Sequential Shortest Path Interdiction with Incomplete Information

Juan S. Borrero

Department of Industrial Engineering, University of Pittsburgh, jsb81@pitt.edu

Oleg A. Prokopyev

Department of Industrial Engineering, University of Pittsburgh, droleg@pitt.edu

Denis Sauré

Department of Industrial Engineering, Universidad de Chile, dsaure@dii.uchile.cl

We study sequential interdiction when the *interdictor* has incomplete initial information about the network, and the *evader* has complete knowledge of the network, including its structure and arc costs. In each time period, the interdictor blocks at most k arcs from the network observed up to that period, after which the evader travels along a shortest path between two (fixed) nodes in the interdicted network. By observing the evader's actions, the interdictor learns about the network structure and arc costs, and adjusts its actions so as to maximize the cumulative cost incurred by the evader. A salient feature of our work is that the feedback in each period is deterministic and adversarial. In addition to studying the *regret* minimization problem, we also discuss *time-stability* of a policy, which is the number of time periods until the interdictor's actions match those of an *oracle* interdictor with prior knowledge of the network. We propose a class of simple interdiction policies that have a finite regret and detect when the instantaneous regret reaches zero in real time. More importantly, we establish that this class of policies belongs to the set of *efficient* policies.

Key words: Network interdiction, shortest path, k-most vital arcs, learning, incomplete information

1. Introduction

In network interdiction, an *interdictor* or leader selects a series of interdiction measures that change the structure of a network with the objective of disrupting or stopping an *evader*'s (follower's) movement through the network. The problem, initially studied in the context of military applications, has received considerable attention during the past decades, considering various objectives for the interdictor, from minimizing the maximum flow between given nodes as in Wollmer (1964), McMasters and Mustin (1970), Ghare et al. (1971), Corley and Chang (1974), Ratliff et al. (1975), Wood (1993), to maximizing the shortest path in the interdicted network, see, for instance, Fulkerson and Harding (1977), Corley and Sha (1982), Malik et al. (1989), Israeli and Wood (2002), to minimizing the maximum probability of successful evasion as in, e.g., Washburn and Wood (1995), Morton et al. (2007) and Pan and Morton (2008). Stochastic variations of these problems have also been considered by Cormican et al. (1998), Hemmecke et al. (2003), Janjarassuk and Linderoth (2008), as well as multicommodity versions, by Lim and Smith (2007). See also Smith and Lim (2008) for a survey on several types of interdiction and fortification models.

In addition to these classical settings, recent work in the area focuses on critical node/edge deletion problems, where the objective is to remove a set of nodes/edges in order to maximally degrade some connectivity measure of the remaining network. See, for instance, Shen et al. (2012), Granata et al. (2013), Veremyev, Boginski and Pasiliao (2014), Veremyev, Prokopyev and Pasiliao (2014), a survey in Walteros and Pardalos (2012) and the references given therein.

While most studies assume complete knowledge of the network structure and costs, in many applications areas, like in military settings, the interdictor operates (at least initially) with limited information about the conditions on the ground. The same holds true for applications in drug and nuclear material smuggling, which have been casted as interdiction problems, see, e.g., Wood (1993) and Morton et al. (2007). In this paper, we envision network interdiction as a sequential process, where the interdictor initially has *partial* knowledge about the structure and costs of the network, and may adapt the interdiction actions as new information is collected from observing the evader's reaction to previous actions. More specifically, we focus our attention on a specific setting, where the interdictor knows that (the evader's) arc costs are *deterministic* and belong to a given set. In each time period, the interdictor blocks at most k arcs from the network (only for the duration of the period), and the evader then travels along a shortest $1 - n$ path of the interdicted network, where nodes 1 and n are assumed to be the same in all time periods, arbitrary and fixed. Subsequently, the interdictor observes each arc on said path and its cost, and, hence, learns about the structure and costs of the network, and adjusts its actions so as to maximize the cumulative cost incurred by the evader.

Our modeling approach is motivated by the dynamics (between the interdictor and the evader) arising, for instance, when monitoring and patrolling the flow of drugs, illegal materials and illegal immigration. Gift (2010), for example, considers an interdictor (e.g., a U.S. law-enforcement or military task force) that has to periodically re-allocate its resources such as ships, planes, helicopters and land units, over different geographical zones of known routes to capture drug smugglers. The smugglers, on the other hand, for any given allocation, learn the route with highest probability of success by trial and error. The problem is modeled as a bi-level program, where the lower level corresponds to the best evader solution to a given allocation, and in the higher level the interdictor aims to minimize the long-run success probability of the evader.

The aforementioned dynamics also arise in Morton et al. (2007) and Brown et al. (2006), which consider the problem of detecting illegal material or immigrants entering through a border (their base models assume that the probabilities of successful evasion are known upfront); the interdictor allocates surveillance resources modifying the probability of detection throughout a network, and the evader, who observes such an allocation, chooses a path of minimum detection probability. The problem is formulated as a bi-level program where the interdictor minimizes the probability

of evasion, and solved using mixed integer-programming techniques. In a similar setting, Malaviya et al. (2012) study urban drug-traffic monitoring, where a police force sequentially allocates officers to monitor criminals and their trade links so as to minimize the flow of illegal drugs within a time horizon. Considering deterministic data and common (shared) information, the authors study the effect of different interdiction actions on drug flow, and learning of the network is modeled by including additional constraints, such as allowing a criminal to be arrested only if it has been monitored for certain time epochs and/or if lower ranked criminals have denounce him. (The problem is formulated as a bi-level program and solved using mixed-integer programming.)

Sequential attacker-defender and defender-attacker problems have been analyzed using game theory. Assuming perfect information, Hausken and Zhuang (2011) study how the government should balance defensive investments over time. Zhuang et al. (2010) study to what extent a defender with private information must be deceptive or secretive towards an attacker who updates its beliefs about the defender's "toughness" whenever a confrontation takes place. In a similar setting, albeit single-period, Xu and Zhuang (2014) study the attacker's trade-off between investing resources in either attacking or learning the defender's vulnerability. Non-sequential adversarial decision models where the attacker is uncertain about the defender's actions have been considered as well, see, e.g., McLay et al. (2012), and the references given therein.

The attacker-defender models above do not capture the particular model-learning component that arises in our setting. In this sense, sequential decision-making problems that involve both generic model uncertainty and learning are usually casted as multi-armed bandits (Robbins (1952)). However, the typical bandit formulation focuses on stochastic feedback, as in Lai and Robbins (1985), Auer et al. (2002), or models of adversarial nature, such as Auer et al. (2003), consider a trivial mapping between decisions and feedback. We also refer to the work by Modaresi et al. (2012) and Cesa-Bianchi and Lugosi (2012) for bandit settings with combinatorial structure. A salient feature of our model, which distinguishes it from previous work, is that the feedback collected is not directly selected by the decision-maker, but can be used to infer the cost structure of the setting.

Our model makes several assumptions. Consistent with the literature discussed above (see, e.g., Israeli and Wood (2002)), we assume that the interdictor's objective is to maximize the cumulative cost incurred by the evader, thus implying that both agents perceive costs equally (a notable exception is Bayrak and Bailey (2008)). This might accommodate settings, for example, where costs represent travel times, and the interdictor adjusts time estimates after observing the evader go through a particular route. Note that we implicitly assume that the evader does not react strategically to the interdictor's actions, i.e., he/she always chooses a shortest $1 - n$ path. We assume that all costs are deterministic and that arc costs are observed by the interdictor once these

are used by the evader. These assumptions aim at isolating a first-order effect of model uncertainty: the cost of recovering the optimal interdiction action.[1]

In this paper we analyze the performance of a simple interdiction policy that in each period interdicts a set of *k-most vital arcs*[2] (see Malik et al. (1989)) of the *observed* network, and separate the analysis for the cases when: (*i*) the costs of all initially observed arcs are known; and (*ii*) the costs of some initially observed arcs are unknown, but are in a known range. In this regard, the proposed policies are *greedy*, and *pessimistic* in that they assume a worst-case realization of the costs in setting (*ii*). The k-most vital arcs problem is *NP*-hard, as shown in Ball et al. (1989), but effective solution approaches are available in practice, see, e.g., Israeli and Wood (2002). Following similar work in sequential decision-making under uncertainty (see, e.g., Cesa-Bianchi and Lugosi (2006)), we measure policy performance in terms of the *regret*, which is the cumulative loss in cost incurred by a policy relative to that achieved by an oracle interdictor with prior knowledge of the network's structure and arc costs, and on the *time stability* of a policy, which is the number of periods before the interdiction actions match those taken by said oracle.

The contribution of the paper is two-fold. First, we show that the proposed class of policies is *efficient* (we define the concept of *efficiency* in the next section). In doing so, we identify attractive features of these policies: their regrets admit a finite horizon-independent upper bound; and they detect in real time when their actions match those taken by the oracle policy (thus, indicating that both regret and time stability not longer grow with time). In addition, we show that the pessimistic nature of the proposed policies in the case of uncertain costs on observed arcs is key to attain a finite regret. Second, we propose a semi-oracle performance benchmark that contrasts cumulative cost against that induced by an oracle with advance knowledge of the cost vector, but that must not signal that such knowledge is available. We argue that this measure provides a better fit to the setting, relative to the cumulative regret, which is arguably impractical when feedback is deterministic. In addition, we perform numerical experiments to assess efficiency of the proposed policies.

The remainder of this paper is organized as follows. Next, we introduce some necessary notation. Section 2 provides a detailed and more formal description of the problem as well as a definition of efficiency of an interdiction policy. In Section 3 we propose a simple class of interdiction policies and establish their efficiency. Section 4 develops fundamental lower bounds for policy performance,

---

[1] Consider that under stochastic feedback, repeated implementation of an action should lead to reliable cost estimates, thus our model can be viewed as a certainty equivalent version of a model with stochastic feedback where actions are changed at a maximum frequency.

[2] A set of k-most vital arcs in graph G consists of (at most) k arcs whose removal from G results in the greatest increase in the length of the shortest path between two specified nodes.

and Section 5 presents our numerical experiments. Finally, Section 6 presents our conclusions and highlights possible directions for future research.

Notation. Consider a directed graph (network) $G := (N, A, C)$, where N and A denote nodes and arcs, respectively, and $C := (c_a)_{a \in A}$ is a nonnegative cost vector, where c_a denotes the cost or length of arc $a \in A$. Let $n = |N|$. For $A' \subseteq A$, we define the graph $G[A'] := (N, A', C)$, where it is understood that only the information in C about arcs in A' is available. We denote by $\mathcal{S}(G)$ the set of all shortest $1 - n$ paths in G, where nodes 1 and n are arbitrary, fixed and given, i.e.,

$$\mathcal{S}(G) := \arg\min \left\{ \sum_{a \in P} c_a : P \text{ is an } 1 - n \text{ path in } G \right\},$$

and let $z(G)$ denote the cost of a path in $\mathcal{S}(G)$. Finally, for any path P in G, let $\ell(P)$ denote the cost of the path, i.e., $\ell(P) := \sum_{a \in P} c_a$.

2. Problem Formulation

Consider an interdictor that initially observes a subnetwork $G[A_0]$ of $G = (N, A, C)$, and knows that the cost vector C lies in the set

$$\mathcal{C}_0 := \left\{ \left(c'_1, \cdots, c'_{|N \times N|} \right) \in \mathbb{R}_+^{|N \times N|} : c'_a = c_a \text{ for } a \in \widetilde{A}_0 \text{ and } \ell_a \leq c'_a \leq u_a \text{ for } a \in \widehat{A}_0 \right\},$$

for given sets $\widetilde{A}_0 \subseteq A_0$ and $\widehat{A}_0 \subseteq A_0$, where $0 \leq \ell_a < u_a < \infty$ for all $a \in \widehat{A}_0$ and $A_0 \subseteq A$ (with $\widetilde{A}_0 \cap \widehat{A}_0 = \emptyset$). That is, the interdictor is aware of arcs in A_0, she/he knows the costs of those in $\widetilde{A}_0$, and has some prior information about the costs (specifically, lower and upper bounds) of arcs in $\widehat{A}_0$. We refer to $\mathcal{C}_0$ as the initial information available to the interdictor, as it contains her/his initial knowledge about the structure and costs of the network (we assume that the set of nodes N is known to the interdictor upfront including the evader's source and destination nodes 1 and n, respectively).

In each time period $t \in \mathcal{T} := \{0, 1, \ldots, T\}$, the following sequence of events takes place:

(1) The interdictor blocks a set of arcs $I_t \subseteq A_t$ only for the duration of period t, with $|I_t| \leq k$, where A_t denotes the set of arcs the interdictor is aware of at the beginning of period t, and the constant k denotes the maximum number of arcs that can be removed in any time period.

(2) The evader traverses through path $P_t \in \mathcal{S}(G[A \setminus I_t])$, incurring a cost of $z(G[A \setminus I_t])$ and revealing the arcs in P_t as well as their costs to the interdictor, so that $A_{t+1} := A_t \cup P_t$ and

$$\mathcal{C}_{t+1} := \{ C' \in \mathcal{C}_t : c'_a = c_a \text{ for } a \in P_t \},$$

where $\mathcal{C}_t$ denotes the set of cost vectors that are consistent with the information available to the interdictor at the beginning of period $t \in \mathcal{T}$.

In the above, and throughout the paper, we made the following assumptions:

A1. Each time period $t \in \mathcal{T}$ the interdictor observes path P_t and cost c_a of each arc $a \in P_t$ used by the evader.

A2. The evader acts myopically, always selecting a shortest $1 - n$ path in the interdicted network. Also, the evader observes the interdictor's actions before choosing a path.

A3. If there is more than one possible choice for P_t, then the evader chooses a path following a well-defined deterministic rule. Furthermore, this rule is *consistent*, in the sense that if P_t is chosen from a collection of paths $\mathcal{P}$, then it is also chosen from any collection $\widetilde{\mathcal{P}} \subseteq \mathcal{P}$ containing P_t.

A4. $I_0 = \emptyset$ and $I_t \neq \emptyset$ for all $t \geq 1$; furthermore, $A_0 = \widetilde{A}_0 \cup \widehat{A}_0$, and G is not "trivially" k-separable.[3]

Assumption **A1** can be viewed as an instance of *perfect (or transparent) feedback*, which is common in the learning theory literature, see, e.g., Cesa-Bianchi and Lugosi (2006) and requires some degree of monitoring of the evaders actions, thus its validity ultimately depends on the particulars of an application. In this regard, one can interpret this assumption in the context of a repeated interaction between the evader and the interdictor in an stochastic environment, as mentioned in the previous section: for example, consider the application to drug smuggling or ilegal immigration detection, where evaders repeatedly choose a path of minimum detection probability (which can be formulated as shortest-path interdiction): there, repeated interaction would account for successful as well as failed smuggling/trespassing attempts, thus providing the interdictor with some information regarding the success probability of the aforementioned path. While this assumption simplifies our theoretical analysis, it is somewhat limiting as it also implies that our model cannot be applied directly to some practical settings where there are limited monitoring capabilities, e.g., when the interdictor observes only the total length of the path used by the evader, or a subset of the arcs used. Nonetheless, relaxing this assumption is an interesting topic of future research (see our additional discussion in Section 6). For example, one could potentially adapt the concept of the *barycentric spanners* used by Awerbuch and Kleinberg (2004) for such generalizations (referred to as the *opaque feedback* case).

With regard to the first part of **A2**, this assumption imposes a rather simple behavior on the evader. However, one can show that the proposed policies are robust (with respect to their convergence) in settings with strategic evaders. Regarding the second part of **A2**, the assumption is that the evader has some degree of monitoring of the interdictor actions. As outlined previously, it is possible to interpret this assumption in the context of repeated interactions in an stochastic setting, in which such monitoring might arise naturally from a process of learning by trial and error on the evader's side. Please see Section 6 for further discussion.

---

[3] We refer to a directed network G as "trivially" k-separable if any set of k arcs in G forms an 1-n cut.

Assumption **A3** ensures that the evader's decisions are consistent with his/her past decisions. Intuitively, one can think that the evader ranks all paths in the network based on their costs (resolving ties according to any criteria, or even randomly) *in advance.* In each time period the evader selects the highest-rank unblocked (shortest) path from such a list. Generally speaking, this assumption prevents the evader from using randomized algorithms *during* the evasion process.

Assumption **A4** is technical and made without loss of generality. Its first part implies that the evader acts first and always interdicts at least one arc. The second part simply states that the interdictor knows valid lower and upper bounds on the cost of any arc he/she is initially aware of (note that this assumption is not limiting as such lower and upper bounds can be set at zero and at an arbitrarily large value, respectively). Finally, the non-k-separability condition implies that the problem does not admit a trivial solution.

Considering **A1** and **A4**, for $t \in \mathcal{T} \setminus \{T\}$, we define recursively $\widetilde{A}_{t+1} := \widetilde{A}_t \cup P_t$, and $\widehat{A}_{t+1} := \widehat{A}_t \setminus P_t$, hence $A_t = \widetilde{A}_t \cup \widehat{A}_t$ for all $t \in \mathcal{T}$.

An *interdiction policy* is a deterministic sequence of set functions $\pi := (\pi_t, \ t \in \mathcal{T})$, such that for each $t \geq 1$, $I_t^\pi = \pi_t(\mathcal{F}_t^\pi)$ represents the set of arcs blocked in period t, and $I_t^\pi \subseteq A_t$, where $\mathcal{F}_t^\pi$ summarizes the initial information and history of the interdiction process up to time $t-1$. That is,

$$\mathcal{F}_t^\pi := \left(\mathcal{C}_0, I_0^\pi, P_0, I_1^\pi, P_1, \cdots, I_{t-1}^\pi, P_{t-1} \right),$$

where $I_0^\pi = \emptyset$ by Assumption **A4**. As P_t, $\widehat{A}_t$, $\widetilde{A}_t$, and A_t depend on I_s^π for all $s < t$, we add a π superscript to these sets to denote dependence in policy π, when necessary.

Let Π denote the set of all *feasible interdiction policies*. Given G and $\mathcal{C}_0$, from assumption **A3**, applying policy $\pi \in \Pi$ results in a *unique* sequence $\{(I_t^\pi, P_t^\pi) \colon \ t \in \mathcal{T}\}$ of *blocking* and *evasion* decisions. We define the *cumulative regret* incurred by policy π by time t as

$$R_t^\pi(G, \mathcal{C}_0) := \sum_{s \leq t} \left(z^*(G) - z\left(G[A \setminus I_s^\pi]\right) \right),$$

where $z^*(G)$ denotes the optimal cost in the k-most vital arcs problem on G, i.e.

$$z^*(G) := \max \left\{ z\left(G[A \setminus I]\right) \colon \ I \subseteq A \ \text{s.t.} \ |I| \leq k \right\}.$$

The regret represents the cumulative loss in cost incurred by a policy, relative to that of an oracle interdictor with prior knowledge of G. For a given graph G, regret minimization is equivalent to cumulative cost maximization. We say that $(z^*(G) - z\left(G[A \setminus I_t^\pi]\right))$ is the *instantaneous regret* incurred by policy π at time $t \in \mathcal{T}$. Note that when G is k-separable, then $z^*(G) = +\infty$, and $z\left(G[A \setminus I_t^\pi]\right) = +\infty$ when I_t^π is an 1-n cut. Thus, in such cases, we take the convention that $(z^*(G) - z\left(G[A \setminus I_t^\pi]\right)) = 0$.

(a) Network G

(b) Network G'

(c) What π^1 and π^2 observe for both G and G' at time $t = 1$

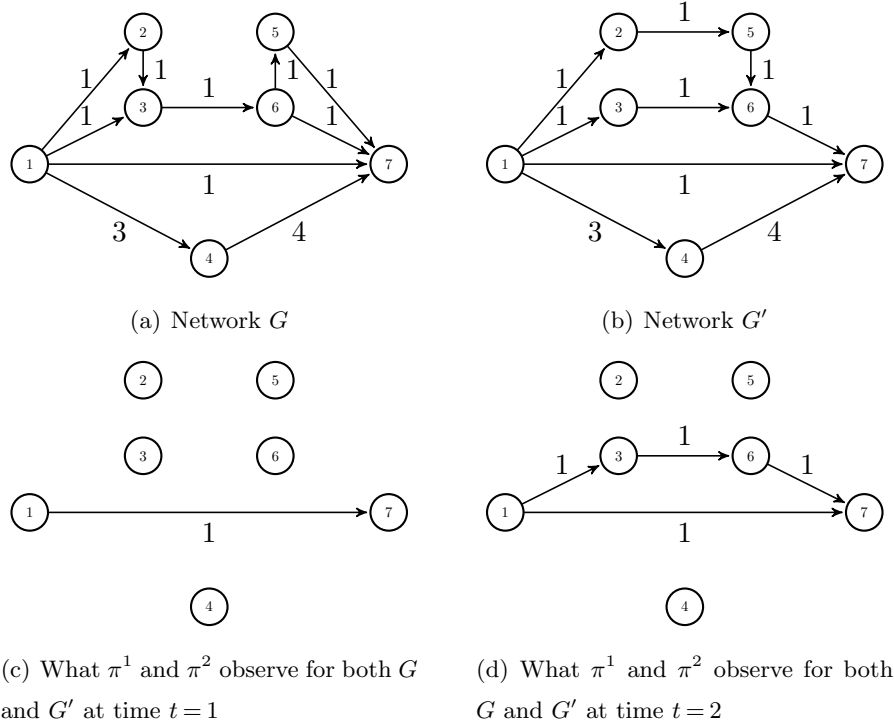(d) What π^1 and π^2 observe for both G and G' at time $t = 2$

Figure 1   Networks used in Remark 1.

Alternatively, one might instead focus on recovering the solution to the underlying k-most vital arcs problem as soon as possible, which is not necessarily aligned with the goal of regret minimization. Hence, we define the *time-stability* of policy $\pi \in \Pi$ as

$$\tau^\pi (G, \mathcal{C}_0) := \min \{t \in \mathcal{T}: \ z(G[A \setminus I_s^\pi]) = z^*(G) \text{ for all } s \geq t\},$$

where we assume the convention that $\min\{\emptyset\} = T + 1$. (Table 1 summarizes the notation in the paper.)

Table 1   Brief summary of the key notation used in the paper.

G	Underlying directed graph	$\ell(P)$	Cost of $1 - n$ path P
$G[A']$	Subgraph including only the arcs in A'	T	Time horizon
$\mathcal{S}(G)$	Set of all $1 - n$ shortest paths in G	P_t	Path chosen by the evader in period t
$z(G)$	Cost of a shortest $1 - n$ path in G	I_t	Set of arcs removed during period t
$z^*(G)$	Optimal cost of the k-most vital arcs problem on G	(l_a, u_a)	Lower and upper bounds (known to evader) on cost of arc $a \in A$
$\widetilde{A}_t$	Arcs with known cost in period t	Π_μ	*Efficient* policies with respect to μ
$\widehat{A}_t$	Arcs with known cost interval in period t	$\mathcal{F}_t^\pi$	History up to time t under policy π
$\mathcal{C}_t$	Set of cost vectors consistent with information in period t	k	Maximum number of arcs that can be interdicted in a time period
x^π	First period evader incurs a cost predicted by π	τ_t^π	Time-stability of policy π by time t
$\mathbb{G}(\mathcal{C}_0)$	Graphs compatible with initial information $\mathcal{C}_0$	R_t^π	Regret of policy π by time t

Ideally, we would like to find a policy $\pi' \in \Pi$ that performs better than any other policy for any graph G that is consistent with the initial information in $\mathcal{C}_0$. That is, given $\mathcal{C}_0$, we aim to find π' such that $R_T^{\pi'}(G, \mathcal{C}_0) \leq R_T^{\pi}(G, \mathcal{C}_0)$ and $\tau^{\pi'}(G, \mathcal{C}_0) \leq \tau^{\pi}(G, \mathcal{C}_0)$, for all $\pi \in \Pi$ and $G \in \mathbb{G}(\mathcal{C}_0)$, where

$$\mathbb{G}(\mathcal{C}_0) := \{G : G = (N, A, C), A \subseteq N \times N, C \in \mathcal{C}_0\}.$$

As shown in Remark 1 below, this is not always possible.

REMARK 1. Consider networks $G = (N, A, C)$ and $G' = (N, A', C')$ depicted in Figures 1(a) and 1(b), respectively. Set $k = 2$, $T = 2$, and assume that $A_0 = A'_0 = \emptyset$ (thus, $\mathcal{C}_0 = \mathcal{C}'_0 = \mathbb{R}_+^{|N \times N|}$). Because $I_0 = \emptyset$ and $I_1 \neq \emptyset$ (by assumption **A4**), then $P_0 = 1 - 7$, $I_1 = \{(1, 7)\}$ and $P_1 = 1 - 3 - 6 - 7$ for both networks under all policies. Define π^1 so that $I_2^{\pi^1} = \{(1, 7), (3, 6)\}$, and π^2 so that $I_2^{\pi^2} = \{(1, 7), (6, 7)\}$. Figures 1(c) and 1(d) depict the networks observed by the interdictor at times $t = 1$ and $t = 2$ for both G and G'. Observe that $z^*(G) = z^*(G') = 7$, and thus for policy π^1 we have that on $(G, \mathcal{C}_0)$ the total regret is $R_T^{\pi^1}(G, \mathcal{C}_0) = (7-1) + (7-3) + (7-7) = 10$, while on $(G', \mathcal{C}'_0)$ the total regret is $R_T^{\pi^1}(G', \mathcal{C}'_0) = (7-1) + (7-3) + (7-4) = 13$. Similarly, for policy π^2, $R_T^{\pi^2}(G, \mathcal{C}_0) = 13$ and $R_T^{\pi^2}(G', \mathcal{C}'_0) = 10$. Moreover, one can check that, for any policy $\pi \in \Pi$, $R_T^{\pi^1}(G, \mathcal{C}_0) = 10 \leq R_T^{\pi}(G, \mathcal{C}_0)$ and $R_T^{\pi^2}(G', \mathcal{C}'_0) = 10 \leq R_T^{\pi}(G', \mathcal{C}'_0)$. In particular, $R_T^{\pi^1}(G, \mathcal{C}_0) < R_T^{\pi^2}(G, \mathcal{C}_0)$ and $R_T^{\pi^2}(G', \mathcal{C}'_0) < R_T^{\pi^1}(G', \mathcal{C}'_0)$. Similar arguments can also be applied to time-stability. ∎

In light of the discussion above, consider the properties that one would expect efficient policies to have. Generally speaking, for any policy π, let $\mu_T^{\pi}(G, \mathcal{C}_0)$ be a measure of performance (e.g., cumulative regret, time-stability) that depends on T, G and $\mathcal{C}_0$. We say that a subset of feasible policies $\Pi_\mu^* \subseteq \Pi$ is *efficient with respect to* μ if the following conditions hold:

C1: Any policy $\pi \in \Pi_\mu^*$ eventually finds and maintains a solution to the underlying k-most vital arcs problem for all T above some finite instance-dependent threshold.

C2: Π_μ^* is a *homogeneous* set in the sense that for any policy in Π_μ^* there is no other policy in Π_μ^* that is better, or worse, across all instances. Formally, for any policy $\pi \in \Pi_\mu^*$ there exist another policy $\pi' \in \Pi_\mu^*$, $\mathcal{C}_0$ and networks $G, G' \in \mathbb{G}(\mathcal{C}_0)$ such that $\mu_T^{\pi}(G, \mathcal{C}_0) < \mu_T^{\pi'}(G, \mathcal{C}_0)$ and $\mu_T^{\pi}(G', \mathcal{C}_0) > \mu_T^{\pi'}(G', \mathcal{C}_0)$.

C3: Π_μ^* is not *dominated* by another class of policies. That is, for any $\mathcal{C}_0$ and $\pi' \in \Pi \setminus \Pi_\mu^*$, there exist $\pi \in \Pi_\mu^*$, $G \in \mathbb{G}(\mathcal{C}_0)$ and T such that $\mu_T^{\pi}(G, \mathcal{C}_0) < \mu_T^{\pi'}(G, \mathcal{C}_0)$.

In the next section we show that such class of policies exists for the case when μ is either cumulative regret $R_T^{\pi}(G, \mathcal{C}_0)$ or time-stability $\tau^{\pi}(G, \mathcal{C}_0)$. Moreover, we show that $\Pi_R^* \cap \Pi_\tau^* \neq \emptyset$.

3. Efficient Interdiction Policies

Guided by the discussion in the previous section, in this section we analyze a class of policies that are efficient with respect to regret and time-stability. First, in Section 3.1 we analyze the somewhat simpler case of $\widehat{A}_0 = \emptyset$, i.e., there is no uncertainty with respect to the costs of arcs known to the interdictor at time $t = 0$. The setting reveals the *greedy* nature of the proposed policies: in each period they remove a set of k-most vital arcs from the observed network. Later, in Section 3.2 we analyze settings where $\widehat{A}_0 \neq \emptyset$, i.e., the case with possible cost uncertainty of the initially known arcs. The setting now reveals the *pessimistic* nature of the proposed policies: when faced with uncertain costs on observed arcs, they operate as in the case where $\widehat{A}_0 = \emptyset$ by replacing said costs by their upper bounds.

We complement the study of these policies with numerical experiments in Section 5. We show that the theoretical efficiency of these policies also translates into good regret and time-stability performance across different instances when compared to other benchmark policies.

3.1. Efficient Policies When $\widehat{A}_0 = \emptyset$

Assume that $\widehat{A}_0 = \emptyset$, which implies that $A_t = \widetilde{A}_t$ for all $t \in \mathcal{T}$. For any policy $\pi \in \Pi$, $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$, define

$$x^\pi(G, \mathcal{C}_0) := \min\{t \in \mathcal{T}: \ z(G[A_t^\pi \setminus I_t^\pi]) = z(G[A \setminus I_t^\pi])\}, \tag{1}$$

the first time period in which the evader uses a path whose length is expected by the interdictor (who follows policy π). Additionally, observe that, by the end of period t, the interdictor is aware of whether or not any time period t corresponds to x^π.

Define Γ as the class of policies that at any time period t prior to x^π interdict a set of k-most vital arcs of $G[A_t^\pi]$, and then keep removing the same set of k-most vital arcs (used at time x^π) until T. That is, $\gamma \in \Gamma \subset \Pi$ if and only if

$$I_t^\gamma \in \arg\max\left\{z(G[A_t^\gamma \setminus I]): \ I \subseteq A_t^\gamma, |I| \leq k\right\} \text{ for } t \leq x^\gamma, \qquad I_t^\gamma = I_{x^\gamma}^\gamma \text{ for } t > x^\gamma. \tag{2}$$

As we discuss later (see Lemma 1), regardless of any new information provided by path P_t^γ for $t > x^\gamma$, $I_{x^\gamma}^\gamma$ remains a set of k-most vital arcs of $G[A_t^\gamma]$ for $t > x^\gamma$. Hence, the policies in Γ always interdict a set of k-most vital arcs of the observed network. Furthermore, by the definition of an interdiction policy given in Section 2, I_t^γ is a deterministic function. Thus, whenever a policy $\gamma \in \Gamma$ faces a tie at some time period (i.e., there are multiple sets of k-most vital arcs in the observed network), it is broken in a deterministic fashion. This observation is similar in spirit to assumption **A3** describing the evader's behavior, in the sense that it prevents the use of randomized decisions.
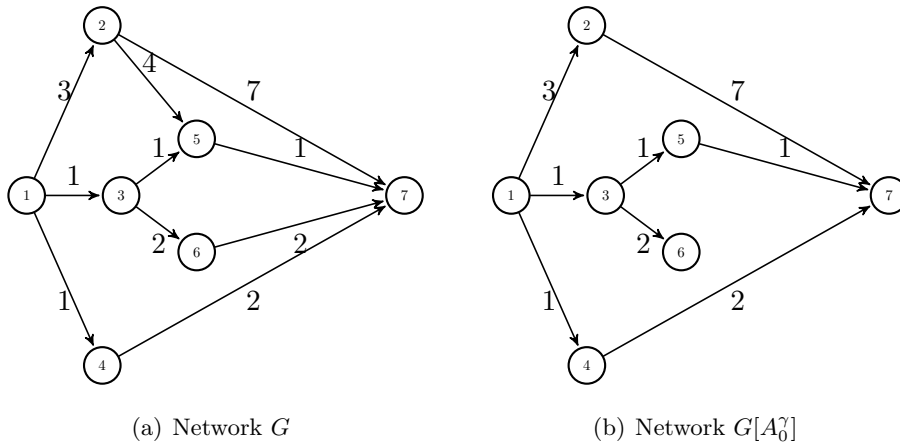
(a) Network G (b) Network $G[A_0^\gamma]$

Figure 2   **Networks used in Remark 2.**

REMARK 2. One might expect that if the interdictor uses policies in Γ, then the lengths of the shortest paths used by the evader, i.e., $\{z(G[A \setminus I_t^\gamma]) : t \in \mathcal{T}\}$, define a non-decreasing sequence in t. However, it turns out not to be the case in general. For example, let $k = 2$, and assume that $G = (N, A, C)$ is as depicted in Figure 2(a), while $G[A_0^\gamma]$ is given in Figure 2(b). Observe that $P_0^\gamma = 1 - 3 - 5 - 7$ (this is a shortest path in $G = G[A \setminus I_0^\gamma]$, recall that $I_0^\gamma = \emptyset$), and that $I_1^\gamma = \{(1,3), (1,4)\}$ (this is a 2-most vital arc solution for $G[A_1^\gamma]$). Next, $P_1^\gamma = 1 - 2 - 5 - 7$ is a unique shortest path in $G[A \setminus I_1^\gamma]$. Suppose now that $I_2^\gamma = \{(1,4), (5,7)\}$ (this is a 2-most vital arc solution in $G[A_2^\gamma]$), which implies that $P_2^\gamma = 1 - 3 - 6 - 7$ (this is a unique shortest path in $G[A \setminus I_2^\gamma]$). Therefore, we have that $z(G[A \setminus I_0^\gamma]) = 3$, $z(G[A \setminus I_1^\gamma]) = 8$, and $z(G[A \setminus I_2^\gamma]) = 5$, yielding the desired counterexample. ∎

In general, removing the same subset of arcs from a network with fewer arcs results in longer shortest paths, i.e. if $L \subseteq A' \subseteq A$, then $z(G[A' \setminus L]) \geq z(G[A \setminus L])$ (because $(A' \setminus L) \subseteq (A \setminus L)$). However, it is possible that such an action results in the same shortest path lengths in both networks. This observation motivates the following definition.

DEFINITION 1. Given $G = (N, A, C)$, L and A' such that $L \subseteq A' \subseteq A$, the network $G[A']$ is called *L-spare* (with respect to G) if $z(G[A' \setminus L]) = z(G[A \setminus L])$. Moreover, if L is also a set of k-most vital arcs of $G[A']$, then the pair $(G[A'], L)$ is called *k-complete* (with respect to G), and L is referred to as a *k-set* of $G[A']$.

Observe that if, for a given policy $\pi \in \Pi$, time t is the first period in which $G[A_t^\pi]$ is I_t^π-spare, then $t = x^\pi$. The importance of the notion of k-completeness is illustrated by the following result.

LEMMA 1. *Given $G = (N, A, C)$, let L and A' be such that $L \subseteq A' \subseteq A$ and $(G[A'], L)$ is k-complete. Then L is a set of k-most vital arcs of $G[U]$ for all U such that $A' \subseteq U \subseteq A$.*
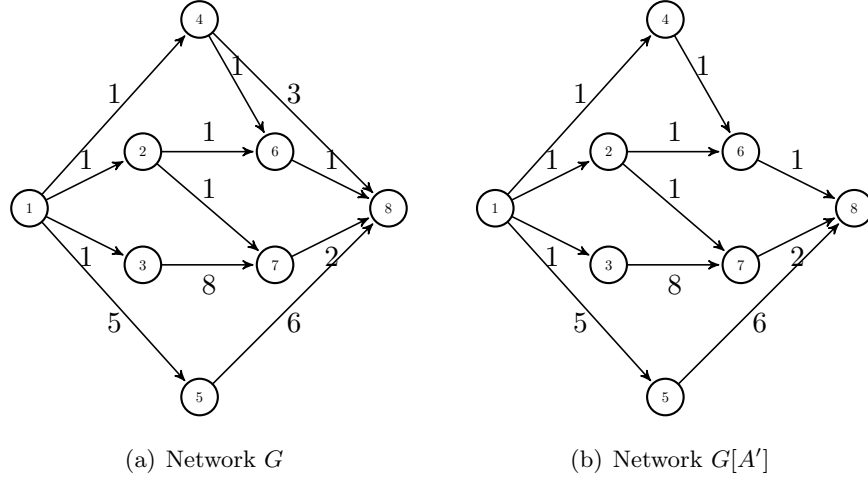
(a) Network G (b) Network $G[A']$

Figure 3   **Networks used in Remark 3.**

Proof. See Appendix B.   ∎

The practical importance of Lemma 1 lies in the fact that if one is to discover a k-complete solution of a partially observed graph, then one has indeed found a k-most vital arcs solution for the full network. This observation will play a role in showing the efficiency of the proposed policies.

REMARK 3. Note that when $(G[A'], L)$ is k-complete it is not necessarily the case that every set of k-most vital arcs of $G[A']$ is a k-set of $G[A']$ (see Definition 1). That is, if $(G[A'], L)$ is k-complete there might exist a set of k-most vital arcs $\widetilde{L}$ of $G[A']$ such that $G[A']$ is not $\widetilde{L}$-spare, i.e., such that $z(G[A' \setminus \widetilde{L}]) > z(G[A \setminus \widetilde{L}])$. For example, consider $k = 2$ and $G = (N, A, C)$ in Figure 3(a), and assume that $G[A']$ is as shown in Figure 3(b). Set $L = \{(1, 2), (1, 4)\}$ and observe that $(G[A'], L)$ is 2-complete. On the other hand, $\widetilde{L} = \{(6, 8), (7, 8)\}$ is a set of 2-most vital arcs of $G[A']$, but $\widetilde{L}$ is not a 2-set as $G[A']$ is not $\widetilde{L}$-spare. Indeed, $1 - 4 - 8$ is a shortest path in $G[A \setminus \widetilde{L}]$ and $z(G[A \setminus \widetilde{L}]) = 4$, while $1 - 5 - 8$ is a shortest path in $G[A' \setminus \widetilde{L}]$ and $z(G[A' \setminus \widetilde{L}]) = 11$.   ∎

The next two lemmas establish that the class Γ defined by (2) satisfies properties **C1** and **C2** (both with respect to cumulative regret and with respect to time-stability).

LEMMA 2. *Let $\gamma \in \Gamma$. Then for any $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$:*

1. $\tau^\gamma(G, \mathcal{C}_0) \leq x^\gamma(G, \mathcal{C}_0)$;
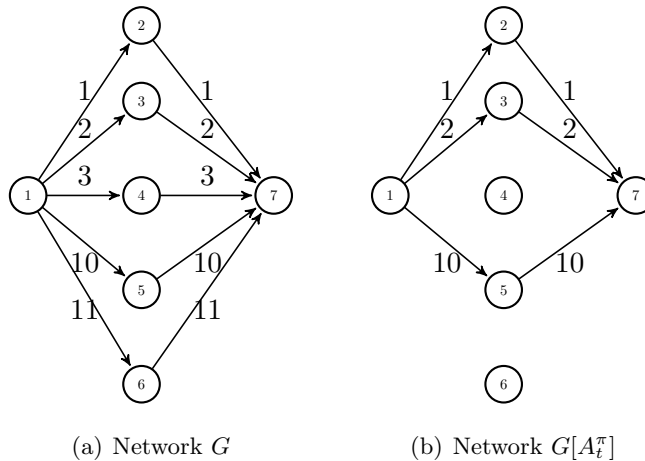2. *if $T > |A|$ then $\tau^\gamma(G, \mathcal{C}_0) \leq |A|$.*

(a) Network G (b) Network $G[A_t^\pi]$

Figure 4   **Networks used in Remark 4.**

Proof. See Appendix B. ∎

Loosely speaking, the results above follow from noting that: (i) if k-completeness is satisfied, then one has found a k-most vital arcs solution, per Lemma 1; and (ii) while k-completeness is not met, new arcs are discovered in each period.

REMARK 4. In general, for $\pi \in \Pi$ the fact that $G[A_t^\pi]$ is I_t^π-spare does not necessarily guarantee that $z(G[A \setminus I_t^\pi]) = z^*(G)$. To see this, consider the following example. Set $k = 2$, and consider G and $G[A_t^\pi]$ in Figures 4(a) and 4(b), respectively. Suppose that $I_t^\pi = \{(1,2),(1,5)\}$ (so that $P_t^\pi = 1 - 3 - 7$). Note that $G[A_t^\pi]$ is I_t^π-spare as $z(G[A_t^\pi \setminus I_t^\pi]) = z(G[A \setminus I_t^\pi]) = 4$, but $z(G[A \setminus I_t^\pi]) < z^*(G) = 6$. The reason for this is that I_t^π is not a set of 2-most vital arcs of $G[A_t^\pi]$. This observation further highlights the necessity of interdicting a set of k-most vital arcs in order to achieve an instantaneous regret of zero. ∎

LEMMA 3. *If $\widehat{A}_0 = \emptyset$, then Γ is a homogenous set both with respect to cumulative regret and with respect to time-stability.*

Proof. Let $k \geq 2$, $|N| \geq k + 2$ and $\mathcal{C}_0$ be given by Figure 5(c),[4] where for simplicity we only show $k + 2$ nodes. Consider networks G and G' depicted in Figures 5(a) and 5(b), respectively, and observe that $G, G' \in \mathbb{G}(\mathcal{C}_0)$. Clearly, $P_0^\pi = \{(1,n)\}$.

Let $\mathcal{F}_1^\pi = (\mathcal{C}_0, \emptyset, P_0^\pi)$. Observe that for the considered networks the set $\mathcal{F}_1^\pi$ is the same for all policies and the dependence on π can be dropped. Therefore, the set of policies Γ can be partitioned as $\Gamma = \Gamma^1 \cup \Gamma^2$, where $\Gamma^1 \cap \Gamma^2 = \emptyset$, $\Gamma^1 = \left\{ \gamma \colon (3,n) \in I_1^\gamma = \pi_1(\mathcal{F}_1) \right\}$ and $\Gamma^2 = \left\{ \gamma \colon (1,3) \in I_1^\gamma = \pi_1(\mathcal{F}_1) \right\}$.

---

[4] for $k = 1$ the same arguments apply after removing arc $(1,n)$ from $\mathcal{C}_0$, G and G'.

(a) Network G (b) Network $G^{'}$ (c) Initial information $\mathcal{C}_0$
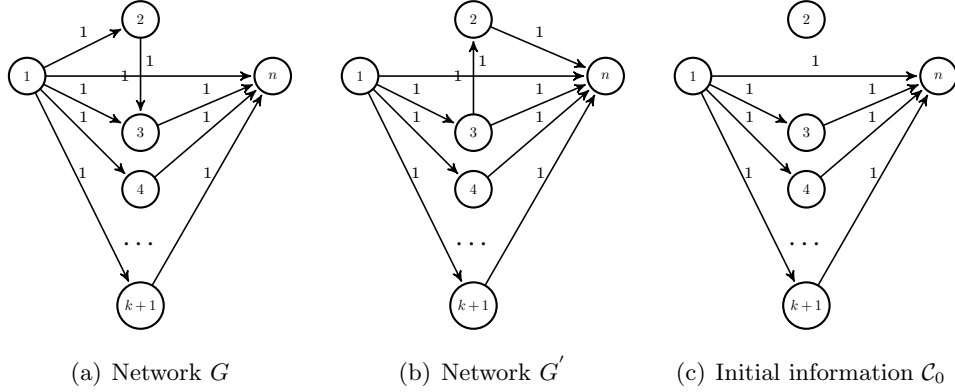
Figure 5   Networks used in the proof of Lemma 3.

Note that for any $\gamma \in \Gamma^1$, $\tau^\gamma(G, \mathcal{C}_0) = 1$ and $\tau^\gamma(G', \mathcal{C}_0) = 2$, while for any $\gamma \in \Gamma_2$, $\tau^\gamma(G, \mathcal{C}_0) = 2$ and $\tau^\gamma(G', \mathcal{C}_0) = 1$. Likewise, if $\gamma \in \Gamma_1$ then $R_1^\gamma(G, \mathcal{C}_0) = 0$ and $R_1^\gamma(G', \mathcal{C}_0) = +\infty$, and if $\gamma \in \Gamma_2$, then $R_1^\gamma(G, \mathcal{C}_0) = +\infty$ and $R_1^\gamma(G', \mathcal{C}_0) = 0$. These observations provide the result. ∎

We have proven that there exist sets of initial information $\mathcal{C}_0$ for which there is no policy in Γ that is better (or worse) than all other policies in Γ across all $G \in \mathbb{G}(\mathcal{C}_0)$. A natural question at this point is if this result can be extended for any given $\mathcal{C}_0$. The answer is negative for both regret and time-stability as illustrated by Remark 5.
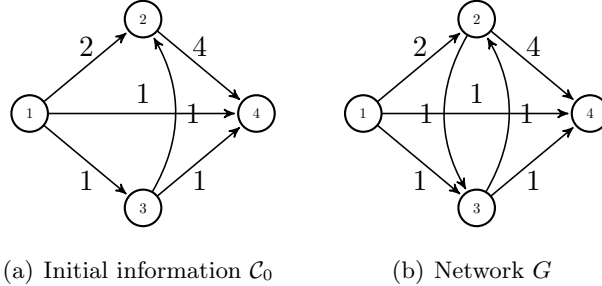

(a) Initial information $\mathcal{C}_0$ (b) Network G

Figure 6   Networks used in Remark 5.

REMARK 5. Consider $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$ as given in Figure 6. Let $T = 2$. At time $t = 1$ there are two sets of 2-most vital arcs: $I_1^\gamma = \{(3,4), (1,4)\}$ and $I_1^{\gamma'} = \{(1,3), (1,4)\}$. Observe that $P_1^\gamma = 1 - 3 - 2 - 4$ with cost $\ell(P_1^\gamma) = 6$, while $P_1^{\gamma'} = 1 - 2 - 3 - 4$ with cost $\ell(P_1^{\gamma'}) = 4$. Moreover, $x^\gamma(G, \mathcal{C}_0) = 1$ (thus, $\tau^\gamma(G, \mathcal{C}_0) = 1$) and $R_2^\gamma(G, \mathcal{C}_0) = 5$. On the other hand, $\tau^{\gamma'}(G, \mathcal{C}_0) = 2$, and $R_2^{\gamma'}(G, \mathcal{C}_0) = 7$.

Consider any other $G' \in \mathbb{G}(\mathcal{C}_0)$ different from G. Note that adding arcs $(3,1)$, $(2,1)$, $(4,2)$ and/or $(4,3)$ to G does not affect P_1^γ and $P_1^{\gamma'}$. Therefore, the only possible modification is to change the cost of $(2,3)$. However, independent of this cost, $z(G'[A \setminus I_1^{\gamma'}]) \leq 6$. Thus, for any other $G' \in \mathbb{G}(\mathcal{C}_0)$

it follows that $\tau^{\gamma'}(G', \mathcal{C}_0) \geq \tau^\gamma(G', \mathcal{C}_0)$ and $R_2^{\gamma'}(G', \mathcal{C}_0) \geq R_2^\gamma(G', \mathcal{C}_0)$. Accordingly, we conclude that $\gamma \in \Gamma$ is better than (or at least as good as) any other policy in Γ for all $G \in \mathbb{G}(\mathcal{C}_0)$. ∎

In view of the discussion above, it seems reasonable to define, for any given initial information $\mathcal{C}_0$, a subset of policies $\Gamma^* \subseteq \Gamma$ that contains all the policies that best resolve ties. Formally, for any set $\mathcal{C}_0$ we define

$$\Gamma^*(\mathcal{C}_0) = \left\{ \gamma \in \Gamma \colon R_T^\gamma(G, \mathcal{C}_0) \leq R_T^{\gamma'}(G, \mathcal{C}_0), \ \forall G \in \mathbb{G}(\mathcal{C}_0), \ \forall \gamma' \in \Gamma \right\}. \tag{3}$$

Observe that depending on $\mathcal{C}_0$ the set $\Gamma^*(\mathcal{C}_0)$ might be equal to Γ. Setting $\mathbb{C}^* = \{ \mathcal{C}_0 \colon \Gamma^*(\mathcal{C}_0) \neq \Gamma \}$, define Γ^* as the set of policies that interdict using any k-most vital arc set of the observed network if $\mathcal{C}_0 \notin \mathbb{C}^*$, and that use the element of $\Gamma^*(\mathcal{C}_0)$ if $\mathcal{C}_0 \in \mathbb{C}^*$.

We note, however, that the set Γ^* is devoid of interest from a practical perspective. This follows as for any $t \in \mathcal{T}$ breaking a tie in advance requires the interdictor to consider, at least, all the potential replies P_t^γ over $(N, N \times N)$ that are consistent with $\mathcal{F}_t^\gamma$. Clearly, this is a task that is computationally prohibitive in general.

Next, we establish the main result of this section:

THEOREM 1. *If $\widehat{A}_0 = \emptyset$, then $\Gamma \subseteq \Pi_\tau^* \cap \Pi_R^*$.*

Proof. From Lemmas 2 and 3, Γ satisfies **C1** and **C2** (both with respect to cumulative regret and with respect to time-stability). We show next that Γ also satisfies **C3**.

Specifically, fix $\pi \in \Pi \setminus \Gamma$ and $\mathcal{C}_0$, and select T and $G \in \mathbb{G}(\mathcal{C}_0)$ such that at some $t_0 \in \mathcal{T}$ the set $I_{t_0}^\pi$ is *not* a set of k-most vital arcs of $G[A_{t_0}^\pi]$. Let t_0 denote the earliest among such periods. Define $\bar{G} := G[A_{t_0}^\pi]$, i.e., the arc set of $\bar{G}$ is given by $\bar{A} := A_{t_0}^\pi$, and note that $\bar{G} \in \mathbb{G}(\mathcal{C}_0)$. Also, let $(I_t^\pi, P_t^\pi)_{t \in \mathcal{T}}$ and $(\bar{I}_t^\pi, \bar{P}_t^\pi)_{t \in \mathcal{T}}$ be the unique sequences of blocking and evasion decisions generated by π for graphs G and $\bar{G}$, respectively. By the consistency assumption, namely, **A3**, it must hold that $P_t^\pi = \bar{P}_t^\pi$ and $I_t^\pi = \bar{I}_t^\pi$ for $0 \leq t \leq t_0 - 1$. Thus, $G[A_t^\pi] = \bar{G}[\bar{A}_t^\pi]$ for all $0 \leq t \leq t_0$. Moreover, as the interdictor acts first, then $I_{t0}^\pi = \bar{I}_{t0}^\pi$. Finally, set $\bar{T} = t_0$ and define $\bar{\mathcal{T}} = \{0, 1, \ldots, \bar{T}\}$.

By our construction there exists $\gamma \in \Gamma$ such that $\bar{I}_t^\gamma = I_t^\pi$ for $0 \leq t \leq t_0 - 1$, which also implies that $\bar{A}_t^\gamma = \bar{A}_t^\pi$ for $0 \leq t \leq t_0$. Also, π is such that set $I_{t_0}^\pi$ (which coincides with $\bar{I}_{t_0}^\pi$) is not a set of k-most vital arcs of $\bar{G}$. Therefore:

$$z(\bar{G}[\bar{A}_{t0}^\pi \setminus \bar{I}_{t_0}^\pi]) = z(\bar{G}[\bar{A} \setminus \bar{I}_{t_0}^\pi]) < z^*(\bar{G}). \tag{4}$$

Note that, because $\gamma \in \Gamma$, one has that $\bar{I}_{t_0}^\gamma$ is a set of k-most vital arcs of $\bar{G}[\bar{A}_{t_0}^\gamma]$. Therefore, one has that $z(\bar{G}[\bar{A}_{t_0}^\gamma \setminus \bar{I}_{t_0}^\gamma]) = z^*(\bar{G})$. Moreover, $x^\gamma(\bar{G}, \mathcal{C}_0) \leq t_0$ and, by Lemma 2, $\tau^\gamma(\bar{G}, \mathcal{C}_0) \leq t_0$. In addition,

from equation (4), we have that $\tau^\pi(\bar{G}, \mathcal{C}_0) > t_0$. Thus, $\tau^\gamma(\bar{G}, \mathcal{C}_0) < \tau^\pi(\bar{G}, \mathcal{C}_0)$. Therefore, **C3** holds for Γ with respect to time-stability. Finally, as the regret incurred for both π and γ from $t = 0$ to $t = t_0 - 1$ is the same, the previous observations also imply that $R_{\bar{T}}^\gamma(\bar{G}, \mathcal{C}_0) < R_{\bar{T}}^\pi(\bar{G}, \mathcal{C}_0)$ and Γ satisfies **C3** with respect to cumulative regret. ∎

REMARK 6. We note that in addition of having $\Gamma \subseteq \Pi_\tau^* \cap \Pi_R^*$, an implicit important feature of policies in Γ is that they provide the interdictor with a *certificate of optimality*, i.e., whenever $t = x^\gamma$ (which by Lemma 2 happens at a time period bounded from above by $|A|$) the interdictor has the certificate that I_t^γ is a set of k-most vital arcs of G. ∎

Theorem 1 states that, despite its relative simplicity, the class of policies Γ defined by (2) is efficient with respect to regret and time-stability (i.e., it satisfies conditions **C1-C3**). In particular, such policies eventually attain an instantaneous regret of zero for sufficiently large values of T. However, as demonstrated next, the speed of convergence of policies in Γ may not be fast, and the bound implied in Lemma 3 may actually be tight.

PROPOSITION 1. *There exists $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and $\zeta > 0$ such that if $T \geq |A|$, then $\tau^\gamma(G, \mathcal{C}_0) \geq \zeta |A|$. Moreover, the value of $R_T^\gamma(G, \mathcal{C}_0)$ can be made arbitrarily large.*

Proof. Consider G in Figure 7. There, we have that $|A| = 2(k+1)u$, for some positive integer u.
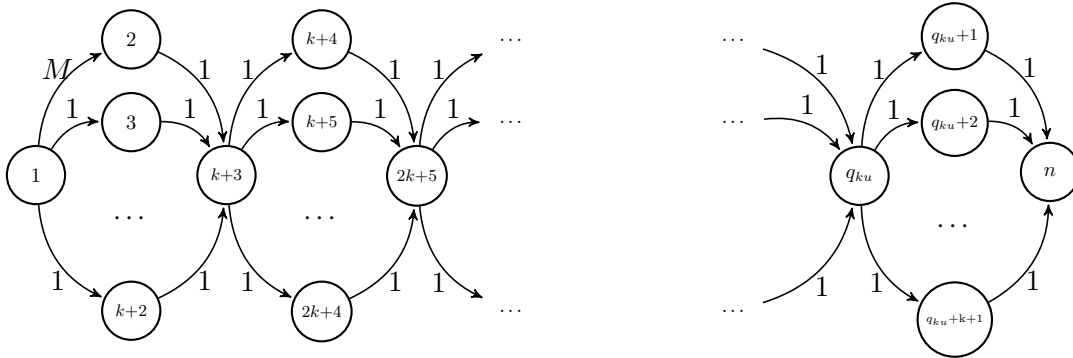
Figure 7   **Network G used for the proof of Proposition 1,** $q_{ku} = (u-1)(k+2) + 1$.

Suppose that $M > 1$ and $A_0 = \emptyset$. Without loss of generality assume that $P_0^\gamma = 1\text{-}(k+2)\text{-}(k+3)\text{-}(2k+4)\text{-}\cdots\text{-}(q_{ku}+k+1)\text{-}n$ (note that $(1,2) \notin P_0^\gamma$), so that $A_1^\gamma = P_0^\gamma$. Suppose that we select $I_1^\gamma = \{(q_{ku}+k+1, n)\}$ (which is a set of k-most vital arcs) and that $P_1^\gamma = 1\text{-}(k+2)\text{-}(k+3)\text{-}(2k+4)\text{-}\cdots\text{-}(q_{ku}+k)\text{-}n$, so that $A_2^\gamma = P_0^\gamma \cup \{(q_{ku}, q_{ku}+k), (q_{ku}+k, n)\}$. Next, we select $I_2^\gamma = \{(q_{ku}+k+1, n), (q_{ku}+k, n)\}$, a set of k-most vital arcs, and P_3^γ is a new shortest path. Proceeding in this way,

by the kth period one has that $A_k^\gamma = P_0^\gamma \cup \{(q_{ku}, q_{ku}+k), (q_{ku}+k, n), \cdots, (q_{ku}, q_{ku}+2), (q_{ku}+2, n)\}$, so we select $I_k^\gamma = \{(q_{ku}+k+1, n), (q_{ku}+k, n), \cdots, (q_{ku}+2, n)\}$, which is a set of k-most vital because it is an 1-n cut with k arcs in A_k^γ.

Note that at this point, P_k^γ includes arc $(q_{ku}+1, n)$, and, thus, $A_{k+1}^\gamma = A_k^\gamma \cup \{(q_{ku}, q_{ku}+1), (q_{ku}+1, n)\}$. While $I_k^\gamma \cup \{(q_{ku}+1, n)\}$ is an 1-n cut, it is no longer a feasible interdiction as this set contains $k+1$ arcs. However, we can select $I_{k+1}^\gamma = I_k^\gamma \cup \{(q_{ku}-1, q_{ku})\} \setminus \{(q_{ku}+k+1, n)\}$, which is an 1-n cut with k arcs, and thus, a set of k-most vital arcs. After removing those arcs, P_{k+1}^γ is of the form 1-$(k+2)$-$(k+3)$-$(2k+4)$-$\cdots$-$(q_{ku}-2)$-q_{ku}-$(q_{ku}+k+1)$-n, so that $A_{k+2}^\gamma = A_{k+1}^\gamma \cup \{(q_{k(u-1)}, q_{ku}-2), (q_{ku}-2, q_{ku})\}$. Then, we can select $I_{k+2}^\gamma = I_{k+1}^\gamma \cup \{(q_{ku}-2, q_{ku})\} \setminus \{(q_{ku}+k, n)\}$, which is a set k-most vital arcs. Proceeding in this fashion, each period (except the first one) we recover the costs of two arcs. The cost of the arc $(1, 2)$ is recovered at period ku. This implies that $\tau^\gamma(G, \mathcal{C}_0) = ku$ and hence, if $\zeta \leq k/(2k+2)$, then $\tau^\gamma(G, \mathcal{C}_0) \geq \zeta|A|$. While the latter fact depends on the selection of $\{(I_t, P_t) : t < uk\}$, one can see that, for any such selection, one can simply assign the cost M to the arc discovered last. Thus, the result holds true, independent of our selection. Finally, we observe that $R_T^\gamma(G, \mathcal{C}_0)$ can be made arbitrarily large by choosing the proper value of M. ∎

Recall from Lemma 2 that $\tau^\gamma(G, \mathcal{C}_0) \leq |A|$ for any $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$. This result, in conjunction with Proposition 1, implies, loosely speaking, that $\tau^\gamma(G, \mathcal{C}_0)$ is a $\Theta(|A|)$ function.[5]

3.2. Efficient Policies When $\widehat{A}_0 \neq \emptyset$

In this section we assume that, in addition to $\widetilde{A}_0$, the interdictor is aware of another subset of arcs $\widehat{A}_0 \subseteq A_0$ ($\widetilde{A}_0 \cap \widehat{A}_0 = \emptyset$) for which only partial cost information is available. Specifically, the interdictor knows that $c_a \in [l_a, u_a]$ for some l_a and u_a, with $l_a < u_a$, known upfront, for all $a \in \widehat{A}_0$.

Fix $\mathcal{C}_0$ and $\pi \in \Pi$, and let $(I_t^\pi, P_t^\pi)_{t \in \mathcal{T}}$ denote the unique sequence of blocking and evasion actions associated with $(G, \mathcal{C}_0)$ and π. Define a sequence of networks $\left\{ G_t^\pi := (N, A_t^\pi, \widehat{C}_t^\pi) : t \in \mathcal{T} \right\}$, where for $t \in \mathcal{T}$, $\widehat{C}_t^\pi := \{\hat{c}_a, \ a \in A_t^\pi\}$ is given by

$$\hat{c}_a := \begin{cases} c_a & \text{if } a \in \widetilde{A}_t^\pi, \\ u_a & \text{if } a \in \widehat{A}_t^\pi. \end{cases} \tag{5}$$

In other words, for network G_t^π, the costs of arcs in $\widetilde{A}_t^\pi$ are at their known values, while the costs of arcs in $\widehat{A}_t^\pi$ are at their upper bounds (this information is part of $\mathcal{C}_0$). Note that, in general, $G_t^\pi \neq G[A_t^\pi]$. Similar to Section 3.1, for any policy $\pi \in \Pi$, $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$, define $\widehat{x}^\pi(G, \mathcal{C}_0)$ as

$$\widehat{x}^\pi(G, \mathcal{C}_0) := \min\{t \in \mathcal{T} : \ z(G_t^\pi[A_t^\pi \setminus I_t^\pi]) = z(G[A \setminus I_t^\pi])\}. \tag{6}$$

---

[5] The definition of the Θ notation is given, for instance, in Ahuja et al. (1993).

Also, as in the previous section, define Λ as the class of policies that at time t interdict a set of k-most vital arcs of G_t^π. That is, $\lambda \in \Lambda$ if and only if

$$I_t^\lambda \in \arg\max\left\{z(G_t^\lambda[A_t^\lambda \setminus I]) \colon I \subseteq A_t^\lambda, |I| \leq k\right\} \text{ for } t \leq \widehat{x}^\lambda, \qquad I_t^\lambda = I_{\widehat{x}^\lambda}^\lambda \text{ for } t > \widehat{x}^\lambda. \tag{7}$$

Note that $\widehat{x}^\pi(G, \mathcal{C}_0)$ and I_t^λ are obtained by replacing the terms $G[A_t^\pi \setminus I_t^\pi]$ and $G[A_t^\gamma \setminus I]$ by $G_t^\pi[A_t^\pi \setminus I_t^\pi]$ and $G_t^\lambda[A_t^\lambda \setminus I]$ in equations (1) and (2), respectively. Simply speaking, according to the policies in Λ the interdictor should act conservatively by assuming that the costs of arcs in $\widehat{A}_t^\lambda$ are at their upper bounds, and then apply the same approach as in Section 3.1. Next, we show that Λ preserves the attractive features (namely, properties **C1**-**C3**) of the policies described in Section 3.1. For this, we need the following technical lemma, whose proof is given in Appendix B.

LEMMA 4. *Suppose that $t \in \mathcal{T}$ is such that $z(G_t^\lambda[A_t^\lambda \setminus I_t^\lambda]) = z(G[A \setminus I_t^\lambda])$, then $(G[A_t^\lambda], I_t^\lambda)$ is k-complete (with respect to G). Moreover, I_t^λ is a set of k-most vital arcs of $G[U]$ for all U such that $A_t^\lambda \subseteq U \subseteq A$.*

The next results, namely, Lemmas 5 and 6 and Theorem 2 generalize the results of Lemmas 2 and 3, and Theorem 1, respectively, for $\widehat{A} \neq \emptyset$ and the class of policies Λ. The proofs of the lemmas are similar to those of their counterparts in the previous section: see the details in Appendix B.

LEMMA 5. *Let $\lambda \in \Lambda$. Then for any $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$:*
1. $\tau^\lambda(G, \mathcal{C}_0) \leq \widehat{x}^\lambda(G, \mathcal{C}_0)$;
2. *if $T \geq |A|$, then $\tau^\lambda(G, \mathcal{C}_0) \leq |A|$.*

LEMMA 6. *Λ is a homogeneous set both with respect to cumulative regret and with respect to time-stability.*

THEOREM 2. *$\Lambda \subseteq \Pi_\tau^* \cap \Pi_R^*$.*

Proof. From Lemmas 5 and 6, Λ satisfies conditions **C1** and **C2** (both with respect to cumulative regret and with respect to time-stability). To show that Λ also satisfies condition **C3**, we use the same construction as in the proof of Theorem 1. However, there is a subtle difference (explained in detail below, see equation (8) and the related discussion) due to existence of uncertain arc costs.

Fix $\pi \in \Pi \setminus \Lambda$ and $\mathcal{C}_0$. As in the proof of Theorem 1, select T such that at some $t_0 \in \mathcal{T}$ the set $I_{t_0}^\pi$ is *not* a set of k-most vital arcs of $G_{t0}^\pi[A_{t_0}^\pi]$, and let t_0 be the earliest time period among such periods. Define $\bar{G} := G_{t0}^\pi[A_{t_0}^\pi]$, i.e., the arc set of $\bar{G}$ is given by $\bar{A} := A_{t_0}^\pi$, and its costs by $\widehat{C}_{t_0}^\pi$. (Note that $\bar{G} \in \mathbb{G}(\mathcal{C}_0)$.) Also, let $(I_t^\pi, P_t^\pi)_{t \in \mathcal{T}}$ and $(\bar{I}_t^\pi, \bar{P}_t^\pi)_{t \in \mathcal{T}}$ be the unique sequences of blocking and evasions decisions generated by π for graphs G and $\bar{G}$, respectively. Next, we need to show that

$$I_t^\pi = \bar{I}_t^\pi \quad \text{and} \quad P_t^\pi = \bar{P}_t^\pi \quad \text{for } 0 \leq t \leq t_0 - 1. \tag{8}$$
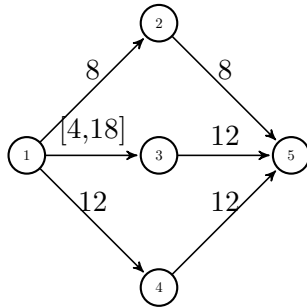
Figure 8   **Network used in Remark 7.**

However, unlike in the proof of Theorem 1, the costs of arcs in $\widehat{A}_{t_0}^{\pi}$ for G do not necessarily coincide with those for $\bar{G}$. We prove next (by contradiction) that (8) holds.

Let t' be the latest time period such that $0 \le t' \le t_0 - 1$ and (8) holds only for $0 \le t \le t' - 1$. Because the interdictor acts first, it follows that $\bar{I}_{t'}^{\pi} = I_{t'}^{\pi}$. Then it must be the case that $P_{t'}^{\pi} \ne \bar{P}_{t'}^{\pi}$. As $t' \le t_0 - 1$, $P_{t'}^{\pi}$ does not contain arcs from $\widehat{A}_{t_0}^{\pi}$. Therefore, $P_{t'}^{\pi}$ is also a shortest path in $\bar{G}[\bar{A} \setminus \bar{I}_{t'}^{\pi}]$ (recall that in $\bar{G}$ we increase only the costs of arcs in $\widehat{A}_{t_0}^{\pi}$). However, by Assumption **A3**, this implies that $P_{t'}^{\pi} = \bar{P}_{t'}^{\pi}$, and we arrive at a contradiction. Thus, (8) holds. The remainder of the proof is similar to the proof of Theorem 1. ∎

REMARK 7. Suppose that instead of a conservative approach, we set the costs of arcs in $\widehat{A}_t$ to some other values (i.e., not upper bounds) when defining $\widehat{C}_t$. As we illustrate next, this leads to policies that do not necessarily converge, thus, violating **C1**. Consider the network depicted in Figure 8, and assume that the interdictor is aware of all the costs except for $c_{13} = 16$. However, its lower and upper bounds are known to the interdictor and given by $\ell_{13} = 4$ and $u_{13} = 18$, respectively. Set $k = 2$ and consider the policies that assign either (i) the lower bound, or (ii) the average of the upper and lower bounds to arcs with unknown costs. Observe that such policies assign the cost of 16 and 23 to path $1 - 3 - 5$, respectively. Thus, the interdictor would always remove one arc of this path and another one of path $1 - 2 - 5$. Note that the instantaneous regret associated with such interdiction is 4 and that if the interdiction decision does not change, the real cost of arc $(1, 3)$ would never be revealed, implying that condition **C1** is not satisfied when the actual cost of $(1, 3)$ is $c_{13} > 12$. Furthermore, one can see that the example extends to any value used by the interdictor, other than the upper bound. ∎

One can show that Remarks 2, 4, and 5 also hold in this setting (via similar counterexamples); likewise policies in Λ also provide a certificate of optimality for the interdictor, as in Remark 6. In addition, Proposition 1 can also be extended to the case of $\widehat{A}_0 \ne \emptyset$, as we show next.

PROPOSITION 2. *There exists $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and $\zeta > 0$ such that if $T \geq |A|$, then $\tau^\lambda(G, \mathcal{C}_0) \geq \zeta |A|$. Moreover, the value of $R_T^\lambda(G, \mathcal{C}_0)$ can be made arbitrarily large.*

Proof. See Appendix B.   ∎

4. Lower Bounds for Policy Performance

To measure policy performance, the classical theory in sequential prediction often postulates a probabilistic model of uncertainty and computes, for example, the expected cost incurred by a policy, thus allowing one to search for an efficient policy based on this well-defined criterion. Assuming that such probabilistic model is initially unknown, the learning literature focuses instead, for the most part, on the concept of regret, namely, the cumulative loss in cost relative to that of an oracle (i.e., an oracle interdictor in our setting) with advanced knowledge of the underlying probabilistic model, see Cesa-Bianchi and Lugosi (2006). In our problem such information consists of the structure and arc costs of the network.

In most situations the assumption of existence of such an oracle is impractical, as feasible policies do not posses such advanced information. Nonetheless, an oracle-based benchmark can be used for normalization (i.e., preventing optimal performance to grow with the time horizon), and also for bounding the opportunity cost of missing information. While the first use also applies to our problem, the second one can be improved upon. Indeed, the oracle policy in our setting would block a set of k-most vital arcs of G in each period, i.e.,

$$I_t^{\text{oracle}} \in \arg\max \left\{ z\left(G[A \setminus I]\right) : \ I \subseteq A \text{ s.t. } |I| \leq k \right\}, \quad t \in \mathcal{T}.$$

However, this might imply using information that is not available to *any* feasible policy because, in general, $A_0 \neq A$. Thus, in our setting the bound on the opportunity cost of information is trivial (note, for example, that $\tau^{\text{oracle}}(G, \mathcal{C}_0) = 0$) and not particularly meaningful. Fortunately, it is possible to tighten such a bound by asking the oracle not to *signal* the availability of advanced information through its actions. That is, we impose that $I_t^{\text{oracle}} \subseteq A_t$ for all $t \in \mathcal{T}$. We refer to such interdictor as a semi-oracle interdictor.

Simply speaking, the semi-oracle is an interdictor that although knowing all the arcs and costs in the network, at any given time period can only remove arcs that have been observed so far, i.e., those used in earlier time periods by the evader, or that lie within A_0. Note that because the semi-oracle interdictor knows G, it is capable of anticipate the evader's actions and feedback for any sequence of blocking decisions, and, thus is also capable of evaluating with certainty the cost incurred by the evader across all time periods if said sequence is blocked, and as a result his

actions are not necessarily adapted to the history of the process (see Remark 8). Moreover, one can show that the semi-oracle will in fact use the best non-adapted policy. Note that, while the existence of such an semi-oracle is still impractical, its actions for a given instance will be matched by some feasible adapted policies, and thus it serves as a more reasonable benchmark, relative to a traditional oracle. While the semi-oracle actions are more complex than those of an oracle, one is still able to formulate the decision problem faced by the semi-oracle interdictor, and to reconstruct its actions.

Next, we formulate such a problem as a mixed integer program (MIP), first, for the case of cumulative regret (Section 4.1), and then for time-stability (Section 4.2). In the remainder of the section we assume that $\widehat{A}_0 = \emptyset$. Also, to simplify the exposition and shorten the notation we refer to the semi-oracle interdictor as an oracle interdictor, or simply, an oracle. Similarly, the resulting policies are referred to as oracle-based instead of semi-oracle-based.

4.1. Lower Bound for Regret

Given $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and T, suppose the oracle interdictor knows G at time $t = 0$, aims at maximizing the evader's cumulative cost over $\mathcal{T}$, but is restricted to selecting $I_t \subseteq A_t$ for $0 \leq t \leq T$. That is, the interdictor solves the following bilevel (max-min) optimization problem:

$$LB(G, \mathcal{C}_0, T): \quad \max \sum_{t \in \mathcal{T}} \ell(P_t) \tag{9a}$$

$$\text{s.t.} \quad P_t \in \arg\min \left\{ \sum_{a \in P} c_a : P \text{ is a path in } G[A \setminus I_t] \right\} \quad \forall t \in \mathcal{T}, \tag{9b}$$

$$I_0 = \emptyset, \ I_t \subseteq A_t, \ |I_t| \leq k, \ A_t = A_{t-1} \cup P_t \quad \forall t \in \mathcal{T} \setminus \{0\}. \tag{9c}$$

Note that in order to produce a valid lower bound, the bilevel problem (9) is *optimistic* in the sense that the interdictor has some degree of control over the decisions made by the evader. Specifically, if (9b) has multiple optimal solutions (i.e., multiple shortest paths), then the evader delegates the decision to the oracle interdictor (otherwise, it is potentially possible to improve upon the interdictor's actions). While this modeling assumption is common in the bilevel optimization literature (see, e.g., Beheshti et al. (2014), Colson et al. (2007) and references therein), in our setting it is necessary to obtain valid lower bounds for the performance (with respect to regret) of *any* policy in Π.

In order to solve (9) we observe that the evader's (lower-level) problem (9b) is the shortest path problem, which admits a compact linear programming (LP) formulation, see Ahuja et al. (1993). Consequently, the initial bilevel problem (9) can be reformulated as a single-level mixed integer program by exploiting the LP duality. We should note that this is a standard approach in the bilevel optimization literature (Colson et al. 2007), which can be applied as long as the lower-level optimization problem can be replaced by its necessary and sufficient optimality conditions.

For any arc $(i,j) \in A$ and period $t \in \mathcal{T}$, define $r_{ij}^t = 1$ if arc (i,j) is blocked at time t, and $r_{ij}^t = 0$ otherwise. Similarly, define $p_{ij}^t = 1$ if the evader travels along arc (i,j) at time t, and $p_{ij}^t = 0$, otherwise. For $t \in \mathcal{T}$, define $\{y_i^t\}_{i \in N}$ as the variables in the dual of the LP formulation of (9b). Then we have the following alternative formulation of (9):

$$LB(G, \mathcal{C}_0, T): \quad \max \sum_{t=0}^{T} (y_1^t - y_n^t)$$

$$\text{s.t.} \qquad\qquad y_i^t - y_j^t \leq c_{ij} + M\, r_{ij}^t \qquad \forall t \in \mathcal{T}, \forall (i,j) \in A, \tag{10a}$$

$$\sum_{j:(i,j)\in A} p_{ij}^t - \sum_{j:(j,i)\in A} p_{ji}^t = \begin{cases} -1 & i = 1 \\ 1 & i = n \\ 0 & \text{otherwise} \end{cases} \qquad \forall t \in \mathcal{T}, \tag{10b}$$

$$\sum_{(i,j)\in A} c_{ij}\, p_{ij}^t = y_1^t - y_n^t \qquad \forall t \in \mathcal{T}, \tag{10c}$$

$$r_{ij}^t \leq \sum_{s=0}^{t-1} p_{ij}^s \qquad \forall t \in \mathcal{T} \setminus \{0\}, \forall (i,j) \in A \setminus A_0, \tag{10d}$$

$$p_{ij}^t \leq 1 - r_{ij}^t \qquad \forall (i,j) \in A, \forall t \in \mathcal{T}, \tag{10e}$$

$$\sum_{(i,j)\in A} r_{ij}^t \leq k \qquad \forall t \in \mathcal{T}, \tag{10f}$$

$$r_{ij}^0 = 0 \qquad \forall (i,j) \in A, \tag{10g}$$

$$r_{ij}^t, p_{ij}^t \in \{0,1\} \qquad \forall (i,j) \in A, \forall t \in \mathcal{T}, \tag{10h}$$

where M is a sufficiently large constant parameter. Constraints (10a) and (10b) correspond to the dual and primal constraints of the LP formulation of the shortest path problem in the interdicted network, respectively. Note that the right-hand side of (10c) is the length of the shortest path in $G[A \setminus I_t]$. Thus, if suffices to consider $M = (n-1) \cdot \max\{c_a \mid a \in A\}$ (we do so in our numerical experiments). Constraints (10c) enforce strong duality at all times, so that $\{p_{ij}^t : (i,j) \in A\}$ corresponds to a shortest path in the interdicted network at time t. Constraints (10d) ensure that the blocking decision at time t includes only the arcs that have been observed prior to time t, and constraints (10e) prevent the evader from using blocked arcs. Finally, constraints (10f) impose that at most k arcs are interdicted in any period.

Let (r^*, p^*, y^*) denote an optimal solution to (10). We have that

$$I_t^*(G, \mathcal{C}_0, T) := \left\{(i,j) \in A : r_{ij}^{*t} = 1\right\} \quad \text{and} \quad P_t^*(G, \mathcal{C}_0, T) := \left\{(i,j) \in A : p_{ij}^{*t} = 1\right\}, \quad t \in \mathcal{T}$$

is a feasible sequence of blocking and evasion decisions, which we refer to as the *oracle-based* policy.

Note that the oracle-based policy does not belong to Π, as it is not *adapted* to the history of the process, see Remark 8. However, its performance serves as a lower-bound for regret of any policy because each policy in Π can be mapped to a feasible solution to (10). Therefore:

$$\sum_{t \leq T} (z^*(G) - z(G[A \setminus I_t^*(G, \mathcal{C}_0, T)])) \leq R_T^\pi(G, \mathcal{C}_0), \quad \pi \in \Pi, \tag{11}$$

for any $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and T.

REMARK 8. Recall that for any $\pi \in \Pi$ there is a unique set I_t^π associated with each sequence $\mathcal{F}_t^\pi$. The oracle-based policy, however, might choose different sets I_t^* for the same sequence $\mathcal{F}_t^*$ because its actions are allowed to depend on G. For instance, consider the networks $G = (N, A, C)$ and $G' = (N, A', C')$ depicted in Figure 1(a) and (b), respectively, set $k = 2$, and assume that $A_0 = \{(1,7)\}$. For both G and G' the oracle-based policy yields $I_1^* = \{(1,7)\}$, which implies that $\mathcal{F}_2^* = (\mathcal{C}_0, \emptyset, P_0, (1,7), P_1)$, where $P_0 = 1 - 7$, and $P_1 = 1 - 3 - 6 - 7$. One can check that the oracle policy satisfies that $I_2^* = \{(1,7),(3,6)\}$ for $(G, \mathcal{C}_0)$, while $I_2^* = \{(1,7),(6,7)\}$ for $(G', \mathcal{C}_0)$. Hence, the oracle policy determines two different interdiction actions at time $t = 2$ for the same $\mathcal{F}_2^*$. ∎

The oracle problem LB can be shown to be *NP*-hard using the reduction from the k-most vital arcs problem (Ball et al. 1989). Nevertheless, MIP formulation (10) can be effectively tackled by state-of-the-art solvers for small values of T and k (see Section 5). However, for larger values of T and k, a significant portion of the solver's running-time is invested into finding a feasible solution to LB. With this in mind, and considering that the total solution time typically depends on the quality of such solutions, we develop Algorithm 1, which constructs an initial feasible solution of (10).

The algorithm begins by finding a set of k-most vital arcs in G, and then solves a sequence of at most k shortest path problems. Thus, its practical complexity is that of the k-most vital arcs problem, for which there exist effective solution algorithms (Israeli and Wood 2002). The intuition behind the algorithm is based on the following observation. Suppose $\mathcal{I}^*$ is a set of k-most vital arcs of G. Then, starting from the set $I_1 = A_0 \cap \mathcal{I}^*$, the evader's response each time period must reveal at least one arc in $\mathcal{I}^* \setminus I_t$. Thus, one can reconstruct $\mathcal{I}^*$ in at most k time periods simply by solving a shortest path in each period, and adding the newly revealed elements of $\mathcal{I}^*$ into the blocking action.

The pseudo-code of the approach is provided in Algorithm 1, where MostVitalArcs(G,k) returns a set of k most vital arcs in G and the length of the optimal solution, and ShortestPath(G) returns the primal and dual solution to the LP formulation of the shortest path problem, as well as the optimal path length. Also, $\mathbf{1}\{\cdot\}$ denotes the indicator function.

Feeding the initial feasible solution given by Algorithm 1 to the MIP solver decreases the overall solution time of LB (see results in Section 5). However, for sufficiently large values of T any MIP-oriented solution approach is not effective.[6] Note, however, that Algorithm 1 provides an approach for finding a feasible solution of LB, which identifies a set of k-most vital arcs within k time periods,

---

[6] Increasing T by Δt increases the number of variables and constraints of LB by a factor of $\Theta(\Delta t \times |A|)$, which translates into an exponential increase in the worst-case time performance for any MIP solver based on branch-and-bound ideas.

Algorithm 1 Finding a feasible solution for $LB(G, \mathcal{C}_0, T)$

Require: $G = (N, A, C)$, A_0, k, and T

  $[\mathcal{I}^*, z^*] =$ MostVitalArcs(G, k)

  $I_0 = \emptyset$, $[p^0, y^0, z^0] =$ ShortestPath$(G \setminus I_0)$, $t = 0$

  while $z^* > z^t$ and $t \leq T$ **do**

    $t = t + 1$

    $I_t = I_{t-1} \cup \left(\{ (i, j) \in \mathcal{I}^* : \ p_{ij}^{t-1} = 1 \} \right)$

    $[p^t, y^t, z^t] =$ ShortestPath$(G \setminus I_t)$

  end while

  Set $r_{ij}^s := \mathbf{1} \{ (i, j) \in I_s \} \ \forall s \leq t$, and $r^s = r^t$, $p^s = p^t$, $y^s = y^t \ \forall s > t$.

  return $\ \{ (r^t, p^t, y^t) : \ t \in \mathcal{T} \}$

---

and then, these k-most vital arcs are successively repeated until time T. This suggests that it might be possible to solve LB for a relatively short horizon and extend the solution to a larger horizon. (Note that this is not always possible, as there exist networks such that the optimal solution of LB does not involve discovering a set of k-most vital arcs sufficiently early, or even at all.)

Algorithm 2 incorporates the ideas above. There, T_0 corresponds to the time period in which a k-most vital arc solution is first discovered in Algorithm 1. The algorithm iterates from time $T' = T_0$ to time $T' = T$, solving $LB(G, \mathcal{C}_0, T')$ at each time. If the solution of $LB(G, \mathcal{C}_0, T')$ involves discovering a set of k-most vital arcs, then it is optimal to extend such set up to time T. Otherwise, the algorithm sets $T' = T' + 1$, and $LB(G, \mathcal{C}_0, T')$ is solved again.

(Note: efficiency of the algorithm is improved by providing an initial feasible solution of (10) to the MIP solver each time it is called. Such solutions can easily be constructed initially from Algorithm 1, and later from the solution of (10) in the previous iteration.)

PROPOSITION 3. *Algorithm 2 correctly solves $LB(G, \mathcal{C}_0, T)$.*

Proof. See Appendix B. ∎

4.2. Lower Bound for Time-Stability

We extend the ideas in the previous section to the case of time-stability. In particular, for $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and T, the oracle interdictor solves the following MIP:

$$TS(G, \mathcal{C}_0, T): \quad \min \sum_{t=0}^{T} w_t$$

$$\text{s.t.} \quad z^*(1 - w^t) \leq y_1^t - y_n^t \quad \forall t \in \mathcal{T}, \tag{12a}$$

$$w^t \in \{0, 1\} \qquad\qquad \forall t \in \mathcal{T}, \tag{12b}$$

Algorithm 2 Solving $LB(G, \mathcal{C}_0, T)$

Require: $G = (N, A, C)$, A_0, k, T

  Use Algorithm 1 to find z^* and T_0

  if $T_0 \geq T$ **then**

    Solve $LB(G, \mathcal{C}_0, T)$ via formulation (10)

  else

    Set $T' = T_0$

    Solve $LB(G, \mathcal{C}_0, T')$ via formulation (10) and denote by $\left\{ (\tilde{r}^t, \tilde{p}^t, \tilde{y}^t) : \ \forall t \leq T' \right\}$ its solution

    while $z^* > y_1^{T'} - y_n^{T'}$ and $T' < T$ **do**

      Set $T' = T' + 1$

      Solve $LB(G, \mathcal{C}_0, T')$ via formulation (10) and denote by $\left\{ (\tilde{r}^t, \tilde{p}^t, \tilde{y}^t) : t \leq T' \right\}$ its solution

    end while

    Set $(r^t, p^t, y^t) = (\tilde{r}^t, \tilde{p}^t, \tilde{y}^t)$ for all $t \leq T'$, and $(r^t, p^t, y^t) = (\tilde{r}^{T'}, \tilde{p}^{T'}, \tilde{y}^{T'})$ for all $t > T'$

  end if

  return $\ \left\{ (r^t, p^t, y^t) : \ t \in \mathcal{T} \right\}$

$$\text{and constraints (10a) to (10h).}$$

In this formulation, w^t indicates whether the blocking decision in period t, $t \in \mathcal{T}$, is a set of k-most vital arcs for G or not. Note that constraints (12a) force $w^t = 1$ when the evader's path length at period t is lower than z^* (otherwise, $w^t = 0$ due to the objective function) for $t \in \mathcal{T}$. Let (r^*, p^*, y^*, w^*) denote an optimal solution to (12). As in the previous section, denote by

$$I_t^*(G, \mathcal{C}_0, T) := \left\{ (i,j) \in A : \ r_{ij}^{*t} = 1 \right\} \quad \text{and} \quad P_t^*(G, \mathcal{C}_0, T) := \left\{ (i,j) \in A : \ p_{ij}^{*t} = 1 \right\}, \quad t \in \mathcal{T}$$

the *oracle-based* policy (for time-stability). (Note that this policy is not necessarily in Π; recall Remark 8, which can be extended for the case of time-stability). Also, as in the previous section, we have that time-stability of the oracle based policy is a lower bound for time-stability of policies in Π. That is, for any $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and T:

$$\sum_{t \in \mathcal{T}} w^{*t} \leq \tau^\pi(G, \mathcal{C}_0), \qquad \pi \in \Pi. \tag{13}$$

Solving (12) entails the same difficulties that are faced when solving LB. In this regard, Algorithm 1 also provides a feasible solution to the formulation above, provided one sets $w^t = 1$ for all $t < T_0$, and $w^t = 0$, otherwise. Note, however, that unlike in the case of regret minimization, T_0 provides an upper bound to the time-stability of the oracle-based policy. Thus, it is sufficient to solve $TS(G, \mathcal{C}_0, T_0)$ to generate an optimal solution for $TS(G, \mathcal{C}_0, T)$, where $T \geq T_0$. The solution procedure is summarized in Algorithm 3. Its correctness follows from the fact that T_0 is an upper bound for time-stability of the oracle-based policy.

Algorithm 3 Solving $TS(G, \mathcal{C}_0, T)$

Require: $G = (N, A, C)$, A_0, k, T

   Use Algorithm 1 to find z^* and T_0

   if $T_0 \geq T$ **then**

      Solve $TS(G, \mathcal{C}_0, T)$ via formulation (12)

   else

      Solve $TS(G, \mathcal{C}_0, T_0)$ via formulation (12) and denote by $\{(\tilde{r}^t, \tilde{p}^t, \tilde{y}^t, \tilde{w}^t) \colon t \leq T_0\}$ its solution

      Set $(r^t, p^t, y^t, w^t) = (\tilde{r}^t, \tilde{p}^t, \tilde{y}^t, \tilde{w}^t)$ for $t \leq T_0$, and $(r^t, p^t, y^t, w^t) = (\tilde{r}^{T_0}, \tilde{p}^{T_0}, \tilde{y}^{T_0}, 0)$ for $t > T_0$

   end if

   return $\{(r^t, p^t, y^t, w^t) \colon t \in \mathcal{T}\}$

---

5. Computational Study

In this section, we study the practical performance of the proposed policies and algorithms. First, in Section 5.1 we describe our test instances, three additional benchmark policies and the implementation details. Then in Section 5.2 we briefly analyze the performance of Algorithms 1 and 2 for computation of the oracle-based policies. In Section 5.3 we compare the performance of the proposed policies (namely, Γ and Λ) against benchmark policies. We also conduct sensitivity analysis of policies with respect to the amount and the *quality* of information initially available to the interdictor in Sections 5.4 and 5.5, respectively.

5.1. Test Instances, Benchmark Policies and Implementation Details

Network Structure and Arc Costs. We test our policies using the class of uniform random graphs (uniform graphs) of Erdös and Rényi (1959). The cost structure of each graph instance is generated as follows. First, for each arc $a \in A$, the bounds l_a and u_a are drawn randomly (in sequence) from uniform distributions $U(0, 500)$ and $U(l_a, 500)$, respectively. Then cost c_a is set to $l_a + (u_a - l_a)x_a$, where x_a is drawn from a Beta(α, β) distribution. As policy performance might be sensitive to the relative location of c_a in $[l_a, u_a]$, we consider *left-skewed*, *symmetric* and *right-skewed* cost distributions by setting (α, β) to $(2, 10)$, $(10, 10)$ and $(10, 2)$, respectively.

Benchmark Policies. We consider three additional benchmark policies:

   (i) The *lower bound* policy π_L interdicts a set of k-most vital arcs in the observed network, assuming that the cost of $a \in \widehat{A}_t$ is $c_a = l_a$. That is, the policy operates as a policy in Λ, but uses the lower bound l_a instead of u_a in equation (5).

   (ii) The *mean bound* policy π_M interdicts a set of k-most vital arcs in the observed network, assuming that the cost of $a \in \widehat{A}_t$ is $c_a = (l_a + u_a)/2$.

   (iii) The *random bound* policy π_R interdicts a set of k-most vital arcs in the observed network, assuming that the cost of $a \in \widehat{A}_t$ is either $c_a = l_a$ or $c_a = u_a$ with equal probability.

Initial Information. We design instances, where a fraction $p_a \in [0, 1]$ of the arcs from A is included into A_0, and a fraction $p_c \in [0, 1]$ of A_0 is included into $\widetilde{A}_0$. Specifically, for a given pair (p_a, p_c), starting from $A_0 = \emptyset$, arcs are selected randomly (uniformly) without replacement from A and added to A_0, until $|A_0| = \lfloor |A| \times p_a \rfloor$. Then, starting from $\widetilde{A}_0 = \emptyset$, arcs are selected randomly (uniformly) without replacement from A_0 and added to $\widetilde{A}_0$, until $|\widetilde{A}_0| = \lfloor |A_0| \times p_c \rfloor$. We construct these sets in a nested fashion, i.e., the set A_0 generated for p_a is a subset of that generated for $p'_a > p_a$. The same applies for $\widetilde{A}_0$.

Implementation Details. The algorithms are coded in Matlab R2012b and all the experiments are performed on a Windows PC with 3.7GHz CPU and 32GB RAM. We solve (10) and (12) using CPLEX 12.4. For finding a k-most vital arc solution we use the *basic covering decomposition algorithm* from Israeli and Wood (2002).

5.2. Computation of the Oracle-based Policy

In this section we demonstrate performance of Algorithms 1 and 2 by comparing three different approaches for solving LB. In the first approach, referred to as "MIP," we feed formulation (10) directly to a state-of-the-art MIP solver. In the second, which we denote as "MIP i.s.," we use Algorithm 1 to generate an initial feasible solution to LB, which is then fed to the MIP solver together with (10). In the third, which we refer to as "Alg.," we use Algorithm 2.

We test efficiency of these three solution procedures using 10 randomly generated uniform graphs, each considering $n = 40$, $p = 0.5$, $p_a = p_c = 0$ and $T = 15$. Costs are drawn from the symmetric distribution. Table 2 summarizes the running-time (in seconds) of each solution approach, where $k \in \{2, 4, 6, 8\}$. We set a time limit of an hour for all methods.

Table 2   Running times (in seconds) to solve LB. The entry "-" implies that an optimal solution was not found within one hour.

k=2			k=4			k=6			k=8		
Alg.	MIP i.s.	MIP	Alg.	MIP i.s.	MIP	Alg.	MIP i.s.	MIP	Alg.	MIP i.s.	MIP
1.5	17.9	29.9	5.4	59.0	124.5	35.9	707.4	-	-	-	-
0.6	4.8	40.1	2.2	22.1	63.3	4.7	169.4	205.6	-	-	-
1.5	9.6	33.3	3.1	122.1	641.1	22.5	1658.8	-	534.9	-	-
1.4	215.5	-	14.7	1799.5	2626.7	15.1	-	-	93.1	-	-
1.2	6.4	21.0	3.1	19.3	57.0	8.9	285.9	697.0	29.2	729.7	711.8
1.3	6.4	90.1	5.1	48.8	74.3	15.1	447.2	447.8	41.4	-	-
1.1	5.5	24.3	8.2	476.2	1742.1	36.7	-	-	-	-	-
1.0	4.5	20.7	1.8	12.9	95.7	3.1	115.4	591.3	130.6	-	2244.6
1.4	9.4	72.3	16.1	685.4	1425.1	33.5	-	-	190.4	-	-
1.3	35.3	2842.9	9.5	1104.3	-	9.4	1166.1	-	18.5	-	-

We observe that, in general, the solution time of Algorithm 2 is between one or two orders of magnitude less than that of the MIP solver. Also, we note that while providing an initial feasible solution

to the solver improves its performance, Algorithm 2 is still significantly faster than the other methods. This suggests that the practical efficiency of Algorithm 2 can be mostly attributed to the idea of extending the solution of LB with a shorter time horizon (recall our discussion in Section 4.1).

5.3. Comparison to Benchmark Policies

In this section we compare the performance of $\lambda \in \Lambda$ and $\gamma \in \Gamma$ against the benchmark policies. For each pair (p_a, p_c), where $p_a, p_c \in \{0, 1/3, 2/3, 1\}$, we generate 20 random networks along with subsets A_0 and $\widetilde{A}_0$, and for each of them we generate three different cost vectors corresponding to each cost distribution. We also set $n = 40$, $p = 0.5$, $k = 6$ and $T = 21$.

First, we consider the settings from Section 3.1, where $\widehat{A}_0 = \emptyset$ (i.e., $p_c = 1$). (Note that in this case, classes Γ, Λ, π_L, π_M and π_R are equivalent.) Tables 3 and 4 summarize the regret and time-stability for each setting, respectively. There, γ denotes a policy in Γ, and each entry represents the average performance among all 20 instances and the mean absolute deviation (MAD), in parenthesis. To quantify the value of the initial information (in particular, the size of A_0), we include the performance of the oracle-based policy for the corresponding performance metrics (regret and time-stability), π^{oracle}.

Table 3   Average cumulative regret $(\times 10^2)$ and MAD (in parenthesis) for $k = 6$.

	Left-Skewed		Symmetric		Right-Skewed	
p_a	γ	π^{oracle}	γ	π^{oracle}	γ	π^{oracle}
0	5.83 (2.12)	2.27 (0.64)	6.55 (2.57)	3.72 (1.20)	9.42 (3.32)	4.70 (1.46)
1/3	4.28 (1.19)	1.19 (0.49)	5.41 (1.94)	1.95 (0.82)	8.10 (2.72)	2.39 (1.03)
2/3	1.78 (1.03)	0.31 (0.34)	2.29 (1.04)	0.29 (0.32)	3.47 (1.49)	0.87 (0.76)

Table 4   Average time-stability and MAD (in parenthesis) for $k = 6$.

	Left-Skewed		Symmetric		Right-Skewed	
p_a	γ	π^{oracle}	γ	π^{oracle}	γ	π^{oracle}
0	12.55 (2.55)	4.95 (0.57)	9.60 (1.86)	5.40 (0.60)	10.50 (1.10)	5.05 (0.48)
1/3	9.75 (2.05)	3.15 (0.70)	8.30 (1.66)	3.65 (0.86)	8.95 (1.76)	3.70 (0.94)
2/3	5.15 (1.87)	1.80 (0.56)	4.30 (1.53)	1.65 (0.52)	5.10 (1.22)	2.2 (0.86)

We observe that the performance of policy $\gamma \in \Gamma$ is roughly between 2 to 4 times that of the oracle-based policy. Given that π^{oracle} has complete information about the network structure and arc costs, the difference in the performance of the policies is reasonably modest. (For example, in the experiments discussed below $\lambda \in \Lambda$ performs orders of magnitude better than the other benchmark policies, namely, π_L, π_M and π_R, in a number of test instances.) As one would expect, as the value of p_a increases (i.e., the size of the initially known arc subset A_0 increases), the performance of both policies improves. Also, it is worth noting that the performance of the oracle policy differs

significantly from zero, the value it would obtain if it was possible to signal availability of complete initial information.

Table 5 depicts the running-time statistics for computation of π^{oracle} (under the regret performance metric, thus, using Algorithm 2), which correspond to the results reported in Table 3. Similarly, in Table 6 we report the running-time statistics for computation of π^{oracle} (under the time-stability performance metric, thus, using Algorithm 3), which correspond to the results reported in Table 4. (The running time for computation of γ is not reported as for all instances it is less than 5 and 1 seconds for the regret and time-stability metrics, respectively.) It can be observed that, in general, the left-skewed case is significantly more time consuming. Furthermore, it is interesting to note that computing the oracle-based policy is more challenging for the regret performance metric than for time-stability.

Table 5   Average running times (in seconds) and MAD (in parenthesis) for computing π^{oracle} using Algorithm 2.

p_a	Left Skewed	Symmetric	Right Skewed
0	177.81 (255.77)	99.87 (153.27)	27.05 (21.24)
1/3	447.02 (756.74)	59.46 (32.97)	66.93 (30.91)
2/3	639.82 (1175.40)	22.88 (13.41)	48.92 (36.51)

Table 6   Average running times (in seconds) and MAD (in parenthesis) for computing π^{oracle} using Algorithm 3.

p_a	Left Skewed	Symmetric	Right Skewed
0	41.86 (50.50)	13.53 (9.77)	12.16 (7.52)
1/3	25.95 (19.60)	29.57 (17.02)	32.32 (17.69)
2/3	10.69 (5.71)	10.32 (5.32)	18.52 (10.84)

Next, we consider instances with $\widehat{A}_0 \neq \emptyset$ and $p_c \in \{0, 1/3, 2/3\}$. Tables 7 and 8 summarize the regret and time-stability for each setting, respectively. There, γ and λ denote policies in Γ and Λ, respectively. As before, each entry represents the average performance among all 20 instances and MAD, in parenthesis. Note that because policies in Γ are defined for settings where $\widehat{A}_0 = \emptyset$, these policies discard any information in $\widehat{A}_0$.

Observe that, in general, policies λ and π_M yield the best results with respect to regret and time-stability, while the performance of π_L and π_R is significantly worse. For left-skewed and symmetric structures policy π_M is generally the best one, while performance of λ is not far behind and very close to π_M. For the right-skewed cost structure λ is the best policy, getting a significant difference for regret with respect to all other policies (including π_M). The difference is highly amplified when considering time-stability, as the other policies, except γ, get very close to the worst possible value ($\tau \approx 22$). Moreover, for many instances policies π_M, π_L and π_R fail to identify a set of k-most vital arcs. Recall that unless one assumes that arc costs are at their upper bounds, there is no guarantee that a policy achieves zero instantaneous regret (see Remark 7).

Borrero, Prokopyev, and Sauré: *Sequential Shortest Path Interdiction with Incomplete Information*

Table 7 **Average regret** $(\times 10^2)$ **and MAD (in parenthesis) for** $k = 6$. **Among the entries denoting regret, the entry in bold is the best value, and the other entries indicate the difference with respect to the best value.**

(p_a, p_c)	Left-Skewed					Symmetric					Right-Skewed				
	λ	π_L	π_M	π_R	γ	λ	π_L	π_M	π_R	γ	λ	π_L	π_M	π_R	γ
$(\frac{1}{3},0)$	0.54	0.93	**5.17**	0.87	0.67	0.32	7.70	**5.58**	2.98	0.97	**8.52**	17.26	3.03	6.98	0.90
	(1.42)	(3.54)	(1.49)	(1.93)	(2.12)	(2.04)	(4.80)	(1.88)	(2.54)	(2.57)	(3.10)	(9.85)	(4.40)	(3.49)	(3.32)
$(\frac{1}{3},\frac{1}{3})$	0.27	0.26	**5.16**	0.27	0.35	0.52	3.96	**5.53**	1.61	1.37	8.37	6.97	1.61	3.43	1.15
	(1.51)	(2.39)	(1.46)	(1.79)	(1.81)	(2.11)	(4.12)	(1.95)	(2.05)	(2.29)	(2.95)	(5.93)	(3.92)	(4.20)	(3.62)
$(\frac{1}{3},\frac{2}{3})$	0.72	**4.28**	0.48	0.62	1.01	0.13	1.76	**5.62**	0.93	0.44	8.58	3.37	0.81	2.29	0.63
	(1.44)	(1.09)	(1.29)	(1.53)	(1.82)	(1.71)	(2.17)	(1.94)	(1.74)	(1.96)	(3.14)	(5.13)	(3.81)	(3.78)	(3.72)
$(\frac{2}{3},0)$	0.62	5.51	**4.68**	1.49	1.16	1.63	23.18	**3.05**	11.05	3.50	4.71	35.69	6.58	23.53	4.71
	(1.59)	(5.05)	(1.70)	(2.06)	(2.12)	(1.47)	(9.64)	(1.35)	(5.12)	(2.57)	(1.49)	(14.14)	(5.93)	(5.44)	(3.32)
$(\frac{2}{3},\frac{1}{3})$	0.61	3.44	**3.63**	0.90	1.51	1.22	15.13	**2.99**	7.77	3.49	4.16	28.47	4.73	15.76	4.74
	(1.57)	(3.63)	(1.44)	(1.75)	(1.77)	(1.63)	(5.00)	(1.46)	(3.25)	(2.21)	(1.73)	(9.88)	(5.57)	(5.21)	(3.21)
$(\frac{2}{3},\frac{2}{3})$	0.53	1.82	**2.82**	0.90	0.87	0.59	9.09	**2.52**	3.69	1.82	3.84	16.24	1.49	8.74	3.05
	(1.45)	(2.77)	(1.33)	(2.09)	(1.51)	(1.31)	(5.87)	(1.16)	(2.83)	(2.04)	(1.76)	(8.20)	(2.80)	(4.68)	(3.04)
$(1,0)$	1.32	10.20	**3.27**	3.99	2.57	1.87	28.72	**1.30**	17.21	5.24	**0.89**	46.13	10.08	33.05	8.53
	(1.27)	(5.37)	(1.03)	(2.65)	(2.12)	(1.14)	(8.87)	(1.48)	(5.20)	(2.57)	(0.46)	(15.06)	(7.11)	(7.15)	(3.32)
$(1,\frac{1}{3})$	1.35	6.15	**2.38**	2.65	2.37	1.35	21.22	**1.04**	12.77	4.66	**0.45**	32.11	9.87	25.58	7.08
	(1.38)	(4.37)	(0.98)	(2.32)	(1.59)	(1.03)	(9.07)	(1.38)	(4.55)	(2.14)	(0.29)	(12.83)	(5.74)	(5.87)	(2.52)
$(1,\frac{2}{3})$	0.46	4.03	**1.05**	1.29	0.52	0.45	15.17	**0.47**	7.90	1.97	**0.25**	25.46	5.88	14.87	3.11
	(0.58)	(3.72)	(0.55)	(1.00)	(0.76)	(0.61)	(5.31)	(0.63)	(3.60)	(1.28)	(0.26)	(10.09)	(5.16)	(5.80)	(1.69)

Table 8 **Average time-stability and MAD (in parenthesis) for** $k = 6$. **Among entries denoting time-stability, the entry in bold is the best value, and the other entries indicate the difference with respect to the best value. The entries in italic and "$-$" mean that the policy did not attain time-stability for some instances.**

(p_a, p_c)	Left-Skewed					Symmetric					Right-Skewed				
	λ	π_L	π_M	π_R	γ	λ	π_L	π_M	π_R	γ	λ	π_L	π_M	π_R	γ
$(\frac{1}{3},0)$	1.1	3.35	**11.9**	3.25	0.65	0.5	*12.1*	**9.05**	9.1	0.55	**9.75**	*12.25*	7.8	*12.25*	0.75
	(1.60)	(6.83)	(1.73)	(3.68)	(2.55)	(1.31)	(-)	(2.27)	(4.04)	(1.86)	(1.85)	(-)	(6.23)	(-)	(1.10)
$(\frac{1}{3},\frac{1}{3})$	0.55	2.4	**11.7**	1.9	1.3	0.8	9.35	**8.95**	6.3	1.3	9.35	*12.65*	5.8	10.1	1.3
	(1.95)	(6.41)	(2.10)	(3.34)	(1.96)	(1.65)	(5.55)	(2.64)	(5.48)	(1.65)	(1.86)	(-)	(6.85)	(3.58)	(1.78)
$(\frac{1}{3},\frac{2}{3})$	1.05	**10.5**	0.5	1.3	1.35	0.3	6.05	**9.05**	4.55	0.5	**9.4**	10.35	4.4	9.25	0.85
	(2.05)	(3.40)	(1.80)	(2.04)	(2.28)	(1.49)	(6.90)	(2.37)	(4.68)	(1.96)	(1.78)	(3.83)	(6.58)	(5.03)	(2.08)
$(\frac{2}{3},0)$	0.8	9.55	**11.4**	6.8	1.15	2.35	*15.25*	**6.75**	*15.25*	2.85	**7.3**	*14.7*	12.75	*14.7*	3.2
	(1.72)	(2.00)	(2.20)	(4.36)	(2.55)	(1.23)	(-)	(2.58)	(-)	(1.86)	(1.29)	(-)	(-)	(-)	(1.10)
$(\frac{2}{3},\frac{1}{3})$	1.05	10.9	**9.2**	7	2.5	1.05	*14.9*	**7.1**	*14.2*	2.7	**6.6**	*15.4*	11.65	*15.4*	3.25
	(2.23)	(3.42)	(1.90)	(5.06)	(1.76)	(1.20)	(-)	(3.44)	(-)	(1.72)	(1.26)	(-)	(6.00)	(-)	(1.94)
$(\frac{2}{3},\frac{2}{3})$	0.55	9.8	**7.7**	4	1.92	1.15	*16*	**5**	11.65	2.3	**5.75**	*15.35*	6.65	*15.8*	2.15
	(1.75)	(6.75)	(1.77)	(5.84)	(1.92)	(1.60)	(-)	(1.50)	(6.76)	(1.66)	(1.45)	(-)	(8.64)	(-)	(1.51)
$(1,0)$	2.1	*12.8*	**9.2**	*11.1*	3.35	**7.8**	*14.2*	1.2	*13.6*	1.8	**4.4**	*17.6*	15.5	*17.6*	6.1
	(1.39)	(-)	(1.70)	(-)	(2.55)	(1.26)	(-)	(9.10)	(-)	(1.86)	(1.30)	(-)	(3.78)	(-)	(1.10)
$(1,\frac{1}{3})$	1.85	*14.7*	**7.3**	11.15	3.35	**6.3**	*15.7*	1.4	*15.7*	2.4	**3**	*17.95*	15.85	*19*	6.05
	(1.35)	(-)	(1.60)	(4.27)	(1.75)	(1.47)	(-)	(8.58)	(-)	(1.97)	(1.10)	(-)	(5.36)	(-)	(1.27)
$(1,\frac{2}{3})$	0.95	*17.25*	**3.7**	11.85	0.75	**3.45**	*18.55*	1.2	17.4	1.15	**1.8**	*20.2*	14.95	*20.2*	2.9
	(1.52)	(-)	(1.44)	(5.90)	(1.64)	(1.23)	(-)	(5.21)	(1.97)	(1.12)	(0.72)	(-)	(7.88)	(-)	(1.33)

Our experimental observations corroborate our theoretical results: it is crucial for the interdictor to have a pessimistic attitude regarding the unobserved costs, i.e., it is better to overestimate the real costs of the arcs whose real costs are unknown rather than underestimate them. The results also suggest that policy λ is robust in the sense that it has a consistently good performance across all cost settings, although not always yields the best result. We also note that policy γ has the same type of robust behavior, although it is typically outperformed by λ, which signals that it is valuable to exploit the cost bounds information.

5.4. Policy Performance: Sensitivity with Respect to $|\widetilde{A}_0|$

In this section we study the performance of policies in Λ as a function of the number of arcs for which the real cost is initially known (i.e., the amount of initial information). We also consider policy π_M as a benchmark, due to its consistent performance in the experiments in Section 5.3.

We set $p_a = 1/2$ and consider $p_c \in \{i/10 \colon 1 \le i \le 10\}$. As before, for every pair (p_a, p_c) we generate 20 networks with different cost structures (right-skewed, symmetric and left-skewed). Also, we set $n = 40$, $p = 0.7$ (giving us an average of 1089.8 arcs), $k = 8$ and $T = 28$. Figures 9 and 10 depict the results for the right-skewed and symmetric cost structure, respectively. The results for the left-skewed case are in Figure 12 in Appendix C.
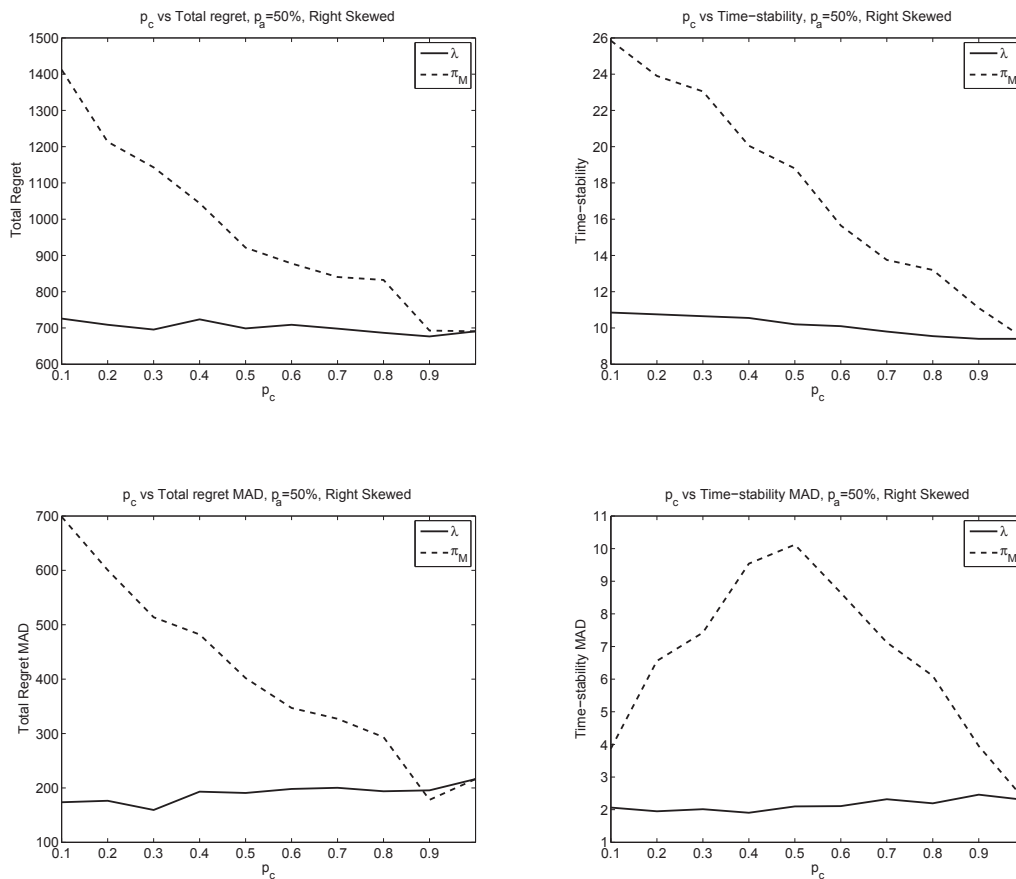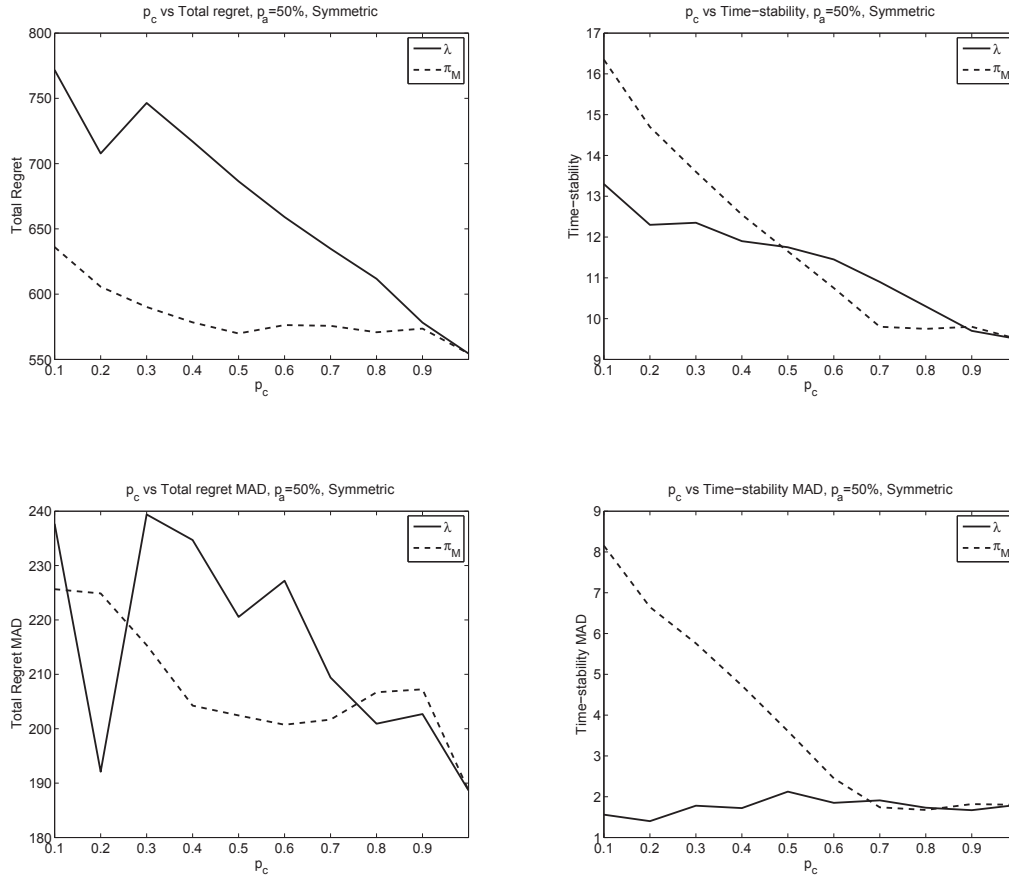

Figure 9   Behavior of the average time-stability, average total regret, time-stability MAD and total regret MAD as p_c increases. The cost distribution is right-skewed and $p_a = 1/2$.

For the right-skewed case it is observed that λ is roughly constant at low values across all measures, indicating good and consistent performance. On the other hand, performance of π_M improves as there is more initial information of the network available, but is significantly worse

Figure 10   **Behavior of the average time-stability, average total regret, time-stability MAD and total regret MAD as p_c increases. The cost distribution is symmetric and $p_a = 1/2$.**

than the one from λ. Also, note that time-stability MAD for π_M has a parabolic behavior, which points out that for low (high) amounts of initial information the time-stability of π_M is consistently high (low), while high variability is observed for intermediate values of initial information.

For the symmetric distribution, both λ and π_M average regret and time-stability decrease as more information is available (λ regret has a subtle increase at one point due to the occurrence of an instance in which it performed significantly bad). In this setting π_M outperforms λ in terms of regret, however their difference is not so great when compared to the right-skewed case (observe the scale of the y-axis). Regarding time-stability, it is seen that π_M is slightly better than λ for low amounts of initial information while the opposite is observed for large amounts. On the other hand both λ and π_M regret MAD don't show any particular pattern (however it can be considered relatively constant after noting its scale is significantly smaller than the one in the right skewed case), while λ time-stability MAD is constant at low values and that of π_M tends to decrease.

These observations confirm our previous conclusions regarding the robustness property of policy λ. They also suggest that different from π_M, performance of λ is highly consistent with respect to changes in the amount of initial information. Moreover, it is observed that the sensitivity of λ to the amount of initial information can depend on the location of the true cost between the lower and upper bounds: if it close to the upper bound (right-skewed), the performance tends to be fairly unsensitive, while for the other cases it tends to decrease as more information is initially available.

It is important to note that, unlike for policies in Λ, the regret of the benchmark policies π_L, π_M and π_R may not converge in some instances (recall our Remark 7), thus in the long run (i.e., for sufficiently large values of T, e.g., $T \geq |A|$) the proposed policies might outperform the benchmark policies with certainty. However, such a feature does not rule out that: (i) the benchmark policies may converge for many instances and (ii) the regret of the benchmark policies may be smaller than the regret attained by the convergent policies, particularly in the case of finite horizons, as it can be observed for some instances in Tables 7 and 8 as well as in Figures 10 and 12.

5.5. Policy Performance: Sensitivity with Respect to Quality of Bounds in $\widehat{A}_0$

We conclude our numerical experiments by studying the performance of the policies in Λ and π_M as the quality of the initial information deteriorates, i.e., as $u_a - l_a$ increases for all $a \in \widehat{A}_0$. To this end, in this set of experiments we generate a cost vector by drawing c_a uniformly from $U(500, 1000)$ for each $a \in A$. We consider three sets of cost bounds: $(c_a - x_a, c_a + y_a)$, $(c_a - 5x_a, c_a + 5y_a)$ and $(c_a - 25x_a, c_a + 25y_a)$, where x_a and y_a are drawn uniformly from $[1, 20]$ for all $a \in A$. We refer to these three intervals as "I. #1," "I. #2," and "I. #3," respectively.

As in the previous experiments, for each pair (p_a, p_c) we generate 20 random networks along with subsets A_0 and $\widetilde{A}_0$, and for each of them we generate the cost vector $(c_a)_{a \in A}$ and x_a, y_a for all $a \in \widehat{A}_0$. We consider $n = 50$ nodes, $p = 0.5$ (the mean number of arcs is 1216.35.), $k = 15$ and $T = 53$. Table 9 summarizes the obtained results.

We observe that both λ and π_M are sensitive to changes in the quality of the information: as the intervals widen, performance deteriorates. Note that this effect is significantly more pronounced if the interdictor has more initial information available, i.e., for larger values of p_a and p_c. Policy π_M tends to be better than λ in regret for narrow cost intervals, i.e., in the cases with *good* information quality. However, λ is better (in particular, with respect to MAD values) as the intervals widen and the values of p_a increase. Similarly, π_M and λ are roughly similar for time-stability for narrow cost intervals, but λ is significantly better as these intervals widen and p_a increases. These results point out again at the *robust* behavior of the policies in Λ, i.e., they have a good performance across all instances, although they are not always the best.

In order to further validate the aforementioned conclusions, we design a similar experiment with more quality levels. Specifically, we pick $p_a = 2/3$, $p_c \in \{0, 1/3, 2/3\}$ and generate the cost vector

Table 9 Average regret ($\times 10^3$) and time-stability, and MAD (in parenthesis) for $k = 15$. The entries in bold denote the best value.

	Mean regret						Mean time-stability					
	Policy λ			Policy π_M			Policy λ			Policy π_M		
(p_a, p_c)	I. #1	I. #2	I. #3	I. #1	I. #2	I. #3	I. #1	I. #2	I. #3	I. #1	I. #2	I. #3
$(0,0)$	**49.95**	**49.95**	**49.95**	**49.95**	**49.95**	**49.95**	**20.65**	**20.65**	**20.65**	**20.65**	**20.65**	**20.65**
	(8.83)	(8.83)	(8.83)	(8.83)	(8.83)	(8.83)	(3.15)	(3.15)	(3.15)	(3.15)	(3.15)	(3.15)
$(\frac{1}{3},0)$	**44.52**	45.31	48.24	44.61	**44.61**	47.37	**19.40**	20.15	22.00	**19.00**	20.45	31.90
	(9.07)	(9.93)	(9.36)	(9.07)	(9.23)	(11.13)	(3.08)	(3.27)	(3.40)	(2.90)	(4.83)	(16.88)
$(\frac{1}{3},\frac{1}{3})$	**44.39**	44.80	46.69	44.71	**44.46**	47.39	**19.20**	19.80	20.70	**20.40**	**20.40**	31.75
	(8.96)	(9.46)	(9.67)	(9.16)	(9.26)	(11.64)	(3.26)	(3.18)	(3.24)	(4.52)	(4.86)	(17.00)
$(\frac{1}{3},\frac{2}{3})$	44.58	44.42	44.69	44.84	**44.70**	45.01	**19.20**	19.40	19.30	19.25	**19.05**	22.30
	(9.19)	(8.88)	(9.00)	(9.25)	(9.10)	(9.82)	(3.26)	(2.88)	(3.29)	(3.20)	(2.86)	(7.68)
$(\frac{1}{3},1)$	**44.74**	**44.74**	**44.74**	**44.74**	**44.74**	**44.74**	**19.15**	**19.15**	**19.15**	**19.15**	**19.15**	**19.15**
	(9.14)	(9.14)	(9.14)	(9.14)	(9.14)	(9.14)	(3.20)	(3.20)	(3.20)	(3.20)	(3.20)	(3.20)
$(\frac{2}{3},0)$	**25.99**	27.94	39.32	**25.49**	25.79	38.34	**12.55**	15.30	20.00	**13.40**	**13.40**	38.60
	(8.23)	(8.60)	(7.55)	(8.25)	(8.22)	(16.29)	(2.85)	(3.13)	(3.00)	(4.76)	(4.70)	(18.72)
$(\frac{2}{3},\frac{1}{3})$	**25.74**	26.58	34.02	**25.42**	26.97	37.29	**12.15**	13.30	16.80	**13.30**	17.60	36.30
	(8.65)	(7.99)	(8.35)	(8.19)	(9.53)	(13.69)	(2.95)	(2.84)	(2.60)	(4.62)	(10.76)	(20.04)
$(\frac{2}{3},\frac{2}{3})$	**25.67**	26.14	30.25	**25.35**	25.38	29.04	**11.75**	12.25	14.15	**11.15**	13.20	25.60
	(8.11)	(8.78)	(8.63)	(8.34)	(8.51)	(9.58)	(2.50)	(2.65)	(2.62)	(2.45)	(4.60)	(19.18)
$(\frac{2}{3},1)$	**25.31**	**25.31**	**25.31**	**25.31**	**25.31**	**25.31**	**11.00**	**11.00**	**11.00**	**11.00**	**11.00**	**11.00**
	(8.36)	(8.36)	(8.36)	(8.36)	(8.36)	(8.36)	(2.30)	(2.30)	(2.30)	(2.30)	(2.30)	(2.30)
$(1,0)$	**0.20**	3.78	25.61	**0.19**	1.58	32.84	**3.15**	7.60	17.05	**6.35**	12.10	43.35
	(0.17)	(1.59)	(3.69)	(0.33)	(2.25)	(21.15)	(1.47)	(2.04)	(2.07)	(9.33)	(16.36)	(15.44)
$(1,\frac{1}{3})$	**0.07**	1.91	17.90	0.21	2.34	23.60	**2.25**	5.15	12.65	**9.00**	19.45	38.05
	(0.06)	(1.00)	(3.53)	(0.35)	(3.17)	(19.03)	(1.10)	(1.90)	(2.22)	(13.20)	(23.49)	(20.93)
$(1,\frac{2}{3})$	**0.05**	0.93	8.66	**0.08**	0.63	8.33	**1.75**	3.60	7.05	**6.35**	11.65	22.40
	(0.05)	(0.63)	(3.44)	(0.14)	(0.97)	(10.91)	(0.68)	(1.44)	(1.85)	(9.33)	(16.54)	(24.48)

by drawing c_a from $U(100, 200)$. We consider 10 sets of costs bounds $(c_a - m x_a, c_a + m y_a)$, where m is referred to as the interval-width multiplier. We change the value of m from 1 to 10, while x_a and y_a are drawn uniformly from $[1, 10]$.

For each pair (p_a, p_c) we generate 30 different networks along with A_0, $\widetilde{A}_0$, the cost vector $(c_a)_{a \in A}$ and x_a, y_a for all $a \in \widehat{A}_0$. We set $n = 40$, $p = 0.8$ (the mean number of arcs of 1248.73), $k = 10$ and $T = 36$. Figure 11 depicts the results for the case $p_c = 1/3$. Additional results for other values of p_a and p_c are available in Appendix C.

The results in Figure 11 illustrate in greater detail the performance of λ and π_M as the quality of cost information worsens. In particular, one observes that, with respect to the total regret, the performance of both policies degrades at a similar rate. This conclusion is not true, however, when considering the total regret MAD as it tends to increase for π_M, while remaining relatively constant for λ. With respect to the time-stability metric, it is observed that although the performance of λ deteriorates, it does so at a much slower rate than π_M, and that its time-stability MAD remains virtually constant as the intervals widens. These results reinforce the conclusions of our previous experiments regarding robustness of the proposed policies. Moreover, they show that λ can be considered somewhat insensitive to the quality of the initial information, while the performance of π_M can significantly degrade if the quality of the initial information is not sufficiently good.

6. Concluding Remarks

In this paper we study sequential interdiction of a directed network when the interdictor has incomplete initial information about the network. By observing the evader's actions (who travels along shortest paths in each time period), the interdictor learns about the structure and costs of
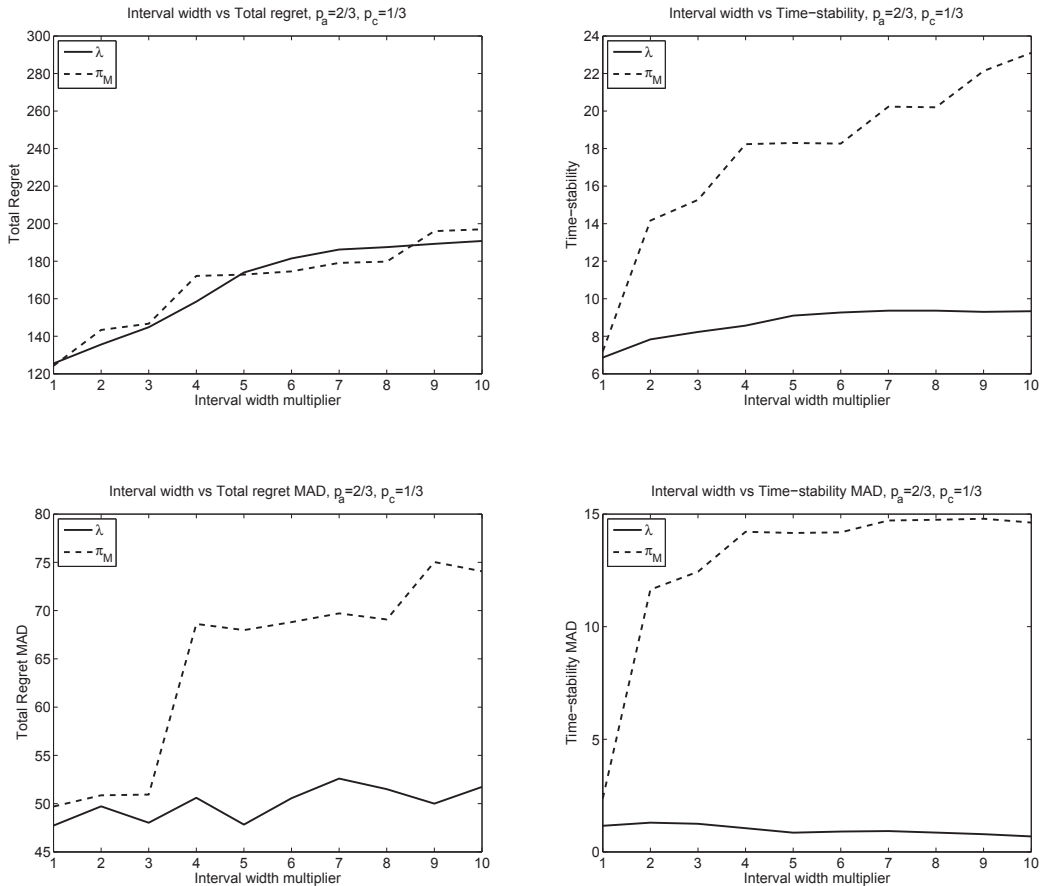
Figure 11 **Behavior of the average time-stability, average total regret, time-stability MAD and total regret MAD as the cost intervals widen for the case of $p_a = 2/3$ and $p_c = 1/3$. Given the interval-width multiplier m, the lower and upper bounds of the arc costs in $\widehat{A}_0$ are $l_a = c_a - m x_a$ and $u_a = c_a + m y_a$, respectively.**

the network and adjusts its actions so as to maximize the cumulative cost incurred by the evader. We formally define the concept of efficient interdiction policies and propose a class of simple interdiction policies that are efficient both with respect to regret and time-stability.

Our theoretical results are supported by numerical experiments which suggest that the proposed policies are robust, in the sense that they provide consistently good results across various levels of the initial information. Aligned with intuition, our interdiction policies yield better results as the quality of the information improves. One important conclusion of our work is that it is crucial for the interdictor to have a pessimistic attitude regarding unobserved arc costs, i.e., it is better to overestimate the real costs of the arcs whose real costs are unknown rather than underestimate them. Otherwise, as we demonstrate both theoretically and computationally, the interdictor is not guaranteed to converge to an optimal k-most vital arc solution, i.e., an interdiction solution with an instantaneous regret of zero.

Finally, we propose a semi-oracle benchmark policy that serves as a lower bound on the performance of any feasible interdiction policy. We formulate such policy as a mixed integer program and describe an algorithmic approach for its computation.

Our work considers *myopic* evaders (recall assumption **A2**), who always traverse along shortest paths of the interdicted network. Consider now a *strategic* evader who does not necessarily travel via shortest paths in each time period, but rather desires to minimize the total costs of moving through the network over time horizon $\mathcal{T}$, i.e., $\sum_{t \in \mathcal{T}} \ell(P_t)$, and suppose that $\widehat{A}_0 = \emptyset$ and that the interdictor uses policies in Γ (similar arguments apply for the more general case). The same type of analysis applies to analyze setting in which the evader does not observe the actions of the interdictor upfront, and needs to learn them in real time.

Using an approach similar to the one used in the proof of Lemma 2, it can be shown that if the evader desires to incur a cost less than $z(G[A_t^\gamma \setminus I_t^\gamma])$ at time period t, then at least one new arc must be revealed to the interdictor in P_t^γ. Thus, we conclude that Lemma 2 holds for any reasonable decision-making process of the evader, i.e., the evader cannot avoid the convergence of policies in Γ.

Therefore, our initial assumption that the evader moves through shortest paths in $G[A \setminus I_t^\gamma]$ turns out to be not too restrictive (at least for the property **C1** to hold for the proposed policies). Moreover, for many instances, this myopic approach might yield the best performance. However, in general there may be a trade-off for the evader between using shortest paths and using alternative paths that, although are not the shortest ones, might improve the performance over the whole time horizon. While the evader's decision-making problem can be casted as sequential mixed-integer bilevel program, as such, it might be intractable in practice. Nevertheless, it presents an interesting avenue for future research.

There are several other research directions that remain open at this point. An immediate one relates to whether our results can be extended or serve as the basis for studying settings, where we relax assumptions **A1** and **A3**. In particular, relaxing the former one results in decision-making problems with alternative feedbacks, where, for example, only a noisy signal of the evader cost is revealed in each time period, or only some components of the evader action are revealed.

Note that if the source and destination nodes are not known initially to the interdictor, then they would be inferred immediately from the evader's actions due to our assumptions of the perfect feedback. Thus, our model also accommodates the setting where the interdictor is not initially aware of such information. Similarly, our model would directly extend to settings, where source and destination nodes are chosen from a given set. In this regard, an interesting extension would be cases, where the source and destination nodes are chosen randomly in each time period according to some distribution. In particular, we believe that our methods are not trivially extendable to

problems, where the source and destination nodes location distribution is initially unknown (observe that the interdictor would not be able to state the objective function upfront under this scenario).

We should also note that our methods can be extended to the setting, where each blocking action does not fully eliminate an arc but rather increases its cost. In such a case, the interdictor might have a budget that must be allocated across arcs. Consequently, in each time period our policies would solve an instance of the *Maximizing the Shortest Path* (MXSP) problem in Israeli and Wood (2002).

Finally, with regard to more general interdiction models in the literature, it is possible to device computational procedures similar to the one proposed in the paper. In the interdiction literature, such setups might include evaders that aim to maximize the flow of illegal materials (or desire to minimize their transportation costs) in the network with capacity constraints on its arcs. Ultimately, the suitability of our approach would rest on the nature of each specific interdiction model (e.g., whether it admits a tractable solution approach in the case of complete information), and, more importantly, the type of the feedback obtained by the interdictor. Generalizing our approach for generic bilevel interdiction models is outside the scope of this paper, and constitutes an interesting line for future research.

Acknowledgments

References

Ahuja, R., Magnanti, T. and Orlin, J. (1993), *Network flows: Theory, algorithms, and applications*, Prentice-Hall, Englewood Cliffs, New Jersey.

Auer, P., Cesa-Bianchi, N. and Fischer, P. (2002), 'Finite-time analysis of the multiarmed bandit problem', *Machine Learning* **47**(2-3), 235–256.

Auer, P., Cesa-Bianchi, N., Freund, Y. and Schapire, R. E. (2003), 'The non-stochastic multi-armed bandit problem', *SIAM Journal on Computing* **32**, 48–77.

Awerbuch, B. and Kleinberg, R. D. (2004), Adaptive routing with end-to-end feedback: distributed learning and geometric approaches, *in* 'Proceedings of the thirty-sixth annual ACM symposium on Theory of computing', STOC '04, ACM, New York, NY, USA, pp. 45–53.

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

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

Beheshti, B., Özaltın, O. Y., Zare, M. H. and Prokopyev, O. A. (2014), 'Exact solution approach for a class of nonlinear bilevel knapsack problems.', *Journal of Global Optimization* . Forthcoming.

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

Cesa-Bianchi, N. and Lugosi, G. (2006), *Prediction, Learning, and Games*, Cambridge University Press, Cambridge, U.K.

Cesa-Bianchi, N. and Lugosi, G. (2012), 'Combinatorial bandits', *Journal of Computer and System Sciences* **78**(5), 1404–1422.

Colson, B., Marcotte, P. and Savard, G. (2007), 'An overview of bilevel optimization', *Annals of Operations Research* **153**(1), 235–256.

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

Corley, H. and Sha, D. (1982), 'Most vital links and nodes in weighted networks', *Operations Research Letters* **1**(4), 157–160.

Cormican, K., Morton, D. and Wood, R. (1998), 'Stochastic network interdiction', *Operations Research* **46**(2), 184–197.

Erdös, P. and Rényi, A. (1959), 'On random graphs, I', *Publicationes Mathematicae (Debrecen)* **6**, 290–297.

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

Ghare, P., Montgomery, D. and Turner, W. (1971), 'Optimal interdiction policy for a flow network', *Naval Research Logistics Quarterly* **18**(1), 37–45.

Gift, P. D. (2010), Planning for an adaptive evader with application to drug interdiction operations, Master's thesis, Monterey, California. Naval Postgraduate School.

Granata, D., Steeger, G. and Rebennack, S. (2013), 'Network interdiction via a critical disruption path: Branch-and-price algorithms', *Computers & Operations Research* **40**(11), 2689–2702.

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

Hemmecke, R., Schultz, R. and Woodruff, D. (2003), Interdicting stochastic networks with binary interdiction effort, *in* D. L. Woodruff, ed., 'Network interdiction and stochastic integer programming', Kluwer Academic Publishers, Boston,MA, pp. 69–84.

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

Janjarassuk, U. and Linderoth, J. (2008), 'Reformulation and sampling to solve a stochastic network interdiction problem', *Networks* **52**(3), 120–132.

Lai, T. L. and Robbins, H. (1985), 'Asymptotically efficient adaptive allocation rules', *Advances in Applied Mathematics* **6**(1), 4–22.

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

Malaviya, A., Rainwater, C. and Sharkey, T. (2012), 'Multi-period network interdiction problems with applications to city-level drug enforcement', *IIE Transactions* **44**(5), 368–380.

Malik, K., Mittal, A. and Gupta, S. (1989), 'The k-most vital arcs in the shortest path problem', *Operations Research Letters* **8**(4), 223–227.

McLay, L., Rothschild, C. and Guikema, S. (2012), 'Robust adversarial risk analysis: a level-k approach', *Decision Analysis* **9**(1), 41–54.

McMasters, A. and Mustin, T. (1970), 'Optimal interdiction of a supply network', *Naval Research Logistics Quarterly* **17**(3), 261–268.

Modaresi, S., Saure, D. and Vielma, J. (2012), Learning in combinatorial optimization: What and how to explore. Working paper.

Morton, D., Pan, F. and Saeger, K. (2007), 'Models for nuclear smuggling interdiction', *IIE Transactions* **39**(1), 3–14.

Pan, F. and Morton, D. (2008), 'Minimizing a stochastic maximum-reliability path', *Networks* **52**(3), 111–119.

Ratliff, H., Sicilia, G. and Lubore, S. (1975), 'Finding the n most vital links in flow networks', *Management Science* **21**(5), 531–539.

Robbins, H. (1952), 'Some aspects of the sequential design of experiments', *Bulletin of the American Mathematical Society* **58**, 527–535.

Shen, S., Smith, J. C. and Goli, R. (2012), 'Exact interdiction models and algorithms for disconnecting networks via node deletions', *Discrete Optimization* **9**(3), 172–188.

Smith, J. C. and Lim, C. (2008), Algorithms for network interdiction and fortification games, *in* 'Pareto optimality, game theory and equilibria', Springer, pp. 609–644.

Veremyev, A., Boginski, V. and Pasiliao, E. L. (2014), 'Exact identification of critical nodes in sparse networks via new compact formulations', *Optimization Letters* **8**(4), 1245–1259.

Veremyev, A., Prokopyev, O. A. and Pasiliao, E. L. (2014), 'An integer programming framework for critical elements detection in graphs', *Journal of Combinatorial Optimization* **28**(1), 233–273.

Walteros, J. L. and Pardalos, P. M. (2012), Selected topics in critical element detection, *in* N. J. Daras, ed., 'Applications of Mathematics and Informatics in Military Science', Vol. 71 of *Springer Optimization and Its Applications*, Springer New York, New York, NY, pp. 9–26.

Washburn, A. and Wood, R. (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, R. (1993), 'Deterministic network interdiction', *Mathematical and Computer Modeling* **17**(2), 1–18.

Xu, J. and Zhuang, J. (2014), 'Modeling costly learning and counter-learning in a defender-attacker game with private defender information', *Annals of Operations Research* . Forthcoming.

Zhuang, J., Bier, V. M. and Alagoz, O. (2010), 'Modeling secrecy and deception in a multiple-period attacker–defender signaling game', *European Journal of Operational Research* **203**(2), 409–418.

Appendix A: Basic properties of k-most vital arcs

Let $G = (N, A, C)$ be a directed network. Recall that a set of k-most vital arcs of G is a subset $L \subseteq A$ that satisfies

$$L \in \underset{\{L \subseteq A :\ |L| \leq k\}}{\arg\max}\ z(G[A \setminus L]).$$

PROPOSITION 4. *Given $G = (N, A, C)$, let $Y \subseteq X \subseteq A$. If L_X and L_Y are sets of k-most vital arcs of $G[X]$ and $G[Y]$, respectively, then $z(G[X \setminus L_X]) \leq z(G[Y \setminus L_Y])$.*

Proof. Let $U = Y \cap L_X$. Then by the definition of L_X and L_Y:

$$z(G[Y \setminus L_Y]) \geq z(G[Y \setminus U]) = z(G[Y \setminus L_X]) \geq z(G[X \setminus L_X]), \tag{14}$$

which concludes the proof. ∎

PROPOSITION 5. *Let $G = (N, A, C)$ and $G' = (N, A, C')$ be networks such that $c_a \leq c'_a$ for all $a \in A$. If L_A and L'_A are sets of k-most vital arcs of G and G', respectively, then*

$$z(G[A \setminus L_A]) \leq z(G'[A \setminus L'_A]). \tag{15}$$

Proof. Since $c_a \leq c'_a$ for all $a \in A$, then $z(G[A \setminus L_A]) \leq z(G'[A \setminus L_A])$. Then by the definition of L'_A, it must hold that $z(G'[A \setminus L_A]) \leq z(G'[A \setminus L'_A])$, which concludes the proof. ∎

Appendix B: Additional proofs

In this appendix we provide proofs for some of the lemmas and for Propositions 2 and 3.

Lemma 1. *Given $G = (N, A, C)$, let L and A' be such that $L \subseteq A' \subseteq A$ and $(G[A'], L)$ is k-complete. Then L is a set of k-most vital arcs of $G[U]$ for all U such that $A' \subseteq U \subseteq A$.*

Proof. Since $G[A']$ is L-spare, we have that

$$z(G[A' \setminus L]) = z(G[A \setminus L]). \tag{16}$$

Let L_U be a set of k-most vital arcs of $G[U]$, where $A' \subseteq U \subseteq A$. Then

$$z(G[A \setminus L]) \leq z(G[U \setminus L]) \leq z(G[U \setminus L_U]) \leq z(G[A' \setminus L]). \tag{17}$$

The first inequality is due to the fact that $(U \setminus L) \subseteq (A \setminus L)$, the second from the definition of L_U, and the last one from Proposition 4 (in Appendix A) and the fact that L is a set of k-most vital arcs of $G[A']$. Therefore, from equations (16) and (17) we have that $z(G[U \setminus L]) = z(G[U \setminus L_U])$, which implies that L is a set of k-most vital arcs of $G[U]$. ∎

Lemma 2. *Let $\gamma \in \Gamma$. Then for any $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$:*
  1. $\tau^\gamma(G, \mathcal{C}_0) \leq x^\gamma(G, \mathcal{C}_0);$

2. *if $T > |A|$ then $\tau^\gamma(G, \mathcal{C}_0) \le |A|$.*

Proof. To simplify the notation, let $x \equiv x^\gamma$. Note that by the definition of x, $G[A_x^\gamma]$ is a I_x^γ-spare network. As I_x^γ is also a set of k-most vital arcs of $G[A_x^\gamma]$, it follows that $(G[A_x^\gamma], I_x^\gamma)$ is k-complete and $z(G[A_x^\gamma \setminus I_x^\gamma]) = z^*(G)$ by Lemma 1. Moreover, by the definition of Γ, we have that $I_t^\gamma = I_x^\gamma$ for $t > x$. Hence, $z(G[A_t^\gamma \setminus I_t^\gamma]) = z^*(G)$ for all $t > x$ and the first claim of the proposition follows.

To prove the second claim, we consider two possible cases. Specifically, one has that $G[A_t^\gamma]$ is either I_t^γ-spare for some $t \le |A|$ or not. In the former case, the arguments above imply that $\tau^\gamma(G, \mathcal{C}_0) \le |A|$ and the result follows. In the latter case, because $\ell(P_t^\gamma) = z(G[A \setminus I_t^\gamma])$ for all t, by equation (17) we have that

$$z(G[A_t^\gamma \setminus I_t^\gamma]) > z(G[A \setminus I_t^\gamma]) \quad t \le |A|. \tag{18}$$

The above implies that $P_t^\gamma \not\subseteq A_t^\gamma$ for all $t \le |A|$. Indeed, suppose that this is not the case. Since $P_t^\gamma \cap I_t^\gamma = \emptyset$ and I_t^γ is a set of k-most vital arcs of $G[A_t^\gamma]$, then $P_t^\gamma \subseteq A_t^\gamma$ implies that $\ell(P_t^\gamma) \ge z(G[A_t^\gamma \setminus I_t^\gamma])$. Thus, from equation (18) we have that

$$\ell(P_t^\gamma) > z(G[A \setminus I_t^\gamma]), \tag{19}$$

which contradicts the fact that P_t^γ is a shortest path in $G[A \setminus I_t^\gamma]$. We conclude that $A_t^\gamma \subset A_{t+1}^\gamma$, which implies the required result. ∎

Lemma 4. *Suppose that $t \in \mathcal{T}$ is such that $z(G_t^\lambda[A_t^\lambda \setminus I_t^\lambda]) = z(G[A \setminus I_t^\lambda])$, then $(G[A_t^\lambda], I_t^\lambda)$ is k-complete (with respect to G). Moreover, I_t^λ is a set of k-most vital arcs of $G[U]$ for all U such that $A_t^\lambda \subseteq U \subseteq A$.*

Proof. Let L_A and L be sets of k-most vital arcs of G and $G[A_t^\lambda]$, respectively. We have that

$$z(G[A \setminus I_t^\lambda]) \le z(G[A \setminus L_A]) \overset{(a)}{\le} z(G[A_t^\lambda \setminus L]) \overset{(b)}{\le} z(G_t^\lambda[A_t^\lambda \setminus I_t^\lambda]), \tag{20}$$

where (a) and (b) follow from Propositions 4 and 5, respectively (see Appendix A). This, and the condition that $z(G[A \setminus I_t^\lambda]) = z(G_t^\lambda[A_t^\lambda \setminus I_t^\lambda])$ implies that

$$z(G[A \setminus I_t^\lambda]) = z(G[A \setminus L_A]) = z(G[A_t^\lambda \setminus L]) = z(G_t^\lambda[A_t^\lambda \setminus I_t^\lambda]). \tag{21}$$

On the other hand, because $A_t^\lambda \subseteq A$ the following inequalities hold

$$z(G[A \setminus I_t^\lambda]) \le z(G[A_t^\lambda \setminus I_t^\lambda]) \le z(G[A_t^\lambda \setminus L]), \tag{22}$$

thus, implying that $z(G[A \setminus I_t^\lambda]) = z(G[A_t^\lambda \setminus I_t^\lambda]) = z(G[A_t^\lambda \setminus L])$. Hence, $G[A_t^\lambda]$ is I_t^λ-spare. Moreover, because L is a set of k-most vital arcs of $G[A_t^\lambda]$, the fact that $z(G[A_t^\lambda \setminus I_t^\lambda]) = z(G[A_t^\lambda \setminus L])$ implies that I_t^λ is also a set of k-most vital arcs of $G[A_t^\lambda]$, and the first statement of the proposition follows. The second statement follows directly from Lemma 1. ∎

Lemma 5. *Let $\lambda \in \Lambda$. Then for any $\mathcal{C}_0$ and $G \in \mathbb{G}(\mathcal{C}_0)$:*

1. $\tau^\lambda(G, \mathcal{C}_0) \le \hat{x}^\lambda(G, \mathcal{C}_0)$;

2. *if $T \geq |A|$, then $\tau^\lambda(G, \mathcal{C}_0) \leq |A|$.*

Proof. To simplify the notation let $x = \widehat{x}^\lambda$. Because $z(G_x^\lambda[A_x^\lambda \setminus I_x^\lambda]) = z(G[A \setminus I_x^\lambda])$, Lemma 4 implies that I_x^λ is a set of k-most vital arcs of G. Thus, $z(G[A \setminus I_x^\lambda]) = z^*(G)$. Moreover, from the definition of Λ, we have that $I_t^\lambda = I_x^\lambda$ for all $t > x$, and the first claim of the proposition follows. The proof of the second statement follows from the arguments in proof of the second statement in Lemma 2, with (20) playing the role of (17).

∎

Lemma 6. Λ *is a homogeneous set both with respect to cumulative regret and with respect to time-stability.*

Proof. We slightly modify the proof of Lemma 3. Consider again $\mathcal{C}_0$, G and G' as given by Figure 5. However, assume that the cost of arc $(1, n)$ in $\mathcal{C}_0$ is not known exactly, but is known to be within range $[1, M]$, where $M \geq 2$, i.e., $(1, n) \notin \widetilde{A}_0$, but $(1, n) \in \widehat{A}_0$. At the first time period the evader travels along the arc $(1, n)$ and its cost becomes known to the interdictor. The result then follows by mimicking the proof of Lemma 3. ∎

Proposition 2. *There exists $\mathcal{C}_0$, $G \in \mathbb{G}(\mathcal{C}_0)$ and $\zeta > 0$ such that if $T \geq |A|$, then $\tau^\lambda(G, \mathcal{C}_0) \geq \zeta|A|$. Moreover, the value of $R_T^\lambda(G, \mathcal{C}_0)$ can be made arbitrarily large.*

Proof. Consider the same network as in Proposition 1, but with $\widetilde{A}_0^\lambda = \emptyset$, and $\widehat{A}_0^\lambda = A$. Let $c_{ij} = 1$ for all $(i, j) \in A$, except for $(1, 2)$ with $c_{12} = M \geq 2$, and assume that $\ell_{ij} = 1$ and $u_{ij} = M + 1$ for all $(i, j) \in A$. Proceeding in a similar fashion as in Proposition 1, arc $(1, 2)$ is revealed after $\zeta|A|$ time periods, where $\zeta \leq k/(2k + 2)$. ∎

Proposition 3. *Algorithm 2 correctly solves $LB(G, \mathcal{C}_0, T)$.*

Proof. Note that only the case when $T_0 < T$ is relevant. Upon convergence one has that either $T' = T$ or $T' < T$. In the first case, $LB(G, \mathcal{C}_0, T)$ is solved using formulation (10) and the result follows. For the second case suppose that $\{(r^t, p^t, y^t) : t \in \mathcal{T}\}$ is not optimal, and consider a solution $\{(\bar{r}^t, \bar{p}^t, \bar{y}^t) : t \in \mathcal{T}\}$ feasible for $LB(G, \mathcal{C}_0, T)$ such that

$$\sum_{t=0}^{T} \bar{y}_1^t - \bar{y}_n^t > \sum_{t=0}^{T} y_1^t - y_n^t. \tag{23}$$

Because $T' < T$ it is necessarily the case that $z^* = y_s^t - y_t^t$ for all $t \geq T'$. Therefore, (23) implies that

$$\sum_{t=0}^{T'} \bar{y}_1^t - \bar{y}_n^t > \sum_{t=0}^{T'} y_1^t - y_n^t.$$

However, $\{(\bar{r}^t, \bar{p}^t, \bar{y}^t) : t \leq T'\}$ is feasible for $LB(G, \mathcal{C}_0, T')$. Thus, the equation above contradicts the fact that $\{(r^t, p^t, y^t) : t \leq T'\}$ is an optimal solution of $LB(G, \mathcal{C}_0, T')$. This proves the result. ∎

Appendix C: Additional graphs

We provide additional results for the computational experiments described in Section 5. Figure 12 corresponds to the discussion in Section 5.4 for the left-skewed distributed costs. Figures 13 and 14 are the complementary figures for the second experiment in Section 5.5, setting $p_c = 0$ and $p_c = 2/3$ respectively.
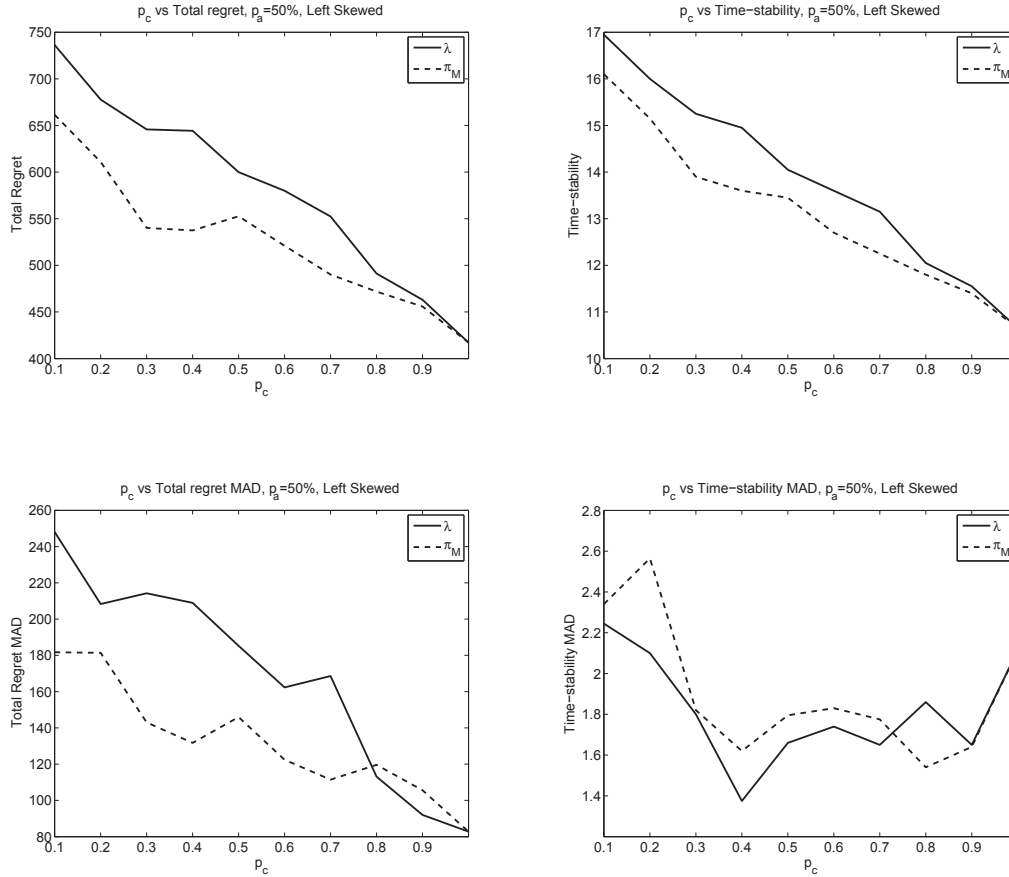

Figure 12   **Behavior of the average time-stability, average total regret, time-stability MAD and total regret MAD as p_c increases. The cost distribution is left-skewed and $p_a = 1/2$.**
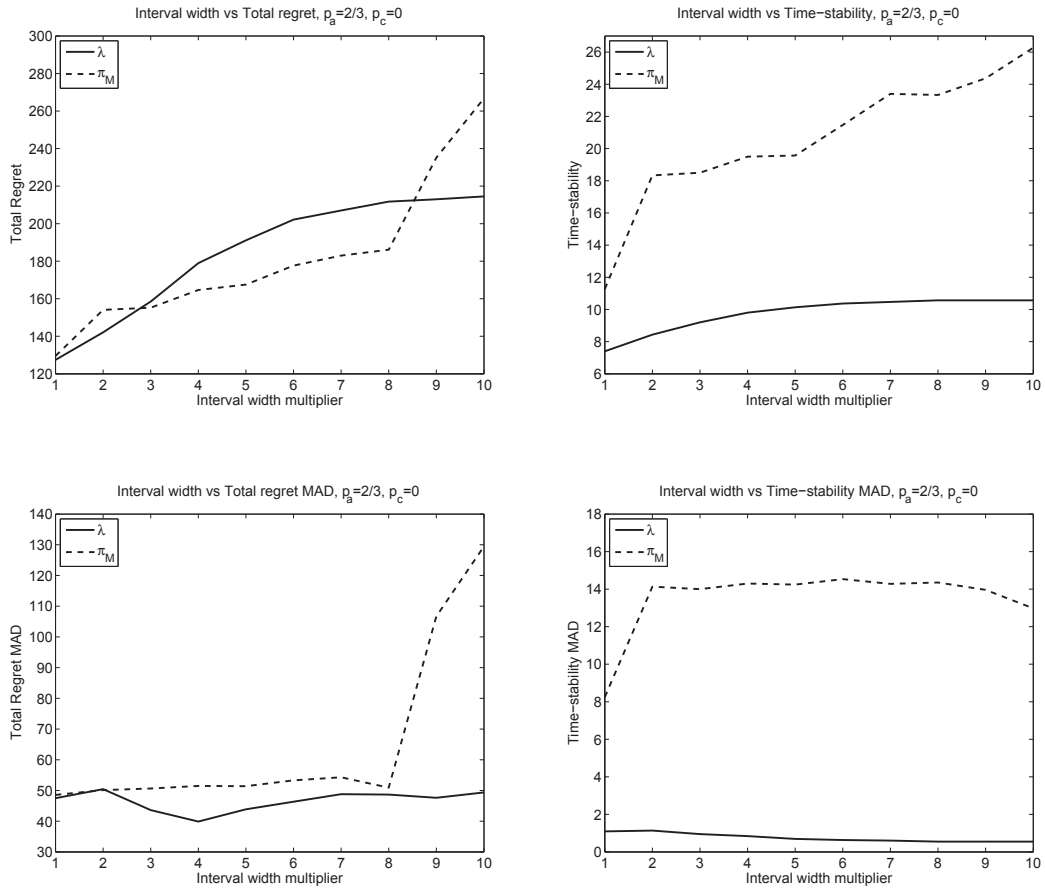
Figure 13 **Behavior of the average time-stability, average total regret, time-stability MAD and total regret MAD as the cost intervals widen for the case of** $p_a = 2/3$ **and** $p_c = 0$. **Given the interval-width multiplier** m, **the lower and upper bounds of the arc cost in** $\widehat{A}_0$ **are** $l_a = c_a - m x_a$ **and** $u_a = c_a + m y_a$, **respectively.**
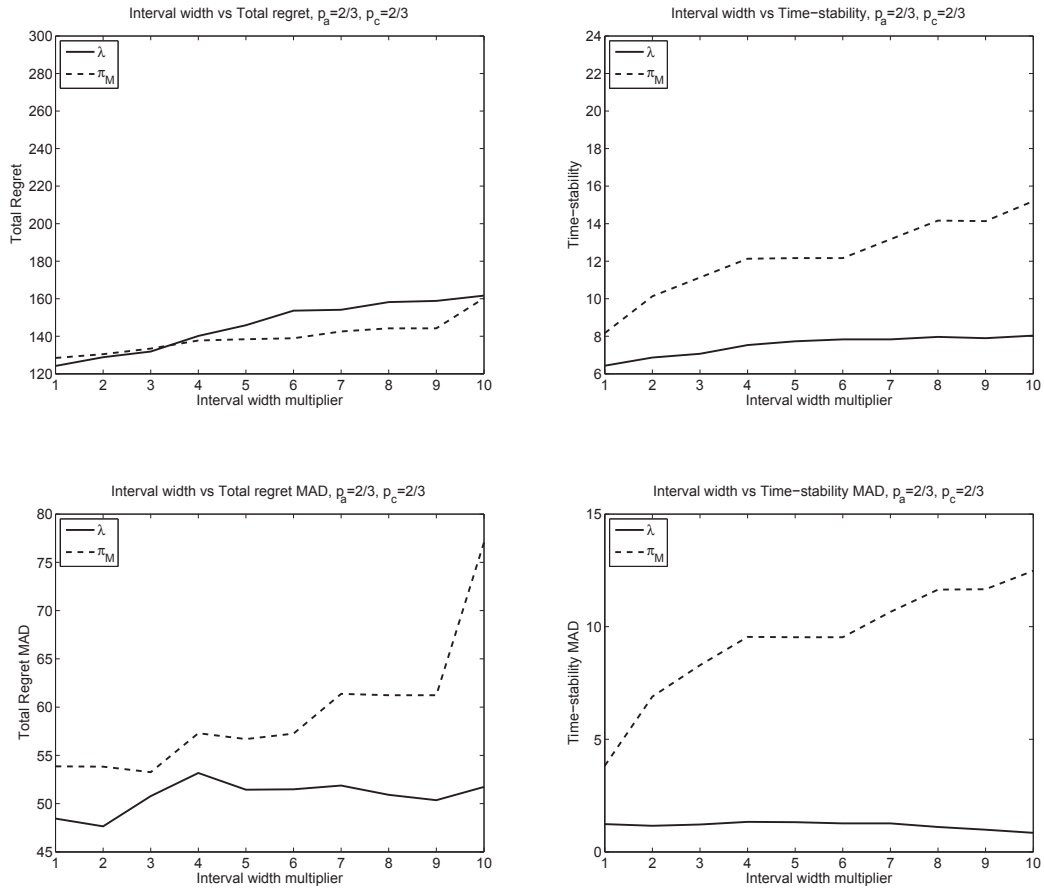
Figure 14   **Behavior of the average time-stability, average total regret, time-stability MAD and total regret MAD as the cost intervals widen for the case of** $p_a = 2/3$ **and** $p_c = 2/3$**. Given the interval-width multiplier** m**, the lower and upper bounds of the arc costs in** $\widehat{A}_0$ **are** $l_a = c_a - m x_a$ **and** $u_a = c_a + m y_a$**, respectively.**