Ann Oper Res (2016) 236:271–289
DOI 10.1007/s10479-014-1722-3

Modeling costly learning and counter-learning in a defender-attacker game with private defender information

Jie Xu · Jun Zhuang

Published online: 17 September 2014

Abstract In asymmetric war scenarios (e.g., counter-terrorism), the adversary usually invests a significant time to learn the system structure and identify vulnerable components, before launching attacks. Traditional game-theoretic defender-attacker models either ignore such learning periods or the entailed costs. This paper fills the gap by analyzing the strategic interactions of the terrorist's costly learning and defender's counter-learning and defense strategies in a game with private defender information. Our model allows six possible attacker strategies: (a) attack immediately; (b) learn and attack; (c) learn and not attack; (d) learn and attack when appearing vulnerable and not attack when appearing invulnerable; (e) learn and not attack when appearing vulnerable and attack when appearing invulnerable; and (f) not attack. Our results show that four of the six strategies (a, d, e, f) are possible at equilibrium and the other two (b, c) are strictly dominated. Interestingly, we find that the counterintuitive strategy (e) could be at equilibrium, especially when the probability that the target appears vulnerable given it is invulnerable is sufficiently high. Our results also show that the attacker's learning cost has a significant impact on both the attacker's best responses and the defender's equilibrium deception and defense strategies. Finally, we study the attacker's values of perfect information and imperfect information, which provide additional insights for defense and counter-learning strategies.

Keywords Defender-attacker games · Costly learning · Counter-learning · Game theory · Value of perfect information · Value of imperfect information

This research was partially supported by the United States Department of Homeland Security (DHS) through the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) under award number 2010-ST-061-RE0001. This research was also patricianly supported by the United States National Science Foundation under award numbers 1200899 and 1334930. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the DHS, CREATE, or NSF.

J. Xu · J. Zhuang (✉)
Department of Industrial and System Engineering, SUNY at Buffalo,
Buffalo, NY 14260, USA
e-mail: jzhuang@buffalo.edu

1 Introduction

Since the 9/11 attacks, many researchers have studied protection against terrorism. Some works address the defender's optimization problem with exogenous attacker effort levels (Bier et al. 2005; Sandler and Siqueira 2006; Zhuang 2010; Bier and Haphuriwat 2011). However, attackers can strategically respond to government's defensive strategies. Therefore, some researchers have considered endogenous attacker efforts (Zhuang and Bier 2007; Hausken and Zhuang 2011), using game theory (Alpern et al. 2011) or adversarial risk analysis (Insua et al. 2009). Since information asymmetry is common in conflict scenarios, defender-attacker games of incomplete information have been studied in the literature. For example, Powell (2007) studies a defender-attacker, multi-target game where the defender has private information about asset vulnerabilities. Roberson (2006) considers symmetric and asymmetric configurations of the players' aggregate levels of force and constructs equilibrium variate distributions in Colonel Blotto games. Powell (2009) studies the allocation problem of confronting a defender who decides how to distribute limited resources across multiple sites before an attacker chooses where to strike. Bohme and Moore (2009) devise a model for dynamic interaction between a defender and an attacker, compare optimal security investments over multiple periods, and explore the delicate balance between proactive and reactive security investments.

In addition to endogenous attack efforts, an important reflection of attacker intelligence and adaptiveness is that the attacker can learn and test the system. In general, attackers could both monitor government defense activities (e.g., using cameras, taking notes, drawing diagrams, eliciting information from phone calls, mails, or in person), and test security (e.g., measuring reaction times to security breaches or penetrating physical security barriers or procedures).

Table 1 summarizes several real stories in which attackers made serious efforts to acquire information. For example, CNN (2010) reports that two passengers on a flight from Chicago to Amsterdam were arrested for testing the airport security by putting electronic devices inside bottles in their checked baggage, which was a clear attacker learning endeavor; two gunmen spent three days in monitoring and learning the Isabela election supervisor activities before killing the supervisor (Global Terrorism Database 2013); a terrorist named Hanjour received training on a Boeing 737 simulator in Arizona to prepare for the 9/11 attack (NCTA 2004). Al Qaeda carried out a test in September, 2010 to see how long it would take for a "package containing books, a computer disc and religious literature" to arrive in the U.S; by running such a test, bombmakers were able to work out "the timing to trigger the device" in order to maximize damage (Mail Online 2010). All the above activities costed the attackers significant amount of effort, even to the extent of risking arrest or death.

On the other hand, the defender could use disclosure to counter attacker's efforts. "Defender discloses only a subset of the defenses, in an attempt to route attacks to heavily-defended locations" is defined as "deception" (Hespanha et al. 2000; Brown et al. 2005;

Table 1 Examples of terrorists' learning activities

Year	Country	Activity	Target type	Source
2010	UK	Time testing	Transportation	Mail Online (2010)
2010	US/Netherlands	Telephone bomb and testing security	Transportation	CNN (2010)
2009	Philippines	Monitoring	Government	GTD (2013)
2000–2001	US	Flight training	Transportation	NCTA (2004)

Zhuang et al. 2010). The attackers' uncertainty about the defender's private information can create opportunities for either defender secrecy or deception (Zhuang and Bier 2010, 2011; Zhuang et al. 2010), in order to mislead the attackers. Secrecy and deception have been widely studied in military analysis (Joint Chiefs of Staff 1996), psychology (DePaulo et al. 2003), and computer science (Swire 2001), as well as economics and political science. Brown et al. (2005) discuss the benefits of secrecy in a zero-sum defender-attacker game in the context of ballistic missile deployment. Powell (2007) studies that the defender has private information about the vulnerability of targets in a defender-attacker game. Hausken and Levitin (2009) examine a defender-attacker game in which the defender builds both genuine and false targets while the attacker chooses the targets to attack in such a way as to maximize the system vulnerability. Cobb and Basuchoudhary (2009) define a modified decision-theoretic approach to solve games of strategic interaction between two players whose choices are modeled with separate decision trees comprised entirely of chance nodes.

To the best of our knowledge, no previous study has investigated the attackers' costly learning behavior before launching an attack, although this scenario is important and realistic. This paper aims at supporting the defender to help her making optimal or better decisions. One of the critical findings are how optimal attack and deception levels depend on various system parameters. This paper fills the gap by modeling the attacker's costly learning, and the defender's counter-learning and strategies in a sequential game. The next section describes the notation, the models, and a decision tree for the attacker. Section 3 provides the attacker's best response strategy, and the values of perfect and imperfect information. Section 4 provides the defender's optimal deception effort and defense level, and conducts sensitivity analyses. Section 5 summarizes the paper and discusses future research directions. An appendix provides definitions of all conditions, and sensitivity analysis of equilibrium strategies as functions of different parameters.

2 The model

Table 2 lists the notations that are used throughout this paper, including parameters, decision variables, functions and the notation for six attacker strategies.

The sequence of moves and the decision tree are illustrated in Fig. 1. We consider a game between one attacker and one defender. Common knowledge about the rules of the game is assumed among the players (Dutta 1999). In particular, as shown in Fig. 1, at the beginning of the game, nature randomly chooses the defender's type: being vulnerable and being invulnerable with probabilities P_V and $1 - P_V$, respectively.[1] Both players know the value of P_V and do not know the defender's actual type. The defender could use the deception effort d and the defense level l, at unit costs α and β respectively, to strengthen the target and deceive the attacker. The deception effort would be used to make the target appear less vulnerable so as to confuse the attacker. The defense level would be used to increase the strength of the target. After costly learning and updating, the attacker decides whether to attack.[2] In particular, as shown in the right-hand side of Fig. 1, the attacker has six possible

[1] This assumption is reasonable in many security scenarios (especially those involving new and less-than-fully tested technology), where even the defender could be uncertain about the system vulnerability (US Department of Homeland Security 2011).

[2] For simplicity, we focus on binary attacker effort (i.e., attack or not attack). This might be relevant in some high-level strategic decision-making situations, concerning which targets are likely to be attacked (rather than the level of attack effort on each targets). However, we acknowledge that the attack effort may be different among attacked targets and future work could consider continuous-level attack.

Table 2 Notations used in this paper

Notation	Explanation
Six attacker strategies	
A	Attack
NA	Not attack
$L_{A,A}$	Learn and A
$L_{A,NA}$	Learn and A when 'vulnerable' and NA when 'invulnerable'
$L_{NA,A}$	Learn and NA when 'vulnerable' and A when 'invulnerable'
$L_{NA,NA}$	Learn and NA
Parameters	
V_D, V_A	Defender's and attacker's target valuations, respectively
W	Attacker's cost of attacking invulnerable target
α	Defender's unit cost of deception effort
λ	Effectiveness coefficient of defender's deception effort
β	Defender's unit cost of defense level
k	Attacker's unit cost of learning
γ	Effectiveness coefficient for defender's defense level
P_V	Probability of target being vulnerable on the attacker's belief
$P_{V'\|NV}$	Given the target is invulnerable, the probability that it appears vulnerable
Decision variables	
a	Attacker's strategy, $a \in \{A,\ NA,\ L_{A,A},\ L_{A,NA},\ L_{NA,A},\ L_{NA,NA}\}$
d	Defender's deception effort
l	Defender's defense level
Functions	
$C(d)$	Attacker's cost of learning
$P_0(l)$	Probability of success when attacking the vulnerable defender
$P_{V'}(d)$	Probability that the target appears vulnerable when d is the deception effort
$P_{V\|V'}(d)$	Given the target appears vulnerable, the probability that it is vulnerable
$P_{V'\|V}(d)$	Given the target is vulnerable, the probability that it appears vulnerable
$P_{V\|NV'}(d)$	Given the target appears invulnerable, the probability that it is vulnerable
$U_D(a,d,l)$	Defender's expected utility
$U_A(a,d,l)$	Attacker's expected utility
$V_{PI}(d,l)$	Attacker's value of perfect information
$V_{II}(d,l)$	Attacker's value of imperfect information

strategies: attack immediately (A), not attack (NA), learn and launch an attack ($L_{A,A}$), learn and not attack ($L_{NA,NA}$), learn and attack when perceived as vulnerable and not attack when perceived as invulnerable ($L_{A,NA}$), and learn and not attack when perceived as vulnerable and attack when perceived as invulnerable ($L_{NA,A}$).

We assume that the attacker will succeed with probability $P_0(l)$, which depends on the defense level l, when attacking a vulnerable target and will not succeed when attacking an invulnerable target. If the attacker chooses to learn the system structure and identify its vulnerability, he needs to pay for the cost of learning $C(d) > 0$. In this paper, learning is referred to (a) obtaining knowledge on how to best attack a given target by recognizing the

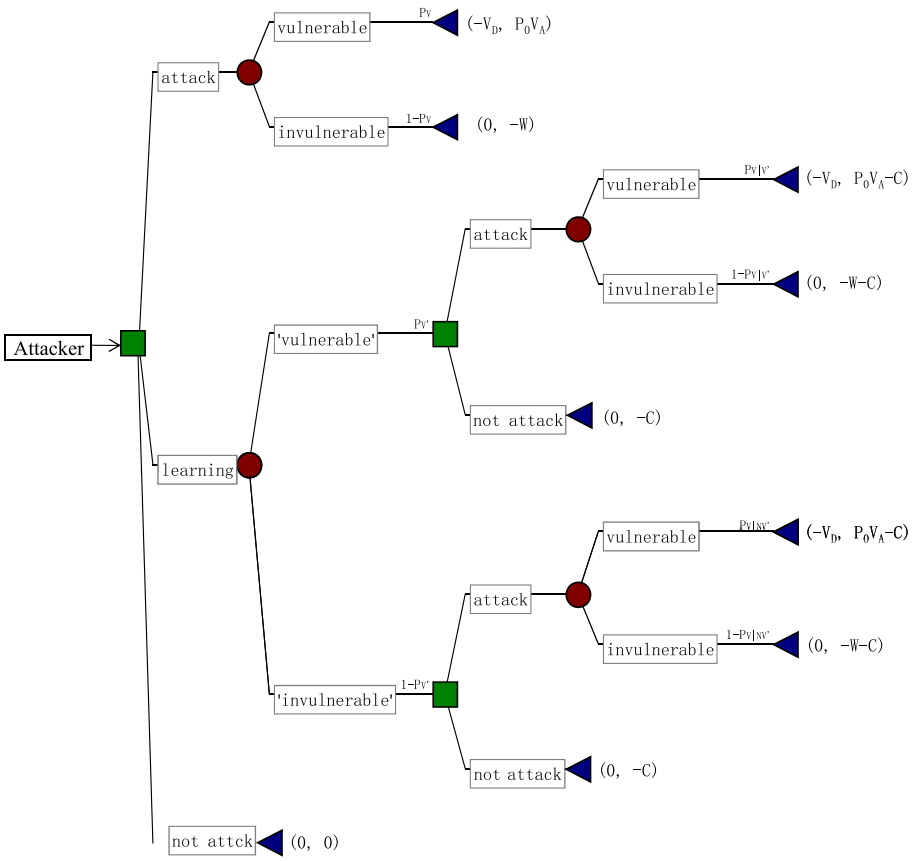

Fig. 1 Decision tree

target's vulnerability. We acknowledge that the term "learning" could have other meanings, including (b) acquiring new, or modifying existing, knowledge, behaviors, skills, values, or preferences; and (c) exploring the equilibrium in some game theoretic models. The attacker tries to maximize the expected damage to the target subtracting his learning cost $C(d)$ and expected cost of attacking the invulnerable target W. Prior to the learning period, the attacker has beliefs $P_V(d)$ about the target being vulnerable, which get updated, using Bayes' rule, to $P_{V|V'}(d)$ (or $P_{V|NV'}(d)$) depending on the observation. In this paper, we use the concept of subgame perfect Nash Equilibrium (SPNE, see Mas-Colell et al. 1995) and backwards induction to solve the model. In particular, Sect. 3 solves the attacker's best response functions and values of perfect and imperfect information, and then Sect. 4 derives the defender's equilibrium strategies considering the attacker's best responses.

3 Attacker's best response strategy

As discussed in Sect. 2, the attacker has six strategies: $a \in \{A, NA, L_{A,A}, L_{A,NA}, L_{NA,A}, L_{NA,NA}\}$. Evaluating the decision tree in Fig. 1, the attacker's expected utility (when the defender's deception effort is d and the defense level is l) for each strategy is:

$$U_A(a,l,d)=\begin{cases}[P_0(l)V_A+W]P_V-W & \text{if } a=A\\ [P_0(l)V_A+W]P_V-W-C(d) & \text{if } a=L_{A,A}\\ [(P_0(l)V_A+W)P_{V|V'}(d)-W]P_{V'}(d)-C(d) & \text{if } a=L_{A,NA}\\ [(P_0(l)V_A+W)P_{V|NV'}(d)-W][1-P_{V'}(d)]-C(d) & \text{if } a=L_{NA,A}\\ -C(d) & \text{if } a=L_{NA,NA}\\ 0 & \text{if } a=NA\end{cases} \tag{1}$$

Although the terrorist has six potential strategies, it is clear that strategies $L_{A,A}$ and $L_{NA,NA}$ are strictly dominated by strategies A and NA, respectively. For simplicity, in the equilibrium analysis, we shall assume: (a) when the attacker is indifferent between A and $L_{NA,A}$, he chooses $L_{NA,A}$; (b) when the attacker is indifferent between $L_{NA,A}$ and $L_{A,NA}$, he chooses $L_{A,NA}$; and (c) when the attacker is indifferent between $L_{A,NA}$ and NA, he chooses NA.

3.1 Attacker's best response

In this subsection, we derive the attacker's best response and corresponding optimal utilities. By comparing the attacker's expected utility of choosing the four possible non-dominated strategies, we have his best responses as follows:

$$a^*(d,l)=\begin{cases}A & \text{if } C_1(d,l) \text{ holds}\\ L_{A,NA} & \text{if } C_2(d,l) \text{ holds}\\ L_{NA,A} & \text{if } C_3(d,l) \text{ holds}\\ NA & \text{if } C_4(d,l) \text{ holds}\end{cases} \tag{2}$$

where the conditions $C_i(d,l)$, $i=1,\ldots,4$, are defined in "Appendix". Then, the corresponding optimal attacker's expected utilities are:

$$U_A^*(d,l)=\begin{cases}(P_0(l)V_A+W)P_V-W & \text{if } C_1(d,l) \text{ holds}\\ [(P_0(l)V_A+W)P_{V|V'}(d)-W]P_{V'}(d)-C(d) & \text{if } C_2(d,l) \text{ holds}\\ [(P_0(l)V_A+W)P_{V|NV'}(d)-W][1-P_{V'}(d)]-C(d) & \text{if } C_3(d,l) \text{ holds}\\ 0 & \text{if } C_4(d,l) \text{ holds}\end{cases} \tag{3}$$

We use the following baseline parameter values to illustrate the model. In realistic applications, such parameters could be estimated based on historical data (Shan and Zhuang 2013b), or expert elicitation (Wang and Bier 2013). In addition, the players' target valuation could be approximated using the expected economic losses or casualties (Shan and Zhuang 2013a); unit costs of defending, learning, and attacking could be based on the cost estimations on labor, equipment, supply, and other costs necessary for the corresponding operations; and the coefficients for defense effectiveness, probabilities and likelihood functions could be estimated using experiments. The defender and the attacker have the same target valuation ($V_A=V_D=15$), which is higher than the cost of attacking ($W=8$). The defender's unit cost of the deception effort is set at $\alpha=0.5$; the attacker's unit cost of learning is set at $k=0.1$; the effectiveness coefficient of the defender's deception effort is set at $\lambda=0.1$; the defender's unit cost of the defense level is set at $\beta=0.1$; the effectiveness coefficient of the defender's defense level is set at $\gamma=0.01$; the probability of the target being vulnerable is set at $P_V=0.4$; and the probability that the target appears vulnerable, given it is invulnerable, is set at $P_{V'|NV}=0.4$.

3.1.1 Attacker's best responses to different deception efforts

Figure 2 shows how the attacker's best responses change when the defense level is low ($l=0$). It shows that any of the four strategies (A, $L_{A,NA}$, $L_{NA,A}$ and NA) can be used by

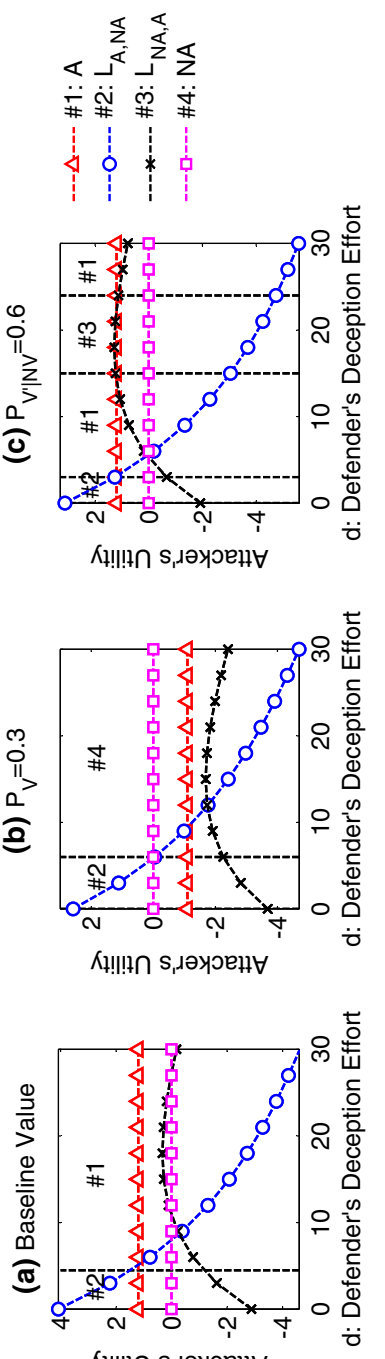

Fig. 2 Comparison of attacker's expected utilities when using four possible strategies with different deception effort and $l = 0$

the attacker as a best response to the deception effort when the defense level is low ($l = 0$). As the defender's deception effort increases, we see that the attacker's expected utility of using strategy $L_{A,NA}$ (case #2) decreases, the attacker's expected utility of using strategy $L_{A,NA}$ (case #3) increases, and the expected utilities of using strategies A (case #1) and NA (case #4) remain constant.

The probability of successfully attacking the vulnerable defender when the defense level is l, $P_0(l)$, will be modeled as follows:

$$P_0(l) = e^{-\gamma l} \tag{4}$$

where γ is a coefficient representing the effectiveness of the defender's defense level. When $l = 0$, we have $P_0(l) = 1$; and when $l \to \infty$, we have $\lim_{l\to\infty} P_0(l) = 0$.

The attacker's learning cost is highly correlated with the defender's deception effort d, i.e., the attacker's learning cost increases with the defender's deception effort. We assume that:

$$C(d) = kd \tag{5}$$

where k is attacker's unit cost of learning based on the defender's deception effort d.

We use the following exponential function to model the probability that the target appears vulnerable when it is vulnerable, and the defender's deception effort is d:

$$P_{V'|V}(d) = e^{-\lambda d} \tag{6}$$

where λ is a coefficient representing the effectiveness of the defender's deception effort. When $d = 0$, we have $P_{V'|V}(d) = 1$ (the vulnerable target would appear vulnerable with certainty). When $d \to \infty$, we have $\lim_{d\to\infty} P_{V'|V}(d) = 0$ (the vulnerable target would appear invulnerable with certainty).

In particular, Fig. 2a shows that when the defender's deception effort level is low ($d < 4.5$), meaning that the attacker has a belief close to the truth about the system vulnerability ($P_{NV'|V}$ is low), the attacker would choose to learn the system first, and attack if it appears vulnerable and not to attack if it appears invulnerable ($L_{A,NA}$, case #2). As d increases, the probability that the attacker's false belief on the system vulnerability ($P_{NV'|V}(d)$) increases. In this case, the attacker would choose to attack immediately (A, case #1) because learning is not beneficial in the presence of much deception.

In Fig. 2b, the probability of the target being vulnerable, P_V, decreases from the baseline value 0.4 to 0.3. When the defender's deception effort level is high ($d \geq 6$), we no longer have case #1, but have case #4 (not attack, NA) instead. This is because the expected utility of using strategy A increases in P_V. When P_V decreases, the expected utility of using strategy A decreases, while the expected utility of using strategy NA remains constant at zero which is greater than the expected utility of using strategy A. Therefore, the attacker would use NA (#4) when the deception effort d becomes sufficiently high.

In Fig. 2c, the probability that the target appears vulnerable when it is invulnerable, $P_{V'|NV}(d)$, increases from the baseline value 0.4 to 0.6. This implies that when the target is invulnerable, it is more likely to appear vulnerable. Interestingly, the attacker would use the strategy $L_{NA,A}$ (case #3; learn and choose to attack when it appears vulnerable and not to attack when it appears vulnerable) when $15 \leq d \geq 24$. This is because the more vulnerable the target appears, the more invulnerable it is. However, as the deception effort increases, the attacker's learning cost increases. Thus, when d is sufficiently high ($d > 24$), the attacker would choose case #1 instead of case #3 due to the increasing learning cost.

3.1.2 Attacker's best responses to different defense levels

Figure 3 shows the attacker's best response to the defense level when the deception effort is low ($d = 0$). As the defender's defense level increases, we see that the attacker's expected utilities of choosing strategies A (case #1) and $L_{A,NA}$ (case #2) decrease, and the attacker's expected utilities of choosing strategies $L_{A,NA}$ (case #3) and NA (case #4) remain constant. In Fig. 4, the attacker's expected utility of choosing strategies $L_{A,NA}$ (case #3) decreases in the defense level when the deception effort is high ($d = 10$).

Comparing Figs. 3a and 4a, and 3b and 4b, we see that case #2 ($L_{A,NA}$) disappears and the ranges of other cases get larger when the deception effort is higher ($d = 10$). This is because the probability that the attacker's false belief on the system vulnerability increases in the defender's deception effort (see Eq. 6); thus the attacker is more likely to be deceived.

Figure 5 shows the attacker's best response when the deception effort d and the defense level l vary. As shown in Fig. 5a, when the deception effort d is low ($d < 5$), the attacker would use the strategy $L_{A,NA}$ (case #2) because he is more likely to successfully attack after learning when $P_{V'|NV}$ is higher; e.g., $P_{V'|NV} = 0.4$. When l is lower ($l < 23$) and d is higher ($d > 6$), the attacker would use the strategy A (case #1) because he is more likely to be deceived, and thus he gives up learning and chooses the strategy of attacking directly. When both l and d are sufficiently high, the attacker would give up immediately (case #4) because, in this case, he is less likely to succeed in either learning or attacking. By comparing Fig. 5a, b, we see that when the defender becomes more likely to be invulnerable, the attacker would give up attacking. By comparing Fig. 5a, c, we see that $L_{NA,A}$ (case #3) appears when l is relatively low ($l < 30$) and d is in the middle ($14 < d < 23$). This is because the attacker would like to choose strategies $L_{A,NA}$ (case #3) when l is relatively low and d is in the middle range due to the increasing learning cost C caused by the high deception effort.

Figure 6a–c show the contour of the attacker's optimal expected utilities when d and l vary. We see that the attacker's expected utility decreases in l because the probability of successful attack decreases in l. In Fig. 6a, b, the attacker's expected utility decreases in d; while in Fig. 6c, the attacker's expected utility increases in d. This is because the attacker is more likely to be deceived when $P_{V'|NV}$ is lower.

3.2 Values of perfect and imperfect information

Sections 3.2.1 and 3.2.2 calculate the value of perfect information and of imperfect information, respectively. These values provide benchmark guidelines. For example, the defender could compare them with the (expected) attacker's learning costs. If the values are significantly high, the defender should expect the attacker to use the costly learning strategies and vice versa.

3.2.1 Value of perfect information

Figure 7 shows the decision tree under the option for the attacker to purchase perfect information about the target vulnerability. Based on Fig. 1, a new branch of purchasing perfect information is added in Fig. 7. After purchasing perfect information, the attacker would know whether the target is vulnerable or not. Then, the attacker makes a decision: attack, learn or not attack, although learning is always dominated after purchasing perfect information. In the calculation from the game tree in Fig. 7, we get the attacker's value of perfect information as shown in Eq. (7) subtracting the value without perfect information from the value when purchasing perfect information:

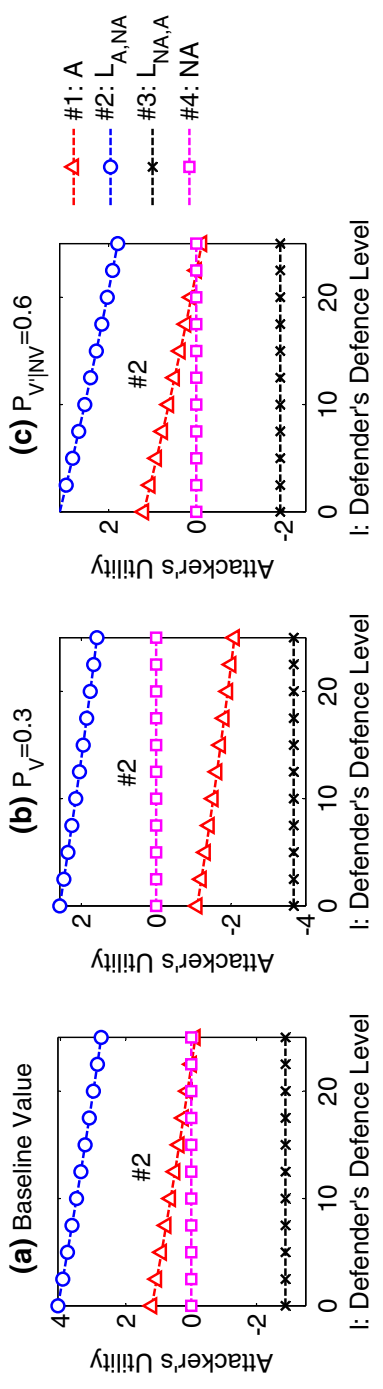

Fig. 3 Comparison of attacker's expected utilities when using four possible strategies with different defense level and $d = 0$

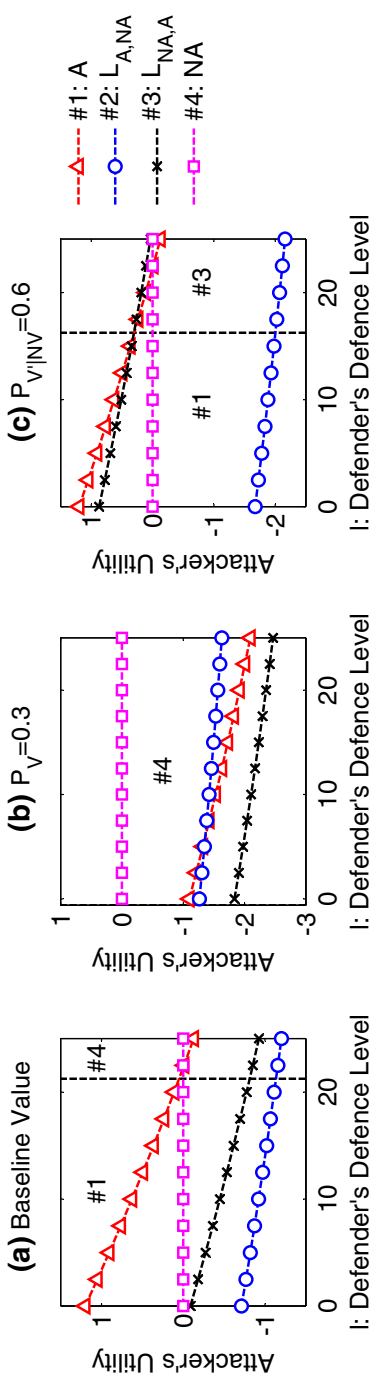

Fig. 4 Comparison of attacker's expected utilities when using four possible strategies with different defense level and $d = 10$

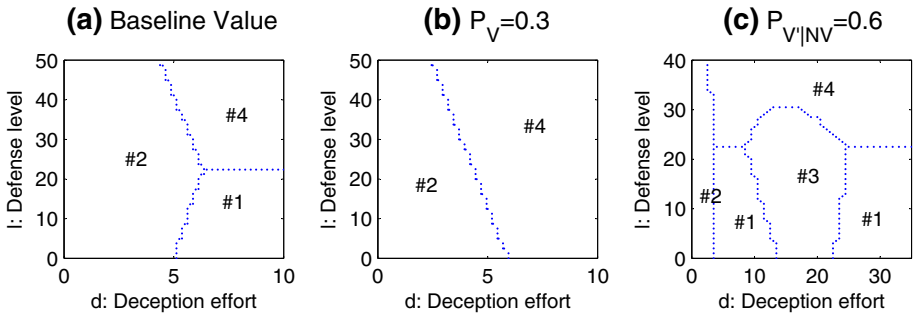

Fig. 5 Attacker's best response strategies with perfect information as a function of deception effort d and defense level l

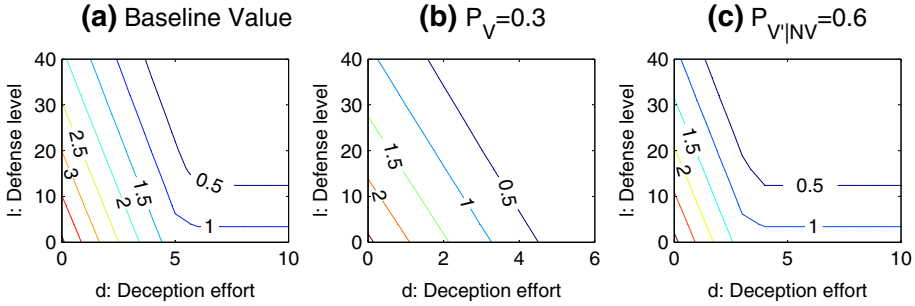

Fig. 6 Contours: attacker's optimal expected utility as a function of deception effort d and defense level l

$$V_{PI}(l,d) = \underbrace{P_V P_0(l) V_A}_{\text{Value of purchasing perfect information}} - \underbrace{\max\left\{\underbrace{(P_0(l)V_A+W)P_V - W}_{A}, \underbrace{[(P_0(l)V_A+W)P_{V|V'}(d)-W]P_{V'}(d)-C}_{L_{A,NA}}, \underbrace{[(P_0(l)V_A+W)P_{V|NV'}(d)-W][1-P_{V'}(d)]-C}_{L_{NA,A}}, \underbrace{0}_{NA}\right\}}_{\text{Value without perfect information}}. \tag{7}$$

Using the definition of $C_i(d,l)$, $i = 1,\ldots,4$ in "Appendix", Eq. (7) becomes:

$$V_{PI}(d,l) = \begin{cases} W(1-P_V) & \text{if } C_1(d,l) \text{ holds} \\ P_0(l)P_V V_A[1-P_{V'|V}(d)] + W P_{V'|NV}(1-P_V) + C(d) & \text{if } C_2(d,l) \text{ holds} \\ P_0(l)P_V V_A P_{V'|V}(d) + W(1-P_{V'|NV})(1-P_V) + C(d) & \text{if } C_3(d,l) \text{ holds} \\ P_0(l)P_V V_A & \text{if } C_4(d,l) \text{ holds} \end{cases} \tag{8}$$

Figure 8a–c show the contours of the attacker's value of perfect information as functions of d and l for three scenarios (baseline value, lower P_V, and higher $P_{V'|NV}$). We see that the value of perfect information decreases in l. This is because when l is high enough, the probability of successful attack is low and the attacker is more likely to not attack. In Fig. 8a–c, the value of perfect information increases in d. This is because the attacker is more likely to be deceived and thus the value difference between purchasing perfect information and without perfect information becomes larger.

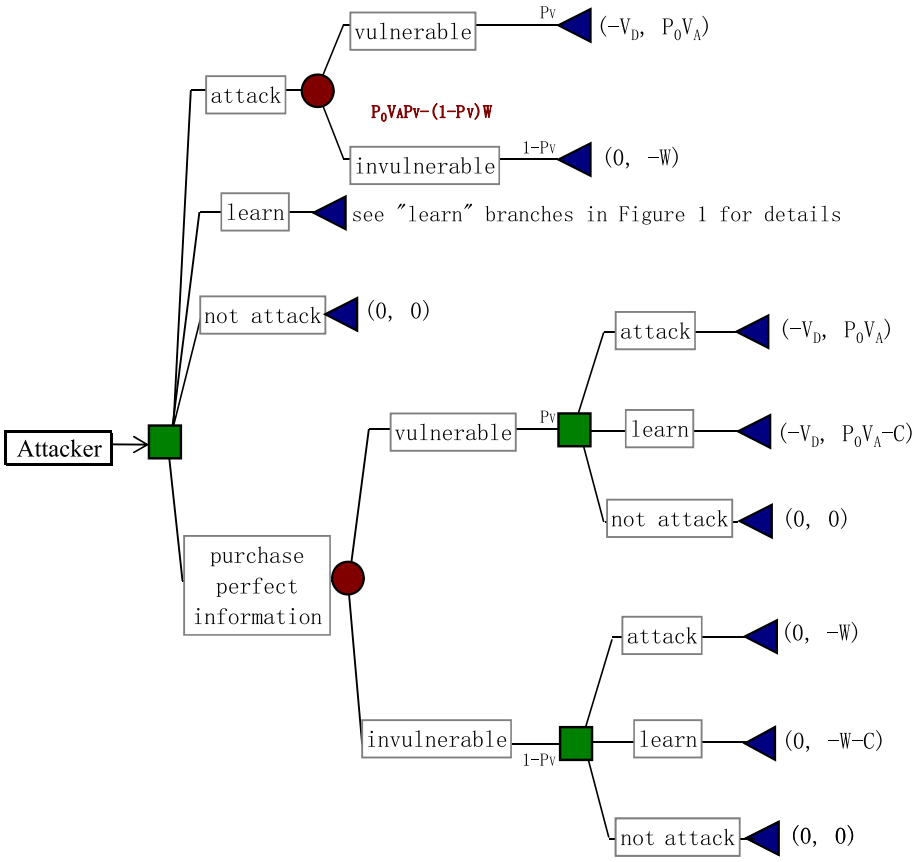

Fig. 7 Decision tree under the option of purchasing perfect information

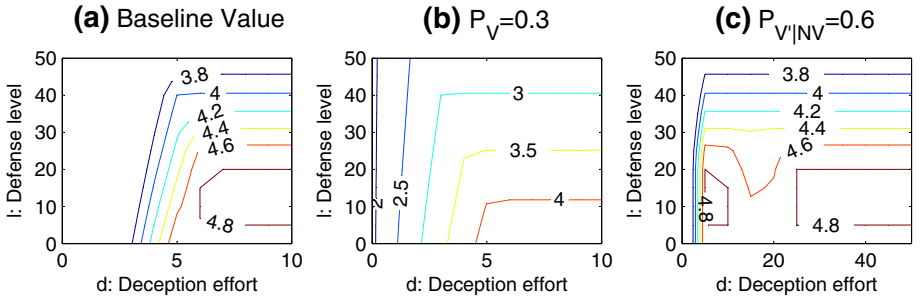

Fig. 8 Contours: value of perfect information as a function of deception effort d and defense level l

3.2.2 Value of imperfect information

Figure 9 shows the decision tree under the option for the attacker to purchase imperfect information about the vulnerability of the attacker's target. Based on Fig. 1, the corresponding new branch is added in Fig. 9. After purchasing imperfect information, the attacker would have his own belief on the vulnerability of the target. The attacker could choose one of these

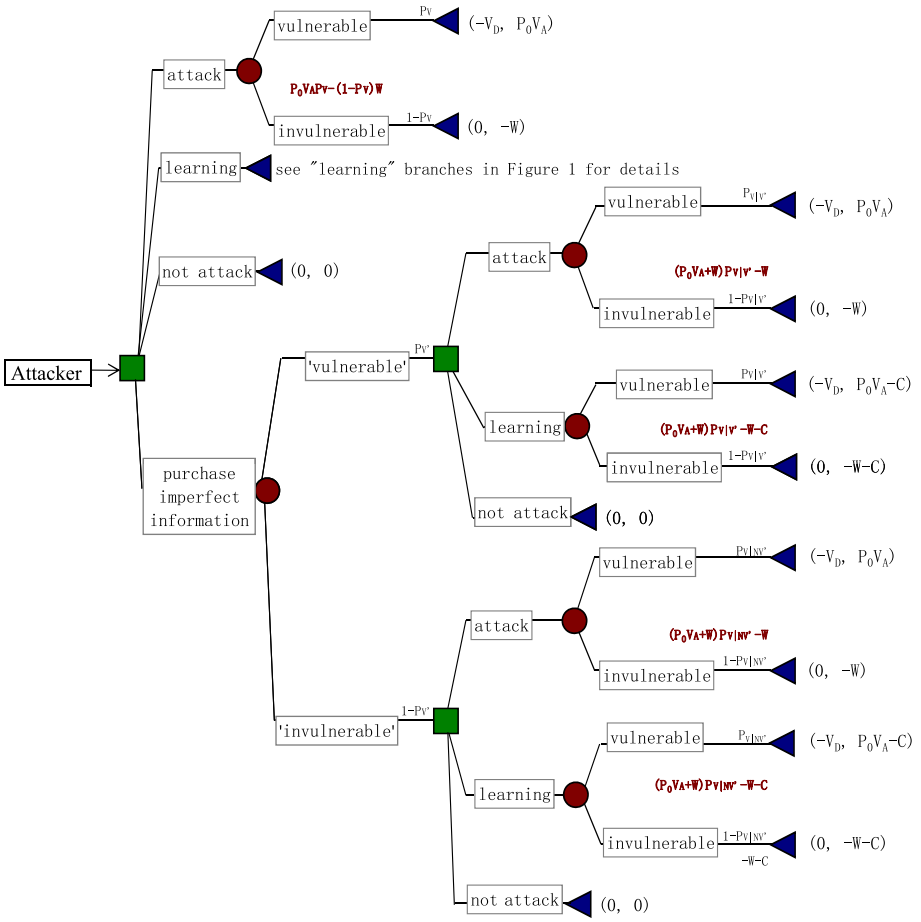

Fig. 9 Decision tree under imperfect information

three decisions: attack immediately, learning, and not attack. By subtracting the value without imperfect information from the value of purchasing imperfect information, we could get the value of imperfect information as shown in Eq. (9):

$$
\begin{aligned}
&V_{II}(d,l)\\
&\underbrace{= \max\left\{\underbrace{(P_0(l)V_A+W)P_V-W}_{\text{attack}}, \underbrace{[(P_0(l)V_A+W)P_{V|V'}(d)-W]P_{V'}(d)}_{\text{attack when it appears vulnerable}}, \underbrace{[(P_0(l)V_A+W)P_{V|NV'}(d)-W](1-P_{V'}(d))}_{\text{attack when it appears invulnerable}}, \underbrace{0}_{\text{not attack}}\right\}}_{\text{Value of purchasing imperfect information}}\\
&\underbrace{-\max\left\{\underbrace{(P_0(l)V_A+W)P_V-W}_{A}, \underbrace{[(P_0(l)V_A+W)P_{V|V'}(d)-W]P_{V'}(d)-C}_{L_{A,NA}}, \underbrace{[(P_0(l)V_A+W)P_{V|NV'}(d)-W][1-P_{V'}(d)]-C}_{L_{NA,A}}, \underbrace{0}_{NA}\right\}}_{\text{Value without imperfect information}}.
\end{aligned}
\tag{9}
$$

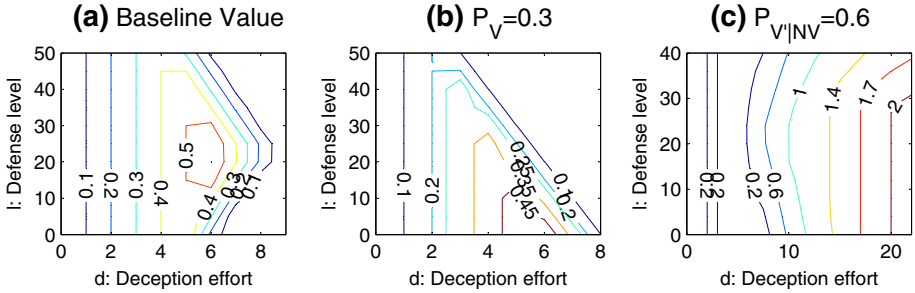

Fig. 10 Contours: value of imperfect information as a function of deception effort d and defense level l

Substituting Eqs. (4) and (6) to (9), we get Eq. (10):

$$V_{II}(d,l) = \begin{cases} 0 & \text{if } C_5(d,l) \text{ or } C_6(d,l) \text{ holds} \\ C(d) & \text{if } C_7(d,l) \text{ or } C_8(d,l) \text{ holds} \\ P_0(l)V_A P_V(1-e^{-\lambda d}) - W(1-P_V)(1+P_{V'|NV}) & \text{if } C_9(d,l) \text{ holds} \\ P_0(l)V_A P_V e^{-\lambda d} + W(1-P_V)P_{V'|NV} & \text{if } C_{10}(d,l) \text{ holds} \\ [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) & \text{if } C_{11}(d,l) \text{ holds} \\ [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1-P_{V'}(d)] & \text{if } C_{12}(d,l) \text{ holds} \end{cases} \tag{10}$$

where conditions $C_i(d,l)$, $i = 5,\ldots,12$ are defined in "Appendix".

Figure 10a–c show the contours of the value of imperfect information as functions of the deception effort d and the defense level l for three scenarios (baseline value, lower P_V, and higher $P_{V'|NV}$) in this simplified example. We see that the value of imperfect information first increases and then decreases in d in Fig. 10a, b, while the value of imperfect information increases in d in Fig. 10c. The value of imperfect information first increases and then decreases in l in Fig. 10a, c, while the value of imperfect information decreases in l in Fig. 10b.

4 Defender's optimal deception effort and defense level

Using the attacker's best response function Eq. (2), the defender's expected utility as a function of d and l is:

$$U_D(d,l) = U_D(a^*(d,l),d,l) = \begin{cases} -d\alpha - l\beta - V_D P_V & \text{if } C_1(d,l) \text{ holds} \\ -d\alpha - l\beta - V_D P_{V'}(d) P_V & \text{if } C_2(d,l) \text{ holds} \\ -d\alpha - l\beta - V_D P_{NV'}(d) P_V & \text{if } C_3(d,l) \text{ holds} \\ -d\alpha - l\beta & \text{if } C_4(d,l) \text{ holds} \end{cases} \tag{11}$$

and the optimal defender's deception effort and defense level are given as follows:

$$(d^*, l^*) = \arg\max_{d\geq 0, l\geq 0} U_D(d,l)$$

Figure 11a shows the equilibrium defender effort d^*, the equilibrium defense level l^*, the attacker's strategies a^*, and the two utilities U_D^* and U_A^*, as the parameter $P_{V'|NV}$ varies from its baseline value as introduced in Sect. 3.1. The vertical dashed lines indicate different equilibrium attacker strategies. The vertical solid lines indicate baseline values for each parameter. Figure 11b shows the attacker's values of perfect information and imperfect information at equilibrium. The first case before the arrow (→) means the strategy that the attacker actually

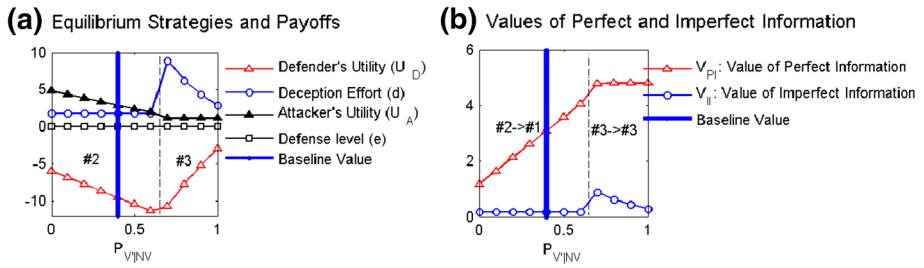

Fig. 11 Equilibrium strategies and value of perfect and imperfect information

takes in the game, while the seconde case after the arrow (→) means the strategy that the attacker would take if he has imperfect information about the system vulnerability. Note that when the attacker has the option of purchasing perfect information, he would always attack when it appears vulnerable and not attack otherwise.

In Fig. 11a, when $P_{V'|NV}$ is low ($P_{V'|NV} < 0.65$), the attacker would choose to learn and attack when it appears vulnerable and not to attack when it appears invulnerable (case #2, as defined in Sect. 3.1). The defender would use relatively low deception effort and defense level since the attacker would learn (using case #2) and update his belief. The defender's and the attacker's utilities decrease in $P_{V'|NV}$ because the attacker is more likely to be confused even without the deception with the increase of $P_{V'|NV}$. When $P_{V'|NV}$ is sufficiently high ($P_{V'|NV} > 0.65$), it is more likely that the target appears vulnerable when in fact it is invulnerable. Interestingly, the attacker would learn and attack when it appears invulnerable and not attack when it appears vulnerable (case #3), which seems counter-intuitive. In this case, the defender would decrease her deception effort to decrease the probability of being vulnerable in the attacker's belief.

In Fig. 11b, when $P_{V'|NV}$ is low ($P_{V'|NV} < 0.65$), the value of perfect information increases because $P_{V|V'}$ decreases in $P_{V'|NV}$ when the attacker chooses to learn the system first, and attack if it appears vulnerable and not to attack if it appears invulnerable (case #2) while the value of imperfect information remains constant. When $P_{V'|NV}$ is sufficiently high ($P_{V'|NV} > 0.65$), the value of imperfect information decreases.

The defender would be better off by knowing how she should balance efforts between counterintelligence and direct defenses for critical infrastructures, facing strategic adversaries with options of learning. For example, the defender should increase her deception effort to confuse the attacker in cases #1, #2, and #4. However, in the counterintuitive case #3, the defender should lower her deception effort.

5 Conclusion and future research directions

Asymmetric player relationship in modern terrorism is a key difference from the one in the cold war literature (see Schelling 1966, for an overview). In particular, such asymmetric player relationship in modern terrorism literature includes the different choices and information.

In this paper, we consider the scenario in which the attacker spends a considerable amount of time and effort learning the system vulnerability before launching real attacks. We contribute to the literature by analyzing the terrorist's costly learning and the defender's simultaneously counter-learning (using deception) and defense strategies in a sequential game of imperfect information. Interestingly, we find that four out of six possible strategies for

the attacker are at equilibrium; and the strategy "learn and not attack when apparently vulnerable and attack when apparently invulnerable" could be at equilibrium, which might be counterintuitive. In addition, we study the values of perfect information and imperfect information.

This paper would help the defender make better decisions by providing insights on: (a) how the attacker's learning and attacking strategies depend on defense and deception efforts and other system parameters; and (b) how to decide the optimal defense and deception strategies facing with uncertainties and adaptive adversaries; e.g., the defender would know when she should increase her deception effort and/or defense level to confuse the attacker and induce him to make wrong decisions.

Future research direction includes a continuous-time game. In particular, the theory of learning curves (Zangwill and Kantor 1998) could be used to model the real-time attacker's belief about the system vulnerability and how long the attacker would spend on learning, as well as the corresponding real-time defender's counter-learning strategies. Another research direction is the study of costly learning in repeated games, where the attacker would have to balance short-term loss and long-term gains. It is also promising to investigate a multiple-target game, and study which target the attacker might want to learn first, and the government's corresponding counter-learning and defense strategies. In addition, how would the attacker's behavior be affected by the attacker's or the defender's risk preferences (e.g., risk averse or risk seeking) could be further discussed.

Appendix: Definitions for the $C_i(d,l)$, $i = 1, \ldots, 12$

$$
\begin{aligned}
C_1(d,l) \equiv \big\{ & [P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d); \\
& [P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d); \\
& [P_0(l)V_A + W]P_V - W > 0 \big\}; \\
C_2(d,l) \equiv \big\{ & [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d) \geq W(1 - P_V)(1 - P_{V'|NV}); \\
& [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \geq [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)]; \\
& [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d) > 0 \big\}; \\
C_3(d,l) \equiv \big\{ & [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) \geq W(1 - P_V)P_{V'|NV}; \\
& [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] > [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d); \\
& [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) > 0 \big\}; \\
C_4(d,l) \equiv \big\{ & [P_0(l)V_A + W]P_V - W \geq 0; [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d) \geq 0; \\
& [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) \geq 0 \big\}; \\
C_5(d,l) \equiv \big\{ & [P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d); \\
& [P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|NV'}(d) - W](1 - P_{V'}(d)); \\
& [P_0(l)V_A + W]P_V - W > 0 \big\}; \\
C_6(d,l) \equiv \big\{ & [P_0(l)V_A + W]P_V - W \leq 0; [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \leq 0; \\
& [(P_0(l)V_A + W)P_{V|NV'}(d) - W](1 - P_{V'}(d)) \leq 0 \big\}; \\
C_7(d,l) \equiv \big\{ & [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d) \geq [P_0(l)V_A + W]P_V - W; \\
& [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \geq [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)]; \\
& [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d) > 0 \big\}; \\
C_8(d,l) \equiv \big\{ & [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) \geq [P_0(l)V_A + W]P_V - W; \\
& [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) < [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)];
\end{aligned}
$$

$$
\begin{aligned}
&[(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) > 0\big\};\\
C_9(d,l) \equiv \big\{&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \geq [P_0(l)V_A + W]P_V - W;\\
&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \geq [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)];\\
&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) > 0;\\
&[P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d);\\
&[P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d);\\
&[P_0(l)V_A + W]P_V - W > 0\big\};\\
C_{10}(d,l) \equiv \big\{&[(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) \geq [P_0(l)V_A + W]P_V - W;\\
&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) < [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)];\\
&[(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) > 0;\\
&[P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) - C(d);\\
&[P_0(l)V_A + W]P_V - W > [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d);\\
&[P_0(l)V_A + W]P_V - W > 0\big\};\\
C_{11}(d,l) \equiv \big\{&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \geq [P_0(l)V_A + W]P_V - W;\\
&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \geq [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)];\\
&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) > 0; [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \leq 0;\\
&[P_0(l)V_A + W]P_V - W \leq 0; [(P_0(l)V_A + W)P_{V|NV'}(d) - W](1 - P_{V'}(d)) \leq 0\big\};\\
C_{12}(d,l) \equiv \big\{&[(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) \geq [P_0(l)V_A + W]P_V - W;\\
&[(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) < [(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)];\\
&[(P_0(l)V_A + W)P_{V|NV'}(d) - W][1 - P_{V'}(d)] - C(d) > 0;\\
&[P_0(l)V_A + W]P_V - W \leq 0; [(P_0(l)V_A + W)P_{V|V'}(d) - W]P_{V'}(d) \leq 0;\\
&[(P_0(l)V_A + W)P_{V|NV'}(d) - W](1 - P_{V'}(d)) \leq 0\big\}.
\end{aligned}
$$

References

Alpern, S., Morton, A., & Papadaki, K. (2011). Patrolling games. *Operations Research*, *59*(5), 1246–1257.

Bier, V. M., & Haphuriwat, N. (2011). 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.

Bier, V. M., Nagaraj, A., & Abhichandani, V. (2005). Protection of simple series and parallel systems with components of different values. *Reliability Engineering and System Safety*, *87*(3), 315–323.

Bohme, R., & Moore, T. (2009). The iterated weakest link—a model of adaptive security investment. In *Workshop on the economics of information security (WEIS)*, University College, London, UK. Available at http://weis09.infosecon.net/files/152/paper152.pdf. Accessed in August, 2014.

Brown, G., Carlyle, M., Diehl, D., Kline, J., & Wood, K. (2005). A two-sided optimization for theater ballistic missile defense. *Operations Research*, *53*(5), 745–763.

CNN. (2010). Dutch arrest two men after flight from US Available at http://news.blogs.cnn.com/2010/08/30/two-men-arrested-at-amsterdam-airport/. Accessed in August, 2014.

Cobb, B. R., & Basuchoudhary, A. (2009). A decision analysis approach to solving the signaling game. *Decision Analysis*, *6*(4), 239–255.

DePaulo, B. M., Wetzel, C., Sternglanz, R. W., & Wilson, M. J. W. (2003). Verbal and nonverbal dynamics of privacy, secrecy, and deceit. *Journal of Social Issues*, *59*(2), 391–410.

Dutta, P. K. (1999). *Strategies and games: Theory and practice*. Cambridge, Massachusetts: MIT Press.

Global Terrorism Database. (2013). Available at http://www.start.umd.edu/gtd/search/IncidentSummary.aspx?gtdid=200911140007. Accessed in August, 2014.

Hausken, K., & Levitin, G. (2009). Protection vs. false targets in series systems. *Reliability Engineering and System Safety*, *94*(5), 973–981.

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

Hespanha, J. P., Ateskan, Y. S., & Kizilocak, H. H. (2000). Deception in non-cooperative games with partial information. In *Proceedings of the 2nd DARPA-JFACC symposium on advances in enterprise control*. Citeseer. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.158.4664&rep=rep1&type=pdf. Accessed in August, 2014.

Insua, D. R., Rios, J., & Banks, D. (2009). Adversarial risk analysis. *Journal of the American Statistical Association*, *104*(486), 841–854.

Joint Chiefs of Staff. (1996). Joint doctrine for military deception. Joint Publication 3–13.4 Available at http://www.dtic.mil/doctrine/jel/new_pubs/jp3_13_4.pdf. Accessed in August, 2014.

Mail Online. (2010). Ink bomb defused 'with 17 minutes to spare': Device at UK airport was ready to explode. Available at http://www.dailymail.co.uk/news/article-1326552/Yemen-ink-bomb-defused-17-minutes-spare-Device-ready-explode.html. Accessed in August, 2014.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. New York, NY: Oxford University Press.

National Commission on Terrorist Attacks Upon the United States. (2004). *The 9/11 commission report: Final report of the national commission on terrorist attacks upon the United States*. W. W. Norton and Company, New York, NY.

Powell, R. (2007). Allocating defensive resources with private information about vulnerability. *The American Political Science Review*, *101*(4), 799–809.

Powell, R. (2009). Sequential, nonzero-sum "Blotto": Allocating defensive resources prior to attack. *Games and Economic Behavior*, *67*(2), 611–615.

Roberson, B. (2006). The colonel blotto game. *Economic Theory*, *29*(1), 1–24.

Sandler, T., & Siqueira, K. (2006). Global terrorism: Deterrence versus pre-emption. *Canadian Journal of Economics*, *39*(4), 1370–1387.

Schelling, T. C. (Ed.). (1966). *Arms and influence*. Yale University Press, New Haven, CT.

Shan, X., & Zhuang, J. (2013a). Cost of equity in homeland security resource allocation in the face of a strategic attacker. *Risk Analysis*, *33*(6), 1083–1099.

Shan, X., & Zhuang, J. (2013b). Hybrid defensive resource allocations in the face of partially strategic attackers in a sequential defender-attacker game. *European Journal of Operational Research*, *228*(1), 262–272.

Swire, P. P. (2001). What should be hidden and open in computer security: Lessons from deception, the art of war, law, and economic theory. ArXiv Computer Science e-prints cs/0109089.

US Department of Homeland Security. (2011). Risk management fundamentals—homeland security risk management doctrine. Available at http://www.dhs.gov/xlibrary/assets/rma-risk-management-fundamentals.pdf. Accessed in August, 2014.

Wang, C., & Bier, V. M. (2013). Expert elicitation of adversary preferences using ordinal judgments. *Operations Research*, *61*(2), 372–385.

Zangwill, W. I., & Kantor, P. B. (1998). Toward a theory of continuous improvement and the learning curve. *Management Science*, *44*(7), 910–920.

Zhuang, J. (2010). Impacts of subsidized security on stability and total social costs of equilibrium solutions in an n-player game with errors. *The Engineering Economist*, *55*(2), 131–149.

Zhuang, J., & Bier, V. M. (2007). Balancing terrorism and natural disasters-defensive strategy with endogenous attacker effort. *Operations Research*, *55*(5), 976–991.

Zhuang, J., & Bier, V. M. (2010). Reasons for secrecy and deception in homeland-security resource allocation. *Risk Analysis*, *30*(12), 1737–1743.

Zhuang, J., & Bier, V. M. (2011). Secrecy and deception at equilibrium, with applications to anti-terrorism resource allocation. *Defence and Peace Economics*, *22*(1), 43–61.

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