Stop Nuclear Smuggling Through Efficient Container Inspection[*]

Xinrun Wang[1], Qingyu Guo[2], Bo An[3]
[1,3]School of Computer Science and Engineering, Nanyang Technological University, Singapore
[2]Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly, NTU, Singapore
[1,2]{xwang033, qguo005}@e.ntu.edu.sg,[3]boan@ntu.edu.sg

ABSTRACT

Since 2003, the U.S. government has spent $850 million on the Megaport Initiative which aims at stopping the nuclear smuggling in international container shipping through advanced inspection facilities including Non-Intrusive Inspection (NII) and Mobile Radiation Detection and Identification System (MRDIS). Unfortunately, it remains a significant challenge to efficiently inspect more than 11.7 million containers imported to the U.S. due to the limited inspection resources. Moreover, existing work in container inspection neglects the sophisticated behavior of the smuggler who can surveil the inspector's strategy and decide the optimal (sequential) smuggling plan. This paper is the first to tackle this challenging container inspection problem, where a novel Container Inspection Model (CIM) is proposed, which models the interaction between the inspector and the smuggler as a leader-follower Stackelberg game and formulates the smuggler's sequential decision behavior as a Markov Decision Process (MDP). The special structure of the CIM results in a non-convex optimization problem, which cannot be addressed by existing approaches. We make several key contributions including: i) a linear relaxation approximation with guarantee of solution quality which reformulates the model as a bilinear optimization problem, ii) an algorithm inspired by the Multipleparametric Disaggregation Technique (MDT) to solve the reformulated bilinear optimization, and iii) a novel iterative algorithm to further improve the scalability. Extensive experimental evaluation shows that our approach can scale up to realistic-sized problems with robust enough solutions outperforming heuristic baselines significantly.

CCS Concepts

•**Computing methodologies** → **Multi-agent systems;**

Keywords

Game Theory; Container Inspection; Nuclear Smuggling

1. INTRODUCTION

Maritime container shipping has been a critical measure for terrorists and smugglers to transport illegal goods including weapons of mass destruction (WMD) and even nuclear materials. To prevent nuclear smuggling activities, various initiatives are deployed by governments at ports, such as Container Security Initiative (CSI) [9], Megaports initiative [21] and Secure Freight Initiative (SFT) [11] of the U.S. government, which inspect containers at foreign and domestic ports with advanced inspection facilities and security officers. However, since there are around 17.5 million containers imported into the U.S. per year [19], only a small percentage (less than 20%) can be inspected thoroughly by security agencies with sophisticated facilities such as radiation and spectroscopic portal monitors, while most containers are under non-intrusive inspection[1] which cannot ensure the detection of nuclear material whose amount is under some threshold, especially when they are shielded. Therefore, it is extremely critical to decide how to allocate the limited inspection resources over shipping lines and ports.

However, the inspection strategy at ports can be learnt by the sophisticated smuggler through extensive surveillance and the shipping lines with the minimal risk will be chosen. Therefore, how to optimally deploy the limited inspection resources becomes an extremely challenging task for security agencies for several reasons: i) the smuggler can choose multiple containers to transport illegal items through several shipping lines, which causes an exponentially large number of possible actions; ii) it is not necessary for the smuggler to ship all the illegal containers at the same time due to the high risk of being detected, rather, the long-term plan and sequential decisions are preferred, making the resulting decision process of the smuggler even more difficult to infer; iii) the inspection may operate under different modes to quickly respond to an emergency when illegal containers are interdicted; and iv) the security resource allocation has complicated impact on the smuggler's decision-making process and the large-scale non-convex optimization, a notorious class of hard problems, is unavoidable to design the optimal inspection strategy.

Container inspection has drawn the attention of researchers from several fields. Some research applied the optimization models to design more efficient inspection apparatus to provide more reliable inspection results [7, 24]. Some proposed advanced inspection protocols to maximally

[*] This work will also be published in AAMAS 2017.

[1] Non-intrusive inspection uses X-rays or gamma rays to scan a container and creates images of the container's contents without opening it to help the inspector to identify anomalies among other goods.

Appears in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

utilize the limited resources at ports to speed up the inspection process [4] while others applied game theoretical methods to help the inspector to select which containers to be inspected [3, 6, 15]. However, previous works ignore the smuggler's strategic sequential decision-making process and the allocation of inspection resources may be far from optimal.

Different from previous works, we tackle the challenge of optimally preventing sophisticated smugglers in a novel way and make several key contributions in this paper. First, we propose a realistic container inspection model (CIM) where the inspector randomly chooses containers from different shipping lines to inspect while the smuggler makes a long-term plan with the knowledge of the inspection strategy, and the corresponding decision-making problem of the smuggler is modeled as a Markov Decision Process (MDP). Second, we formulate the inspector's optimization problem as a nonconvex program with an exponential number of constraints. Third, to address the nonconvexity and scalability issues, several novel approaches are proposed, including a linear relaxation approximation with guarantee of solution quality, an algorithm inspired by the Multipleparametric Disaggregation Technique (MDT) [26] to address the bilinear terms, and a novel heuristic iterative method to deal with the exponential number of constraints. Finally, we conduct extensive experimental evaluations on both simulated and real-world shipping networks and show that our approaches can solve realistic-sized problem instances with good enough and robust solutions.

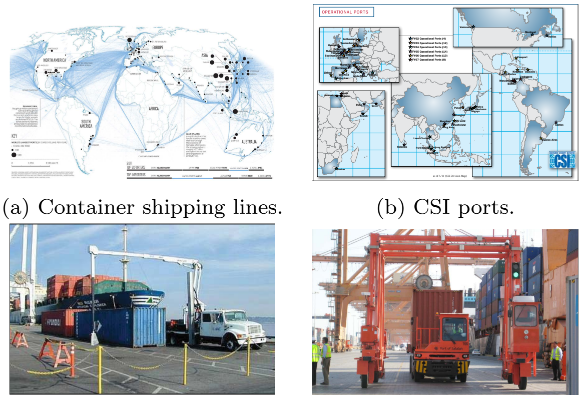
2. RELATED WORK

Container inspection has been investigated in many aspects [3, 4, 6, 7, 15]. Bakshi *et al.* [4] estimated the operational impact of container inspection at the ports where the inspector inspects all containers at ports. Bakir [3] proposed a Stackelberg game model for resource allocation on one or multiple but independent routes in cargo container security. Haphuriwat *et al.* [6, 15] identified the number of containers to be inspected where the inspector inspects containers uniformly. However, in the real-world, only a small percentage of containers is inspected at ports and both the inspector and the smuggler can behave strategically to maximize their utility.

There is also abundant literature in the *network interdiction problem* [10], where we are given a weighted, directed or undirected graph, and various objectives are studied, including minimizing the network flow, maximizing the shortest path and increasing detection probability via deletion of edges/nodes or decreasing edge capacity [1, 5, 16, 30]. Other works study *stochastic network interdiction* where the interdiction action is successful with some known probability p [23]. Recently, the randomized resource allocation to interdict the escape path or illegal network flow has drawn extensive attention of *security game* researchers [14, 17, 27]. Unfortunately, none of them tackles the sophisticated smuggling activities and sequential decisions, and the smuggling activity is assumed to be one-shot.

Several recent works in security games study the long-term planning and model the sequential decision-making of attackers as an MDP [2, 31]. An *et al.* [2] studied the adversary's sequential observations of the realization of the defender's random allocation before taking the attacking action where the number of states of the smuggler's MDP is linear to the number of defender's pure strategies and the time horizon; Zhao *et al.* [31] computed the optimal thresholds of different users for the email filtering system to prevent the long-term sequential cyber attacks where the state of the smuggler's MDP is linear to the number of targets. Both problems are categorized in target protection scenarios where the smuggler's actions in states is simple. Therefore, the MDPs in both works can be solved efficiently by the dynamic programming algorithm. While in our network security domain, both the state space and action space of the smuggler's MDP are exponentially large, which requires us to come up with novel and efficient algorithms to address the scalability issues.

3. MOTIVATION



(a) Container shipping lines.

(b) CSI ports.

(c) Gamma-ray for NII.

(d) MRDIS.

Figure 1: Nuclear smuggling and its prevention.

In this section, we use the nuclear smuggling as a motivating example, while our model can be applied to a variety of illegal container smuggling scenarios. Figure 1a shows the global shipping network where more than 34 million containers are transported all around the world per year [19]. To prevent nuclear smuggling through maritime container shipping, The U.S. government has launched several international initiatives, such as Container Security Initiative (CSI), Megaport Initiative (MI) and Security Fright Initiative (SFI) to inspect containers at ports. Figure 1b shows the worldwide 58 CSI ports, which cover more than 80 percent of containers imported into the U.S. [9]. As there are enormous containers arriving at ports per day, most containers are under non-intrusive inspection (NII) using the system shown in Figure 1c. However, if the amount of nuclear material is small[2] or the material is shielded by other goods, NII cannot provide reliable inspections.

Therefore, the U.S. government installed Mobile Radiation Detection and Identification System (MRDIS) as shown in Figure 1d, which is more sensitive and reliable. However, as using MRDIS is time consuming, only less than 20% containers are inspected by MRDIS [21]. Besides, the government has developed emergency plans to quickly respond

[2] Different from other illegal goods such as guns or cigarettes, small amount of nuclear material is of concern. As little as 25 kilograms of highly enriched uranium or 8 kilograms of plutonium could be used to build a nuclear weapon known as an improvised nuclear device [13].

to the emergency when illegal containers are interdicted by adding more manpower to execute the inspection. For example, seven different government agencies will coordinate to respond to a radioactive emergency in Jamaica in its emergency plan [12].

On the other hand, in order to avoid the risk of being interdicted by NII, the smuggler would divide his illegal material into several small units (e.g., 1 kg) and shield them with other goods [22]. Furthermore, the sophisticated smuggler may choose multiple shipping lines after making enough surveillance to the inspector's strategy and stop to smuggle when the government's emergency plan is triggered. Thus, a long-term sequential plan is preferred by the smuggler, which makes it even more difficult for the inspector to detect the smuggling activities. Therefore, optimally allocating the limited reliable inspection resources such as MRDIS and responding to the emergency is an extremely challenging task. In this paper, we aim to compute the optimal allocation of inspector's resources to combat the nuclear smuggling after the inspector knows a number of containers will be shipped to some ports.

4. MODEL

We now illustrate our Container Inspection Model (CIM) which models the interaction between the inspector and smuggler as a Stackelberg game where the inspector moves first and decides her allocation of inspection resources at various ports and shipping lines, while the smuggler chooses the optimal trafficking plan with knowledge of the inspector's inspection strategy with extensive surveillance.

CIM models the ocean shipping network as a tuple $\mathcal{N} = \langle \mathcal{L}, \mathcal{P} \rangle$ where $\mathcal{L}$ is the set of shipping lines and $\mathcal{P}$ is the set of ports. Let $\alpha : \mathcal{L} \times \mathcal{P} \to \{0, 1\}$ be the indicator of ports that a shipping line passes through, where $\alpha_{lp} = 1$ if port $p \in \mathcal{P}$ is on shipping line $l \in \mathcal{L}$ and $\alpha_{lp} = 0$ otherwise. The containers flow on the network is represented by $\mathbf{f} = \langle f_l \rangle$ where f_l denotes the number of containers shipping through line $l \in \mathcal{L}$ within certain time period. As we consider the sequential actions of both the smuggler and the inspector, we set τ as the unit time period of shipping an illegal container[3] and the smuggler makes decisions at times $\{0, \tau, ..., t \cdot \tau, ...\}$ where t is the *time step*.

The smuggler with m ($m \leq |\mathcal{L}|$) illegal containers can strategically ship them over different shipping lines where m can be estimated by the amount of nuclear material lost by the governments and institutions. Analogous to existing literatures in security games [14, 17, 27, 29], we focus on the game where the inspector is minimizing the smuggler's utility. Assume that the payoffs are zero for both players if an illegal container shipped through l is interdicted. We denote by u_l^a and u_l^d the payoffs for the smuggler and the inspector respectively when an illegal container is successfully shipped through shipping line l and $u_l^d = -u_l^a$. W.l.o.g., assume $u_l^a > 0$ for all shipping lines[4].

Inspector's Strategy: The inspector with limited inspection resources decides the proportion of containers inspected of each shipping line on each port. In realistic shipping scenarios, extra inspection power can be implemented when illegal containers are interdicted [12]. However, the emergency inspection cannot last long due to the high cost of inspection facilities and officers. Therefore, the inspector's strategy consists of two modes: *normal mode* without emergency inspection and *emergency mode* with extra resources. Let $\Theta = \{normal, emergency\}$ be the two modes and the transition of different modes is as follows: if an illegal container is interdicted at normal mode at time step t, the emergency mode is triggered at $t + 1$ and will last for one time step as long as no illegal container is interdicted. Otherwise, the emergency mode will last for another time step. Figure 2 illustrates the transitions between modes.

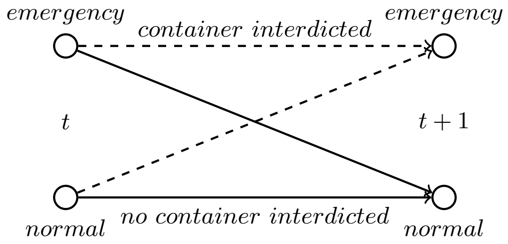


Figure 2: Mode transition.

Let n_p denote the inspection capability of port $p \in \mathcal{P}$ which represents the maximum number of containers inspected at port p within τ. The capabilities of ports are fixed in both modes, as the inspector will maximally utilize the existing equipment to inspect containers. We denote by $C : \Theta \times \mathcal{P} \times \mathcal{L} \to [0, 1]$ the allocation of security resources at ports such that C_{pl}^{θ} represents the proportion of containers on shipping line l inspected at port p at mode θ, and $C_{pl}^{\theta} = 0$ for the pair (p, l) with $\alpha_{lp} = 0$. A valid allocation C satisfies the capability constraint at each port:

$$\sum_{l \in \mathcal{L}} C_{pl}^{\theta} f_l \leq n_p \quad \forall p \in \mathcal{P}, \theta \in \Theta. \tag{1}$$

Let $\mathbf{c} : \mathcal{L} \to [0, 1]$ represent the emergency inspection strategy such that c_l denotes the proportion of containers on shipping line l being inspected by the emergency inspection resources. Let n^e denote the emergency inspection capability, i.e., the maximal number of containers inspected by emergency resources within τ. Emergency resource allocation $\mathbf{c}$ satisfies:

$$\sum_{l \in \mathcal{L}} c_l f_l \leq n^e \tag{2}$$

Overall, we denote by $X : \Theta \times \mathcal{L} \to [0, 1]$ the inspector's strategy:

$$X_l^{\theta} = \sum_{p \in \mathcal{P}} C_{pl}^{\theta} + c_l \mathbb{I}\{\theta = emergency\} \quad \forall l \in \mathcal{L}, \tag{3}$$

where X_l^{θ} represents the proportion of containers on shipping line l inspected at mode θ and $\mathbb{I}\{\theta = emergency\}$ is the indicator function indicating whether θ is emergency[5].

[3] The time to smuggle a container is different for different shipping lines varying from a week to more than a month. For easy of analysis, we take the average time as the value of τ and the assumption can be easily relaxed.

[4] The smuggler will try to smuggle the nuclear material in a container as much as possible without triggering the alarm of NII inspection system. Therefore, we assume the containers are identical and the payoff of a container could depend on the price in the black market where the smuggler tries to sell the nuclear material.

[5] As the inspection resources are limited and the inspection result is reliable, we assume the container is inspected at most once through the shipping line. Even if we add some failure probability of the inspection, the inspector will prefer to inspect the containers not inspected before.

Smuggler's strategy: The smuggler may not ship all illegal containers immediately to avoid the risk of being interdicted in emergency mode, instead, sequential plans are preferred. On the other hand, the smuggler is willing to ship the illegal containers as soon as possible with the fear of being detected by the local security agency. Therefore, we denote by γ the *discount factor* such that the payoff of successful shipped container at time step t is discounted by γ^t. We model the long-term planning process of smugglers as a Markov Decision Process (MDP), which is represented as a tuple $(\mathcal{S}, \mathcal{A}, T, R, \pi)$. $\mathcal{S}$ is the state space of the MDP and each state $s \in \mathcal{S}$ is denoted by $s = \langle t, \theta, \tilde{m} \rangle$ where t is the time step, θ is the current mode of inspection strategy at state s and $\tilde{m}$ is the number of illegal containers waiting to be shipped. The process is initiated at $s_0 = \langle 0, normal, m \rangle$ and terminated at $s = \langle *, *, 0 \rangle$ where all illegal containers are shipped, which is called a terminal state. Let $\mathcal{S}^T \subset \mathcal{S}$ denote the set of all terminal states. $\mathcal{A}$ represents the smuggler's action space. In particular, an action $\mathbf{a} = \langle a_l \rangle$ denotes the allocation of illegal containers over shipping lines such that $a_l \in \{0, 1\}$ represents the number of containers shipped on l[6]. Assume that the probabilities of being detected are independent among different shipping lines. We denote by $\mathcal{A}_s \subseteq \mathcal{A}$ the set of all possible actions at state $s = \langle t, \theta, \tilde{m} \rangle$:

$$\mathcal{A}_s = \{\mathbf{a} \in \{0,1\}^{|\mathcal{L}|} : \textstyle\sum_{l \in \mathcal{L}} a_l \leq \tilde{m}\} \quad \forall s \in \mathcal{S}.$$

$\mathcal{A} = \mathcal{A}_{s_0}$ and we denote $\mathbf{a}^s$ when needed.

There are two reachable states by taking action $\mathbf{a} \in \mathcal{A}_s$ at state $s = \langle t, \theta, \tilde{m} \rangle$: $s' = \langle t+1, normal, \tilde{m}' \rangle$ and $s' = \langle t+1, emergency, \tilde{m}' \rangle$ where $\tilde{m}' = \tilde{m} - \sum_{l \in \mathcal{L}} a_l$. Given the inspector's strategy X, $T(s, \mathbf{a}, s')$ represents the transition probability of reaching state $s' = \langle t+1, \theta', \tilde{m}' \rangle$ from state $s = \langle t, \theta, \tilde{m} \rangle$ by taking action $\mathbf{a} \in \mathcal{A}_s$:

$$T(s, \mathbf{a}, s') = \left\{ \begin{array}{ll} \Phi(\theta, X, \mathbf{a}) & \theta' = normal; \\ 1 - \Phi(\theta, X, \mathbf{a}) & \theta' = emergency. \end{array} \right. \tag{4}$$

where $\Phi(\theta, X, \mathbf{a}) = \prod_{l \in \mathcal{L}} (1 - X_l^\theta)^{a_l}$ represents the probability that smuggling action $\mathbf{a}$ is successful and no shipped illegal container is inspected when inspection strategy is at mode θ. $R(s, \mathbf{a}, s')$ denotes the reward of taking action $\mathbf{a}$ at state s and reaching state s', taking into account the discount factor:

$$R(s, \mathbf{a}, s') = \left\{ \begin{array}{ll} \gamma^t \sum_{l \in \mathcal{L}} a_l u_l^a & \theta' = normal; \\ \gamma^t \frac{\sum_{l \in \mathcal{L}} (1 - X_l^\theta - \Phi(\theta, X, \mathbf{a})) a_l u_l^a}{1 - \Phi(\theta, X, \mathbf{a})} & \theta' = emergency. \end{array} \right. \tag{5}$$

The reward for $\theta' = emergency$ is deduced by the fact that the expected reward of taking action $\mathbf{a}$ at state s equals to $\gamma^t \sum_{l \in \mathcal{L}} (1 - X_l^\theta) a_l u_l^a$.

The smuggler's policy is denoted by $\pi : \mathcal{S} \to \mathcal{A}$ such that $\pi(s)$ returns the action $\mathbf{a} \in \mathcal{A}_s$ to play at state s which is not the terminal state. We denote by the value function $V^\pi : \mathcal{S} \to \mathbb{R}$ such that $V^\pi(s)$ represents the expected utility of the smuggler following policy π when current state is s.

Utility & Equilibrium: Given the profile of both players' strategies $\langle X, \pi \rangle$, the expected utility $U_a(X, \pi)$ of the smuggler is defined as: $U_a(X, \pi) = V^\pi(s_0)$; Given the zero-sum assumption, the expected utility of the inspector is $U_d(X, \pi) = -U_a(X, \pi)$. We adopt the Stackelberg equilibrium as our solution concept [18, 28], which is the strategy profile $\langle \mathbf{x}^*, \pi^* \rangle$ which satisfies:

1. $U_a(X^*, \pi^*) \geq U_a(X^*, \pi)$ for any other policy π,
2. $U_d(X^*, \pi^*) \geq U_d(X, \pi)$ for any other strategy X where π is the best response policy against X.

5. SOLUTION APPROACH

In this section, we illustrate our solution approach to solve the equilibrium $\langle X^*, \pi^* \rangle$ efficiently. First, a large-scale nonconvex program is proposed based on the linear program of solving the smuggler's MDP, where the number of constraints grows exponentially with respect to the size of game instance. To address the nonconvexity of the formulation, a linear relaxation approximation of transition probability T is adopted and the Multiparametric Disaggregation Technique with binary base is developed to relax the remaining bilinear terms and the linear formulation approximation is obtained, with theoretical guarantees of approximation quality. Finally, a novel constraint generation approach is provided to improve the scalability of the formulation.

5.1 LP for smuggler's MDP

Given the inspector's strategy X, the smuggler's MDP can be solved by the following linear program which enumerates all states and actions and adds them as constraints [25]:

$$\min_V V(s_0) \tag{6a}$$

$$\text{s.t. } V(s) \geq \textstyle\sum_{s' \in \mathcal{S}} T(s, \mathbf{a}, s')[R(s, \mathbf{a}, s') + V(s')] \quad \forall \mathbf{a} \in \mathcal{A}_s, \forall s \in \mathcal{S} \setminus \mathcal{S}^T \tag{6b}$$

$$V(s) = 0, \quad \forall s \in \mathcal{S}^T \tag{6c}$$

Let V^* denote the optimal solution of formulation (6). The smuggler's optimal policy π^* can be obtained by:

$$\pi^*(s) = \arg\max\nolimits_{\mathbf{a} \in \mathcal{A}_s} Q(s, \mathbf{a}), \quad \forall s \in \mathcal{S} \setminus \mathcal{S}^T$$

where $Q(s, \mathbf{a}) = \sum_{s' \in \mathcal{S}} T(s, \mathbf{a}, s')[R(s, \mathbf{a}, s') + V^*(s')]$. We say $\mathbf{a} \in \pi^*$ if π^* selects the action $\mathbf{a}$.

5.2 Nonconvex program for optimal inspection

Recall that the inspector's strategy X consists of the allocation C of resources at each port and the deployment of emergency inspection $\mathbf{c}$, as defined in Eq.(3). We propose a non-convex optimization formulation for solving the optimal inspection strategy as follows:

$$\min_{C, \mathbf{c}, X, V} V(s_0) \tag{7a}$$

$$\text{s.t. Eqs.(1)–(3)} \tag{7b}$$

$$\text{Eqs.(6b)–(6c)} \tag{7c}$$

Constraint (7b) ensures that the inspector's strategy X is valid. Constraint (7c) restricts value function V to be optimal given the minimization objective of $V(s_0)$.

To make the formulation (7) practical to solve, we first need to know the horizon of the smuggler's MDP in advance. Lemma 1 implies that in the optimal policy π^*, the smuggler's long-term and sequential planning terminates in finite number of steps.

Lemma 1. *The smuggler will smuggle all containers no more than $2m$ time steps.*

[6] Normally, the goods are packed by the shipping companies at container fright stations. To avoid the occasion where two or more units of nuclear material are packed into a container, which can be easily interdicted by the NII inspection, the smuggler will only smuggle at most an unit through a shipping line at a time step.

Proof sketch. The proof is based on the fact that the smuggler will smuggle at least a container at each time step under normal state. If not, the inspection will stay in normal state and the smuggler's utility will decrease because of the discount factor. Therefore, the maximum number of time steps for the smuggler to smuggle all containers is $2m$ where the smuggler smuggles at least a container every 2 time steps. □

Unfortunately, even if we restrict the smuggler's MDP horizon within $2m$ time steps, the nonconvexity of formulation (7) makes it impossible to solve for the optimal solution for large-scale game instances, which originates from two terms: $\Phi(\theta, X, \mathbf{a})$ in transition probability T and reward function R and the term $T(s, \mathbf{a}, s') \cdot V(s')$ in Eq.(6b). To make the formulation scalable, we first relax the smuggler's MDP which is the same as original MDP expect that $\Phi(\theta, X, \mathbf{a})$ is replaced with its first order Taylor expansion in transition probability (4) and reward function (5), and the program (6) for solving such a relaxed MDP becomes a linear program when X is fixed.

5.3 Linear relaxation of Φ

Notice that with huge number of containers shipping per day, the proportion of inspected containers is very small. Therefore, the first order Taylor expansion of $\Phi(\theta, X, \mathbf{a})$ gives a good approximation which is linear to X:

$$\Phi(\theta, X, \mathbf{a}) = \prod\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l} \approx 1 - \sum\nolimits_{l\in\mathcal{L}} a_l \cdot X_l^\theta$$

Replacing $\Phi(\theta, X, \mathbf{a})$ in transition probability (4) and reward function (5) with above linear approximation, the attacker's MDP is relaxed, and the formulation (7) for computing the equilibrium becomes:

$$\min_{C,\mathbf{c},X,V} \; V(s_0) \tag{8a}$$

$$\text{s.t. Eqs.(1)–(3)} \tag{8b}$$

$$\begin{aligned} V(s) \geq \gamma^t \sum\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)a_l u_l^a + (1 - \sum\nolimits_{l\in\mathcal{L}} a_l X_l^\theta)V(s^n) \\ + \sum\nolimits_{l\in\mathcal{L}} a_l X_l^\theta V(s^e) \quad \forall \mathbf{a} \in \mathcal{A}_s, \forall s \in \mathcal{S}\setminus\mathcal{S}^T \end{aligned} \tag{8c}$$

$$V(s) = 0, \quad \forall s \in \mathcal{S}^T \tag{8d}$$

In Eq.(8c), state $s = \langle t, \theta, \tilde{m}\rangle$. s^n and s^e are the two states reachable from taking action $\mathbf{a}$ in s, where $s^n = \langle t+1, \mathit{normal}, \tilde{m}'\rangle$ and $s^e = \langle t+1, \mathit{emergency}, \tilde{m}'\rangle$, and $\tilde{m}' = \tilde{m} - \sum_{l\in\mathcal{L}} a_l$. Theorem 2 provides a bound of the utility computed by the approximation formulation (8) compared with the optimal utility.

Theorem 2. *Given the inspector's strategy X, let π^* be the optimal attacker policy and π be the smuggler policy which is optimal for the relaxed MDP. Let V^* and V be the value functions corresponding to π^* and π in original MDP and relaxed MDP respectively. The following inequality holds:*

$$V^*(s_0) - V(s_0) \leq m^2 \cdot \frac{\kappa^2 \cdot \gamma}{1-\gamma} \cdot \bar{V}$$

where $\bar{V} = m \cdot \max_{l\in\mathcal{L}} u_l^a$ and $\kappa = \max_{l\in\mathcal{L},\theta\in\Theta} X_l^\theta$.

Proof. According to the Taylor theorem, given the inspector's strategy $X_l^\theta \in [0, 1]$, we have:

$$\begin{aligned} &\prod\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l} - (1 - \sum\nolimits_{l\in\mathcal{L}} a_l \cdot X_l^\theta) \\ &\approx \sum_{l,l'\in\mathcal{L}, l\neq l'} a_l a_{l'} \cdot X_l^\theta X_{l'}^\theta \\ &\leq m^2 \cdot \kappa^2 \end{aligned}$$

where three and higher order terms are neglected. Given the inspector's strategy, there are two cases **C1** and **C2** to consider.

C1: Assume that the smuggler has the same optimal policy π in regardless of whether transition probability Φ is linear relaxed. Further, we assume that the policy terminates at T time step, i.e., the smuggler transports all containers in T time steps. It is easy to verify that $V^*(s) \geq V(s)$ when the smuggler follows the same policy from current state s, as in original MDP, the probability to transit to the state of normal mode is higher than that in the relaxed MDP, while the reward of transition to the state of normal mode and the expected reward of taking an action remain the same in both MDPs. We use induction method to iteratively bound the value $V^*(s) - V(s)$ at each time step from $T - 1$ and finally give the bound of $V^*(s_0) - V'(s_0)$.

(1) It is obvious that $V^*(s) = V(s)$ where $s = \langle T-1, \theta, \tilde{m}\rangle$.
(2) We assume $0 \leq V^*(s) - V(s) \leq \varepsilon$ where $s = \langle t+1, \theta, \tilde{m}\rangle$. then we can write down the explicit expressions:

$$\begin{aligned} V^*(s) =& \prod\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l} \cdot V^*(s^n) \\ &+(1 - \prod\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l}) \cdot V^*(s^e) + \gamma^t \sum\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)a_l u_l^a \\ V(s) =& (1 - \sum\nolimits_{l\in\mathcal{L}} a_l \cdot X_l^\theta) \cdot V(s^n) \\ &+(\sum\nolimits_{l\in\mathcal{L}} a_l \cdot X_l^\theta) \cdot V(s^e) + \gamma^t \sum\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)a_l u_l^a \end{aligned}$$

where $s = \langle t, \theta, \tilde{m}\rangle$, $s^n = \langle t+1, \mathit{normal}, \tilde{m}'\rangle$, $s^e = \langle t+1, \mathit{emergency}, \tilde{m}'\rangle$ and $\tilde{m}' = \tilde{m} - \sum_{l\in\mathcal{L}} a_l$. Then,

$$\begin{aligned} V^*(s) - V(s) =& (\prod\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l}) \cdot (V^*(s^n) - V(s^n)) \\ &+ (1 - \prod\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l})) \cdot (V^*(s^e) - V(s^e))) \\ &+ \delta^\theta(V(s^n) - V(s^e)) \\ \leq& \varepsilon + \delta^\theta(\gamma^{t+1}\bar{V}) \\ \leq& \varepsilon + m^2 \cdot \kappa^2 \cdot \gamma^{t+1}\bar{V} \end{aligned}$$

where $\delta^\theta = \prod_{l\in\mathcal{L}}(1 - X_l^\theta)^{a_l} - (1 - \sum_{l\in\mathcal{L}} a_l \cdot X_l^\theta)$.

Then we have:

$$\begin{aligned} &V^*(s_0) - V'(s_0) \\ &\leq m^2 \cdot \kappa^2 \sum\nolimits_{t=1}^{T-1} \gamma^t \bar{V} \\ &\leq m^2 \cdot \frac{\kappa^2 \cdot \gamma}{1-\gamma}\bar{V} \end{aligned} \tag{9}$$

C2: If the smuggler's optimal policy π in the relaxed MDP differs from the optimal policy π^* in original MDP, which implies $V^*(\pi^*) - V(\pi) \leq V^*(\pi^*) - V(\pi^*)$ where $V^*(\pi)$ and $V(\pi)$ represent the values at initial state following policy π in the original MDP and relaxed MDP respectively. Thus, Eq.(9) also holds, which concludes the proof. □

Although formulation (8) is much simpler than the original formulation (7), it is still a non-convex program due to the bilinear terms $X_l^\theta \cdot V(s^n)$ and $X_l^\theta \cdot V(s^e)$ in Eq.(8d). To further linearize the formulation, we adopt the Multiparametric Disaggregation Technique (MDT).

5.4 Linearization based on MDT

The basic idea of MDT is to replace $X_l^\theta \cdot V(s)$ with an auxiliary variable $w_l^\theta(s)$ and add several linear constraints involving $w_l^\theta(s)$, X_l^θ and $V(s)$ to approximate the equality relationship $w_l^\theta(s) = X_l^\theta \cdot V(s)$. In particular, since $X_l^\theta \in [0,1]$[7], we approximate X_l^θ with a binary number with Z digits located at powers $\{-Z, ..., -1\}$. We define one binary variable λ_{lz}^θ for each power $-z$ and X_l^θ can be represented as follows:

$$X_l^Z = \sum\nolimits_{z=1}^{Z} 2^{-z}\lambda_{lz}^\theta + \tilde{X}_l^\theta, \tag{10}$$

where $\tilde{X}_l^\theta \in [0, 2^{-Z}]$ is the slack variable. Since $w_l^\theta(s) = X_l^\theta \cdot V(s)$, we have:

$$w_l^\theta(s) = \sum\nolimits_{z=1}^{Z} 2^{-z}\eta_{lz}^\theta(s) + \tilde{w}_l^\theta(s), \tag{11}$$

where $\eta_{lz}^\theta(s) = \lambda_{lz}^\theta \cdot V(s)$, which can be ensured by the following constraints with a large enough constant M:

$$\begin{aligned} 0 \leq & \eta_{lz}^\theta(s) \leq V(s) \\ V(s) - (1 - \lambda_{lh}^\theta)M \leq & \eta_{lz}^\theta(s) \leq \lambda_{lz}^\theta M \end{aligned} \tag{12}$$

$\tilde{w}_l^\theta(s) = \tilde{X}_l^\theta \cdot V(s)$ which, however, cannot be represented exactly with finite linear constraints. Therefore, the McCormick Envelope [20] is adopted to approximate the relationship $\tilde{w}_l^\theta(s) = \tilde{X}_l^\theta \cdot V(s)$ with following linear constraints:

$$\begin{aligned} 0 \leq & \tilde{w}_l^\theta(s) \leq 2^{-Z}V(s) \\ 2^{-Z}V(s) + \bar{V}(\tilde{X}_l^\theta - 2^{-Z}) \leq & \tilde{w}_l^\theta(s) \leq \bar{V}\tilde{X}_l^\theta \end{aligned} \tag{13}$$

where $\bar{V}$ is an upper bound of $V(s)$ which can be roughly estimated as the maximal possible utility $m \cdot \max_{l\in\mathcal{L}} u_l^a$. Overall, MDT applies the linear system (10)–(13) to approximate all the bilinear terms $X_l^\theta \cdot V(s)$ with $w_l^\theta(s)$ in formulation (8), and the resulting MILP is as follows:

$$\min_{\substack{C,\, \mathbf{c},\, X,\, V \\ \lambda,\, \eta,\, \tilde{X},\, w}} \quad V(s_0) \tag{14a}$$

$$\text{s.t. Eqs.(1)–(3)} \tag{14b}$$

$$\begin{aligned} V(s) \geq \gamma^t & \sum\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)a_l u_l^a + V(s^n) \\ & - \sum\nolimits_{l\in\mathcal{L}} a_l w_l^\theta(s^n) + \sum_{l\in\mathcal{L}} a_l w_l^\theta(s^e) \\ & \forall \mathbf{a} \in \mathcal{A}_s, \forall s \in \mathcal{S} \setminus \mathcal{S}^T \end{aligned} \tag{14c}$$

$$V(s) = 0, \quad \forall s \in \mathcal{S}^T \tag{14d}$$

$$\text{Eqs.(10)–(13).} \tag{14e}$$

Notice that since slack variable $\tilde{w}_l^\theta(s)$ is allowed to take any value in $[0, 2^{-Z}V(s)]$ in the worst case, $w_l^\theta(s)$ is not strictly required to be equal to $X_l^\theta \cdot V(s)$, rather, it is restricted to take values close to $X_l^\theta \cdot V(s)$:

$$|w_l^\theta(s) - X_l^\theta \cdot \tilde{V}(s)| \leq \frac{\tilde{V}(s)}{2^Z}. \tag{15}$$

Hence, given the minimization objective of program 8, the solution $\tilde{V}$ returned by MILP 14 serves as a lower bound of the optimal value function V of program 8. Furthermore, with a larger number of digits Z, $w_l^\theta(s)$ takes values closer to $X_l^\theta \cdot V(s)$ and the lower bound $\tilde{V}$ will approach the optimal value function V. On the other hand, for each valid inspector's strategy X, Program (6) computes an upper bound $\hat{V}$ of V. Thus, we propose Algorithm 1, **B**inary-**B**ased MDT for **C**IM (BBC), which iteratively increases the number of digits Z in MDT until the upper bound and lower bound are close enough. Theorem 3 analyzes the relationship between the gap $\hat{V}(s_0) - \tilde{V}(s_0)$ and the number of digits Z.

THEOREM 3. *The bounds obtained by Algorithm 1 satisfy the following inequality:*

$$V(s_0) - \tilde{V}(s_0) \leq \frac{|\mathcal{L}|}{2^{Z-1}} \cdot \frac{\gamma(1-(2\gamma)^{2m})}{1-2\gamma}\bar{V}$$

where $\bar{V} = m \cdot \max_{l\in\mathcal{L}} u_l^a$.

PROOF. The proof utilizes the bound between $w_l^\theta(s)$ and $X_l^\theta \cdot \tilde{V}(s)$ shown in (15).

It is obvious that $V(s) = \tilde{V}(s)$ where $s \in \mathcal{S}^T$. For the non-terminal states, we assume $0 < V(s) - \tilde{V}(s) \leq \varepsilon$ where $s = \langle t+1, \theta, \tilde{m}\rangle$. For the state $s = \langle t, \theta, \tilde{m}\rangle$, the related constraints in Program (8) and Program (14) are as follows:

$$\begin{aligned} V(s) \geq \gamma^t & \sum\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)a_l u_l^a + (1 - \sum\nolimits_{l\in\mathcal{L}} a_l X_l^\theta)V(s^n) \\ & + \sum\nolimits_{l\in\mathcal{L}} a_l X_l^\theta V(s^e) \end{aligned} \tag{16}$$

$$\begin{aligned} \tilde{V}(s) \geq \gamma^t & \sum\nolimits_{l\in\mathcal{L}}(1 - X_l^\theta)a_l u_l^a + \tilde{V}(s^n) \\ & - \sum\nolimits_{l\in\mathcal{L}} a_l w_l^\theta(s^n) + \sum\nolimits_{l\in\mathcal{L}} a_l w_l^\theta(s^e) \end{aligned} \tag{17}$$

where $s^n = \langle t+1, normal, \tilde{m}'\rangle$, $s^e = \langle t+1, emergency, \tilde{m}'\rangle$ and $\tilde{m}' = \tilde{m} - \sum_{l\in\mathcal{L}} a_l$. The difference between the right hands of Eq.(16) and Eq.(17) is upper bounded by:

$$\begin{aligned} & \varepsilon - \sum\nolimits_{l\in\mathcal{L}} a_l(X_l^\theta V(s^n) - w_l^\theta(s^n)) + \sum\nolimits_{l\in\mathcal{L}} a_l(X_l^\theta V(s^e) - w_l^\theta(s^e)) \\ & \leq \varepsilon + \sum\nolimits_{l\in\mathcal{L}} a_l(\frac{1}{2^Z}(\tilde{V}(s^n) + \tilde{V}(s^e))) \\ & \qquad + \sum\nolimits_{l\in\mathcal{L}} a_l \cdot X_l^\theta((V(s^e) - \tilde{V}(s^e)) - (V(s^n) - \tilde{V}(s^n))) \\ & \leq 2\varepsilon + \frac{|\mathcal{L}|}{2^{Z-1}}\gamma^{t+1}\bar{V} \end{aligned} \tag{18}$$

Note that $\sum_{l\in\mathcal{L}} a_l \cdot X_l^\theta \leq 1$ is naturally ensured by the linear approximation because the transition probability always be positive. As our problem is a minimization problem, the increment from $\tilde{V}(s)$ and $V(s)$ is also upper bounded by:

$$V(s) - \tilde{V}(s) \leq 2\varepsilon + \frac{|\mathcal{L}|}{2^{Z-1}}\gamma^{t+1}\bar{V}$$

So that we can find the following inequality of the bounds obtained from Algorithm 1:

$$\begin{aligned} V(s) - \tilde{V}(s) & \leq \frac{|\mathcal{L}|}{2^{Z-1}} \cdot \sum_{t=1}^{2m} 2^{t-1}\gamma^t\bar{V} \\ & \leq \frac{|\mathcal{L}|}{2^{Z-1}} \cdot \frac{\gamma(1-(2\gamma)^{2m})}{1-2\gamma}\bar{V} \end{aligned} \tag{19}$$

Specifically, when $\gamma < 0.5$, Eq.(19) becomes:

$$V(s) - \tilde{V}(s) \leq \frac{|\mathcal{L}|}{2^{Z-1}} \cdot \frac{\gamma}{1-2\gamma}\bar{V}$$

which concludes the proof. □

5.5 Improving the scalability

Program (14) involves too many auxiliary (binary) variables and constraints to relax the problem into a MILP, which makes it unscalable. Therefore, we propose an algorithm, **S**tate and **A**ction **G**eneration for **BBC** (SAG-BBC),

[7] In practice, the upper bound can be tightened by taking into account the capability of each port.

Algorithm 1: Binary-based MDT for CIM

```
Initialize $H$, $\epsilon$;
repeat
    $\langle X, \tilde{V} \rangle \leftarrow$ Program (14);
    $\hat{V} \leftarrow$ Program (6), given $X$;
    $Z = Z + 1$;
until $(\hat{V}(s_0) - \tilde{V}(s_0))/\hat{V}(s_0) < \epsilon$;
return $\langle X, \hat{V} \rangle$;
```

depicted in Algorithm 2 to compute the global optimal solution iteratively, which is based on the observations that the smuggler intends to smuggle all the containers in the first several time steps because of the discount factor and there are many abundant actions for each state which are never selected by the smuggler. The basic idea of Algorithm 2 is as follows: Instead of solving the problem where the horizon of the smuggler's policies is $2m$, a restricted problem where the smuggler's MDP has much smaller horizon h is solved (i.e., $\mathcal{S}_h = \{s | s = \langle t, \theta, m \rangle \in \mathcal{S}, t \leq h\}$). In order to solve the optimal solution $\langle X(h), V(h) \rangle$ of the restricted problem, the algorithm first calls Algorithm 1 to compute the optimal solution $\langle X'(h), V'(h) \rangle$ for the restricted problem where the in the MDP, only a subset of actions $\mathcal{A'}_s \subset \mathcal{A}_s$ are available at state $s \in \mathcal{S}_h$ (i.e., Line 6). Then, the algorithm calls Program (6) to compute the smuggler's optimal policy $\hat{\pi}'_h$ against $X'(h)$ in the restricted problem assuming all actions $\mathcal{A}_s$ are available for $s \in \mathcal{S}_h$ (i.e., Line 7). If there are actions which are selected by $\hat{\pi}'_h$ but not in $\mathcal{A}'_s$, we add them into $\mathcal{A}'_s$ and resolve the restricted problem with action set $\mathcal{A}'_s$. Otherwise, the optimal solution $\langle X'(h), V'(h) \rangle$ for the restricted problem with action set $\mathcal{A}'_s$ is optimal to the restricted problem with all actions $\mathcal{A}_s$ available, i.e., $\langle X'(h), V'(h) \rangle = \langle X(h), V(h) \rangle$. Given the inspector's strategy $X(h)$, the algorithm calls Program (6) to compute the smuggler's optimal value $\hat{V}_h$ in the original MDP with horizon $2m$ and action set $\mathcal{A}$ (i.e., Line 14). If the value of $V(h)$ at initial state $V_h(s_0)$ equals $\hat{V}_h(s_0)$, the algorithm terminates and the optimal solution is obtained; Otherwise, increase the horizon h of restricted MDP by fixed time steps (e.g., 1) and resolve the restricted problem. Theorem 4 ensures that the computed solution is the optimal solution.

THEOREM 4. *Algorithm 2 returns the optimal solution for Program (14).*

PROOF SKETCH. We divide the proof into two parts according to the two loops in Algorithm 2.

Inner loop: when all actions selected by $\hat{\pi}'_h$ belong to $\mathcal{A}'_s$, all the constraints of Program 6 corresponding to actions in $\mathcal{A}_s$ are satisfied, so that $\langle X'(h), V'(h) \rangle$ is the optimal solution for Program 6 for the restricted MDP.

Outer loop: As the inner loop can obtain the optimal solution of the restricted problem where we restrict the smuggler's policies with in h time step. Line 14 of Algorithm 2 computes the real smuggler's optimal value $\hat{V}_h$ to the inspector $X(h)$ in the original MDP. If the value of $V(h)$ at initial state equals $\hat{V}_h(s_0)$, V_h is the optimal utility against X, which implies that $\langle X(h), V(h) \rangle$ is optimal to the original MDP. Thus, Algorithm 2 obtains the optimal solution. □

Algorithm 2: State and Action Generation for BBC

```
1  Initialize $h$;
2  repeat
3      $\mathcal{S}_h = \{s | s = \langle t, \theta, m \rangle \in \mathcal{S}, t \leq h\}$;
4      Arbitrarily select $\mathbf{a}^s \in \mathcal{A}_s$ to form $\mathcal{A}'_s \subset \mathcal{A}_s, \forall s \in \mathcal{S}_h$;
5      while true do
6          $\langle X'(h), V'(h) \rangle \leftarrow$ solution of Algorithm 1 by
           substituting $\mathcal{S}$ and $\mathcal{A}_s, \forall s \in \mathcal{S}$ with $\mathcal{S}_h$ and
           $\mathcal{A}'_s, \forall s \in \mathcal{S}_h$, respectively;
7          $\hat{V}'_h \leftarrow$ solution of Program (6) by substituting $\mathcal{S}$ with
           $\mathcal{S}_h$, given $X'(h)$;
8          Find the smuggler's optimal policies $\hat{\pi}'_h$, given $\hat{V}'_h$;
9          if $\exists s \in \mathcal{S}_s : \pi'(s) \notin \mathcal{A}'_s$ then
10             $\mathcal{A}'_s = \mathcal{A}'_s \cup \{\pi'(s)\}$;
11         else
12             $\langle X(h), V(h) \rangle \leftarrow \langle X'(h), V'(h) \rangle$
13             break;
14     $\hat{V}_h \leftarrow$ solution of Program (6), given $X(h)$;
15     $h = h + 1$;
16 until the value of $V(h)$ at initial state equals $\hat{V}(s_0)$;
17 return $\langle X(h), V(h) \rangle$;
```

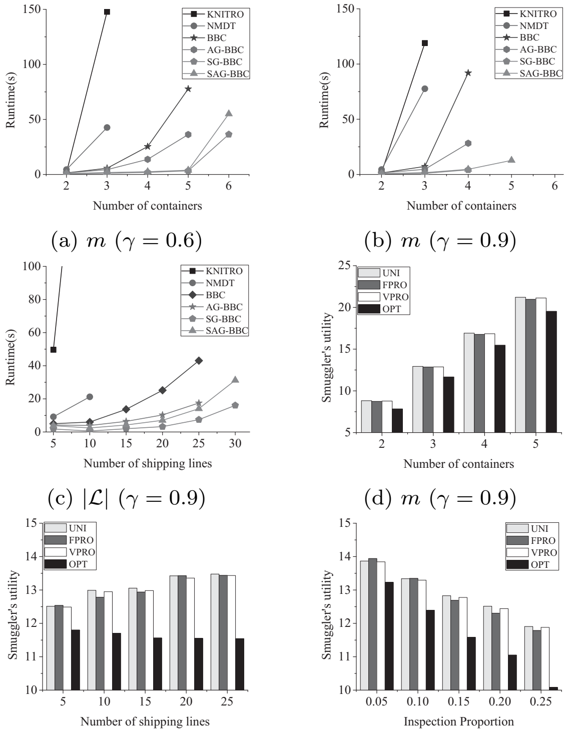
6. EXPERIMENTAL EVALUATION

We evaluate the performance of our approaches through extensive experiments. We use CPLEX (version 12.6) to solve linear programs and KNITRO (version 9.0.0) to solve nonlinear programs. All computations were performed on a 64-bit PC with 16.0 GB RAM and a 12-core 3.50 GHz processor. All values are averaged over 30 instances unless otherwise specified. All shipping networks are generated uniformly: for each shipping line, each port has a fixed probability of being visited. The payoff u^a_l and the flow f_l for each shipping line are generated from uniform distributions between [4, 5] and [0, 5], respectively. We use $\zeta \cdot \sum_{l \in \mathcal{L}} f_l$ to denote the total number of inspection resources, among which the number of the emergency inspection resources is fixed as $0.01 \cdot \sum_{l \in \mathcal{L}} f_l$ and other resources are randomly assigned to the ports. Thus, ζ is the proportion of containers inspected through all shipping lines at the emergency mode, which ranges in $[0.05, 0.25]$. The optimality tolerance ϵ is 0.001. The default setting of the experiments is $\langle |\mathcal{L}|, |\mathcal{P}|, m, \gamma, \zeta \rangle = \langle 10, 10, 3, 0.9, 0.15 \rangle$.

We compare the scalability of four versions of our algorithms: i) BBC depicted in Algorithm 1; ii) AG-BBC: Algorithm 2 with $T = 2m$; iii) SG-BBC: Algorithm 2 with $\mathcal{A}'_s = \mathcal{A}_s, \forall s \in \mathcal{S}_h$; iv) SAG-BBC: Algorithm 2. All versions return the same global optimal solution, which is denoted by OPT. The benchmarks are: i) KNITRO which is widely used in solving nonlinear program and ii) Normalized MDT (NMDT) proposed in [8] to solve the bilinear problems.

We compare the solution quality of our solution with three heuristic allocation strategies: i) UNI where the inspection resources at a port are uniformly assigned to shipping lines which pass this port and the emergency resources are assigned to all shipping lines uniformly; ii) FPRO where the number of inspection resources assigned to shipping line l is proportional to its container flow f_l; iii) VPRO where the number of inspection resources assigned to shipping line l is proportional to the payoff value u^a_l.

Scalability analysis. We compare the scalability of six methods on the generated shipping network. The experiment results are displayed in Figures 3a-3c. We range the



(a) m $(\gamma = 0.6)$

(b) m $(\gamma = 0.9)$

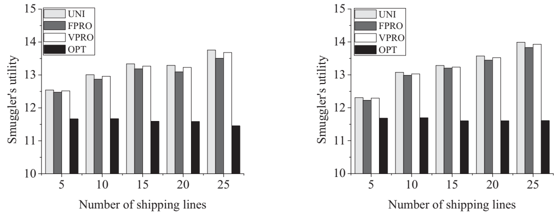
(c) $|\mathcal{L}|$ $(\gamma = 0.9)$

(d) m $(\gamma = 0.9)$

(e) $|\mathcal{L}|$ $(\gamma = 0.9)$

(f) ζ $(\gamma = 0.9)$

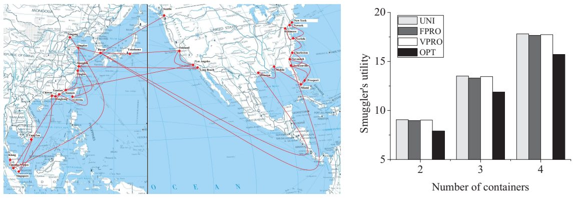
Figure 3: Scalability and Solution quality.



(a) γ $(\delta = 0.1)$

(b) $u_l^a (\rho = 10\%)$

Figure 4: Robustness.



(a) Real shipping network.

(b) Smuggler's utility.

Figure 5: Application on a real shipping network.

number of containers m under $\gamma \in \{0.6, 0.9\}$. The results show that our approaches significantly outperform KNITRO and NMDT and SG-BBC has a better performance when γ is small, while SAG-BBC performs better when γ becomes larger. This is because when γ increases, the smuggler prefers a policy with longer time horizon, i.e., the algorithm needs to consider more time steps to reach the global optimal solution. SG-BBC adds all states and actions within the time steps, while SAG-BBC selectively adds actions of each state into consideration, which limits the number of constraints in the program and makes the program scalable. We also range the number of shipping lines and the result is displayed in Figure 3c. Compared with the number of shipping lines, the number of containers has much more influence on the scalability because it influences both the number of actions in each state and the time horizon.

Solution quality. We compare the quality of our solution with three baseline strategies with varying the number of containers, shipping lines and the value of ζ. The results are showed in Figure 3d-3f. As the inspector is minimizing the smuggler's utility, our solution, denoted by OPT, outperforms the heuristic strategies in all settings. Besides, when the number of shipping lines, containers and the inspection proportion increases, our solution has a greater advantage over the heuristic strategies, which implies the effectiveness of strategically allocation of inspection resources.

Robustness. In reality, it is difficult for the inspector to know the smuggler's discount factor γ and the payoff u_l^a of each shipping line. In this section of experiments, we assume that the real value of the discount factor $\hat{\gamma}$ may range in $[\gamma - \delta, \gamma + \delta]$ where γ is the value from the inspector's perspective and $\delta < \min\{\gamma, 1-\gamma\}$. We assume the user uses γ to compute her optimal strategy while the smuggler uses $\hat{\gamma}$ to compute his optimal policy and find the smuggler's utility. Figure 4a shows the smuggler's utility where our solution still outperforms the heuristic strategies with $\gamma = 0.9$ and $\delta = 0.1$. Analogously, we assume the real payoff of shipping line $\hat{u}_l^a$ may range in $u_l^a \cdot [1-\rho, 1+\rho]$ where u_l^a is the payoff from the inspector's perspective and $0 < \rho < 1$. Figure 4b shows that our solution is robust enough to outperform the baselines with a $\rho = 10\%$ error under the default setting with three containers.

Application on a real shipping network. We also conduct experiments on a real shipping network displayed in Figure 1a, which includes 32 ports and 23 shipping lines operated by the three largest shipping companies[8] from Asia to North America. The smuggler's utility depicted in Figure 5b shows that our solution outperforms the heuristic strategies for the real shipping network, especially when the number of containers becomes larger.

7. CONCLUSION

This paper studies the problem of optimal inspection to prevent nuclear smuggling by containers. We introduce a novel container inspection model (CIM) and propose several efficient algorithms to compute the near-optimal solution, including a linear relaxation approximation which solves the near-optimal solution with a bilinear program, a novel approach inspired by MDT to obtain the optimal solution of the bilinear program and an iterative method of state and action generation to further improve the scalability. Extensive experiments show that our algorithms significantly outperform the existing methods and can obtain a robust enough solution better than heuristic strategies and can scale up to realistic-sized problems.

Acknowledgments

This research is supported by NRF2015NCR-NCR003-004.

[8] https://www.maerskline.com/, https://www.msc.com/sgp, http://www.cma-cgm.com/

REFERENCES

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

[2] B. An, M. Brown, Y. Vorobeychik, and M. Tambe. Security games with surveillance cost and optimal timing of attack execution. In *Proceedings of the 12th International Joint Conference on Autonomous Agents and Multi-Agent Systems(AAMAS)*, pages 223–230, 2013.

[3] N. O. Bakır. A Stackelberg game model for resource allocation in cargo container security. *Annals of Operations Research*, 187(1):5–22, 2011.

[4] N. Bakshi, S. E. Flynn, and N. Gans. Estimating the operational impact of container inspections at international ports. *Management Science*, 57(1):1–20, 2011.

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

[6] V. M. Bier and N. Haphuriwat. Analytical method to identify the number of containers to inspect at US ports to deter terrorist attacks. *Annals of Operations Research*, 187(1):137–158, 2011.

[7] E. Boros, Fedzhora, P. Kantor, K. Saeger, and P. Stroud. A large-scale linear programming model for finding optimal container inspection strategies. *Naval Research Logistics(NRL)*, 56(5):404–420, 2009.

[8] P. M. Castro. Normalized multiparametric disaggregation: An efficient relaxation for mixed-integer bilinear problems. *Journal of Global Optimization*, 64(4):765–784, 2016.

[9] CBP. Container Security Initiative: In Summary. Report, U. S. Customs and Border Protection, 2011.

[10] R. Church, M. Scaparra, and R. Middleton. Identifying critical infrastructure: The median and covering facility interdiction problems. *Annals of the Association of American Geographers*, 94(3):491–502, 2004.

[11] DHS. Secure Freight Initiative. `http://www.dhs.gov/secure-freight-initiative`.

[12] GAO. Combating nuclear smuggling: Megaports Initiative faces funding and sustainability challenges. Technical report, United States Government Accountability Office, 2012.

[13] GAO. Combating nuclear smuggling: DHS research and development on radiation detection technology could be strengthened. Technical report, United States Government Accountability Office, 2015.

[14] Q. Guo, B. An, Y. Zick, and C. Miao. Optimal interdiction of illegal network flow. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence(IJCAI)*, pages 2507–2513, 2016.

[15] N. Haphuriwat, V. M. Bier, and H. H. Willis. Deterring the smuggling of nuclear weapons in container freight through detection and retaliation. *Decision Analysis*, 8(2):88–102, 2011.

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

[17] M. Jain, D. Korzhyk, O. Vanek, V. Conitzer, M. Pechoucek, and M. Tambe. A double oracle algorithm for zero-sum security games on graphs. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 327–334, 2011.

[18] G. Leitmann. On generalized Stackelberg strategies. *Journal of Optimization Theory and Applications*, 26(4):637–643, 1978.

[19] T. J. Leonard, P. Gallo, and S. Véronneau. Security challenges in United States sea ports: an overview. *Journal of Transportation Security*, 8(1-2):41–49, 2015.

[20] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I - convex underestimating problems. *Mathematical programming*, 10(1):147–175, 1976.

[21] NNSA. Megaports Initiative. `https://nnsa.energy.gov/aboutus/ourprograms/nonproliferation/programoffices/internationalmaterialprotectionandcooperation/-5`.

[22] NTI. Illicit trafficking in weapons-useable nuclear material: Still more questions than answers. `http://www.nti.org/analysis/articles/illicit-trafficking-weapons-useable-nuclear-material-still-more-questions-answers/`, 2011.

[23] F. Pan, W. Charlton, and D. Morton. Stochastic network interdiction of nuclear material smuggling. *Network Interdiction and Stochastic Integer Programming*, pages 1–19, 2001.

[24] J. E. Ramirez-Marquez. Port-of-entry safety via the reliability optimization of container inspection strategy through an evolutionary approach. *Reliability Engineering & System Safety*, 93(11):1698–1709, 2008.

[25] P. J. Schweitzer and A. Seidmann. Generalized polynomial approximations in Markovian decision processes. *Journal of Mathematical Analysis and Applications*, 110(2):568–582, 1985.

[26] J. P. Teles, P. M. Castro, and H. A. Matos. Global optimization of water networks design using multiparametric disaggregation. *Computers & Chemical Engineering*, 40:132–147, 2012.

[27] J. Tsai, Z. Yin, J. Kwak, D. Kempe, C. Kiekintveld, and M. Tambe. Urban security: Game-theoretic resource allocation in networked domains. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, pages 881–886, 2010.

[28] H. von Stackelberg. *Marktform und Gleichgewicht*. J. Springer, 1934.

[29] Z. Wang, Y. Yin, and B. An. Computing optimal monitoring strategy for detecting terrorist plots. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pages 637–643, 2016.

[30] R. K. Wood. Deterministic network interdiction. *Mathematical and Computer Modelling*, 17(2):1–18, 1993.

[31] M. Zhao, B. An, and C. Kiekintveld. Optimizing personalized email filtering thresholds to mitigate sequential spear phishing attacks. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pages 658–664, 2016.