977) were the first that studied the impact of partial interdictions on arcs, to maximise the shortest path between two nodes. Other variations of shortest path interdiction problem include Golden (1978), Corely and David (1982), and Ball et al. (1989). Israeli and Wood (2002) proposed a bi-level formulation for the interdiction problem on shortest path networks. They proposed several solution approaches based on Benders' and covering decompositions, and Super Valid Inequalities (SVI). An extension of this work was introduced by Bayrak and Bailey (2008) who assumed that attacker and system user have different knowledge of the arcs' costs and delays. They proposed a solution methodology based on dualization and linearization of the problem. Yates and Sanjeevi (2013) studied the shortest path interdiction problem from a different perspective, where the attacker aims to travel to a specific target, without being detected. They considered two different types of sensors for detection and provided a case study based on California highway sub-network.

Focusing on the supply chain context, Church et al. (2004) introduced the r-interdiction problem for the p-median and max-cover problems. The problem aims at finding the r facilities that, when removed, would maximise the service's cost (p-median) or minimise demand coverage (max-cover). After this paper, the attention on interdiction models for supply chain systems increased significantly. Losada et al. (2012a) studied a problem where the outcome of a disruptions is uncertain. The authors assumed that the probability that a facility is disabled is dependent on the intensity of the disruption. Different deterministic reformulations of the problem are proposed to solve problems of realistic size to optimality. Aksen et al. (2014) extended the r-interdiction median problem by considering capacitated facilities and demand outsourcing. The capacity of each facility is dependent on the interdiction resources used to disrupt the facility itself. The excess of the demand is outsourced to an external supplier. Two solution algorithms are proposed: a progressive grid search which is targeted for small-medium size instances and a more efficient multi-start simplex search. Zhang et al. (2015) studied the interdiction problem when facilities have limited capacity, the disruption is partial, and customers can be served by multiple facilities. They developed a specific heuristic to solve the bi-level problem.

Protection models Protection or fortification models address the problem of identifying the optimal protection plans (i.e., the optimal allocation of protection resources) so that the consequences of worst case disruptions are minimised. Church and Scaparra (2007) proposed first the fortification model for the p-median problem. The same problem is formulated as a bilevel mixed-integer model in Scaparra and Church (2008a). Aksen et al. (2010) extended previous works by introducing a budget constraint for the protection. They further assumed that the capacity of a facility can be increased at a fixed known cost. The problem is solved with an implicit enumeration algorithm implemented on a binary tree. Liberatore et al. (2011) extended the p-median fortification problem by considering the number of disrupted facilities uncertain. They proposed a max-covering based formulation and an efficient solution approach using upper and lower bounds. A heuristic, based on heuristic concentration-type rules, is also presented. Liberatore et al. (2012) studied the protection of a p-median network subject to disruption with ripple effect. Specifically, they used a correlation matrix which indicates the level of disruption of each facility when a specific facility in the vicinity is targeted for disruption. They solved the problem using an exact approach based on a three search. Scaparra and Church (2012) introduced facility capacity constraints in the fortification problem. An implicit enumeration algorithm is proposed to solve the problem to optimality. Losada et al. (2012) considered the issue of recovery time in the context of the p-median fortification problem. The problem is studied for a number of time periods and each facility requires a recovery time to return to its normal functioning, after an interdiction. They solved the model by using three different decomposition methods. Aksen et al. (2013) considered the problem where the defender must decide both the location of facilities and the allocation of protective resources to minimise worst case disruptions. They proposed an exact algorithm, a tabu search heuristic and a sequential solution method. Zhu et al. (2013) presented a fortification model focused on distributing anti-aircraft protection resources on a supply system. Each resource has a probability to intercept an attack. A target can be disrupted only if all the protective resources assigned to that target fail to intercept the attack. They proposed a greedy search to obtain good approximations of the optimal solutions. Aliakbarian et al. (2015) studied the fortification problem for the p-median problem where facilities are organized hierarchically. They implemented an exhaustive enumeration method and three meta-heuristics to solve the problem.

Less attention has been given to fortification problems for distance-based and flow-based networks. Cappanera and Scaparra (2011) added the fortification layer to the shortest path interdiction problem. They proposed an implicit enumeration algorithm and a heuristic solution for the resulting fortification problem. Alguacil et al. (2014) proposed a model to find the optimal allocation of defensive resources on an electric power grid subject to intelligent attacks. They solved the problem using dualization and an implicit enumeration algorithm. Jenelius et al. (2010) studied the protection problem for a generic infrastructure when the attacker has an imprecise perception of the system. They showed that taking into account the attacker's perception leads to different protection investment plans. Sarhadi et al. (2015) introduced a tri-level fortification formulation for rail inter-modal terminal networks. Their model embeds a capacitated multi-commodity flow problem with delivery times and penalty costs. To solve the model, they studied explicit and implicit enumeration algorithms and a traffic based heuristic. Jin et al. (2015)

developed a model to allocate protection resources optimally over an urban rail network threatened by intentional attacks. They considered that rail stations suffer disruptions of different levels, resulting in a decrease of their inbound and outbound capacity. They solved the model with a nested variable neighborhood search algorithm.

2.1.2 Stochastic models

A different approach to the study of infrastructure vulnerabilities and protections entails abandoning the worst-case approach in favour of a stochastic approach. Stochastic models assume that one or more variables of the problem are random. The aim of the planner becomes, in general, to optimize an expected cost or system value. He and Liu (2012) modelled the evolution of daily traffic when the network is subject to disruptions. They implemented a prediction-correction model to evaluate how the traffic flow evolves depending on drivers knowledge. Liu et al. (2009) studied the problem of fortifying transportation links to reduce a system's loss when probabilistic disruptions occur. They solved the model using an algorithm that combines an L-shaped method with Benders' decomposition. Fan and Liu (2010) introduced a two stage stochastic model to find the protection plan that minimizes the expected physical and social losses following transportation network disruptions. In their model users choose the best route according to their perception. They used the progressive hedging method to decompose and subsequently solve the problem. Peeta et al. (2010) addressed the problem of distributing pre-disaster resources in a shortest path based network subject to random failures. They reformulated the objective function and obtained a local optimum using a knapsack formulation. Qin et al. (2013) introduced a two-stage stochastic model to plan fortifications for a capacitated logistic problem. They focused on minimizing the emergency inventory and transportation costs. They proposed a solution algorithm based on disjunctive decomposition-based branch-and-cut (D2-BAC). Du and Peeta (2014) presented a model to reduce the response time of transportation networks by strengthening network links. The transportation network is subject to random disasters. They also studied a two-stage heuristic to solve the model. Faturechi and Miller-Hooks (2014) analysed the road network resilience under probabilistic disasters. They considered a budget available to implement both preparedness and response actions. To evaluate the traffic flow, they added partial user equilibrium constraints, where only users directly affected by a disruption are likely to change decisions. Medal et al. (2015) proposed a two-stage stochastic problem to study the effects of disruptions on a transportation model where facilities provide relief goods to customers. They assumed that the post-hazard state of a facility is dependent on the intensity level of the hazard and on the amount of protection resources allocated to it. They introduced a greedy algorithm to obtain approximate solutions.

2.1.3 Robust models

When no probability assumptions can be made regarding to the parameters of a problem, robust optimization is frequently used. It is particularly popular when it comes to protecting against natural events. In this context, in fact, it is usually easy to find historical data that, once combined with other problem specific data, can be used to simulate different disruption scenarios. These models generally aim at finding the solutions that minimise the expected cost across all scenarios. A few more metrics have been studied and applied to model reliable infrastructures. Huang et al. (2007) studied a max-cover model for the optimal allocation of emergency vehicles over a transportation infrastructure. Scenarios were considered to contemplate variations in travel times and demand values. They finally conducted a sensitivity analysis to the budget on a case-study based on the Singapore metropolitan area. Yin (2008) studied the problem of finding the optimal allocation of tow trucks so as to allow a faster response to incident and minimise disruptions to motorways. They used a scenario-based approach where each scenario is composed of a set of incidents occurring on the network. Zhang et al. (2014) proposed a scenario based model to protect supply systems against random disruption. They also proposed an hybrid model aimed at finding the optimal protection strategies against both strategic and random attacks. They finally proposed an implicit enumeration based algorithm to solve their model.

2.2 Design models

As opposed to protection models which deal with a system that is already in place, in this category we group works that study design models to increase reliability. More specifically, we list papers that deal with the problem of identifying a design configuration that makes the system inherently robust to disruptions.

2.2.1 Worst-case scenario models

O'Hanley and Church (2011) developed a design model to identify facility locations so that a combination of pre-disaster and post-disaster coverage is maximised. They also propose two decomposition approaches based on Benders and Super Valid Inequalities. Parvaresh et al. (2013) studied the problem of locating p hubs in a network threatened by an intelligent attacker. They developed two solution algorithms based on simulated annealing. Hernandez et al. (2014) proposed a framework to design a supply chain robust to worst case disruptions. They proposed a multi-objective formulation and solved the model with Multi-Objective Evolutionary Algorithms. Finally, they displayed computational experiments run on Swain and London datasets.

2.2.2 Stochastic models

Snyder and Daskin (2005) considered the problem of identifying inexpensive and reliable facility locations for the uncapacitated fixed-charge location problem (UFLP). They modelled the problem as a multi-objective problem to consider operating and expected failure costs. They proposed a Lagrangian relaxation solution algorithm and performed an analysis to evaluate the trade-off between operating and expected costs. The same problem was studied by Cui et al. (2010), which proposed a compact mixed integer formulation and a continuum approximation. They also used Lagrangian relaxation for their solution approach. Li and Ouyang (2010) further inspected reliability issues in the UFLP considering spatially correlated disruptions. Berman et al. (2007) analysed the impact of centralization and co-location on the problem of identifying facility locations robust to disruptions. They proposed exact and heuristic approaches and provided an illustrative example to locate hospitals in Toronto. Chen et al. (2011) studied a location problem where facilities are under risk of probabilistic disruption. Among other costs to minimise, they considered the expected inventory and holding costs. They developed a Lagrangian relaxation based solution algorithm. O'Hanley et al. (2013) introduced a general approach to linearise locations problems with site-dependent failures probability. They used a flow network structure to compute compound probabilities. Li et al. (2013a) proposed a facility location model to design infrastructure systems which are reliable to inter-dependent disruptions. They introduced the concept of supporting stations to simulate spatially-correlated disruptions. Facilities are connected to at least one station and stations are subject to independent disruptions. A facility is disrupted if all the stations connected to it are disrupted. Rawls and Turnquist (2010) studied emergency response planning for hurricanes and other natural disasters. They formulated a model to determine the quantity, type, and location of supplies to be pre-positioned, while taking into account the availability of transportation links following a disruption event. Li et al. (2013b) proposed reliable formulations for the p-median and the fixed-charge location problems. They incorporated facility failure probabilities and the fortification concept, along with a budget constraint. They solved the model using a solution algorithm based on Lagrangian relaxation.

2.2.3 Robust models

An alternative way for increasing reliability is to design systems which are inherently robust to disruption. Snyder and Daskin (2006) introduced the p-robustness metric to find the network design which minimise the overall expected cost and at the same time is robust enough for each disruption scenario. Chen et al. (2006) proposed another metric, called α-reliability, aimed at minimising the expected regret by considering a subset of scenarios endogenously chosen according to their probability of occurrence. Peng et al. (2011) applied the p-robustness metric to design supply chain networks which are robust to disruption scenarios. They solved the model using a hybrid meta-heuristic combining genetic, local improvement and shortest augmenting path algorithms. Rawls and Turnquist (2012) extended Rawls and Turnquist (2010) by incorporating a regret-based reliability metric and illustrate their approach using a case study hurricanes threatening the North Carolina area. They developed an L-shaped method to solve efficiently large problems. Baghalian et al. (2013) introduced a robust model to design a supply system where demand is uncertain and manufacturers are subject to disruptions. In order to solve the model they obtained a transformation and linearization of the model. They discussed some practical results on a agri-food industry case study.

2.3 Other

In this section I review works that, despite dealing with vulnerability and protection evaluation for critical infrastructure, do not clearly fit in any of the previous categories. Chang (2003) developed a methodology to assess the performances of a transportation system in the aftermath of a disaster. They used an accessibility metric estimated in terms of the ratio between the pre and post-disaster minimum distances. They tested their methodology on the 1995 Kobe earthquake in Japan and on a hypothetical earthquake striking the Seattle area. Sohn et al. (2003) studied an approach to identify a retrofit priority for the links of a transport network. They consider disrupting each link individually and solve a multi-commodity flow problem to identify the damage in terms of demand loss and transport cost. Scott et al. (2006) introduced a new index to identify the most critical links in a transportation network. They proposed a global approach to overcome the limits of the popular volume/capacity criteria, which evaluates each element locally. Their index is computed as the difference between the transportation cost when a link is non operational and the base case cost (all links are functioning). Sun et al. (2015) proposed a model to evaluate vulnerabilities of a urban rail network. They introduced an algorithm that ranks the nodes according to the network topological efficiency (the mean of the reciprocal of the shortest paths) and passenger flow. The algorithm is tested on the Shanghai metro. The results highlight that the network is quite vulnerable to attacks to stations with larger node degree.

2.4 Conclusions

This brief literature review highlights several research opportunities. The great majority of works reviewed, tackle generic problems (p-median, max-flow, shortest path etc.) based on generic assumptions (binary interdictions and protections). This motivated my idea of moving from a generic critical infrastructure protection context to a more specific railway and road protection context. As a consequence, models presented in this thesis incorporate problem-specific issues to enhance their applicability to realistic problems.

Another interesting gap in the literature can be found by considering that post-disruption user-behaviour is often represented in a simplistic way. In Chapter 6 a way to address this issue is proposed, nonetheless more research opportunities can be pursued by linking disruption literature with non/disruption literature (e.g., revenue management).

Although a few works focusing on multiple disruptions can be found, the temporal component has been generally neglected. This is particularly true when we focus on dynamic aspects from the system planner perspective. For example, concepts such as defences deterioration over time and protection budget spread over a planning horizon have been barely inspected.

Furthermore, in network-based fortification problems it is always assumed that the system planner decides on single assets (arc/node) protections. This decision flexibility can be unrealistic in certain contexts such as flooding protection where the most effective defence measures (dikes, river diversion etc.) are not implemented along the specific asset and generally have impact on multiple assets.

One of the challenges of this research area is the trade off between complexity and applicability of these models. Adding realistic features to these optimization models might increase their complexity up to the point where they can no longer be applied to problem instances of realistic size. Conversely, over-simplifying the models can lead to inaccurate or sub-optimal results. An example of this trade off is represented by most of the works grouped as *other.* These papers study the protection problem using an algorithmic approach. They generally evaluate and rank independently and sequentially each asset of the system. On the one hand, this can obviously lead to suboptimal solutions because interactions between assets are not considered. On the other hand, the approaches introduced do not share the computational issues typical of comprehensive optimization models, therefore they are easily applicable to large real life problems.

Finally, from the literature review it is obvious that significant efforts have been done to develop efficient solution approaches. A wide and heterogeneous range of exact and approximate algorithms have been studied. Nonetheless, the likely increase in the models' complexity will require a continuous refinement of existing methodologies and the development of new solution techniques.

To summarize and conclude, I think that the literature reviewed provides a solid methodological ground that should be used by future researchers to target much more realistic problems.

3 Methodology

The common idea behind the models proposed in this thesis is to find the allocation of protective resources which minimizes the consequence of disruptions on an infrastructure system. This kind of problem can be effectively studied using mathematical programming, where the purpose is to optimise a function subject to constraints. Therefore, mathematical programming is the broad methodology used in this thesis to implement a systematic approach to the problem of protecting critical infrastructures. Specifically, we use two mathematical programming frameworks: multi-level and scenario-based models.

3.1 Multi-level models

Multi-level formulations have been widely used in the critical infrastructure protection context. Much like hypergame analysis (Bennett, 1977), these models simulate a game between two intelligent actors that make their choices to maximise their opposite benefits, while observing each others' strategies. Due to their structure, hypergames can be generally applied to small, simple problems. On the other hand, multi-level models studied in this thesis can be used on realistic size problems. As hypergame problems, multi-level models have been firstly applied to military and anti-terrorism context. Using a framework where two actors have diametrically opposed aims is indeed particularly helpful to study such problems. Nonetheless, recent literature applied this approach to the problem of protection of critical infrastructure against *generic* threats. The reason why a framework where two actors share the objective function with opposite aims can also be applied to study non man-made disruptions lays in the definition of critical infrastructure. As explained in the introduction chapter, well-being of people is highly dependent on how well such critical systems work. Consequently, risk-adverse approaches are frequent when protecting against both intelligent and random disruptions.

Furthermore, planning protection strategies for critical infrastructure is obviously a complex task which cannot be summarized by a single optimization model. A range of methods and structuring issues (Rosenhead and Mingers, 2001) must be considered. In this context, solving worst-case scenario models should be seen as one of those methods which provide information that can be used to support decision making. Brown et al. (2006) provided a detailed analysis on how these models are used for critical infrastructure defence.

Two frameworks have been generally used in the literature: *attacker-user* or *interdiction* models and *defender-attacker-user* or *protection* models.

The *attacker-user* models can be formulated as follows:

$$\max_{x \in \mathbb{X}} \min_{w \in \mathbb{W}(x)} \mathbf{c^t w}. \tag{1}$$

Variables w are decision variables used to estimate the system's cost. Variables x represent interdiction strategies. Problem (1) models a game between two actors with opposite aims. The *user* level aims at finding the system's configuration that minimizes the overall costs. On the other hand, the *attacker* seeks to identify the most disruptive interdiction strategy. Let $(\hat{x}, \hat{w})$ be the optimal solution of problem (1), interdiction strategy $\hat{x}$ represents the disruption scenario that generates the highest possible damage in terms of cost increment. The solution sheds light on what are the vulnerabilities of the infrastructure.

If the aim is to optimally distribute protection resources among the assets of the system, $\hat{x}$ is generally different from the optimal protection plan (Church and Scaparra, 2007). To find the best protection strategy a further level needs to be added to the interdiction problem so as to explicitly model the allocation of protection resources. These models are called *defender-attacker-user* models and can be formulated as follows:

$$\min_{y \in \mathbb{Y}} \max_{x \in \mathbb{X}(y)} \min_{w \in \mathbb{W}(x)} \mathbf{c^t w}. \tag{2}$$

Variables y identify the protection strategies. Problem (2) introduces a third actor, the *defender*, whose aim is to distribute protection resources among the system's assets so that the impact of worst-case disruptions is minimised.

3.2 Scenario-based models

When dealing with natural events, worst-case models can be overcautious by leading to protection plans aimed at very unlikely events. Generally, risk analysis for events such as earthquakes, flooding, tornadoes etc., highly relies on past experience. Consequently, it can be convenient to incorporate the risk in the problem as a set of disruption scenarios. Scenario-based models (also referred to as robust models) can be formulated as follows:

$$\min_{y \in \mathbb{Y},\, w \in \mathbb{W}} \sum_{s \in S} \pi_s \mathbf{c^t}_s(y)\mathbf{w}. \tag{3}$$

S is a set of disruption scenarios and each scenario s occurs with probability π_s. Costs c_s are now scenario-indexed to consider the effect that scenario s has on the system's cost. The aim of the model is to find the optimal allocation of protection resources so that the expected system's cost is minimised.

A popular metric often used for this kind of problems was introduced by Snyder and Daskin (2006). Their aim was to obtain solutions that minimise the expected system's cost and at the same time bound the relative regret of each scenario. The metric is called p-robustness and is implemented by adding the following constraints:

$$\mathbf{c^t}_s(y)\mathbf{w} \leq (1+p)Z_s^* \qquad \forall s \in S. \tag{4}$$

where Z_s^* is the optimal value of problem (3) when only scenario s is considered. The relative regret for solution $\mathbf{w}$ is defined as follows:

$$\frac{\mathbf{c^t}_s(y)\mathbf{w} - Z_s^*}{Z_s^*}. \tag{5}$$

Therefore, constraints (4) bound the regret of each scenario to be less or equal than p.

Regret is a popular metric used to drive decisions when some parameters of a problem are uncertain. Several other regret-based metrics have been studied in the past. Ghosh and McLafferty (1982) investigated a scenario-based location problem where they minimize the sum of the regrets or the sum of the square regrets. Daskin et al. (1997) and Chen et al. (2006) proposed the α-Reliable Minimax regret and the α-reliable mean-excess regret metrics respectively. Both these metrics are implemented to minimize the expected regret of a subset of scenarios that are selected based on their probabilities.

3.3 Solution approaches

Solving multi-level models can be particularly challenging. Algorithms like branch and cut cannot be applied directly due to the nested structure of these models. Two approaches can be used:

- The structure of the problem can be modified so that general purpose algorithms can be applied,
- Ad-hoc exact methods or approximate heuristics are devised.

Here a brief discussion of the approaches that have been used to solve the models proposed in this dissertation is provided. To better explain the methodologies I will introduce the shortest path interdiction problem (Israeli and Wood, 2002) and discuss how those methodologies can be applied to this problem.
The network we consider is connected (i.e., there is at least one path connecting origin to destination). The notation used for the formulation is as follows:

- $i \in N$ is the index for the nodes,
- $k \in A$ is the index for the arcs,
- $FS(i)$ is the forward star of node i,
- $RS(i)$ is the reverse star of node i,
- $o \in N$ is the origin node,
- $d \in N$ is the destination node,
- c_k is the nominal length of arc k,

- d_k is the delay ($< \infty$) when k is interdicted,
- P is the maximum number of disrupted arcs,
- $w_k = 1$ if arc k is in the shortest path; 0 otherwise,
- $x_k = 1$ if arc k is disrupted; 0 otherwise.
- $\hat{W}$ is a set of paths

The shortest path interdiction problem is formulated as follows:

$$[\text{SPI}] \quad Z = \max_{\mathbf{x}} \min_{\mathbf{w}} \sum_{k \in A} (c_k + x_k d_k) w_k \tag{6}$$

$$\text{s. t.} \quad \sum_{k \in RS(i)} w_k - \sum_{k \in FS(i)} w_k = \begin{cases} 1 & \text{if } i = o \\ 0 & \forall i \in N \setminus \{o, d\} \\ -1 & \text{if } i = d \end{cases} \tag{7}$$

$$\sum_{k \in A} x_k \leq P \tag{8}$$

$$w_k \geq 0 \qquad \forall k \in A \tag{9}$$

$$x_k \in \{0, 1\} \qquad \forall k \in A \tag{10}$$

The aim of the interdictor is to disrupt at most P arcs so that the shortest path length (6) between o and d is maximised. Constraints (7) are the flow balance constraints. Constraint (8) imposes that the maximum number of disrupted arcs is P. Finally, constraints (9) and (10) are the domain restrictions for the variables $\mathbf{w}$ and $\mathbf{x}$.

3.3.1 Decomposition approaches

Both decomposition approaches discussed in this thesis can be put into the more generic context of Combinatorial Benders' cuts (Codato and Fischetti, 2006). These cuts have been introduced to efficiently solve Mixed-Integer Programs with constraints using big-M coefficients. The approached used is to decompose the program into a master (with binary only variables x) and a slave (with continuous variables w). Details of these approach are beyond the scope of this work, nonetheless it is interesting to briefly summarize the main concept underlying this methodology so that the reader can appreciate the similarities with the decomposition approaches introduced for our models. An iterative algorithm based on Combinatorial Benders' cuts firstly solves the master problem obtaining an optimal solution $\hat{x}$. If the slave model is feasible with regard to $\hat{x}$, then it can be solved to obtain $\hat{w}$ and $(\hat{x}, \hat{w})$ is an optimal solution for the original problem. Conversely, if the slave model is infeasible then the algorithm adds a cut to eliminate solution $\hat{x}$ from the master problem.

Benders' decomposition Benders decomposition (Benders, 1962) is a technique that has been devised to solve large linear programming problems. Same technique was subsequently generalized to tackle non linear problems (Geoffrion, 1972). The idea behind this approach is to decompose a large problem by identifying a partition of the variables and invoking the dual representation of the sub-problems generated.

It can be also successfully used for multi-level problems such as SPI. In this thesis the decomposition approach is slightly different from the generic one in that we do not consider dual representations of the sub-problems because of the binary requirements of some variables. Other than that, the algorithms introduced are quite standard and focus on solving decomposed problems iteratively until an optimal solution is found. Focusing on the Shortest Path Interdiction problem, the model is decomposed as follows:

$$[\ \text{INNER}(\hat{\mathbf{x}})\ ] \quad Z_i = \min_{\mathbf{w}} \sum_{k \in A} (c_k + \hat{x}_k d_k) w_k \tag{11}$$

$$(7), (9)$$

$$[\ \text{MASTER}(\hat{W})\ ] \quad Z_m = \max_{\mathbf{x}} z \tag{12}$$

$$z \leq \sum_{k \in A} (c_k \hat{w}_k + x_k d_k \hat{w}_k) \qquad \forall\, \hat{\mathbf{w}} \in \hat{W} \tag{13}$$

$$(8), (10)$$

The decomposition algorithm is summarized in figure 4:

INNER($\hat{\mathbf{x}}$) finds the shortest path given an interdiction strategy $\hat{\mathbf{x}}$. The solu-

Algorithm 1 Benders decomposition algorithm

```
$\hat{W} \leftarrow \emptyset$; $Z_{inf} = -\infty; Z_{sup} = \infty; \hat{\mathbf{x}} \leftarrow \mathbf{0}; \mathbf{x}^{\mathbf{best}} \leftarrow \mathbf{0}$
while $Z_{sup} - Z_{inf} > 0$ do
    Solve INNER($\hat{\mathbf{x}}$) to find $Z_i$ and $\hat{\mathbf{w}}$;
    $\hat{W} = \hat{W} \cup \{\hat{\mathbf{w}}\}$;
    if $Z_{inf} < Z_i$ then
        $Z_{inf} = Z_i$ and $\mathbf{x}^{\mathbf{best}} \leftarrow \hat{\mathbf{x}}$;
    end if
    if $Z_{sup} - Z_{inf} < 0$ then
        goto END;
    end if
    Solve MASTER($\hat{W}$) to find $Z_m$ and $\hat{\mathbf{x}}$;
    $Z_{sup} = Z_m$;
end while
END    return ($\mathbf{x}^{\mathbf{best}}$, $Z_{inf}$);
```

Figure 4: Benders decomposition algorithm.

tions is a lower bound on the interdiction problem. On the contrary, MASTER($\hat{W}$) is the interdiction problem applied to a subset of paths identified by set $\hat{W}$. Its solution is an upper bound for SPI. The algorithm (4) solves iteratively the two problems, updating the set of discovered paths and the bounds. The procedures converges to the optimal solution when the lower and upper bound become equal. Examples of papers that implemented Benders' decomposition solution approaches are Israeli and Wood (2002), Losada et al. (2012), and Rad and Kakhki (2013).

Super Valid Inequalities (SVIs) Another popular decomposition approach for multi-level models is based on the use of super-valid inequalities. A super-valid inequality is a cut that eliminates the incumbent solution and may eliminate other feasible solutions as well. Nonetheless, this cut does not eliminate any optimal solution unless an optimum has already been found.

A SVI for the SPI problem can be devised by noticing that an optimal interdiction must include at least one arc belonging to the shortest path. This observation can be expressed mathematically as the following inequality:

$$\mathrm{SVI}(\hat{\mathbf{w}}) : \sum_{k \in A} \hat{w}_k x_k \geq 1. \tag{14}$$

Proposition. $SVI(\hat{\mathbf{w}})$ is supervalid.

Proof. : if solution $(\hat{\mathbf{x}}, \hat{\mathbf{w}})$ is not optimal, (14) will yield a new interdiction strategy $\hat{\mathbf{x}}' \neq \hat{\mathbf{x}}$. As a consequence a new shortest path $\hat{\mathbf{w}'}$ is generated. The inequality eliminates the incumbent solution, in fact $(\hat{\mathbf{x}}, \hat{\mathbf{w}}) \neq (\hat{\mathbf{x}}', \hat{\mathbf{w}}')$. Furthermore, if the incumbent solution is optimal, the inequality is super-valid by definition. This proves, therefore, that (14) is super-valid.

For this approach, the problem is again decomposed into an inner and master problem. The inner problem is the same as the one introduced earlier. The master problem is a feasibility seeking problem, with an empty objective function and is constrained by (8) and (10).

The solution algorithm using the SVIs is shown in figure 5.

At each step, the algorithm solves INNER($\hat{\mathbf{x}}$) and adds the correspondent SVI to the MASTER sub-problem. The procedure stops when the MASTER becomes unfeasible (i.e. there are not enough interdiction resources to thwart all the shortest paths identified in previous iterations). Examples of papers that implemented SVI based solution approaches are Israeli and Wood (2002), O'Hanley and Church (2011), and Losada et al. (2012).

Algorithm 2 SVIs decomposition algorithm

```
$Z_{opt} = -\infty; \hat{\mathbf{x}} \leftarrow \mathbf{0}; \mathbf{x}^{\mathbf{best}} \leftarrow \mathbf{0}$
while MASTER is feasible do
    Solve INNER($\hat{\mathbf{x}}$) to find $Z_i$ and $\hat{\mathbf{w}}$;
    add SVI($\hat{\mathbf{w}}$) to MASTER problem
    if $Z_{opt} < Z_i$ then
        $Z_{opt} = Z_i$ and $\mathbf{x}^{\mathbf{best}} \leftarrow \hat{\mathbf{x}}$;
    end if
end while
END    return ($\mathbf{x}^{\mathbf{best}}$, $Z_{opt}$);
```

Figure 5: SVIs decomposition algorithm.

3.3.2 Greedy heuristic

Sometimes the structure and the complexity of a problem makes almost impossible to implement exact algorithms. Often it is impossible to find the optimal value because of excessively long computing time or memory requirement impossible to meet. Consequently, non-exact solution approaches sometimes are necessary. A greedy algorithm is a basic heuristic which builds the solution by repeatedly making choices that are locally optimal. Generally these algorithms are not precise enough and need to be complemented with other more extensive search procedures.

Given a vector $\mathbf{x}$, let us define $\mathbf{x}^k$ as follows:

$$x_j^k = \begin{cases} x_j & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases} \tag{15}$$

In Figure 6 the greedy approach is applied to the shortest path interdiction problem.

Algorithm 3 Greedy procedure

```
$\mathbf{x} \leftarrow \mathbf{0}; \mathbf{gf} \leftarrow \mathbf{0}$;
for disr = 0; disr < P; disr++ do
   for each $k \in A$ do
      Solve INNER($\mathbf{x^k}$) to find $Z_i$;
      $gf_k = Z_i$
   end for
   Select $\hat{k}$ such that $gf_{\hat{k}} \geq gf_k \; \forall k \in A$;
   $x_{\hat{k}} = 1$;
end for
Solve INNER($\mathbf{x}$) to find $Z_i$;
$Z_{best} = Z_i$;
END   return ($\mathbf{x}$, $Z_{best}$);
```

Figure 6: Greedy procedure.

Specifically, the greedy procedure always adds to the solution the arc whose disruption will cause the highest increase in the shortest path length.

3.3.3 Greedy Randomized Adaptive Search Procedure (GRASP)

Here, I discuss a meta-heuristic known as GRASP, which proved to be efficient when applied to some protection problems proposed in this thesis.
GRASP was introduced by Feo and Resende (1995) and provides a more flexible and sophisticated approach compared with a pure greedy algorithm. The idea is to use a greedy metric not to select an element to add to the solution but to identify a set of candidates, called restricted candidate list (RCL). The element to add to the solution is chosen randomly from this list. Because of its random nature, the algorithm is iterated for a fixed number of steps so that a larger space of solutions is investigated (Figure 7). Finally, a local search is performed to improve the best solution found.

```
Algorithm 4 GRASP procedure
 $Z_{best} = -\infty; \mathbf{x^{best}} \leftarrow \mathbf{0}$
 for iter = 0; iter < MAX_ITER; iter++ do
   $\hat{\mathbf{x}} \leftarrow \mathbf{0}$;
   for disr = 0; disr < P; disr++ do
     RCL = $buildRCL()$;
     arc $k^*$ is selected randomly from RCL;
     $\hat{x}_{k^*} = 1$;
   end for
   Solve INNER($\hat{\mathbf{x}}$) to find $Z_i$;
   if $Z_i > Z_{best}$ then
     $Z_{best} = Z_i$ and $\mathbf{x^{best}} \leftarrow \hat{\mathbf{x}}$;
   end if
 end for
 ($\mathbf{x^{best}}$, $Z_{best}$) = $localSearch()$;
 END   return ($\mathbf{x^{best}}$, $Z_{best}$);
```

Figure 7: GRASP procedure.

In figure 7 there is an example of GRASP applied to the SPI problem.

The RCL is built using a parameter α which is initialized to a value smaller than 1 (for $\alpha = 1$ GRASP becomes a pure greedy approach, $\alpha = 0$ a random approach). Different types of local search procedures can be implemented, from single arc swap to more sophisticated searches based on the network topology.

3.3.4 Other solution approaches

To provide a better understanding of the solution algorithms available to solve multi-level models, here I list some papers where solution approaches different from the ones previously introduced have been implemented. Scaparra and Church (2008a) and Cappanera and Scaparra (2011) developed implicit enumeration algorithms to solve two and tri-level models. Losada et al. (2012a) solved their stochas-

Algorithm 5 buildRCL procedure

```
$\mathbf{gf} \leftarrow \mathbf{0}$ and $gf_{max} = -\infty$;
for each $k \in A$ do
    Solve INNER($\hat{\mathbf{x}}^k$) to find $Z_i$;
    $gf_k = Z_i$;
    if $gf_k > gf_{max}$ then
        $gf_{max} = gf_k$;
    end if
end for
add to RCL all arcs $k$ such that $gf_k \geq \alpha gf_{max}$;
END    return RCL;
```

Figure 8: Build RCL procedure.

tic interdiction median problem by reformulating into a single-level deterministic model. Lim and Smith (2007) proposed a partitioning algorithm. Aliakbarian et al. (2015) studied a Simulated Annealing (SA), a Variable Depth Neighborhood Search (VDNS) and a combination of SA and VDNS to solve their bi-level model. Liberatore et al. (2012) reformulated their tri-level protection model as a single level model by using dualization. Aksen and Aras (2012) implemented a Tabu search heuristic to solve their model.

4 Optimizing investment decisions for railway systems protection

This chapter presents a mathematical model for identifying the optimal allocation of protective resources among the components of a railway network. The aim is to minimize the impact on passenger flow of worst-case disruptions which might affect both railway stations and tracks. The proposed model is tested on an Italian railway system network to demonstrate how the model results can be used to inform policy making and protection investment decisions.

In light of numerous recent terrorist attacks to transportation systems, the issue of protecting critical transportation infrastructures has become a necessity. Railways, in particular, have often been the target of terrorist activity. Examples include the 1995 Paris metro bombing, the 2004 Madrid train bombing, the 2005 London underground suicide attacks, and the 2010 Moscow bombing. These events have demonstrated that rail systems are a crucial yet sensitive component of a nation's infrastructure and that disruptions in railway system services can have a significant adverse impact not only on the economy but also on public health and safety.

In some countries like the US, the rail industry and the government have undertaken extensive efforts to protect the movement of freight and passenger trains. Nevertheless, rail security remains an exercise in risk mitigation, as opposed to risk prevention, and protection efforts are mostly undermanned and underfunded (Hartong et al., 2008). Undoubtedly, railway protection presents some inherent difficulties, due to the specific characteristics of rail systems. First of all, railways are geographically extensive, open and easily accessible infrastructures. As an example, the Italian railroad comprises 16,741 kilometers of operational rail lines, and 2,260 passenger stations. Strengthening all these assets to targeted safety levels may require unacceptable expenditures. In addition, effective security improvements specific to rail transport are difficult to identify and implement. Security mechanisms used by other transportation modes (e.g., aviation passenger screening) cannot be readily applied in the rail environment. Given these difficulties, it is key that protection expenditures are invested wisely in a manner that optimises both service efficiency and public safety.

Railway security can be improved by optimizing the allocation of protection devices within a single asset (e.g., security cameras in a station) but also through a cost efficient allocation of protective resources across the entire railway network. This involves identifying the most critical network components whose loss or temporary closure might have the greatest impact on daily service provision and allocating protection resources among these components so as to make the overall system as robust as possible to external disruptions.

This chapter considers a bilevel optimization model to deal with security resource allocation in railway systems. We model the rail system as a network of nodes and links, where the nodes represent the stations and the links are the track segments. A limited budget is available for increasing the system security through the protection of nodes and or links. Different security measures can be employed, depending upon the asset to be protected. For example, a link containing a bridge or a tunnel can be protected through monitoring devices or structural reinforcement. A station can be protected by increasing surveillance and patrolling, or installing security cameras. Obviously, different costs are incurred for protecting different components (e.g., protecting a high-traffic commuter station requires significantly more protective resources than protecting a small station or a secondary rail track). Costs also depend on the type of security measure adopted. We assume that a protected component becomes completely invulnerable to possible disruptions. Likewise, if a failure occurs, the affected component becomes completely inoperable and unable to provide service. These assumptions can be considered strong but are common in the literature. In Chapter 7 we will relax both these binary assumptions. The aim of the model is to identify a cost-efficient allocation of the available budget so as to minimize the impact of worst-case scenario disruptions to the system. We focus, in particular, on passenger traffic and measure the disruption impact in terms of lost customer flow or demand. More specifically, we assume that if a node or a link fails, traffic must be rerouted through alternative paths on the network. However, detour routes may not exist or be too long from a user point of view. In this case, passengers may resort to different transport modes or abandon the trip altogether. The amount of customer flow which is lost provides an indication of the disruption extent. To evaluate the worst-case amount of disrupted flow, we use an adaptation of the flow interdiction model proposed by Murray et al. (2007). A common assumption in interdiction modeling is that there is a limit to the number of components that can be lost simultaneously. Without loss of generality, we also assume that interdiction resources are limited and that the amount of resources needed to disable a component varies according to the component size and topology.

4.1 The Railway Protection Investment Model

To formulate the railway protection investment problem mathematically, we consider a railway network as composed of a set of nodes N (the stations) and a set of arc A (the track segments). We assume that the daily traffic flow between any two stations s and t is known and that, in case of disruption, passengers are willing to use alternative railroad routes only if they are not significantly longer than their normal journey time. These information can be obtained by elaborating data on historical usage and through passengers' surveys. We call these routes *acceptable paths* and we compute them in a pre-processing phase. This evaluation is done by comparing each alternative path between an origin and a destination node with the shortest path: all the paths whose length exceeds a given threshold are discarded. The threshold is computed by adding a tolerance parameter to the length of the shortest path.

The other model assumptions can be summarized as follows:

- An interdicted element is excluded from the network.
- Both arcs and nodes can be interdicted. This assumption is made to simulate the disruptions of tunnels, bridges and stations at the same time.
- All the arcs directly linked to an interdicted node are interdicted as well.
- A protected element cannot be interdicted.
- A limited amount of interdiction resources is available.

The mathematical model uses the following notation.

Sets and Indices

N = set of nodes

A = set of arcs

$s \in N$ = index used for flow sources

$t \in N$ = index used for flow destinations

$i \in N$ = index used for network nodes

$j \in A$ = index used for network arcs

f_{st} = traffic demand between s and t

N_{st} = set of acceptable paths that connect s and t

$\beta \in N_{st}$ = index used for network paths

$N(\beta)$ = set of nodes along path β

$A(\beta)$ = set of arcs along path β

B = protection budget (or amount of resources available to the defender)

P = amount of resources available to the attacker

q_i^n = estimate of the amount of resources needed to protect node i

p_i^n = estimate of the amount of resources needed to disrupt node i

q_j^a = estimate of the amount of resources needed to protect arc j

p_j^a = estimate of the amount of resources needed to disrupt arc j.

Decision variables:

$$X_i^n = \begin{cases} 1 & \text{if node } i \text{ is disabled} \\ 0 & \text{otherwise;} \end{cases}$$

$$X_j^a = \begin{cases} 1 & \text{if arc } j \text{ is disabled} \\ 0 & \text{otherwise;} \end{cases}$$

$$Y_i^n = \begin{cases} 1 & \text{if node } i \text{ is protected} \\ 0 & \text{otherwise;} \end{cases}$$

$$Y_j^a = \begin{cases} 1 & \text{if arc } j \text{ is protected} \\ 0 & \text{otherwise;} \end{cases}$$

$$Z_{st} = \begin{cases} 1 & \text{if the flow between } s \text{ and } t \text{ is lost} \\ 0 & \text{otherwise;} \end{cases}$$

The railway protection investment model can be formulated as the following bilevel problem:

$$\min_{\mathbf{Y}} F(\mathbf{Y}) \tag{16}$$

$$\text{s.t.} \quad \sum_i q_i^n Y_i^n + \sum_j q_j^a Y_j^a \leq B, \tag{17}$$

$$Y_i^n \in \{0,1\} \qquad \forall i \in N, \tag{18}$$

$$Y_j^a \in \{0,1\} \qquad \forall j \in A, \tag{19}$$

$$\text{where} \quad F(\mathbf{Y}) = \max_{\mathbf{X}} \sum_s \sum_t f_{st} Z_{st}, \tag{20}$$

$$\text{s. t.} \quad \sum_i p_i^n X_i^n + \sum_j p_j^a X_j^a \leq P, \tag{21}$$

$$X_i^n \leq 1 - Y_i^n \qquad \forall i \in N, \tag{22}$$

$$X_j^a \leq 1 - Y_j^a \qquad \forall j \in A, \tag{23}$$

$$\sum_{i \in N(\beta)} X_i^n + \sum_{j \in A(\beta)} X_j^a \geq Z_{st} \qquad \forall s, t, \beta \in N_{st}, \tag{24}$$

$$X_i^n \in \{0,1\} \qquad \forall i \in N, \tag{25}$$

$$X_j^a \in \{0,1\} \qquad \forall j \in A, \tag{26}$$

$$Z_{st} \in \{0,1\} \qquad \forall s,t \in N. \tag{27}$$

In this leader-follower model the leader chooses the optimal strategy to minimize the objective function F (16), that is the amount of flow that cannot be served after the interdiction. Constraint (17) is the budget constraint: the leader can allocate at most q protection resources among the nodes and arcs of the network. Constraints (18) and (19) are the binary restrictions on the protection variables. The lower level program (20) $-$ (27) is the interdiction model used to evaluate worst-case losses. The aim of the follower is to choose the attack strategy that maximizes the amount of flow disrupted (20). Constraint (21) is the follower resource constraint: the attacker has at most p resources to interdict the nodes and arcs of the network. Constraints (22) state that protected nodes cannot be disrupted. Similarly, constraints (23) state that protected arcs cannot be disrupted. Constraints (24) state that the flow between s and t can be considered disrupted ($Z_{st} = 1$) only if all the acceptable paths between s and t are disrupted, i.e., at least one of their nodes or arcs is interdicted. If there is at least one acceptable path without interdicted components, the value of the variable Z_{st} is forced to be zero. Finally, constraints (25) $-$ (27) are binary restrictions on the interdiction and path variables.

4.2 Solution methodology

Multi-level models are generally very difficult to solve. Hansen et al. (1992) proved that even the simplest bilevel models, the ones with continuous variables on every level, are strongly NP-hard. Several solution approaches have been studied in the literature, including both heuristic techniques and exact methods. Examples of heuristic approaches can be found in Aksen and Aras (2013), Aksen et al. (2013, 2014), Parvaresh et al. (2013). Exact methods can be broadly classified into reformulation, enumeration and decomposition methods (Saharidis and Ierapetritou, 2009). Reformulation and enumeration techniques are usually only applicable to bilevel problems with linear lower level programs. A few exceptions to this are the reformulation of the p-median interdiction problem with fortification (Scaparra and Church, 2008b) and the implicit enumeration algorithm used to solve several protection-interdiction problems (Cappanera and Scaparra, 2011, Liberatore et al., 2012). Solution methods based on dualization cannot be applied to this problem because of the non continuous nature of the protection and interdiction variables. In general, the most effective methods for tackling problems with discrete variables in both levels are decomposition methods. These directly exploit the decomposable structure of the model and solve a series of smaller subproblems to find an overall optimal solution.

We used a decomposition method based on super valid inequalities. As described in Chapter 3, the bilevel model is split into two interlinked subproblems: an upper level protection problem (MASTER), and a lower level interdiction subproblem (INNER). Each protection strategy identified by the master problem is fed into the subproblem to determine an optimal interdiction plan. Special cuts, called supervalid inequalities (SVI), are then generated based on the solution to the interdiction problem and added to the master problem, which then computes a new protection strategy. The process is iterated until a sufficient number of SVIs has been added to make the protection problem unfeasible. The SVI is built on the idea that, for a given interdiction strategy $\hat{X}$, an effective protection must include at least one disrupted element.

Formally:

$$\text{SVI}(\hat{\mathbf{X}}) \quad \sum_i \hat{X}_i^n Y_i^n + \sum_j \hat{X}_j^a Y_j^a \geq 1 \tag{28}$$

$$\Big[\text{INNER}(\hat{\mathbf{Y}})\Big] \quad Z = \max_{\mathbf{X}} \sum_s \sum_t f_{st} Z_{st} \tag{29}$$

$$\text{s.t.} \quad X_i^n \leq 1 - \hat{Y}_i^n \qquad \forall i \in N, \tag{30}$$

$$X_j^a \leq 1 - \hat{Y}_j^a \qquad \forall j \in A, \tag{31}$$

$$(21), (24) - (27) \tag{32}$$

For each generic iteration of the algorithm, MASTER problem is a feasibility seeking problem subject to constraints (17) – (19) and to the set of SVIs added up to that iteration.

Figure 9 shows the pseudo-code of the implemented algorithm.

The decomposition algorithm was implemented in C++ inside the Visual Studio environment. At each iteration, both the master problem and the sub-problem were solved using the IBM ILOG optimization software Cplex 12.5.

Algorithm 6 SVIs decomposition algorithm

```
$Z_{opt} = -\infty; \hat{\mathbf{Y}} \leftarrow \mathbf{0}; \mathbf{Y}^{\mathbf{best}} \leftarrow \mathbf{0}$
while MASTER is feasible do
    Solve INNER($\hat{\mathbf{Y}}$) to find $Z$ and $\hat{\mathbf{X}}$;
    add SVI($\hat{\mathbf{X}}$) to MASTER problem
    if $Z_{opt} > Z$ then
        $Z_{opt} = Z$ and $\mathbf{Y}^{\mathbf{best}} \leftarrow \hat{\mathbf{Y}}$;
    end if
end while
END    return ($\mathbf{Y}^{\mathbf{best}}$, $Z_{opt}$);
```

Figure 9: SVIs decomposition algorithm.

4.3 Case study and Analysis

To demonstrate the practical applicability of our approach, we applied the model to the railway network infrastructure of Campania, a region in Southern Italy. The region Campania is populated by almost 6 million people, making it the second-most-populous region of Italy. Its capital city is Naples. The railway network under consideration is composed by a primary network which connects major cities in Italy and has high traffic (high speed and inter-regional rail services), a secondary network which connects an highly populated urban centre to outer suburbs (Cumana, Circumflegrea, Circumvesuviana and north-east metro services), and some complementary lines which connect small regional centres. The overall network is depicted in Fig. 10. The network has 26 nodes, corresponding to cities and towns in the region, and 37 arcs.

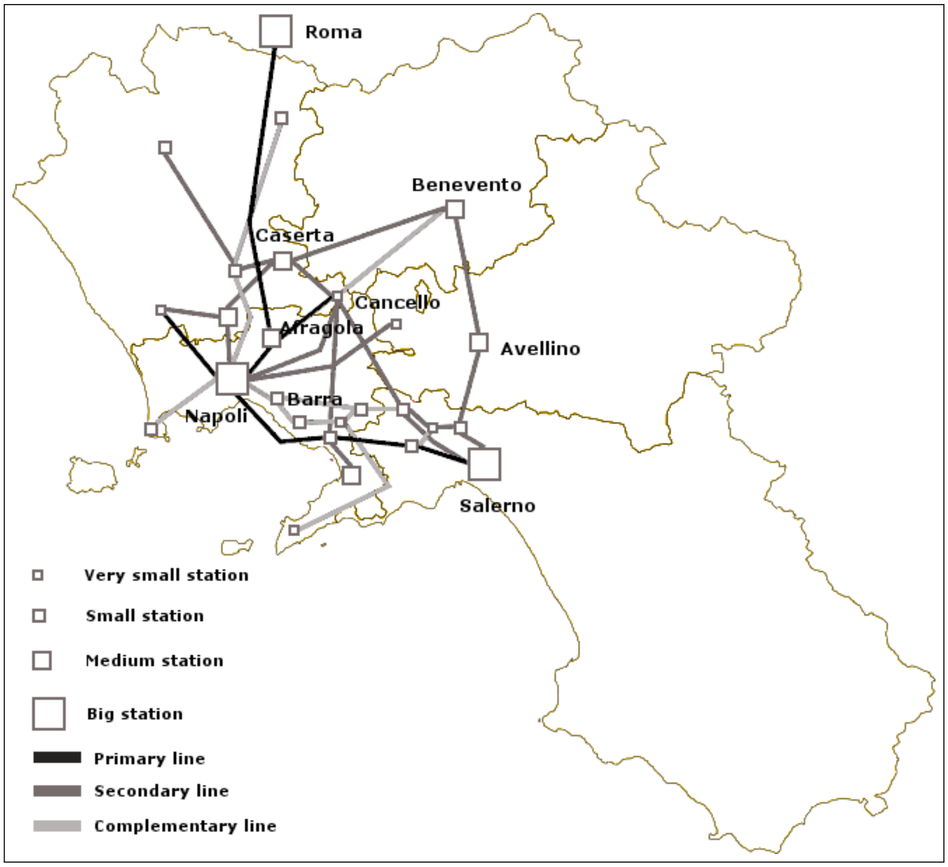


Figure 10: Campania rail network.

In the absence of real data on passenger traffic between pairs of stations, we have generated estimates of the origin-destination flows as a function of the size and distance of the connected cities. We assumed that disrupting an arc requires one unit of resource ($p_j^a = 1$), whereas the cost of protecting an arc, q_j^a, depends upon the number of tunnels and bridges along the arc. We do not consider the protection of arcs without tunnels or bridges. To generate realistic values for the interdiction and protection resources associated with the nodes (q_i^n and p_i^n), we have divided the stations in four groups according to their dimension. The values chosen for the stations in each group are shown in Table 1. Obviously, bigger stations require more resources to be protected/disrupted. As an example, Caianello is a very small station and only requires 2 units, whereas Naples is the biggest station and requires 12 units.

Table 1: Resources needed to protect/interdict a node

Node dimension	Interdiction/Protection resources
Very small	2
Small	4
Medium	8
Big	12

In our empirical study, we have analysed and compared protection strategies to hedge against disruptions of different magnitudes. Specifically, we considered small, medium, large and very large disruptions. The amount of interdiction resources associated with each event size are displayed in Table 2. With this choice, a small disruption can only affect a very small station, whereas a very large event is able to interdict a big station and a few other smaller assets.

Table 2: Disruption scenarios

Size	Resource units
Small	2
Medium	5
Large	10
Very large	20

The analysis also considers different budget levels. These were chosen as a percentage of the budget needed to protect the whole network.

Some preliminary results are displayed in Table 3, which shows the total amount of flow which is lost in different disruption scenarios and for different protection investment levels. It can be seen that even a small disruption can have a considerable impact on traffic flow if protective measures are not carried out: the worst-case loss after a small disruption can result in a loss of 38% of the total flow. This can reach 67%, 88%, and 98% for medium, large, and very large disruptions respectively. The effect of protecting even as little as 1% of the assets can be considerable, if protection resources are allocated optimally. This is true especially for small and medium size disruption scenarios, where the total losses can be reduced from 38% to 18% for small events and from 67% to 39% for medium events. We run some tests to assess the benefits of allocating resources optimally. As an example, for small scenarios, with a 5% budget, a random allocation resulted in a 17% (average) flow loss increment, compared with the optimal allocation. For large and very large events, greater protection investments are needed to get significant reductions in flow losses. As an example, an optimal investment equal to 5% of the protection cost of the total network, can more than halve the flow loss resulting from a large disruption (from 88% to 35%).

Table 3: Percentage of lost flow for different disruption scenarios and protection budget levels.

	No protection	1%	5%	10%
small	38%	18%	10%	5%
medium	67%	39%	20%	16%
large	88%	75%	35%	28%
very large	98%	96%	77%	54%

To provide a better understanding of how increasing budget levels may affect the system losses in case of disruption, in 11 we show the percentage marginal reduction in flow losses for each percentage point increase in protection resources. We let the budget vary between 1% and 10% of the protection cost of the whole network.

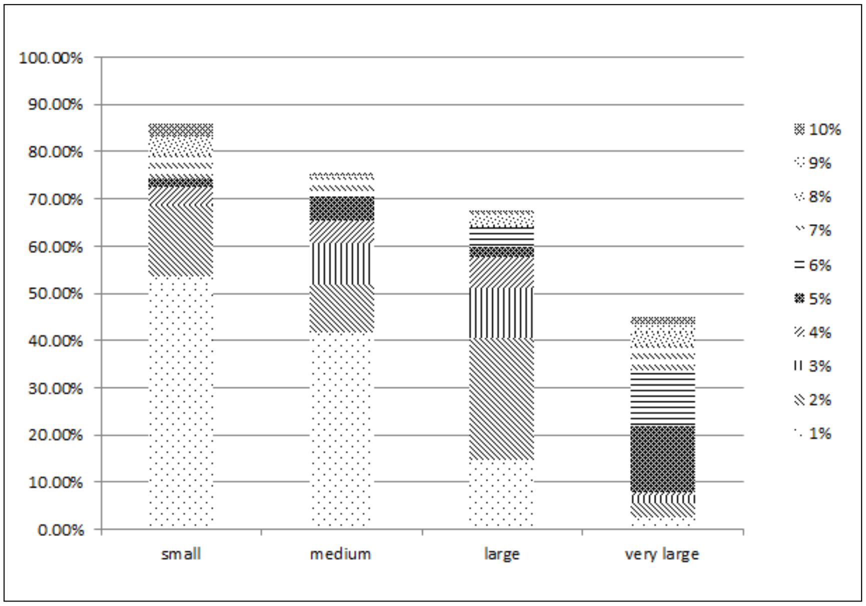


Figure 11: Marginal percentage decrease in flow loss due to percentage point increments of the protection budget.

This analysis sheds light on possible tradeoffs between protection expenditures and flow loss reductions in case of worst-case system disruptions. As an example, if a large disruptive event is considered, a 1% investment results in a worst-case loss reduction of about 15% (first segment of the third bar in the chart). However, if an investment of 2% can be made, the benefit is more than doubled, bringing an additional 25% flow loss reduction and an overall reduction of 40%.

The differences between the four disruption scenarios can be further analysed through the graphs plotted in Fig. 12, Fig. 13, Fig. 14, and Fig.15. For each scenario, the corresponding graph displays the contribution of a percentage point increment in protection resources on the overall objective improvement.

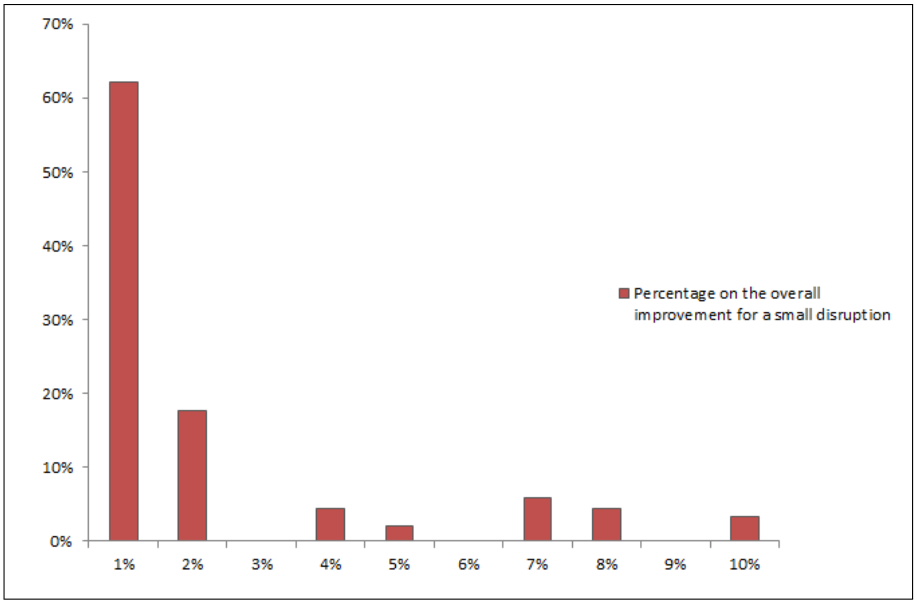


Figure 12: Analysis of the contribution of percentage point increases of the protection resources on the overall improvement for small disruptions.

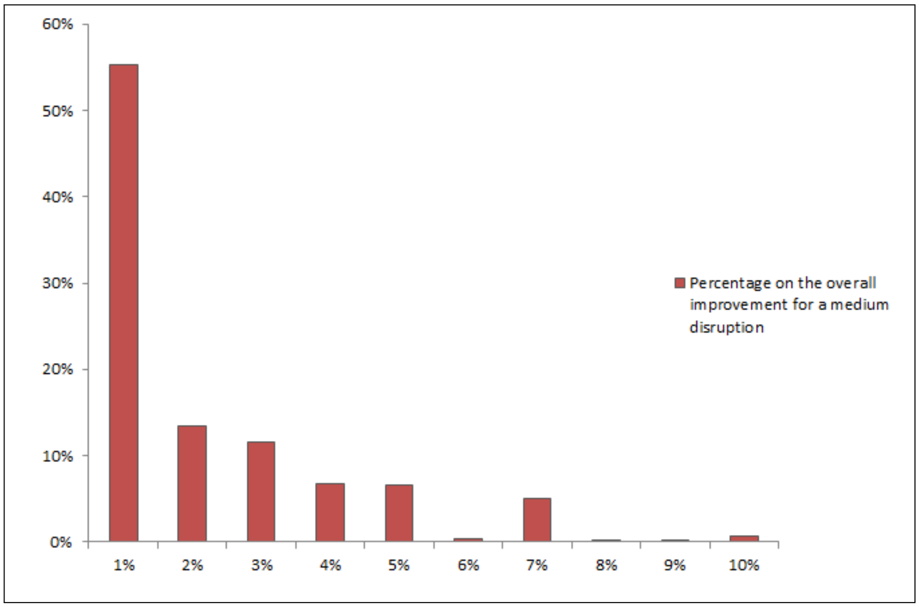


Figure 13: Analysis of the contribution of percentage point increases of the protection resources on the overall improvement for medium disruptions.

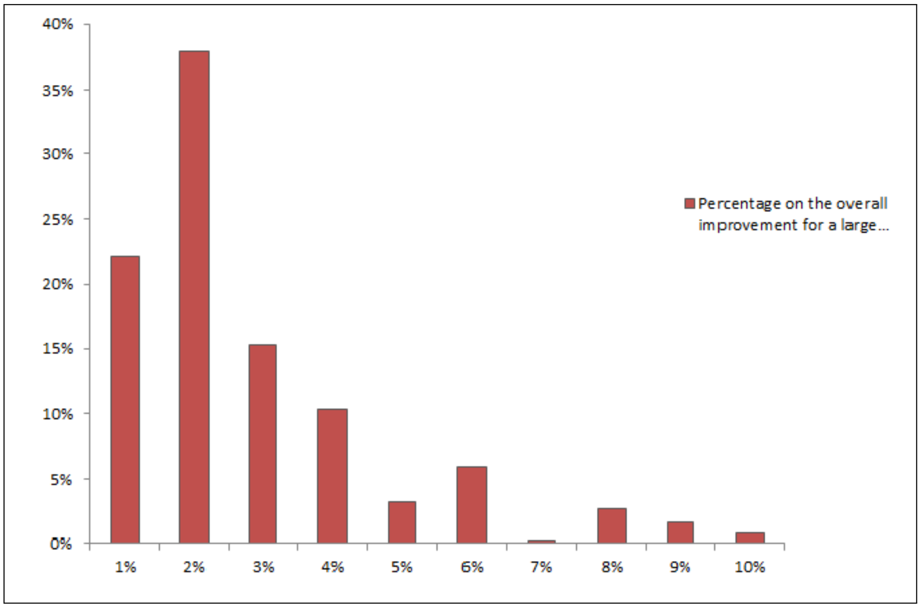


Figure 14: Analysis of the contribution of percentage point increases of the protection resources on the overall improvement for large disruptions.

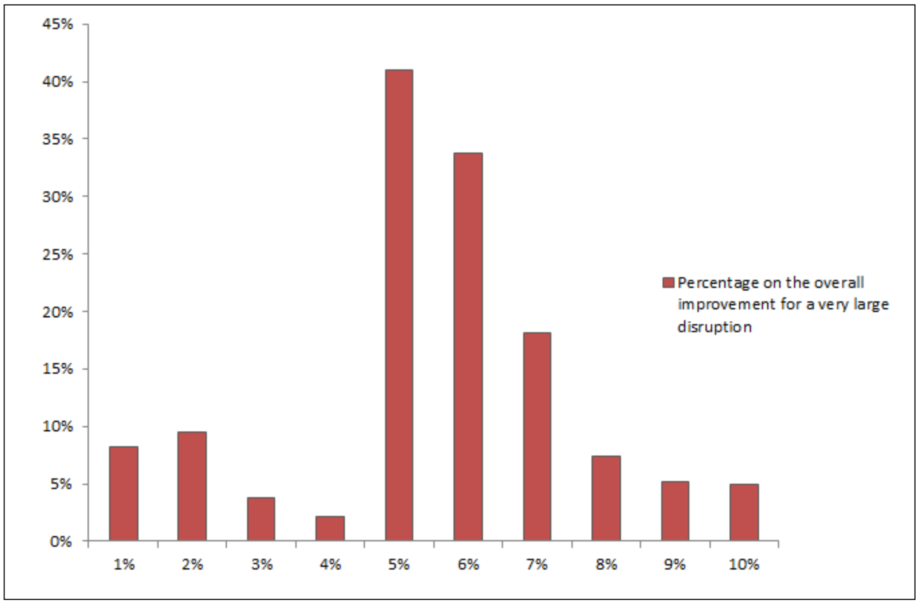


Figure 15: Analysis of the contribution of percentage point increases of the protection resources on the overall improvement for very large disruptions.

The first clear difference is that in the scenarios with low and medium level disruptions (Fig. 12 and Fig. 13) the first percentage point increase is responsible for more than half of the overall benefit. To reach similar results for large disruptions, a two point increment is needed (Fig. 14). When very large disruptions are considered, the first few increments have a somewhat limited effect on reducing flow losses whereas a peak can be noticed in correspondence of a 5% investment (Fig. 15). An additional percentage point increase, results in another significant flow loss reduction. This seems to indicate that if large disruptions are anticipated, a protection budget in this range (5%-6% of the total protection costs) should be warranted to maximize the benefits of security investments.

It is clear that the protection strategies identified by the model may differ quite significantly, depending on the magnitude of the disruption given in input to the model (parameter p). Our next analysis aims at identifying protection plans which are robust across all scenarios, so as to hedge against the uncertainty characterizing the size and extent of disruptive events. To this end, we evaluate how the optimal solution identified for a given disruption size performs in all the other scenarios.

The results of this analysis are shown in Table 4 and Table 5 for two budget levels, equal to 5% and 10% of the resources needed to protect the whole network. These cases correspond to values of q equal to 17 and 35 respectively. The tables show the percentage flow loss increase which is observed when the optimal protection strategy computed for a given scenario (*supposed scenario*) is used in a different scenario (*actual scenario*). The last two columns display the maximum and average increase across all the other scenarios. From the analysis of Table 4 ($q = 17$), it is clear that the optimal solution for medium and large events is the same. It is also the solution that works better across the different scenarios, with an average error of 11.1% and a maximum error of 22.5%. In the second case (Table 5), all the solutions are different and the best choice, in terms of average percentage increase of disrupted flow, is the optimal protection strategy computed for very large disruptions. Nevertheless, assuming a medium size disruption may result in a better compromise solution: the average percentage increase is really close to the one obtained for very large events (41.5% vs. 41.4%) but the maximum value is considerably smaller (85.7% vs. 126.1%). Overall this analysis indicates that the assumptions made on the disruption size may have a significant impact on the identification of effective protection strategies. In general, avoiding the extreme cases and assuming medium to large disruptions leads to the most robust defensive plans.

Table 4: Cross-comparison of different optimal protection plans. Relative flow loss increase in percentage. Case: $q = 17$.

Supposed	Actual scenario				MAX	AVG
scenario	small	medium	large	very large		
small	0%	98.8%	75.2%	20.3%	98.8%	48.6%
medium	21.8%	0%	0%	22.5%	22.5%	11.1%
large	21.8%	0%	0%	22.5%	22.5%	11.1%
very large	133.8%	104%	78.1%	0%	133.8%	79%

Table 5: Cross-comparison of different optimal protection plans. Relative flow loss increase in percentage. Case: $q = 35$.

Supposed	Actual scenario				MAX	AVG
scenario	small	medium	large	very large		
small	0%	138.4%	106.3%	74.8%	138.4%	79.9%
medium	85.7%	0%	8.6%	71.6%	85.7%	41.5%
large	98.1%	3.2%	0%	74%	98.1%	43.8%
very large	126.1%	19.9%	19.7%	0%	126.1%	41.4%

Finally, in Table 6 and Table 7 we display the solutions to the model for different disruption scenarios and protection budget levels. Table 6 shows the network components chosen for protection, whereas Table 7 shows the interdiction plans (i.e., the worst-case losses) after protection.

We can see that Afragola and Barra appear quite often in the protection and disruption strategies. This can be explained by noticing that the first station is a crucial node of the high speed service and its disruption affects the connection between Rome and Naples; the second station belongs to the Circumvesuviana railway network and intercepts a huge portion of the traffic generated by that service. It is interesting to note that Cancello appears very frequently among the components to be interdicted, in spite of being a very small station. This may be due to its very central position. Cancello, in fact, intercepts the flow between the largest cities of the region and this makes it an attractive target for an intelligent attacker. Finally, it can be noted that Naples only appears in a few solutions probably because, although it is the most important station, is also the most difficult and expensive asset to protect and/or disrupt.

Table 6: Optimal protection plans for different disruption scenarios and different protection budgets

Disruption	Protection resources			
size	$q = 1\%$	$q = 2\%$	$q = 5\%$	$q = 10\%$
small	Naples-Barra Naples-Afragola	Naples-Barra Naples-Afragola T.Annunziata-C.Stabia	Naples-T.Annunziata Torregaveta-Naples Naples-Afragola Naples-Barra S.Maria C.V.-Caserta T.Annunziata-C.Stabia T.Annunziata-Nocera	Torregaveta-Naples Rome-Afragola Naples-Afragola Naples-Barra Naples-T.Annunziata T.Annunziata-C.Stabia T.Annunziata-Nocera
medium	Naples-Barra T.Annunziata-C.Stabia	Barra Naples-Barra T.Annunziata-C.Stabia	Barra T.Annunziata Naples-Aversa Naples-Afragola Naples-Barra Naples-T.Annunziata T.Annunziata-C.Stabia T.Annunziata-Nocera	Barra T.Annunziata S. Giorgio a C. Torregaveta-Naples Naples-Aversa Naples-Afragola Naples-Barra Naples-T.Annunziata S.Maria C.V.-Caserta S. Giorgio a C.-T.Annunziata O. T.Annunziata-C.Stabia T.Annunziata-Nocera
large	Naples-Barra T.Annunziata-Nocera	Barra Naples-Afragola Naples-Barra	Barra T.Annunziata Naples-Aversa Naples-Afragola Naples-Barra Naples-T.Annunziata T.Annunziata-C.Stabia T.Annunziata-Nocera	Barra T.Annunziata Cancello Torregaveta-Naples Naples-Aversa Naples-Afragola Naples-Barra Naples-T.Annunziata Barra-P.Marino T.Annunziata-C.Stabia T.Annunziata-Nocera T.Annunziata O.-P.Marino
very large	Naples-Afragola Sarno-Codola	Naples-Aversa S.Maria C.V.-Caserta T.Annunziata-C.Stabia T.Annunziata-Nocera	Naples Barra Naples-Barra	Naples Barra T.Annunziata Nocera Naples-Aversa Naples-Afragola Naples-Barra Naples-T.Annunziata T.Annunziata-C.Stabia T.Annunziata-Nocera Nocera-Codola Sarno-Codola

Table 7: Post-protection worst-case losses in different disruption scenarios and for different protection budgets

Disruption	Protection resources			
size	$q = 1\%$	$q = 2\%$	$q = 5\%$	$q = 10\%$
small	Rome-Afragola T.Annunziata-C.Stabia	Rome-Afragola Naples-T.Annunziata	Rome-Afragola Sorrento-T.Annunziata O.	Naples-S.Maria C.V. S.Maria C.V.-Caserta
medium	Barra P.Marino-Sarno	V.Literno-Naples Rome-Afragola Naples-Aversa Aversa-Caserta T.Annunziata-Nocera	Torregaveta-Naples Rome-Afragola Barra-P.Marino S.Giorgio a C.-T.Annunziata O. P.Marino-Sarno	Nocera Rome-Afragola
large	Barra Torregaveta-Naples Naples-Afragola S.Maria C.V.-Caserta Afragola-Cancello T.Annunziata-C.Stabia P.Marino-Sarno	Cancello V.Literno-Naples Torregaveta-Naples Rome-Afragola Naples-Aversa Naples-T.Annunziata S.Maria C.V.-Caserta T.Annunziata-Nocera P.Marino-Sarno	Cancello Torregaveta-Naples Rome-Afragola S.Maria C.V.-Caserta Aversa-Caserta Barra-P.Marino S.Giorgio a C.-T.Annunziata O. Nocera-Codola Nocera-Salerno	S.Giorgio a C. Nocera Rome-Afragola S.Maria C.V.-Caserta
very large	Naples Cancello S.Maria C.V.-Caserta Aversa-Caserta T.Annunziata-C.Stabia Nocera-Codola Nocera-Salerno Mercato-Salerno	Naples Cancello Nocera Aversa-Caserta Mercato-Salerno	Cancello S.Giorgio a C. V.Literno-Naples Torregaveta-Naples Naples-S.Maria C.V. Naples-Aversa Naples-Afragola Naples-T.Annunziata S.Maria C.V.-Caserta Aversa-Caserta Barra-P.Marino T.Annunziata-C.Stabia P.Marino-Sarno Nocera-Codola Nocera-Salerno Mercato-Salerno	Cancello S.Giorgio a C. Afragola Torregaveta-Naples S.Maria C.V.-Caserta Aversa-Caserta Barra-P.Marino P.Marino-Sarno Nocera-Salerno

5 Optimizing dynamic investment decisions for railway systems protection

In this chapter we further extend the previous model by adding a temporal component. An important issue that should be taken into account when modelling protection efforts is that protection resources can become available at different times. Our model addresses this issue by including a temporal component whereby the available budget for protection is spread over a planning horizon. This choice renders the model more applicable to real situations. In fact, public expenditures to protect and modernize critical infrastructures are usually set in spending reviews that cover a number of years. For instance, the last UK spending review (HM Treasury, 2013) allocated £100bn for the modernisation of the energy and transportation sectors. This budget is spread over a five-year period (2015-2020). Furthermore, after 2013-2014 floods in UK, £130m were allocated by the government. Of the whole budget, £30m were made available in 2014, the rest in 2015 (Carrington and Weaver, 2014).

In addition,the UK Department for Environment, Food and Rural Affairs (DEFRA) set out a six-year programme of capital investment to improve flood defences up to 2021, of 2.3bn. Fixed capital settlements were allocated for each year, although flexibility to move funds between years was allowed for effective delivery (DEFRA, 2015). These examples demonstrate that funds availability is often time-related. Consequently, prioritizing expenditures over time is key to the development of long-term, effective strategies for improving infrastructures security and resiliency. To respond to the practical planning needs of railway stakeholders and operators, we therefore propose a protection model that optimizes the allocation of scarce protection resources over time.

The allocation of defensive and offensive resources over time has recently been analysed by a few researchers within a game theoretic framework. The majority of these models focused on the protection and disruption of a single target. For example, Levitin and Hausken (2010) proposed a defender-attacker model where the attacker can launch sequential attacks. Hausken and Zhuang (2011a) considered a government-terrorist game over multiple time periods, where the terrorist can stockpile its resources for later attacks and the government can allocate resources for defending the asset or attack the terrorist's resources. Other single-asset sequential defender-attacker problems can be found in Hausken and Zhuang (2011b), Hausken and Zhuang (2012), and Levitin and Hausken (2012a). A multiple-target version of these problems has been considered in Levitin and Hausken (2009), who studied the problem of protecting identical elements in a parallel system against two sequential attacks. Levitin and Hausken (2012b) extended this model by including the possibility of imperfect detection of the first attack outcomes. In Levitin and Hausken (2013) both the attacker and the defender can stockpile their resources over a planning horizon. Note that the game-theoretic approach used in this literature stream is quite different from our approach in that the problems are represented as a two-stage game and require a closed-form analytic solution for the identification of Nash equilibria. Therefore, the application of these models is limited to simple problems (i.e. two time periods only, single target, small system). On the contrary, our approach is able to solve problem with several time periods and networks of realistic size.

Our dynamic network protection problem (DNP) is formulated as a bilevel model where the aim of the upper level is to find the best allocation of protection resources, over a planning horizon, to minimize the amount of disrupted flow. The lower level is used to evaluate worst-case losses in response to a given protection plan. The resulting model is complicated and requires ad-hoc expedients tailored to its dynamic structure. We propose two decomposition approaches that are tested using randomly generated networks. In both the approaches, the dynamic structure of the problem can lead to solve the same lower problem more than once. We use hash structures to avoid recomputation, obtaining significant improvement in the computing time. Finally insights of the problem are provided using a network representing the railway infrastructure of Kent (UK).

5.1 The Dynamic Network Protection Problem (DNP)

We consider a directed graph $G = (N, A)$ representing the transportation network. In a railway network, the nodes represent the stations and the arcs are the tracks connecting the nodes. Assumptions, parameters and decision variables are introduced below:

(a) The problem is studied over a planning horizon represented by the set $T = \{0, 1, ..., \hat{T}\}$.

(b) Interdiction is complete (i.e., an interdicted component is completely unusable in the time period when interdiction takes place).

(c) A protected element becomes completely immune to interdiction. Therefore the same element does not need to be protected more than once in the planning horizon. Both arcs and nodes can be disrupted and protected.

(d) Each element has a different protection cost and there is a limited protection budget in each time period. Any unutilized budget can be carried forward to the next time period.

(e) In each time period, interdiction resources are limited and the amount of resources needed to disrupt a component varies according to the component size and topology. Interdiction resurces cannot be carried forward.

(f) In case of disruption, system users are willing to use alternative paths to reach their destinations only if they are not significantly longer than their shortest route. We refer to these alternative routes as *acceptable* paths. All the paths that establish connectivity between two nodes s and d are computed in a preprocessing phase. The paths that are too long from a user perspective are then removed from further considerations. This evaluation is done by comparing each path with the shortest one: all paths exceeding a given length threshold are discarded.

(g) The daily traffic flow between any two nodes is known with certainty and the flow matrix is symmetric.

The bilevel model for DNP uses the following notation.

Indices, sets and parameters

$s \in N$: index used for flow sources.

$d \in N$: index used for flow destinations.

$i \in N$: index used for network nodes.

$j \in A$: index used for network arcs.

$t, u \in T$: index used for time periods.

f_{sd} : traffic demand between s and d.

N_{sd} : set of *acceptable* paths that connect s and d.

$r \in N_{sd}$: index used for network paths.

$N(r)$: set of nodes along path r.

$A(r)$: set of arcs along path r.

B_t : cumulative protection budget available up to period t.

P_t : amount of interdiction resources in period t.

q_i^n : estimate of the amount of resources needed to protect node i.

p_i^n : estimate of the amount of resources needed to disrupt node i.

q_j^a : estimate of the amount of resources needed to protect arc j.

p_j^a : estimate of the amount of resources needed to disrupt arc j.

λ_t : weight used in the objective function to give different importance to the time periods. The parameter is normalized (i.e., $\sum_t \lambda_t = 1$).

Decision variables

$X_{it}^n = 1$ if node i is disabled in period t; 0 otherwise.

$X_{jt}^a = 1$ if arc j is disabled in period t; 0 otherwise.

$Y_{it}^n = 1$ if node i is protected in period t; 0 otherwise.

$Y_{jt}^a = 1$ if arc j is protected in period t; 0 otherwise.

$Z_{sdt} = 1$ if the flow between s and d is unserved in period t; 0 otherwise.

The DNP can be formulated as follows.

$$[\text{DNP}] \quad \min_{\mathbf{Y}} F(\mathbf{Y}) \tag{33}$$

$$\text{s.t.} \quad \sum_{u=1}^{t} (\sum_{i \in N} q_i^n Y_{iu}^n + \sum_{j \in A} q_j^a Y_{ju}^a) \leq B_t \qquad \forall t \in T \tag{34}$$

$$Y_{it}^n \in \{0,1\} \qquad \forall i \in N, \forall t \in T \tag{35}$$

$$Y_{jt}^{a} \in \{0,1\} \qquad \forall j \in A, \forall t \in T \tag{36}$$

$$\text{where} \quad F\left(\mathbf{Y}\right) = \max_{\mathbf{X}} \sum_{t \in T} \lambda_t \sum_{s} \sum_{d} f_{sd} Z_{sdt} \tag{37}$$

$$\text{s.t.} \quad X_{it}^{n} \leq 1 - \sum_{u=1}^{t} Y_{iu}^{n} \qquad \forall i \in N, \forall t \in T \tag{38}$$

$$X_{jt}^{a} \leq 1 - \sum_{u=1}^{t} Y_{ju}^{a} \qquad \forall j \in A, \forall t \in T \tag{39}$$

$$\sum_{i \in N} p_i^n X_{it}^n + \sum_{j \in A} p_j^a X_{jt}^a \leq P_t \qquad \forall t \in T \tag{40}$$

$$\sum_{i \in N(r)} X_{it}^{n} + \sum_{j \in A(r)} X_{jt}^{a} \geq Z_{sdt} \qquad \forall s, d \in N, r \in N_{sd}, \forall t \in T \tag{41}$$

$$X_{it}^{n} \in \{0,1\} \qquad \forall i \in N, \forall t \in T \tag{42}$$

$$X_{jt}^{a} \in \{0,1\} \qquad \forall j \in A, \forall t \in T \tag{43}$$

$$Z_{sdt} \in \{0,1\} \qquad \forall s, d \in N, \forall t \in T. \tag{44}$$

In the bilevel model above, the leader seeks the optimal protection strategy to minimize the function F (33), which represents the weighted sum of demand that cannot be served after interdiction, over the planning horizon. Constraint (34) represents the budget limit: the amount of resources utilized up to period t for nodes and arcs protection cannot exceed the available cumulative budget B_t. Constraints (35) and (36) are the binary requirements for the protection variables. The lower level program (37)-(44) is the interdiction model. The follower seeks the attack strategy that maximizes the overall amount of unserved demand (37). Constraints (38) state that a node cannot be disrupted at period t, if it is protected in the time window $\{1, ..., t\}$. Similarly, constraints (39) state that an arc cannot be disrupted at period t, if it is protected in the time window $\{1, ..., t\}$. Constraints (40) set a limit on the interdiction resources available in each time period. Constraints (41) state that the demand between s and d is unserved in period t ($Z_{sdt} = 1$), only if all the *acceptable* paths connecting the two nodes are disrupted at period t. This occurs if at least one node or arc on each path is disabled. Finally, constraints (42)-(44) enforce binary restrictions on the lower level variables.

5.2 Solution methodology

In this chapter, we present two different decomposition approaches for DNP. The first is based on the use of Benders cuts. Benders decomposition has been widely used in the literature to deal with large-scale MILP problems (Benders, 1962). More recently, the use of Benders-like decomposition algorithms has been extended to multi-level programs (Israeli and Wood, 2002, O'Hanley and Church, 2011, Losada et al., 2012). The second approach utilizes special cutting planes known as Super Valid Inequalities (SVIs). A SVI is a cutting plane that reduces the feasible region without excluding any optimal solution unless the incumbent solution is itself optimal. SVIs were initially introduced by Israeli and Wood (2002) to speed up a Benders decomposition approach. SVIs were also used explicitly as a stand alone solution method in O'Hanley and Church (2011) and in Losada et al. (2012).

In all our decomposition approaches, DNP is split into two connected subproblems referred to as the *Restricted Master Problem* (RMP) and the *SubProblem* (SP). These subproblems are solved alternatively until the algorithms converge to an optimal solution. The RMP entails decisions about what to protect to thwart the most disruptive interdiction plans identified in previous iterations. At each iteration, the most disruptive interdiction plan in response to a given protection strategy is identified by solving SP, which is the interdiction problem (37)-(44) with the protection variables fixed to the feasible values identified by the current RMP's solution. The solution to the SP is then used to generate either Benders or SVIs cuts to be appended to the RMP and the process is iterated.

The description of the decomposition methods uses the following additional notation.

w : iterations index.

$\hat{\mathbf{Y}}_{\mathbf{w}} = [\hat{\mathbf{Y}}^{\mathbf{n}}_{\mathbf{w}}, \hat{\mathbf{Y}}^{\mathbf{a}}_{\mathbf{w}}]$: RMP's solution at iteration w. This vector holds the values of the protection variables Y^n_{it} and Y^a_{jt}.

$\hat{\mathbf{Z}}_{\mathbf{w}}\hat{\mathbf{X}}_{\mathbf{w}}$: SP's optimal response plan given protection strategy $\hat{\mathbf{Y}}_{\mathbf{w}}$. This vector holds the values of the variables Z_{sdt}, X^n_{it}, and X^a_{jt}.

$\hat{\mathbf{Z}}_{\mathbf{w}}$: sub-vector of $\hat{\mathbf{Z}}_{\mathbf{w}}\hat{\mathbf{X}}_{\mathbf{w}}$ holding the variables Z_{sdt}.

$\hat{\mathbf{X}}_{\mathbf{w}} = [\hat{\mathbf{X}}^{\mathbf{n}}_{\mathbf{w}}, \hat{\mathbf{X}}^{\mathbf{a}}_{\mathbf{w}}]$: sub-vector of $\hat{\mathbf{Z}}_{\mathbf{w}}\hat{\mathbf{X}}_{\mathbf{w}}$ holding the variables X^n_{it} and X^a_{jt}.

Given a protection strategy $\hat{\mathbf{Y}}_{\mathbf{w}}$, the subproblem SP, which is the same for both the approaches, is simply:

$\left[\text{SP}(\hat{\mathbf{Y}}_{\mathbf{w}})\right]$

$$\max_{\mathbf{X}} \sum_{t \in T} \lambda_t \sum_s \sum_d f_{sd} Z_{sdt} \tag{45}$$

$$\text{s.t.} \quad X^n_{it} \leq 1 - \sum_{u=1}^{t} \hat{Y}^n_{iuw} \qquad \forall i \in N, \forall t \in T \tag{46}$$

$$X^a_{jt} \leq 1 - \sum_{u=1}^{t} \hat{Y}^a_{juw} \qquad \forall j \in A, \forall t \in T \tag{47}$$

$$(40) - (44)$$

By solving this model to optimality, we obtain a feasible solution, $[\hat{\mathbf{Y}}_{\mathbf{w}}, \hat{\mathbf{Z}}_{\mathbf{w}}\hat{\mathbf{X}}_{\mathbf{w}}]$, for DNP and an upper bound to its objective. Additionally, the optimal response strategy $\hat{\mathbf{X}}_{\mathbf{w}}$ can be used to generate cutting planes for the RMP, as described in the following sections.

5.3 Benders Decomposition (BND-D)

The Benders decomposition algorithm uses the following additional notation.

a_{rtw} : number of different elements along path r which are interdicted at time t in the interdiction plan identified at iteration w.

$\bar{Z}_w = \{(s, d, t) \in N \times N \times T \mid Z_{sdtw} = 1\}$: indices of the disrupted flows at iteration w.

Q_{rtw} : binary variable which takes value 1 if the interdiction of path r at time t in iteration w is thwarted; 0 otherwise.

Q_{sdtw} : binary variable which takes value 1 if the interdiction of the flow from s to d at time t in iteration w is thwarted; 0 otherwise.

The Q variables are introduced to *reconstruct* the link between a protection strategy and the correspondent thwarted interdictions. In BND-D, the RMP at iteration $\bar{w}$ is a mixed-integer program defined as follows:

$[RMP(\bar{w})]$

$$\min_{\mathbf{Y}} z \tag{48}$$

$$\text{s. t.} \quad (34) - (36)$$

$$z \geq \sum_{(s,d,t)\in\bar{Z}_w} \lambda_t \left(f_{sd}(1 - Q_{sdtw})\right) \qquad \forall w \in [1, \bar{w}] \tag{49}$$

$$\sum_{i \in N(r)} \hat{X}^n_{itw} \sum_{u=1}^{t} Y^n_{iu} + \sum_{j \in A(r)} \hat{X}^a_{jtw} \sum_{u=1}^{t} Y^a_{ju} \geq \; a_{rtw} \, Q_{rtw} \tag{50}$$

$$\forall s \in N, d \in N, t \in T : (s,d,t) \in \bar{Z}_w, \forall r \in N_{sd}, \forall w \in [1, \bar{w}]$$

$$\sum_{r \in N_{sd}} Q_{rtw} \geq \; Q_{sdtw} \qquad \forall s \in N, d \in N, t \in T : (s,d,t) \in \bar{Z}_w, \forall w \in [1, \bar{w}] \tag{51}$$

$$Q_{rtw} \in \{0,1\} \qquad \forall s \in N, d \in N, t \in T : (s,d,t) \in \bar{Z}_w, \forall r \in N_{sd}, \forall w \in [1, \bar{w}] \tag{52}$$

$$Q_{sdtw} \in \{0,1\} \qquad \forall (s,d,t) \in \; \bar{Z}_w, \forall w \in [1, \bar{w}] \tag{53}$$

$$z \in \mathbb{R}^+. \tag{54}$$

The aim of the objective function (48) is to find the best protection strategy that thwarts the interdiction plans identified in the previous iterations. Constraints (49) are called *Benders cuts*. They are lower bounds to the objective function z generated by all the interdiction strategies found in the previous iterations. Constraints (50) represent the relationship between the variables Q_{rtw} and the protection variables. Specifically, they state that a path r connecting s and d can no longer be disrupted at time t by the interdiction strategy $\hat{\mathbf{X}}_w$ (i.e., $Q_{rtw} = 1$), if all its interdicted arcs and nodes are protected either at time t or in some time period prior to t. Constraints (51) state that the interdiction of the flow between s and d at time t in iteration w can be thwarted (i.e., $Q_{sdtw} = 1$) only if the protection strategy thwarts the interdiction of at least one acceptable path r connecting s and d at time t. If at least one path is not disrupted, then the objective function pushes the variable Q_{sdtw} to take value 1 and the flow f_{sd} at time t is no longer considered unserved in (49). Finally, constraints (52) and (53) represent the binary requirements for the variables Q_{rtw} and Q_{sdtw} and constraint

(54) states that variable z is a non negative real.

The pseudo-code of BND-D is displayed below.

Algorithm 7 Bender decomposition

```
Set $w = 1$, $\hat{\mathbf{Y}}_w = \mathbf{0}$, $\mathbf{Y_{opt}} = \mathbf{0}$, $z_{sup} = \infty$ and $z_{inf} = -\infty$
MAINSTEP
Solve SP($\hat{\mathbf{Y}}_\mathbf{w}$) to obtain $\hat{\mathbf{Z}}_\mathbf{w}\hat{\mathbf{X}}_\mathbf{w}$ and the objective value $\hat{z}$
if $\hat{z} < z_{sup}$ then
  $z_{sup} = \hat{z}$ and $\mathbf{Y_{opt}} \leftarrow \hat{\mathbf{Y}}_\mathbf{w}$
end if
if $z_{sup} - z_{inf} = 0$ then
  goto TERMINATE
end if
$w = w + 1$
Solve RMP($w$) to obtain $\hat{\mathbf{Y}}_\mathbf{w}$ and $z_{inf}$
if $z_{sup} - z_{inf} > 0$ then
  goto MAINSTEP
end if
TERMINATE
Return($\mathbf{Y_{opt}}$)
```

The solution of the SP provides an upper bound to the DNP. Conversely, the solution of the RMP is a lower bound for the DNP (the RMP is in fact a relaxation of DNP as it only includes a subset of all possible interdiction plans). When the two sub-problems have the same objective function value, the algorithm stops. It is easy to prove that BND-D converges in a finite number of iterations. The resource constraints, in fact, guarantee that the number of interdiction and protection strategies is finite.

5.3.1 SVI Decomposition (SVI-D)

The basic idea behind this approach is that, to thwart a worst-case interdiction and hence lower the objective function value of the follower, the protection strategy must include at least one element belonging to the optimal interdiction set (Church and Scaparra, 2007). Our SVIs embed this idea by enforcing the protection of at least one of the arcs or one of the nodes interdicted in the current follower response $\hat{\mathbf{X}}_w$. More specifically, the SVI generated at each iteration w is:

$$SVI\left(\hat{\mathbf{X}}_{\mathbf{w}}\right) : \sum_i \sum_t \hat{X}^n_{itw} \sum_{u=1}^{t} Y^n_{iu} + \sum_j \sum_t \hat{X}^a_{jtw} \sum_{u=1}^{t} Y^a_{ju} \geq 1. \qquad (55)$$

This inequality states that at least one interdicted component in $\hat{\mathbf{X}}_w$ must be protected, either at time t or in a previous time period.

At each iteration w, the RMP for SVI-D is simply a feasibility seeking problem, including constraints $(34) - (36)$ and all the SVIs generated up to the current iteration. If a feasible solution to the RMP can be identified, SP is solved again with the new protection strategy $\hat{\mathbf{Y}}_w$ as input and the process is repeated. The algorithm stops when in the master model the protection resources are insufficient to thwart all the interdiction strategies discovered in the previous iterations, and thus the RMP becomes infeasible. Considering that the protection and interdiction resources are limited, the number of possible strategies is finite. Consequently, the RMP will become infeasible after a finite number of iterations.

The fact that inequalities (55) are supervalid is proven in the following proposition.

Proposition. $SVI(\hat{\mathbf{X}}_{\mathbf{w}})$ is supervalid.

Proof. : Let $[\hat{\mathbf{Y}}_{\mathbf{w}}, \hat{\mathbf{Z}}_{\mathbf{w}}\hat{\mathbf{X}}_{\mathbf{w}}]$ be the feasible solution of DNP found at iteration w. If this solution is optimal, then by definition inequality (55) is super-valid. If the solution is sub-optimal, adding inequality (55) to the RMP problem will generate a new protection strategy $\hat{\mathbf{Y}}_{\mathbf{w+1}} \neq \hat{\mathbf{Y}}_{\mathbf{w}}$. This strategy will in turn lead to a solution $\mathbf{\hat{Z}_{w+1}\hat{X}_{w+1}}$ of the SP that is different from the previous one because of constraints (46) and (47). Thus, for every w, the inequality is super-valid because it eliminates the incumbent solution, i.e.:

$$[\mathbf{\hat{Y}_{w+1}},\, \mathbf{\hat{Z}_{w+1}\hat{X}_{w+1}}] \neq [\mathbf{\hat{Y}_w},\, \mathbf{\hat{Z}_w\hat{X}_w}]. \quad \square$$

The main steps of the SVI-D algorithm are outlined below:

Algorithm 8 SVI-D

```
Set $w = 1$, $\hat{\mathbf{Y}}_w = \mathbf{0}$, $\mathbf{Y_{opt}} = \mathbf{0}$, $z_{opt} = \infty$.
MAINSTEP
Solve $SP(\hat{\mathbf{Y}}_w)$ to obtain $\mathbf{\hat{Z}_w\hat{X}_w}$ and the objective value $\hat{z}$
if $\hat{z} < z_{opt}$ then
  $z_{opt} = \hat{z}$ and $\mathbf{Y_{opt}} \leftarrow \hat{\mathbf{Y}}_w$.
end if
Add $SVI(\mathbf{\hat{X}_w})$ to $RMP$.
$w = w + 1$
Solve $RMP$ to obtain $\hat{\mathbf{Y}}_w$.
if $RMP$ is feasible then
  goto MAINSTEP
end if
TERMINATE
$Return(\mathbf{Y_{opt}})$
```

5.4 Results and Analysis

In this section, we investigate the computational efficiency of solving the dynamic network protection problem using BND-D and SVI-D. Both algorithms were implemented in C and run on a 64-bit machine with a quad-core 3.4GHz processor and 4GB of RAM. The *Restricted Master Problems* and the *SubProblems* were solved using the IBM ILOG CPLEX version 12.5 callable library. In our computational analysis, we set a time limit of $10,000$ seconds. In the algorithms' implementation, we used specialized data structures to store and retrieve information efficiently. Specifically, we observed that, given a protection strategy, each *Sub-Problem* could be decomposed into $|T|$ independent interdiction problems. Some of these sub-problems recurred multiple times across different iterations. For instance, consider a protection planning problem over two time periods and assume that in two different iterations of the algorithm we have to solve the interdiction problem with the two protection strategies listed in Table 8 as inputs.

Table 8: Protection strategies at two different iterations

T	Arcs	Nodes
0	1, 3	7
1	2	5

T	Arcs	Nodes
0	1, 2	5
1	3	7

The two protection strategies are different but, since protections are permanent, the sub-problem solved when $t = 1$ is the same in both iterations. In fact, in both cases, the arcs protected in the second period are $1, 2, 3$ and the nodes are $5, 7$.

It is possible to improve the algorithm by storing the solutions without solving the same problem multiple times. To store and retrieve the solutions efficiently we use a hash table. In fact, by using this data structure, insertions and search functions have a time complexity equal to $O(1)$ (Cormen et al., 2001). A hash table is an associative array in which keys are mapped to values, using a function known as hash function. For this problem we are using the protection strategy as key and the modulo as the hash function. As showed in Figure 16, the binary string representing the protection strategy is first converted into a decimal number and then is modulated by the function H. The protection strategy identifies a cell that holds the objective function value and optimal interdiction strategy of the lower level model. In order to consider the possibility of collisions, each cell of the table is a list.

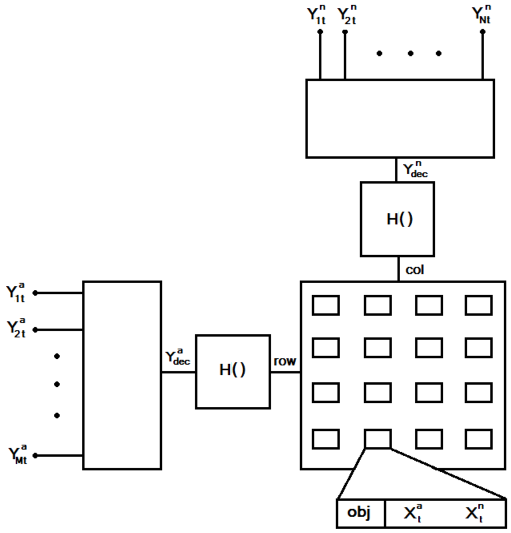


Figure 16: Hash data structure.

The initial testing was performed on two sets of randomly generated problems. Specifically, we generated 5 undirected networks with 10 nodes and 15 arcs, and 5 undirected networks with 20 nodes and 25 arcs. Distances were chosen uniformly from the set $\{1, 2, ..., 6\}$. The flow demand matrix was generated by drawing each value uniformly from $\{0, 1, ..., 100\}$. Each unit of flow can be interpreted as $10,000$ passengers. The costs of protecting / disrupting a node (q^n_i and p^n_i) were drawn uniformly among the values $\{2, 4, 6\}$. These three values were chosen to model stations of different size (small, medium and large). We also assumed $p^a_j = 1$. This choice was driven by the observation that in real life disrupting an arc is usually easier than disrupting a station. Tracks, in fact, are highly vulnerable because of their length and the presence of accessible and easily attackable structures (overpasses, bridges, tunnels). Just hitting one of these structures would impair the full link. On the other hand, the complete protection of a track can be an expensive task. Therefore, the values q_j^a were chosen uniformly from the set $\{1, 2, ..., 6\}$.

One of the assumptions of our model is that there is a limit to the number of arcs and nodes that can be disrupted. This budget limit is introduced to model disruptions of different magnitude. For example, a small interdiction budget indicates that the disruptive event only affects a few small components of the network. Conversely, a large disruption can affect a larger number of elements of the network and/or big assets. In our analysis, we consider three disruption scenarios. The interdiction resources associated with each scenario are shown in Table 9. Specifically, we assume that a small event is able to interdict only a small station or two arcs, whereas a large event is able to completely disrupt a big station or a combination of small components. The protection budget is assumed to be a percentage α of the total amount of resources needed to protect the full network, denoted by B. Namely, $B_{|T|} = \alpha B$. We consider values of α equal to 5% and 10%. The protection resources are spread in a 5-period planning horizon. The time periods are all weighted equally ($\lambda = [0.2, 0.2, 0.2, 0.2, 0.2]$).

Table 9: Disruption scenarios

Size	Resource units
Small	2
Medium	4
Large	6

The results for the two data sets are displayed in Table 10 and Table 11, respectively.

Table 10: Computational comparison between BND-D and SVI-D for the 10-15-x networks

Network	Disr units	Prot	Objective	Computing time (sec)		Prot	Prot	Disr	Disr
name	per period	budget	value	BND-D	SVI-D	arcs	nodes	arcs	nodes
10-15-1	2	5%	6886	0.52	0.01	1	0	6	2
	2	10%	6414	0.75	0.05	1	1	8	1
	4	5%	8513	0.59	0.05	1	1	18	1
	4	10%	8179	8.01	0.89	2	1	20	0
	6	5%	9992	0.64	0.03	1	1	9	5
	6	10%	9516	7.39	0.70	1	3	22	2
10-15-2	2	5%	5336	0.61	0.04	1	1	6	2
	2	10%	4884	4.26	0.36	2	1	8	1
	4	5%	7782	0.94	0.09	2	1	16	2
	4	10%	7449	66.79	2.02	4	1	16	2
	6	5%	9126	1.77	0.41	2	1	26	2
	6	10%	8829	> 10000	25.96	4	1	24	2
10-15-3	2	5%	5235	0.52	0.02	0	1	8	1
	2	10%	4611	0.76	0.04	2	1	10	0
	4	5%	7760	0.57	0.04	1	1	16	2
	4	10%	7200	2.14	0.44	2	1	18	1
	6	5%	9196	0.83	0.08	0	2	18	6
	6	10%	8796	62.24	3.16	1	3	26	2
10-15-4	2	5%	5122	0.71	0.04	3	0	10	0
	2	10%	4398	2.35	0.53	4	0	10	0
	4	5%	7280	2.09	0.39	2	0	16	2
	4	10%	6797	208.82	5.93	4	1	12	4
	6	5%	8832	7.74	0.62	3	0	20	5
	6	10%	8405	3479.83	33.82	4	1	16	6
10-15-5	2	5%	4642	0.52	0.02	1	0	10	0
	2	10%	4270	0.83	0.08	1	1	8	1
	4	5%	7270	0.74	0.05	1	1	14	3
	4	10%	6892	12.47	0.61	2	2	18	1
	6	5%	8811	1.43	0.43	1	2	22	4
	6	10%	8385	144.07	7.81	2	2	28	1
AVG			7226.93	138.65	2.82	1.87	1.07	15.30	2.03

For each network, disruption scenario and protection budget level, the tables show the DNP's objective function values, i.e., the worst-case disrupted flow over the planning horizon, the computing times of the two algorithms, and the number of network elements protected and disrupted in the optimal solutions. In these initial tests, the threshold for the path choice was fixed to 1, i.e., only the shortest paths are considered acceptable.

Table 11: Computational comparison between BND-D and SVI-D for the 20-25-x networks

Network	Disr units	Prot	Objective	Computing time (sec)		Prot	Prot	Disr	Disr
name	per period	budget	value	BND-D	SVI-D	arcs	nodes	arcs	nodes
20-25-1	2	5%	21807	2.28	0.11	2	0	10	0
	2	10%	19367	204.54	1.68	4	1	8	1
	4	5%	32436	1.89	0.16	2	0	4	4
	4	10%	31257	> 10000	11.43	3	2	12	2
	6	5%	37113	16.75	0.73	3	0	10	5
	6	10%	36580	> 10000	232.22	3	2	26	1
20-25-2	2	5%	22773	1.05	0.03	1	1	8	1
	2	10%	20241	20.15	1.11	4	1	8	1
	4	5%	31581	12.27	0.56	2	1	14	3
	4	10%	29370	9430.88	32.80	4	2	12	4
	6	5%	36833	41.37	2.56	1	2	12	9
	6	10%	35242	> 10000	1965.73	4	2	24	3
20-25-3	2	5%	24297	1.10	0.03	1	1	8	1
	2	10%	21534	14.03	0.88	3	1	8	1
	4	5%	32058	2.15	0.22	1	1	16	2
	4	10%	30800	487.11	8.39	5	1	14	3
	6	5%	37155	2.84	0.65	2	1	20	5
	6	10%	35646	> 10000	128.42	5	1	18	6
20-25-4	2	5%	27102	1.15	0.06	1	1	8	1
	2	10%	24255	21.37	0.64	3	1	8	1
	4	5%	34934	2.06	0.17	1	1	18	1
	4	10%	33105	853.55	9.52	4	1	12	4
	6	5%	39368	5.87	0.96	2	1	18	6
	6	10%	37842	> 10000	393.51	3	1	20	5
20-25-5	2	5%	27791	829.00	0.04	1	1	8	1
	2	10%	26065	43.73	1.45	4	2	6	2
	4	5%	36440	15.19	0.48	2	1	12	4
	4	10%	34949	6093.55	11.99	4	2	8	6
	6	5%	40817	82.36	1.82	1	2	24	3
	6	10%	39344	> 10000	3444.29	4	2	24	3
AVG			31270.07	757.76	208.42	2.67	1.20	13.27	2.97

The tables clearly show that SVI-D outperforms BND-D in every case. This is mostly due to the fact the RMP in the SVI-D approach does not have an objective function and, upon solving it, one stops as soon as a feasible solution is identified. As a consequence, the RMPs can be solved very quickly. The drawback is that without an objective to drive the protection strategy selection, the algorithm takes a considerable number of iterations (thousands for the complex instances) before converging to an optimal solution. Conversely, finding a solution to each RMP in the BND-D algorithm is quite time-consuming. Although this algorithm converges in a much smaller number of iterations compared to SVI-D, this is not sufficient to offset the greater difficulty of solving each RMP and its overall computing time is considerably higher. Furthermore, using both the approaches in one combined algorithm is inefficient. Results for this experiment are not reported because they are very similar to the results obtained by BND-D.

The impact of the size of the network is evident by comparing the two tables. Nonetheless, the high variability in the computing time suggests that the complexity of the problem depends on a combination of several factors, including the network topology. For instance, networks 20-25-4 and 20-25-5, although of equal size, have very different computing times.

In Table 12, we report some additional results for the largest data set using different path thresholds. The threshold value determines the number of acceptable paths, which in turn affects the size of the problems in terms of number of variables and constraints. Table 4 shows the impact of three different threshold values on the number of available paths and the computing time. A threshold value equal to 1.5 indicates that the users are willing to accept a 50% increase on their normal travel time, before switching to other transportation services or abandoning the trip. Similarly, a value equal to 2 indicates that a travel delay up to 100% is considered acceptable. Given the superiority of SVI-D, the computing times are reported for this algorithm only. In the analysis, we consider two protection levels ($\alpha = 5\%, 10\%$) and three disruption scenarios (2, 4, and 6 disruption units).

Although in most of the cases an increase in the path threshold value results

Table 12: Computational results for different path threshold values

Network	Threshold	Paths	Computing time (sec.)					
Protection level (α)			5%			10%		
Disruption scenario (p_t)			2	4	6	2	4	6
20-25-1	1	172	0.11	0.16	0.73	1.68	11.43	232.22
	1.5	247	0.19	1.44	4.00	1.64	198.13	1684.57
	2	327	0.18	4.66	6.29	3.86	114.65	1002.78
20-25-2	1	166	0.03	0.56	2.56	1.11	32.80	1965.73
	1.5	229	0.07	0.60	3.95	1.35	25.40	2730.99
	2	299	0.11	0.99	4.60	2.78	31.22	2469.20
20-25-3	1	172	0.03	0.22	0.65	0.86	8.39	128.42
	1.5	220	0.04	0.97	1.23	0.74	28.25	543.99
	2	317	0.04	1.04	4.02	1.75	41.19	1089.08
20-25-4	1	170	0.06	0.17	0.96	0.64	9.52	393.51
	1.5	245	0.06	0.77	1.10	0.95	26.88	395.96
	2	323	0.13	1.55	3.08	1.81	87.27	2587.95
20-25-5	1	175	0.04	0.48	1.82	1.45	11.99	3444.29
	1.5	269	0.11	0.78	1.71	1.59	14.27	4183.99
	2	331	0.26	0.75	6.18	1.91	24.76	1224.40

in an increase in computing time, there are some exceptions to this general trend, especially for large instances ($\alpha = 10\%$ and $p_t = 6$). For these instances, the most critical threshold value seems to be 1.5. As previously noted, these results point out that the performance of the algorithms are influenced by an interaction of different elements, such as the protection and disruption budgets, the network topological structure and the flow demand matrix. In general, increasing the number of acceptable paths increases the number of elements that must be targeted to disrupt a flow. As a consequence, the interdiction problems may become more difficult to solve. However, an increment in the path threshold value may also render some flows too difficult or even impossible to disrupt, thus reducing the number of possible interdiction plans and, consequently, the overall solution time.

To highlight how the path threshold affects the interdiction and protection optimal plans, in Figure 17 we compare the number of arcs and nodes protected and interdicted over the planning horizon.

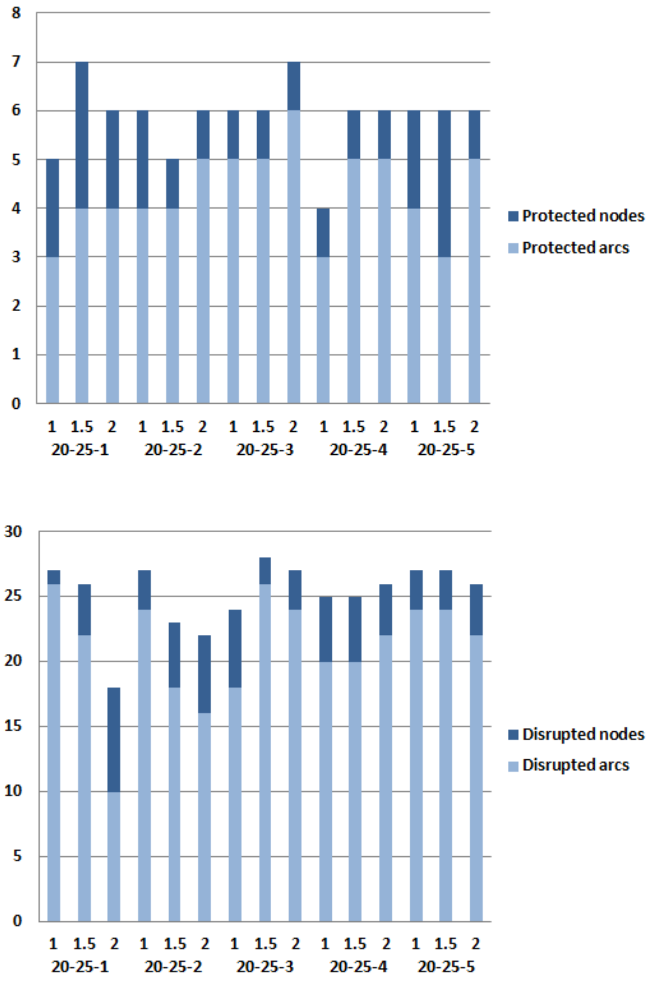


Figure 17: Impact of the path threshold on the number of protected/interdicted elements ($p_t = 6, \forall t$, $\alpha = 10\%$).

Changing the threshold value almost always results in different protection and interdiction plans. In some cases the changes are small, in others can be significant. For example, consider network 20-25-1. When the threshold is increased from 1 to 2, the number of interdicted arcs drops from 26 to 10, whereas the number of nodes increases from 1 to 8. This indicates that the interdiction plans are significantly different.

In summary, this analysis shows that changes to the path threshold parameter can have significant effects on both the problem complexity and the optimal solutions. Consequently, modelling users' behaviour accurately is a critical issue when solving this type of protection models for service systems.

5.5 Case study analysis

In this section, we test the efficiency of the decomposition approaches and analyse the results using a case study which represent the railway network of Kent (UK). The strategic position of this county makes the case study particularly interesting. Kent has a nominal border with France and, therefore, intercepts all the passenger flow from and to France. Although most of the traffic flow is represented by London commuters, Kent's railway has also a considerable traffic of tourists, attracted by historical places like Canterbury and Rochester. The overall network comprises 18 nodes, corresponding to cities and towns of the region, and 22 undirected arcs. The actual railway network is more complex, having more nodes and arcs. We simplified it by aggregating neighbouring stations and the corresponding flow generated/attracted by them. A graphic representation of the network is showed in Figure 18.

In the absence of real flow data, we generated the flow matrix as a function of the dimension of the connected towns. As in the computational result section, we assumed that disrupting an arc requires one unit of resources ($p_j^a = 1$). We also used the same assumption made for the protection/disruption of nodes: we divided the stations into three groups according to their annual passenger usage (Table 13). For example, Battle which is a small touristic town with less than

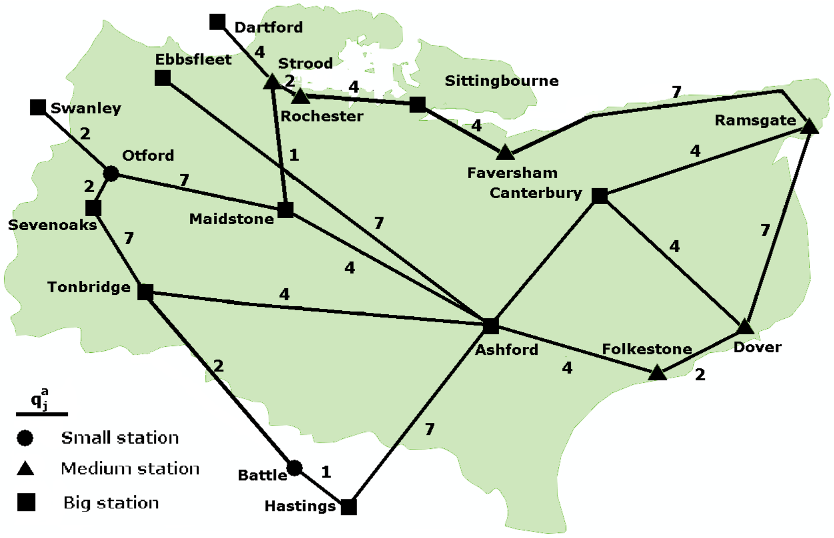


Figure 18: Railway network in Kent (UK).

half a million annual passengers, needs two units to be disrupted/protected. On the other hand, Ashford, which is a town of considerable size with more than 2 million annual passengers, requires six units. The number of protection units, q_j^a, needed to fully protect an arc depends on the number of tunnels and bridges that can be found on that arc. These numbers are displayed along the arcs in Figure 18. The disruption scenarios are the same as the ones used in the previous section (see Table 9).

Table 13: Resources needed to disrupt/protect a node

Node dimension	Disr/Prot resources
Small (annual passengers $<$ 0.5 M)	2
Medium (0.5 M $\leq$ annual passengers $<$ 1.5 M)	4
Big (annual passengers $\geq$ 1.5 M)	6

To compute the *acceptable* paths, we choose a threshold value equal to 1.5 (i.e., increases up to 50% of the normal travel time are considered acceptable). We focus on a 5-period planning horizon. In our initial investigation each period is weighted equally.

5.5.1 Impact of protection investments

In this section, we analyse the impact that different levels of protection resources have on the amount of flow loss, for different disruption scenarios.

Table 14: Percentage amount of flow loss for different disruption scenarios and protection budgets.

Scenario	No protection	5%	6%	7%	8%	9%	10%
small	27.37%	25.38%	24.48%	24.48%	23.97%	23.21%	22.34%
medium	34.50%	34.11%	32.92%	32.65%	32.65%	31.51%	31.12%
large	41.24%	38.18%	37.93%	36.50%	36.23%	35.69%	34.77%

For each disruption scenario and protection budget level, Table 14 displays the worst-case percentage flow loss when the optimal protection strategy for that scenario is implemented. The results show that even a small disruption can result in a loss of traffic flow as high as 27.37% of the total traffic, if no protection is carried out. This suggests that the network under study is highly vulnerable: even small, but possibly frequent, disruptive events can affect a significant portion of the flow. Obviously, the impact of disruption is more pronounced for medium and large disruption scenarios, with a flow loss of 34.50% and 41.24%, respectively. Investing in protection measures brings notable benefits. In particular, with a protection investment equal to 10%, the worst-case percentage flow loss can be reduced by about 18%, 10% and 16% in the three scenarios.

Figure 19 displays the marginal percentage decrease in flow loss, for each percent unit increment in protection resources. This graph provides in depth information on how each budget increment affects potential system losses. This analysis is useful to highlight the trade-off between protection expenditures and flow loss reductions in case of disruption. As an example, if small disruptions are considered, a 5% investment results in a worst-case flow loss reduction of about 7% (first segment of the first bar in the chart). If protection investments can be increased to 10%, the benefit is more than doubled with an overall flow loss reduction of about 18%. The graph also highlights possible investment inefficiencies. For example, for medium disruptive scenarios (second bar), increasing the budget from 7% to 8% has no impact on the worst-case flow loss, to denote that this added budget, although optimally allocated, is insufficient to thwart any additional interdiction plan.

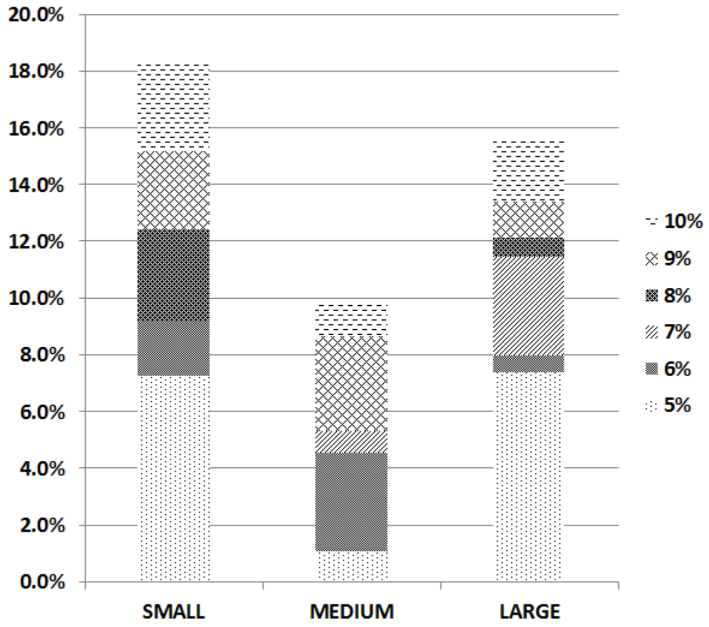


Figure 19: Marginal percentage decrease in flow loss due to unit increments of the protection budget.

5.5.2 Uncertainty of disruption events

One of the main issues involved in infrastructure protection planning is the intrinsic uncertainty of the disruption events. It is difficult and sometimes impossible to forecast when a disruption will happen and what its magnitude will be. The aim of the protection planner is to make the network as robust as possible, which means identifying a strategy that works well in all the possible scenarios. To this end, we consider how the optimal solution found for a given scenario, works if a different scenario occurs. The analysis is performed considering five equally weighted time periods and a protection budget equal to 10% of the resources needed to protect the full network. The results are shown in Table 15.

Table 15: Cross-comparison of different optimal protection plans. Relative percentage flow loss increase.

	Actual scenario			MAX	AVG
Supposed scenario	small	medium	large		
small	0%	4.3%	8.0%	8.0%	4.1%
medium	13.8%	0%	14.4%	14.4%	9.4%
large	12.7%	5.6%	0%	12.7%	6.1%

The table shows the percentage increase in unserved flow when an optimal strategy, obtained with a fixed scenario (*supposed scenario*), is used in a different scenario (*actual scenario*). The table also shows the maximum and the average increase across all different scenarios. Both solutions obtained for medium and large disruptions can be highly sub-optimal if a small disruption takes place, resulting in a flow loss increase of 13.8% and 12.7%, respectively. The best solution, in terms of both maximum and average values, is the one obtained for small disruptions. This seems to indicate that planning for a small disruption is overall a more robust strategy for this railway network. This result is further confirmed when the analysis takes into account that the three scenarios may have a different probability of occurring. For example, let us suppose that small, medium and large scenarios have a probability of occurrence equal to $0.7, 0.25, 0.05$, respectively. These probabilities can be incorporated in the decision process, for example by considering the expected percentage increase in flow loss. This can be computed as follows: $EV(small) = 0.25 * 4.3 + 0.05 * 8.0 = 1.5$, $EV(medium) = 0.7*13.8+0.05*14.4 = 10.4$, and $EV(large) = 0.7*12.7+0.25*5.6 = 10.3$. Using this criterion amplifies the benefit of adopting the protection strategy identified for small scenarios. Ultimately, which strategy to implement depends on the risk attitude of the decision maker. A risk adverse decision maker would probably opt for the optimal solution obtained for large disruptions, whereas a risk neutral decision maker may select the strategy found for small disruptions. Obviously, a thorough analysis of this issue would require the development of more sophisticated optimization models, which account for the probability of occurrence of different scenarios and explicitly incorporate robustness measures (Snyder and Daskin , 2006).

5.5.3 Dynamic investments

When dynamic investments are considered, a key questions is whether to opt for a protection strategy which renders the network as robust as possible at the end of the planning horizon, or for a strategy which guarantees high levels of protection as soon as possible (although this may decreases the overall efficiency of the final protection plan). In this subsection we investigate how the protection strategies change when the time periods are weighted differently. Specifically, we consider three different cases:

- CASE 1: $\lambda = [0.8, 0.05, 0.05, 0.05, 0.05]$. Here the aim of the protection planner is to obtain a good level of protection from the very first time period.

- CASE 2: $\lambda = [0.2, 0.2, 0.2, 0.2, 0.2]$. Here all the periods are equally weighted.
- CASE 3: $\lambda = [0.05, 0.05, 0.05, 0.05, 0.8]$. Here the aim of the protection planner is to maximise the safety level achieved when the protection strategy is fully implemented.

The protection budget used in this analysis is equal to 10% of the budget needed to protect all the assets in the network.

Table 16: Impact of the time weights on the worst-case percentage flow loss.

		CASE1			CASE2			CASE3	
Scenario	TL	IL	FL	TL	IL	FL	TL	IL	FL
Small	22.34%	26.40%	17.76%	22.34%	26.40%	17.76%	22.21%	27.01%	16.55%
Medium	31.12%	36.01%	28.10%	31.12%	36.01%	28.10%	31.51%	36.01%	27.86%
Large	34.77%	39.78%	30.90%	34.77%	39.78%	30.90%	34.77%	39.78%	30.90%

Table 16 has three columns for each case. The first one represents the total worst-case percentage flow loss over all the time periods (TL). The second represents the worst-case percentage flow loss in the initial time period (IL). The third represents the worst-case percentage flow loss in the final period (FL). This gives an indication of the protection level reached by the network at the end of the planning horizon. The analysis is done for three different disruption scenarios. Interestingly, the optimal protection strategy identified for large disruptions is the same, independently on the weights used in the objective function. Conversely, the other two scenarios present differences in the optimal protection strategies when more importance is given to the last time period (the first two cases are still equal). For small disruptions, giving more importance to the last period results in a more resilient final network, with a drop of the worst-case flow loss from 17.76% to 16.55%. Also the total flow loss slightly reduces from 22.34% to 22.21%. This indicates that aiming for the safest possible network after 5 years also results in a more resilient network during the transitory periods in which protections are implemented. For medium disruptions, the strategy to obtain a good level of protection in the last period results in higher losses of traffic (from 31.12% to 31.51%) throughout the planning horizon. Overall, for this case study, the weights given to the different time periods do not seem to have a massive impact on the protection strategies and on the network resiliency achieved of the end of the planning horizon.

5.5.4 Robust analysis

Robust analysis is a popular and powerful tool that has been extensively used to support decision making under uncertainty (Caplin and Kornbluth (1975), Best et al. (1986), Gupta and Rosenhead (1968)). Here we perform a robust analysis following the approach used by Rosenhead and Mingers (2001). The decision process is organized as a three-stage planning problem (Fig. 20)

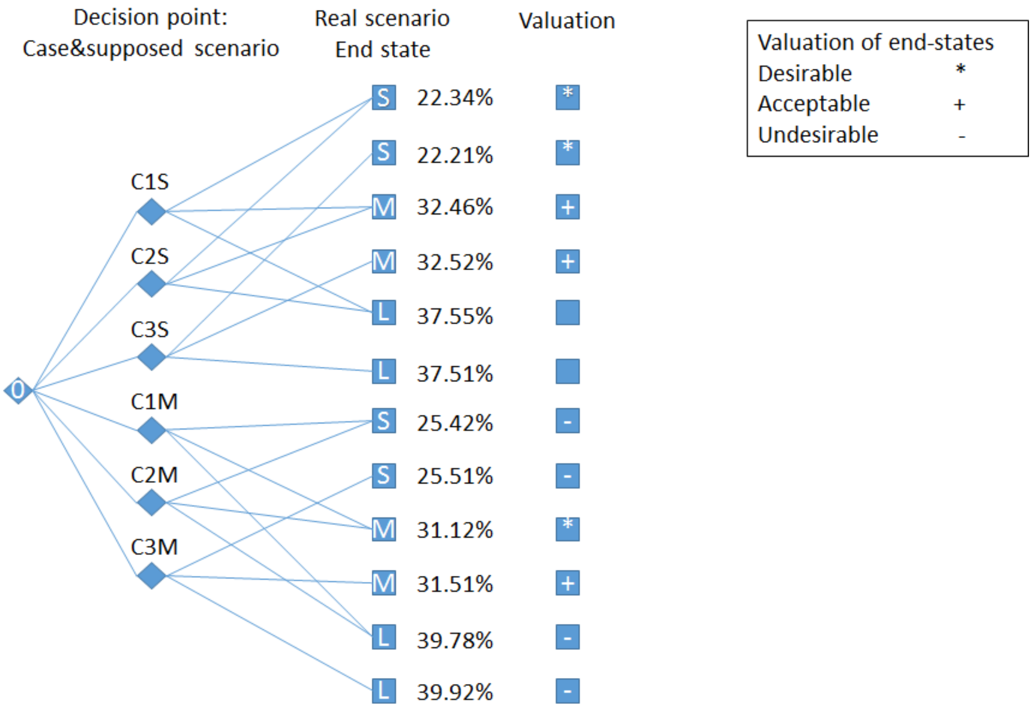


Figure 20: A three-stage planning problem with end-state valued.

A decision point represents a couple identifying the investment policy (case1, case2, and case3) and the scenario anticipated (small, medium , large). For example, decision C1M means that the decision planner is targeting medium scenarios aiming to obtain good level of protection from the very first time period (i.e., case 1). For sake of clarity, we have excluded large scenarios from the figure. Real scenario column is evaluated in terms of percentage flow loss. Finally, valuation column uses three possible options: desirable, acceptable and undesirable. The options for each state have been set using a reasonable empirical approach aiming at penalizing flow loss increments. Results are further summarized in Table 17.

Table 17: Summary of robust analysis.

	Options left open		
Initial decision	$*$	$+$	$-$
C1S	1	1	0
C2S	1	1	0
C3S	1	1	0
C1M	1	0	2
C2M	1	0	2
C3M	0	1	2
C1L	1	0	0
C2L	1	0	0
C3L	1	0	0

The analysis further confirms the results obtained in the last two sections: Solutions targeting small scenarios are more robust and weighting different time periods has only a small impact on this case study.

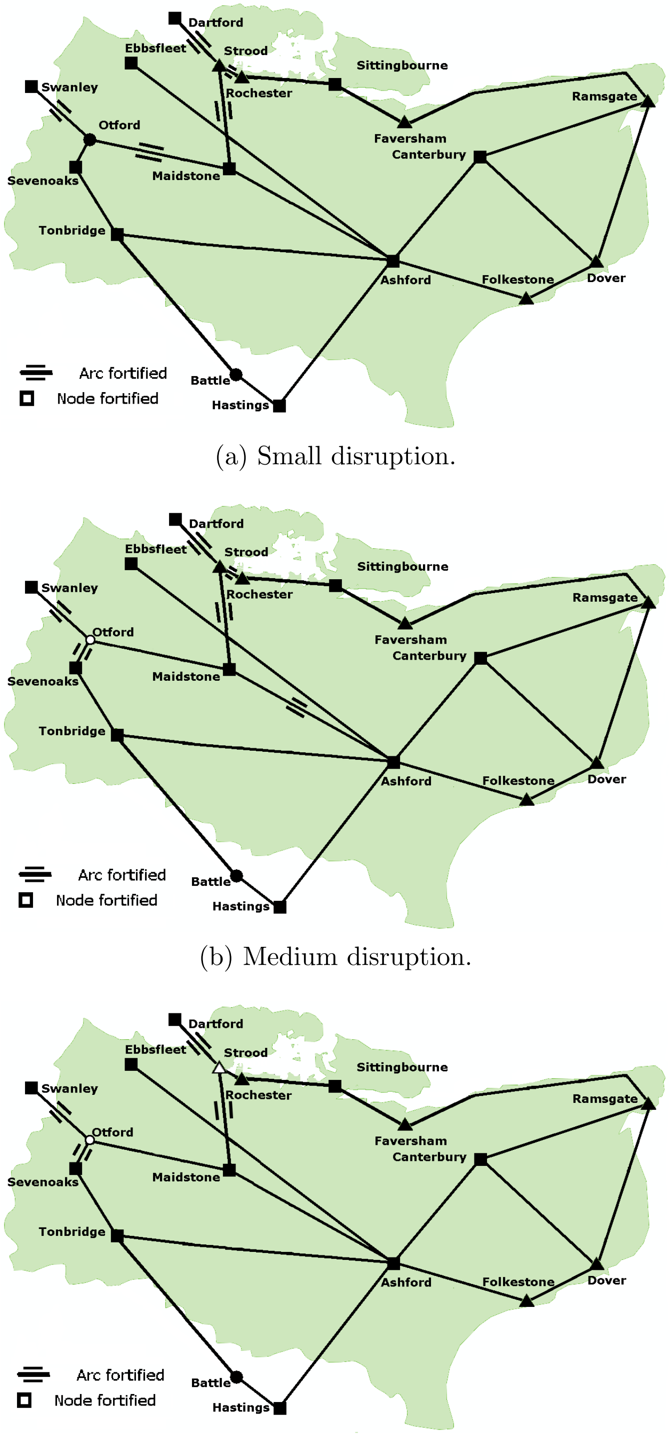
5.5.5 Solution analysis

In this section we show a sample solution of the proposed model. In particular, Figure 21 displays the assets chosen in the optimal protection plans, over the planning horizon, for the three disruption scenarios. The protection budget is again equal to 10% of the budget necessary to protect the entire network and all the time periods have equal weights.

It is clear that protecting the traffic to and from London is of strategic importance. In fact, both the arcs connected to Swanley and Dartford are chosen for protection. Also some arcs connected to Maidstone and Ashford are protected. These towns are among the most populated in Kent and therefore generate and attract high volumes of traffic. It is also interesting to notice that two relatively small stations like Otford and Strood are protected. This is a consequence of their strategic position. They intercept the traffic to and from London and are also directly connected to Maidstone. The main difference between the three graphs is that when the extent of a possible disruption increases (Figure 21c), more stations can be disrupted. Consequently, more stations appear in the optimal protection strategy.

Finally, Figure 22 shows the network components involved in a worst case disruption, after the implementation of the optimal protection strategies displayed in Figure 21. The interdiction strategies follow a pattern similar to the one identified in the protection plans. The affected components are, in fact, on the paths to and from London (link connected to Ebbsfleet), and on the paths to big or touristic stations (Maidstone, Ashford, Canterbury and Hastings).

Table 18 provides the details of how the optimal protection strategies are implemented over the planning horizon and, for each time period, displays the worst case interdictions. It can be noticed that the three protection plans share several targets to protect. Nonetheless the periods in which these targets are protected are usually different. Interestingly, in the second time period no protection is implemented for the small disruption scenario. This is because the resources available in this time period are saved to protect a larger asset (Maidstone-Otford link) in the successive period. This table highlights how, in a real protection planning situation, not only it is critical to choose what to protect but also when to protect the different assets.

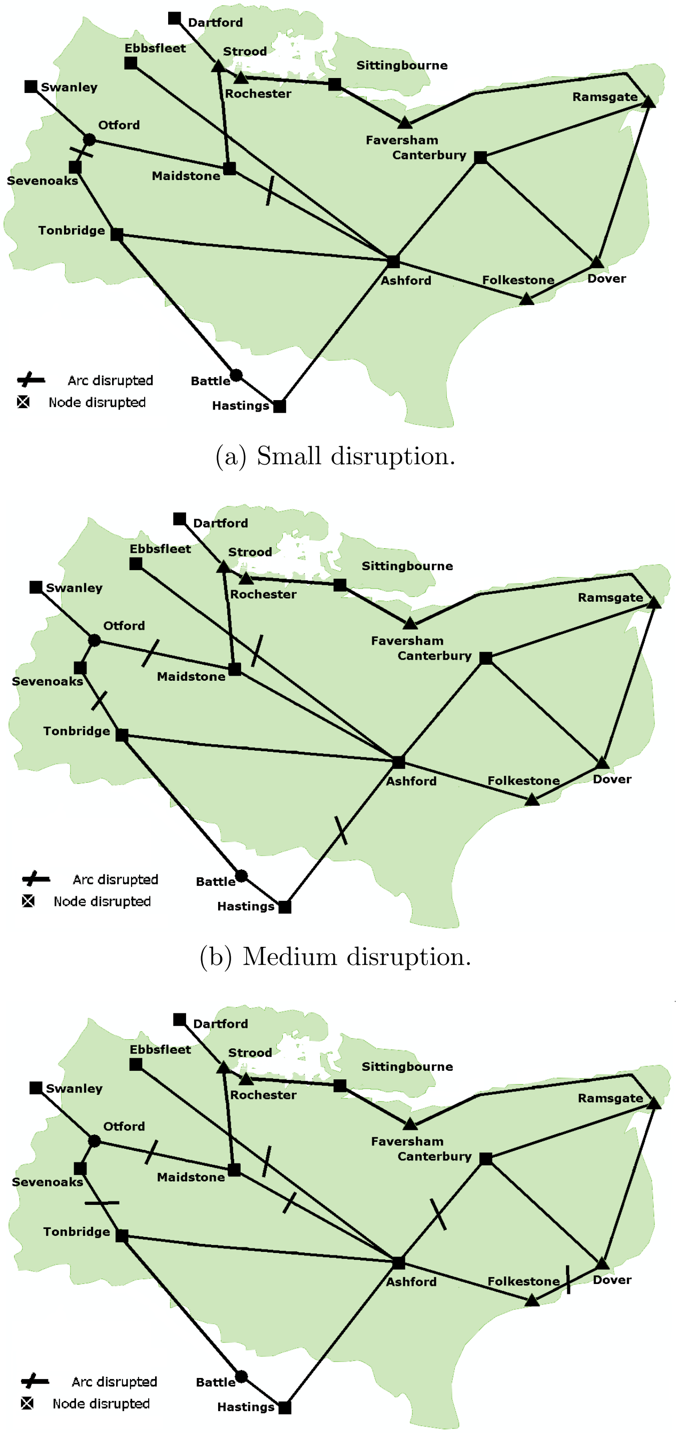


(a) Small disruption.

(b) Medium disruption.

(c) Large disruption.

Figure 21: Optimal protection plans for different disruption scenarios.



(a) Small disruption.

(b) Medium disruption.

(c) Large disruption.

Figure 22: Post-protection worst case losses in different disruption scenarios.

Table 18: Optimal protection plans and worst case losses over the planning horizon.

		PROTECTIONS		INTERDICTIONS	
Scenario	T	Arcs	Nodes	Arcs	Nodes
SMALL	0	Dartford-Strood		Ashford-Ebbsfleet Maidstone-Strood	
	1	Maidstone-Strood Otford-Swanley		Ashford-Ebbslfeet Maidstone-Otford	
	2			Ashford-Ebbslfeet Maidstone-Otford	
	3	Maidstone-Otford		Ashford-Ebbsfleet Rochester-Strood	
	4	Rochester-Strood		Ashford-Maidstone Otford-Sevenoaks	
MEDIUM	0	Dartford-Strood		Ashford-Ebbslfeet Maidstone-Strood Otford-Swanley Rochester-Strood	
	1	Maidstone-Strood Rochester-Strood		Ashford-Ebbslfeet Ashford-Maidstone	Otford
	2	Ashford-Maidstone		Ashford-Ebbsfleet Ashford-Hastings Maidstone-Otford Otford-Swanley	
	3	Otford-Sevenoaks		Ashford-Ebbsfleet Ashford-Hastings Maidstone-Otford Otford-Swanley	
	4	Otford-Swanley	Otford	Ashford-Ebbsfleet Ashford-Hastings Maidstone-Otford Sevenoaks-Tonbridge	
LARGE	0	Dartford-Strood		Ashford-Ebbslfeet Otford-Swanley	Strood
	1		Strood	Ashford-Canterbury Ashford-Ebbsfleet Dover-Folkestone Maidstone-Strood Otford-Swanley Rochester-Strood	
	2	Maidstone-Strood Rochester-Strood		Ashford-Canterbury Ashford-Ebbsfleet Ashford-Maidstone Dover-Folkestone Maidstone-Otford Otford-Sevenoaks	
	3	Otford-Swanley		Ashford-Canterbury Ashford-Ebbsfleet Ashford-Maidstone Dover-Folkestone Maidstone-Otford Otford-Sevenoaks	
	4	Otford-Sevenoaks	Otford	Ashford-Canterbury Ashford-Ebbsfleet Ashford-Maidstone Dover-Folkestone Maidstone-Otford Sevenoaks-Tonbridge	

5.6 Conclusions

In this chapter we introduced a bi-level fortification model to identify the best allocation of protection resources against worst case scenario disruptions in transportation networks. This model includes the important issue of considering dynamic investments. Two decomposition methods to find optimal solutions to the model were proposed and compared. The method based on super-valid inequalities clearly outperformed a classic Benders decomposition approach in terms of computational efficiency. Our analysis showed how the model results can be used to identify the optimal investment level to achieve a desirable degree of protection, and highlighted possible trade-offs between protection expenditures and traffic flow preserved in case of disruption. We applied the modelling approach to the Kent railway network and showed the optimal protection strategies for different disruption scenarios (small, medium and large). For this particular case study, the weights given to the different time periods in the objective of our dynamic model did not seem to have a significant impact on the optimal protection plans. Tests on some randomly generated problems indicated that a critical problem parameter is the path threshold value. This parameter is used to model the users' behaviour and identify the *acceptable* paths from a user perspective.

6 A model with variable post-disruption demand service for railway protection

In this chapter, we present a new optimization problem for railway network protection. The model is an extension of the first protection model introduced in chapters 4. A practical limitation of previous interdiction models for flow-based networks is that the flow between two nodes is considered *lost* or *unserved* only if the two nodes are completely disconnected after interdiction. Although this assumption simplifies the mathematical representation of the problems and their solution, it also limits their practical applicability, especially to the context of transportation networks. In transport systems, in fact, passenger demand between two nodes may be lost even if a connection does exist but the service is significantly deteriorated. As an example, if an interdiction causes long delays, travellers may resort to different modes of transportation or even abandon the trip. In the models introduced in chapter 4 and 5 we make a first attempt at redressing this shortcoming by introducing the concept of *acceptable paths*, i.e., paths whose length does not exceed the length of a shortest path by more than a given threshold. User demand is considered unserved if, after a disruption, no acceptable path is available between the two origin-destination nodes. However, these models still present some limitations in that each origin-destination path is either acceptable or not, and the resulting solutions are highly sensitive to the path threshold parameter used to define acceptable paths.

Following, we propose a bi-level protection model for railway infrastructures where the post-disruption user behaviour is modelled in a more accurate and realistic way. We refer to the proposed problem as the Network Protection Problem with

Variable Demand Loss (NPVDL). The aim is to find the optimal allocation of protection resources among railway assets (stations, tunnels, bridges, flyovers, rail tracks, etc.) so as to minimize the impact of worst-case disruptions on the service provision to rail passengers. The disruption impact is measured in terms of passenger flow (or demand) loss. A key aspect of our model is that it takes into account the system users' behaviour after a disruption. Travel time is one of the most important factors influencing route choice behaviour (Wang et al., 2014). After a disruption, increased travel times may cause some passengers to abandon the trip or resort to other means of transportation, with associated user disutility and system-wide costs. Failure to capture flow demand adjustments as a response to increased travel time in a mathematical model may lead to the identification of inaccurate and or suboptimal protection plans. Our protection model is unique by virtue of its inclusion of flow demand adjustments in response to increased travel time. Our model also considers the different costs of protection measures associated with the various assets.

6.1 Problem statement and formulation

6.1.1 Model assumptions

The NPVDL problem is formulated as a bi-level linear mixed integer model. The transportation network is modelled as a graph $G(N, A)$, where N is the set of nodes (e.g., stations) and A is the set of arcs (e.g., rail tracks). A limited budget is available for protecting the network. Interdiction resources are also assumed to be limited. This is a common assumption in interdiction modelling, where interdiction resources are used as a surrogate for the disruption magnitude. The demand for service between any two nodes is known and entirely served by the shortest path. If the shortest path becomes unavailable, the amount of flow loss depends on the length of the alternative routes. A disrupted element is completely unusable and, therefore, is removed from the network. An element, once protected, is immune to any disruption. Both arcs and nodes can be disrupted/protected. We also assume that the amount of resources needed to protect/disrupt an element is known.

6.1.2 A bilevel formulation for the NPVDL problem

The bilevel model for NPVDL uses the following notation.

Indices, sets and parameters

$s \in N$: index used for flow sources.

$d \in N$: index used for flow destinations.

$i \in N$: index used for network nodes.

$j \in A$: index used for network arcs.

c_j : nominal length of arc j.

f_{sd} : passenger flow between s and d.

N_{sd} : set of paths that connect s and d.

$r, t \in N_{sd}$: indexes used for network paths.

$A(r)$: set of arcs along path r.

$N(r)$: set of nodes along path r.

α_r : percentage of passenger flow using the service when path r is the shortest available path.

$P(r) = \{t \in N_{sd} : length(t) < length(r)\}$.

B : protection budget.

P : amount of interdiction resources available.

p_i^n, p_j^a : resource units needed to disrupt node i and arc j, respectively.

q_i^n, q_j^a : resource units needed to protect node i and arc j, respectively.

Decision variables

$X_i^n = 1$ if node i is disabled; 0 otherwise.

$X_j^a = 1$ if arc j is disabled; 0 otherwise.

$Y_i^n = 1$ if node i is protected; 0 otherwise.

$Y_j^a = 1$ if arc j is protected; 0 otherwise.

$\omega_r = 1$ if path r is available; 0 otherwise.

$S_r = 1$ if path r is the shortest non-disrupted path between a given origin and destination; 0 otherwise.

Z_{sd} = percentage of disrupted (or lost or unserved) flow, between s and d.

The problem is formulated as follows:

$$\text{[NPVDL]} \quad \min_{\mathbf{Y}} F(\mathbf{Y}) \tag{56}$$

$$\text{s.t.} \quad \sum_{i \in N} q_i^n Y_i^n + \sum_{j \in A} q_j^a Y_j^a \leq B \tag{57}$$

$$Y_i^n \in \{0,1\} \qquad \forall i \in N \tag{58}$$

$$Y_j^a \in \{0,1\} \qquad \forall j \in A \tag{59}$$

$$\text{where} \quad F(\mathbf{Y}) = \max_{\mathbf{X}} \sum_s \sum_d f_{sd} Z_{sd} \tag{60}$$

$$\text{s.t.} \quad X_i^n \leq 1 - Y_i^n \qquad \forall i \in N \tag{61}$$

$$X_j^a \leq 1 - Y_j^a \qquad \forall j \in A \tag{62}$$

$$\sum_{i \in N} p_i^n X_i^n + \sum_{j \in N} p_j^a X_j^a \leq P \tag{63}$$

$$Z_{sd} = 1 - \sum_{r \in N_{sd}} S_r \alpha_r \qquad \forall s, d \in N \tag{64}$$

$$\sum_{r \in N_{sd}} S_r \leq 1 \qquad \forall s, d \in N \tag{65}$$

$$\sum_{r \in N_{sd}} S_r \geq \sum_{r \in N_{sd}} \omega_r / |N_{sd}| \qquad \forall s, d \in N \tag{66}$$

$$S_r \leq \omega_r \qquad \forall s, d \in N, r \in N_{sd} \tag{67}$$

$$S_r \leq 1 - \sum_{t \in P(r)} \omega_t / |N_{sd}| \qquad \forall s, d \in N, r \in N_{sd} \tag{68}$$

$$\omega_r \geq 1 - \sum_{j \in A(r)} X_j^a - \sum_{i \in N(r)} X_i^n \qquad \forall s, d \in N, r \in N_{sd} \tag{69}$$

$$X_i^n \in \{0, 1\} \qquad \forall i \in N \tag{70}$$

$$X_j^a \in \{0, 1\} \qquad \forall j \in A \tag{71}$$

$$0 \leq Z_{sd} \leq 1 \qquad \forall s, d \in N \tag{72}$$

$$S_r \in \{0, 1\} \qquad \forall s, d \in N, r \in N_{sd} \tag{73}$$

$$\omega_r \in \{0, 1\} \qquad \forall s, d \in N, r \in N_{sd} \tag{74}$$

The aim of the *defender* is to minimize the overall flow loss (56), by distributing protection resources over the elements of the network. Constraint (57) represents the protection budget limit. The aim of the *attacker* is to maximize the flow loss

(60), by targeting the unprotected elements of the network. Constraints (61) and (62) state that nodes and arcs cannot be disrupted if they are protected. Constraint (63) limits the number of elements that can be interdicted. Constraints (64) define the percentage amount of flow between s and d which is lost. This is computed based on the available shortest path between the two nodes. Constraints (65) state that, for each pair of nodes, there can be at most one shortest path available. Constraints (66) state that, if there is at least one path available between s and d, there must also be a non-disrupted shortest path connecting the two nodes. Constraints (67) impose that a path can be the shortest available path only if it is available (i.e., not disrupted). Constraints (68) ensure that, given an origin-destination pair s and d, a path r can be the shortest available one connecting the two nodes only if all the paths shorter than r are unavailable. Constraints (69) state that a path r is disrupted only if at least one element (node or arc), belonging to that path, is interdicted. Finally, constraints (58)-(59) and (70)-(74) represents the domain restrictions of the decision variables.

6.2 SVI decomposition algorithm

To solve the problem, we propose a decomposition algorithm based on the use of Super Valid Inequalities (SVI).

The proposed approach involves decomposing the NPVDL into two sub-problems which are solved alternately: the Relaxed Master Problem (RMP) and the Sub-Problem (SP). The RMP is simply a feasibility seeking problem, consisting of a set of SVIs and constraints (57), (58), and (59). At each iteration, the RMP is solved to identify a feasible protection strategy $\hat{\mathbf{Y}}$. The Sub-Problem is subsequently solved to obtain the most disruptive interdiction strategy, $\hat{\mathbf{X}}$, in response to protection plan $\hat{\mathbf{Y}}$. Namely, SP is the lower level interdiction problem where the protection variables are fixed. SP is defined as follows:

$$\left[SP(\hat{\mathbf{Y}})\right] \quad z_{sp} = \max_{\mathbf{X}} \sum_{s} \sum_{d} f_{sd} Z_{sd} \tag{75}$$

$$\text{s.t.} \quad X_i^n \leq 1 - \hat{Y}_i^n \qquad \forall i \in N \tag{76}$$

$$X_j^a \leq 1 - \hat{Y}_j^a \qquad \forall j \in A \tag{77}$$

$$(63) - (74)$$

By solving SP, we obtain a feasible solution $(\hat{Y}, \hat{X})$ to the NPVDL and an upper bound to the problem. In addition, the new interdiction plan $\hat{\mathbf{X}}$ is used to generate an SVI, which is appended to the RMP in the next iteration. For a given $\hat{\mathbf{X}}$, the corresponding SVI is defined as:

$$SVI(\hat{\mathbf{X}}) : \sum_{i \in N} Y_i^n \hat{X}_i^n + \sum_{j \in A} Y_j^a \hat{X}_j^a \geq 1 \tag{78}$$

Inequality (78) simply states that to thwart an interdiction strategy, at least one element of that strategy must be protected.

The algorithm starts with an empty protection strategy and solves the SPs and the RMPs alternately. The iterative process terminates when the RMP becomes infeasible, i.e., the available protection resources are insufficient to thwart all the interdiction strategies discovered up to that iteration. Given that protection resources are limited, the algorithm is guaranteed to converge to an optimal solution in a finite number of iterations.

The following proposition states that constraint (78) satisfies the conditions of being super-valid (O'Hanley and Church, 2011).

Proposition. $SVI(\hat{\mathbf{X}})$ is super-valid.

Proof. : Let $\hat{\mathbf{X}}$ be the solution of $SP(\hat{\mathbf{Y}})$ found during a generic step of the algorithm. Then $(\hat{\mathbf{Y}}, \hat{\mathbf{X}})$ is the incumbent solution for the bi-level problem. If $(\hat{\mathbf{Y}}, \hat{\mathbf{X}})$ is optimal, the inequality is super-valid by definition. If it is sub-optimal, the inclusion of inequality (78) in the RMP will generate a new protection plan $\hat{\mathbf{Y}}\prime \neq \hat{\mathbf{Y}}$ to block interdiction strategy $\hat{\mathbf{X}}$. In turn, a new response $\hat{\mathbf{X}}\prime$ to $\hat{\mathbf{Y}}\prime$ will be generated. Therefore, at each iteration $(\hat{\mathbf{Y}}, \hat{\mathbf{X}}) \neq (\hat{\mathbf{Y}}\prime, \hat{\mathbf{X}}\prime)$. Thus, the incumbent solution is eliminated, making the inequality super-valid.

6.3 Heuristic approach

In the next section, we will show that the SVI-D decomposition method can only be employed to solve NPVDL instances of modest size. We, therefore, propose a heuristic approach that can be used as a good approximation of the exact algorithm and can tackle bigger size networks. The heuristic, referred to as GLS-H, is composed of a greedy-based construction phase, followed by a local search procedure. The algorithm uses two auxiliary models. The first model, called $USER(\hat{\mathbf{X}})$, is the system user sub-model. It computes the system's value (i.e., the amount of disrupted flow) associated with a specific interdiction plan $\hat{\mathbf{X}}$.

$$\left[USER(\hat{\mathbf{X}})\right] z_{user}(\hat{\mathbf{X}}) = \quad \max_{\mathbf{Z}} \sum_{s} \sum_{d} f_{sd} Z_{sd} \tag{79}$$

$$\text{s.t.} \quad (64) - (68)$$

$$\omega_r \geq 1 - \sum_{j \in A(r)} \hat{X}_j^a - \sum_{i \in N(r)} \hat{X}_i^n \qquad \forall s, d \in N, r \in N_{sd} \tag{80}$$

$$(72) - (74)$$

The second model, called $SP2(\hat{\mathbf{Y}})$, is a simplified version of $SP(\hat{\mathbf{Y}})$. This model assumes that the flow between two nodes is entirely lost only if all the paths connecting the two nodes are disrupted. If at least one path is available, independently on its length, all the flow is preserved. The mathematical formulation of $SP2(\hat{\mathbf{Y}})$, therefore, no longer requires the path variables S_r and ω_r and all the constraints associated with these variables. Also, the variables Z_{sd} are redefined as binary.

$$\left[SP2(\hat{\mathbf{Y}})\right] \quad z_{sp2}(\hat{\mathbf{Y}}) = \max_{\mathbf{X}} \sum_s \sum_d f_{sd} Z_{sd} \tag{81}$$

$$\text{s.t.} \quad (63)(70)(71)(76)(77)$$

$$Z_{sd} \leq \sum_{j \in A(r)} X_j^a + \sum_{i \in N(r)} X_i^n \qquad \forall s, d \in N, r \in N_{sd} \tag{82}$$

$$Z_{sd} \in \{0, 1\} \qquad \forall s, d \in N \tag{83}$$

Constraints (82) state that the flow between an origin s and a destination d is disrupted only if at least one element on each path connecting s and d is interdicted.

6.3.1 Greedy construction phase

In the initial step of the heuristic, we estimate how *important* each element of the network is, from the *attacker* point of view. For the sake of clarity, in this section we will ignore the difference between nodes and arcs and we will refer to them as network elements, belonging to the set E. $\mathbf{X}$ and $\mathbf{Y}$ will represent the interdiction and protection variables, respectively. Also, the disruption and protection resource vectors $\mathbf{p^n}, \mathbf{p^a}, \mathbf{q^n}$ and $\mathbf{q^a}$ are merged into two vectors $\mathbf{p}$ and $\mathbf{q}$.

Let $\mathbf{X^i}$ be a vector such that $X^i_j = 1$ if $i = j$ and 0 otherwise. Namely, $\mathbf{X^i}$ is an interdiction strategy where only element $i \in E$ is interdicted. The *importance* ρ_i of element i is computed by first solving $USER(\mathbf{X^i})$ to obtain $z_{user}(\mathbf{X^i})$. This value is then weighted by the resources needed to disrupt i. Formally, $\rho_i = z_{user}(\mathbf{X^i})/p_i$. This parameter is used as an estimate of the likelihood that element i appears in an interdiction plan and, consequently, in a protection plan. From now on, all the *greedy* choices will be made with reference to this parameter.

The greedy construction phase is a knapsack based heuristic. At each iteration the best element from a set of candidates is chosen and added to the solution. Let $\mathbf{Y^g_0}$ represent the initial protection plan. This plan is initialized with the best T elements, with respect to ρ. The exact value of T depends on the network size and protection budget. Let $\mathbf{Y^g_k}$ be the greedy protection plan after k iterations. $\mathbf{Y^g_k}$ is build from $\mathbf{Y^g_{k-1}}$ by adding the best element from the subset $\bar{E}$. $\bar{E}$ is the set of disrupted elements in the optimal interdiction plan, obtained by solving $SP(\mathbf{Y^g_{k-1}})$. The constructive algorithm ends when no element can be added to the solution, without violating the budget constraint.

6.3.2 Local search phase

After the greedy initial solution is built, a local search procedure is used to explore the solution space in the attempt to find improving solutions in the neighbourhoods of the current solution. The neighbourhood is defined by one-to-many swap moves. Let B^g and obj^g be the protection cost and the objective value of the greedy solution $\mathbf{Y^g}$, respectively. The pseudo-code of the local search procedure is shown below.

Algorithm 9 *LocalSearch*()

```
$obj_{sp} = obj^g, obj_{sp2} = obj^g, \mathbf{Y^{best}} \leftarrow \mathbf{Y^g}, B^{best} = B^g$
**for** $i = 1$ **to** $|E|$ **do**
    $\mathbf{Y^s} \leftarrow \mathbf{Y^{best}}, B^s = B^{best}$
    **if** $Y_i^s == 1$ **and** $\rho_i < \rho^H$ **then**
        $Y_i^s = 0$
        $swapin(0, B^s - q_i)$
    **end if**
**end for**
return $\mathbf{Y^{best}}, obj_{sp}$
```

This procedure simply scans the list of protected elements in the incumbent solution and tries to swap them out. Subsequently, the subroutine *swapin*() is called to identify the best elements to replace the outgoing element in the solution. To reduce the computational effort of the search phase, we restrict the set of elements that can be swapped out, by only considering those elements i with an importance factor less than a given threshold ρ^H ($\rho_i < \rho^H$).

Algorithm 10 *swapin*(ind, B^s)

```
while $B^s < B$ do
  for $j = ind + 1$ to $|E|$ do
    if $Y_j^s == 0$ and $\rho_j > \rho^L$ and $B^s + q_j \leq B$ then
      $Y_j^s = 1$
      Solve $SP2(\mathbf{Y^s})$ to obtain $z_{sp2}$
      if $z_{sp2} \leq obj_{sp2} + \epsilon$ then
        Solve $SP(\mathbf{Y^s})$ to obtain $z_{sp}$
        if $z_{sp2} < obj_{sp2}$ then
          $obj_{sp2} = z_{sp2}$
        end if
        if $z_{sp} < obj_{sp}$ then
          $obj_{sp} = z_{sp}$, $\mathbf{Y^{best}} \leftarrow \mathbf{Y^s}$, and $B^{best} = B^s + q_j$
        end if
      end if
      $B^s = B^s + q_j$
      $swapin(j, B^s)$
    end if
  end for
end while
```

The recursive routine *swapin* examines all the possible combinations of elements that can be swapped in. Every time a feasible swap is identified, the objective value of the new solution must be computed. In order to do this efficiently, we first solve the model $SP2$, which is significantly easier to solve than SP and may provide an indication of the quality of the new protection plan. If the new solution to $SP2$ improves or is close enough (ϵ is used to define how close) to the best solution found for $SP2$ (from the defender perspective), then there are good chances that the same solution may improve SP as well. Otherwise, the problem SP is not solved and another swap is attempted. This expedient reduces the number of times that SP is solved and, consequently, the overall computing time of the algorithm. As for the outgoing elements, we use a parameter ρ^L to reduce the number of elements that can be swapped in. The specific values of ρ^H

and ρ^L will be defined in the next section.

6.4 Results and analysis

In this section, the two solution approaches SVI-D and GLS-H are tested and compared on some randomly generated instances.

6.4.1 Data sets and problem parameters

We evaluate the algorithms' performances on a set of undirected networks of different size. We call the networks *n*-*m*-*x*, where *n* denotes the number of nodes, *m* the number of arcs and *x* is used to differentiate networks of same size. We focus on three different network sizes: 16-24, 25-40 and 36-60. For each size, 5 instances with the same grid topology were generated. Figure 23 illustrates the topology of a 16-24-x network. The networks 25-40 and 36-60 have the same square structure.

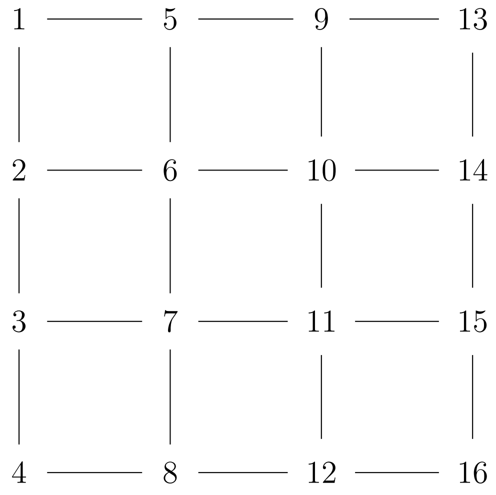


Figure 23: 16-24 grid topology

The other parameters of the problems are generated as follows:

– The length (or travel distance) of each arc c_j is drawn uniformly from the interval $[1, 100]$. We assume that the cost for protecting an arc depends on its length and set $q_j^a = c_j$.

– We model the presence of three different station sizes: small, medium and big stations. We assume that 40% of the stations are small, 40% are medium and 20% are big.

– The cost of protecting a node q_i^n, depends on the station's size. Namely, we assume that small stations require 20 units of protection resources to be fully protected, medium stations require 40 units and big stations require 60 units.

– The cost of interdicting a node p_i^n, depends on the station's size. Namely, we assume that small stations require 2 units of interdiction resources to be disrupted, medium stations require 4 units and big stations require 6 units.

– The cost p_j^a of disrupting any arc is set to 1. Tracks, in fact, are highly vulnerable and easy to disrupt because of their length and the presence of accessible and easily attackable structures (overpasses, bridges, tunnels).

– The flow matrix is assumed to be symmetrical and its values are drawn uniformly from $[0, 100]$.

– The disruption budget, P, is initially chosen to be equal to 6. This indicates that a disruption can disable a big station, 6 different links, or a combination of smaller assets (e.g., one small station and 4 links).

– The protection budget B is defined as a percentage of the budget T needed to protect the entire network, i.e. $B = QT$. We initially consider values of Q equal to 15% and 20%.

Finally, the values of the parameters used to model the travellers' behaviour, α_r, are given in Table 19. For each origin-destination pair, we use 4 different values which depend on the shortest path length increase. For instance, if the shortest path r connecting two nodes after a disruption is less than 20% longer than the shortest path connecting the two nodes before the disruption, then all the passenger demand is preserved ($\alpha_r = 1$). In contrast, an increase of the shortest path length over 100% (i.e., the new shortest path is more than twice as long as the initial shortest path) results in the loss of the entire demand ($\alpha_r = 0$).

Length increment	$\leq 20\%$	$> 20\%$ and $\leq 50\%$	$> 50\%$ and $\leq 100\%$	$> 100\%$
α_r	1	0.5	0.1	0

Table 19: Values of α_r as a function of the shortest path length increase.

6.4.2 Solution algorithms' setting

Both the exact and heuristic approaches are implemented using Cplex 12.5 embedded in a C++ program. Tests were run on a computer with 3.4 GHz quad-core processor and 8GB of RAM. The SVI-D algorithm uses Cplex default parameters. We enforce a time limit of $10,000$ seconds for its execution. The values of the parameters used by GLS-H were chosen empirically after some preliminary tests. Their setting is as follows:

- $\epsilon = 30$.
- ρ^H is the T^{th} highest value of ρ.

– ρ^L is chosen such that $|E^L| = |E|/4$, where $E^L = \left\{ i \in E : \rho_i < \rho^L \right\}$.

The values of T, shown in Table 20, were chosen empirically. These values depend on the network size and the protection budget. For example, for the 25-40-x networks and with a 15% budget, $\mathbf{Y^g}$ is initialized with the first 5 best elements.

	Network name		
Q	16-24-x	25-40-x	36-60-x
15%	3	5	10
20%	6	10	16

Table 20: Values of T for each combination of network size and protection budget values.

6.4.3 Performance comparison

In Table 21 and Table 22 we compare the performance of SVI-D and GLS-H for two protection budget levels: 15% and 20%, respectively. For both algorithms, the tables list the objective values and the computing times. The gap column shows the percentage error of the GLS-H solutions compared with the SVI-D solutions. For the heuristic approach, we also show the execution time for the two phases of the algorithm.

	SVI-D		GLS-H				
Network Name	Objective value	Computing time (s)	Objective value	Gap	Greedy time (s)	Search time (s)	Overall time (s)
16-24-1	3076.0	25.24	3076.0	0%	0.78	1.64	2.42
16-24-2	3582.0	17.75	3634.7	1.5%	1.45	3.82	5.23
16-24-3	3390.8	9.83	3390.8	0%	0.40	1.31	1.71
16-24-4	3201.2	24.36	3208.6	0.2%	0.51	4.45	4.98
16-24-5	3143.0	12.33	3143.0	0%	0.61	1.36	1.99
25-40-1	5590.3	2316.24	5590.3	0%	25.90	88.94	114.84
25-40-2	5435.2	1958.49	5435.2	0%	16.94	133.81	170.76
25-40-3	5579.8	1611.01	5609.2	0.5%	10.43	67.73	78.16
25-40-4	5825.8	2156.72	5848.9	0.4%	19.05	95.04	114.09
25-40-5	5528.0	1821.43	5528.0	0%	24.95	158.22	183.17
36-60-1	11195.7[⋆]	10000	10441.1	-6.7%	634.26	1349.67	1983.94
36-60-2	12834.4[⋆]	10000	10556.5	-17.7%	160.19	399.46	559.66
36-60-3	12714.8[⋆]	10000	9962.5	-21.6%	114.19	499.19	613.39
36-60-4	12027.7[⋆]	10000	10711.0	-10.9%	284.17	1557.88	1842.06
36-60-5	11997.4[⋆]	10000	10325.5	-13.9%	174.04	769.05	943.10
AVG	7008.14	3996.89	6430.75	0.3%[†]	97.86	342.10	441.30

[⋆] Objective value obtained after 10,000 sec.

[†] The average is computed excluding the cases where the gaps are negative.

Table 21: Computational results ($Q = 15\%$ and $P = 6$)

	SVI decomposition				Heuristic		
Network Name	Objective value	Computing time (s))	Objective value	Gap	Greedy time(s)	Search time(s)	Overall time(s)
16-24-1	2725.7	46.37	2725.7	0%	0.65	1.69	2.34
16-24-2	3313.0	53.2	3313.0	0%	0.61	2.47	3.08
16-24-3	2791.8	29.28	2791.8	0%	0.63	2.79	3.47
16-24-4	2836.3	46.57	2836.3	0%	0.53	4.60	5.19
16-24-5	2859.8	26.3	2859.8	0%	0.84	2.56	3.41
25-40-1	5388.1	3385.82	5439.3	1.0%	24.35	391.7	416.05
25-40-2	4864.3	3339.65	4864.3	0%	31.75	37.52	69.28
25-40-3	5051.7	3173.4	5051.7	0%	31.09	93.19	124.23
25-40-4	5322.0	3619.15	5322.0	0%	26.59	50.13	76.73
25-40-5	4818.0	3826.99	4818.0	0%	39.51	120.00	159.52
36-60-1	10322.4*	10000	9079.0	-12.0%	318.03	689.05	1007.50
36-60-2	10843.0*	10000	9503.0	-12.4%	98.71	387.32	486.04
36-60-3	11627.8*	10000	8999.9	-22.6%	79.80	318.93	398.75
36-60-4	11204.4*	10000	8437.6	-24.7%	1820.16	3055.88	4876.05
36-60-5	12423.3*	10000	9729.4	-21.7%	166.25	288.61	454.87
AVG	6426.11	4503.12	5718.05	0.1%†	175.97	363.10	539.10

* Objective value obtained after 10,000 sec.

† The average is computed excluding the cases where the gaps are negative.

Table 22: Computational results ($Q = 20\%$ and $P = 6$)

From the analysis of the tables, it is clear that SVI-D is able to solve to optimality only small and medium problem instances. The algorithm not only does not converge in any of the 36-60-x cases, but its solutions, after a considerable amount of time (10,000 secs), are always very far from the solutions obtained heuristically. GLS-H is able to identify good approximate solutions. This is true particularly when the protection budget is 20%. In this case the optimal solution is found for all small and medium networks, except the 25-40-1 network, where there is a 1% gap. The average gap, for proven optimal solutions, is 0.3% when $Q = 15\%$ and 0.1% when $Q = 20\%$. The heuristic seems to be both accurate and efficient. Its computing time is always considerably smaller than SVI-D's computing time. As an example, for the network 25-40-5 and $Q = 20\%$, the heuristic is 96% faster than the exact algorithm (Table 22). In the next section, we will show that the heuristic algorithm can be successfully used to identify cost-efficient protection plans for even larger networks.

6.5 Case study

In this section, we present a case study on the central London Tube. Fig. 24 displays the central portion of the London tube. The corresponding network, with 51 nodes and 70 undirected arcs, is shown in Fig. 25.

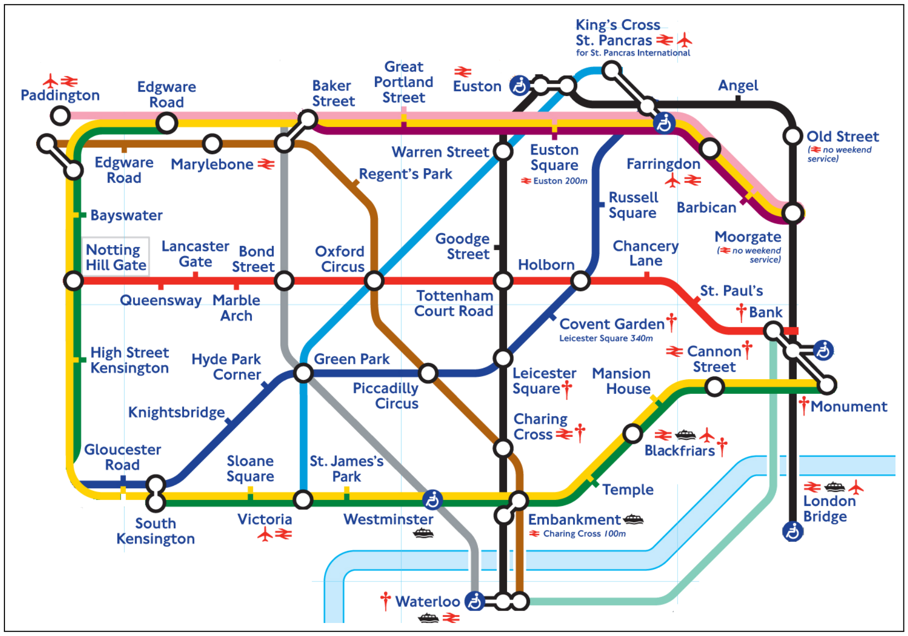


Figure 24: Central London tube map

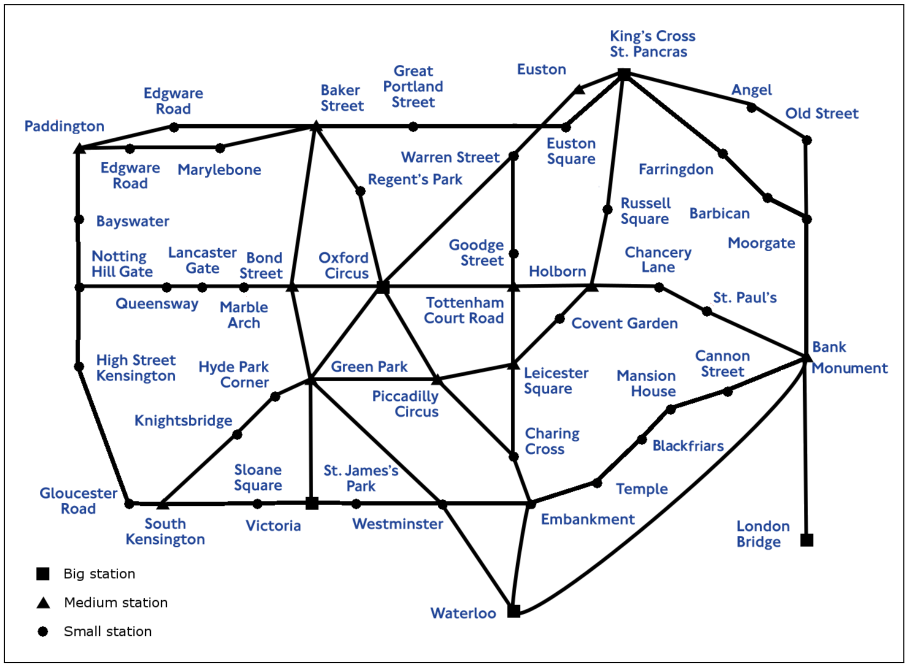


Figure 25: Central London tube network

To set the parameters of the problem, we use the open data available on the Transport For London website (TFL) (www.tfl.gov.uk/info-for/open-data-users). We use the running time between two directly connected stations to set the nominal cost of each arc, which is then used for computing the paths' length. The length of each path includes a 10-minute delay for each line change along the path. The physical distance of a connection is used to estimate its protection cost. This choice is motivated by the fact that typical protection strategies, such as digging draining pits, fortifying water pipes and sewers, and installing video surveillance, are all dependent on the length of the link. TFL also provides information regarding the flow from all the origins to all the destinations. This is used to build the demand matrix. We categorize the stations into three groups based on their sizes: small ($p_i^n = 2$), medium ($p_i^n = 4$) and big ($p_i^n = 6$). The annual flow of passengers is used to identify the category of each station. Namely, a station is *small* if the amount of annual passengers going through it is less than 25 millions, *medium* if it is between 25 and 50 millions, *big* otherwise. The rest of the parameters are set as explained in the previous section. We analyse different scenarios which varies in terms of protection budget $(0\%, 5\%, 10\%, 15\%, 20\%)$ and amount of disruption resources $(1, 2, 3, 4, 5, 6)$.

6.5.1 Impact of the protection budget on the flow loss

The impact of different protection budget levels on the system worst-case flow loss is displayed in Fig. 26. The analysis is performed for six scenarios, which differ in terms of the disruption magnitude, defined by the parameter P.

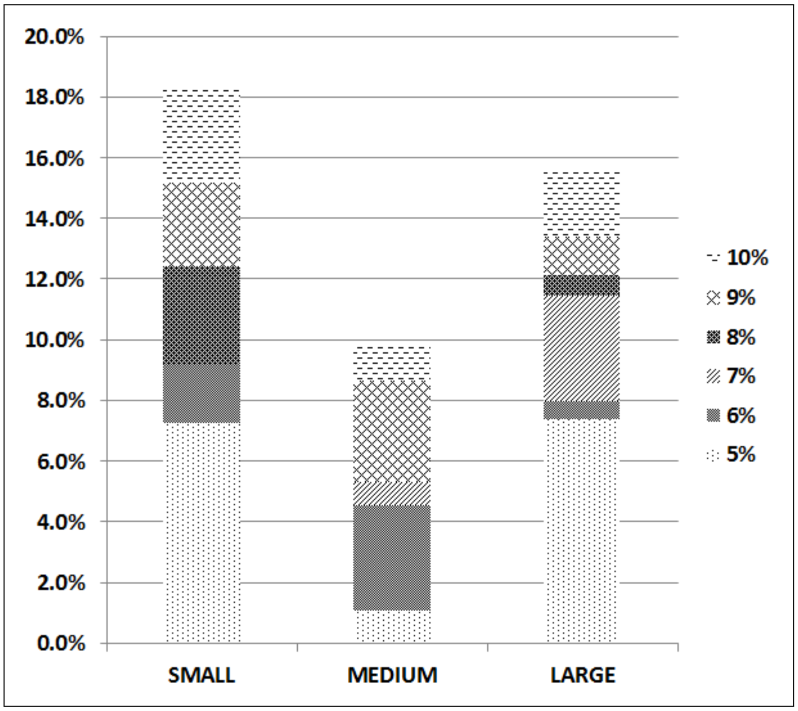


Figure 26: Flow loss for different budget levels and disruption scenarios

Clearly, increasing the protection resources from 0% to 20% can reduce the worst-case flow loss quite significantly, in every disruption scenario. For the largest disruption scenario ($P = 6$), the flow loss drops from more than 56% to about 34%. This means that, without any protection, a *large* disruption can potentially affect more than half of the entire traffic on the network. If 20% of the network is protected in a cost-efficient way, then the worst-case scenario flow loss drops to about one third of the total flow. Generally, all the budget increments prove to have a beneficial impact on the demand losses, although the marginal benefit due to the last increment decreases for small disruption scenarios (i.e., P = 1).

6.5.2 Optimal protection plans analysis

In this section, we analyse the protection plans identified by the model in different scenarios. We consider the same six disruption scenarios used in the previous section and four protection budget levels (5%, 10%, 15%, 20%). Tables 23 and 24 show the most frequently protected nodes and links of the network across the 24 scenarios.

Station	No. of protections
Westminster	18
Notting Hill	16
St. Paul's	10
Chancery Lane	9
Moorgate	9
Old Street	8
Marble Arch	7
Lancaster Gate	6
Bank/Monument	5
Holborn	3

Table 23: Frequency of protections for stations

Link	No. of protections
Holborn-Tottenham Court Road	24
Chancery Lane-St. Pauls's	24
Bank/Monument-St. Pauls's	24
Bond Street-Mable Arch	23
Chancery Lane-Holborn	23
Oxford Circus-Tottenham Court Road	21
Lancaster Gate-Marble Arch	19
Bond Street-Oxford Circus	19
Notting Hill-Queensway	18
Queensway-Lancaster Gate	18

Table 24: Frequency of protections for links

Table 24 shows that some key links (the first three) appear in every single protection plan. This is a clear evidence of how critical these assets are for the network: independently on the disruption scenario and available protection resources, these links must be protected to minimize the system's losses in case of disruption. Among the stations, Westminster and Notting Hill are clearly the most critical: they are protected in 18 and 16 out of the 24 cases, respectively (Table 23).

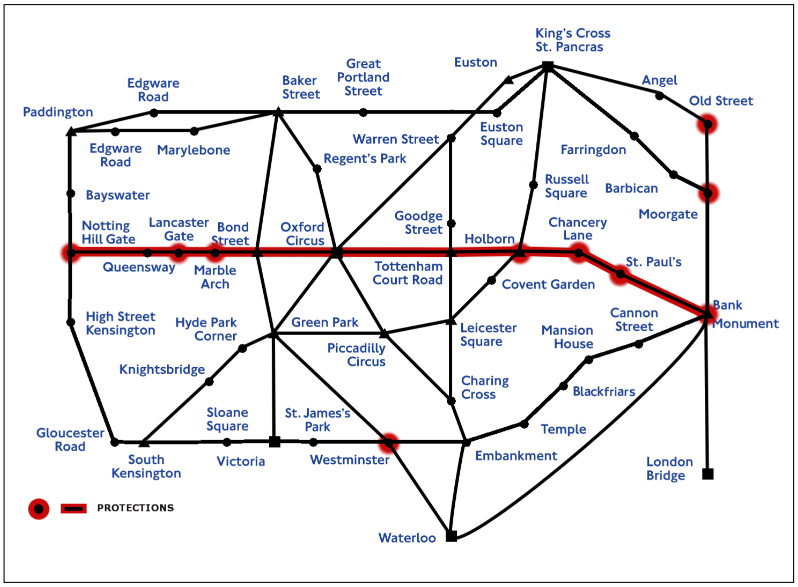


Figure 27: Most frequently protected elements

In Fig. 27 the most frequently protected elements of the network are highlighted. The figure shows a clear pattern in the protection. All the protected arcs and most of the protected nodes belong to the Central line. This is a plausible result, considering that the Central line is the busiest tube line, with over 260 million annual passengers.

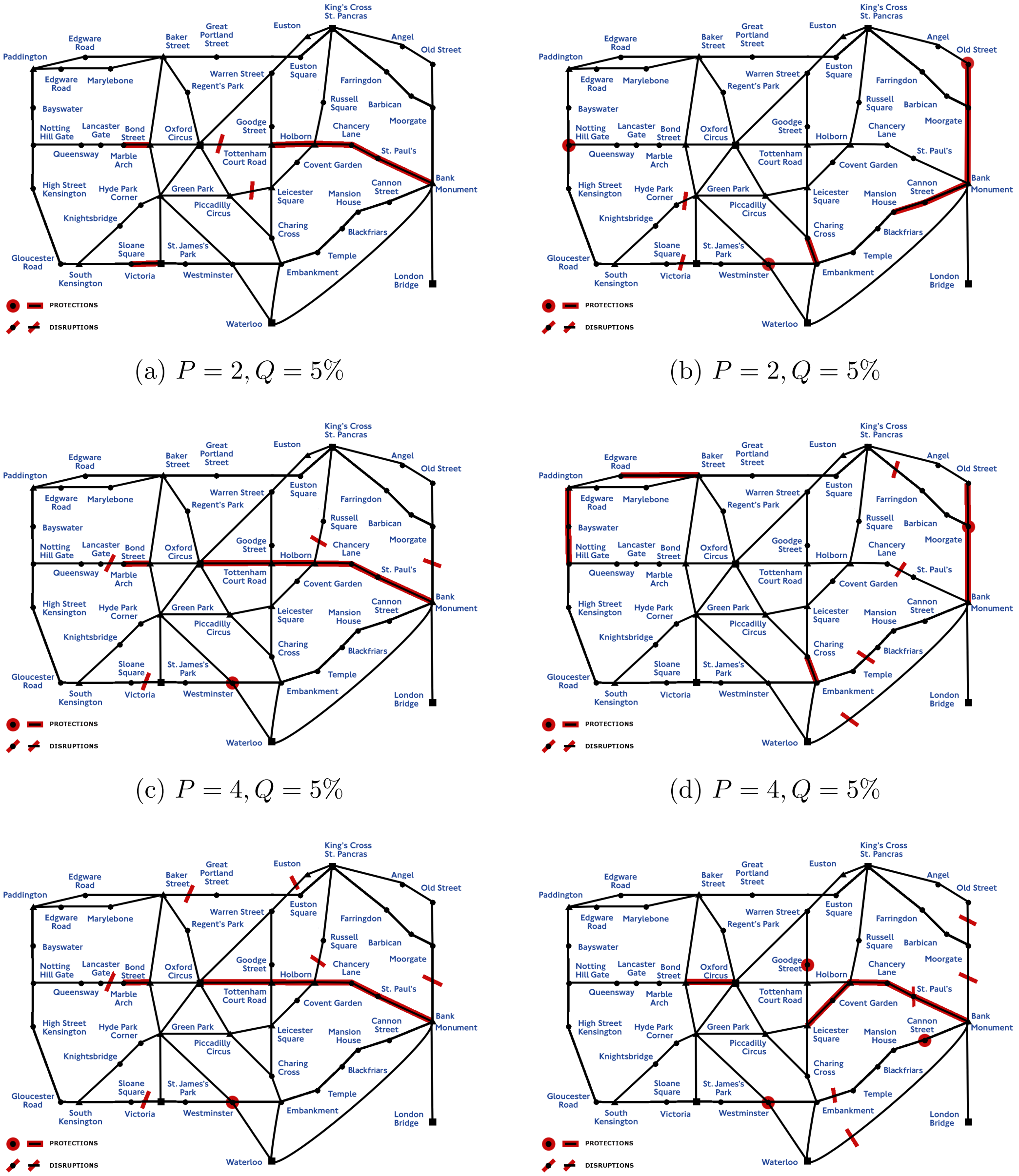
6.5.3 The importance of considering variable post-disruption demand

To prove the importance of modelling the user behaviour in an accurate way when planning protection efforts, we propose a comparison between the solutions obtained with NPVDL and the solutions obtained with the model introduced in Section 4 (referred to as NPCDL). As mentioned in previous Sections, this model relies on the definition of acceptable paths. In the following analysis, we consider two threshold values to define the acceptable paths: 2 (meaning that a path twice as long as the shortest path or less is considered acceptable) and 1.5 (meaning that a path up to 50% longer than the shortest path is considered acceptable).

In Fig. 28, we display the optimal protection plans and the post-protection, worst-case interdictions identified by the two models (NPVDL and NPCDL with threshold 2) for three different disruption scenarios ($P = 2, 4, 6$). For the sake of clarity, the figure displays the results for a protection budget level equal to 5% (protection plans involving fewer elements can be better visualized in the pictures). The pictures in the left column (28a, 28c, 28e) show the NPVDL solutions, whereas the pictures in the right column (28b, 28d, 28f) show the NPCDL solutions. It is evident that the solutions identified by the two models differ quite significantly, both in terms of protected elements and in terms of post-protection, worst-case disruptions. For example, when $P = 2$ and $P = 4$ the protection plans are completely different. Substantial differences were also noted for other values of the parameter Q.

To evaluate the impact that overlooking the user behaviour may have on the evaluation of worst-case demand losses, we use the optimal protection plans identified by NPCDL to compute the worst-case losses in our modelling framework (i.e., when the post-disruption passengers demand varies with the extent of the travel delay according to the pattern displayed in Table 19). Tables 25 and 26 display the percentage demand loss increase for different disruption scenarios and protection budget levels, when the threshold is equal to 2 and 1.5, respectively.

By observing the tables, it can be noticed that the solutions found by NPCDL are strongly suboptimal, especially for large values of Q and small disruptions. In this case study, the demand loss increase for both threshold values can be as high as 153% when $Q = 20\%$ and $P = 1$. Although the increase is less substantial for other combinations of the parameters P and Q, in all cases but two, solving the



(a) $P = 2, Q = 5\%$

(b) $P = 2, Q = 5\%$

(c) $P = 4, Q = 5\%$

(d) $P = 4, Q = 5\%$

(e) $P = 6, Q = 5\%$

(f) $P = 6, Q = 5\%$

Figure 28: Optimal protection plans and post-protection interdictions for NPVDL and NPCDL

more simplistic NPCDL model results in a misestimation of the real worst-case scenario losses in case of disruption. Note that the negative increase in Table 26, observed when $P = 3$ and $Q = 5\%$, is due to the fact that the heuristic solution to the NPVDL is not the optimal one.

	Q			
P	5%	10%	15%	20%
1	16%	57%	129%	153%
2	12%	12%	58%	71%
3	14%	19%	38%	56%
4	11%	19%	32%	51%
5	6%	12%	3%	8%
6	4%	9%	7%	16%
AVG	11%	21%	45%	59%

Table 25: Demand loss increase when using NPCDL with threshold 2

	Q			
P	5%	10%	15%	20%
1	16%	59%	139%	153%
2	11%	10%	49%	85%
3	−1%	16%	48%	99%
4	1%	15%	30%	24%
5	4%	2%	21%	8%
6	0%	5%	11%	16%
AVG	5%	18%	50%	64%

Table 26: Demand loss increase when using NPCDL with threshold 1.5

These results show empirically that failure to accurately represent the user behaviour into a modelling framework may lead to highly sub-optimal protection strategies, where limited protection resources are not allocated in a cost-effective way.

6.5.4 Solution analysis with a different user behaviour

In the previous section, we have demonstrated that considering the user behaviour is crucial to identify sound protection strategies. In this section, we analyse the sensitivity of our solutions to different values of the parameter α_r, used to capture the users behaviour. To this end, we run a new set of experiments where the values of α_r have been changed to the values shown in Tab. 27.

Length increment	$\leq 40\%$	$> 40\%$ and $\leq 70\%$	$> 70\%$ and $\leq 100\%$	$> 100\%$
α_r	1	0.5	0.1	0

Table 27: New values of α as a function of the shortest path increase.

These values indicate that users are willing to accept longer travel delays, as compared to the ones used in the previous analysis. For instance, a travel time increase up to 40% does not cause any flow loss, whereas previously a 40% increment would have led to the loss of 50% of the flow.

Tables 28 and 29 show the most frequently protected stations and links, across all the proposed scenarios.

Station	No. of protections
Westminster	15
Notting Hill	15
Old Street	10
Chancery Lane	9
St. Paul's	9
Marble Arch	8
Moorgate	7
Lancaster Gate	6
Bank/Monument	4
Holborn	3

Table 28: Frequency of protections for stations, with new α values.

Link	No. of protections
Bond Street-Mable Arch	24
Chancery Lane-Holborn	24
Chancery Lane-St. Pauls's	22
Holborn-Tottenham Court Road	22
Bank/Monument-St. Pauls's	20
Lancaster Gate-Marble Arch	20
Oxford Circus-Tottenham Court Road	18
Notting Hill-Queensway	18
Queensway-Lancaster Gate	18
Bond Street-Oxford Circus	14

Table 29: Frequency of protections for links, with new α values.

Changing α has an obvious impact on the objective function. There is, in fact, an average 4.4% decrease in the flow loss. Nonetheless, it seems that the protection plans have not changed significantly. Tables 28 and 29 show the same patterns highlighted by Tables 23 and 24. No new element appears in the protection plans and there are only small variations in the frequency of the protected elements. This suggests that, for this particular case, the solutions identified by our model are quite robust to variations of the parameter α. As mentioned in the introduction, estimates of this parameter can be obtained by surveying a sample of the railway system users. A small misestimation of this figure should not have a major impact on the protection strategies identified by the model.

6.6 Conclusions

In this chapter we introduced a new modelling approach for increasing the reliability and security of flow-based networks. Our focus is on railway systems. The proposed approach overcomes some of the limitations of pre-existing models, by capturing the user behaviour in a post-disruption period. Specifically, our model assumes that the demand for service after a disruption depends upon the extent of travel delay of each origin-destination route on the network. Results show that failing to consider the user behaviour may lead to sub-optimal protection plans and an underestimation of disruption consequences.

The inclusion of the post-disruption user behaviour into a mathematical model significantly increases the model's complexity and tractability. To identify optimal or near-optimal solutions to the problem, we developed an exact method and a heuristic solution approach. The exact algorithm is an iterative procedure based on the concept of Super-Valid Inequalities. The heuristic algorithm is composed by a greedy construction phase followed by a local search procedure. Computational tests on some randomly generated networks show that the exact method, although useful to assess the accuracy of the heuristic on small problems, can only tackle networks of modest size. In contrast, the heuristic proves to be both efficient and effective in identifying high quality solutions. The application of the modelling approach to a real rail network (the London tube) provides a practical demonstration of how limited protection resources can be allocated in a cost-efficient way among the most vulnerable assets of a rail system. It also highlights how some key elements must be protected in every disruption scenario to achieve high level of network security. Finally, the case study highlights the fact that neglecting the post-disruption user behaviour may lead to the identification of highly inefficient protection strategies, with worst-case disruption losses 150% higher than those obtained with our model.

7 A model to protect road network infrastructure against flooding

This chapter focuses on the problem of identifying optimal protection strategies to reduce the impact of flooding on a road network. We propose a dynamic mixed-integer programming model that extends the classic concept of road network protection by shifting away from single-arc fortifications to a more general and realistic approach involving protection plans that cover multiple components. We also consider multiple disruption scenarios of varying magnitude. To efficiently solve large problem instances, we introduce a customized GRASP heuristic. Finally, we provide some analysis and insights from a case study of the Hertfordshire road network in the East of England. Results show that optimal protection strategies mainly involve safeguarding against flooding events that are small and likely to occur, whereas implementing higher protection standards are not considered cost-effective.

Natural hazards can have serious impacts in terms of economic losses and human casualties. Floods can be particularly dangerous due to the high number of people living in at-risk areas and due to the high frequency of occurrence. Both coastal and inland areas can be affected by floods caused by the overflow of rivers and the sea. Even areas that are far from a water course or coastal zone can still be at risk of flooding caused by heavy rain. The World Resources Institute estimates that climate change and population rise will more than double (from 21 to 54 millions) the number of people exposed to floods (Luo et al., 2015). In the UK, one in six properties are at risk of flooding (Environmental Agency, 2009). The 2005 flood in New Orleans caused by Hurricane Katrina is one of the most tragic examples of the disruptive power of such events. Even less dramatic floods can still lead to major disruption of vital services and significant economic losses. During the winter of 2013-14 in the UK, heavy rainfall triggered numerous floods that left major roadways under water and disrupted several train lines. The government subsequently allocated £130M to repair and maintain flood defences (Carrington and Weaver, 2014).

The first step to mitigate flood risk has to be done at the policy-level so as to regulate urban development in hazardous areas. Subsequently, risk analysis should be implemented to identify vulnerabilities and suitable protection measures. Resources should also be used to devise post-disaster strategies to reduce loss of life and economic damages.

Flood damages can be split into three categories: urban, rural, and infrastructure (Dutta et al., 2003). In this chapter we deal with the issue of protecting the road transportation system. Transportation systems are often highly exposed to flood risk due to their considerable size, which makes implementing fully comprehensive protection measures all but impossible. In order to cope with the limited availability of resources for securing road networks, we propose a multi-period optimization model for selecting multiple asset protection plans to guard against a specified set of flood scenarios. Flood events are usually classified using the "return period" concept (Gumbel, 1941). The return period is the time interval between two events of the same disruption size and can be used to estimate the probability of a given scenario occurring. A 100 year flood, for instance, it is estimated to happen with a 1% chance in any given year.

The same concept is normally used to define the standard of protection offered by an engineered flood defence measure, which may include building/replacing dikes, sluices, slopes, embankments, and culverts. These structures differ in terms of the standard of protection guaranteed, the cost to implement them, and the size of the area afforded protection. Relatively inexpensive defences like vertical walls or concrete structures can often be quickly built to protect a small, targeted area. More ambitious plans, such as the construction of diversion canals or large dams can impact larger areas, but normally at a much higher cost.

Although the literature dealing with transportation disruption is vast (Chang (2003), Huang et al. (2007), Fan and Liu (2010), He and Liu (2012)), very little attention has been given to the effects of extreme weather conditions on transportation services. Suarez et al. (2005) propose a model to assess how flooding and climate change impact the performances of urban transportation. They use an Urban Transportation Modelling System (UTMS) to simulate traffic flows and apply the model to the Boston Metro Area. Sohn (2006) assess the critical links in a highway network under flood damage. Criticality is estimated according to an accessibility metric based on shortest distances and traffic flows.

7.1 Model Formulation

Let a road network be represented as a graph $G(N, A)$, where N is the set of nodes (junctions) and A the set of arcs (road links). Suppose a set of flood disruption scenarios S is under consideration. A given flood disruption scenario $s \in S$, which occurs with probability π_s, will cause a subset of arcs $A_s \subseteq A$ to be disrupted. The disruption caused by any scenario s can vary spatially such that arcs may face a range of discrete flood magnitudes represented by set L. The return period of a given flood size $\ell \in L$ is given by r_ℓ and expresses the average time interval (normally expressed in years) between two floods of the same magnitude. A flood's magnitudes and return period are directly linked. Specifically, flood events with longer return periods have higher magnitude and vice versa. Set $R = \{r_\ell : \ell \in L\}$ denotes all possible return periods associated with flood sizes L.

Under normal conditions, travel time along an arc $k \in A$ is given by d_k. Assuming no additional protection is provided to an arc, flood scenario $s \in S$ will cause a delay in travel time of Δd_{ks} for arc $k \in A_s$ as a result of facing a flooding event with a return period of $\rho_{ks} \in R$. We assume that protection of any arc k to a standard sufficient to withstand a flood of magnitude $\ell \in L$ will protect it against any flood of lesser magnitude (i.e., no delay will occur for any flood $\ell' \in L$ having a return period $r_{\ell'} < r_\ell$). With this in mind, let P be a set of protection plans. An individual protection plan $p \in P$, which costs c_p to implement, can protect multiple arcs $\hat{A}_p \subseteq A$ to various standards. Specifically, plan p will protect arc $k \in \hat{A}_p$ against any flood with a return period of $\sigma_{kp} \in R$ or less.

Now the aim of our problem is to select a subset of protection plans in P over a specified planning horizon of length T in order to minimise, across disruption scenarios S, the expected all-pairs shortest path from a defined origin $o \in N$. Each time period is subject to the same set of scenarios S. It is assumed that there is a budget b_t available in each period for implementing protection plans and that unused portions of the budget can be carried forward to subsequent periods. It is further assumed that protection of any arc in time $t = 1, ..., T$ lasts for the remainder of the planning horizon $t, ..., T$. In cases where two or more protection plans are implemented and would provide overlapping protection to the same arc, the arc will be protected to the highest standard among the plans (e.g., if an arc has already been protected to a standard ℓ and a subsequent plan would protect it to a lower standard ℓ' such that $r_{\ell'} < r_\ell$, then the previous higher standard ℓ

would be retained).

7.2 A Dynamic model for Road Protection against Flooding

To formulate our Dynamic Road Protection against Flooding (DRPF) model, consider the following additional notation:

$$FS(i) = \text{the forward star of node } i$$

$$RS(i) = \text{the reverse star of node } i$$

$$M = \text{some very large constant } (\geq \text{ the largest return period in set } R)$$

We also introduce the following decision variables:

y^o_{kst} = the number of times arc k is included in a shortest path originating from node o to any other node if scenarios s occurs in period t

$$z_{pt} = \begin{cases} 1 & \text{if plan } p \text{ is implemented in period } t \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{kst} = \begin{cases} 1 & \text{if arc } k \text{ is disrupted by scenario } s \text{ in period } t \\ 0 & \text{otherwise} \end{cases}$$

β_{kt} = the largest return period that arc k would be protected against in period t

With this in place, a non-linear formulation of DRPF, based on the well known single origin to multiple destination problem, is given below.

$$[\text{DRPF}] \quad \min \sum_{s \in S} \pi_s \sum_{t=1}^{T} \sum_{o \in N} \sum_{k \in A} (d_k y_{kst}^o + \Delta d_{ks} \alpha_{kst} y_{kst}^o) \tag{84}$$

$$s.t.$$

$$\sum_{u=1}^{t} \sum_{p \in P} c_p z_{pu} \leq B_t \quad t = 1, ..., T \tag{85}$$

$$\sum_{k \in RS(i)} y_{kst}^o - \sum_{k \in FS(i)} y_{kst}^o = \begin{cases} -|N| + 1 & \text{if } i = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$\forall o \in N, \forall i \in N, \forall s \in S, t = 1, ..., T \tag{86}$$

$$\beta_{kt} = \max_{u \leq t,\, p \in P} \{\sigma_{kp} z_{pu}\} \quad \forall k \in A, t = 1, ..., T \tag{87}$$

$$M\alpha_{kst} \geq \rho_{ks} - \beta_{kt} \quad \forall k \in A, \forall s \in S, t = 1, ..., T \tag{88}$$

$$y_{kst}^o \geq 0 \quad \forall o \in N, \forall k \in A, \forall s \in S, t = 1, ..., T \tag{89}$$

$$z_{pt} \in \{0, 1\} \quad \forall p \in P, t = 1, ..., T \tag{90}$$

$$\alpha_{kst} \in \{0, 1\} \quad \forall k \in M, \forall s \in S, t = 1, ..., T \tag{91}$$

$$\beta_{kt} \in \mathbb{R} \quad \forall k \in M, \forall t = 1, ..., T. \tag{92}$$

The aim of DRPF model is to choose protection plans that minimize the expected all-pairs shortest path cost over all disruption scenarios (84). Inequalities (85) impose a budget restriction on the cost of protection plans in each time period t. Equations (86) are the flow-balance constraints for an all-pairs shortest-path problem. Equations (87) set β_{kt} to the largest return period that arc k would be protected against taking into account all of the protection plans covering arc k that have been implemented in periods $u \leq t$. Constraints (88) require variable α_{kst} to be equal to 1 whenever the highest safety standard implemented for arc k by period t is insufficient to protect against scenario s (i.e., if $\rho_{ks} > \beta_{kt}$). Finally, constraints (89)-(92) impose necessary the restrictions on the decision variables.

7.3 Linearization of the objective function

The second term of the objective function $d_k \alpha_{kst} y^o_{kst}$ is non-linear. To linearize it, we introduce auxiliary variables γ^o_{kst} as well as a large constant M'.
Objective function (84) can then be replaced by:

$$\min \sum_{s \in S} \pi_s \sum_{t=1}^{T} \sum_{k \in A} \sum_{o \in N} (d_k y^o_{kst} + \Delta d_k \gamma^o_{kst}) \tag{93}$$

subject to (85)-(92) and the following.

$$\gamma^o_{kst} \geq y^o_{kst} + M'(\alpha_{kst} - 1) \quad \forall o \in N, \forall k \in A, \forall s \in S, t = 1, ..., T \tag{94}$$

$$\gamma^o_{kst} \geq 0 \quad \forall o \in N, \quad \forall k \in A, \forall s \in S, t = 1, ..., T \tag{95}$$

Constraints (94) are introduced so that $\gamma^o_{kst} = y^o_{kst}$ whenever $\alpha_{kst} = 1$, 0 otherwise. Constant M' needs to be larger than the total number of origin-destination couples. Constraints (95) are the non-negativity requirements for the γ^o_{kst} variables.

7.4 Linearization of the protection level constraints

Constraint (87) can be linearized by introducing additional decision variables $x_{kt\ell}$ defined as:

$$x_{kt\ell} = \begin{cases} 1 & \text{if the highest protection standard afforded to arc } k \text{ in period } t \\ & \text{is sufficient to withstand a flood of size } \ell \\ 0 & \text{otherwise} \end{cases}$$

Further, let $P_{k\ell}$ be the subset of plans in P that would protect arc k to a standard sufficient to withstand a flood of size ℓ. With this in place, equalities (87) can be replaced with the following set of linear constraints.

$$\beta_{kt} = \sum_{\ell \in L} r_\ell x_{kt\ell} \quad \forall k \in A, t = 1, ..., T \tag{96}$$

$$\sum_{\ell \in L} x_{kt\ell} \leq 1 \quad \forall k \in A, t = 1, ..., T \tag{97}$$

$$x_{kt\ell} \leq \sum_{u=1}^{t} \sum_{p \in P_{k\ell}} z_{pu} \quad \forall k \in A, \forall \ell \in L, t = 1, ..., T \tag{98}$$

$$x_{kt\ell} \in \{0, 1\} \quad \forall k \in A, t = 1, ..., T \tag{99}$$

Constraints (96) are used to set β_{kt} to the highest protection standard implemented for arc k by period t. Inequalities (97) stipulate arc k can be protected to at most one safety standard ℓ in period t. Constraints (98) state that arc k can be protected to a safety standard of ℓ in period t if a plan $p \in P_{k\ell}$ has been implemented in the time window $1, ..., t$. Constraints (99) are the binary requirements for variables $x_{kt\ell}$.

7.5 Solution approach

The proposed model can be solved directly with a general purpose mixed integer linear programming (MILP) solver like CPLEX. The formulation, however, requires a very large number of variables and constraints,even when moderate numbers of scenarios, protection plans, and time periods are considered. This has a strong impact on the performance of the MILP solver, both in terms of computing time and memory requirements. Furthermore, shortest-path problems can be solved very efficiently using specialized algorithms. These considerations suggest that more effective solution approaches can be devised. Here, we present a heuristic algorithm based on the GRASP meta-heuristic for building an initial solution combined with a local search to further explore the feasible space.

7.5.1 GRASP step

The GRASP (Greedy Randomised Adaptive Search Procedure) meta-heuristic was introduced by Feo and Resende (1995) to overcome the limits of purely greedy construction algorithms. It is an iterative procedure that at each step builds up a solution by randomly choosing elements from a dynamically constructed restricted candidate list (RCL). The RCL consists of a subset of elements which are selected according to a greedy scheme. Once GRASP has produced a starting solution, a local search can be carried out to find an improved solution. The entire process can be repeated multiple times in an effort to more fully explore the solution space.

The following is the notation used in our GRASP implementation:

- *itr* is the iteration index

- $plans_{itr}$ and obj_{itr} are the solution and the corresponding objective value obtained at iteration itr
- $bestPlans$ and $bestObj$ refer to the best solution found and the corresponding objective value
- $MAXITER$ is maximum the number of iterations
- $RCLSIZE$ is the size of the restricted candidate list
- Φ_{pt} is a greedy metric used to rank the pair (p,t) (i.e., the importance of plan p implemented at time t)
- $S_{k\ell}$ is the set of scenarios that disrupt arc k at level ℓ
- Δ_s is the increment in the all-pairs shortest path expected cost generated by disruption scenario s

We further introduce some crucial subroutines used in the algorithm:

- $computeShortestPath(plans_{itr})$ finds a solution to the all-pairs shortest path problem, subject to the disruption scenarios and with protection strategy $plans_{itr}$ as input. The function is implemented efficiently using the Floyd-Warshall algorithm (Floyd, 1962).
- $updateImportance(plans_{it})$ updates the metric Φ_{pt} given the $plans_{it}$ as input. Before defining Φ_{pt} mathematically, we need to introduce an additional parameter $\lambda_{k\ell}$ for each arc and flood size pair (k,ℓ). Specifically, let $\lambda_{k\ell} = \sum_{\ell' \leq \ell} \sum_{s \in S_{k\ell}} \Delta_s$. Parameter $\lambda_{k\ell}$ which gives the cumulative increment in the shortest path expected travel cost generated by all scenarios

disrupting arc k at level $\ell' \leq \ell$. For the sake of clarity, we define $fit = computeShortestPath(plans_{itr})$ and $fit(p,t) = computeShortestPath(plans_{itr} \cup (p,t))$. The difference between these two values $fit - fit(p,t)$ represents the net decrease in overall travel costs obtained when adding (p,t) to the solution. The importance metric is then computed by the following formula:

$$\Phi_{pt} = \frac{\sum_{k \in \hat{A}_p} \lambda_{k\sigma_{kp}}(fit - fit(p,t))}{c_p} \tag{100}$$

Looking at (100), importance is defined as the benefit obtained by including plan p at time t to the set of implemented protections weighted by parameter $\lambda_{k\sigma_{kp}}$ and the inverse of the cost c_p of implementing that plan.

- $buildRCL(plans_{itr})$ creates the restricted candidate list of size $RCLSIZE$ by selecting the best pairs (p,t) according to their Φ_{pt} values.
- $updateFeasibePlans(plans_{itr})$ is used to keep track of plans that can be added to $plans_{itr}$ without violating the budget constraints.

In Algorithm 11, we provide pseudo-code of the GRASP heuristic.

7.5.2 Improvements to the GRASP algorithm

Here, we briefly discuss some expedients that have been adopted to improve the GRASP algorithm both in terms of efficiency and accuracy. The first improvement takes advantage of multi-core processors by implementing a multi-thread version of the algorithm. Using concurrent programming allows for a more in-depth exploration of the solution space without an increase in the execution time. Furthermore, threads can work with different algorithm settings to implement

Algorithm 11 GRASP pseudo-code.

```
$bestPlans \leftarrow \emptyset$; $bestObj = \infty$; $itr = 1$
while $itr \leq MAXITER$ do
  $plans_{itr} = \emptyset$; $obj_{itr} = \infty$
  $updateImportance(plans_{itr})$
  $updateFeasibePlans(plans_{itr})$
  while there are feasible plans do
    $buildRCL(plans_{itr})$
    randomly select a plan $\left(\hat{p}, \hat{t}\right)$ from RCL
    $plans_{itr} = plans_{itr} \cup \left(\hat{p}, \hat{t}\right)$
    $updateImportance(plans_{itr})$
    $updateFeasibePlans(plans_{itr})$
  end while
  $obj_{itr} = computeShortestPath(plans_{itr})$
  if $obj_{itr} < bestObj$ then
    $bestPlans \leftarrow plans_{itr}$; $bestObj = obj_{itr}$
  end if
  $itr = itr + 1$
end while
return $bestPlans$; $bestObj$
```

different policies so that the probability of exploring the same solution is reduced. To provide this flexibility, we extend the definition of Φ_{pt} as follows:

$$\Phi_{pt} = \frac{\mu(p) \sum_{k \in M} b_{kp} \lambda_{k\sigma_p}(fit - fit(p,t))}{r_p} \tag{101}$$

The function $\mu(p)$ is used to weight each plan according to the protection level offered. Specifying different functional forms of μ for each thread will result in different solutions explored, thus increasing the chance of finding an optimal solution. For example, $\mu(p) = 1$ weights all plans p equally. Alternatively, $\mu(p) = \ell_p^{max}$, where ℓ_p^{max} is the index number of the maximum protection standard afforded to any arc in plan p (i.e., $\ell_p^{max} = \arg\max_{k \in \hat{A}_p} \sigma_{kp}$), puts higher weight on plans offering potentially higher levels of protection. This, in turn, increases the probability of having such plans in the solutions.

The algorithm can be further improved by optimizing the routine $updateImportance()$. The idea behind the improvement is as follows. If there are available protection resources at time t, the most important plans (the ones with higher Φ) will more likely be the plans (p, t') with $t' \in [t, t+1]$. As a consequence, we can reduce the number of plans for which Φ needs to be recomputed by excluding the plans (p, t') such that $t' \in [t+2, T]$.

Finally, if we keep track of the highest possible level of disruption threatening each arc, we can ignore any plans that over protect. Formally, a plan p is excluded from consideration if:

$$\begin{aligned}\sigma_p > \max_{s \in S} \{\rho_{ks}\} \ \forall k \in \hat{A}_p \quad \wedge \quad \exists p' \text{ s.t. } c'_p < c_p, \hat{A}'_p \subseteq \hat{A}_p, \\ \sigma_{kp} \geq \sigma_{kp'} \geq \max_{s \in S} \{\rho_{ks}\} \ \forall k \in \hat{A}_{p'}\end{aligned} \tag{102}$$

7.5.3 Local search step

The best solution found by the GRASP procedure is used as the starting point for a local search. The temporal component of the model makes the implementation of an effective local search challenging. In fact, a search procedure should take into account the fact that the GRASP solution generated might include plans that are not included in the optimal solution and/or plans that are in the optimal solution but should be implemented at different time periods. To tackle this issue, we implemented two types of swap moves: internal and external. For each type, two approaches were devised. The four different swaps are carried out sequentially in the same order as they are introduced in the next sections.

Internal swaps can perform single or multiple swaps between the elements of a solution. Specifically, these swaps change only the time when plans are implemented, not the type of plans chosen. Swaps are done only among plans implemented in consecutive time periods. We consider two combinations of swaps:

- 1-to-2 swap: A plan (p, t) with $t \in [1, T-1]$ is postponed one period. Simultaneously, up to two plans implemented at time $t+1$ are brought forward one period. Every combination involving 1 or 2 plans is considered. Non-feasible solutions are discarded. The pseudo-code of this approach is shown in Algorithm 12.

- 2-to-3 swap: This approach is almost identical to the previous one. The only difference is the number of plans that are postponed (2) and those that are brought forward (up to 3).

Table 30 shows an example of 1-to-2 internal swap, where the plans in bold are those being swapped. The code is implemented to avoid unnecessary computation. For instance, assuming that the swap $p_2 \leftrightarrow (p_3, p_5)$ generates a feasible solution, then swaps $p_2 \leftrightarrow p_3$ and $p_2 \leftrightarrow p_5$ are not considered.

Before				After			
t	Plans			t	Plans		
0	p_1	p_2		0	p_1	$\boldsymbol{p_3}$	$\boldsymbol{p_5}$
1	p_3	p_4	p_5	1	$\boldsymbol{p_2}$	p_4	
2	p_6	p_7		2	p_6	p_7	

Table 30: Example of 1-to-2 internal swap.

Algorithm 12 Internal 1-to-2 swap

```
$plans \leftarrow bestPlans$
for $t = 1; t < T; t++$ do
   for $(p, t) \in plans^t$ do
      $plans \leftarrow plans \setminus (p, t) \cup (p, t + 1)$
      for $(p', t')^1 \in plans^{t+1}$ do
         $plans \leftarrow plans \setminus (p', t')^1 \cup (p', t' - 1)^1$
         if $plans$ is a feasible solution then
            $obj = computeShortestPath(plans)$
            if $obj < bestObj$ then
               $bestPlans \leftarrow plans$; $bestObj = obj$
            end if
            for $(p', t')^2 \in plans^{t+1} \wedge (p', t')^2 \neq (p', t')^1$ do
               $plans \leftarrow plans \setminus (p', t')^2 \cup (p', t' - 1)^2$
               if $plans$ is a feasible solution then
                  $obj = computeShortestPath(plans)$
                  if $obj < bestObj$ then
                     $bestPlans \leftarrow plans$; $bestObj = obj$
                  end if
               end if
               $plans \leftarrow plans \setminus (p', t' - 1)^2 \cup (p', t')^2$
            end for
         end if
         $plans \leftarrow plans \setminus (p', t' - 1)^1 \cup (p', t')^1$
      end for
      $plans \leftarrow plans \setminus (p, t + 1) \cup (p, t)$
   end for
end for
```

External swaps remove one or more plans from the best solution and replace them with one or more plans not currently included in the solution. As with internal swaps, we consider two cases:

– 1-to-2 swap: A plan (p, t) is removed from the solution. Simultaneously, up to two plans are added in its place. Only plans that would generate a feasible solution are considered. The pseudo-code of this approach is shown in Algorithm 13.

- 2-to-3 swap: This approach is nearly identical to the previous one. The only difference is the number of plans swapped out (2) and swapped in (up to 3).

Algorithm 13 External 1-to-2 swap

```
$plans \leftarrow bestPlans$
for $(p, t) \in plans$ do
    $plans \leftarrow plans \setminus (p, t)$
    $updateFeasibePlans(plans)$
    for $(p', t')^1 \in FeasiblePlans$ do
        $plans \leftarrow plans \cup (p', t')^1$
        $updateFeasibePlans(plans)$
        if $FeasiblePlans \setminus (p, t) \neq \emptyset$ then
            for $(p', t')^2 \in FeasiblePlans$ do
                $plans \leftarrow plans \cup (p', t')^2$
                $obj = computeShortestPath(plans)$
                if $obj < bestObj$ then
                    $bestPlans \leftarrow plans$; $bestObj = obj$
                end if
                $plans \leftarrow plans \setminus (p', t')^2$
            end for
        else
            $obj = computeShortestPath(plans)$
            if $obj < bestObj$ then
                $bestPlans \leftarrow plans$; $bestObj = obj$
            end if
        end if
        $plans \leftarrow plans \setminus (p', t')^1$
    end for
    $plans \leftarrow plans \cup (p, t)$
end for
```

Table 31 shows an example of 1-to-2 external swap with swapped plans shown in bold.

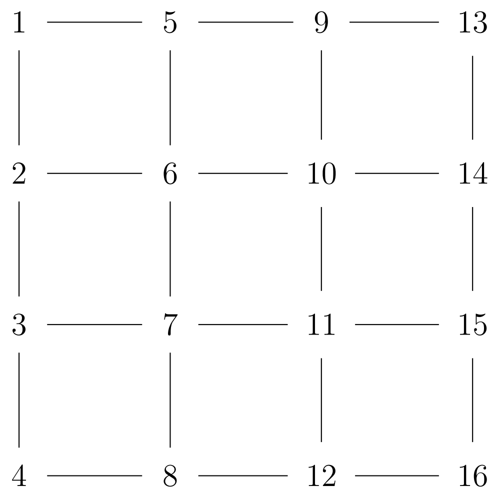


Figure 29: An example 16 × 24 grid.

Before							After					
t	Internal plans			External plans			t	Internal plans			External plans	
0	p_1	p_2		p_8	p_9	p_{10}	0	p_1	p_2		$\boldsymbol{p_7}$	p_8
1	p_3	p_4	p_5	p_8	p_9	p_{10}	1	p_3	p_4	p_5	$\boldsymbol{p_7}$	p_8
2	p_6	p_7		p_8	p_9	p_{10}	2	p_6	$\boldsymbol{p_9}$	$\boldsymbol{p_{10}}$	$\boldsymbol{p_7}$	p_8

Table 31: Example of 1-to-2 external swap.

7.6 Computational tests

For our initial computational tests, we generated non-directed square networks of size $n \times m$, where n is the number of nodes and m is the number of arcs. We work on non directed-networks because we assume that the travel time of a link is independent of the travel direction and that protections and disruptions affect simultaneously and equally both directions. In Figure 29, we show the topology of an example 16×24 grid. Arcs travel times d_k were drawn uniformly in the range $[1, 10]$.

We consider 4 different flood sizes categorized in terms of their return period:

20, 50, 100, and 250 years. The same values are used to identify the levels of protection. Therefore, $R = \{20, 50, 100, 250\}$. According to the disruption level of each scenario, the delay Δd_{ks} is set to $2.5, 5, 10$, and 20 times the baseline travel time d_k. We examined three types of protection plans: arc, node and, row/column plans. Arc plans protect a single arc. Node plans protect all the arcs incident to a node. Finally, row/column plans include all the arcs forming a row or a column in the grid. For example, if we consider the grid in Figure 29, the node plan corresponding to node 6 will include arcs 2-6, 5-6, 6-7 and 6-10. The row plan corresponding to the second row will include arcs 2-6, 6-10 and 10-14. Overall, we can identify $m + 3n$ different plans using this scheme. Each plan can be implemented at 4 levels of protection. Consequently, for each instance we use $4(m + 3n)$ protection plans. In Table 32, we show the criteria used to estimate the costs of implementing the protection plans.

The discount factor f was set to 1.3. It is incorporated so that adopting a node or row-column plan p will be cheaper than independently protecting all the arcs included in p.

We assume that all arcs face the same level of disruption flooding for any given scenario (i.e., $\rho_{ks} = \rho_s$, $\forall k \in A_s$) and that the number of arcs disrupted in each scenario is dependent on the scenario's return period. The least damaging scenario is capable of disrupting only 1 arc. The number of disrupted arcs increases with

	Return period (yrs)			
Type of plan	20	50	100	250
Arc	d_k	$2d_k$	$3d_k$	$4d_k$
Node	$[f^{-1}\sum_{k\in A} b_{kp}c_k]$	$[2f^{-1}\sum_{k\in A} b_{kp}c_k]$	$[3f^{-1}\sum_{k\in A} b_{kp}c_k]$	$[4f^{-1}\sum_{k\in A} b_{kp}c_k]$
Row/Column	$[f^{-1}\sum_{k\in A} b_{kp}c_k]$	$[2f^{-1}\sum_{k\in A} b_{kp}c_k]$	$[3f^{-1}\sum_{k\in A} b_{kp}c_k]$	$[4f^{-1}\sum_{k\in A} b_{kp}c_k]$

Table 32: Cost of implementing a given type of plan according to its protection level.

the scenario's size from 1 to $m/4$ (i.e., the most severe disruption scenario can disrupt 25% of the network). In Table 33, we list the number of disrupted arcs for each of the return periods considered.

	Return period (yrs)			
	20	50	100	250
No. disrupted arcs	1	$\lceil (m+20)/24 \rceil$	$\lceil (m+5)/9 \rceil$	$\lceil m/4 \rceil$

Table 33: Number of disrupted arcs based on a scenario's return period.

Scenarios were built in such a way that arcs affected by a given scenario formed a connected sub-network. This was done so that the elements affected by a flood event were contiguous, as is often the case in real life. Finally, we report in Table 34 the number of scenarios generated for each level of flooding. The probability of occurrence for a scenario was set as the inverse of the scenario's return period. We generated 30 instances using networks of size 9 × 12, 16 × 24, and 25 × 40.

	Return period (yrs)			
	20	50	100	250
No. scenarios	m	$\lceil m/2 \rceil$	$\lceil m/4 \rceil$	$\lceil m/8 \rceil$

Table 34: Number of scenarios generated for each return period.

We considered a 4-periods planning horizon (i.e., $T = 4$). We chose a total protection budget equal to 10% of the resources needed to protect the entire network at the highest safety level. The budget was equally spread over the time periods. We ran the heuristic with the number of threads set to 4. Each thread uses a different μ vector to compute Φ based on equation (101). The values for the different policies used in each thread are listed in Table 35.

	Return period (yrs)			
Thread ID	20	50	100	250
0	1	1	1	1
1	1	2	3	4
2	1	1.3	1.6	1.9
3	1	2	2	1

Table 35: Set of μ values used for each thread.

Thread 0 treats all the protection plans equally. Threads 1 and 2 give more importance to plans that grant higher levels of protection. Finally, thread 4 weights more heavily medium level protection plans. These values were chosen empirically to maximize the probability of discovering an optimal solution. The size of the RCL is dependent on the problem size. It was set to 3 for 9×12 and 16×24 networks and to 4 for 25×40 networks. The number of iterations was set to 100 and we specified a 6-hour limit on the run time of the algorithm.

In Table 36, we compare the performance of CPLEX against our GRASP heuristic on the 30 test problem instances. The "Obj" and "Time" columns show the objective value and the solution time obtained by the two algorithms. "Gap" is the percentage gap between the best feasible solution obtained with CPLEX and the value returned by the heuristic. A negative value for "Gap" indicates that the heuristic found a better solution. Finally, "Time diff" is the percentage difference between the computing time of the two algorithms. A negative value for "Time diff" indicates that the heuristic was faster than CPLEX. The table shows that CPLEX can find a verified optimal solution within 6 hours only for small and mediums sized instances. In fact, on the 25×40 networks, CPLEX never converges to an optimal solution. The heuristic, on the other hand, finds optimal or near-optimal solutions much more efficiently. It is on average at least 82% faster than CPLEX and always returns better solutions for the 25×40 networks.

		CPLEX		Heuristic			
Network size	Instance	Obj	Time(s)	Obj	Time(s)	Gap(%)	Time diff(%)
9 × 12	1	2168.6	7	2168.6	2.5	0.00%	-64.29%
	2	2595.5	6.9	2595.5	3.3	0.00%	-52.17%
	3	2282.4	9.7	2282.4	3.8	0.00%	-60.82%
	4	2831.9	6.5	2831.9	2.9	0.00%	-55.38%
	5	2476.4	4.7	2476.4	3.1	0.00%	-34.04%
	6	3107.9	6.6	3107.9	3.3	0.00%	-50.00%
	7	2360.5	8.8	2360.5	3.8	0.00%	-56.82%
	8	2600.1	6.9	2600.1	3.6	0.00%	-47.83%
	9	2505.0	5.6	2505.0	3.3	0.00%	-41.07%
	10	1755.5	8.9	1756.1	2.7	0.03%	-69.66%
	Avg		**7.2**		**3.2**	**0.00%**	**-53.21%**
16 × 24	1	17020.9	1900.1	17020.9	176.9	0.00%	-90.69%
	2	19487.3	3581.8	19487.3	164.7	0.00%	-95.40%
	3	21347.0	1714.5	21347.0	131.6	0.00%	-92.32%
	4	16744.5	1553.9	16747.5	166.6	0.02%	-89.28%
	5	18646.8	1354.2	18646.8	129.4	0.00%	-90.44%
	6	16661.8	4395.2	16662.7	187.8	0.01%	-95.73%
	7	15346.2	4322.8	15349.1	135.1	0.02%	-96.87%
	8	14288.9	387.1	14288.9	207.4	0.00%	-46.42%
	9	18783.4	2177.7	18783.4	192.2	0.00%	-91.17%
	10	12225.7	1971.92	12225.7	180.9	0.00%	-90.83%
	Avg		**2335.9**		**167.3**	**0.00%**	**-87.92%**
25 × 40	1	106278.9†	21600.0	106064.0	3183.7	-0.20%	-85.26%
	2	88124.5†	21600.0	87981.4	3665.6	-0.16%	-83.03%
	3	97700.2†	21600.0	97684.1	3676.1	-0.02%	-82.98%
	4	91102.4†	21600.0	91025.2	3448.5	-0.08%	-84.03%
	5	109276.9†	21600.0	109091.0	3876.9	-0.17%	-82.05%
	6	80102.2†	21600.0	79972.0	4384.8	-0.16%	-79.70%
	7	88891.3†	21600.0	88729.7	4102.1	-0.18%	-81.01%
	8	97849.8†	21600.0	97818.4	4315.6	-0.03%	-80.02%
	9	115763.9†	21600.0	115681.0	3937.2	-0.07%	-81.77%
	10	83557.7†	21600.0	83530.9	3988.9	-0.03%	-81.53%
	Avg		**21600.0**		**3857.9**	**-0.11%**	**-82.14%**

† Best objective value found after 6 hours of running time.

Table 36: Computational results for test problem instances.

7.7 Case study

In this section, we present a case study based on the A-road network infrastructure of Hertfordshire in the UK. The network is composed of 36 nodes and 47 undirected arcs. We used historical data on floods, which are publicly available through the Environmental Agency (EA). Using geographic information system

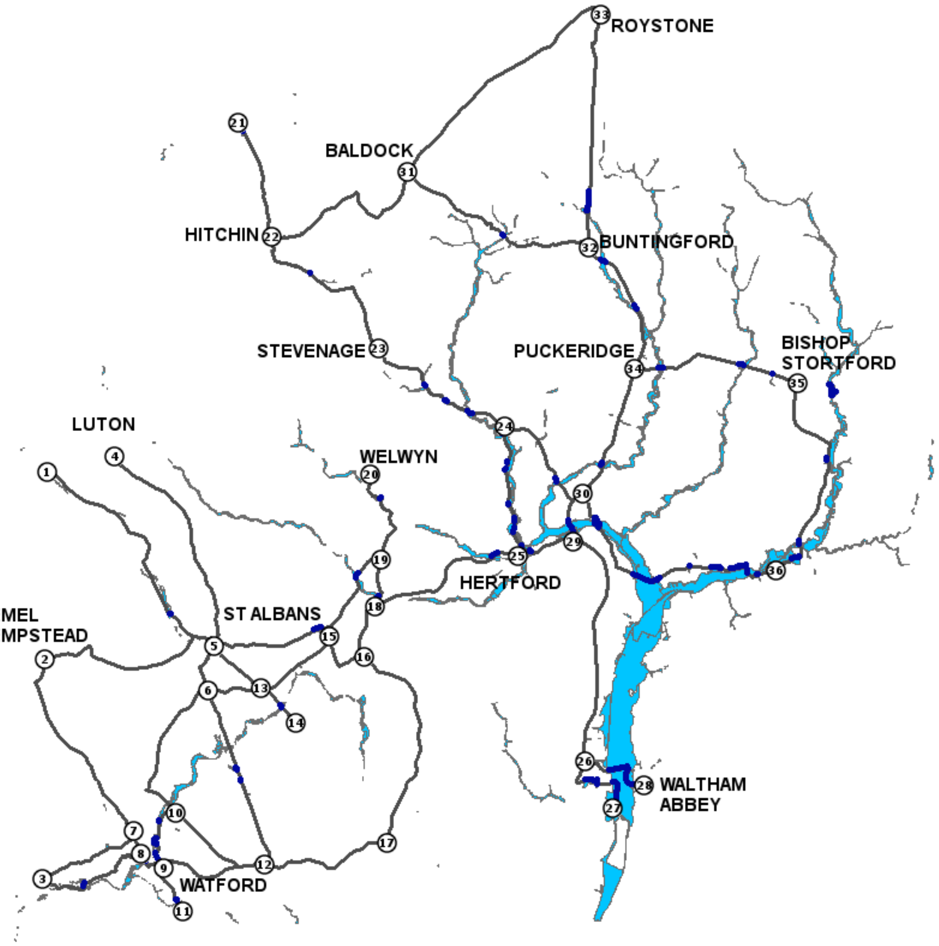


Figure 30: Hertfordshire A road network.

(GIS) software (ArcGIS), we generated disruption scenarios by identifying regions where floods and the road network overlapped. Over the past 70 years, 29 floods in Hertfordshire are recorded in the EA database. Among them, 16 overlapped with the A-road network.

In Figure 30, we show the road network considered in our case study. Recorded floods are shown in light blue; in dark blue are shown events that affected the A-road network. Unfortunately, floods are not ranked in the database. Consequently, we rank them into three size categories (small, medium, and large) according to the extent of the area flooded. The approach used is explained in Table 37. Coming up with precise estimates for the return periods and, consequently, the likelihood for each flood category requires a detailed hydrogeological analysis that goes beyond the scope of this chapter. For simplicity, we use the number of events that occurred over the past 70 years to estimate their return period (see Table 37).

	Area flooded (km^2)		
	$(0, 2.5]$	$(2.5, 10]$	$(10, 50]$
Category	Small	Medium	Large
No. observed events	13	2	1
Return period (yrs)	5	35	70

Table 37: Return period estimation.

We consider all single arc protection plans. To this, we add some multiple arc plans based on a proximity criteria. The costs of implementing the protection plans are again dependent on arc lengths in the same way as explained in the computational results for the test instances. The protection budget is computed as a percentage of the resources needed to implement the highest level of protection for the entire network (equivalent to 9,825 protection units).

Table 38 reports how the shortest path cost decreases when the protection budget is increased. The first two increments (from 0% to 2% and from 2% to 4%) are the most significant. Investing a budget of 10% results in a 26% reduction in the expected shortest path cost over the planning horizon.

	Budget					
	0%	2%	4%	6%	8%	10%
ΔObj	0%	8%	16%	21%	23%	26%

Table 38: Change in objective function value (ΔObj) versus budget.

In Figure 31, we show the optimal protection plans when the protection budget is equal to 10%. The map shows, along with the protected links, the protection level of each plan. The majority of protections are designed to thwart small floods. No link is fortified at the highest possible level of protection. The results indicate that the central area of the network (close to the town of Hertford) is the most critical. Interestingly, arc 29-30 is initially protected at a small level in the first period and subsequently upgraded to a medium level in the last period.

Next, we compare the results of our base model with those obtained by adopting the p-robustness criteria introduced by Snyder and Daskin (2006). The p-robustness measure is combined with our model by adding the following set of constraints:

$$\sum_{t=1}^{T}\sum_{k\in A}\sum_{o\in N}(d_k y^o_{kst} + \Delta d_k \gamma^o_{kst}) \leq (1+p)D^*_s \quad \forall s \in S \tag{103}$$

where D^*_s is the objective value obtained by solving our base model with only scenario s as input.

The results of this analysis are summarized in Figure 32. The values of p are obtained following the same approach used by Snyder and Daskin (2006). We solve the problem with p = ∞ and then set p to the maximum regret minus 0.00001. Subsequently, p is decremented by 0.00001 until the problem becomes infeasible. Figure 32 shows that improvements in robustness are not justified

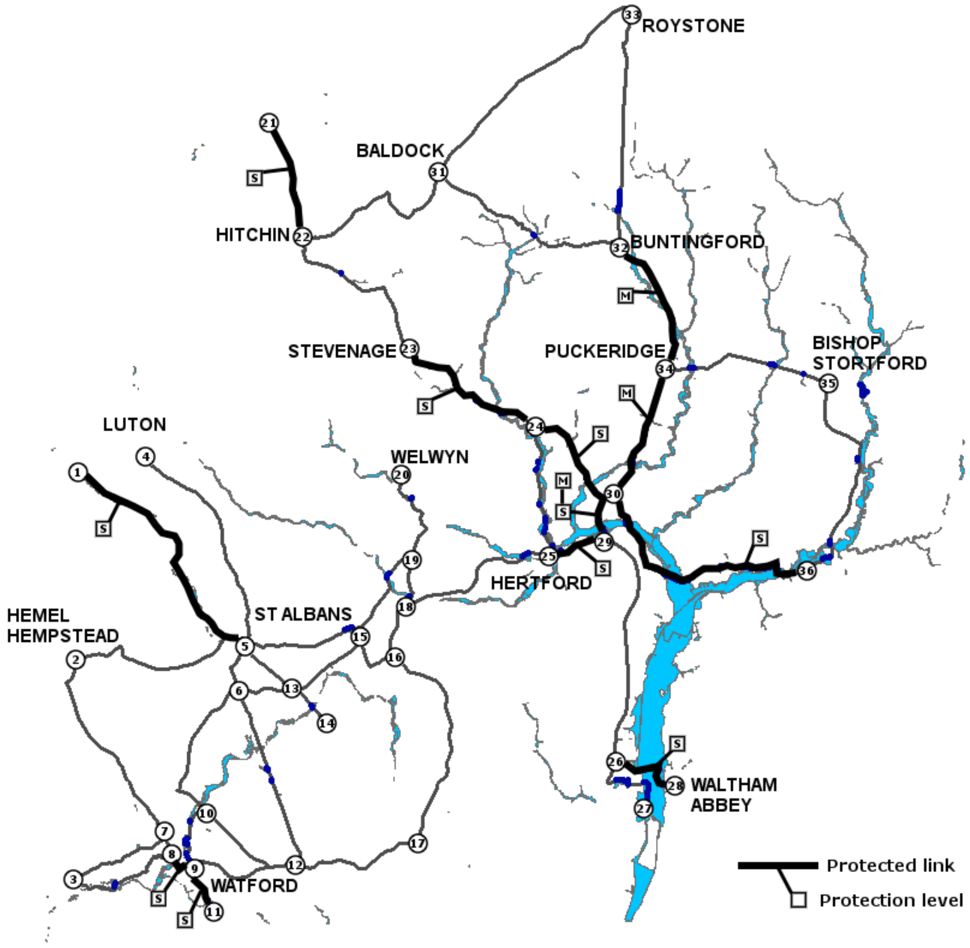


Figure 31: Optimal protections with the protection budget equal to 10%.

by the costs. At best, a 0.00018 (0.23%) decrease in relative regret (p) can be achieved for a 7.5% increase in expected travel cost when the protection budget is 5%. Using higher budgets and or different values of p for different scenarios did not have any appreciable impact on robustness, therefore results have not been included.

We also carried out a sentivity analysis to understand how uncertainty in the estimation of return periods might impact optimal protection strategies. Specifically, we allowed the return period to vary by ±20% from the base case for small,

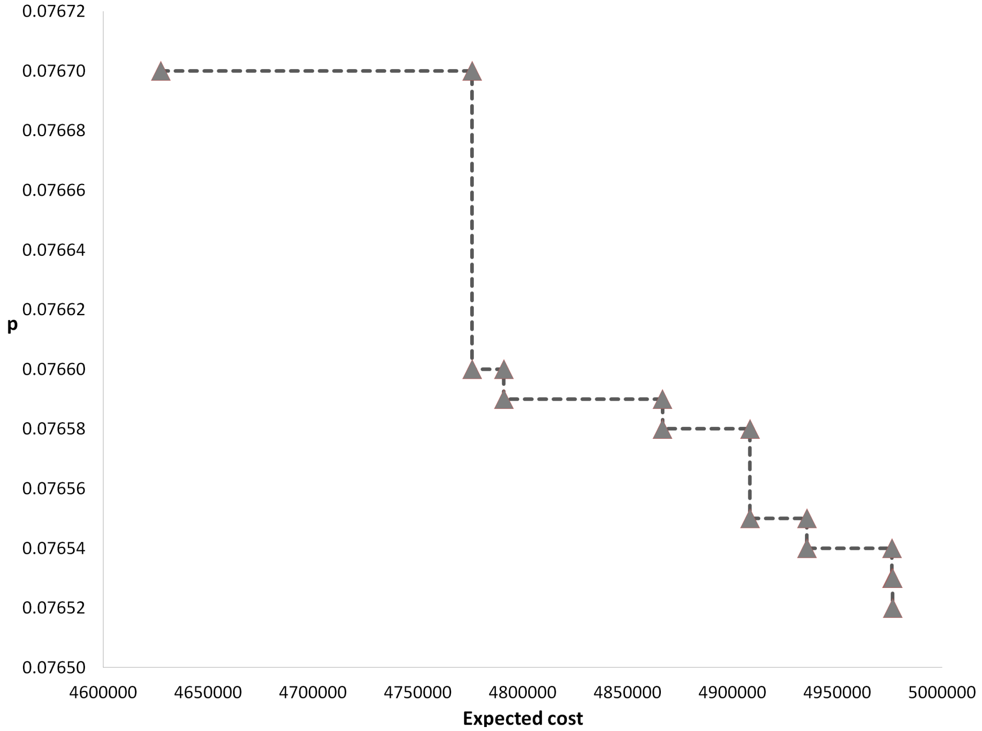


Figure 32: Maximum regret vs expected travel cost when the protection budget is equal to 5%.

	Return period (yrs)				Return period (yrs)				Return period (yrs)		
Case	Small	Medium	Large	Case	Small	Medium	Large	Case	Small	Medium	Large
1	4	28	56	**10**	5	28	56	**19**	6	28	56
2	4	28	70	**11**	5	28	70	**20**	6	28	70
3	4	28	84	**12**	5	28	84	**21**	6	28	84
4	4	35	56	**13**	5	35	56	**22**	6	35	56
5	4	35	70	**14**	5	35	70	**23**	6	35	70
6	4	35	84	**15**	5	35	84	**24**	6	35	84
7	4	42	56	**16**	5	42	56	**25**	6	42	56
8	4	42	70	**17**	5	42	70	**26**	6	42	70
9	4	42	84	**18**	5	42	84	**27**	6	42	84

Table 39: List of different return period value permutations for small, medium, large floods.

medium, and large floods, resulting in a total of 27 cases as shown in Table 39. The budget used in this analysis was set to 10%.

Figure 33 shows the net percentage change in expected travel cost relative to a

0% protection budget. The figure clearly highlights that it is the return period of small floods that has the strongest impact on the objective value. In particular, the different flood scenarios clearly separate into three sub-categories (cases 1-9, cases 10-18, and cases 19-27) based entirely on the return period for small floods. Within these three sub-categories, only relatively small changes in the objective value are observed depending on the return periods for medium and large floods. For example, increasing the return period of small floods from just 4 to 5 years (the equivalent of going from a 0.25 to a 0.2 chance of occuring in any given year), causes net expected travel cost to drop by more than 2-fold, with values in the range $[-9\%, -12\%]$ versus $[-25\%, -28\%]$, respectively. What this suggests is that errors in return periods estimates for medium and large floods are of less concern compared to small floods.

Finally, we evaluate the robustness of optimal protection strategies to uncertainty in flood return periods. Table 40 reports the percentage increase in the

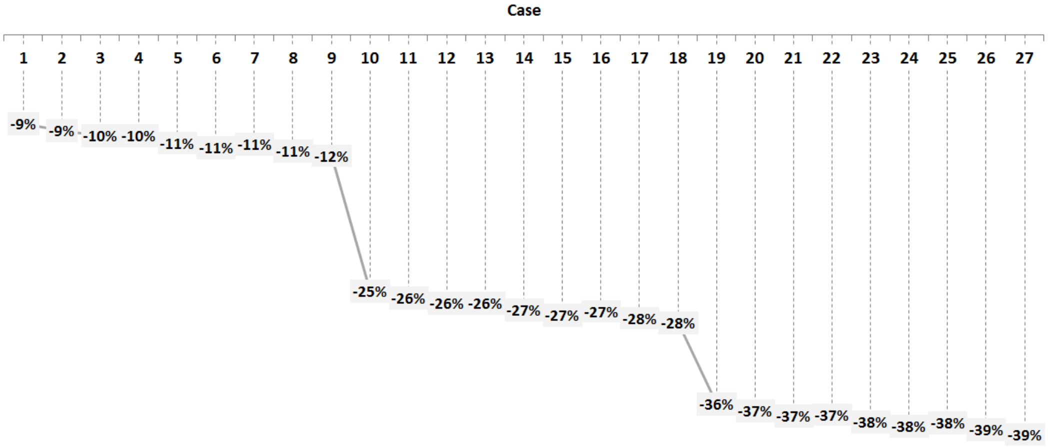


Figure 33: Net percentage change in expected travel cost give a protecton budget of 10%.

objective function for the base case solution (i.e., case 14) across all 27 permutations for small, medium, and large flood return period values. The table indicates that base case optimal solution is highly robust even when the estimation error for return periods is as high as 20%. Indeed, the base case solution produced no increase in expected travel cost for 21 out the 27 cases. In the remaining 6 cases, the increase in cost was marginal (i.e., under 0.3%).

Case	Inc (%)	Case	Inc (%)	Case	Inc (%)
1	0.00%	**10**	0.00%	**19**	0.17%
2	0.00%	**11**	0.00%	**20**	0.24%
3	0.00%	**12**	0.04%	**21**	0.29%
4	0.00%	**13**	0.00%	**22**	0.00%
5	0.00%	**14**	0.00%	**23**	0.00%
6	0.00%	**15**	0.00%	**24**	0.00%
7	0.00%	**16**	0.00%	**25**	0.00%
8	0.01%	**17**	0.00%	**26**	0.00%
9	0.03%	**18**	0.00%	**27**	0.00%
Avg	0.01%		0.00%		0.08%

Table 40: Robustness of the base case optimal solution to flood return period uncertainty.

7.8 Conclusions

In this chapter, we introduce a scenario-based model to identify the optimal set of plans to adopt for protecting a road infrastructure against flooding. When dealing with flood protection, a large number of measures can be considered. These measures may differ significantly in terms of cost and protection standard guaranteed. We incorporate this issue by considering different safety levels and using protection plans involving potentially many arcs as opposed to single arc fortifications. A multi-period planning horizon is adopted to take into account the possibility that protection resources are distributed over time. The consequences of disruptions are estimated in terms of the expected all-pairs shortest path travel cost.

We find that solving the problem with a general purpose solver is challenging. A GRASP heuristic is, therefore, introduced. Tests on randomly generated networks of different size demonstrate the efficiency of the heuristic. Finally, a case study of the Hertfordshire A-road network provides some useful insights.

8 Conclusions

8.1 Summary

In this thesis, new models to protect transportation infrastructure against disruption are studied. The first model considers a fortification framework type for optimally allocating protection resources to a railway network subject to worst-case disruptions. This model is extended in two different ways.

Firstly, a temporal component is introduced. This is done to consider the more realistic case of having protection resources allocated over a planning horizon. Two decomposition approaches are presented to solve this bi-level model.

Secondly, user-behaviour is modelled in a more accurate way. Specifically, it is assumed that post disruption delays directly affect the number of passengers using the rail system. To solve the model, an exact and a heuristic approach are introduced and compared.

Finally, the last problem focuses on protecting a road-network against random flooding. In contrast with the models proposed in the previous chapters, for this problem we use a scenario-based optimization model so as to identify robust solutions across a range of disruption scenarios. A heuristic solution approach is proposed to overcome the limitations of general purpose solvers.

8.2 Suggestions

While the study of critical infrastructure protection has increased significantly in the last decade, there are still many open questions.

The literature review shows that the great majority of works has focused on generic problems (e.g., protecting a network against non specified disruptions). Although these works have been fundamental in building a base knowledge on how to model and solve critical infrastructure protection problem, now research must narrow the focus to specific problems driven by real life data.

The problems analysed in this thesis are significantly complex. A comprehensive study analysis of counter-terrorism plans cannot ignore the geographical, sociological and historical issues involved. Similarly, protecting against natural catastrophes involves multiple topics (forecasting, structural engineering etc.). To this aim, an interesting extension of the work proposed here would be a multidisciplinary research project focusing on flood protection. The project will aim to support decision making both at a strategic and operational level. Optimization models can be integrated with accurate weather forecasting (driven by historical data) and structural engineering analysis to assess the impact of flooding on both infrastructure's assets and defence structures. This multi-disciplinary knowledge can be used in two ways:

- To develop cost-efficient and accurate long-term protection plan to decide what defence measures should be implemented and where.
- To support a live decision system capable of predicting, to some extent, the occurring of a flooding and suggesting a set of mitigation strategies, such as traffic re-routing, evacuation, distribution of emergency supplies etc.

The assumption that disruptor and system protector in fortification problem share same information and objective function although common can be considered as a strong one. Very few attempts (Bayrak and Bailey (2008); Jenelius et al. (2010)) have been made to stir away from it. Nonetheless, these attempts simply resulted in models where actors estimate a few parameters differently. A more interesting line of research would be to study multi-objective models where actors have asymmetric knowledge of the system and evaluate their benefits using different functions.

More efforts can be made into studying dynamic aspects of protection problems. The challenge is that a protection problem is generally characterized by several dynamic components which require significantly different time units. For instance, protection planning is usually driven by spending review spread over years, whereas assessing the impact of disruptions may require to evaluate the system over a hour or even a minute scale. Harmonizing these different time components in comprehensive models is a difficult yet interesting research direction.

A more general research direction to undertake focuses on relaxing as many as possible of the following strong assumptions commonly made in the literature:

- Binary interdictions
- Binary fortifications
- Binary user decisions
- Perfect information
- Same objective function for attacker and defender

For example, when dealing with earthquakes and flooding, protection measures are generally characterized by fragility curves (Hall et al., 2003). These curves represent the relationship between conditional probabilities of failure and levels of stress acting on the defences. One interesting line of research would be to integrate this concept to describe protection efforts.

I also plan on further investigating the user behaviour in post-disruption contexts. In Chapter 6, we model user decisions using a parameter that can be pre-computed by doing surveys. A logic extension would be to represent travel behaviour as an endogenous element to the model. Customer choice models such as multinomial logit (MNL) and ranking based models have been successfully used in the revenue management context to represent the behaviour of a customer facing different products. I think that these models can be applied to evaluate user behaviour in a post-disruption context. In particular, MNL has been already used in the past in a transportation context (Ben-Akiva and Lerman, 1985), therefore I will consider the possibility of using it to extend the model described in Chapter 6.

As explained in the introduction, one of the challenge in protecting Critical Infrastructure is the degree of inter-dependence of these systems. For example, if we focus on transportation, railway and roadway systems cannot be considered as two clearly separated and independent systems. Their infrastructure often overlap and so does their demand. Failing into considering these systems' interactions can lead to less realistic results. Nonetheless, very few works have considered the impact of disruptions on more than one system.

Focusing on road-transportation protection, models that better represent user behaviour are needed. For example, static and dynamic user-equilibrium models could be used to better estimate the impact of disruptions on road-transportation. Further, an integrated model for protecting urban areas, rural areas, and or infrastructure against flooding would be particularly useful.

Finally, this thesis highlights how complex is the task of devising efficient and accurate solution methods. Additional work might focus on refining current solution approaches or propose new decomposition, heuristic or hybrid methods able to cope with the complexity of bi-level, probabilistic, dynamic and multi-objective protection models.

References

Aksen D, Aras N, and Piyade N (2013). A bilevel p-median model for the planning and protection of critical facilities. *Journal of Heuristics* **19(2)** 373-398.

Aksen D, Piyade N, and Aras N (2010). The budget constrained r-interdiction median problem with capacity expansion. *Central European Journal of Operations Research 18(3)* 269-291.

Aksen D, and Aras N (2012). A bilevel fixed charge location model for facilities under imminent attack. Computers & Operations Research **39(7)** 1364-1381.

Aksen D, Aras N (2013) A matheuristic for leader-follower games involving facility location-protection-interdiction decisions. *Studies in Computational Intelligence* **482** 115-151.

Aksen D, Şengul Akca S, and Aras N (2014) A bilevel partial interdiction problem with capacitated facilities and demand outsourcing. *Computers & Operations Research* **41 (1)** 346-358.

Alguacil N, Delgadillo A, and Arroyo, J M (2014). A trilevel programming approach for electric grid defense planning. *Computers & Operations Research* **41** 282-290.

Aliakbarian N, Dehghanian F, and Salari M (2015). A bi-level programming model for protection of hierarchical facilities under imminent attacks. *Computers & Operations Research* **64** 210-224.

Altner D S, Ergun Ö, and Uhan N A (2010). The maximum flow network interdiction problem: valid inequalities, integrality gaps, and approximability. *Operations Research Letters* **38(1)** 33-38.

Baghalian A, Rezapour S, and Farahani R Z (2013). Robust supply chain network design with service level against disruptions and demand uncertainties: A real-life case. *European Journal of Operational Research* **227(1)** 199-215.

Ball M O, Golden B L, and Vohra R V (1989). Finding the most vital arcs in a network. *Operations Research Letters* **8(2)** 73-76.

Bayrak H, Bailey M D(2008) Shortest path network interdiction with asymmetric information. *Networks* **52(3)** 133-140.

Bell M G, Kanturska U, Schmöcker, J D, and Fonzone A (2008). Attacker-defender models and road network vulnerability. *Philosophical Transactions of the Royal Society of London A: Mathematical Physical and Engineering Sciences* **366(1872)** 1893-1906.

Ben-Akiva M E and Lerman S R (1985). Discrete choice analysis: theory and application to travel demand. *MIT press* 9.

Benders J F (1962) Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik* **4(1)** 238–252.

Bennett P G (1977) Toward a Theory of Hypergames. *Omega* **5** 749751

Berman O, Krass D, and Menezes M B (2007). Facility reliability issues in network p-median problems: strategic centralization and co-location effects. *Operations Research* **55(2)** 332-350.

Best G, Parston G, and Rosenhead J (1986). Robustness in practice-the regional planning of health services. *Journal of the Operational Research Society* 463-478.

Brown G, Carlyle M, Salmeron J, and Wood K (2006) Defending critical infrastructure. *Interfaces* **36(6)** 530-544.

Caplin D A, and Kornbluth J S H (1975). Multiobjective investment planning under uncertainty. *Omega* **3(4)** 423-441.

Cappanera P, and Scaparra MP (2011) Optimal allocation of protective resources in shortest-path networks. *Transportation Science* **45** 64-80.

Carrington D, and Weaver M (2014) Emergency funding to repair damaged UK flood defences raised to 130m. Available at http://www.theguardian.com Last accessed: June 2015.

Chen G, Daskin M S, Shen Z J M, and Uryasev S (2006). The α-reliable mean-excess regret model for stochastic facility location modeling. *Naval Research Logistics* **53(7)** 617-626.

Chen Q, Li X, Ouyang Y (2011). Joint inventory-location problem under the risk of probabilistic facility disruptions. *Transportation Research Part B: Methodological* **45(7)** 991-1003.

COMMISSION OF THE EUROPEAN COMMUNITIES (2006). European Programme for Critical Infrastructure Protection. Available at: http://eur-lex.europa.eu Last accessed: June 2015.

Cormen T. H., Leiserson C. E., Rivest R. L., and Stein C. (2001) Introduction to algorithms. *Cambridge: MIT press.*

Cormican K J, Morton D P, Wood R K (1998) Stochastic network interdiction. *Operations Research* **46(2)**, 184-197.

Chang S E (2003). Transportation planning for disasters: an accessibility approach. *Environment and Planning A* **35(6)** 1051-1072.

Church R L, Scaparra M P, and Middleton RS (2004) Identifying critical infrastructure: The median and covering facility interdiction problems. *Annals of the Association of American Geographers* **94** 491–502.

Church R L, Scaparra M P (2007) Protecting Critical Assets: The r-Interdiction Median Problem with Fortification. *Geographical Analysis* **39(2)** 129-146.

Church R L, and Scaparra M P (2007). Protecting critical assets: the r-interdiction median problem with fortification. *Geographical Analysis* **39(2)** 129-146.

Codato G, and Fischetti M (2006). Combinatorial Benders' cuts for mixed-integer linear programming. *Operations Research* **54(4)** 756-766.

Corley H W, and David Y S (1982). Most vital links and nodes in weighted networks. *Operations Research Letters* **1(4)** 157-160.

Cui T, Ouyang Y, and Shen Z J M (2010). Reliable facility location design under the risk of disruptions. *Operations Research* **58(4-part-1)** 998-1011.

Daskin M S, Hesse S M, and Revelle C S (1997). α-reliable p-minimax regret: a new model for strategic facility location modeling. *Location Science* **5(4)** 227-246.

DEFRA (2015). Central Government Funding for Flood and Coastal Erosion Risk Management in England. Available at: https://www.gov.uk. Last accessed: June 2015.

Dutta D, Herath S, and Musiake K (2003). A mathematical model for flood loss estimation. *Journal of Hydrology* **277(1)** 24-49.

Du L, and Peeta S (2014). A stochastic optimization model to reduce expected post-disaster response time through pre-disaster investment decisions. *Networks and spatial economics* **14(2)** 271-295.

Environmental Agency (2009) Flooding in England: A National Assessment of Flood Risk. Available at: `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/292928/geho0609bqds-e-e.pdf` [Accessed 10 Nov 2015]

Fan Y, and Liu C (2010). Solving stochastic transportation network protection problems using the progressive hedging-based method. *Networks and Spatial Economics* **10(2)** 193-208.

Faturechi R, and Miller-Hooks E (2014). Travel time resilience of roadway networks under disaster. *Transportation research part B: methodological* **70** 47-64.

Feo T A, and Resende M G (1995). Greedy randomized adaptive search procedures. *Journal of global optimization* **6(2)** 109-133.

Floyd R W (1962). Algorithm 97: shortest path. *Communications of the ACM* **5(6)** 345.

Fulkerson D R, Harding G C (1977) Maximizing the minimum source-sink path subject to a budget constraint. *Mathematical Programming* **13(1)** 116-118.

Geoffrion A M (1972). Generalized benders decomposition. *Journal of optimization theory and applications* **10(4)** 237-260.

Golden B (1978). A problem in network interdiction. *Naval Research Logistics Quarterly* **25(4)** 711-713.

Ghosh A, and McLafferty S L (1982). Locating Stores in Uncertain Environments-a Scenario Planning Approach. *Journal of Retailing* **58(4)** 5-22.

Gumbel E J (1941). The return period of flood flows. *The Annals of Mathematical Statistics* **12(2)** 163-190.

Gupta S K, and Rosenhead J (1968). Robustness in sequential investment decisions. *Management science* **15(2)** B-18.

Hall J W, Dawson R J, Sayers P B, Rosu C, Chatterton J B, and Deakin R. (2003). A methodology for national-scale flood risk assessment. *Proceedings of the Institution of Civil Engineers-Water Maritime and Engineering* **156(3)** 235-248

Hansen P, Jaumard B, and Savard G (1992) New branch-and-bound rules for linear bilevel programming. *SIAM Journal on Scientific and Statistical Computing* **13(5)** 1194-1217.

Hartong M, Goel R, Wijesekera D (2008) Security and the US rail infrastructure. *International Journal of Critical Infrastructure Protection* **1** 15-28.

Hausken K, and Zhuang J (2011a) Defending against a stockpiling terrorist. *The Engineering Economist* **56(4)** 321-353.

Hausken, K, and Zhuang J (2011b) Governments' and terrorists' defense and attack in a T-period game. *Decision Analysis* **8(1)** 46-70.

Hausken K, and Zhuang J (2012) The timing and deterrence of terrorist attacks due to exogenous dynamics. *Journal of the Operational Research Society* **63(6)** 726-735.

He X, and Liu H X (2012). Modeling the day-to-day traffic evolution process after an unexpected network disruption. *Transportation Research Part B: Methodological* **46(1)** 50-71.

Hernandez I, Ramirez-Marquez J E, Rainwater C, Pohl E, and Medal H (2014). Robust facility location: Hedging against failures. *Reliability Engineering & System Safety* **123** 73-80.

Huang Y, Fan Y, and Cheu R (2007). Optimal allocation of multiple emergency service resources for protection of critical transportation infrastructure. *Transportation Research Record: Journal of the Transportation Research Board* **(2022)** 1-8.

HM Treasury: UK Spending review 2013. Available at https://www.gov.uk/government/topical-events/spending-round-2013.

Israeli E, and Wood R K (2002) Shortest path network interdiction. *Networks* **40(2)** 97-111.

Ishfaq R (2013). Intermodal shipments as recourse in logistics disruptions. *Journal of the Operational Research Society* **64(2)** 229-240.

Jenelius E, Westin J, and Holmgren Å J (2010). Critical infrastructure protection under imperfect attacker perception. *International Journal of Critical Infrastructure Protection* **3(1)** 16-26.

Jin J G, Lu L, Sun L, and Yin J (2015). Optimal allocation of protective resources in urban rail transit networks against intentional attacks. *Transportation Research Part E: Logistics and Transportation Review* **84** 73-87.

Levitin G, and Hausken K (2009) Parallel systems under two sequential attacks. *Reliability Engineering & System Safety* **94(3)** 763-772.

Levitin, G, and Hausken, K (2010) Resource distribution in multiple attacks against a single target. *Risk Analysis* **30(8)**, 1231-1239.

Levitin G, and Hausken K (2012a) Resource distribution in multiple attacks with imperfect detection of the attack outcome. *Risk Analysis* **32(2)** 304-318.

Levitin G, and Hausken K (2012b) Parallel systems under two sequential attacks with imperfect detection of the first attack outcome. *Journal of the Operational Research Society* **63(11)** 1545-1555.

Levitin G, and Hausken K (2013) Defence resource distribution between protection and decoys for constant resource stockpiling pace. *Journal of the Operational Research Society* **64(9)** 1409-1417.

Li X, and Ouyang, Y. (2010). A continuum approximation approach to reliable facility location design under correlated probabilistic disruptions. *Transportation research part B: methodological* **44(4)** 535-548.

Li A C, Nozick L, Xu N, and Davidson R (2012). Shelter location and transportation planning under hurricane conditions. *Transportation Research Part E: Logistics and Transportation Review* **48(4)** 715-729.

Li X, Ouyang Y, and Peng F (2013a). A supporting station model for reliable infrastructure location design under interdependent disruptions. *Transportation Research Part E: Logistics and Transportation Review* **60** 80-93.

Li Q, Zeng B, and Savachkin A (2013b). Reliable facility location design under disruptions. *Computers & Operations Research* **40(4)** 901-909.

Liberatore F, Scaparra M P, Daskin M S (2012). Hedging against disruptions with ripple effects in location analysis. *Omega* **40(1)** 21-30.

Liberatore F, Scaparra M P, Daskin M S (2011) Analysis of facility protection strategies against an uncertain number of attacks: the stochastic R-interdiction median problem with fortification. *Computers & Operations Research* **38(1)**, 357-366.

Liberatore F, Scaparra M P, and Daskin M S (2012) Hedging against disruptions with ripple effects in location analysis. *Omega* **40(1)** 21-30.

Lim C, and Smith J C (2007) Algorithms for discrete and continuous multicommodity flow network interdiction problems. *IIE Transactions* **39(1)** 15-26.

Liu C, Fan Y, and Ordóñez, F. (2009). A two-stage stochastic programming model for transportation network protection. *Computers & Operations Research* **36(5)** 1582-1590.

Losada C, Scaparra M P, Church R L, Daskin M (2012a) The stochastic interdiction median problem with disruption intensity levels. *Annals of Operations Research* **201(1)** 345-365.

Losada C, Scaparra M P, O'Hanley J R (2012). Optimizing system resilience: a facility protection model with recovery time. *European Journal of Operational Research* **217(3)** 519-530.

Luo T, Maddocks A, Iceland C, Ward P, and Winsemius H (2015). World's 15 Countries with the Most People Exposed to River Floods. *World Resources Institute*. Available at: `http://www.wri.org/blog/2015/03/world\%E2\%80\%99s-15-countries-most-people-exposed-river-floods` [Accessed 10 Nov 2015]

Matisziw T C, Murray A T (2009) Modeling $s - t$ path availability to support disaster vulnerability assessment of network infrastructure. *Computers & Operations Research* **36(1)** 16-26.

Medal H R, Pohl E A, and Rossetti M D (2015). Allocating Protection Resources to Facilities When the Effect of Protection is Uncertain. *IIE Transactions* (just-accepted).

Murray A T, Matisziw T C, Grubesic T H (2007) Critical network infrastructure analysis: interdiction and system flow. *Journal of Geographical Systems* **9(2)** 103-117.

Myung Y S, Kim H J (2004) A cutting plane algorithm for computing k-edge survivability of a network. *European Journal of Operational Research* **156(3)** 579-589.

O'Hanley J R, and Church R L (2011) Designing robust coverage networks to hedge against worst-case facility losses. *European Journal of Operational Research.* **209(1)** 23-36.

O'Hanley J R, Scaparra M P, and García S (2013). Probability chains: A general linearization technique for modeling reliability in facility location and related problems. *European Journal of Operational Research* **230(1)** 63-75.

Parvaresh F, Hashemi Golpayegany SA, Moattar Husseini SM, Karimi B (2013) Solving the p-hub Median Problem Under Intentional Disruptions Using Simulated Annealing. *Networks and Spatial Economics* **13 (4)** 445-470.

Peeta S, Salman F S, Gunnec D, Viswanath K (2010). Pre-disaster investment decisions for strengthening a highway network. *Computers & Operations Research* **37(10)** 1708-1719.

Peng P, Snyder L V, Lim A, and Liu Z (2011). Reliable logistics networks design with facility disruptions. *Transportation Research Part B: Methodological* **45(8)** 1190-1211.

Qin X, Liu X, and Tang L (2013). A two-stage stochastic mixed-integer program for the capacitated logistics fortification planning under accidental disruptions. *Computers & Industrial Engineering* **65(4)** 614-623.

Rad M A, Kakhki H T (2013). Maximum dynamic network flow interdiction problem: New formulation and solution procedures. *Computers & Industrial Engineering* **65(4)** 531-536.

Rawls C G, and Turnquist M A (2010). Pre-positioning of emergency supplies for disaster response. *Transportation research part B: Methodological* **44(4)** 521-534.

Rawls C G, and Turnquist M A (2012). Pre-positioning and dynamic delivery planning for short-term response following a natural disaster. *Socio-Economic Planning Sciences* **46(1)** 46-54.

Rosenhead J, and Mingers J (2001). Rational analysis for a problematic world revisited. *John Wiley and Sons.*

Royset J O, and Wood R K (2007). Solving the bi-objective maximum-flow network-interdiction problem. *INFORMS Journal on Computing* 19(2) 175-184.

Saharidis G K, Ierapetritou M G (2009) Resolution method for mixed integer bi-level linear problems based on decomposition techniques. *Journal of Global Optimization* **44 (1)** 29-51.

Sarhadi H, Tulett D M, and Verma M (2015). A defender-attacker-defender approach to the optimal fortification of a rail intermodal terminal network. *Journal of Transportation Security* **8(1-2)** 17-32.

Scaparra M P, and Church R L (2008a) A bilevel mixed-integer program for critical infrastructure protection planning. *Computers & Operations Research* **35(6)** 1905-1923.

Scaparra M P, and Church R L (2008b) An exact solution approach for the interdiction median problem with fortification. *European Journal of Operational Research* **189 (1)** 76-92.

Scaparra M P, and Church R (2012). Protecting supply systems to mitigate potential disaster a model to fortify capacitated facilities. *International Regional Science Review* **35(2)** 188-210.

Scott D M, Novak D C, Aultman-Hall L, and Guo F (2006). Network robustness index: a new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography* **14(3)** 215-227.

Snyder L V and Daskin M S (2005). Reliability models for facility location: the expected failure cost case. *Transportation Science* **39(3)** 400-416.

Snyder L V and Daskin M S (2006). Stochastic p-robust location problems. *IIE Transactions* **38(11)** 971-985.

Sohn J, Kim T J, Hewings G J, Lee J S, and Jang S G (2003). Retrofit priority of transport network links under an earthquake. *Journal of Urban Planning and Development.*

Sohn J (2006). Evaluating the significance of highway network links under the flood damage: An accessibility approach. *Transportation Research Part A: Policy and Practice* **40(6)** 491-506.

Suarez P, Anderson W, Mahal V, and Lakshmanan T R (2005). Impacts of flooding and climate change on urban transportation: A systemwide performance assessment of the Boston Metro Area. *Transportation Research Part D: Transport and Environment* **10(3)** 231-244.

Sun D J, Zhao Y, and Lu Q C (2015). Vulnerability Analysis of Urban Rail Transit Networks: A Case Study of Shanghai, China. *Sustainability* **7(6)** 6919-6936.

Wang J. Y., Ehrgott M., and Chen A. (2014). A bi-objective user equilibrium model of travel time reliability in a road network. *Transportation Research Part B: Methodological* **66** 4-15.

Wintour P and Topham G (2014) UK rail service disruptions continue after high winds. Available at *http://www.theguardian.com/uk-news/2014/feb/13/uk-rail-services-disrupted-high-winds-storm*. Last accessed: May 2014.

Wollmer R. (1964) Removing arcs from a network. *Operations Research* **12** 934–940.

Wood R K (1993) Deterministic network interdiction. *Mathematical and Computer Modelling* **17(2)** (1993) 1-18.

Yates J, and Sanjeevi S (2013) A length-based, multiple-resource formulation for shortest path network interdiction problems in the transportation sector. *International Journal of Critical Infrastructure Protection*, **6(2)** 107-119.

Yin Y (2008). A scenario-based model for fleet allocation of freeway service patrols. *Networks and Spatial Economics* **8(4)** 407-417.

Zenklusen R (2010). Network flow interdiction on planar graphs. *Discrete Applied Mathematics* **158(13)** 1441-1455.

Zhang X, Zheng Z, Zhu Y, and Cai K Y (2014). Protection issues for supply systems involving random attacks. *Computers & Operations Research* **43** 137-156.

Zhang X, Zheng Z, Zhang S, and Du W (2015). Partial interdiction median models for multi-sourcing supply systems. *The International Journal of Advanced Manufacturing Technology* **1-17**.

Zhu Y, Zheng Z, Zhang X, Cai K (2013). The r-interdiction median problem with probabilistic protection and its solution algorithm. *Computers & Operations Research* **40(1)** 451-462.