Operations Research
manuscript

Mitigating Interdiction Risk with Fortification *

Le Thi Khanh Hien
Department of Mathematics and Operational Research, University of Mons, Belgium, thikhanhhien.le@umons.ac.be

Melvyn Sim
Department of Analytics & Operations, NUS Business School, National University of Singapore, Singapore 119245, melvynsim@gmail.com

Huan Xu
School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Drive, NW, Atlanta GA, 30332 huan.xu@isye.gatech.edu

We study a network fortification problem on a directed network that channels single-commodity resources to fulfill random demands delivered to a subset of the nodes. For given a realization of demands, the malicious interdictor would disrupt the network in a manner that would maximize the total demand shortfalls subject to the interdictor's constraints. To mitigate the risk of such shortfalls, a network's operator can fortify it by providing additional network capacity and/or protecting the nominal capacity. Given the stochastic nature of the demand uncertainty, the goal is to fortify the network, within the operator's budget constraint, that would minimize the expected disutility of the shortfalls in events of interdiction. We model this as a three-level, nonlinear stochastic optimization problem that can be solved via a robust stochastic approximation approach under which each iteration involves solving a linear mixed-integer program. We provide favourable computational results that demonstrate how our fortification strategy effectively mitigates interdiction risks. We also extend the model to multi-commodity network with multiple sources and multiple sinks.

Key words: Fortification model; random demand; robust stochastic approximation; multiple sources and sinks.

1. Introduction

Network resilience is an important operational problem in many applications including military, infectious disease control, computer networks, emergency service protection – to name just a few. Consider a transportation network (e.g., a city subway system). If even a few arcs (a.k.a. bridges,

* This research was supported by A*STAR Grant 1421200078

edges) break down, then the network's transportation ability may be significantly weakened. How to respond to, reduce the number of, or even prevent such events are therefore important questions faced by the network's designer and operators. This network "interdiction and protection" problem was first posed in the 1960s (see Wollmer 1964) and has since then been extensively studied, well documented, and broadly applied. A typical interdiction problem includes two players: the attacker (interdictor) and the defender (operator). The attacker is modeled as taking actions to destroy the system in the sense of lowering the operator's optimal value, while the defender operates the network and responds to the attack by finding the optimal configuration for the post-attack network (see e.g. Fulkerson and Harding 1977; Washburn and Wood 1995; Wood 1993). The network operator may also take preventive measures to reduce the likelihood of attacks and to ameliorate their effects. These actions characterize the *fortification problem*, whereby the operator is given a budget to reinforce the network: to improve its robustness, reliability, and sustainability. In the fortification problem, the attacker can be modeled as being either a purposely malicious interdiction (e.g., man-made threats) or as an unintentional and uncertain disruption (e.g., natural disasters).

Most fortification models described in the literature fall into one of two main streams: networked system protection (a.k.a. linkage fortification) and location analysis. Fortification models in *location analysis* address the problem of finding the location at which defensive resources can be optimally allocated to minimize the effects of possible destruction of infrastructure systems; see, for example, Carr et al. (2006), Gendreau et al. (2000), Hodgson et al. (1996), James and Salhi (2002), Liberatore et al. (2011), Scaparra and Church (2008, 2012), and the references therein. In contrast, fortification models in *networked system protection* seek to identify (within a limited budget) the most critical arcs of the network so that they can be protected or augmented with secondary capacity. In this paper we propose a linkage fortification model that optimally distributes a limited amount of protective resources among a set of arcs in a directed network – which channels single-commodity flows to a subset of the nodes in satisfying random demands – in order to mitigate a worst-case disruption event. We show how this model can help inform planners in disruption management.

Note that the fortification problem differs from the network design (or capacity planning) problem. The latter amounts to building a network while assuming that no facilities currently exist; the former involves deciding which network facilities, among those that are already in place, are the best candidates for protection or expanded capacity. Although a variety of fortification models in location analysis have been proposed and thoroughly studied in terms of different objectives and underlying network structures, there are very few linkage fortification models in the literature. Hence we shall start by briefly reviewing all the linkage fortification models of which we are aware.

The operator tends to analyze the most disruptive situation associated with an interdiction event, whether natural or intended, so as to provide a safe guarantee of how the network behaves when disrupted. It follows that worst-case analysis is an appropriate and reasonable tool for modeling the interdiction events in a fortification model. See Brown et al. (2005) and Salmeron et al. (2004) for brief discussions of the concept of protection against worst-case losses. Using worst-case analysis in a fortification model often leads to a defender–attacker–defender (DAD) sequential game model, which is a three-level optimization problem. We remark that there exists no universal algorithm for solving multi-level optimization problems. Brown et al. (2006) propose a DAD model for the purpose of defending infrastructure against terrorist attacks. They illustrate the model's use in an electrical transmission system; the model helps identify the set of critical power lines that must be defended to minimize the worst-case disruption – where disruption may be viewed as the penalty for the demand shortage accumulated in different customer sectors – that is caused by a group of terrorists with limited offensive resources. However, Brown et al. (2006) do not actually solve the DAD model. Alderson et al. (2011), who consider a municipal transportation network subject to minimizing the total user travel time for a single period, use a DAD model to protect and improve the resilience of the network against intelligent attacks by identifying the key links to be defended under a limited total defense budget; these authors then develop a decomposition algorithm to solve their model. In Scaparra and Cappanera (2005; see also Snyder et al. 2006), the DAD approach is used to describe a shortest-path interdiction problem with fortification; in their setup, optimal fortification plans are identified to mitigate the destruction caused by an opponent that is attempting to disable network linkages. Scaparra and Cappanera (2005) also propose a maximum-flow interdiction problem with fortification. In this case, the network is fortified so as to maximize commodity flow following worst-case damage to some network linkages.

In most of the fortification models just cited, demands are deterministic. Yet demands are known to exhibit considerable variation in many networks, so it makes more sense to capture them by positing an uncertain environment than a deterministic one. Furthermore, there are often sufficient historical data to estimate the demand's probability distribution. Stochastic programming (SP) is tool widely used for explicitly modeling uncertain quantities, and it has proved to be both flexible and effective in many areas of science and engineering. In network problems that use SP to model demand uncertainty, demand is viewed as a random variable characterized by a known probability distribution. Sen et al. (1994), for instance, model a planning problem with demand uncertainty – for telecommunication networks that provide private line services – by using a two-stage stochastic linear program with recourse. Other examples include Santoso et al. (2005), who use a two-stage stochastic model to represent a supply chain network design problem with uncertain demand, and Riis and Andersen (2002), who develop a stochastic integer programming formulation to model the telecommunication network design problem of distributing capacity across network links in a manner that satisfies uncertain customer demand while minimizing the total incurred cost. For more examples of using SP to model demand uncertainty in network problems, interested readers are referred to Baghalian et al. (2013), Laporte et al. (1994), Olinick and Rosenberger (2008), Shapiro et al. (2014), Terblanche et al. (2011), and the references therein.

Our own paper's main contributions may be summarized as follows. We incorporate the DAD sequential game model with SP to study optimal allocation of the defensive resource to fortify a transportation network. Our model accounts for three crucial factors and thus models interdiction risks in a setting that is more realistic than those discussed in much of the previous research.

(i) *Adversarial disruption:* Adversarial agents seek to maximize the effects of their attack.

(ii) *Stochastic demand:* Demand is random and does not depend on the actions of adversarial agents.

(iii) *Risk-aware decision maker:* Different decision makers have different perceptions of the trade-offs between the risks involved and the associated costs incurred by fortification plans.

In our model, the damage that a disruption inflicts on the network is measured by *shortage*, or the amount of unfulfilled demand. Because the interdictor's resources are limited, its attack aims to maximize that shortage. At the same time, the operator attempts to fortify the network (within a limited budget) by choosing appropriate arcs so as to protect (partial or entire) nominal arc capacities or by reinforcing them with secondary (and indestructible) capacities. Since demand is stochastic, it follows that shortage is a random variable. Our model allows the operator both to protect and to expand nominal arc capacities; hence the strategies for using protection resources are not only practical but also flexible. For instance: if closing a bridge could lead to traffic jams, then construction of a second bridge is probably imminent; in this way, then, the old bridge's capacity is expanded. Relocating a facility to a higher ground (to protect against flooding) and building a storm wall are examples of protecting a network's nominal arc capacities.

Other than being risk neutral, our model allows the decision maker to articulate a particular risk attitude via expected utility, which is the axiomatized choice preference associated with a rational agent (Neumann and Morgenstern 2007). To guide the decision maker's assessment of the trade-offs involved, a common approach is to assume (say) a family of utility functions and then to vary the risk attitude associated with those utilities, thereby obtaining a range of fortification plans as a function of risk attitudes. This approach ensures that the fortification plan is consistent with the choice of a rational decision maker. Observe that we can – by appropriately augmenting the underlying graph – transform problems for which *nodes* are the critical components to problems for which *arcs* are the critical components (cf. Corley and Chang 1974). Our linkage fortification model is general enough that its scope can encompass the corresponding location (or node) fortification model.

The rest of this paper is organized as follows. In Section 2, we describe the proposed fortification model under stochastic demand and adversarial attacks. Our model leads to a three-level optimization, and we show that the (inner) two-level minimax optimization problem can be converted to a mixed-integer problem. Section 3 applies a robust stochastic approximation approach (cf. Nemirovski et al. 2009) to find an ε-optimal solution of the model. Preliminary computational results are reported in Section 4, and we conclude in Section 5 with a brief summary. Technical proofs are given in Appendix A, and Appendix B extends our model to the case of continuous interdiction and a multi-commodity network flow.

2. A network fortification model

2.1. Network description

Here we consider a directed, single-commodity *transshipment network* $(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ represents the set of nodes and $\mathcal{A}$ the set of directed arcs. (In Section 2.3 we extend this model to a generic network.) We denote the set of (pure) transshipment nodes by $\mathcal{T}$, the set of demand nodes by $\mathcal{D}$, and the set of source nodes by $\mathcal{S}$. The sets $\mathcal{T}$, $\mathcal{D}$, and $\mathcal{S}$ are disjoint, and $\mathcal{N} = \mathcal{T} \cup \mathcal{D} \cup \mathcal{S}$. Let c_{ij} be the nominal capacity of the arc $(i,j) \in \mathcal{A}$. At each demand node $d \in \mathcal{D}$, the demand is a random variable denoted by ξ_d. Here ξ is a random vector consisting of the random variables ξ_d of demand nodes $d \in \mathcal{D}$.

Now we use the defender–attacker–defender sequential game to describe our fortification model. In a normal operating environment without interdiction, there exists a feasible flow $x = (x_{ij})_{(i,j)\in\mathcal{A}}$ along the network that channels single-commodity resources from the source nodes $\mathcal{S}$ to fulfil random demands at the demand nodes. In an event of an interdiction, however, some arcs may be eliminated and hence affect the network's ability to fulfill demand. From the operator's perspective, the network can be fortified by adding secondary capacity $y_{ij}^{(2)}$ to an arc $(i,j) \in \mathcal{A}$ or by protecting some or all of the nominal arc capacity. So if we let $y_{ij}^{(1)}$ be the capacity of the arc (i,j) to be protected, then $0 \le y_{ij}^{(1)} \le c_{ij}$. These fortified capacity, denoted by $y = \big(y_{ij}^{(1)}, y_{ij}^{(2)}\big)_{(i,j)\in\mathcal{A}}$, are assumed to be indestructible. We also assume that interdiction is binary in the sense that an arc under attack will lose all of its unfortified capacity. The new capacity of an arc $(i,j)\in\mathcal{A}$ is written as $\bar{c}_{ij}=(c_{ij}-y_{ij}^{(1)})(1-v_{ij})+y_{ij}^{(1)}+y_{ij}^{(2)}$. This formulation reflects that, if the arc (i,j) is attacked (i.e., if $v_{ij}=1$), then the protected extent $y_{ij}^{(1)}$ of the nominal arc capacity together with the additional arc capacity $y_{ij}^{(2)}$ are still operable; if the arc (i,j) is not attacked (i.e., if $v_{ij}=0$) then its total working capacity is $c_{ij}+y_{ij}^{(2)}$. In the *first* stage of the DAD sequential game, the defender/operator attempts to minimize the demand fulfillment shortfall that corresponds to a given realization ξ of demand over the network with capacity $\bar{c}=(\bar{c}_{ij})_{(i,j)\in\mathcal{A}}$ as follows (to streamline the notation, we use ξ to signify also the *realization* of the random variable representing the random demand):

$$\begin{aligned}\phi(\bar{c},\xi):=\min_x \sum_{d\in\mathcal{D}}&\max\left\{\xi_d-\sum_{i:(i,d)\in\mathcal{A}}x_{id},0\right\}\\ \text{s.t.}\ &\sum_{j:(i,j)\in\mathcal{A}}x_{ij}-\sum_{j:(j,i)\in\mathcal{A}}x_{ji}=0,\quad \forall i\in\mathcal{T},\\ &0\le x_{ij}\le\bar{c}_{ij},\qquad\qquad \forall (i,j)\in\mathcal{A},\end{aligned}\tag{1}$$

which can be reformulated as

$$\begin{aligned}\phi(\bar{c},\xi)=\sum_{d\in\mathcal{D}}\xi_d-\max_x&\sum_{d\in\mathcal{D}}\sum_{i:(i,d)\in\mathcal{A}}x_{id}\\ \text{s.t.}\quad &\sum_{j:(i,j)\in\mathcal{A}}x_{ij}-\sum_{j:(j,i)\in\mathcal{A}}x_{ji}=0,\ \forall i\in\mathcal{T},\\ &0\le x_{ij}\le\bar{c}_{ij},\qquad\qquad \forall (i,j)\in\mathcal{A},\\ &\sum_{i:(i,d)\in\mathcal{A}}x_{id}\le\xi_d,\qquad\quad \forall d\in\mathcal{D}.\end{aligned}\tag{2}$$

We assume that the interdictor is malicious and has knowledge of the fortified network capacity y. Hence, the interdiction would disrupt the network in a manner that maximizes the total demand shortfalls. We emphasize that the randomness we are considering in our fortification model is due to time; that is, there are different levels of demand at different times of the day. The effect of any operator decision is long-lasting, and the operator can repeat, under identical conditions, the experiments that led to her final decision. The time frame for making a decision is much shorter from the interdictor's perspective. So to ensure full protection of the network in the event of a worst-case event, we assume that the interdictor can observe the realization of demand uncertainty whereas the operator forms expectations about that uncertainty (i.e., the randomness of demand). We assume that the interdictor is operating within a budget, and hence can destroy at most R arcs. In the *second* stage of the DAD sequential game, the attacker/interdictor solves the following optimization problem to inflict the greatest disruption on the network:

$$\begin{aligned}
\rho(y,\xi) := \max_{v} \ & \phi(\bar{c},\xi) \\
\text{s.t.} \ & \bar{c}_{ij} = (c_{ij} - y^{(1)}_{ij})(1 - v_{ij}) + y^{(1)}_{ij} + y^{(2)}_{ij}, \ \forall (i,j) \in \mathcal{A}, \\
& \sum_{(i,j)\in\mathcal{A}} v_{ij} \le R, \\
& v_{ij} \in \{0,1\}, \qquad \forall (i,j) \in \mathcal{A}.
\end{aligned} \tag{3}$$

Thus, for a given decision about the fortification capacity y in an operating environment with interdiction, the demand shortfall associated with the demand realization ξ is given by $\rho(y,\xi)$. We assume that the operator has a budget M; we also assume that it costs $e^{(1)}_{ij}$ to implement a unit of protected capacity for the arc (i,j) and $e^{(2)}_{ij}$ to construct a unit of additional capacity for that arc. We evaluate the risk of shortfall by taking the expected disutility $\mathbb{E}\left[u\left(\rho\left(y,\xi\right)\right)\right]$, where $u(\cdot)$ is a convex, increasing disutility function. Then, in the *third* stage of the DAD sequential game, the defender/operator acts to mitigate the risk of network interdiction by solving the optimization problem

$$\begin{aligned}
\min_{y} \ & \mathbb{E}\left[u\left(\rho\left(y,\xi\right)\right)\right] \\
\text{s.t.} \ & \sum_{(i,j)\in\mathcal{A}} e^{(1)}_{ij} y^{(1)}_{ij} + \sum_{(i,j)\in\mathcal{A}} e^{(2)}_{ij} y^{(2)}_{ij} \le M, \\
& 0 \le y^{(1)}_{ij} \le c_{ij}, \qquad \forall (i,j) \in \mathcal{A}, \\
& 0 \le y^{(2)}_{ij}, \qquad \forall (i,j) \in \mathcal{A}.
\end{aligned} \tag{4}$$

2.2. A reformulation

We now solve Problem (4) by deriving an equivalent form that is more amenable to analysis. We start with the maximization problem stipulated in (2). It is a linear programming problem;

therefore, by linear programming duality, its dual problem is also a linear programming problem. The latter may be written as follows:

$$
\begin{aligned}
\min_{\alpha,\beta,\gamma} \ & \sum_{(i,j)\in\mathcal{A}} \bar{c}_{ij}\beta_{ij} + \sum_{d\in\mathcal{D}} \xi_d\gamma_d \\
\text{s.t.} \ & \alpha_i + \beta_{id} + \gamma_d \geq 1, && \forall (i,d)\in\mathcal{A}, i\in\mathcal{T}, d\in\mathcal{D}, \\
& \beta_{sd} + \gamma_d \geq 1, && \forall (s,d)\in\mathcal{A}, s\in\mathcal{S}, d\in\mathcal{D}, \\
& \alpha_i - \alpha_j + \beta_{ij} \geq 0, && \forall (i,j)\in\mathcal{A}, i,j\in\mathcal{T}, \\
& -\alpha_j + \beta_{sj} \geq 0, && \forall (s,j)\in\mathcal{A}, s\in\mathcal{S}, j\in\mathcal{T}, \\
& \beta_{ij} \geq 0, && \forall (i,j)\in\mathcal{A}, \\
& \gamma_d \geq 0, && \forall d\in\mathcal{D}.
\end{aligned}
\tag{5}
$$

We observe that Problem (5), together with the maximization over v in the second stage (see (3)), leads to a bilinear programming problem due to the product of v and β. We adapt the methodology established in Wood (1993) to prove the following lemma, which helps restrict the dual variables β in (5) to binary values. This restriction is used later to eliminate constraints and thereby reduce the dimension of Problem (20), which is crucial for obtaining our main result in Theorem 1 (see the proof of that theorem in Appendix A.2). The restriction will also be used to apply our linearization technique and thus reduce the dimension of Problem (20) for the case of *partial* interdiction (see Appendix B.1).

Proposition 1. *Problem* (5) *has an optimal solution such that* (a) β_{ij}, $(i,j)\in\mathcal{A}$, *takes value in* $\{0,1\}$, *and* (b) *the constraints*

$$
\begin{aligned}
\beta_{sd} + \gamma_d &\geq 1, \ \forall (s,d)\in\mathcal{A}, s\in\mathcal{S}, d\in\mathcal{D}, \\
-\alpha_j + \beta_{sj} &\geq 0, \ \forall (s,j)\in\mathcal{A}, s\in\mathcal{S}, j\in\mathcal{T},
\end{aligned}
$$

are satisfied with equality.

We can now use Proposition 1 to rewrite (5) as follows:

$$\begin{aligned}
\min_{\alpha,\beta,\gamma} \ & \sum_{(i,j)\in\mathcal{A}} \bar{c}_{ij}\beta_{ij} + \sum_{d\in\mathcal{D}} \xi_d\gamma_d \\
\text{s.t.} \ & \alpha_i + \beta_{id} + \gamma_d \geq 1, && \forall (i,d)\in\mathcal{A}, i\in\mathcal{T}, d\in\mathcal{D}, \\
& \beta_{sd} + \gamma_d = 1, && \forall (s,d)\in\mathcal{A}, s\in\mathcal{S}, d\in\mathcal{D}, \\
& \alpha_i - \alpha_j + \beta_{ij} \geq 0, && \forall (i,j)\in\mathcal{A}, i,j\in\mathcal{T}, \\
& -\alpha_j + \beta_{sj} = 0, && \forall (s,j)\in\mathcal{A}, s\in\mathcal{S}, j\in\mathcal{T}, \\
& 0\leq \alpha_i \leq 1, && \forall i\in\mathcal{T}, \\
& 0\leq \gamma_d \leq 1, && \forall d\in\mathcal{D}, \\
& \beta_{ij}\in\{0,1\}, && \forall (i,j)\in\mathcal{A}.
\end{aligned} \tag{6}$$

We use Θ to denote the feasible set that contains all feasible (α,β,γ) of Problem (6), and rewrite $\phi(\bar{c},\xi)$ in (2) as follows

$$\phi(\bar{c},\xi) = \max_{(\alpha,\beta,\gamma)\in\Theta} \left\{ \sum_{d\in\mathcal{D}} \xi_d(1-\gamma_d) - \sum_{(i,j)\in\mathcal{A}} \bar{c}_{ij}\beta_{ij} \right\}.$$

Next we denote

$$Y := \left\{ y = \left(y_{ij}^{(1)}, y_{ij}^{(2)}\right)_{(i,j)\in\mathcal{A}}, 0\leq y_{ij}^{(1)} \leq c_{ij}, 0\leq y_{ij}^{(2)}, \sum_{(i,j)\in\mathcal{A}} e_{ij}^{(1)} y_{ij}^{(1)} + \sum_{(i,j)\in\mathcal{A}} e_{ij}^{(2)} y_{ij}^{(2)} \leq M \right\} \tag{7}$$

as the set of feasible fortification decisions, and

$$V := \left\{ v = (v_{ij})_{(i,j)\in\mathcal{A}}, v_{ij}\in\{0,1\}, \sum_{(i,j)\in\mathcal{A}} v_{ij} \leq R \right\}, \tag{8}$$

as the set of feasible interdiction decisions. Then Problem (4) can be written in the following more simplified form:

$$\min_{y\in Y} \mathbb{E}\left[u\left(\max_{v\in V} \max_{(\alpha,\beta,\gamma)\in\Theta} \left(\sum_{d\in\mathcal{D}} \xi_d(1-\gamma_d) - \sum_{(i,j)\in\mathcal{A}} \left((c_{ij} - y_{ij}^{(1)})(1-v_{ij}) + y_{ij}^{(1)} + y_{ij}^{(2)} \right) \beta_{ij} \right) \right) \right]. \tag{9}$$

We are now ready to state Theorem 1 which provides a stochastic programming formulation to solve Problem (9).

THEOREM 1. *The optimal solution y^* of the following problem solves Problem* (9):

$$\min_{y\in Y} \mathbb{E}\big[u(g(y,\xi))\big], \tag{10}$$

where

$$\begin{aligned}
g(y,\xi) = \max_{v,\alpha,\gamma,\eta} & \sum_{d\in\mathcal{D}} \xi_d(1-\gamma_d) - \sum_{(i,j)\in\mathcal{A}} (c_{ij}+y^{(2)}_{ij})\eta_{ij} + (y^{(1)}_{ij}+y^{(2)}_{ij})v_{ij} \\
\text{s.t.} \quad & \alpha_j + \eta_{jd} + v_{jd} + \gamma_d \geq 1, && \forall (j,d)\in\mathcal{A}, j\in\mathcal{T}, d\in\mathcal{D}, \\
& \eta_{sd} + v_{sd} + \gamma_d = 1, && \forall (s,d)\in\mathcal{A}, s\in\mathcal{S}, d\in\mathcal{D}, \\
& \alpha_i - \alpha_j + \eta_{ij} + v_{ij} \geq 0, && \forall (i,j)\in\mathcal{A}, i,j\in\mathcal{T}, \\
& -\alpha_j + \eta_{sj} + v_{sj} = 0, && \forall (s,j)\in\mathcal{A}, s\in\mathcal{S}, j\in\mathcal{T}, \\
& \sum_{(i,j)\in\mathcal{A}} v_{ij} \leq R, \\
& 0\leq \alpha_i \leq 1, && \forall i\in\mathcal{T}, \\
& 0\leq \gamma_d \leq 1, && \forall d\in\mathcal{D}, \\
& \eta_{ij}, v_{ij} \in \{0,1\}, && \forall (i,j)\in\mathcal{A}.
\end{aligned} \tag{11}$$

2.3. Extension to generic network flow problem

In this section we consider a directed, single-commodity *generic network* flow in which a demand node can also be used to transfer the commodity. We use the same notation as in Section 2.1 for the network attributes. We assume that the network's source nodes are "pure"; that is, the commodity cannot be transferred via a source node. The demand fulfillment shortfall for a given realization ξ of demand over the generic network with capacity $\bar{c} = (\bar{c}_{ij})_{(i,j)\in\mathcal{A}}$ is now defined by

$$\begin{aligned}
\phi_G(\bar{c},\xi) := \min_x & \sum_{d\in\mathcal{D}} \max\left\{\xi_d - \left(\sum_{i:(i,d)\in\mathcal{A}} x_{id} - \sum_{j:(d,j)\in\mathcal{A}} x_{dj}\right), 0\right\} \\
\text{s.t.} \quad & \sum_{j:(i,j)\in\mathcal{A}} x_{ij} - \sum_{j:(j,i)\in\mathcal{A}} x_{ji} = 0, && \forall i\in\mathcal{T}, \\
& \sum_{i:(i,d)\in\mathcal{A}} x_{id} - \sum_{j:(d,j)\in\mathcal{A}} x_{dj} \geq 0, && \forall d\in\mathcal{D}, \\
& 0\leq x_{ij} \leq \bar{c}_{ij}, && \forall (i,j)\in\mathcal{A}.
\end{aligned} \tag{12}$$

Next we define the demand shortfall $\rho(y,\xi)$ for a given fortification decision y as in (3) and derive the fortification problem for the generic network as in (4). Formally, we have

$$\min_{y\in Y}\mathbb{E}\Big[u\Big(\max_{v\in V}\phi_G\Big(\Big[(c_{ij}-y^{(1)}_{ij})(1-v_{ij})+y^{(1)}_{ij}+y^{(2)}_{ij}\Big]_{(i,j)\in\mathcal{A}},\xi\Big)\Big)\Big]; \tag{13}$$

here Y and V are defined as in (7) and (8), respectively. Recall that Problem (9) can be rewritten in terms of (1)'s function ϕ as follows:

$$\min_{y\in Y}\mathbb{E}\Big[u\Big(\max_{v\in V}\phi\Big(\Big[(c_{ij}-y^{(1)}_{ij})(1-v_{ij})+y^{(1)}_{ij}+y^{(2)}_{ij}\Big]_{(i,j)\in\mathcal{A}},\xi\Big)\Big)\Big].$$

We now construct a counterpart transshipment network – with respect to which we can deploy the corresponding fortification problem in Section 2.1 – and show that the fortification Problem (13) of the generic network can be solved by solving the fortification problem of the corresponding transshipment network. More specifically, for each demand node $d\in\mathcal{D}$ we introduce a corresponding fictitious demand node d' and then connect d and d' to create a directed arc (d,d'). Let $\mathcal{D}'$ denote the set of all fictitious demand nodes d' and let $\mathcal{A}'$ be the set of all new arcs (d,d'). We define a new transshipment network $\Big(\hat{\mathcal{N}},\hat{\mathcal{A}}\Big)$, where the set of nodes $\hat{\mathcal{N}}=\mathcal{N}\cup\mathcal{D}'$ and the set of arcs $\hat{\mathcal{A}}=\mathcal{A}\cup\mathcal{A}'$. In the new network, the set of pure source nodes is still $\mathcal{S}$, the set of pure transshipment nodes $\hat{\mathcal{T}}$ is defined by $\hat{\mathcal{T}}=\mathcal{T}\cup\mathcal{D}$ and the set of pure demand nodes is $\hat{\mathcal{D}}=\mathcal{D}'$. Let $y^{(2)}_{\max}=\max_{y\in Y,(i,j)\in\mathcal{A}}y^{(2)}_{ij}$; that is, $y^{(2)}_{\max}$ is the most capacity that can be added to an arc of the network $(\mathcal{N},\mathcal{A})$.

The arc capacities of the new network are defined as follows. The capacity of arc $(i,j)\in\hat{\mathcal{A}}$ is $\hat{c}_{ij}=c_{ij}$ if $(i,j)\in\mathcal{A}$, and $\hat{c}_{dd'}=\sum_{i:(i,d)\in\mathcal{A}}\big(c_{id}+y^{(2)}_{\max}\big)$ for $(d,d')\in\mathcal{A}'$. The cost to fortify the artificial arc (d,d') is $\hat{e}^{(1)}_{dd'}=\hat{e}^{(2)}_{dd'}=0$ (i.e., the arc (d,d') is always fortified). We let

$$\hat{V}=\{\hat{v}=(\hat{v}_{ij})_{(i,j)\in\hat{\mathcal{A}}},\hat{v}_{ij}\in\{0,1\},\sum_{(i,j)\in\hat{\mathcal{A}}}\hat{r}_{ij}\hat{v}_{ij}\le R\};$$

here $\hat{r}_{ij}=1$ if $(i,j)\neq(d,d')$ and $\hat{r}_{dd'}=R+1$. This reflects that the attacker can destroy at most R arcs, and the cost to destroy the arc (d,d') is so great that the arc can not be destroyed by the interdictor. We deploy the fortification model described in Section 2.1 for our new transshipment network $(\hat{\mathcal{N}},\hat{\mathcal{A}})$ and express Problem (4) with regard to $(\hat{\mathcal{N}},\hat{\mathcal{A}})$ as follows:

$$\min_{\hat{y}\in\hat{Y}}\mathbb{E}\Big[u\Big(\max_{\hat{v}\in\hat{V}}\hat{\phi}\Big(\Big[(\hat{c}_{ij}-\hat{y}^{(1)}_{ij})(1-\hat{v}_{ij})+\hat{y}^{(1)}_{ij}+\hat{y}^{(2)}_{ij}\Big]_{(i,j)\in\hat{\mathcal{A}}},\xi\Big)\Big)\Big]. \tag{14}$$

Here

$$\begin{aligned}\hat{\phi}(\bar{c},\xi) = \min_{\hat{x}} & \sum_{d'\in\hat{\mathcal{D}},(d,d')\in\hat{\mathcal{A}}} \max\left\{\xi_{d'} - \hat{x}_{dd'}, 0\right\} \\ \text{s.t.} & \sum_{j:(i,j)\in\hat{\mathcal{A}}} \hat{x}_{ij} - \sum_{j:(j,i)\in\hat{\mathcal{A}}} \hat{x}_{ji} = 0, \quad \forall i \in \hat{\mathcal{T}}, \\ & 0 \le \hat{x}_{ij} \le \bar{c}_{ij}, \quad \forall (i,j)\in\hat{\mathcal{A}},\end{aligned} \tag{15}$$

and

$$\hat{Y} = \left\{ \hat{y} = \left(\hat{y}^{(1)}_{ij}, \hat{y}^{(2)}_{ij}\right)_{(i,j)\in\hat{\mathcal{A}}}, 0 \le \hat{y}^{(1)}_{ij} \le \hat{c}_{ij}, 0 \le \hat{y}^{(2)}_{ij}, \, \hat{y}^{(2)}_{dd'} \le \hat{c}_{dd'}, \sum_{(i,j)\in\hat{\mathcal{A}}} \hat{e}^{(1)}_{ij}\hat{y}^{(1)}_{ij} + \sum_{(i,j)\in\hat{\mathcal{A}}} \hat{e}^{(2)}_{ij}\hat{y}^{(2)}_{ij} \le M \right\}.$$

For each arc $(d,d')\in\hat{\mathcal{A}}$ directed to the pure demand node $d'\in\hat{\mathcal{D}}$, it follows from $\hat{r}_{dd'} = R+1$ that $\hat{v}_{dd'} = 0$. In the definition of $\hat{Y}$ we include the constraint $\hat{y}^{(2)}_{dd'} \le \hat{c}_{dd'}$ so that $\hat{Y}$ will be a compact set. Our next proposition shows that we can solve Problem (13) by solving the corresponding Problem (14). See its proof in Appendix A.3.

Proposition 2. *Let $\hat{y}^* = \left(\hat{y}^{*(1)}_{ij}, \hat{y}^{*(2)}_{ij}\right)_{(i,j)\in\hat{\mathcal{A}}}$ be an optimal solution of Problem (14). For $(i,j)\in\mathcal{A}$, we set $y^{*(1)}_{ij} = \hat{y}^{*(1)}_{ij}$ and $y^{*(2)}_{ij} = \hat{y}^{*(2)}_{ij}$. Then $y^* = \left(y^{*(1)}_{ij}, y^{*(2)}_{ij}\right)_{(i,j)\in\mathcal{A}}$ is an optimal solution of Problem (13).*

3. Robust stochastic approximation

Put $F(y,\xi) = u(g(y,\xi))$. Then Problem (10) can be rewritten as

$$\min_{y\in Y} \left\{\mathbb{E}[F(y,\xi)] := f(y)\right\}. \tag{16}$$

Stochastic approximation and sample average approximation are two approaches often used to solve a stochastic optimization problem. A short introduction to these methods – and an insightful comparison between them – may be found in Nemirovski et al. (2009). In that paper, a modified stochastic approximation, which the authors refer to as *robust* stochastic approximation, is proposed and shown to perform as well or even better than the sample average approximation for a certain class of stochastic problems. In this modified approach, the number of iterations is determined before running the algorithm and a "constant step size" policy is used when updating the iteration. In Proposition 3 (see its proof in Appendix A.4), we establish all the results needed for applying the robust stochastic approximation algorithm (as in Nemirovski et al. 2009, Sec. 2.2) to solve Problem (10) numerically.

Proposition 3. *Let $(v^*, \alpha^*, \gamma^*, \eta^*,)$ be an optimal solution of the maximization problem in* (11). *Then the following statements hold.*

1. *The function $y \mapsto F(y,\xi)$ is convex.*
2. *Let*

$$F_1'(y,\xi) = -u'(g(y,\xi)) \left(\left[-v_{ij}^*\right]_{(i,j)\in\mathcal{A}}, \left[-\eta_{ij}^* - v_{ij}^*\right]_{(i,j)\in\mathcal{A}} \right), \tag{17}$$

where $u'(g(y,\xi))$ is a subgradient of $u(\cdot)$ at $g(y,\xi)$; then $F_1'(y,\xi)$ is a subgradient of the map $y \mapsto F(y,\xi)$ at y.

We are now ready to present the robust stochastic approximation method for solving Problem (10). Algorithm 1 gives a full description of the method. Proposition 3 shows that, as long as the utility function is convex and increasing, Problem (10) is a convex optimization problem that can be efficiently solved via Algorithm 3.1. It follows that the particular utility functions chosen should not affect the ease of computation. The main difficulty lies in solving Problem (11) to evaluate the subgradients. The rest of this section is devoted to illustrative remarks concerning Algorithm 1.

Algorithm 1 starts with the initial point $y_{(1)}$ in Y. The total number Q of the algorithm's iterations must be determined in advance. In order to calculate Q, we estimate an upper bound D_Y for the distance from $y_{(1)}$ to the points of Y and then calculate an upper bound Ω for the subgradient $F_1'(y,\xi)$. Because the set Y is bounded, we can choose D_Y to be any number greater than the diameter of Y. We can also choose the tightest bound $D_Y = \max_{y\in Y} \left\|y_{(1)} - y\right\|_2$, which is a quadratic programming problem. To estimate the value of Ω, we first observe from Equation (17) that $|v_{ij}^*| \le 1$ and $|\eta_{ij}^*| \le 1$ for all $(i,j) \in \mathcal{A}$. Furthermore, if we assume that $\xi_d \le \mathbf{B}$ for all $d \in \mathcal{D}$ then the function $g(y,\xi)$ – as defined in (11) – is also bounded by $|\mathcal{D}|\mathbf{B}$, where $|\mathcal{D}|$ is the number of sink nodes. Hence it should be easy to estimate the value of Ω if we assume that the subgradient $u'(\cdot)$ is bounded on

Input: the desired accuracy of solution ε.

Initialization: Choose $y_{(1)} \in Y$. Estimate $D_Y \geq \max_{y\in Y} \left\| y_{(1)} - y \right\|_2$.

Let Ω be a constant such that $\mathbb{E}\big[\, \|F_1'(y,\xi)\|_2^2 \,\big] \leq \Omega^2$ for all $y \in Y$.

Find the number of iterations Q such that $\frac{\Omega D_Y}{\sqrt{Q}} \leq \varepsilon$.

Set the step size $\lambda = \frac{D_Y}{\Omega\sqrt{Q}}$.

for $l = 2,\dots,Q$ **do**

Update

$$y_{(l)} := \Pi_Y(y_{(l-1)} - \lambda F_1'(y_{(l-1)}, \xi_{(l-1)})),$$

where $F_1'(y,\xi)$ is defined as in (17) and Π_Y is the projection onto Y.

end for

Output: $\tilde{y}_{(Q)} = \frac{1}{Q}\sum_{i=1}^{Q} y_{(i)}$.

Algorithm 1: A robust stochastic approximation method

its bounded domain. This assumption is satisfied, for example, with regard to the utility functions (piecewise linear, quadratic, and exponential) considered in Section 4.

We assume that it is possible to generate an independent and identically distributed sample $\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(Q)}$ of realizations of the random vector ξ. At Step $l = 2,\dots,Q$, Algorithm 1 finds the subgradient $F_1'(y_{(l-1)}, \xi_{(l-1)})$, as defined by (17), and applies the constant–step size policy $\lambda_{(l)} = \lambda = D_Y/(\Omega\sqrt{Q})$ to update the new iteration $y_{(l)}$. It is worth mentioning that classical stochastic approximation algorithms often use the step size policy $\lambda_{(l)} = \varsigma/l$, where ς is some constant. However, that policy (which is decreasing with respect to l) may result in an extremely slow convergence when the objective function is not strongly convex. The constant step size used in Algorithm 1 improves the convergence performance of the stochastic approximation algorithm, rendering it both robust and eminently applicable to general convex objectives. Note also that this constant step size guarantees the best convergence rate for the algorithm in terms of D_Y and Ω (see Nemirovski et al. 2009, Sec. 2.2).

The output of Algorithm 1 is $\tilde{y}_{(Q)} = \frac{1}{Q}\sum_{i=1}^{Q} y_{(i)}$, where the $\left\{y_{(i)}\right\}_{1\le i\le Q}$ are generated by Algorithm 1. Let y^* be an optimal solution of (16). Proposition 3 allows us to prove (cf. Nemirovski et al. 2009, Sec. 2.2) that

$$\mathbb{E}[f(\tilde{y}_{(Q)})] - \mathbb{E}[f(y^*)] \le \frac{\Omega D_Y}{\sqrt{Q}} \le \varepsilon.$$

Thus the algorithm identifies an ε-optimal solution to Problem (10).

4. Preliminary computational results

In this section, we report the results of our computational experiments using Algorithm 1 to solve (10). We implement the algorithm in Python 3.4 on a desktop computer with an Intel(R) Core(TM) i5-4690 CPU running at 3.50 GHz with 16 GB of RAM in a Windows 8.1 Enterprise (64-bit) environment. We use CPLEX Studio 1262 to solve the mixed-integer linear programming problem of (11). In the experiments we consider the set Y with $e_{ij}^{(1)} = 1$ and $e_{ij}^{(2)} = 3$. That is, implementing a unit of additional capacity to the arc (i,j) costs 3 times as much as protecting a unit of nominal capacity. We consider three types of utility functions (UFs), as follow:

1. Piece-wise linear utility function $u_1(z) = \max\{a_1 + b_1 z, \ldots, a_m + b_m z\}$.
2. Quadratic utility function $u_2(z) = az^2$.
3. Exponential utility function $u_3(z) = \frac{1-e^{-az}}{a}, a < 0$.

The network capacities are integer values that are independently and uniformly generated in $\{1,\ldots,10\}$. To generate the demands ξ_d $(d \in \mathcal{D})$ of sink nodes, we use normal distributions for which (a) the variances are random integers distributed uniformly in $[1,6]$ and (b) the means are positive values that are no less than the maximum flows to each sink node in the “vanilla” case of no interdiction or fortification. We study the effect of fortifying the network under various fortification budgets M. For a sample size fixed at $Q = 10{,}000$, we assess the out-of-sample performance corresponding to different values of M. To find an approximate optimal solution y of Problem (10), we generate a sample set (comprising 10,000 elements) of ξ to estimate the objective value.

In the *first* experiment, we test eight networks whose topologies are taken from SNDlib, the library of survivable fixed telecommunication network design (see Orlowski et al. 2007). The first

Table 1 Network information.

Network	$\|\mathcal{S}\|$	$\|\mathcal{D}\|$	$\|\mathcal{N}\|$	$\|\mathcal{A}\|$
abilene	3	3	12	15
nobel-us	6	5	14	21
atlanta	4	4	15	22
france	3	9	25	45
cost	12	10	37	57
germany50	9	5	50	88
pioro40	7	7	40	89
ta2	11	16	65	108
polblogs	500	160	1,490	19,017
network1	1,000	50	3,050	5,999
network2	500	200	10,700	29,999

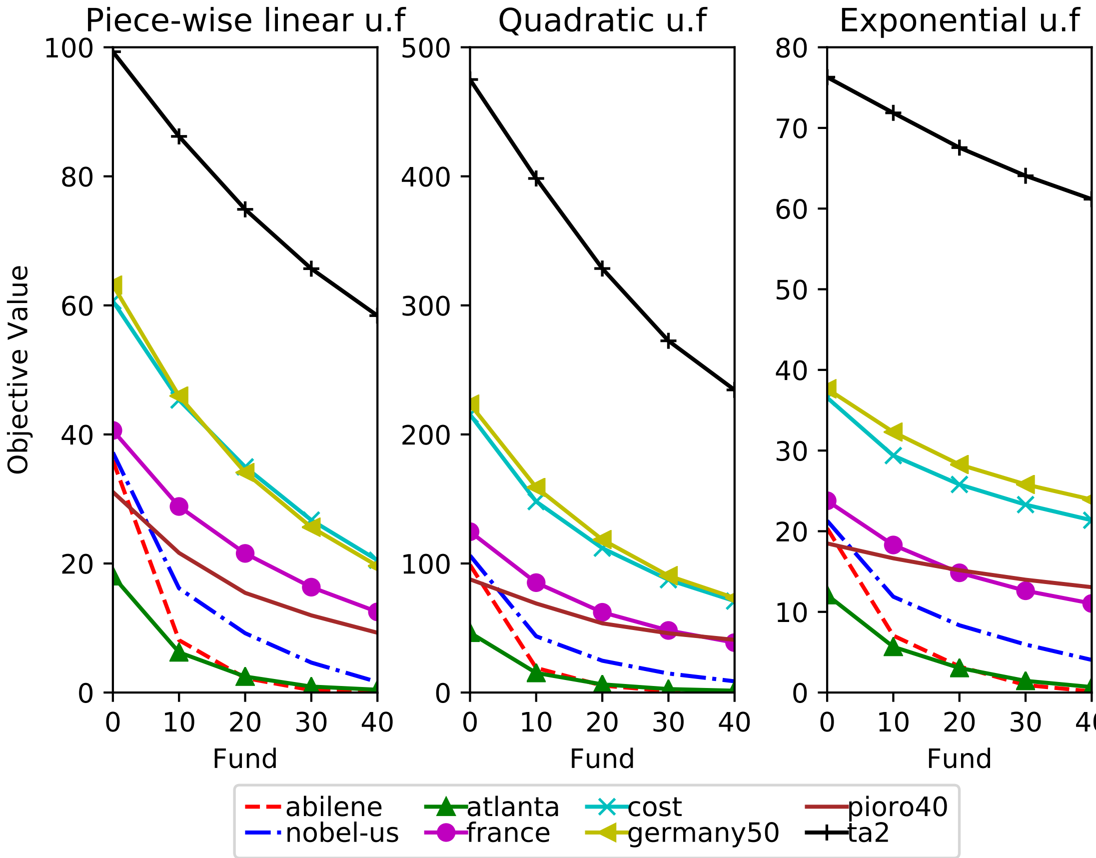
half of Table 1 provides information on these networks. As before, we use $|\mathcal{S}|$, $|\mathcal{D}|$, $|\mathcal{N}|$, and $|\mathcal{A}|$ to denote (respectively) the number of source nodes, the number of sink nodes, the total number of nodes, and the number of edges. We put $V = \big\{v : v_{ij} \in \{0,1\}, \sum_{(i,j)\in\mathcal{A}} v_{ij} \leq 3\big\}$; that is, at most 3 arcs are attacked. Let M take any value in $\{0, 10, 20, 30, 40\}$. We test for 3 utility functions $u_1(z) = \max\{\frac{1}{4}z, \frac{1}{2}z - 3, 2z - 4, 4z - 20\}$, $u_2(z) = \frac{1}{2}z^2$ and $u_3(z) = \frac{1-e^{0.05z}}{(-0.05)}$. The results of this first experiment are illustrated in Figure 1, where the horizontal axes represent the fund (budget) M and the vertical axes represent the corresponding objective value – that is, the unmet demand.

The sudden decline in unmet demand, as shown by the steep curves, clearly establishes that fortification is extremely effective. The numerical results serve also to indicate how vulnerable or stable the networks are. Thus, for instance, pioro40 is the most vulnerable network – especially in the case of an exponential utility function – because it exhibits, when fortified, the lowest decreasing rate of unmet demand. Although ta2 has the greatest number of arcs, it is quite stable because its corresponding rate (in response to fortification) falls dramatically for all three utility functions.

In the *second* experiment we test higher-dimensional networks, information for which is given in the second half of Table 1. The topology of polblogs is taken from a network data [1], whereas the topologies of network1 and network2 are randomly generated. Here the fund M takes a value in $\{0, 50, 100, 500, 1000\}$, and we consider the interdiction set $V = \big\{v : v_{ij} \in \{0,1\},\ \sum_{(i,j)\in\mathcal{A}} v_{ij} \leq 5\big\}$; that is, at most five arcs are attacked. Again we test for three utility functions; the piecewise linear

[1] We use the data from Mark Newman, Network data, URL http://www-personal.umich.edu/ mejn/netdata/

Figure 1 Effect of fortification.



UF $u_1(z)$ and the quadratic UF $u_2(z)$ are the same as in the first experiment, but the exponential UF is now $u_3(z) = (1 - e^{0.005z})/-0.005$ so that we do not end up with an excessively large objective function. Results of the second experiment are given in Table 2. The values reported in Table 2 show that, although network2 has the largest number of arcs among the three tested networks, we can reduce nearly 80% of interdictor's disruption if we fortify it by spending the protective budget $M = 500$. Note that it would be much more expensive to obtain the same level of protection for polblogs and network1.

The computational time for our experiments is a topic that merits some attention. For each network and utility function, we calculate the average time required to run Algorithm 1 over the five values of the fortification budget M; these averages, in seconds of CPU time, are reported in Table 3. We observe that the running time of Algorithm 1 is not monotonic with respect to the number of arcs. For example: we need about one hour (on average) for network2 but network1 requires nearly

Table 2 Approximate objective values (unmet demand) for higher-dimensional networks under various fortification funds (budgets) M

Network	$M=0$	$M=10$	$M=50$	$M=500$	$M=1,000$
Piece-wise linear utility function, $u_1(z)=\max\{\frac{1}{4}z, \frac{1}{2}z-3, 2z-4, 4z-20\}$					
polblogs	387.6540	361.0258	346.2739	277.8344	236.1309
network1	5,356.0431	5,298.8556	5,251.3110	4,880.1810	4,417.3026
network2	907.0864	810.3345	710.3369	186.0351	168.9187
Quadratic utility function, $u_2(z)=\frac{1}{2}z^2$					
polblogs	5,287.6263	5,190.3246	5,172.3757	5,035.8275	4,889.7812
network1	903,558.384	901,668.752	900,897.975	894,645.096	887,705.561
network2	27,113.4182	21,843.5698	16,947.9226	1,870.7179	1,825.7641
Exponential utility function, $u_3(z)=\frac{1-e^{0.005z}}{(-0.005)}$					
polblogs	133.7050	132.4480	132.4313	132.3626	132.3616
network1	167,140.001	166,377.746	166,359.235	166,260.562	165,970.806
network2	441.3497	369.8770	302.9794	72.3556	72.3363

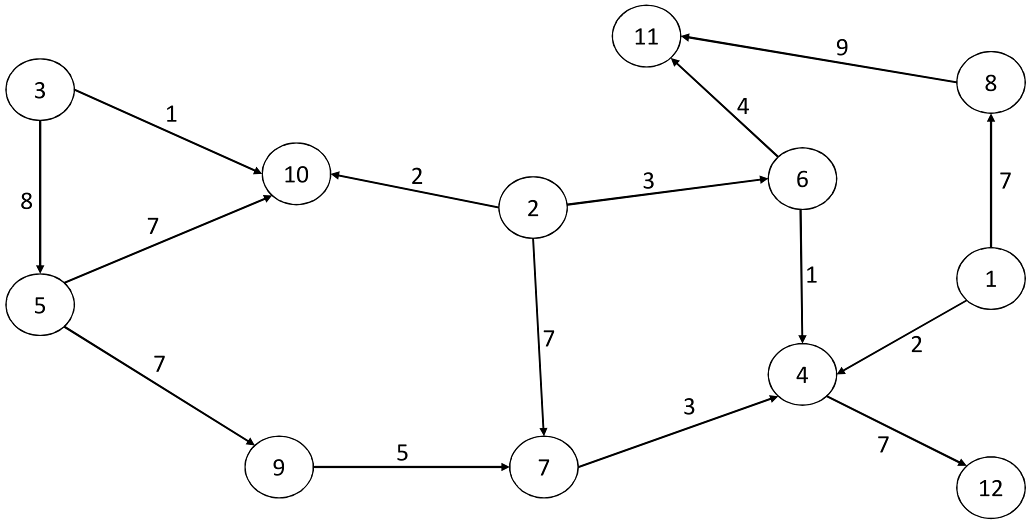
Table 3 Average running time (CPU seconds) of Algorithm 1

Network	Piece-wise linear	Quadratic	Exponential
abilene	135.71	129.55	140.05
nobel-us	147.99	145.67	157.24
atlanta	136.00	136.40	136.77
france	153.40	152.20	155.67
cost	156.07	155.37	158.13
germany50	199.30	197.96	196.40
pioro40	222.22	214.54	210.20
ta2	212.31	213.29	208.66
polblogs	23,031.20	22,498.27	22,550.47
network1	22,963.83	23,139.65	23,183.69
network2	3,721.26	3,734.76	3,601.65

six hours even though it has 5 times less arcs than network2. We remark that a network protection scheme should be carefully considered and thoroughly planned prior to implementation. Hence the algorithm's running time needed to solve the fortification model is far less critical than its stability. Although Algorithm 1's running time is seldom short, the algorithm is extremely stable in our numerical experiments for both medium- and large-size networks.

Finally, we illustrate the effect of utility functions on the fortification solutions by conducting an experiment on network3. The attributes of this network, which has the same topology as the abilene network, are illustrated in Figure 2. We test for the same utility functions as in the first experiment, and assume that at most 3 arcs are attacked. To simplify the presentation – and to demonstrate our proposed model's flexibility – we consider a scenario in which building additional

Figure 2 Attributes of network3.



(indestructible) capacity makes more sense than protecting the nominal capacity of network3's arcs. Examples of such a scenario include transferring the commodity via aircraft, which would render that transport impregnable to road-based attacks, and constructing underground power lines for an electricity grid, which would negate the attempts of interdictors targeting aerial power lines. In this case, we simply let $e_{ij}^{(1)} = 0$ and $y_{ij}^{(1)} = 0$. The operator needs to distribute the fortification budget $M = 30$ across the network. The solutions to this problem, $y = \left(y_{ij}^{(2)}\right)_{(i,j)\in\mathcal{A}}$, are reported in Table 4. The table's first column lists the arcs of network3, and the next three columns give an optimal fortification budget for each arc depending on the type of utility function.

We observe that different utility functions result in fortification plans whose budgets are distributed differently. Thus the first and second operators, whose respective risk attitudes are captured by a piecewise linear and a quadratic utility function, spend similar amounts of the budget

Table 4 A fortification solution

	Utility function		
Arcs	Piece-wise linear	Quadratic	Exponential
$(1,8)$	4.5226	4.9232	4.8287
$(1,4)$	4.0264	3.7211	2.7153
$(2,10)$	4.5218	4.4835	4.0041
$(2,6)$	0.0346	0.0241	0.0730
$(2,7)$	0.0007	0.0000	0.0032
$(3,10)$	4.5218	4.4835	4.0042
$(3,5)$	0.0027	0.0019	0.0034
$(4,12)$	6.8460	6.5812	5.8809
$(5,9)$	0.0063	0.0010	0.6175
$(5,10)$	0.0020	0.0014	0.0024
$(6,11)$	0.0343	0.0239	0.0729
$(6,4)$	1.1092	1.2562	1.7377
$(7,4)$	0.9353	0.9441	1.4162
$(8,11)$	3.4294	3.5534	4.0225
$(9,7)$	0.0062	0.0009	0.6176

when fortifying the arcs $(1,8)$, $(1,4)$, $(2,10)$, $(3,10)$, $(4,12)$, and $(8,11)$; in contrast, the third operator – whose risk attitude is captured by an exponential utility function spends less of the budget on arcs $(1,4)$, $(2,10)$, $(3,10)$, and $(4,12)$ while spending more on arcs $(6,4)$, $(7,4)$, and $(8,11)$. The first operator practically ignores the arc $(2,7)$, and the second operator spends hardly any of the budget on arcs $(2,7)$ and $(9,7)$; yet the third operator devotes nontrivial amounts of the budget to all arcs, although $(2,7)$, $(3,5)$, and $(5,10)$ receive the least fortification funds.

5. Conclusion

In this paper we proposed a preventive approach to achieving network resilience by allocating resources – to build additional capacity or protect extant capacity – and thus fortifying the network against interdiction/disruption. Our model accounts simultaneously for three components: adversarial disruption, stochastic demand, and the decision maker's attitude toward risk. The result is a three-level and nonlinear stochastic optimization problem. We show that such a problem can be solved, in theory and also in numerical experiments, via the robust stochastic approximation method proposed by Nemirovski et al. (2009); under this approach, in each iteration the subgradient to be evaluated can be obtained by solving a linear mixed-integer problem. Our computational results firmly establish the critical importance of fortification. In an appendix, we extend the model to continuous interdiction and a multi-commodity transportation network.

Acknowledgement

We are grateful to the anonymous reviewers for their meticulous comments and insightful suggestions. L. T. K. Hien gives special thanks to Prof. W. B. Haskell for his support.

References

Alderson DL, Brown GG, Carlyle WM, Wood RK (2011) Solving defender-attacker-defender models for infrastructure defense. Wood RK, Dell RF, eds., *Proceedings of the 12th INFORMS Computing Society Conference: Research, Computing, and Homeland Defense*, 28–49 (INFORMS).

Ali I, Barnett D, Farhangian K, Kennington J, Patty B, Shetty B, McCarl B, Wong P (1984) Multicommodity network problems: Applications and computations. *IIE Transactions* 16(2):127–134.

Baghalian A, Rezapour S, Farahani RZ (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.

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

Brown GG, Carlyle WM, Salmerón J, Wood K (2005) Analyzing the vulnerability of critical infrastructure to attack and planning defenses. Greenberg HJ, ed., *Emerging Theory, Methods, and Applications*, 102–123 (INFORMS).

Carr RD, Greenberg HJ, Hart WE, Konjevod G, Lauer E, Lin H, Morrison T, Phillips CA (2006) Robust optimization of contaminant sensor placement for community water systems. *Mathematical Programming* 107(1):337–356.

Corley JHW, Chang H (1974) Finding the n most vital nodes in a flow network. *Management Science* 21(3):362–364.

Folie M, Tiffin J (1976) Solution of a multi-product manufacturing and distribution problem. *Management Science* 23(3):286–296.

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

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

Gendron B, Crainic TG, Frangioni A (1999) Multicommodity capacitated network design. Sansò B, Soriano P, eds., *Telecommunications Network Planning*, 1–19 (Boston, MA: Springer US).

Geoffrion AM, Graves GW (1974) Multicommodity distribution system design by benders decomposition. *Management Science* 20(5):822–844.

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

James J, Salhi S (2002) A tabu search heuristic for the location of multi-type protection devices on electrical supply tree networks. *Journal of Combinatorial Optimization* 6(1):81–98.

Laporte G, Louveaux FV, van Hamme L (1994) Exact solution to a location problem with stochastic demands. *Transportation Science* 28(2):95–103.

Liberatore F, Scaparra MP, Daskin MS (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, project Management and Scheduling.

McCormick G (1976) Computability of global solutions to factorable nonconvex programs: Part i - convex underestimating problems. *Mathematical Programming* 10(1):147–175.

Nemirovski A, Juditsky A, Lan G, Shapiro A (2009) Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization* 19(4):1574–1609.

Neumann JV, Morgenstern O (2007) *Theory of Games and Economic Behavior (commemorative edition)* (Princeton, NJ: Princeton university press).

Olinick EV, Rosenberger JM (2008) Optimizing revenue in cdma networks under demand uncertainty. *European Journal of Operational Research* 186(2):812–825.

Orlowski S, Pióro M, Tomaszewski A, Wessäly R (2007) SNDlib 1.0–Survivable Network Design Library. *Proceedings of the 3rd International Network Optimization Conference (INOC 2007), Spa, Belgium*, URL `http://sndlib.zib.de`, extended version accepted in Networks, 2009.

Riis M, Andersen KA (2002) Capacitated network design with uncertain demand. *INFORMS Journal on Computing* 14(3):247–260.

Salmeron J, Wood K, Baldick R (2004) Analysis of electric grid security under terrorist threat. *Power Systems, IEEE Transactions on* 19(2):905–912.

Santoso T, Ahmed S, Goetschalckx M, Shapiro A (2005) A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research* 167(1):96–115.

Scaparra MP, Cappanera P (2005) Optimizing security investments in transportation and telecommunication networks. INFORMS Annual Meeting 2005, San Francisco, CA.

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

Scaparra MP, Church RL (2012) Protecting supply systems to mitigate potential disaster: A model to fortify capacitated facilities. *International Regional Science Review* 35(2):188–210.

Sen S, Doverspike RD, Cosares S (1994) Network planning with random demand. *Telecommunication Systems* 3(1):11–30.

Shapiro A, Dentcheva D, Ruszczynski A (2014) *Lectures on Stochastic Programming: Modeling and Theory, Second Edition* (Philadelphia, PA, USA: Society for Industrial and Applied Mathematics).

Snyder LV, Scaparra MP, Daskin MS, Church RL (2006) Planning for disruptions in supply chain networks. *Models, Methods, and Applications for Innovative Decision Making*, 234–257 (INFORMS).

Terblanche S, Wessäly R, Hattingh J (2011) Survivable network design with demand uncertainty. *European Journal of Operational Research* 210(1):10–26.

Washburn A, Wood K (1995) Two-person zero-sum games for network interdiction. *Operations Research* 43(2):243–251.

Wollmer R (1964) Removing arcs from a network. *Operations Research* 12(6):934–940.

Wood RK (1993) Deterministic network interdiction. *Mathematical and Computer Modelling* 17(2):1–18.

Appendix

A. Technical proofs

A.1. Proof of Proposition 1

We create a super-source node $\bar{s}$ and connect it to each source node $s \in \mathcal{S}$. The capacity of each arc $(\bar{s}, s)$ is assigned to be ∞. We also create a super-sink node $\bar{d}$ and each demand node $d \in \mathcal{D}$ is connected to $\bar{d}$ with the capacity of the arc $(d, \bar{d})$ being ξ_d. We notice that the optimal value of the maximization problem in (2) is exactly the maximal flow from $\bar{s}$ to $\bar{d}$ in the new network. Let $(\bar{\mathcal{S}}, \bar{\mathcal{D}})$ be the cut of the new network with minimum capacity. We let $\alpha_i = 0$ if $i \in \bar{\mathcal{S}}$ and $\alpha_i = 1$ if $i \in \bar{\mathcal{D}}$. For forward arcs (i,j) and $(d, \bar{d})$, where $d \in \mathcal{D}$, with $i, d \in \bar{\mathcal{S}}, j \in \bar{\mathcal{D}}$ in the cut, we let $\beta_{ij} = 1, \gamma_d = 1$ and let all others $\beta_{ij} = 0, \gamma_d = 0$. We see that the objective value of (5) equals the capacity of the min cut, and all $\alpha_i, \beta_{ij}, \gamma_d$ are in $\{0,1\}$. We now verify the feasibility of

$$\begin{aligned}
&\alpha_i + \beta_{id} + \gamma_d \geq 1, \ \forall (i,d) \in \mathcal{A}, i \in \mathcal{T}, d \in \mathcal{D}, \\
&\alpha_i - \alpha_j + \beta_{ij} \geq 0, \ \forall (i,j) \in \mathcal{A}, i, j \in \mathcal{T},
\end{aligned}$$

and

$$\begin{aligned}
\beta_{sd} + \gamma_d &\geq 1, \ \forall (s,d) \in \mathcal{A}, s \in \mathcal{S}, d \in \mathcal{D}, \\
-\alpha_j + \beta_{sj} &\geq 0, \ \forall (s,j) \in \mathcal{A}, s \in \mathcal{S}, j \in \mathcal{T}
\end{aligned}$$

will hold with equalities.

1. To verify $\alpha_i + \beta_{id} + \gamma_d \geq 1, (i,d) \in \mathcal{A}, i \in \mathcal{T}, d \in \mathcal{D}$, for each arc $(i,d) \in \mathcal{A}, i \in \mathcal{T}, d \in \mathcal{D}$, if $d \in \bar{\mathcal{S}}$ then $\gamma_d = 1$; if $i, d \in \bar{\mathcal{D}}$ then $\alpha_i = 1$; and if $i \in \bar{\mathcal{S}}, d \in \bar{\mathcal{D}}$ then $\beta_{id} = 1$. Hence $\alpha_i + \beta_{id} + \gamma_d \geq 1$.
2. To see $\beta_{sd} + \gamma_d = 1, (s,d) \in \mathcal{A}, s \in \mathcal{S}, d \in \mathcal{D}$, for each arc $(s,d) \in \mathcal{A}, s \in \mathcal{S}, d \in \mathcal{D}$, if $d \in \bar{\mathcal{D}}$ then $\beta_{sd} = 1, \gamma_d = 0$; if $d \in \bar{\mathcal{S}}$ then $\beta_{sd} = 0, \gamma_d = 1$. Hence $\beta_{sd} + \gamma_d = 1$.
3. For each arc $(i,j) \in \mathcal{A}, i, j \in \mathcal{T}$, if $i, j \in \bar{\mathcal{S}}$ or $i, j \in \bar{\mathcal{D}}$ then $\alpha_i = \alpha_j$, $\beta_{ij} = 0$; if $i \in \bar{\mathcal{S}}, j \in \bar{\mathcal{D}}$ then $\alpha_i = 0, \alpha_j = 1, \beta_{ij} = 1$ and if $\alpha_i \in \bar{\mathcal{D}}, j \in \bar{\mathcal{S}}$ then $\alpha_i = 1, \alpha_j = 0, \beta_{ij} = 0$. Therefore, $\alpha_i - \alpha_j + \beta_{ij} \geq 0$.
4. For each arc $(s,j) \in \mathcal{A}, j \in \mathcal{T}, s \in \mathcal{S}$, if $j \in \bar{\mathcal{S}}$ then $\alpha_j = 0, \beta_{sj} = 0$ and if $j \in \bar{\mathcal{D}}$ then $\alpha_j = 1, \beta_{sj} = 1$. Therefore, $-\alpha_j + \beta_{sj} = 0$.

From max flow min cut theorem, we conclude that we have constructed an optimal solution of (5) that has all $\alpha_i (i \in \mathcal{T})$, $\beta_{ij}((i,j) \in \mathcal{A})$, $\gamma_d(d \in \mathcal{D})$ in $\{0,1\}$ and $\beta_{sd} + \gamma_d = 1$ for $(s,d) \in \mathcal{A}$, $s \in \mathcal{S}$, $d \in \mathcal{D}$, and $-\alpha_j + \beta_{sj} = 0$ for $(s,j) \in \mathcal{A}, s \in \mathcal{S}, j \in \mathcal{T}$. □

A.2. Proof of Theorem 1

We rewrite the double maximization problem $\max_v \max_{\alpha,\beta,\gamma}$ in (9) as

$$\sum_{d \in \mathcal{D}} \xi_d - \min_{v \in V, (\alpha,\beta,\gamma) \in \Theta} \left\{ \sum_{d \in \mathcal{D}} \xi_d \gamma_d + \sum_{(i,j) \in \mathcal{A}} \left((c_{ij} - y_{ij}^{(1)})(1 - v_{ij}) + y_{ij}^{(1)} + y_{ij}^{(2)} \right) \beta_{ij} \right\} \tag{18}$$

The minimization problem in (18) is a bilinear mixed integer problem. Using McCormick convex relaxation McCormick (1976), we replace $(1-v_{ij})\beta_{ij}$ by η_{ij} and obtain the following relaxation.

$$\begin{array}{lll}
\min\limits_{v,\alpha,\beta,\gamma,\eta} & \sum\limits_{d\in\mathcal{D}}\xi_d\gamma_d+\sum\limits_{(i,j)\in\mathcal{A}}\left((c_{ij}-y^{(1)}_{ij})\eta_{ij}+\left(y^{(1)}_{ij}+y^{(2)}_{ij}\right)\beta_{ij}\right) & \\
\text{s.t.} & 0\le\eta_{ij}\le\beta_{ij}, & \forall(i,j)\in\mathcal{A},\\
& \eta_{ij}\le 1-v_{ij}, & \forall(i,j)\in\mathcal{A},\\
& \eta_{ij}\ge\beta_{ij}-v_{ij}, & \forall(i,j)\in\mathcal{A},\\
& v\in V, & \\
& (\alpha,\beta,\gamma)\in\Theta. &
\end{array} \tag{19}$$

Solving (19) gives us a lower bound for the minimization problem in (18). If there is an optimal solution of (19) satisfying $\eta_{ij}=(1-v_{ij})\beta_{ij},(i,j)\in\mathcal{A}$, then (19) has the same optimal value as the minimization problem in (18). Now consider an optimal solution of (19). If $\beta_{ij}=0$ then we can choose $\eta_{ij}=v_{ij}=0$ and if $\beta_{ij}=1$ then we can choose $\eta_{ij}=\beta_{ij}-v_{ij}$, since these choices do not affect the feasibility of (19) but minimize the objective function. Therefore, (19) has an optimal solution satisfying $\eta_{ij}=\beta_{ij}-v_{ij}$. We now consider solutions of (19) with $\eta_{ij}=\beta_{ij}-v_{ij}$ for $(i,j)\in\mathcal{A}$. If $\beta_{ij}=0$ then we must have $\eta_{ij}=-v_{ij}=0$. Consequently, $\eta_{ij}=(1-v_{ij})\beta_{ij}$. And if $\beta_{ij}=1$ then we obviously have $\eta_{ij}=(1-v_{ij})\beta_{ij}$. Therefore, system (19) with $\eta_{ij}=\beta_{ij}-v_{ij}$, for $(i,j)\in\mathcal{A}$ solves the minimization problem in (18). Substituting $\beta_{ij}=\eta_{ij}+v_{ij}$ to (19) and eliminating redundant constraints $\eta_{ij}\le\beta_{ij}$ and $\eta_{ij}\le 1-v_{ij}$ lead us to the following problem

$$\begin{array}{llll}
\min\limits_{v,\alpha,\beta,\gamma,\eta} & \sum\limits_{d\in\mathcal{D}}\xi_d\gamma_d+\sum\limits_{(i,j)\in\mathcal{A}}(c_{ij}+y^{(2)}_{ij})\eta_{ij}+(y^{(1)}_{ij}+y^{(2)}_{ij})v_{ij} & & \\
\text{s.t.} & \alpha_j+\eta_{jd}+v_{jd}+\gamma_d\ge 1, & \forall(j,d)\in\mathcal{A},j\in\mathcal{T},d\in\mathcal{D}, & (20a)\\
& \eta_{sd}+v_{sd}+\gamma_d=1, & \forall(s,d)\in\mathcal{A},s\in\mathcal{S},d\in\mathcal{D}, & (20b)\\
& \alpha_i-\alpha_j+\eta_{ij}+v_{ij}\ge 0, & \forall(i,j)\in\mathcal{A},i,j\in\mathcal{T}, & (20c)\\
& -\alpha_j+\eta_{sj}+v_{sj}=0, & \forall(s,j)\in\mathcal{A},s\in\mathcal{S},j\in\mathcal{T}, & (20d)\\
& \sum\limits_{(i,j)\in\mathcal{A}}v_{ij}\le R, & & (20e)\\
& \beta_{ij}=\eta_{ij}+v_{ij}, & \forall(i,j)\in\mathcal{A}, & (20f)\\
& 0\le\eta_{ij}\le 1, & \forall(i,j)\in\mathcal{A}, & (20g)\\
& 0\le\alpha_i\le 1, & \forall i\in\mathcal{T}, & (20h)\\
& 0\le\gamma_d\le 1, & \forall d\in\mathcal{D}, & (20i)\\
& \beta_{ij},v_{ij}\in\{0,1\}, & \forall(i,j)\in\mathcal{A}, & (20j).
\end{array} \tag{20}$$

From $\beta_{ij}=\eta_{ij}+v_{ij}$ and $\beta_{ij},v_{ij}\in\{0,1\}$, we get $\eta_{ij}\in\{0,1\}$. We observe that β in (20) only appears in the constraints (20f) and (20j). We now eliminate β to reduce the dimensions of Problem (20). To do so, we just need to prove that $\eta_{ij}+v_{ij}\in\{0,1\}$ for $(i,j)\in\mathcal{A}$.

Let v,α,γ,η be an optimal solution of Problem (20) without the constraints (20f) and the conditions $\beta_{ij}\in\{0,1\}$. Since $\eta_{ij},v_{ij}\in\{0,1\}$, we have $\eta_{ij}+v_{ij}\in\{0,1,2\}$ for all $(i,j)\in\mathcal{A}$. For $(s,j)\in\mathcal{A},s\in\mathcal{S}$, from constraints (20b), (20d), and (20h), (20i), (20j) we deduce $\eta_{sj}+v_{sj}\in\{0,1\}$. Now we consider $(i,j)\in\mathcal{A},i,j\in\mathcal{T}$. If $\eta_{ij}+v_{ij}=2$ then $v_{ij}=\eta_{ij}=1$. However, when $v_{ij}=1$ we can choose η_{ij} to be 0 to decrease the objective function of (20) while satisfying the constraint (20c) since

$v_{ij} - \alpha_j \geq 0$ then. Hence $\eta_{ij} + v_{ij} \in \{0,1\}$ for $(i,j) \in \mathcal{A}, i,j \in \mathcal{T}$. Similarly, for $(j,d) \in \mathcal{A}, j \in \mathcal{T}, d \in \mathcal{D}$, from (20a) and (20h), (20i) we deduce $\eta_{jd} + v_{jd} \in \{0,1\}$.

We have proved that we can omit β of Problem (20). Consequently, Theorem 1 then follows. □

A.3. Proof of Proposition 2

Considering Problem (13), for a given $y \in Y$, a given $v \in V$ and a given realization of demand, $\xi = (\xi_d)_{d \in \mathcal{D}}$, we set $\hat{y}_{ij} = y_{ij}$ and $\hat{v}_{ij} = v_{ij}$ for $(i,j) \in \mathcal{A}$ and let $\hat{y}_{dd'} = 0$ and $\hat{v}_{dd'} = 0$ for $(d,d') \in \mathcal{A}'$. By substituting $\hat{x}_{dd'} = \sum_{i:(i,d)\in\mathcal{A}} x_{id} - \sum_{j:(d,j)\in\mathcal{A}} x_{dj}$ in the objective function in (12) to correspond to a balance equation of the transshipment node $d \in \hat{\mathcal{T}}$ in (15), we see that

$$\phi_G\left(\left[(c_{ij} - y^{(1)}_{ij})(1 - v_{ij}) + y^{(1)}_{ij} + y^{(2)}_{ij}\right]_{(i,j)\in\mathcal{A}}, \xi\right) = \hat{\phi}\left(\left[(\hat{c}_{ij} - \hat{y}^{(1)}_{ij})(1 - \hat{v}_{ij}) + \hat{y}^{(1)}_{ij} + \hat{y}^{(2)}_{ij}\right]_{(i,j)\in\hat{\mathcal{A}}}, \xi\right).$$

Conversely, considering Problem (14), for a given $\hat{y} \in \hat{Y}$, a given $\hat{v} \in \hat{V}$ and a given realization of demand, $\xi = (\xi_{d'})_{d' \in \hat{\mathcal{D}}}$, we set $y_{ij} = \hat{y}_{ij}$ and $v_{ij} = \hat{v}_{ij}$ for $(i,j) \in \mathcal{A}$. For each transshipment node $d \in \hat{\mathcal{T}}$ that is connected to the fictitious pure demand node $d' \in \hat{\mathcal{D}}$ by the arc $(d,d') \in \hat{\mathcal{A}}$, using the balance equality at d, we replace $\hat{x}_{dd'}$ in (15) by $\sum_{i:(i,d)\in\mathcal{A}} x_{id} - \sum_{j \neq d':(d,j)\in\mathcal{A}} x_{dj}$. Furthermore, note that the arc capacity constraints $\hat{x}_{dd'} \leq \hat{c}_{dd'}$ is always satisfied, thus can be omitted. Then we have

$$\hat{\phi}\left(\left[(\hat{c}_{ij} - \hat{y}^{(1)}_{ij})(1 - \hat{v}_{ij}) + \hat{y}^{(1)}_{ij} + \hat{y}^{(2)}_{ij}\right]_{(i,j)\in\hat{\mathcal{A}}}, \xi\right) = \phi_G\left(\left[(c_{ij} - y^{(1)}_{ij})(1 - v_{ij}) + y^{(1)}_{ij} + y^{(2)}_{ij}\right]_{(i,j)\in\mathcal{A}}, \xi\right).$$

We have shown that solving Problem (13) is equivalent to solving Problem (14). In particular, if $\hat{y}^* = \left(\hat{y}^{*(1)}_{ij}, \hat{y}^{*(2)}_{ij}\right)_{(i,j)\in\hat{\mathcal{A}}}$ is an optimal solution of (14) then y^* with $y^{*(1)}_{ij} = \hat{y}^{*(1)}_{ij}$ and $y^{*(2)}_{ij} = \hat{y}^{*(2)}_{ij}$ is an optimal solution of (13). □

A.4. Proof of Proposition 3

Denote $\theta = (v, \alpha, \gamma, \eta)$, let $h(y,\theta)$ be the objective function, and Ξ be the feasible set of Problem (11). We see that $h(y,\theta)$ is linear with respect to y. Hence, $y \mapsto g(y,\xi) = \max_{\theta\in\Xi} h(y,\theta)$ is convex. On the other hand, by convexity and non-decreasing of $u(\cdot)$ we conclude that for all $y_1, y_2 \in Y$ and $a \in [0,1]$ the following holds

$$\begin{aligned} aF(y_1,\xi) + (1-a)F(y_2,\xi) &= au(g(y_1,\xi)) + (1-a)u(g(y_2,\xi)) \\ &\geq u(ag(y_1,\xi) + (1-a)g(y_2,\xi)) \\ &\geq u(g(ay_1 + (1-a)y_2), \xi) = F(ay_1 + (1-a)y_2, \xi). \end{aligned}$$

Statement 1 of Proposition 3 follows then.

To show Statement 2, let $\theta^*(y) \in \arg\max_{\theta\in\Xi} h(y,\theta)$ and let $h'_1(y,\theta)$ be the derivative of the map $y \mapsto h(y,\theta)$. Then $h'_1(y,\theta^*(y))$ is a subgradient of the map $y \mapsto g(y,d)$ at y. Furthermore, our utility function $u(\cdot)$ is convex and non-decreasing. Hence $u'(g(y,d))h'_1(y,\theta^*(y))$ is a subgradient of the map $y \mapsto u(g(y,d))$ at y. Finally, the result follows from $h'_1(y,\theta^*(y)) = \left(\left[-v^*_{ij}\right]_{(i,j)\in\mathcal{A}}, \left[-\eta^*_{ij} - v^*_{ij}\right]_{(i,j)\in\mathcal{A}}\right)$.
□

B. Extension to partial arc interdiction and multicommodity network flow

B.1. Fortification model with partial arc interdiction

In practice, there are many cases in which the attacker can partially reduce capacities of arcs. For an example, when a highway is attacked, some lanes can be closed but some lanes can be still usable; that is, the capacity of the highway is partially reduced. For another example, consider an electric distribution system in which there are many generators working in parallel to generate electricity for a transmission line. If some of the generators are attacked and would be terminated, the remaining generators can still be working; that is, the electricity capacity of the transmission line is partially reduced. To model partial arc interdiction, we assign the interdiction variables $v_{ij} \in [0,1]$ for $(i,j) \in \mathcal{A}$, i.e., the set of feasible fortification decisions is rewritten as follows.

$$V := \left\{ v = (v_{ij})_{(i,j)\in\mathcal{A}}, 0 \le v_{ij} \le 1, \sum_{(i,j)\in\mathcal{A}} v_{ij} \le R \right\}.$$

This formulation reflects that there may be more than R arcs to be attacked, but the sum of the attack coefficients is at most R. Similarly to Theorem 1, we now give a stochastic programming formulation to solve Problem (9) for the case of partial arc interdiction.

THEOREM 2. *The optimal solution y^* of the following problem solves Problem* (9)*:*

$$\min_{y\in Y} \mathbb{E}[u(g(y,\xi))], \tag{21}$$

where

$$\begin{array}{lll} g(y,\xi) = \max\limits_{v,\alpha,\gamma,\eta} & \sum\limits_{d\in\mathcal{D}} \xi_d(1-\gamma_d) - \sum\limits_{(i,j)\in\mathcal{A}} (c_{ij} + y_{ij}^{(2)})\eta_{ij} + (y_{ij}^{(1)} + y_{ij}^{(2)})v_{ij} & \\ \text{s.t.} & \alpha_j + \eta_{jd} + v_{jd} + \gamma_d \ge 1, & \forall (j,d) \in \mathcal{A}, j \in \mathcal{T}, d \in \mathcal{D}, \\ & \eta_{sd} + v_{sd} + \gamma_d = 1, & \forall (s,d) \in \mathcal{A}, s \in \mathcal{S}, d \in \mathcal{D}, \\ & \alpha_i - \alpha_j + \eta_{ij} + v_{ij} \ge 0, & \forall (i,j) \in \mathcal{A}, i,j \in \mathcal{T}, \\ & -\alpha_j + \eta_{sj} + v_{sj} = 0, & \forall (s,j) \in \mathcal{A}, s \in \mathcal{S}, j \in \mathcal{T}, \\ & \sum\limits_{(i,j)\in\mathcal{A}} v_{ij} \le R, & \\ & 0 \le \eta_{ij} \le 1, & \forall (i,j) \in \mathcal{A}, \\ & 0 \le v_{ij} \le 1, & \forall (i,j) \in \mathcal{A}, \\ & \alpha_i, \gamma_d \in \{0,1\}, & \forall i \in \mathcal{T}, d \in \mathcal{D}. \end{array} \tag{22}$$

Proof of Theorem 2 We note that in the proof of Proposition 1 we gave an optimal solution of (5) with $\alpha_i, \gamma_d \in \{0,1\}$ for $i \in \mathcal{T}, d \in \mathcal{D}$. We thus can restrict α_i and γ_d of Problem (20) to binary variables. Similarly to the case of binary interdiction, we will eliminate the constraints $\beta_{ij} = \eta_{ij} + v_{ij}, (i,j) \in \mathcal{A}$ to reduce the dimensions of Problem (20). Let us rewrite Problem (20) without constraints (20f), with the restriction $\alpha_i, \gamma_d \in \{0,1\}$ and with the condition $v_{ij} \in [0,1]$ as follows.

$$\begin{array}{lll} \min\limits_{v,\alpha,\gamma,\eta} & \sum\limits_{d\in\mathcal{D}} \xi_d\gamma_d + \sum\limits_{(i,j)\in\mathcal{A}} (c_{ij} + y_{ij}^{(2)})\eta_{ij} + (y_{ij}^{(1)} + y_{ij}^{(2)})v_{ij} & \\ \text{s.t.} & (20a), (20b), (20c), (20d), (20e) & \\ & 0 \le \eta_{ij} \le 1, & \forall (i,j) \in \mathcal{A}, \\ & 0 \le v_{ij} \le 1, & \forall (i,j) \in \mathcal{A}, \\ & \alpha_i, \gamma_d \in \{0,1\}, & \forall i \in \mathcal{T}, d \in \mathcal{D}. \end{array} \tag{23}$$

Let v, α, γ, η be the optimal solution of (23) that has the least $\sum_{(i,j)\in\mathcal{A}} v_{ij}$ (note that the set $\{(\alpha_i, \gamma_d), \alpha_i, \gamma_d \in \{0,1\}, i \in \mathcal{T}, d \in \mathcal{D}\}$ is a finite set, and thus the set of the optimal solution v of (23) is a finite intersection of closed convex sets). For $(s,j) \in \mathcal{A}, s \in \mathcal{S}$, from $\eta_{sd} + v_{sd} + \gamma_d = 1, (s,d) \in \mathcal{A}, s \in \mathcal{S}, d \in \mathcal{D}$, and $-\alpha_j + \eta_{sj} + v_{sj} = 0, (s,j) \in \mathcal{A}, s \in \mathcal{S}, j \in \mathcal{T}$, we deduce $\eta_{sj} + v_{sj} \in \{0,1\}$. We now consider $(i,j) \in \mathcal{A}, i, j \in \mathcal{T}$.

- If $\alpha_i \geq \alpha_j$, then from $\alpha_i - \alpha_j + \eta_{ij} + v_{ij} \geq 0, (i,j) \in \mathcal{A}, i, j \in \mathcal{T}$, we get $\eta_{ij} + v_{ij} \geq 0 \geq \alpha_j - \alpha_i$. This yields that $\eta_{ij} + v_{ij} = 0$, otherwise we can decrease η_{ij} or v_{ij} to obtain a smaller objective function of Problem (23) or smaller $\sum_{(i,j)\in\mathcal{A}} v_{ij}$.
- If $\alpha_i < \alpha_j$, then together with $\alpha_i, \alpha_j \in \{0,1\}$ we deduce $\eta_{ij} + v_{ij} \geq \alpha_j - \alpha_i = 1$. Using the same argument as the previous case, we obtain that $\eta_{ij} + v_{ij} = 1$.

Therefore, $\eta_{ij} + v_{ij} \in \{0,1\}$ for each $(i,j) \in \mathcal{A}, i, j \in \mathcal{T}$. Similarly, from $\alpha_j + \eta_{jd} + v_{jd} + \gamma_d \geq 1, (j,d) \in \mathcal{A}, j \in \mathcal{T}, d \in \mathcal{D}$ and the restriction $\alpha_j, \gamma_d \in \{0,1\}$ we can prove that $\eta_{jd} + v_{jd} \in \{0,1\}$ for $(j,d) \in \mathcal{A}, j \in \mathcal{T}, d \in \mathcal{D}$.

We have constructed an optimal solution of (23) that also is a feasible solution of (20) with $\alpha_i, \gamma_d \in \{0,1\}$. As a consequence, we yield that the system (23) and (20) with $\alpha_i, \gamma_t \in \{0,1\}$ has the same optimal value. Theorem 2 then follows. □

Totally similarly to the case of binary interdiction, see Section 3, we can use Algorithm 1 to numerically solve Problem (22).

B.2. Fortification model for multi-commodity transshipment network flow

Although single commodity networks are prevalent in most network fortification models, multi-commodity networks are more advisable in certain circumstances, such as in airline operations, telecommunication applications and supply chain networks. In these networks, the operator is transferring multiple commodities through the networks, see e.g., Ali et al. (1984), Folie and Tiffin (1976), Gendron et al. (1999), Geoffrion and Graves (1974), Wood (1993). In this section , we give an extension of the fortification model described in Section 2.1 to multicommodity network flow with multiple sources and multiple sinks. The variables $\beta_{ij}, (i,j) \in \mathcal{A}$, in our model for multi-commodity network still satisfy the bounds $0 \leq \beta_{ij} \leq 1, (i,j) \in \mathcal{A}$. These bounds are useful for solving the involved bilinear programming in case of continuous interdiction and for the linearization in case of binary interdiction.

We consider a directed network $(\mathcal{N}, \mathcal{A})$ with a set of commodities G. Each commodity $m \in G$ has its own source nodes $\mathcal{S}^{(m)}$, demand nodes $\mathcal{D}^{(m)}$ and transfer nodes $\mathcal{T}^{(m)}$. We remark that some commodities may have common source nodes or sink nodes. The set of nodes of the network is $\mathcal{N} = \bigcup_{m\in G} \mathcal{S}^{(m)} \cup \mathcal{D}^{(m)} \cup \mathcal{T}^{(m)}$. Each arc (i,j) in the set of arcs $\mathcal{A}$ has a capacity $\bar{c}_{ij}$. Since the sets of transshipment nodes can be different for different commodities, it is not necessary that all of commodities in G are transferred through an arc $(i,j)\in\mathcal{A}$. For clarity, we denote G_{ij} to be the set of commodities transferred through an arc (i,j). Our problem for multicommodity network flow is mathematically modelled as follows

$$\min_{y\in Y}\mathbb{E}\left[\max_{v\in V}\min_{x\in\mathcal{X}(y,v)}\sum_{m\in G}\tau_m\sum_{d\in\mathcal{D}^{(m)}}\max\left\{\xi_d^{(m)}-\sum_{j:(j,d)\in\mathcal{A}}x_{jd}^{(m)},0\right\}\right] \tag{24}$$

where Y,V are defined as in (7), (8) and

$$\mathcal{X}(y,v)=\left\{x=\left(x_{ij}^{(m)}:(i,j)\in\mathcal{A},m\in G_{ij}\right)\right\}\text{ satisfying}$$

$$\begin{array}{ll}\sum\limits_{j:(i,j)\in\mathcal{A}}x_{ij}^{(m)}-\sum\limits_{j:(j,i)\in\mathcal{A}}x_{ji}^{(m)}=0, & \forall i\in\mathcal{T}^{(m)},m\in G,\\ \sum_{m\in G_{ij}}x_{ij}^{(m)}\le(c_{ij}-y_{ij}^{(1)})(1-v_{ij})+y_{ij}^{(1)}+y_{ij}^{(2)}, & \forall(i,j)\in\mathcal{A},\\ x_{ij}^{(m)}\ge 0, & \forall(i,j)\in\mathcal{A},m\in G_{ij}.\end{array} \tag{25}$$

Here parameters $\tau_m>0$, $m\in G$, are the penalty coefficients of shortfall of each commodity. The first equalities of (25) are the conservation laws for transshipment nodes of each commodity, and the second inequalities of (25) are the common capacity constraints for each arc of the network. To simplify the problem of minimizing the penalty in the third stage of (24), let us consider the following problem:

$$\begin{array}{lll}\max\limits_{x} & \sum\limits_{m\in G}\tau_m\sum\limits_{\substack{d\in\mathcal{D}^{(m)}\\ j:(j,d)\in\mathcal{A}}}x_{jd}^{(m)} & \\ \text{s.t.} & \sum\limits_{j:(i,j)\in\mathcal{A}}x_{ij}^{(m)}-\sum\limits_{j:(j,i)\in\mathcal{A}}x_{ji}^{(m)}=0, & \forall i\in\mathcal{T}^{(m)},m\in G,\\ & \sum\limits_{m\in G_{ij}}x_{ij}^{(m)}\le\left(c_{ij}-y_{ij}^{(1)}\right)(1-v_{ij})+y_{ij}^{(1)}+y_{ij}^{(2)}, & \forall(i,j)\in\mathcal{A},\\ & \sum\limits_{j:(j,d)\in\mathcal{A}}x_{jd}^{(m)}\le\xi_d^{(m)}, & \forall d\in\mathcal{D}^{(m)},m\in G,\\ & x_{ij}^{(m)}\ge 0, & \forall(i,j)\in\mathcal{A},m\in G_{ij}.\end{array} \tag{26}$$

If $\tau_m=1,\forall m\in G$, following the methodology established in Section 2 we can show a result similar to Theorem 1. For general τ_m, without lost of generality we assume $\tau_m\le 1$ for all $m\in G$, as otherwise we can divide the objective of (26) by $\sum\limits_{m\in G}\tau_m$. Denote $z_{ij}^{(m)}=\tau_m x_{ij}^{(m)}$ for $(i,j)\in\mathcal{A},m\in G_{ij}$, and consider the following auxiliary problem

$$\begin{array}{llll}\max\limits_{z} & \left(\sum\limits_{m\in G}\sum\limits_{\substack{d\in\mathcal{D}^{(m)}\\ j:(j,d)\in\mathcal{A}}}z_{jd}^{(m)}-\sum\limits_{(i,j)\in\mathcal{A}}w_{ij}-\sum\limits_{m\in G,d\in\mathcal{D}^{(m)}}w_d^{(m)}\right) & & \\ \text{s.t.} & \sum\limits_{j:(i,j)\in\mathcal{A}}z_{ij}^{(m)}-\sum\limits_{j:(j,i)\in\mathcal{A}}z_{ji}^{(m)}=0, & \forall i\in\mathcal{T}^{(m)},m\in G, & (27a)\\ & \sum\limits_{m\in G_{ij}}\frac{z_{ij}^{(m)}}{\tau_m}-w_{ij}\le\left(c_{ij}-y_{ij}^{(1)}\right)(1-v_{ij})+y_{ij}^{(1)}+y_{ij}^{(2)}, & \forall(i,j)\in\mathcal{A}, & (27b)\\ & \sum\limits_{j:(j,d)\in\mathcal{A}}z_{jd}^{(m)}-w_d^{(m)}\le\tau_m\xi_d^{(m)}, & \forall d\in\mathcal{D}^{(m)},m\in G, & (27c)\\ & z_{ij}^{(m)}\ge 0, & \forall(i,j)\in\mathcal{A},m\in G_{ij}, & (27d)\\ & w_{ij}\ge 0, & \forall(i,j)\in\mathcal{A}, & (27e)\\ & w_d^{(m)}\ge 0, & \forall d\in\mathcal{D}^{(m)},m\in G, & (27f).\end{array} \tag{27}$$

PROPOSITION 4. *Problem (27) has an optimal solution which satisfies that all $w_{ij}=0$, $(i,j)\in\mathcal{A}$, and $w_d^{(m)}=0$, $m\in G, d\in\mathcal{D}^{(m)}$.*

Proof of Proposition 4 We consider an optimal solution of (27) that among all optimal solutions has the least value of

$$\sum_{(i,j)\in\mathcal{A},m\in G_{ij}} z_{ij}^{(m)} + \sum_{(i,j)\in\mathcal{A}} w_{ij} + \sum_{m\in G,d\in\mathcal{D}^{(m)}} w_d^{(m)}. \tag{28}$$

Consider an arc $(i,j)\in\mathcal{A}$, if $z_{ij}^{(m)}=0, \forall m\in G_{ij}$, then from (27b) we would choose $w_{ij}=0$, otherwise the objective of (27) is not maximized. Assume that $z_{ij}^{(m)}>0$ and $w_{ij}>0$ for some $m\in G_{ij}$. Since $z_{ij}^{(m)}>0$, we deduce from the conservation equalities (27a) that we either find a circle of the commodity m or a simple path containing the arc (i,j) that connects a source $s\in\mathcal{S}^{(m)}$ to a sink $d\in\mathcal{D}^{(m)}$. Suppose we find a circle, then we deduce all arcs in the circle a small value of commodity m, but we do not change w_{ij}. Noting that the circle never contains a sink node of m, it follows that the objective function does not change and all constraints are maintained. However, the sum (28) is smaller, which is a desired contradiction. Suppose we find a simple path, we deduce from all arcs of the path as well as w_{ij} a small value ε of commodity m. This deduction does not change the objective function of (27), and it keeps the feasibility of all constraints since we can choose ε such that $\varepsilon-\frac{\varepsilon}{\tau_m}\le 0$. But we also get a smaller value of (28). We again have a contradiction. Therefore, $w_{ij}=0$. Showing $w_t^{(m)}=0$ follows a similar argument. □

Proposition 4 shows that the optimal value of (26) equals the optimal value of (27). The dual of (27) is the following:

$$\begin{array}{lll}
\min\limits_{\alpha,\beta,\gamma} & \sum\limits_{(i,j)\in\mathcal{A}} \Big(\big(c_{ij}-y_{ij}^{(1)}\big)(1-v_{ij}) + y_{ij}^{(1)} + y_{ij}^{(2)} \Big)\beta_{ij} + \sum\limits_{\substack{m\in G\\ d\in\mathcal{D}^{(m)}}} \tau_m \xi_d^{(m)} \gamma_d^{(m)} & \\
\text{s.t.} & \alpha_j^{(m)} + \frac{\beta_{jd}}{\tau_m} + \gamma_d^{(m)} \ge 1, & \forall (j,d)\in\mathcal{A}, j\in\mathcal{T}^{(m)}, d\in\mathcal{D}^{(m)}, \ (29a)\\
& \frac{\beta_{sd}}{\tau_m} + \gamma_d^{(m)} \ge 1, & \forall (s,d)\in\mathcal{A}, s\in\mathcal{S}^m, d\in\mathcal{D}^{(m)}, \ (29b)\\
& \alpha_i^{(m)} - \alpha_j^{(m)} + \frac{\beta_{ij}}{\tau_m} \ge 0, & \forall (i,j)\in\mathcal{A}, i,j\in\mathcal{T}^{(m)}, \ (29c)\\
& -\alpha_j^{(m)} + \frac{\beta_{sj}}{\tau_m} \ge 0, & \forall (s,j)\in\mathcal{A}, j\in\mathcal{T}^{(m)}, s\in\mathcal{S}^{(m)}, \ (29d)\\
& 0\le \beta_{ij} \le 1, & \forall (i,j)\in\mathcal{A}, \ (29e)\\
& 0\le \gamma_d^{(m)} \le 1, & \forall d\in\mathcal{D}^{(m)}, m\in G \ (29f).
\end{array} \tag{29}$$

Combining this with the second stage of (24), we consider the following mixed integer bilinear optimization problem

$$\begin{array}{ll}
\min\limits_{v,\alpha,\beta,\gamma} & \sum\limits_{(i,j)\in\mathcal{A}} \Big(\big(c_{ij}-y_{ij}^{(1)}\big)(1-v_{ij}) + y_{ij}^{(1)} + y_{ij}^{(2)} \Big)\beta_{ij} + \sum\limits_{\substack{m\in G\\ d\in\mathcal{D}^{(m)}}} \tau_m \xi_d^{(m)} \gamma_d^{(m)} \\
\text{s.t.} & (29a), (29b), (29c), (29d), (29e), (29f) \\
& \sum\limits_{(i,j)\in\mathcal{A}} v_{ij} \le R \\
& 0\le v_{ij}\le 1, \forall (i,j)\in\mathcal{A}.
\end{array} \tag{30}$$

Thanks to constraints $0 \le \beta_{ij} \le 1, (i,j) \in \mathcal{A}$ and $0 \le \gamma_d^{(m)} \le 1, d \in \mathcal{D}^{(m)}, m \in G$, we can solve (30) using existing algorithms for bilinear program. Specifically, in the case of binary interdiction $v_{ij} \in \{0,1\}$, we use linearization techniques to transform (30) to a mixed integer linear program, as shown below. We replace $v_{ij}\beta_{ij}$ by η_{ij} and add constraints

$$\begin{aligned}&\eta_{ij} \le \beta_{ij},\\&\eta_{ij} \le v_{ij},\\&\eta_{ij} \ge \beta_{ij} + v_{ij} - 1,\\&\eta_{ij} \ge 0,\end{aligned}$$

to (30). If $v_{ij} = 0$ then it follows from $\eta_{ij} \le v_{ij}$ that we can choose $\eta_{ij} = 0$ as coefficient of η_{ij} in the objective function of (30) is $-(c_{ij} - y_{ij}^{(1)}) \le 0$. If $v_{ij} = 1$ then, similarly, we deduce from $\eta_{ij} \le \beta_{ij}$ and $\eta_{ij} \le v_{ij}$ that we can choose $\eta_{ij} = \beta_{ij}$. Therefore, there always exists an optimal solution of the linearized problem that satisfies $\eta_{ij} = v_{ij}\beta_{ij}$ and the constraints $\eta_{ij} \ge \beta_{ij} + v_{ij} - 1$, $\eta_{ij} \ge 0$ are redundant. We thus have the following theorem.

THEOREM 3. *The optimal solution y^* of the following problem solves our model* (24),

$$\min_{y \in Y} \mathbb{E}[u(\boldsymbol{g}(y,\xi))], \tag{31}$$

where $\boldsymbol{g}(y,\xi)$ is the optimal value of the following mixed integer linear programming

$$\begin{aligned}\max_{v,\alpha,\beta,\gamma,\eta}\ & \sum_{\substack{m\in G\\ d\in\mathcal{D}^{(m)}}} \tau_m \xi_d^{(m)} - \sum_{(i,j)\in\mathcal{A}} (c_{ij} + y_{ij}^{(2)})\beta_{ij} - \sum_{\substack{m\in G\\ d\in\mathcal{D}^{(m)}}} \tau_m \xi_d^{(m)} \gamma_d^{(m)} + \sum_{(i,j)\in\mathcal{A}} (c_{ij} - y_{ij}^{(1)})\eta_{ij}\\ \text{s.t.}\ & (29a), (29b), (29c), (29d), (29e), (29f)\\ & \eta_{ij} \le \beta_{ij}, \qquad \forall (i,j) \in \mathcal{A},\\ & \eta_{ij} \le v_{ij}, \qquad \forall (i,j) \in \mathcal{A},\\ & \sum_{(i,j)\in\mathcal{A}} v_{ij} \le R,\\ & v_{ij} \in \{0,1\}, \qquad \forall (i,j) \in \mathcal{A}.\end{aligned}$$

for the case of binary interdiction, and $\boldsymbol{g}(y,\xi)$ is the optimal value of the following bilinear programming

$$\begin{aligned}\max_{v,\alpha,\beta,\gamma}\ & \sum_{\substack{m\in G\\ d\in\mathcal{D}^{(m)}}} \tau_m \xi_d^{(m)} - \sum_{(i,j)\in\mathcal{A}} \Big(\big(c_{ij} - y_{ij}^{(1)}\big)(1 - v_{ij}) + y_{ij}^{(1)} + y_{ij}^{(2)} \Big) \beta_{ij} - \sum_{\substack{m\in G\\ d\in\mathcal{D}^{(m)}}} \tau_m \xi_d^{(m)} \gamma_d^{(m)}\\ \text{s.t.}\ & (29a), (29b), (29c), (29d), (29e), (29f)\\ & \sum_{(i,j)\in\mathcal{A}} v_{ij} \le R\\ & 0 \le v_{ij} \le 1, \forall (i,j) \in \mathcal{A}.\end{aligned}$$

for the case of partial arc interdiction.

We remark that for each realization ξ, the function $y \mapsto u(\mathbf{g}(y,\xi))$ in (31) is convex. We thus can apply the robust stochastic approximation approach in Section 3 to solve Problem (31).