Optimizing Dynamic Investment Decisions for Railway Systems Protection

Stefano Starita* Maria Paola Scaparra†

Abstract

Past and recent events have shown that railway infrastructure systems are particularly vulnerable to natural catastrophes, unintentional accidents and terrorist attacks. Protection investments are instrumental in reducing economic losses and preserving public safety. A systematic approach to plan security investments is paramount to guarantee that limited protection resources are utilized in the most efficient manner. In this article, we present an optimization model to identify the railway assets which should be protected to minimize the impact of worst case disruptions on passenger flows. We consider a dynamic investment problem where protection resources become available over a planning horizon. The problem is formulated as a bilevel mixed-integer model and solved using two different decomposition approaches. Random instances of different sizes are generated to compare the solution algorithms. The model is then tested on the Kent railway network to demonstrate how the results can be used to support efficient protection decisions.

Keywords. Strategic planning, transportation, protection, bilevel programming, decomposition.

1 Introduction

Nowadays the social well-being of people highly relies on the well functioning of critical interconnected infrastructures such as transportation, information, telecommunication, and electric power systems. Planning and protecting infrastructure systems is a complex task, especially because of their dimension and interdependence. Even small random disruptions can severely affect the normal functioning of one or more infrastructure. Intelligent attacks

*Kent Business School, University of Kent, CT2 7PE Canterbury, UK.
†Corresponding author. Kent Business School, University of Kent, CT2 7PE Canterbury, UK. Tel: +44-1227-824556. Fax: +44-1227-761187. E-mail: m.p.scaparra@kent.ac.uk

or large natural catastrophes can have even more dramatic consequences in terms of both economic and life losses. Examples of such events include the 1995 Paris metro bombing, the 2004 Madrid train bombing, the 2005 London underground suicide attacks, and the 2010 Moscow bombing. Most recently, severe floods hit some western regions of the UK and forced the Network Rail to pay £12.5m for the suppressed services and further £15m to repair the infrastructure (Wintour and Topham, 2014). It is therefore paramount to protect infrastructure systems so that continuity in service provision and safety for the users can be guaranteed, even when disruptions occur.

A critical aspect in planning infrastructure protection is the scarce availability of protection resources. Protecting all the components of an infrastructure system to targeted safety levels may in fact be 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. Another complicating factor in protecting railways is that they are open and easily accessible systems. This renders them highly vulnerable to all kinds of disruption and requires careful identification of suitable protection measures. These may include the structural reinforcement of vulnerable parts (tunnels and bridges), video surveillance of critical areas (crosses, stations) and the use of a wide range of sensors to detect intrusions and obstacles on tracks. Since resources are limited, it is important to identify and protect the most critical assets of the infrastructure.

In recent years, several mathematical models have been developed to identify systems' vulnerabilities and plan protection strategies for critical infrastructures. Predominantly, interdiction and protection of infrastructure systems have been modelled using multi-level optimization. Multi-level optimization models represent an effective tool to "*model a complete infrastructure system and its value to society, including how losses of the system's assets reduce that value, or how improvements in the system mitigate lost value*"(Brown et al., 2006). These models are also referred to as *defender-attacker* models since they emulate the game between two actors with opposite aims. The defender's aim is to distribute limited defence resources so as to minimize the effects of a worst case disruption. On the other hand, the attacker's aim is to choose the attack plan which minimizes the system's value (or maximize the system's cost). The *attacker* is an intelligent actor who has perfect knowledge of the system and is always able to inflict the maximum damage. In other words, he is a proxy to model worse-case disruptions. Clearly this kind of models are very useful to simulate terrorist attacks and intentional disruptions. Nonetheless, given the criticality of infrastructure systems, protection efforts are often guided by risk-averse decision making criteria, thus making these models extremely valuable also for problems involving natural catastrophes.

A third actor, referred to as the *system user*, is often used in multi-level models to

evaluate the system's value after protection and interdiction. Bilevel *attacker-user* models are typically used to identify the vulnerabilities of a system, by highlighting the outcomes of a worst-case interdiction. Trilevel *defender-attacker-user* models are typically used to identify the system's components that should be hardened or protected. Sometimes the models mirroring the actions of the attacker and the system user can be collapsed into a single model (Church et al., 2004, Losada et al., 2012a), so that a single level model can represent an interdiction problem whereas a bilevel model can represent a protection problem.

The main contribution of this paper is to study a dynamic network protection problem. The model we present is quite general and, albeit designed for railway infrastructures, can be applied in other contexts as well. The model aims at distributing protection resources among the assets of a railway system so as to maximise its survivability after a worst case disruption. Generally, survivability can be described as "*the capability of a system to fulfill its mission, in a timely manner, in the presence of attacks, failures, or accidents*"(Ellison et al., 1997). According to the specific context, network survivability can be measured using different metrics. For instance, for shortest-path based problems, the length of a path is the key measure to assess whether a network is vulnerable and reliable. To evaluate survivability, we use the same metric introduced by Myung and Kim (2004), Murray et al. (2007), Matisziw and Murray (2009) and Scaparra et al. (2015). In these works, network survivability is measured in terms of lost or unserved system flow. In our model, the flow between an origin and a destination node is considered lost if after a disruption affecting some network components, the two nodes are no longer connected or they are connected but the post-disruption service is significantly deteriorated (i.e., alternative routes are too long from a user's perspective). In this case, in fact, rail network users may abandon the trip or resort to another mode of transport.

An important issue that should be taken into account when modelling protection efforts is that protection resources usually 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). Similarly, after the 2013-2014 floods in the UK, £130m were allocated by the government to repair flood defences. 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 infrastructure's 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.

Our model builds upon and extends the static protection model proposed by Scaparra et al. (2015) by considering dynamic investments. The model has a bilevel structure where the aim of the upper level is to find the best allocation of protection resources over a planning horizon so as to minimize the amount of disrupted flow. The lower level is used to evaluate worse case losses in each time period in response to a given protection plan. The resulting multi-period bilevel model is significantly more difficult to solve than its static counterpart. To find optimal solutions to problem instances of realistic size, we propose two decomposition approaches tailored to the dynamic structure of the model and test them on a set of new, randomly generated instances. Given that the multi-period structure of the model can lead to solve the same lower level program multiple times in our iterative approaches, we streamline the algorithms by using efficient data structures, thus avoiding recomputing solutions already found. The use of this expedient proved to be extremely efficient and on some preliminary tests reduced the overall computing time of the decomposition methods by as much as 80%. Finally, some practical insights, including an analysis of dynamic investments, are discussed for a real network representing the railway infrastructure of Kent (UK).

The remainder of this paper is organized as follows. Sec. 2 provides a review of the literature related to this work. In Sec. 3, the bilevel formulation of DNP is introduced. Sec. 4 provides a description of the decomposition algorithms. In Sec. 5, we report computational results on two sets of random problems, while the results on the Kent case study are analysed in Sec. 6. Some conclusive remarks are discussed in Sec. 7.

2 Background

Since the seminal paper of Wollmer (1964), several papers have appeared in the literature which deal with network interdiction problems. The underlying idea of these models is to find the critical arcs and/or nodes of a network. An element is critical if, after interdiction, the performance of the network drops significantly. Therefore a key issue in these models is to identify a suitable metric to evaluate network performance.

One stream of the literature has analysed the problem of finding the network components

that, when interdicted, minimize the maximum flow between pairs of nodes (Wollmer, 1964). Wood (1993) introduced an integer programming formulation for this problem and extended the model by considering multiple resources and multiple commodities. Since real life interdiction problems are intrinsically characterized by uncertainty (what will the interdicted components be? when will they be interdicted? to what extent?), subsequent work has focused on problems including stochastic and probabilistic elements. For instance, Cormican et al. (1998) proposed a stochastic model aiming at minimizing the expected maximum flow on a network while taking into account that interdictions can be unsuccessful. They also considered further extensions in which arc capacities are random variables and multiple attacks on the same arc can be attempted. Lim and Smith (2007) worked on a multi commodity flow network, considering both discrete and continuous interdiction. Other recent network flow interdiction models can be found in Afshari and Kakhki (2013), Altner et al. (2010), and Royset and Wood (2007).

Another stream of the literature has focused on distance based networks. The models developed in this framework aim at identifying the elements that, if disrupted, maximize the distance, travel time or transportation cost between nodes. Fulkerson and Harding (1977) studied the interdiction problem in a shortest path network using continuous variables. Israeli and Wood (2002) formulated the same problem using binary interdiction variables and introduced different solution approaches. Hemmecke et al. (2003) extended the problem to stochastic networks. Bayrak and Bailey (2008) studied the asymmetric interdiction problem, considering that the two actors have different levels of knowledge of the network.

A third stream has dealt with network connectivity. In these problems, the aim of the interdictor is to use offensive resources in order to reduce connectivity. Grubesic et al. (2008) used graph theory to assess network connectivity and vulnerability. Several metrics to estimate the network connectivity have been proposed. Albert et al. (2000) used the network diameter, Grubesic et al. (2003) the degree of a node. Shen et al. (2012) used three different metrics: the number of connected components, the largest component size, and the minimum cost required to reconnect the graph after the loss of some nodes.

Interdiction models are a useful tool to identify infrastructure vulnerabilities and critical assets. However, protection decisions must be explicitly represented into a model to ensure that the most efficient resource allocation is identified. In fact, choosing the elements to protect from the interdiction set obtained by solving an *attacker-user* model may lead to sub-optimal solutions (Church and Scaparra, 2007). The literature stream dealing with protection models has mainly focused on facility protection in supply chain systems (Scaparra and Church, 2008a, Liberatore and Scaparra, 2011, Liberatore et al., 2012, Losada et al., 2012b, Bricha and Nourelfath, 2015). Only a few papers have dealt with protection issues

in transportation networks. Cappanera and Scaparra (2011) analyzed a trilevel protection problem to study the optimal allocation of protection resources in shortest path networks. Yates and Sanjeevi (2013) studied a variation of the shortest path fortification problem where the aim of the interdictor is to travel across a network towards a precise target without being detected. The defender's goal is to find the optimal allocation of protection devices to detect the intrusions. Du and Peeta (2014) presented a stochastic model to identify the optimal allocation of defense resources in order to minimise the post-disaster response time of a transportation network. Talarico et al. (2015) proposed a model to allocate security resources in a multi-level chemical transportation network.

If we narrow the focus to the railway context, the number of papers dealing with protection is quite limited. Peterson and Church (2008) introduced a framework to assess the vulnerability of rail networks and apply it to the State of Washington infrastructure. Laporte et al. (2010) proposed a model to design a robust railway transit network maximizing its minimum utility (e.g., trip coverage) when one link can fail. This model was later extended by Perea and Puerto (2013).

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. These solutions cannot always be easily identified and, consequently, the application of these models is often limited to simple problems (i.e., small systems or problems with a single target or only two time periods). On the contrary, our proposed approaches, based on the use of sophisticated decomposition techniques for solving mixed-integer bilevel programs, are able to solve problems where all of the network

components (nodes and arcs) can be protected and interdicted, the planning horizon can include several time periods, and the networks have realistic sizes.

3 The Dynamic Network Protection Problem (DNP)

We consider an undirected 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. This assumption is made to contemplate the possibility of disruptions of tracks, tunnels, bridges and stations at the same time.

(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.

(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.
q_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.

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.

$$[\text{DNP}] \quad \min_{\mathbf{Y}} F(\mathbf{Y}) \tag{1}$$

$$\text{s.t.} \quad \sum_{u=1}^{t} \Big(\sum_{i \in N} q_i^n Y_{iu}^n + \sum_{j \in A} q_j^a Y_{ju}^a\Big) \leq q_t \qquad \forall t \in T \tag{2}$$

$$Y_{it}^n \in \{0,1\} \qquad \forall i \in N, \forall t \in T \tag{3}$$

$$Y_{jt}^a \in \{0,1\} \qquad \forall j \in A, \forall t \in T \tag{4}$$

$$\text{where} \quad F(\mathbf{Y}) = \max_{\mathbf{X}} \sum_{t \in T} \lambda_t \sum_{s} \sum_{d} f_{sd} Z_{sdt} \tag{5}$$

$$\text{s.t.} \quad X_{it}^n \leq 1 - \sum_{u=1}^{t} Y_{iu}^n \qquad \forall i \in N, \forall t \in T \tag{6}$$

$$X^a_{jt} \leq 1 - \sum_{u=1}^{t} Y^a_{ju} \qquad \forall j \in A, \forall t \in T \tag{7}$$

$$\sum_{i \in N} p^n_i X^n_{it} + \sum_{j \in A} p^a_j X^a_{jt} \leq p_t \qquad \forall t \in T \tag{8}$$

$$\sum_{i \in N(r)} X^n_{it} + \sum_{j \in A(r)} X^a_{jt} \geq Z_{sdt} \qquad \forall s, d \in N, r \in N_{sd}, \forall t \in T \tag{9}$$

$$X^n_{it} \in \{0, 1\} \qquad \forall i \in N, \forall t \in T \tag{10}$$

$$X^a_{jt} \in \{0, 1\} \qquad \forall j \in A, \forall t \in T \tag{11}$$

$$Z_{sdt} \in \{0, 1\} \qquad \forall s, d \in N, \forall t \in T. \tag{12}$$

In the bilevel model above, the leader seeks the optimal protection strategy to minimize the function F (1), which represents the weighted sum of demand that cannot be served after interdiction, over the planning horizon. Constraint (2) represents the budget limit: the amount of resources utilized up to period t for nodes and arcs protection cannot exceed the available cumulative budget q_t. Constraints (3) and (4) are the binary requirements for the protection variables. The lower level program (5)-(12) is the interdiction model. The follower seeks the attack strategy that maximizes the overall amount of unserved demand (5). Constraints (6) state that a node cannot be disrupted at period t, if it is protected in the time window $\{1, ..., t\}$. Similarly, constraints (7) state that an arc cannot be disrupted at period t, if it is protected in the time window $\{1, ..., t\}$. Constraints (8) set a limit on the interdiction resources available in each time period. Constraints (9) 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 (10)-(12) enforce binary restrictions on the lower level variables.

4 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). 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 sub-problems to find an overall optimal solution.

In this paper, 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., 2012b). The second approach utilizes special cutting planes known as Super Valid Inequalities (SVIs). An 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. (2012b).

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 (5)-(12) 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 ap-

proaches, is simply:

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

$$\max_{\mathbf{X}} \sum_{t\in T} \lambda_t \sum_s \sum_d f_{sd} Z_{sdt} \tag{13}$$

$$\text{s.t.} \quad X^n_{it} \leq 1 - \sum_{u=1}^{t} \hat{Y}^n_{iuw} \qquad \forall i \in N, \forall t \in T \tag{14}$$

$$X^a_{jt} \leq 1 - \sum_{u=1}^{t} \hat{Y}^a_{juw} \qquad \forall j \in A, \forall t \in T \tag{15}$$

$$(8) - (12)$$

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.

4.1 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.

In BND-D, the RMP at iteration $\bar{w}$ is a mixed-integer program defined as follows:

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

$$\min_{\mathbf{Y}} z \tag{16}$$

$$\text{s. t.} \quad (2) - (4)$$

$$z \geq \sum_{(s,d,t)\in \bar{Z}_w} \lambda_t \left(f_{sd}(1 - Q_{sdtw})\right) \qquad \forall w \in [1, \bar{w}] \tag{17}$$

$$\sum_{i\in N(r)} \hat{X}^n_{itw} \sum_{u=1}^{t} Y^n_{iu} + \sum_{j\in A(r)} \hat{X}^a_{jtw} \sum_{u=1}^{t} Y^a_{ju} \geq a_{rtw}\, Q_{rtw} \tag{18}$$

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

$$\sum_{r \in N_{sd}} Q_{rtw} \geq Q_{sdtw} \qquad \forall s \in N, d \in N, t \in T : (s,d,t) \in \bar{Z}_w, \forall w \in [1, \bar{w}] \tag{19}$$

$$Q_{rtw} \in \{0,1\} \qquad \forall s \in N, d \in N, t \in T : (s,d,t) \in \bar{Z}_w, \forall r \in N_{sd}, \forall w \in [1, \bar{w}] \tag{20}$$

$$Q_{sdtw} \in \{0,1\} \qquad \forall (s,d,t) \in \bar{Z}_w, \forall w \in [1, \bar{w}] \tag{21}$$

$$z \in \mathbb{R}^+. \tag{22}$$

The aim of the objective function (16) is to find the best protection strategy that thwarts the interdiction plans identified in the previous iterations. Constraints (17) 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 (18) 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 (19) 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 (17). Finally, constraints (20) and (21) represent the binary requirements for the variables Q_{rtw} and Q_{sdtw} and constraint (22) states that variable z is a non negative real.

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

Algorithm 1 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.

4.2 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:

$$SVI\left(\hat{\mathbf{X}}_{\mathbf{w}}\right): \sum_i \sum_t \hat{X}^n_{itw} \sum_{u=1}^{t} Y^n_{iu} + \sum_j \sum_t \hat{X}^a_{jtw} \sum_{u=1}^{t} Y^a_{ju} \geq 1. \tag{23}$$

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 $(2) - (4)$ 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 (23) 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 (23) is super-valid. If the solution is sub-optimal, adding inequality (23) 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 $\hat{\mathbf{Z}}_{\mathbf{w+1}}\hat{\mathbf{X}}_{\mathbf{w+1}}$ of the SP that is different from the previous one because of constraints (14) and (15). Thus, for every w, the

inequality is super-valid because it eliminates the incumbent solution, i.e.:

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

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

Algorithm 2 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 $\hat{\mathbf{Z}}_{\mathbf{w}}\hat{\mathbf{X}}_{\mathbf{w}}$ 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 $SVI(\hat{\mathbf{X}}_{\mathbf{w}})$ 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 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 *SubProblem* could be decomposed into $|T|$ independent interdiction problems. Some of these sub-problems recurred multiple times across different iterations. We therefore used a hash table to store and retrieve their solutions efficiently. On some preliminary tests, this expedient yielded a reduction in computing time as high as 80%.

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_i^n and p_i^n) 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 $p_j^a = 1$. 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 1. 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, $q_{|T|} = \alpha B$. 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 1: Disruption scenarios

Size	Resource units
Small	2
Medium	4
Large	6

The results for the two data sets are displayed in Table 2 and Table 3, respectively.

Table 2: Computational comparison between BND-D and SVI-D for the 10-15-x networks

Network name	Disr units per period	Prot budget	Objective value	Computing time (sec) BND-D	Computing time (sec) SVI-D	Prot arcs	Prot nodes	Disr arcs	Disr 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 3: Computational comparison between BND-D and SVI-D for the 20-25-x networks

Network name	Disr units per period	Prot budget	Objective value	Computing time (sec) BND-D	Computing time (sec) SVI-D	Prot arcs	Prot nodes	Disr arcs	Disr 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 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.

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 4, 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).

Table 4: 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

Although in most of the cases an increase in the path threshold value results 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 Fig. 1 we compare the number of arcs and nodes protected and interdicted over the planning horizon.

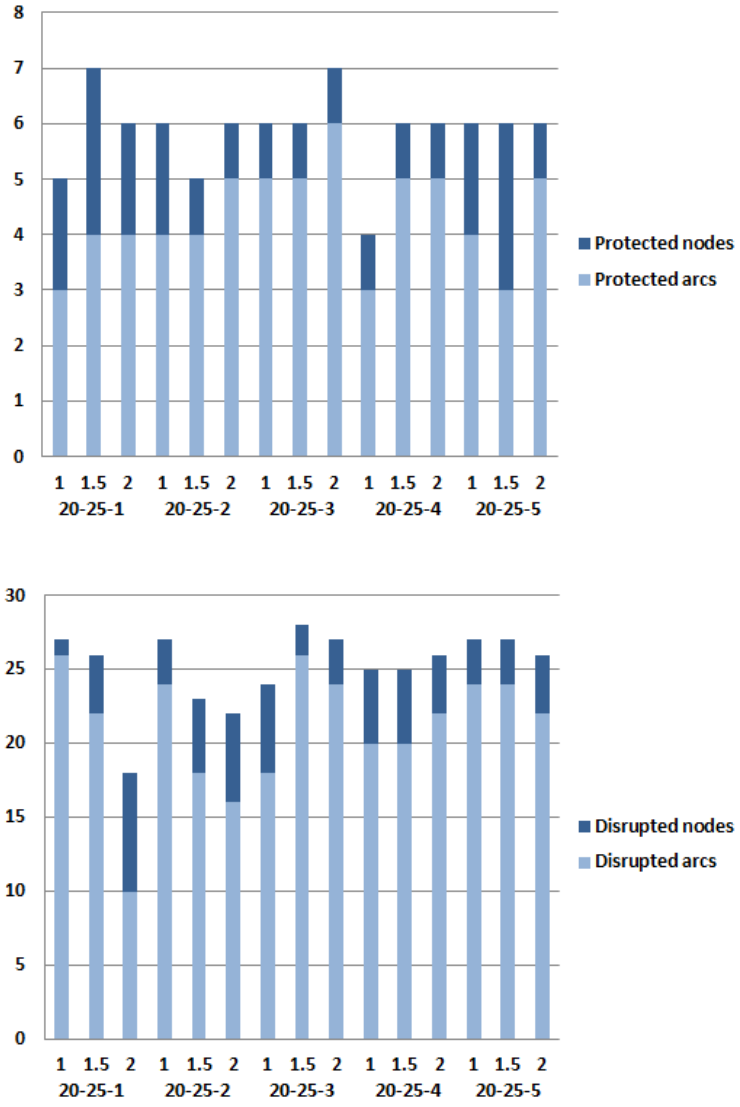


Figure 1: Impact of the path threshold on the number of protected/interdicted elements ($p_t = 6, \forall t$, $\alpha = 10\%$).

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, modeling users' behavior accurately is a critical issue when solving this type of protection models for service systems.

6 Case study analysis

In this section, we test the efficiency of the decomposition approaches and analyze 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 Fig. 2.

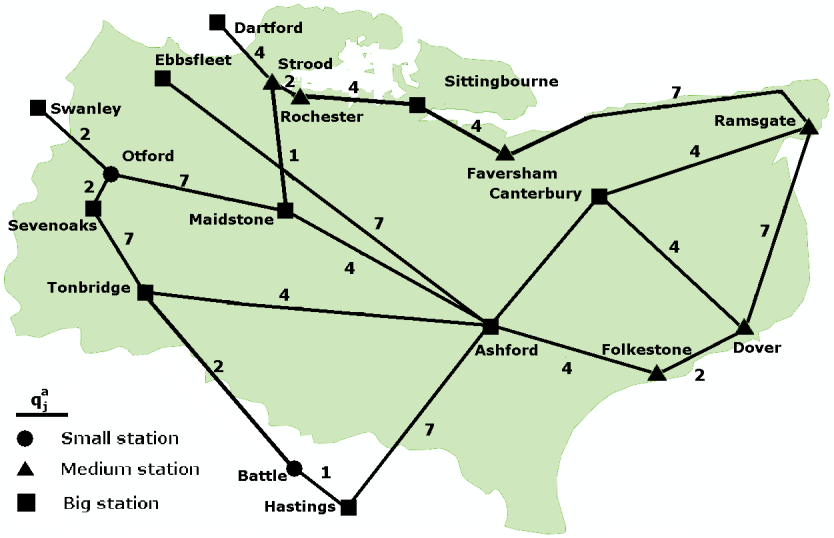


Figure 2: Railway network in Kent (UK).

In the absence of real flow data, we generated the flow matrix as a function of the dimension of the connected towns, and the frequency and capacity of the trains travelling on the network. 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 5). For example, Battle which is a small touristic town with less than half a million annual passengers, needs two units to be disrupted/protected. Differently, 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 Fig. 2. The disruption scenarios are the same as the ones used in the previous section (see Table 1).

Table 5: 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.

6.1 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 6: 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 6 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.

Fig. 3 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 form 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.

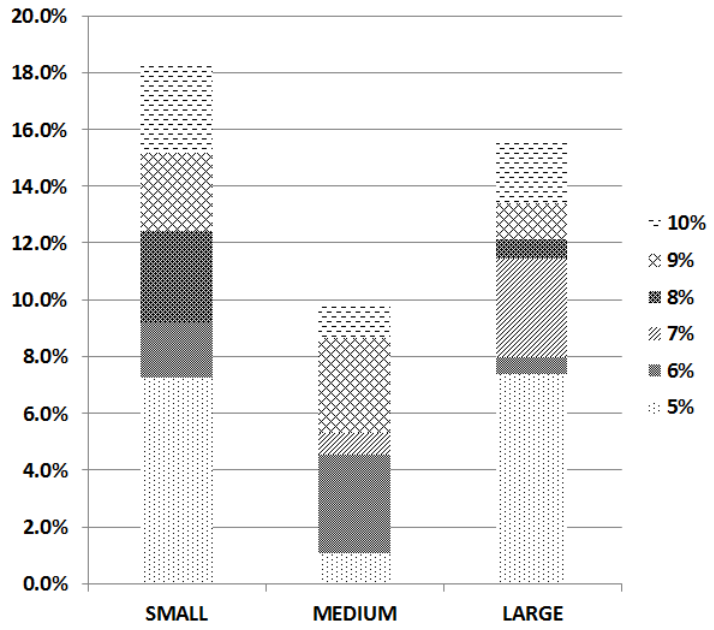


Figure 3: Marginal percentage decrease in flow loss due to unit increments of the protection budget.

6.2 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 7.

Table 7: 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. 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).

6.3 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 decrease 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. Note that the weight λ_t associated with the time periods in the model objective

has some similarity with the discount parameter used in economics, in that it can be used to discount future losses. However, in economics, future values are usually discounted more heavily. In our model, instead, the weights vary and higher values may be associated with the last time period if the ultimate objective is to achieve maximum protection effectiveness at the end of the planning horizon.

In our analysis, 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 8: 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 8 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 at the end of the planning horizon.

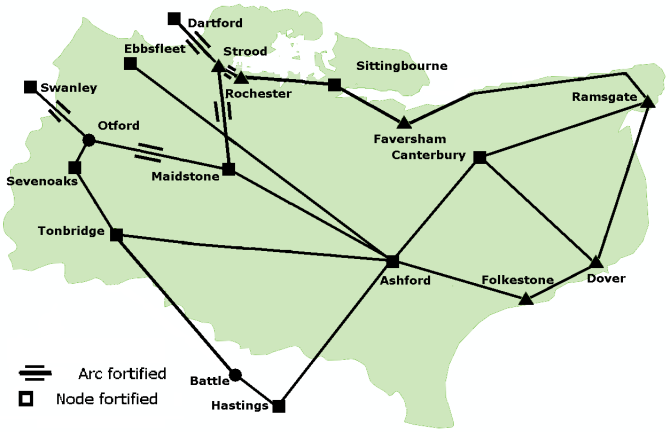
6.4 Solution analysis

In this section we show a sample solution of the proposed model. In particular, Fig. 4 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.

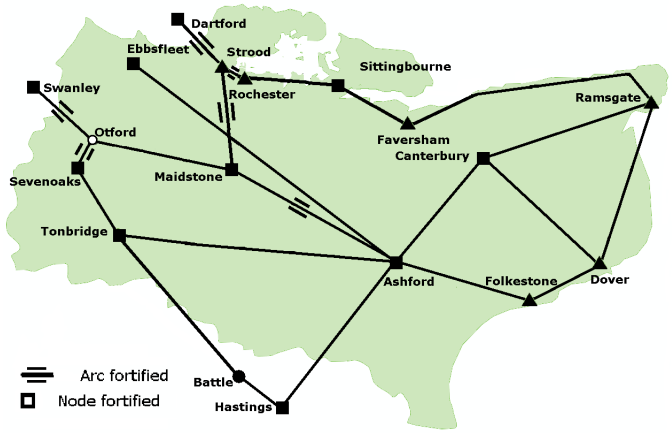
It is clear that protecting the traffic to and from London is of strategic importance. In fact, both the arcs connected to Swanley and Dartford are chosen for protection. Also some arcs connected to Maidstone and Ashford are protected. These towns are among the most populated in Kent and therefore generate 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 (Fig. 4c), more stations can be disrupted. Consequently, more stations appear in the optimal protection strategy.

Finally, Fig. 5 shows the network components involved in a worst case disruption, after the implementation of the optimal protection strategies displayed in Fig. 4. 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).

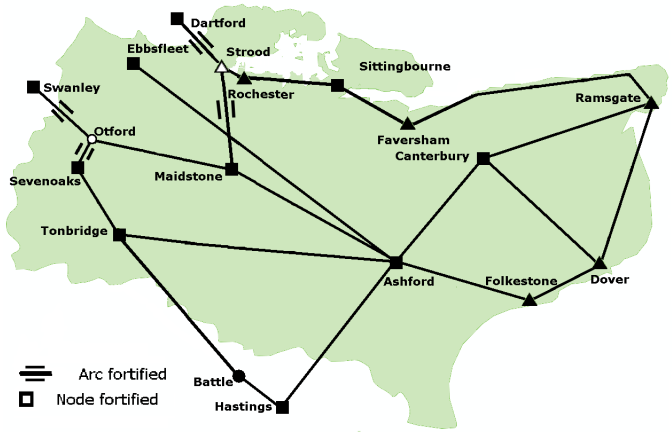
Table 9 provides the details of how the optimal protection strategies are implemented over the planning horizon and, for each time period, displays the worst case interdictions. It can be noticed that the three protection plans share several targets to protect. Nonetheless the periods in which these 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.

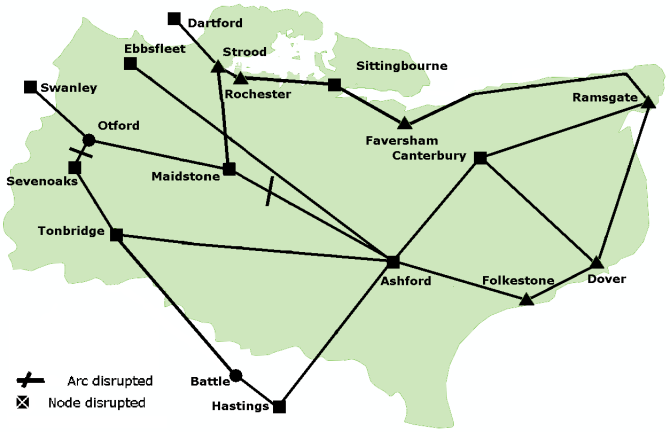


(b) Medium disruption.

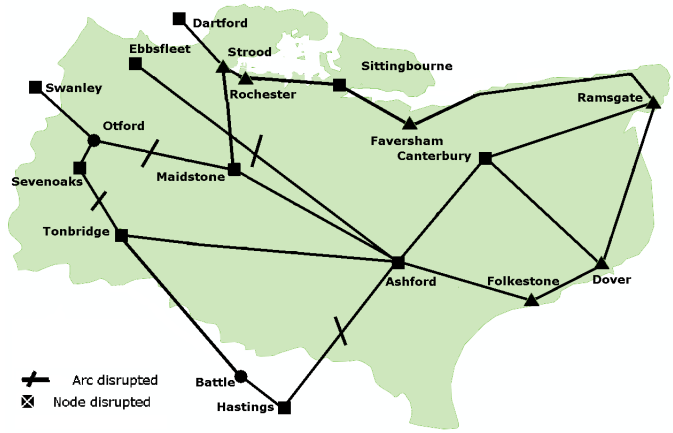


(c) Large disruption.

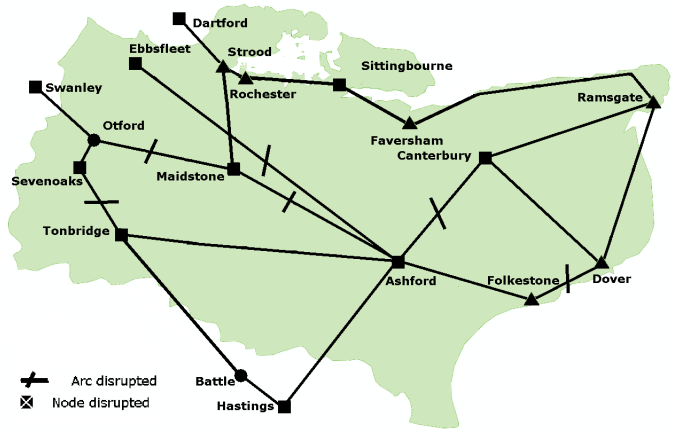
Figure 4: Optimal protection plans for different disruption scenarios.



(a) Small disruption.



(b) Medium disruption.



(c) Large disruption.

Figure 5: Post-protection worst case losses in different disruption scenarios.

Table 9: 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		Ashford-Ebbslfeet	
		Otford-Swanley		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		Ashford-Ebbslfeet	Otford
		Rochester-Strood		Ashford-Maidstone	
	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	Strood
				Otford-Swanley	
	1		Strood	Ashford-Canterbury	
				Ashford-Ebbsfleet	
				Dover-Folkestone	
				Maidstone-Strood	
				Otford-Swanley	
				Rochester-Strood	
	2	Maidstone-Strood		Ashford-Canterbury	
		Rochester-Strood		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	

7 Conclusions and discussion

To protect critical infrastructure systems, it is necessary to distribute limited protection resources in the most effective way. This paper introduced a bilevel 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 modeling 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' behavior and identify the *acceptable* paths from a user perspective. A limitation of the current model is that each origin-destination path is either acceptable or not. Given that the solutions identified by the model are highly sensitive to the path threshold parameter, variations to this basic model should be developed which better capture the users' behavior. A logical extension would be to consider that, following a disruption, the proportion of users taking a different path depends on the extra travel time of this path, compared to the shortest one.

Other possible extensions of this work include the following. Other metrics, such as the system costs, and the duration and frequency of a disruption, should be used to measure the system's performance. These aspects could be merged into a multi-objective model. Our model only considers binary interdictions and protections. Future works may relax this assumption by considering different levels of disruption/protection for each asset at different costs. Protection models against random failures, as opposed to worst-case interdictions, should also be developed and the protection strategies identified by the two modeling approaches should be compared. Finally, an interesting line of research would be to consider scenario-indexed models including robustness measures. These models, which directly capture the intrinsic uncertainty of disruptive events, would be better suited to identify robust solutions across different disruption scenarios.

In terms of methodology, adding a temporal component undoubtedly renders this type of

bilevel protection problems significantly more difficult to solve than their static counterparts. The proposed solution approaches can only be applied to small/medium networks, such as the one used in the Kent case study. Solving larger instances will require developing more sophisticated approaches, including heuristics and hybrid approaches.

Acknowledgements

The authors gratefully acknowledge three anonymous referees for their insightful and constructive comments.

References

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

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

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

Albert R, Jeong H, and Barabasi A L (2000) Error and attack tolerance of complex networks. *Nature* **406** 378–382.

Altner DS, Ergun O, and Uhan NA (2010) The maximum flow network interdiction problem: valid inequalities, integrality gaps, and approximability. *Operations Research Letters* **38(1)** 33–38.

Azaiez MN, and Bier VM (2007) Optimal resource allocation for security in reliability systems. *European Journal of Operational Research* **181(2)** 773–786.

Bayrak H, and Bailey MD (2008) Shortest path network interdiction with asymmetric information. *Networks* **52(3)** 133–140.

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

Bricha N, and Nourelfath M (2015) Protection of warehouses and plants under capacity constraint. *Reliability Engineering & System Safety* **138** 93–104.

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

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/environment/2014/feb/06/emergency-funding-uk-flood-defences-storms-weather*. Last accessed: June 2015.

Church RL, Scaparra MP, 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 RL, and Scaparra MP (2007) Protecting Critical Assets: The r-Interdiction Median Problem with Fortification. *Geographical Analysis* **39(2)** 129–146.

Cormen TH, Leiserson CE, Rivest RL, and Stein C (2001) Introduction to algorithms. *Cambridge: MIT press.*

Cormican KJ, Morton DP, and Wood RK (1998) Stochastic network interdiction. *Operations Research* **46(2)** 184–197.

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

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** 271–295.

Ellison RJ, Fisher DA, Linger RC, Lipson HF, and Longstaff T (1997) Survivable network systems: An emerging discipline (No. CMU/SEI-97-TR-013).

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

Grubesic TH, O'Kelly ME, and Murray AT (2003) A geographic perspective on commercial internet survivability. *Telematics and Informatics* **20** 51–69.

Grubesic TH, Matisziw TC, Murray AT, and Snediker D (2008) Comparative approaches for assessing network vulnerability. *International Regional Science Review* **31(1)** 88–112.

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.

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.

Hemmecke R, Schultz R, and Woodruff DL (2003) Interdicting stochastic networks with binary interdiction effort. In: Network interdiction and stochastic integer programming. *Springer US.*

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

Israeli E, and Wood RK (2002) Shortest path network interdiction. *Networks* **40(2)** 97–111.

Laporte G, Mesa JA, and Perea F (2010) A game theoretic framework for the robust railway transit network design problem. *Transportation Research Part B: Methodologicals* **44(4)** 447–459.

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.

Liberatore F, and Scaparra MP (2011) Optimizing protection strategies for supply chains: comparing classic decision-making criteria in an uncertain environment. *Annals of the Association of American Geographers* **101(6)** 1241–1258.

Liberatore F, Scaparra MP, and Daskin MS (2012) Hedging against disruptions with ripple effects in location analysis. *Omega* **40(1)** 21–30.

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

Losada C, Scaparra MP, Church RL, and Daskin M (2012a) The stochastic interdiction median problem with disruption intensity levels. *Annals of Operations Research* **201(1)** 345–365.

Losada C, Scaparra MP, and O'Hanley JR (2012b) Optimizing system resilience: a facility protection model with recovery time. *European Journal of Operational Research* **217** 519–530.

Matisziw TC, and Murray AT (2009) Modeling *st* path availability to support disaster vulnerability assessment of network infrastructure. *Computers & Operations Research* **36(1)** 16–26.

Murray AT, Matisziw TC, and Grubesic TH (2007) Critical network infrastructure analysis: interdiction and system flow. *Journal of Geographical Systems* **9(2)** 103–117.

Myung YS, and Kim HJ (2004) A cutting plane algorithm for computing k-edge survivability of a network. *European Journal of Operational Research* **156(3)** 579–589.

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

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.

Perea F, and Puerto J (2013) Revisiting a game theoretic framework for the robust railway network design against intentional attacks. *European Journal of Operational Research.* **226** 286–292.

Peterson SK, and Church RL (2008) A Framework for Modeling Rail Transport Vulnerability. *Growth and Change* **39** 617–641.

Royset JO, and Wood RK (2007) Solving the bi-objective maximum-flow network-interdiction problem. *INFORMS Journal on Computing* **19(2)** 175–184.

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

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

Scaparra MP, and Church RL (2008b) An exact solution approach for the interdiction median problem with fortification. *European Journal of Operational Research* **189 (1)** 76–92.

Scaparra MP, Starita S, and Sterle C (2015). Optimizing investment decisions for railway systems protection. In *Railway Infrastructure Security*, pp. 215-233. Setola, Sforza, Vittorini, Pragliola Eds. Springer International Publishing.

Snyder LV, Daskin MS (2006) Stochastic p-robust location problems. *IIE Transactions* **38** 971–985.

Shen MP, Smith J C, and Goli R (2012) Exact interdiction models and algorithms for disconnecting networks via node deletions. *Discrete Optimization* **9** 172–188.

Talarico L, Reniers G, Sörensen K, Springael, J (2015) MISTRAL: A game-theoretical model to allocate security measures in a multi-modal chemical transportation network with adaptive adversaries, *Reliability Engineering & System Safety* **138** 105–114.

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 RK (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.