Deception Tactics for Network Interdiction: A Multiobjective Approach

Javier Salmerón
Department of Operations Research, Naval Postgraduate School, Monterey, California 93943

This article develops defender-attacker network interdiction models with deception. Here, deception refers to a preemptive and intelligent use of concealed interdiction assets and decoys by the defender, in addition to transparent assets commonly employed in modeling defender-attacker problems. These models can help security planners to locate a limited number of checkpoints and sensors of various types to, for example, detect the smuggling of illegal products. The problem is complex, in part, because the objective functions of the defender and the attacker are different, and because the latter (which represents the attacker's behavior) is difficult to predict by the defender. First, we use duality theory and a generalized network flow model to devise an equivalent mixed-integer programming formulation, and develop its Benders decomposition. We extend this formulation with a multiobjective approach to account for several behaviors simultaneously. The computational effort to solve these models is considerable, as exemplified by our testing on a variety of cases for a medium-sized, notional network. **Published 2011 Wiley Periodicals, Inc. NETWORKS, Vol. 60(1), 45–58 2012**

Keywords: network interdiction; asymmetric information; bi-level optimization; multiobjective optimization; generalized networks; Benders decomposition

1. INTRODUCTION

Network interdiction studies how decisions made by two intelligent adversaries, called "defender" and "attacker," affect a network's functionality. One of these players seeks, for example, to maximize flow through the network, or to minimize cost to supply the demand, while the other player tries to disrupt the network by interdicting selected components [18]. The associated problems are commonly referred to as defender-attacker (DA) and attacker-defender (AD) models (see, e.g., [5] and references therein). In many of these problems, players make their decisions in a specified order, as in bi-level programming and Stackelberg games [16].

This article deals with a special case of DA analysis. For exposition's sake, we refer to suicide attacks as the event against which the defender is protecting. However, the techniques we present have applications in other areas such as weapon smuggling, drug traffic, facility security (such as airports), and the like. The attacker comprises, for example, a set of vehicle-born suicide attackers, each of whom may originate at a different location and may seek one or several possible targets. The defender wishes to allocate limited interdiction assets to minimize damage inflicted by the attacker. In this context, interdiction refers to both the detection and physical neutralization of the attacker, thus saving the lives the attack would have otherwise claimed.

Among the assets available to the defender, we distinguish three types: The first type consists of interdiction assets which, once deployed, are visible (or transparent) to the attacker. We also assume the probability of successful interdiction by transparent assets is equally known by the attacker and defender. One example of such an asset would be a check point which randomly inspects one in every k vehicles. The second type consists of trap-like assets, whose location is only known by the defender. This may include, for example, concealed sensors and other detection equipment, or unmanned aircraft. The third type consists of decoy assets, which are visible by the attacker, but perceived as more effective than they actually are, for example, a surveillance camera that is not being monitored. We also assume a nominal (typically low) probability of interdiction always exists on noninterdicted arcs (e.g., by regular law enforcement patrols without detection equipment).

Our models intend to provide insights into similar questions to those in classical (fully transparent) DA models, such as "what is the optimal allocation of a given number of interdiction resources?," and "what are the associated routes for attackers, and the probabilities of successful interdiction for the defender?" They also analyze the tradeoff between defender's resources and the increased probability of detection to determine how many assets of each type must be allocated to ensure a certain probability of success. Additionally, we reckon the potential benefit provided by secrecy and deception, whose effectiveness may depend on the attacker's behavior.

Received January 2010; accepted July 2001
Correspondence to: J. Salmerón; e-mail: jsalmero@nps.edu
DOI 10.1002/net.20458
Published online 19 September 2011 in Wiley Online Library (wileyonlinelibrary.com).
© 2011 Wiley Periodicals, Inc. *This article is a U.S. Government work and is in the public domain in the USA.

Network interdiction has been extended to consider simultaneous decisions and uncertainty. Simultaneous-game versions of the shortest-path interdiction model (see, e.g., [17]) seek mixed strategies for both the defender's interdiction locations and the evader's route selection. In discussing game-theoretic models for defending critical infrastructure against intelligent attackers, Bier et al. [3] believe that "an approach that is more effective and practical than pure game-theoretic analysis is needed." They point out the caveats in assuming subjective beliefs, game structures and payoff matrices, adding that "…for resource allocation, it is essential to take into account an adversary's possible adaptive behaviors, but without necessarily descending into the mathematical quagmire of full game-theoretic modeling." In this sense, the authors favor a simpler approach under the paradigm of AD or DA models. The deterministic maximum-flow interdiction problem has been extended in [6] to account for uncertainty in the attack success and in the initial arc capacities. Another bi-level, two-stage stochastic model [14] determines the optimal allocation of sensors to detect nuclear smuggled materials, where uncertainty is related to the origin and destination of the evader.

The special structure in our models makes our methodology different from most previous work in the literature. Notable exceptions include a similar model with asymmetric information for the case of a single evader (attacker) and a single behavior (shortest distance) [2], and a stochastic model with differing perceptions for defender and attacker [12], which is especially effective in the case of bipartite networks (after adding specialized, valid inequalities). Our work incorporates several attackers, different types of interdiction assets and multiple behaviors simultaneously.

In the remainder of the article, Section 2 introduces our core model formulation and develops an equivalent mixed-integer program (MIP), and its Benders decomposition (BD). We also discuss other related formulations, and our assumptions to represent different attackers' behaviors. Section 3 describes our notional test networks and computational results. These suggest a natural extension of the original formulation as a multiobjective model, which is presented in Section 4. Section 5 presents our conclusions.

2. MODEL DESCRIPTION

In this section, we describe the mathematical formulation of the "nontransparent" DA (NTDA) model. This model is initially stated as a nonlinear model NTDA^{NL} and then converted into an equivalent MIP called NTDA^{MIP}, which can be solved directly or via BD. In addition, we present a heuristic approximation NTDA^{H} which uses a simplified MIP but only guarantees an optimal solution to the NTDA problem in the case of a unique attacker. We also introduce RNTDA^{NL}, which is a "robust" NTDA model that replaces total expected value by "worst-case value" among attackers; then, we develop its MIP version $\text{RNTDA}^{\text{MIP}}$. Finally, we describe the attacker behaviors investigated in this research.

2.1. Notation

The notation used in our models is as follows:

Sets and indices

I, set of nodes in the network, $i, j \in I$;
$A \subset I \times I$, set of directed arcs in the network, $(i,j) \in A$;
N, set of attackers, $n \in N$;
$s_n \in I$, source node for attacker n;
$T_n \subset I$, subset of possible target nodes for attacker n. We assume $s_n \notin T_n$.

Parameters

v_n, value of attacker n (if he succeeds);
q_{nij}, nominal probability of evasion (nondetection) for attacker n while traversing arc (i,j);
$\tilde{q}_{nij}$, probability of evasion for attacker n while traversing arc (i,j) interdicted with a transparent asset;
$\bar{q}_{nij}$, actual probability of evasion for attacker n while traversing arc (i,j) interdicted with a trap asset. (Attackers perceive nominal probability);
$\bar{\bar{q}}_{nij}$, probability of evasion perceived by attacker n while traversing arc (i,j) interdicted with a decoy asset. (Actual probability is nominal);
d_{nij}, measure of "cost" (such as distance or travel time) for arc (i,j) used by attacker n who ignores the defender's strategy. (See "indifferent" behavior in Section 2.7);
$\tilde{R}$, $\bar{R}$, $\bar{\bar{R}}$, number of transparent, trap and decoy assets, respectively.

Decision variables

$\tilde{y}_{ij}$, $\bar{y}_{ij}$, $\bar{\bar{y}}_{ij}$, equals 1 if arc (i,j) is interdicted with a transparent asset, with a trap or with a decoy, respectively; 0 otherwise;
x_{nij}, equals 1 if attacker n traverses arc (i,j); 0 otherwise;
$x^P_{nij}, \tilde{x}^P_{nij}, \bar{x}^P_{nij}, \bar{\bar{x}}^P_{nij}$, generalized flow variables for attacker n, representing the probability of evasion up to (but not including) arc (i,j), when arc (i,j) has either no interdiction asset, a transparent asset, a trap, or a decoy, respectively;
Z_n, overall probability of evasion for attacker n.

Augmented network, derived data, and auxiliary variables

t, artificial "super-sink" node;
I^*, set of nodes augmented with the super-sink node: $I^* = I \cup \{t\}$;
A^*_n, set of directed arcs augmented with arcs from targets for attacker n to the super-sink: $A^*_n = A \cup \{(i,t) | i \in T_n\}$;
f_{ni}, equals 1 if $i = s_n$, -1 if $i = t$, and 0 otherwise, for attacker n and node $i \in I^*$;
$p_{nij}, \tilde{p}_{nij}, \bar{p}_{nij}, \bar{\bar{p}}_{nij}$, derived data: $p_{nij} = 1 - q_{nij}, \tilde{p}_{nij} = 1 - \tilde{q}_{nij}$, $\bar{p}_{nij} = 1 - \bar{q}_{nij}, \bar{\bar{p}}_{nij} = 1 - \bar{\bar{q}}_{nij}$;

$\tilde{r}_{nij}, \bar{r}_{nij}, \bar{\bar{r}}_{nij}$, derived data: $\tilde{r}_{nij} = \frac{\tilde{q}_{nij}}{q_{nij}}$, $\bar{r}_{nij} = \frac{\bar{q}_{nij}}{q_{nij}}$, $\bar{\bar{r}}_{nij} = \frac{\bar{\bar{q}}_{nij}}{q_{nij}}$;

$c_{nij}, \tilde{c}_{nij}, \bar{c}_{nij}, \bar{\bar{c}}_{nij}$, derived data: $c_{nij} = \log q_{nij}$, $\tilde{c}_{nij} = \log \tilde{r}_{nij}$, $\bar{c}_{nij} = \log \bar{r}_{nij}$, $\bar{\bar{c}}_{nij} = \log \bar{\bar{r}}_{nij}$;

$\tilde{x}_{nij}, \bar{x}_{nij}, \bar{\bar{x}}_{nij}$, auxiliary variables: $\tilde{x}_{nij} = x_{nij}\tilde{y}_{ij}$, $\bar{x}_{nij} = x_{nij}\bar{y}_{ij}$, $\bar{\bar{x}}_{nij} = x_{nij}\bar{\bar{y}}_{ij}$.

Remarks.

1. *We use the following vector notation for our decision variables:* $\boldsymbol{Z} = (Z_n)_{n\in N}$; $\tilde{\boldsymbol{y}} = (\tilde{y}_{ij})_{(i,j)\in A}$; $\bar{\boldsymbol{y}} = (\bar{y}_{ij})_{(i,j)\in A}$; $\bar{\bar{\boldsymbol{y}}} = (\bar{\bar{y}}_{ij})_{(i,j)\in A}$; $\boldsymbol{y} = (\tilde{\boldsymbol{y}}, \bar{\boldsymbol{y}}, \bar{\bar{\boldsymbol{y}}})$; $\boldsymbol{x} = (\boldsymbol{x}_n)_{n\in N}$ *where* $\boldsymbol{x}_n = (x_{nij})_{(i,j)\in A_n^*}$; *and so on for* $\boldsymbol{x}^P, \tilde{\boldsymbol{x}}^P, \bar{\boldsymbol{x}}^P, \bar{\bar{\boldsymbol{x}}}^P, \tilde{\boldsymbol{x}}, \bar{\boldsymbol{x}}$, *and* $\bar{\bar{\boldsymbol{x}}}$.
2. *If an attacker may choose among several origins as his starting point, we can create a super-source node connected to those origins by noninterdictable arcs.*
3. v_n *reflects attacker n's capability, independent of his final target. For example, we expect an attack by truck n to be more destructive than another attack* n' *committed with a car. We discuss the extension that accommodates different target values for the same attacker* (v_{ni}) *after introducing the* NTDA$^{\text{MIP}}$ *model. Alternatively,* v_n *could be used to model a one-attacker case similar to* [14] *with unknown origin and/or destination. In this case, N would be the set of potential origin-destination scenarios and* v_n *would be the probability that scenario n arises.*
4. *For simplicity in our model representation, we assume* $0 < q_{nij}, \tilde{q}_{nij}, \bar{q}_{nij}, \bar{\bar{q}}_{nij} < 1$.

2.2. NTDA$^{\text{NL}}$ *Model Formulation*

In the NTDA$^{\text{NL}}$ model, the defender tries to minimize the total expected value not interdicted:

$$\text{NTDA}^{\text{NL}}: \quad \min_{y,Z} \sum_{n\in N} v_n Z_n, \tag{1}$$

$$\text{s.t.} \quad Z_n = \prod_{(i,j)\in A} q_{nij}^{x_{nij}} \tilde{r}_{nij}^{x_{nij}\tilde{y}_{ij}} \bar{r}_{nij}^{x_{nij}\bar{y}_{ij}}, \quad \forall n \in N, \tag{2}$$

$$\boldsymbol{y} \in Y \equiv \begin{cases} \sum_{(i,j)\in A} \tilde{y}_{ij} \le \tilde{R}, \\ \sum_{(i,j)\in A} \bar{y}_{ij} \le \bar{R}, \\ \sum_{(i,j)\in A} \bar{\bar{y}}_{ij} \le \bar{\bar{R}}, \\ \tilde{y}_{ij} + \bar{y}_{ij} + \bar{\bar{y}}_{ij} \le 1, \ \forall (i,j) \in A, \\ \tilde{y}_{ij}, \bar{y}_{ij}, \bar{\bar{y}}_{ij} \in \{0,1\}, \ \forall (i,j) \in A, \end{cases} \tag{3}$$

where $\boldsymbol{x}_n$ solves the flow problem F_n for each attacker $n \in N$:

$$\text{F}_n: \quad \max_{x_n} \prod_{(i,j)\in A} q_{nij}^{x_{nij}} \tilde{r}_{nij}^{x_{nij}\tilde{y}_{ij}} \bar{\bar{r}}_{nij}^{x_{nij}\bar{\bar{y}}_{ij}}, \tag{4}$$

$$\text{s.t.} \quad x_n \in \aleph_n^* \equiv \begin{cases} \sum_{j|(i,j)\in A_n^*} x_{nij} - \sum_{j|(j,i)\in A_n^*} x_{nji} = f_{ni}, \ \forall i \in I^* \quad [u_{ni}], \\ x_{nij} \in \{0,1\}, \quad \forall (i,j) \in A_n^*. \end{cases} \tag{5}$$

Remark. *Below, we will justify that* $x_{nij} \in \{0,1\}$ *can be replaced by* $x_{nij} \ge 0$*; this, in turn, will justify the existence of dual variables* $\boldsymbol{u} = (u_{ni})_{n\in N,\, i\in I^*}$.

We note that the defender's objective function defined through (1) and (2) incorporates the actual interdiction probabilities. The nth attacker tries to maximize his probability of evasion, but his perceptions are not always truthful, as reflected in (4).

Constraints $\boldsymbol{y} \in Y$ represent the decision space for the defender, which (for simplicity) restricts the possible interdictions by a cardinality constraint on each type of asset. We assume that the defender does not place more than one interdiction asset per arc, and that an interdiction asset is not consumed by an attacker (i.e., it can be used on multiple attackers). The decision space for all the attackers, $\boldsymbol{x} \in \aleph^*$, comprises the individual flow balance constraints for each attacker, $\boldsymbol{x}_n \in \aleph_n^*$.

2.3. *Reformulation: The* NTDA$^{\text{MIP}}$ *Model*

Some manipulations to the above formulation are necessary to convert the min-max, bi-level NTDA$^{\text{NL}}$ problem with two different nonlinear objectives into a single-objective MIP. This requires reformulating the problem in three steps which deal with: (a) the attackers' nonlinear objective; (b) the defender's nonlinear objective; and (c) the fact that both players have different objectives. This is done as follows:

a. We first linearize the attacker's objective (4) by maximizing the logarithm of that objective (see, e.g., [9]). Thus, for each fixed $\boldsymbol{y}$, model F_n can be solved by finding the optimal solution to the following linear problem:

$$\max_{x_n \in \aleph_n^*} \sum_{(i,j)\in A} (c_{nij} + \tilde{c}_{nij}\tilde{y}_{ij} + \bar{\bar{c}}_{nij}\bar{\bar{y}}_{ij})\, x_{nij}. \tag{6}$$

This linearization renders another important benefit: By unimodularity, it is possible to replace $x_{nij} \in \{0,1\}$ by $x_{nij} \ge 0$, $\forall (i,j) \in A_n^*$. The replacement will be assumed as part of the definition of $\boldsymbol{x}_n \in \aleph_n^*$ in the remainder of the article.

b. Unfortunately, the same strategy cannot be applied to the defender's objective (1) and (2) because the order of the logarithm and summation operators cannot be not switched, unless $|N| = 1$. Instead, we use the concept of generalized networks (see [1], pp. 566–568) to construct a flow that captures the probability of interdiction along the path chosen by the attacker, similar to the approach in [12].

Specifically, we add four arcs for each $(i,j) \in A_n^*$, with generalized flows denoted $x_{nij}^P, \tilde{x}_{nij}^P, \bar{x}_{nij}^P, \bar{\bar{x}}_{nij}^P$, respectively. If the nth attacker traverses arc (i,j) in the original network

(i.e., when $x_{nij} = 1$), we let one of the above flows represent the probability that the attacker is not detected up to that arc. The chosen arc with positive flow depends on whether (i,j) has not been interdicted or, if it has, on whether the defender has used a transparent asset, a trap, or a decoy, respectively. The below linear constraints (7)–(13) replace (2) by a generalized flow structure that calculates Z_n:

$$\sum_{j|(s_n,j)\in A_n^*} x^P_{n,s_n j} = 1, \quad \forall n \in N, \tag{7}$$

$$\sum_{j|(i,j)\in A_n^*} \left(x^P_{nij} + \tilde{x}^P_{nij} + \bar{x}^P_{nij} + \bar{\bar{x}}^P_{nij}\right) = \sum_{j|(j,i)\in A_n^*} \left(q_{nji}x^P_{nji} + \tilde{q}_{nji}\tilde{x}^P_{nji} + \bar{q}_{nji}\bar{x}^P_{nji} + q_{nji}\bar{\bar{x}}^P_{nji}\right),$$
$$\forall n \in N, \forall i \neq s_n, t, \tag{8}$$

$$\sum_{j|j\in T_n} \left(x^P_{njt} + \tilde{x}^P_{njt} + \bar{x}^P_{njt} + \bar{\bar{x}}^P_{njt}\right) = Z_n, \quad \forall n \in N, \tag{9}$$

$$\tilde{x}^P_{nij} \le \tilde{y}_{ij}, \quad \forall n \in N,\ \forall (i,j) \in A_n^* \quad \left[\tilde{\pi}^P_{nij}\right], \tag{10}$$

$$\bar{x}^P_{nij} \le \bar{y}_{ij}, \quad \forall n \in N, \forall (i,j) \in A_n^* \quad \left[\bar{\pi}^P_{nij}\right], \tag{11}$$

$$\bar{\bar{x}}^P_{nij} \le \bar{\bar{y}}_{ij}, \quad \forall n \in N, \forall (i,j) \in A_n^* \quad \left[\bar{\bar{\pi}}^P_{nij}\right], \tag{12}$$

$$x^P_{nij} + \tilde{x}^P_{nij} + \bar{x}^P_{nij} + \bar{\bar{x}}^P_{nij} \le x_{nij}, \quad \forall n \in N, \forall (i,j) \in A_n^*. \tag{13}$$

Remarks. *Constraints to establish that $x^P_{nij} \le 1 - \tilde{y}_{ij}$, $x^P_{nij} \le 1 - \bar{y}_{ij}$, and $x^P_{nij} \le 1 - \bar{\bar{y}}_{ij}$ are not needed because the defender is minimizing flow and $q_{nij} \ge \tilde{q}_{nij}, \bar{q}_{nij}, \bar{\bar{q}}_{nij}$. The π notation refers to dual variables which will be used to develop a BD approach.*

c. As the objectives of defender and attacker are different, the well-known technique of dualization of the "inner" problem to obtain a min–min formulation (see e.g., [5]) is not applicable. However, for every defender choice $\boldsymbol{y}$, we may use strong duality theory with the attacker's problem (5) and (6) to characterize an optimal $\boldsymbol{x}$, similar to the technique used in [13]. Specifically, we replace (6) by setting the objective value equal to that of its dual counterpart, and by adding all necessary dual constraints:

$$\sum_{(i,j)\in A} \left(c_{nij} + \tilde{c}_{nij}\tilde{y}_{ij} + \bar{\bar{c}}_{nij}\bar{\bar{y}}_{ij}\right) x_{nij} = u_{n,s_n} - u_{nt}, \quad \forall n \in N, \tag{14}$$

$$u_{ni} - u_{nj} \ge c_{nij} + \tilde{c}_{nij}\tilde{y}_{ij} + \bar{\bar{c}}_{nij}\bar{\bar{y}}_{ij}, \quad \forall n \in N, \forall (i,j) \in A \quad [\pi_{nij}], \tag{15}$$

$$u_{ni} - u_{nt} \ge 0, \quad \forall n \in N, \forall i \in T_n. \tag{16}$$

Constraints (14) involve products of binary and continuous variables. A linearization of these terms can be achieved, for example, by replacing every $\tilde{y}x$, $\bar{y}x$, and $\bar{\bar{y}}x$ occurrence by $\tilde{x}$, $\bar{x}$, and $\bar{\bar{x}}$, respectively, and adding appropriate linear constraints:

$$0 \le \tilde{x}_{nij} \le \tilde{y}_{ij} \quad \left[\tilde{\pi}^1_{nij}\right];\ \tilde{x}_{nij} \ge \tilde{y}_{ij} + x_{nij} - 1\left[\tilde{\pi}^2_{nij}\right];\ \tilde{x}_{nij} \le x_{nij};$$
$$0 \le \bar{x}_{nij} \le \bar{y}_{ij};\quad \bar{x}_{nij} \ge \bar{y}_{ij} + x_{nij} - 1;\ \bar{x}_{nij} \le x_{nij}; \tag{17}$$
$$0 \le \bar{\bar{x}}_{nij} \le \bar{\bar{y}}_{ij} \quad \left[\bar{\bar{\pi}}^1_{nij}\right];\quad \bar{\bar{x}}_{nij} \ge \bar{\bar{y}}_{ij} + x_{nij} - 1\left[\bar{\bar{\pi}}^2_{nij}\right];\ \bar{\bar{x}}_{nij} \le x_{nij}, \quad \forall n \in N, \forall (i,j) \in A,$$

so (14) becomes:

$$\sum_{(i,j)\in A} \left(c_{nij}x_{nij} + \tilde{c}_{nij}\tilde{x}_{nij} + \bar{\bar{c}}_{nij}\bar{\bar{x}}_{nij}\right) = u_{n,s_n} - u_{nt}, \quad \forall n \in N. \tag{18}$$

Remark. *$\bar{x} = \bar{y}x$ is not used in the* NTDA$^{\text{MIP}}$ *model, but will be used in the* NTDA$^{\text{H}}$ *and* RNTDA$^{\text{MIP}}$ *models. For conciseness,* (17) *includes its linearization too.*

The NTDA$^{\text{MIP}}$ *reformulation is:*

$$\text{NTDA}^{\text{MIP}}: \quad \min_{\substack{\boldsymbol{y}\in Y,\ \boldsymbol{x}\in\mathbb{X}^*, \boldsymbol{u}, \boldsymbol{Z},\\ \boldsymbol{x}^P, \tilde{\boldsymbol{x}}^P, \bar{\boldsymbol{x}}^P, \bar{\bar{\boldsymbol{x}}}^P,\\ \tilde{\boldsymbol{x}}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}}} \sum_{n\in N} v_n Z_n, \tag{19}$$
$$\text{s.t.} \quad (7)\text{–}(13),$$
$$(15)\text{–}(18).$$

The above model could be extended to include target-dependent values for each attacker. This would require the replacement of the current objective function $\sum_{n\in N} v_n Z_n$ by $\sum_{n\in N}\sum_{(i,t)\in A^*} v_{ni}x_{nit}Z_n$. With this modification, attacker n would take the value of target i with probability Z_n if he traverses the (i,t) arc. The product $x_{nit}Z_n$ can be linearized using a MIP linear construct similar to that in (17).

2.4. *Benders Decomposition of* NTDA$^{\text{MIP}}$

Fixing $y = (\tilde{\boldsymbol{y}}, \bar{\boldsymbol{y}}, \bar{\bar{\boldsymbol{y}}})$ to $\hat{\boldsymbol{y}} = (\hat{\tilde{\boldsymbol{y}}}, \hat{\bar{\boldsymbol{y}}}, \hat{\bar{\bar{\boldsymbol{y}}}})$ in NTDA$^{\text{MIP}}$ produces a linear subproblem which is separable for each attacker, making NTDA$^{\text{MIP}}$ amenable to BD. This technique has been used in other detector-evader problems including uncertainty, which exhibit a special structure that can be exploited via the so-called L-shaped method [15].

In what follows, a superindex k represents the iteration counter for the BD algorithm. For a fixed $\boldsymbol{y} = \hat{\boldsymbol{y}}^k$, the subproblem for the nth attacker, $\text{SP}^k_n(\hat{\boldsymbol{y}}^k)$, calculates his optimal (as perceived) route, and its actual objective value $v_n Z^k_n$. This is equivalent to solving a network shortest path problem using perceived probabilities of interdiction as costs, and then post-processing its solution to calculate the actual interdiction probability. Note that, by construction, $\text{SP}^k_n(\hat{\boldsymbol{y}}^k)$ is always feasible and bounded. Let $Z(\text{SP}^k(\hat{\boldsymbol{y}}^k)) = \sum_n v_n Z^k_n$ be the total objective function of the subproblem at iteration k. The ensuing master problem can be stated as follows:

$$\text{MP}^k: \quad \min_{\boldsymbol{y}\in Y, Z} Z,$$
$$\text{s.t.} \quad Z \ge \hat{\upsilon}^{k'} + \sum_{(i,j)\in A^*} \left(\tilde{\alpha}^{k'}_{ij}\tilde{y}_{ij} + \bar{\alpha}^{k'}_{ij}\bar{y}_{ij} + \bar{\bar{\alpha}}^{k'}_{ij}\bar{\bar{y}}_{ij}\right),$$
$$\forall k' = 1, \ldots, k.$$

The independent term and the coefficients of a generic cut k are calculated from the subproblem's dual solution (10)–(12), (15), and (17) as follows:

$$\tilde{\alpha}_{ij}^k = \sum_{n\in N} \left(\tilde{\pi}_{nij}^{P,k} + \tilde{c}_{nij}\pi_{nij}^k + \tilde{\pi}_{nij}^{1,k} + \tilde{\pi}_{nij}^{2,k}\right),$$

$$\bar{\alpha}_{ij}^k = \sum_{n\in N} \bar{\pi}_{nij}^{P,k},$$

$$\bar{\bar{\alpha}}_{ij}^k = \sum_{n\in N} \left(\bar{\bar{\pi}}_{nij}^{P,k} + \bar{\bar{c}}_{nij}\pi_{nij}^k + \bar{\bar{\pi}}_{nij}^{1,k} + \bar{\bar{\pi}}_{nij}^{2,k}\right), \text{ and}$$

$$\hat{\upsilon}^k = Z(\mathrm{SP}^k(\hat{\boldsymbol{y}}^k)) - \sum_{(i,j)\in A^*} \left(\tilde{\alpha}_{ij}^k\hat{\tilde{y}}_{ij} + \bar{\alpha}_{ij}^k\hat{\bar{y}}_{ij} + \bar{\bar{\alpha}}_{ij}^k\hat{\bar{\bar{y}}}_{ij}\right).$$

The problem's asymmetry can also be observed by inspecting the above coefficients. $\tilde{\alpha}_{ij}^k, \bar{\alpha}_{ij}^k$ and $\bar{\bar{\alpha}}_{ij}^k$ bound the maximum decrease in the defender's objective (at the kth iteration) should he place an asset of the corresponding type on the arc. If arc (i,j) has not been interdicted, and attacker n traverses the arc (i.e., $x_{nij} = 1$), then $\tilde{\pi}_{nij}^P, \bar{\pi}_{nij}^P, \bar{\bar{\pi}}_{nij}^P$ from (10)–(12) may become strictly negative and contribute to that bound. On the attacker's side, (15), the contribution of additional transparent interdictions and decoys to the defender's objective is captured by $\tilde{c}_{nij}\pi_{nij} \le 0$ and $\bar{\bar{c}}_{nij}\pi_{nij} \le 0$, respectively.

2.5. *Heuristic Approximation:* NTDA$^{\mathrm{H}}$

Our NTDA$^{\mathrm{H}}$ model suggests that, instead of (1)–(2), the defender optimizes $\min_{\boldsymbol{y},\boldsymbol{Z}} \sum_{n\in N} v_n \log(Z_n)$, i.e., replacing evasion probabilities by their logarithms in the expected value calculation. Note this model does not rely on the generalized network, therefore it does not require all the $x_{nij}^P, \tilde{x}_{nij}^P, \bar{x}_{nij}^P, \bar{\bar{x}}_{nij}^P$ variables and their associated constraints. This comes at the expense of producing a heuristic solution, which is only guaranteed to be optimal when $|N| = 1$, and otherwise has unknown quality. Intuitively, however, we expect the approximating model works well given the apparent correlation between the objective function values $\sum_{n\in N} v_n Z_n$ and $\sum_{n\in N} v_n \log(Z_n)$ for $0 < Z_n \le 1$, so that solving the latter produces an acceptable solution for the original problem. (Indeed, this is the case in most of our computational experience.) Naturally, we still prefer to solve NTDA$^{\mathrm{MIP}}$ directly, or via BD, whenever possible.

The objective function of NTDA$^{\mathrm{H}}$ becomes:

$$\min_{\boldsymbol{y}\in Y} \sum_{n\in N} v_n \left(\sum_{(i,j)\in A} (c_{nij} + \tilde{c}_{nij}\tilde{y}_{ij} + \bar{c}_{nij}\bar{y}_{ij})x_{nij} \right). \tag{20}$$

Since this objective also involves products of decision variables, we resort to linearizing constraints (17) to state our heuristic model as follows:

$$\mathrm{NTDA}^{\mathrm{H}}:$$

$$\min_{\substack{\boldsymbol{y}\in Y,\, \boldsymbol{x}\in \aleph^*,\, \boldsymbol{u},\\ \tilde{\boldsymbol{x}},\, \bar{\boldsymbol{x}},\, \bar{\bar{\boldsymbol{x}}}}} \sum_{n\in N} v_n \left(\sum_{(i,j)\in A} (c_{nij}x_{nij} + \tilde{c}_{nij}\tilde{x}_{nij} + \bar{c}_{nij}\bar{x}_{nij}) \right),$$

$$\text{s.t.} \quad (15)-(18). \tag{21}$$

2.6. *Robust Models:* RNTDA$^{\mathrm{NL}}$ *and* RNTDA$^{\mathrm{MIP}}$

As an alternative to the models based on total expected value contributed by all attackers, the defender could solve a NTDA model in the spirit of robust optimization. Here, robustness refers to the worst of the expected values allowed by an individual attacker. The nonlinear, robust NTDA model, RNTDA$^{\mathrm{NL}}$, is stated as:

$$\mathrm{RNTDA}^{\mathrm{NL}}: \min_{\boldsymbol{y}\in Y, z} z, \tag{22}$$

$$\text{s.t.} \quad z \ge v_n \prod_{(i,j)\in A} q_{nij}^{x_{nij}} \tilde{r}_{nij}^{x_{nij}\tilde{y}_{ij}} \bar{\bar{r}}_{nij}^{x_{nij}\bar{\bar{y}}_{ij}}, \quad \forall n \in N, \tag{23}$$

$$(15)-(18),$$

where, for each attacker $n \in N$, $\boldsymbol{x}_n$ solves problem F_n stated in (4) and (5). Naturally, we may linearize the attacker's problem as in (5) and (6). To linearize the defender's problem in RNTDA$^{\mathrm{NL}}$, we use $\hat{z} = \log(z)$ in (22) and (23):

$$\mathrm{RNTDA}^{\mathrm{MIP}}: \min_{\substack{\boldsymbol{y}\in Y,\, \boldsymbol{x}\in \aleph^*,\, \boldsymbol{u},\\ \tilde{\boldsymbol{x}},\, \bar{\boldsymbol{x}},\, \bar{\bar{\boldsymbol{x}}},\, \hat{z}}} \hat{z}, \tag{24}$$

$$\text{s.t.} \quad \hat{z} \ge \log(v_n) + \sum_{(i,j)\in A} (c_{nij}x_{nij} + \tilde{c}_{nij}\tilde{x}_{nij} + \bar{c}_{nij}\bar{x}_{nij}), \quad \forall n \in N, \tag{25}$$

$$(15)-(18).$$

2.7. *Attacker Behaviors*

For the purpose of assessing both the benefits and the possible caveats of the deceptive interdiction concept presented in this article, we analyze four attacker behaviors, similar to those proposed in [11]:

- Pseudo-optimal behavior: Attackers follow optimal routes but only according to their perceptions (as expected by the defender). That is, each attacker n determines his route by solving problem F_n given in (4) and (5).
- Skeptic, pseudo-optimal behavior: As in the pseudo-optimal behavior, attacker n solves F_n to find his route. However, instead of following this route, he randomly deletes one arc from it and solves F_n again for a new route. Two variants are implemented: (a) Preemptive: The attacker still uses his origin s_n as the starting point for the recalculated route. (b) Dynamic: The attacker follows the prescribed route up to the removed arc, and then uses the tail node of that arc as an intermediate origin before resolving F_n. In both variants the attacker's route depends on the arc removed; therefore our computational results run all possible cases and report the average.
- Indifferent behavior: Attackers use a geographical information system to follow the shortest route (based, e.g., on distance or travel time) ignoring all defenses. In our examples, we use the Euclidean distance

for all attackers. Thus, if nodes i, j have coordinates (i_x, i_y), (j_x, j_y), respectively, we define: $d_{nij} = \sqrt{(i_x - j_x)^2 + (i_y - j_y)^2}$, $\forall n \in N, \forall (i,j) \in A$.

– Cognizant behavior: attackers realize traps and decoys, and plan their routes according to the actual probabilities of interdiction (i.e., all defender's assets become transparent).

3. COMPUTATIONAL RESULTS

In this section, we illustrate our deception models using a notional network. We provide all the input data so our results can be reproduced. We also test two random variants of the baseline network.

3.1. Test Network Setup

Our baseline network (see Fig. 1) has 53 nodes, 196 directed arcs (shown as undirected connections in the figure), and three attackers, "A," "B," and "C." One fictitious node (labeled "A,B,C") is used to allow each attacker to start at any of four possible locations. (Arcs originating at this node are not bidirectional.) The attackers' values are $v_A = 20$, $v_B = 30$, and $v_C = 50$, and each attacker has two possible targets (labeled "A*", "B*," and "C*," respectively). Target nodes are connected to a super-sink (not shown) in the augmented network used by our models.

Coordinates for the original nodes (i.e., excluding "A,B,C") are all integer-valued, ranging from $(i_x, i_y) = (1, 1)$ for lower-left node "111" to $(i_x, i_y) = (10, 10)$ for upper-right node "338." For simplicity, interdiction probabilities (actual or perceived) have been made identical for all attackers, therefore we drop their subindex n below. In the baseline network, these probabilities are assigned as follows (see examples in Fig. 1):

$$1 - q_{ij} = p_{ij} = \frac{1}{100}\sqrt{(i_x - j_x)^2 + (i_y - j_y)^2} + \frac{1}{500}(i_x + i_y + j_x + j_y) + 0.005\Delta_y,$$

where $\Delta_y = 1$ if $i_y \neq j_y$ (and $\Delta_y = 0$ otherwise). p_{ij} is designed to increase with the arc length and with its proximity to the upper-right corner. The lowest nominal probability of interdiction is $p_{113,114} = p_{114,113} = 1/100 + 9/500 = 0.028$, and the highest is 0.109 for arcs (238, 338) and (338, 238).

$$1 - \tilde{q}_{ij} = \tilde{p}_{ij} = \begin{cases} 0.25, & \text{for long arcs, e.g., (218, 228) or (118, 228),} \\ 0.35, & \text{for oblique arcs, e.g., (218, 225),} \\ 0.50, & \text{for short arcs, e.g., (225, 226) or (223, 225).} \end{cases}$$

$$1 - \bar{q}_{ij} = \bar{p}_{ij} = \begin{cases} 0.25, & \text{for long arcs,} \\ 0.30, & \text{for oblique arcs,} \\ 0.35, & \text{for short arcs.} \end{cases}$$

$$1 - \bar{\bar{q}}_{ij} = \bar{\bar{p}}_{ij} = \begin{cases} 0.40, & \text{for long arcs,} \\ 0.50, & \text{for oblique arcs,} \\ 0.60, & \text{for short arcs.} \end{cases}$$

TABLE 1. Baseline network results for cases with three transparent assets and up to two traps or decoys.

			Attacker's value V under different behaviors				
Case	$(\tilde{R}, \bar{R}, \bar{\bar{R}})$	Fixed $\tilde{x}$?	Pseudo optimal	Skeptical (Preemt.)	Skeptical (Dynamic)	Indifferent	Cognizant
1	(3,0,0)	No	70.6	69.3	67.7	62.3	70.6
2	(3,1,1)	No	57.8	65.9	60.0	62.3	74.3
3	(3,1,1)	Yes (1)	57.8	65.9	60.0	62.3	70.6
4	(3,2,0)	No	59.0	61.2	60.0	56.7	70.4
5	(3,2,0)	Yes (1)	59.0	61.2	60.0	56.7	70.4
6	(3,0,2)	No	62.6	67.1	63.1	60.0	78.3
7	(3,0,2)	Yes (1)	68.0	67.2	62.4	62.3	70.6

Remark. *We set* $q_{\text{ABC},j} = \tilde{q}_{\text{ABC},j} = \bar{q}_{\text{ABC},j} = \bar{\bar{q}}_{\text{ABC},j} = 1$, $\forall j \in \{111, 117, 217, 317\}$ *for the fictitious arcs originating at the "A,B,C" node.*

The best routes for attackers in the non-interdicted network, that is, using only nominal probabilities of interdiction, are as follows (see Fig. 1): 117-118-128-135 for attacker "A," with overall probability of success $Z_A = 82.3\%$; 217-218-228-235 for attacker "B," with probability $Z_B = 79.2\%$; and, 317-318-328-335 for "C," with probability $Z_C = 76.2\%$. After accounting for the attacker's value, the overall expected value for the attacking team is $V = 78.3$ (where the maximum possible is $v_A + v_B + v_C = 100$).

3.2. Baseline Network: Detailed Results for Selected Runs

Table 1 displays results for several runs of an instance with three transparent interdictions, one trap and one decoy, as specified by the asset cardinality vector $(\tilde{R}, \bar{R}, \bar{\bar{R}})$. First, we run NTDA$^{\text{MIP}}$ without the traps and decoys (Case 1), i.e., $(\tilde{R}, \bar{R}, \bar{\bar{R}}) = (3, 0, 0)$. This case is equivalent to a standard DA model with full transparency, as in [5], where the defender's goal is to minimize his worst-case outcome for any possible route the attackers might find. The defender places transparent interdictions on arcs 118–128, 218–228, and 328–335 (indicated by thick arrows in Fig. 2) to decrease the total attackers' expected value from $V = 78.3$ (calculated above with nominal probabilities) to $V = 70.6$. Naturally, given the structure of DA models, the pseudo-optimal routes for the attackers match the cognizant ones, that is, this plan ensures attackers cannot improve this expectation by any other behavior.

Cases 2 and 3 add one trap and one decoy to Case 1. In Case 2, we reoptimize the location of all interdiction assets. The optimal interdiction plan is 112–122, 118–128, and 218–228 for transparent assets, 225 and 226 for the trap, and 328–335 for the decoy. The result is a notable gain, $V = 57.8$, if the attacker's behavior is pseudo-optimal, as the defender expects. The outcomes for skeptical and indifferent behaviors are worse, but the highest risk this plan poses is when attackers behave cognizantly, in which case $V = 74.3$. To hedge against these possible outcomes, we analyze Case 3,

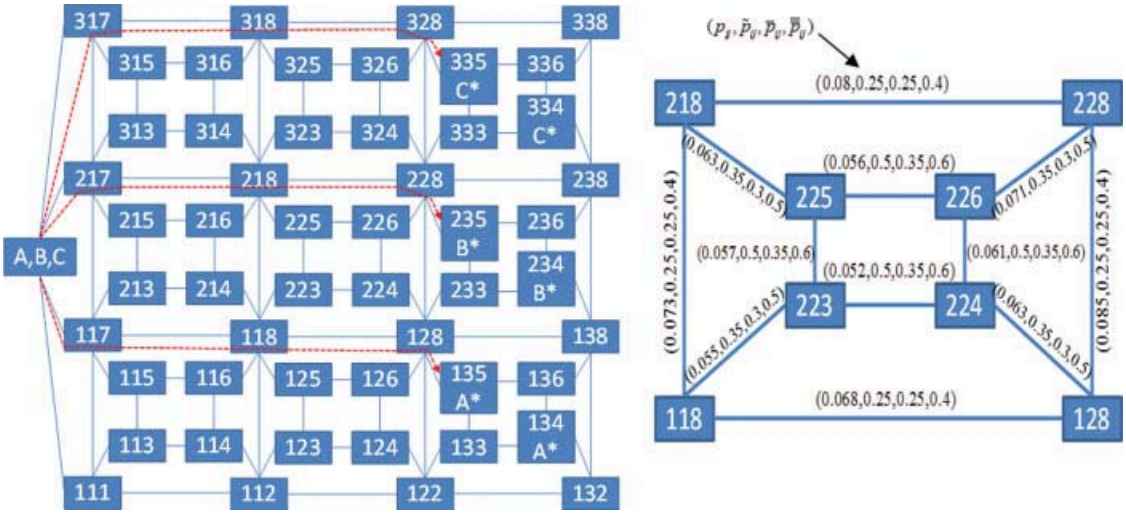

FIG. 1. Baseline test network (left) and detailed probabilities of interdiction for one portion of the network (right). Best nominal routes for attackers are depicted as dotted lines. Nominal probabilities (p_{ij}) shown for oblique arcs are rounded to three decimals. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

where the defender fixes his transparent interdictions as in Case 1, and then deploys the traps and decoys optimally. This idea is similar to that of secrecy in [4]. The trap is again allocated to arc 225 and 226 and the decoy to arc 112 and 122 (Fig. 2). We observe that the attackers' value under the assumption of pseudo-optimal behavior remains $V = 57.8$, but the defender's worst case (cognizant attackers) improves to $V = 70.6$ (as in Case 1). From this, it is clear that Case 3 is superior to Case 2, and that the latter has multiple optimal solutions which should not be overlooked when analyzing multiple behaviors. Reference [2] acknowledges this issue and describes an alternative formulation to choose the worst-possible of these solutions (from the defender's perception).

Cases 4 and 5 explore two traps (and no decoys). Both cases render the same solution whether transparent assets are deployed before or simultaneously with traps. From Cases 2 to 5 it may seem that it could be beneficial to deploy transparent assets first, but the next example demostrates that this result cannot be generalized.

In Cases 6 and 7 (Fig. 3), the defender has two decoys and no traps. Case 6 optimizes the location of all interdiction assets simultaneously. Transparent interdictions occur on arcs 333 and 334, 333–335, and 338–336, whereas 238–334

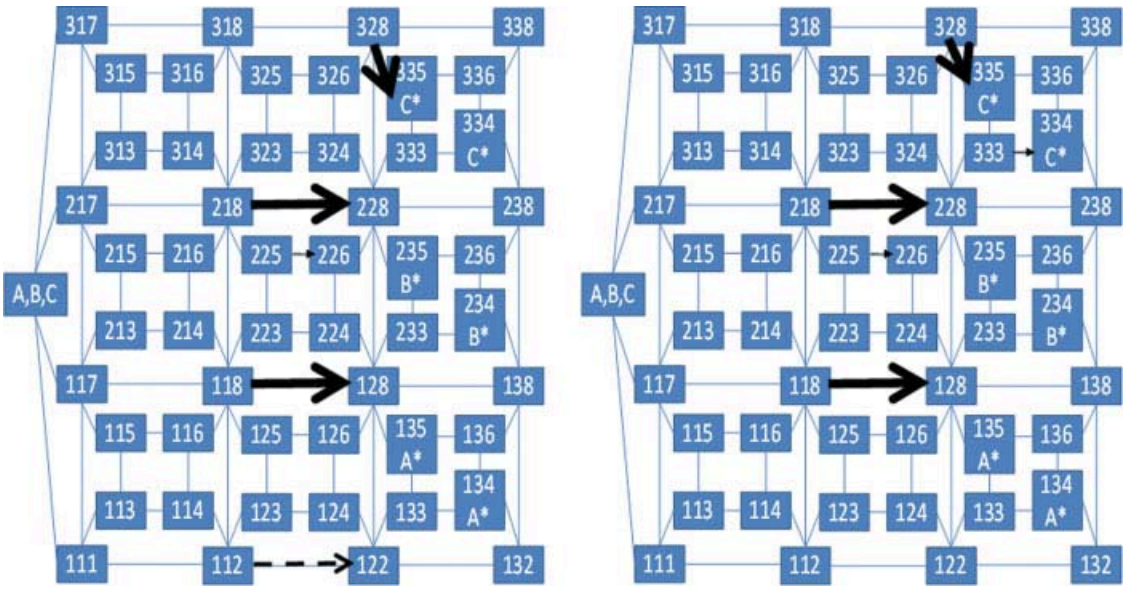

FIG. 2. Cases 1 and 3 (left), and Cases 4 and 5 (right). We use thick, solid arrows to depict transparent assets, solid thin arrows for traps, and dashed arrows for decoys. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and 328–335 become the arcs selected for decoys. Here, the defender is blocking the access to both "C*" targets for attacker "C." Because of the intelligent use of decoys, in fact, the attacker is expected to traverse a transparent interdiction on arc 338–336 as his best choice to arrive at target "335." This strategy produces an overall pseudo-optimal value $V = 62.6$, but if "C" realizes the decoys the worst-case would be $V = 78.3$ (as in the non-interdicted network). In contrast, in Case 7 we fix transparent assets first as in Case 1. Then, we deploy decoys to arcs 112–122 and 228–333, which renders an improvement in the worst-case, $V = 70.6$, at the expense of worsening the pseudo-optimal value to $V = 68.0$. Thus, if we had to choose one of these two strategies, the answer would not be obvious.

The above examples show deceptive strategies may use transparent assets and decoys to funnel attackers into invisible traps. Arguably, an attacker could be suspicious of his perceptions and conjecture traps might have been placed somewhere along his path. In Section 4, we develop a multiobjective model that can help address this issue.

3.3. Testing Multiple Combinations of Assets

In this section, we assess our exact approaches (solving NTDA$^{\text{MIP}}$ directly and via BD) under the assumption of pseudo-optimal behavior, on 135 combinations of $(\tilde{R}, \bar{R}, \bar{\bar{R}})$ with up to eight combined interdiction assets. We also compare these solutions with the heuristic result provided by the NTDA$^{\text{H}}$ model.

The computational experience has been carried out on a 2.60 GHz Dell Precision M6300 dual-processor laptop (but using only one processor), with 3.50 Gb of RAM, running under Windows XP. All the mathematical models and auxiliary code have been implemented in Xpress Mosel 2.4.0 [8]. Models have been solved using Xpress Optimizer 19.00.00 [8] and/or CPLEX 11.2 [10] as the solver engines, using their default settings. In each of these three-attacker cases, the NTDA$^{\text{MIP}}$ model has 9,010 constraints and 5,675 variables (600 binary), with coefficient matrix density 0.06%. Both Xpress and CPLEX presolved matrices exhibit very modest reductions (e.g., Xpress only eliminates 4% of the constraints, 4% of all the variables, and 4% of the discrete variables). The heuristic version NTDA$^{\text{H}}$ has 6,375 constraints and 3,181 variables (600 binary), with coefficient matrix density 0.1%. Presolved matrices eliminate 4% of the constraints and 12% of the variables, but again only 4% of the discrete variables. The CPLEX solver outperforms the Xpress solver in producing 1% near-optimal solutions (when possible) to either of our three approaches (direct MIP, BD, or heuristic). All methods are run for up to 30 minutes, or until the optimality gap reaches 1%. Results are displayed in Table 2.

There are 74 cases where both NTDA$^{\text{MIP}}$ and BD converge within 1%, 40 cases where BD converges but NTDA$^{\text{MIP}}$ does not, and one case where the opposite occurs. In the other 20 cases, neither exact method converges to a near-optimal solution in the allotted time.

It is worth noting that in most cases, the NTDA$^{\text{H}}$ model produces the same (or very close) solutions to those of the exact methods. The largest improvement obtained with exact methods over NTDA$^{\text{H}}$ is 8%, for case $(\tilde{R}, \bar{R}, \bar{\bar{R}}) = (4, 1, 2)$. The largest improvement obtained with NTDA$^{\text{H}}$ is 6%, for cases $(\tilde{R}, \bar{R}, \bar{\bar{R}}) = (5, 1, 2)$ and $(\tilde{R}, \bar{R}, \bar{\bar{R}}) = (3, 1, 4)$. NTDA$^{\text{H}}$ typically reaches these solutions in a fraction of the time needed by the exact methods. (Computational times not listed here; see similar comparison in Sections 3.4 and 4.2). Even though a defender's solution $\boldsymbol{y}$ may be suboptimal, it is important to note that the attackers' routes (for the given $\boldsymbol{y}$) are calculated correctly by our models, because $\boldsymbol{x}$ must follow the explicit constraints derived from strong duality.

In our examples, the objective value of the NTDA$^{\text{MIP}}$ model, $V(\tilde{R}, \bar{R}, \bar{\bar{R}})$, as a function of the number of assets used, ranges from $V(0, 0, 0) = 78.3$ to $V(0, 3, 5) = 37.5$. This range may allow the defender to determine sets of minimum configurations which achieve a pre-specified value. For example, $V(\tilde{R}, \bar{R}, \bar{\bar{R}}) \leq 45$ for the following configurations of $(\tilde{R}, \bar{R}, \bar{\bar{R}})$: (5,1,1), (3,2,1), (4,1,3), (1,5,1), (2,3,2), (5,2,0), (4,3,0), and (1,2,4). The defender may prefer any of these configurations over the others based, for example, on their cost.

The trend of decrease in attacker's value as a function of the number of defender's deceptive assets can be approximated by fitting the above data with two linear regression models, $\hat{V}^0(\tilde{R}, \bar{R}, \bar{\bar{R}})$ and $\hat{V}^1(\tilde{R}, \bar{R}, \bar{\bar{R}})$. The first model considers individual contributions by asset type, whereas the second adds pair-wise interactions.

Remark. *Our models use the best known solution for each case, whether it is provably optimal or not.*

The resulting fitted values produce:

$$\hat{V}^0(\tilde{R}, \bar{R}, \bar{\bar{R}}) = 81.29 - 3.99\tilde{R} - 5.38\bar{R} - 4.05\bar{\bar{R}}$$

$$\hat{V}^1(\tilde{R}, \bar{R}, \bar{\bar{R}}) = 76.92 - 2.82\tilde{R} - 2.62\bar{R} - 2.96\bar{\bar{R}} - 0.92\tilde{R}\bar{R} + 0.06\tilde{R}\bar{\bar{R}} - 0.87\bar{R}\bar{\bar{R}}$$

Coefficients of determination for these models are $R^2 = 0.873$ for $\hat{V}^0$ and $R^2 = 0.941$ for $\hat{V}^1$. Both models point out that gains per transparent interdiction asset and decoy are similar. This is reasonable given that, under the assumed pseudo-optimal behavior, attackers view decoys as actual interdiction points to avoid. In the first model it appears that traps are the most valuable asset. This seems consistent with the fact that, even if they are less effective than transparent assets (if traversed), they are not visible to attackers, and may be strategically placed along their expected routes. The second model makes traps less valuable if individually considered, but again more valuable when used in conjunction with transparent or decoy assets.

3.4. Results for Random Networks

We extend our testing to two random networks, "RN1" and "RN2." Both networks are generated from our baseline network described in Section 3.1, by modifying the interdiction

TABLE 2. Solution comparison for multiple combinations of assets $(\tilde{R}, \bar{R}, \bar{\bar{R}})$. The columns labeled "MIP%" and "BD%" show the optimality gaps for the NTDA^{MIP} solved directly as a MIP and via BD, respectively, after up to 30 minutes of computation.

			MIP	BD	H				MIP	BD	H				MIP	BD	H				MIP	BD	H				MIP	BD	H
$\tilde{R}$	$\bar{R}$	$\bar{\bar{R}}$	%	%	%	$\tilde{R}$	$\bar{R}$	$\bar{\bar{R}}$	%	%	%	$\tilde{R}$	$\bar{R}$	$\bar{\bar{R}}$	%	%	%	$\tilde{R}$	$\bar{R}$	$\bar{\bar{R}}$	%	%	%	$\tilde{R}$	$\bar{R}$	$\bar{\bar{R}}$	%	%	%
0	0	0	–	–	–	1	2	1	–	–	–	0	1	4	–	–	–	5	2	0	–	–	–	5	3	0	>	–	–
1	0	0	–	–	–	1	1	2	25	–	–	0	0	5	–	–	–	5	1	1	>	–	6	5	2	1	>	30	–
0	1	0	–	–	–	1	0	3	–	–	–	5	1	0	–	–	–	5	0	2	>	>	–	5	1	2	>	>	−6
0	0	1	–	–	–	0	4	0	–	–	–	5	0	1	–	–	–	4	3	0	–	–	–	5	0	3	>	>	–
2	0	0	–	–	–	0	3	1	–	–	–	4	2	0	–	–	–	4	2	1	>	–	–	4	4	0	38	–	–
1	1	0	–	–	–	0	2	2	–	–	–	4	1	1	43	–	–	4	1	2	>	–	8	4	3	1	>	–	–
1	0	1	–	–	–	0	1	3	–	–	–	4	0	2	19	>	–	4	0	3	>	>	–	4	2	2	>	>	–
0	2	0	–	–	–	0	0	4	–	–	–	3	3	0	–	–	2	3	4	0	–	–	5	4	1	3	>	>	−3
0	1	1	–	–	–	5	0	0	–	–	–	3	2	1	>	–	–	3	3	1	>	–	–	4	0	4	>	–	–
0	0	2	–	–	–	4	1	0	–	–	–	3	1	2	>	–	–	3	2	2	>	–	–	3	5	0	–	–	7
3	0	0	–	–	–	4	0	1	–	–	–	3	0	3	28	–	–	3	1	3	>	>	4	3	4	1	>	–	–
2	1	0	–	–	–	3	2	0	–	–	–	2	4	0	–	–	–	3	0	4	>	>	–	3	3	2	>	–	–
2	0	1	–	–	–	3	1	1	40	–	–	2	3	1	83	–	2	2	5	0	–	–	–	3	2	3	>	>	–
1	2	0	–	–	–	3	0	2	–	–	–	2	2	2	>	–	–	2	4	1	36	–	5	3	1	4	>	>	−6
1	1	1	–	–	–	2	3	0	–	–	–	2	1	3	>	–	7	2	3	2	>	–	–	3	0	5	>	>	−2
1	0	2	–	–	–	2	2	1	60	–	–	2	0	4	>	–	–	2	2	3	>	–	–	2	5	1	>	2	7
0	3	0	–	–	–	2	1	2	28	–	–	1	5	0	–	–	–	2	1	4	>	>	4	2	4	2	>	–	–
0	2	1	–	–	–	2	0	3	8	–	–	1	4	1	–	–	–	2	0	5	>	–	–	2	3	3	>	27	–
0	1	2	–	–	–	1	4	0	–	–	–	1	3	2	>	–	2	1	5	1	–	–	–	2	2	4	>	–	–
0	0	3	–	–	–	1	3	1	–	–	–	1	2	3	>	–	–	1	4	2	76	–	5	2	1	5	>	>	−5
4	0	0	–	–	–	1	2	2	–	–	–	1	1	4	56	–	7	1	3	3	>	–	–	1	5	2	>	2	7
3	1	0	–	–	–	1	1	3	25	–	–	1	0	5	3	>	–	1	2	4	>	–	–	1	4	3	>	–	–
3	0	1	–	–	–	1	0	4	–	–	–	0	5	1	–	–	–	1	1	5	90	–	4	1	3	4	>	–	–
2	2	0	–	–	–	0	5	0	–	–	–	0	4	2	–	–	–	0	5	2	–	–	–	1	2	5	>	>	–
2	1	1	–	–	–	0	4	1	–	–	–	0	3	3	–	–	2	0	4	3	–	>	5	0	5	3	–	–	7
2	0	2	–	–	–	0	3	2	–	–	–	0	2	4	–	–	–	0	3	4	–	–	–	0	4	4	–	–	–
1	3	0	–	–	–	0	2	3	–	–	–	0	1	5	–	–	–	0	2	5	–	–	–	0	3	5	–	–	–

A dash "-" indicates the gap is less than 1%, and a ">" symbol indicates it is greater than 100%. The column labeled "H %" shows the percent difference between the best solution produced by either exact approach and the NTDA^{H} solution: A dash indicates this difference is under 1%; if positive, either the NTDA^{MIP} direct solution or the BD solution improve the NTDA^{H} solution by that percentage; if negative, NTDA^{H} produces the best solution of the three methods.

probabilities of each interdictable arc as follows: In RN1, we set $p_{nij} = 0.001 + U_{nij}(0, 0.3)$; $\tilde{p}_{nij} = 0.5 + \tilde{U}_{nij}(0, 0.5)$; $\bar{p}_{nij} = 0.5\tilde{p}_{nij}$; and, $\bar{\bar{p}}_{nij} = 0.8\tilde{p}_{nij}$, where both $U_{nij}(a, b)$ and $\tilde{U}_{nij}(a, b)$ denote independently generated random variables, uniformly distributed on the interval (a, b). RN2 also generates p_{nij} randomly as in RN1. However, probabilities for all other assets are replaced by expected values, that is, $\tilde{p}_{nij} = 0.75; \bar{p}_{nij} = 0.375$; and, $\bar{\bar{p}}_{nij} = 0.6$.

For each of six selected configurations of $(\tilde{R}, \bar{R}, \bar{\bar{R}})$ we create ten samples of RN1 and RN2. Each sample instance is run for up to 1 hour, or until the gap of the NTDA^{MIP} and BD solutions reaches 1%. Results are summarized in Tables 3 and 4 for RN1 and RN2, respectively.

In both cases, BD outperforms solving NTDA^{MIP} directly. However, there are sample instances that do not converge in the allotted time, especially for certain combinations of interdiction assets.

For RN1, combinations with one decoy or one transparent asset exhibit better solvability. For these sample cases, BD converges in 34 of 40 instances, and achieves moderate gaps (after one hour of computation) in the other six instances. On the contrary, convergence is poor in the 20 cases from samples with one trap and multiple decoys and transparent assets. In these instances, average gaps are high even after excluding six cases where a relative gap is not available (because the lower bound provided by the master problem is negative after the allotted time.) Solving NTDA^{H} is faster and it provides better solutions than BD in several of the instances where BD does not converge. For RN2 the effectiveness of each approach is very similar to that of RN1 for the same $(\tilde{R}, \bar{R}, \bar{\bar{R}})$. That is, the specific interdiction probabilities do not affect the difficulty of the problem as much as the available number of each type of interdiction asset.

4. THE MULTIOBJECTIVE MODEL

The detailed examples from Section 3.2 show that the defender may need to balance the potential improvement from strategies that assume pseudo-optimal behaviors and the risk should the attackers behave differently and/or be suspicious about apparent "holes" in the defense system. In this section, we address this issue by developing multi-objective optimization extensions of the models introduced in Section 2, and illustrate them with a detailed example.

TABLE 3. Computational results for random network RN1. We run ten samples of each instance of $(\tilde{R}, \bar{R}, \bar{\bar{R}})$ with each approach.

	NTDA$^{\text{MIP}}$			BD			NTDA$^{\text{H}}$		
$(\tilde{R}, \bar{R}, \bar{\bar{R}})$	# conv.	Avg. CPU conv. (s)	Avg. gap not conv.	# conv.	Avg. CPU conv. (s)	Avg. gap not conv.	# worse	# better	Avg. CPU (s)
(5,3,1)	0	n/a	>100%	8	1,252	28.8%	1	0	195
(5,1,3)	0	n/a	>100%	0	n/a	41.8%∗	2	4	1,165
(3,5,1)	7	1,864	14.9%	10	92	n/a	0	0	75
(3,1,5)	0	n/a	>100%	0	n/a	51.8%∗	0	4	1,455
(1,5,3)	9	2,068	7.8%	10	147	n/a	1	0	78
(1,3,5)	0	n/a	>100%	6	1,479	10.6%	0	2	200

For NTDA$^{\text{MIP}}$ and BD we report the number of samples that converge within 1% in one hour of computation, the average computational time for those cases, and the average gap for the cases which did not converge, respectively. For NTDA$^{\text{H}}$ we report the number of samples where the heuristic solution was found to be worse than 1% with respect to the NTDA$^{\text{MIP}}$ or BD solutions, and the number of samples where it was 1% better than both of those solutions, along with the average computational time of all samples. (*) indicates that the average shown disregards at least one case in which BD's gap was not available because the last lower bound provided by the MP was negative.

4.1. Development

We seek to simultaneously minimize all of the objective values associated with each attacker's behavior. However, given that some of these objectives are conflicting with each other, we may only generate nondominated solutions (also known as Pareto or efficient solutions). We use the well-known method of weighted sum objectives (see, e.g., [7], pp. 24, 65–75), even though we cannot guarantee generating the entire efficient frontier because NTDA$^{\text{MIP}}$ is not a convex model.

To develop this extension, we define the set of behaviors, $B = \{\text{pseudo-optimal}, \text{indifferent}, \text{cognizant}\}$, and let the defender specify weights $\lambda^b \geq 0$, $\forall b \in B$, where $\sum_{b\in B} \lambda^b = 1$. (The skeptical behavior is excluded because it cannot be modeled explicitly.) Below, data and variables indexed by b retain their original meaning with the added qualifier "for behavior b." The multiobjective problem we need to solve can be stated as

$$\min_{\boldsymbol{y}\in Y} \sum_{b\in B} \lambda^b f(\boldsymbol{y}; \boldsymbol{x}^b), \tag{26}$$

where $f(\boldsymbol{y}; \boldsymbol{x}^b)$ represents the expected value obtained by the attacker with behavior b and response $\boldsymbol{x}^b$, given the defender's interdiction plan $\boldsymbol{y}$. In its original nonlinear version, we state the above as the following multiobjective NTDL$^{\text{NL}}$ model, MO-NTDA$^{\text{NL}}$:

$$\text{MO-NTDA}^{\text{NL}}: \min_{\boldsymbol{y}\in Y,Z} \sum_{b\in B} \lambda^b \sum_{n\in N} v_n Z_n^b \tag{27}$$

$$\text{s.t.} \quad Z_n^b = \prod_{(i,j)\in A} q_{nij}^{x_{nij}^b} \tilde{r}_{nij}^{x_{nij}^b \tilde{y}_{ij}} \bar{r}_{nij}^{x_{nij}^b \bar{y}_{ij}}, \quad \forall n \in N, \forall b \in B, \tag{28}$$

where $\boldsymbol{x}_n^b$ solves F_n^b, the flow problem for the nth attacker under behavior b:

$$\text{F}_n^b: \begin{cases} \max\limits_{x_n^b \in \aleph_n^*} \prod\limits_{(i,j)\in A} q_{nij}^{x_{nij}^b} \tilde{r}_{nij}^{x_{nij}^b \tilde{y}_{ij}} \bar{\bar{r}}_{nij}^{x_{nij}^b \bar{\bar{y}}_{ij}}, & \text{for } b = \text{pseudo-optimal}, \\ \max\limits_{x_n^b \in \aleph_n^*} \sum\limits_{(i,j)\in A} -d_{nij} x_{nij}^b, & \text{for } b = \text{indifferent}, \\ \max\limits_{x_n^b \in \aleph_n^*} \prod\limits_{(i,j)\in A} q_{nij}^{x_{nij}^b} \tilde{r}_{nij}^{x_{nij}^b \tilde{y}_{ij}} \bar{r}_{nij}^{x_{nij}^b \bar{y}_{ij}}, & \text{for } b = \text{cognizant}. \end{cases} \tag{29}$$

We now proceed as in the development of the NTDA$^{\text{MIP}}$ model to formulate MO-NTDA$^{\text{MIP}}$. For example, this model

TABLE 4. Computational results for random network RN2.

	NTDA$^{\text{MIP}}$			BD			NTDA$^{\text{H}}$		
$(\tilde{R}, \bar{R}, \bar{\bar{R}})$	Conv.	Avg. CPU conv. (s)	Avg. gap not conv.	Conv.	Avg. CPU conv. (s)	Avg. gap not conv.	Worse	Better	Avg. CPU (s)
(5,3,1)	0	n/a	>100%	9	1,119	50.2%	1	1	200
(5,1,3)	0	n/a	>100%	0	n/a	48.3%*	3	3	1,084
(3,5,1)	7	1,751	12.1%	10	114	n/a	0	0	68
(3,1,5)	0	n/a	>100%	0	n/a	54.6%*	0	3	1,316
(1,5,3)	8	2,047	47.2%	10	249	n/a	1	0	67
(1,3,5)	0	n/a	>100%	5	1,059	25.0%	0	2	267

See Table 3 for column description.

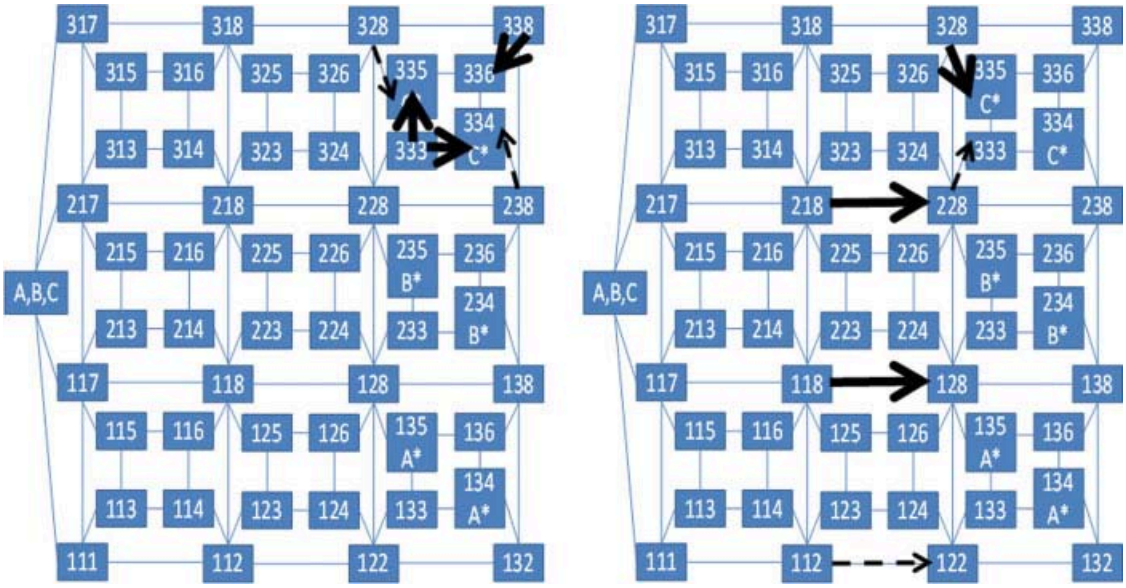

FIG. 3. Cases 6 (left) and 7 (right). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

must contain the dualized version of the attacker's problem (14)–(16) with x_{nij} replaced by x^b_{nij} for each behavior $b \in B$:

$$\sum_{(i,j)\in A} g^b(\tilde{y}_{ij}, \bar{y}_{ij}, \bar{\bar{y}}_{ij})\, x^b_{nij} = u^b_{n,s_n} - u^b_{nt}, \quad \forall n \in N, \forall b \in B, \tag{30}$$

$$u^b_{ni} - u^b_{nj} \geq g^b(\tilde{y}_{ij}, \bar{y}_{ij}, \bar{\bar{y}}_{ij}), \quad \forall n \in N, \forall (i,j) \in A, \forall b \in B \quad \left[\pi^b_{nij}\right], \tag{31}$$

$$u^b_{ni} - u^b_{nt} \geq 0, \quad \forall n \in N, \forall i \in T_n, \forall b \in B, \tag{32}$$

where:

$$g^b(\tilde{y}_{ij}, \bar{y}_{ij}, \bar{\bar{y}}_{ij}) = \begin{cases} c_{nij} + \tilde{c}_{nij}\tilde{y}_{ij} + \bar{\bar{c}}_{nij}\bar{\bar{y}}_{ij}, & \text{for } b = \text{pseudo-optimal}, \\ -d_{nij}, & \text{for } b = \text{indifferent}, \\ c_{nij} + \tilde{c}_{nij}\tilde{y}_{ij} + \bar{c}_{nij}\bar{y}_{ij}, & \text{for } b = \text{cognizant}. \end{cases} \tag{33}$$

To linearize decision variable products in (30) we proceed in the same way as we did for (14) using (17). Specifically, we replace any occurrence of type $\tilde{y}x^b$, $\bar{y}x^b$, and $\bar{\bar{y}}x^b$ by $\tilde{x}^b$, $\bar{x}^b$, and $\bar{\bar{x}}^b$, respectively:

$$\sum_{(i,j)\in A} \left(c_{nij}x_{nij} + \tilde{c}_{nij}\tilde{x}^b_{nij} + \bar{\bar{c}}_{nij}\bar{\bar{x}}^b_{nij}\right) = u^b_{n,s_n} - u^b_{nt}, \quad \forall n \in N,\ b = \text{pseudo-optimal}, \tag{34}$$

$$\sum_{(i,j)\in A} -d_{nij}x^b_{nij} = u^b_{n,s_n} - u^b_{nt}, \quad \forall n \in N,\ b = \text{indifferent}, \tag{35}$$

$$\sum_{(i,j)\in A} \left(c_{nij}x_{nij} + \tilde{c}_{nij}\tilde{x}^b_{nij} + \bar{c}_{nij}\bar{x}^b_{nij}\right) = u^b_{n,s_n} - u^b_{nt}, \quad \forall n \in N,\ b = \text{cognizant}, \tag{36}$$

and use the following set of logical constraints:

$$0 \leq \tilde{x}^b_{nij} \leq \tilde{y}_{ij}\left[\tilde{\pi}^{1,b}_{nij}\right];\ \tilde{x}^b_{nij} \leq x_{nij}\left[\tilde{\pi}^{2,b}_{nij}\right];\ \tilde{x}^b_{nij} \geq \tilde{y}_{ij} + x^b_{nij} - 1;$$
$$0 \leq \bar{x}^b_{nij} \leq \bar{y}_{ij};\ \bar{x}^b_{nij} \leq x_{nij};\ \bar{x}^b_{nij} \geq \bar{y}_{ij} + x^b_{nij} - 1;$$
$$0 \leq \bar{\bar{x}}^b_{nij} \leq \bar{\bar{y}}_{ij}\left[\bar{\bar{\pi}}^{1,b}_{nij}\right];\ \bar{\bar{x}}^b_{nij} \leq x_{nij}\left[\bar{\bar{\pi}}^{2,b}_{nij}\right];\ \bar{\bar{x}}^b_{nij} \geq \bar{\bar{y}}_{ij} + x^b_{nij} - 1, \quad \forall n \in N, \forall (i,j) \in A, \forall b \in B. \tag{37}$$

The last step consists of replicating the generalized network (described by (7)–(13)) for each of the behaviors, therefore replacing any occurrence of $\boldsymbol{x}$ and $\boldsymbol{x}^P$ variables by $\boldsymbol{x}^b$ and $\boldsymbol{x}^{P,b}$, respectively:

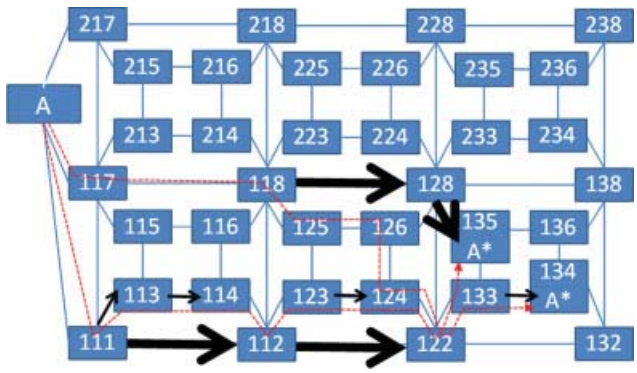

FIG. 4. Interdiction plan for $\boldsymbol{\lambda} = (1, 0)$, and attacker routes for both behaviors. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

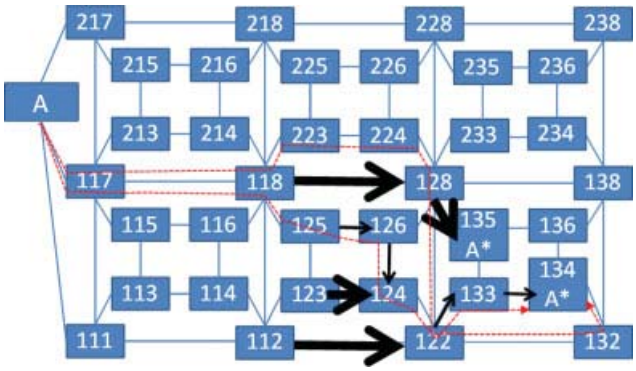

FIG. 5. Interdiction plan for $\boldsymbol{\lambda} = (0.75, 0.25)$, and attacker routes for both behaviors. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$\sum_{j|(s_n,j)\in A_n^*} x_{n,s_n,j}^{P,b} = 1, \quad \forall n \in N, \forall b \in B, \tag{38}$$

$$\sum_{j|(i,j)\in A_n^*} \left(x_{nij}^{P,b} + \tilde{x}_{nij}^{P,b} + \bar{x}_{nij}^{P,b} + \bar{\bar{x}}_{nij}^{P,b}\right) = \sum_{j|(j,i)\in A_n^*} \left(q_{nji} x_{nji}^{P,b} + \tilde{q}_{nji} \tilde{x}_{nji}^{P,b} + \bar{q}_{nji} \bar{x}_{nji}^{P,b} + q_{nji} \bar{\bar{x}}_{nji}^{P,b}\right),$$

$$\forall n \in N, \forall i \neq s_n, t, \forall b \in B, \tag{39}$$

$$\sum_{j|j\in T_n} \left(x_{njt}^{P,b} + \tilde{x}_{njt}^{P,b} + \bar{x}_{njt}^{P,b} + \bar{\bar{x}}_{njt}^{P,b}\right) = Z_n, \quad \forall n \in N, \forall b \in B, \tag{40}$$

$$\tilde{x}_{nij}^{P,b} \leq \tilde{y}_{ij}, \quad \forall n \in N, \forall (i,j) \in A_n^*, \forall b \in B \quad \left[\tilde{\pi}_{nij}^{P,b}\right], \tag{41}$$

$$\bar{x}_{nij}^{P,b} \leq \bar{y}_{ij}, \quad \forall n \in N, \forall (i,j) \in A_n^*, \forall b \in B \quad \left[\bar{\pi}_{nij}^{P,b}\right], \tag{42}$$

$$\bar{\bar{x}}_{nij}^{P,b} \leq \bar{\bar{y}}_{ij}, \quad \forall n \in N, \forall (i,j) \in A_n^*, \forall b \in B \quad \left[\bar{\bar{\pi}}_{nij}^{2,b}\right], \tag{43}$$

$$x_{nij}^{P,b} + \tilde{x}_{nij}^{P,b} + \bar{x}_{nij}^{P,b} + \bar{\bar{x}}_{nij}^{P,b} \leq x_{nij}^{b}, \quad \forall n \in N, \forall (i,j) \in A_n^*, \forall b \in B. \tag{44}$$

The complete MO-NTDA$^{\text{MIP}}$ model becomes:

$$\text{MO-NTDA}^{\text{MIP}}: \min_{\substack{\mathbf{y}\in Y,\, \mathbf{Z},\, \mathbf{x}^b \in \aleph^*,\, \mathbf{u}^b, \\ \mathbf{x}^{P,b}, \tilde{\mathbf{x}}^{P,b}, \bar{\mathbf{x}}^{P,b}, \bar{\bar{\mathbf{x}}}^{P,b}, \\ \tilde{\mathbf{x}}^b, \bar{\mathbf{x}}^b, \bar{\bar{\mathbf{x}}}^b}} \sum_{b\in B} \lambda^b \sum_{n\in N} v_n Z_n^b,$$

$$\text{s.t.} \quad (31)\text{–}(44).$$

Multiobjective versions of the BD approach, the heuristic model and the robust model, MO-BD, MO-NTDA$^{\text{H}}$, and MO-RNTDA$^{\text{MIP}}$, respectively, can be derived analogously. MO-NTDA$^{\text{H}}$ solution cannot be guaranteed to be optimal when $|B| > 1$ or $|N| > 1$.

4.2. Example

We now demonstrate the value of the multiobjective model. To simplify the presentation, we focus on attacker "A" only, whose target nodes are "134" and "135," and the pseudo-optimal and cognizant behaviors (i.e., we set $\lambda^{\text{indifferent}} = 0$). We posit a defender with assets $(\tilde{R}, \bar{R}, \bar{\bar{R}}) = (4, 4, 0)$. We ignore the single attacker's value v_A because it does not affect the optimization, so we report Z_A^b for each behavior b, and redefine $V = \sum_{b\in B} \lambda^b Z_A^b$ as the objective function value.

The size of this multiobjective model is roughly $|B|$ times that of the single-objective (i.e., the number of behaviors simultaneously optimized). However, since the number of binary variables remains the same, solvability is not remarkably more challenging than the single-objective case.

First, we show the optimal interdiction plan when all the weight is given to the pseudo-optimal behavior (Fig. 4), that is, using $\boldsymbol{\lambda} = (\lambda^{\text{pseudo-optimal}}, \lambda^{\text{cognizant}}) = (1, 0)$. Corresponding attacker's routes for both behaviors are depicted as dotted lines. Clearly, under the expected pseudo-optimal behavior, the attacker avoids transparent interdictions but falls into all four traps along the way. He chooses target "134" but actually succeeds with probability $Z_A^{\text{pseudo-optimal}} = 27.8\%$. On the other hand, if his behavior were cognizant, he would be able to devise a route to target "135" that avoids all interdictions, and would succeed with probability $Z_A^{\text{cognizant}} = 68.8\%$.

We now examine the case where $\boldsymbol{\lambda} = (0.75, 0.25)$ (Fig. 5). One transparent interdiction moves from arc 111 and 112 to arc 123 and 124 (where there was a trap in the $\boldsymbol{\lambda} = (1, 0)$ case). Traps move closer to both targets. These changes can be explained because in the previous case a cognizant behavior was completely irrelevant, but as it gains weight the defender must make sure cognizant routes become "longer" (i.e., less likely to succeed with nominal probabilities). For this case, $Z_A^{\text{pseudo-optimal}} = 28.7\%$ and $Z_A^{\text{cognizant}} = 64.1\%$.

The solution we obtain for both cases $\boldsymbol{\lambda} = (0.5, 0.5)$ and $\boldsymbol{\lambda} = (0.25, 0.75)$ (Fig. 6) shows transparent and trap assets surrounding both targets and effectively blocking access to them. Both routes for the attacker are identical up to node "122," and both arrive at target "135." As opposed to previous cases, even the cognizant route is forced to traverse interdicted arcs (three traps in this case). For this case, $Z_A^{\text{pseudo-optimal}} = 38.4\%$ and $Z_A^{\text{cognizant}} = 53.3\%$.

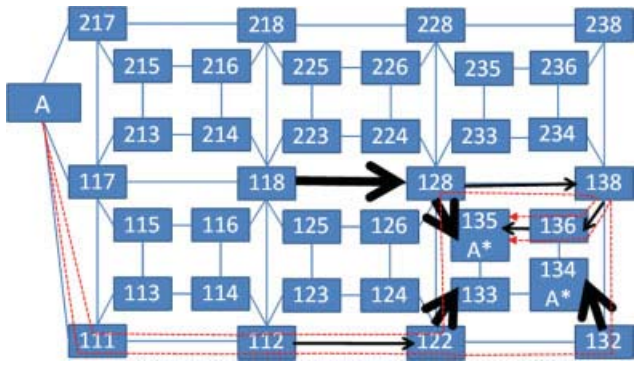

FIG. 6. Interdiction plan for $\boldsymbol{\lambda} = (0.5, 0.5)$ and $\boldsymbol{\lambda} = (0.25, 0.75)$, and attacker routes for both behaviors. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

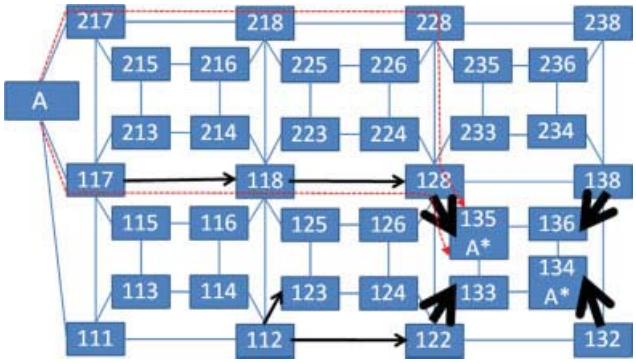

FIG. 7. Interdiction plan for $\boldsymbol{\lambda} = (0, 1)$, and attacker routes for both behaviors. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

In our final case (Fig. 7), only the cognizant behavior prevails, that is, $\boldsymbol{\lambda} = (0, 1)$. All four transparent assets "surround" the targets, forcing the cognizant route to pass through one of them. Here, $Z_A^{\text{pseudo-optimal}} = 46.9\%$ and $Z_A^{\text{cognizant}} = 50.9\%$.

None of the above cases matches the two-phase process of deploying transparent assets first, and then fixing those before deploying the traps for a pseudo-optimal behavior. That process would lead to the same four locations for transparent assets around the targets, and two of the traps on arcs 117 and 118 and 118–128. Since that suffices for the pseudo-optimal route (because such a route uses all interdicted arcs 117 and 118, 118–128, and 128–135), it leaves no guidance on how to locate the two remaining traps.

As in the single-objective case, the MO-NTDA$^{\text{MIP}}$ requires notable computational effort to be solved directly. Table 5 shows results regarding this issue, and compares MO-NTDA$^{\text{MIP}}$ with MO-BD and MO-NTDA$^{\text{H}}$ for weights in increments of 0.25. MO-BD can solve the $\lambda = (1, 0)$ and $\lambda = (0, 1)$ cases in a few seconds, and all others in less than 150s. MO-NTDA$^{\text{H}}$ obtains feasible solutions in less than ten seconds, which, given that $|N| = 1$, are guaranteed to be optimal when $|B| = 1$, that is, for $\lambda = (1, 0)$ and $\lambda = (0, 1)$. Objective values produced by MO-NTDA$^{\text{H}}$ are the same or better than MO-NTDA$^{\text{MIP}}$ after 100 seconds and, except in one case, after 1,000 seconds too. MO-NTDA$^{\text{MIP}}$ requires a long time to converge, as noted by large gaps after 10,000 seconds in some cases. This occurs even in the fully transparent case, $\lambda = (0, 1)$.

TABLE 5. Results for MO-NTDA$^{\text{MIP}}$ (solved directly), MO-BD and MO-NTDA$^{\text{H}}$ after several computational times.

	MO-NTDA$^{\text{MIP}}$			MO-BD	MO-NTDA$^{\text{H}}$
λ	100 s	1,000 s	10,000 s (gap %)	150 s	10 s
(1, 0)	27.8	27.8*		27.8*	27.8*
(0.75, 0.25)	41.2	37.9	37.5*	37.5*	37.9
(0.50, 0.50)	48.8	46.0	45.9 (34%)	45.9*	46.4
(0.25, 0.75)	50.5	49.6	49.6 (64%)	49.6*	49.6
(0, 1)	53.0	51.0	51.0 (80%)	51.0*	51.0*

Objective function values shown are the attacker's weighted evasion probabilities. Solutions marked with an asterisk (*) are guaranteed to be optimal.

To complete the multiobjective analysis, Figure 8 displays an approximation of the efficient frontier of the problem with solutions generated for weights in increments of 0.10, and the corresponding objective values.

5. CONCLUSIONS

This article extends the study of network interdiction with asymmetric information. Under the hypothesis that the defender has credible information about the attackers' behaviors, he can devise deceptive interdiction tactics which are more efficient than those resulting from completely cognizant (i.e., transparent) models. The multiobjective extension strikes an adequate balance among good choices for several behaviors simultaneously considered. It reproduces the cognizant results for a full risk-averse strategy, and otherwise realizes tradeoffs which cannot be obtained with a single-objective model.

The resulting problems are more challenging than those for standard DA models, especially due to the different expressions governing the defender's and attacker's objectives, and the necessary manipulations to convert the model into a MIP. BD helps solve many instances, but further

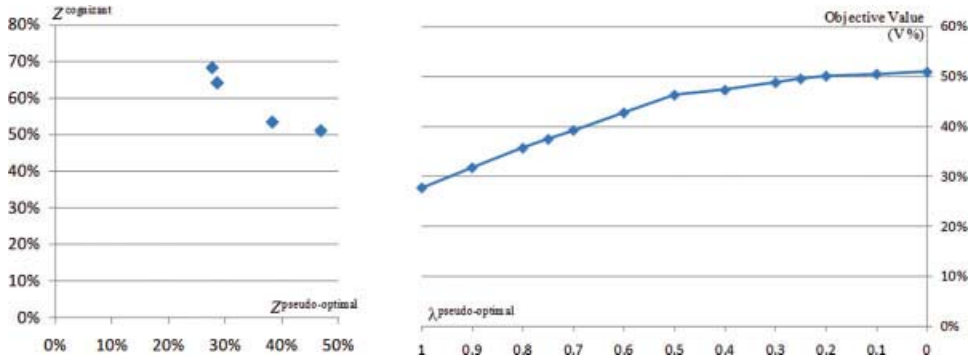

FIG. 8. Plot of approximated efficient frontier and associated objective values. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

research is needed to find stronger representations of the problem that improve convergence. A heuristic approximation has also helped us to obtain fast, empirically good, feasible solutions.

This article considers four classes of behaviors for the attacker, including that of what we call a skeptical attacker. Further research may also represent other behaviors, such as attackers who may realize a fraction of the deceptive assets, or use near-optimal routes, among others. In any case, the behavior's weight is an input to the problem. An interesting extension may consider an attacker whose class of behavior is also influenced by the defender's strategy. For example, if the optimal path (as perceived by the attacker) remains unchanged after the attacker observes the visible defenses, he may become skeptical and randomize his route. Incorporating this information in the defender's and attacker's strategies would require notably complicated data assumptions, and also appears beyond the capabilities of current optimization techniques, except perhaps by combining them with simulation. Alternatively, it would be useful to devise a stochastic optimization model that explicitly accounts for uncertainty in the attacker's behavior and/or estimates of perceived probabilities, producing robust solutions for a modest number of scenarios.

Acknowledgments

The author thanks the U.S. Center for Army Analysis for their partial support of this research. He also expresses great gratitude to two outstanding reviewers who inspired insightful and constructive ideas, as well as subtle thoughts, which became instrumental to the final form of this article.

REFERENCES

[1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network flows: Theory, algorithms and applications, Prentice Hall, Upper Saddle River, NJ, 1993.

[2] H. Bayrak and M. Bailey, Shortest path network interdiction with asymmetric information, Networks 52 (2008), 133–140.

[3] V. Bier, L. Cox, and N. Azaiez, "Why both game theory and reliability theory are important in defending infrastructure against intelligent attacks," Game theoretic risk analysis of security threats, V. Bier and N. Azaiez (Editors), Springer, New York, 2009, pp. 1–11.

[4] G. Brown, M. Carlyle, D. Diehl, J. Kline, and K. Wood, A two-sided optimization for theater ballistic missile defense, Oper Res 53 (2005), 745–763.

[5] G. Brown, W.M. Carlyle, J. Salmerón, and K. Wood, "Analyzing the vulnerability of critical infrastructure to attack, and planning defenses," Tutorials in operations research: Emerging theory, methods, and applications, H. Greenberg and J. Smith (Editors), Institute for Operations Research and Management Science, Hanover, MD, 2005, pp. 102–123.

[6] K. Cormican, D. Morton, and K. Wood, Stochastic network interdiction, Oper Res 46 (1998), 184–197.

[7] M. Ehrgott, "Multicriteria optimization," Lecture Notes in Economics and Mathematical Systems, 2nd edition, Springer, Berlin-Heidelberg, 2005.

[8] FICO (Fair Isaac Corporation), FICO^{TM} Xpress optimization suite 7. Available at http://www.fico.com/en/Products/DMTools/Pages/FICO-Xpress-Optimization-Suite.aspx, (accessed January 2010).

[9] B. Golden, A problem in network interdiction, Nav Res Logist Q 25 (1978), 711–713.

[10] IBM (International Business Machines), CPLEX algorithms. Available at http://www-2000.ibm.com/software/integration/optimization/cplex/algorithms, (accessed January 2010).

[11] M.X. Lugo, Deceptive tactics for protecting cities against vehicle borne improvised explosive devices, M. S. thesis in operations research, Naval Postgraduate School, Monterey, CA, 2008.

[12] D.P. Morton, F. Pan, and K.J. Saeger, Models for nuclear smuggling interdiction, IIE Trans Oper Eng 39 (2007), 3–14.

[13] A.L. Motto, J.M. Arroyo, and F.D. Galiana, A mixed-integer LP procedure for the analysis of electric grid security under terrorist threat, IEEE Trans Power Syst 20 (2005), 1357–1365.

[14] F. Pan, W.S. Charlton, and D.P. Morton, "A stochastic program for interdicting smuggled nuclear material," Network interdiction and stochastic integer programming, D.L. Woodruff (Editor), Kluwer, Norwell, MA, 2003, pp. 1–19.

[15] F. Pan and D.P. Morton, Minimizing a stochastic maximum-reliability path, Networks 52 (2008), 111–119.

[16] H. von Stackelberg, The theory of the market economy, William Hodge & Co., London, 1952.

[17] A. Washburn and K. Wood, Two-person zero-sum games for network interdiction, Oper Res 43 (1995), 243–251.

[18] R.K. Wood, Deterministic network interdiction, Math Comput Model 17 (1993), 1–18.