Two-phase Attacks in Security Games

Andrzej Nagórko[1,2] **Paweł Ciosmak**[2] **Tomasz Michalak**[2,3]

[1]Department of Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland
[2]Ideas NCBR, ul. Chmielna 69, 00-801 Warsaw, Poland
[3]Department of Computer Science, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland

Abstract

A standard model of a security game assumes a one-off assault during which the attacker cannot update their strategy even if new actionable insights are gained in the process. In this paper, we propose a version of a security game that takes into account a possibility of a two-phase attack. Specifically, in the first phase, the attacker makes a preliminary move to gain extra information about this particular instance of the game. Based on this information, the attacker chooses an optimal concluding move. We derive a compact-form mixed-integer linear program that computes an optimal strategy of the defender. Our simulation shows that this strategy mitigates serious losses incurred to the defender by a two-phase attack while still protecting well against less sophisticated attackers.

1 INTRODUCTION

In a classic economic model of a Stackelberg game [Von Stackelberg, 1934], the leader chooses his strategy first, and while doing this, he is observed by the followers, who can adjust their response accordingly. In the last two decades, this model has received significant attention in the context of security applications, where a defender (the leader in the Stackelberg game) distributes limited security resources to guard a set of targets against an attacker (the follower in the Stackelberg game). For instance, Stackelberg games were applied in such domains as infrastructure security (ARMOR [Pita et al., 2009], IRIS [Tsai et al., 2009], PROTECT [Shieh et al., 2012]), green security (PAWS [Yang et al., 2014], MIDAS [Haskell et al., 2014]), opportunistic crimes (TRUSTS [Yin et al., 2012]), as well as cybersecurity [Zhang and Malacaria, 2021]. In all these contexts, Stackelberg games are often called *security games*.

The attack in security games is typically modeled as a one-off assault during which the attacker has no chance to update their strategy even if new valuable information is gained in the process. This, however, does not cover certain tactics that can be applied by ever more agile covert organizations. In particular, given the improvements in border control technologies that result in significant quantities of cocaine being seized in Latin America and Europe, drug cartels have to look for more innovative smuggling methods and routes. Unfortunately, according to a report by the European Monitoring Center for Drugs and Drug Addiction [European Monitoring Center for Drugs and Drug Addiction, 2016, p. 4]: "*These groups are innovative and skilled in switching and modifying both trafficking routes and* modi operandi *to circumvent law enforcement activities. They are quick to identify and exploit new opportunities for cocaine trafficking (...) shift transit routes and storage points to capitalize on the presence of ineffective border controls.*" To look for such new routes and access points, in the first phase of an operation, drug cartels can send "low-profile" couriers that carry small amounts of drugs whose key goal is to gain information. In the second phase, given the extra insight, the decision is made on which routes should be chosen for transports of much larger quantities and value. This paper stems from an observation that most of the existing models are vulnerable to such two-phase attacks which may have significant security repercussions.

Against this background, we propose a security game that takes into account a possibility of a two-phase attack. Specifically, in the first phase, the attacker makes a preliminary move designed to gain extra information on the defender's activities in this particular instance of the game. Next, in the second phase, this insight is used to choose an optimal concluding move. Given this new model we characterize optimal strategies and expected payoffs of both the defender and the attackers. We also derive a compact-form quadratic programming optimization problem to compute optimal strategies, with an exponential reduction in size compared to a possible reduction to a standard Bayesian Stackelberg

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

game. We derive an effective mixed integer linearization of the quadratic formulation. Moreover we show that a strategy computed with our model mitigates serious losses of the defender from a two-phase attack while still protecting well against less sophisticated attackers. Finally we experimentally compare the time complexity of the three solutions of two-phase Bayesian Stackelberg games discussed in this paper: a mixed quadratic linear program, a mixed integer linear program and a "normal-form" transformation to a single-phase Bayesian Stackelberg game.

2 MOTIVATION: PROBING UKRAINIAN BORDER BY BELARUS

A recent real-world example of the tactics that are explicitly modeled in our two-phase game are the actions of Lukashenko's regime in Belarus which exploits immigrants to probe the border with Ukraine. According to Special Operations Forces of the Ukraine's National Resistance Center Romanenko [2022]: *"Belarusian border guards deliberately send refugees from Iran and Pakistan to Ukrainian borders in order to search for vulnerable areas. In this way, the Belarusians check vulnerable and insufficiently protected areas of the border with Ukraine, which can be used for the passage of enemy armed forces. The enemy uses similar tactics on the border with Latvia."* This callus behaviour puts the lives of the immigrants in extreme danger both due to very difficult terrain and the on-going war. In more details, Ukraine's northwestern border of nearly 900 km is a heavily forested area full of forbidding wetlands and the Chernobyl Exclusion Zone. On top of that, the border—that was crossed by the Russian army in February 2022 and then subsequently restored by the Ukrainian counteroffensive—is now heavily fortified with trenches, walls and mine fields.

Unfortunately, despite that the border is now one of the most dangerous in the world, the Belarusian border guards organize and coordinate the groups of immigrants to attempt to cross it. The aim is to uncover and disorganise Ukrainian defences that have to react to any such attempt due to the threat from Russian saboteurs. Given the sophisticated electronic protection measures, most of such border crossing are detected. However, this does not mean that the border is impenetrable as detection does not mean that there is a patrol close enough to prevent the entry. Nevertheless, even if this particular section of the border is unmanned at the moment of entry, the Ukrainian headquarters send a team to the area. This means that a follow-up entry attempt at the same section of the border is hardly possible.

Let us consider a scaled-down version of the problem, with four sections of the Belarus-Ukraine border (S_1, S_2, S_3, and S_4) and two patrol units. This setting can be modelled as a standard security game in the spirit of the one used at the Los Angeles World Airport [Pita et al., 2009]. Pure strategies (moves) of the Ukrainian defenders are possible assignments of patrols to the sections of the border, $I = \{S_1S_2, S_1S_3, S_1S_4, S_2S_3, S_2S_4, S_3S_4\}$.

We assume two possible types of the attacker: low- and high-profile human traffickers (type 1 and 2, respectively). The high-profile type of the attacker inflicts a much larger loss upon the defender as they organize much bigger groups. Both types have the same strategy space, i.e., an attacker of each type can either choose one of the four sections of the border or back off, i.e., $J_1 = J_2 = \{S_1, S_2, S_3, S_4, \emptyset\}$. The payoffs of both parties, depending on the attacker type, increase linearly with S_i: for a high-profile attackers payoffs for successful attack are 50, 100, 150 and 200 respectively and for a low-profile attacker the payoffs are five times smaller. Attacker payoffs for unsuccessful attack are negative at the same scale. The defender payoffs are opposite, with small random noise added uniformly from interval $[-5, 5]$.

Assuming that probabilities of attacks by these two types are $p_1 = 0.8$ for the low-profile attacker and $p_2 = 0.2$ for the high-profile one, an optimal strategy for the defender is:

$$(x_{S_1S_2}, x_{S_1S_3}, x_{S_1S_4}, x_{S_2S_3}, x_{S_2S_4}, x_{S_3S_4}) = (0\%, 50\%, 0\%, 0\%, 50\%, 0\%).$$

According to this strategy, border sections S_1 and S_2 are never protected simultaneously. Such a situation is typical for Stackelberg equilibria in one-phase games and can be easily exploited by performing a two-phase attack.

A two-phase attack: Let us now assume that, unknown to the defender, the attacker has the resources and the capabilities of both the low-profile human trafficker and the high-profile one, and they are able to try two sections of the border sequentially, in phases. Given the optimal strategy derived above, let us assume that, in the first phase, a low-profile human trafficker tries to breach the border at section S_1. This provides valuable information to the attacker, irrespective of how the defender is positioned. This is because the attacker knows now a conditional probability distribution of defender's resources.

In our computation we assumed that the attacker could not attack the same target twice (which was modeled by setting second-phase payoffs for repeating the same attack to minus infinity). This was motivated by the border-patrolling scenario: a small-scale attack (provocation) elicits border patrol's response; the information gained by the attacker is the response time (they learn whether patrol was close by or not) and they could not attack safely at the same place again.

Let $t \in \{0\%, 17\%, 33\%, 50\%, 67\%, 83\%, 100\%\}$ be a chance of encountering a two-phase attacker, $(1-t) \cdot 80\%$ be a probability of encountering a low-profile single-phase attacker and $(1-t) \cdot 20\%$ be a likelihood of encountering a high-profile single-phase attacker. For $t = 0\%$ this is the

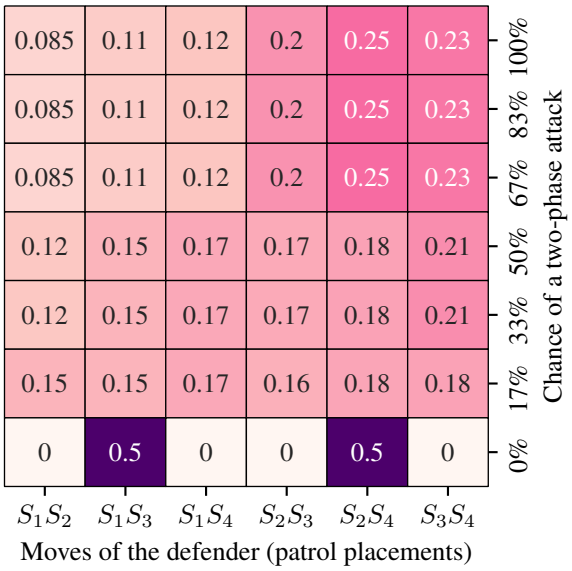

Figure 1: Each row presents an optimal mixed strategy of the defender against a group of attackers with a given chance of encountering a two-phase attack. As we can see in the last row, without presence of two-phase attackers the Stackelberg equilibrium heavily over-fits to the random noise in payoff matrices.

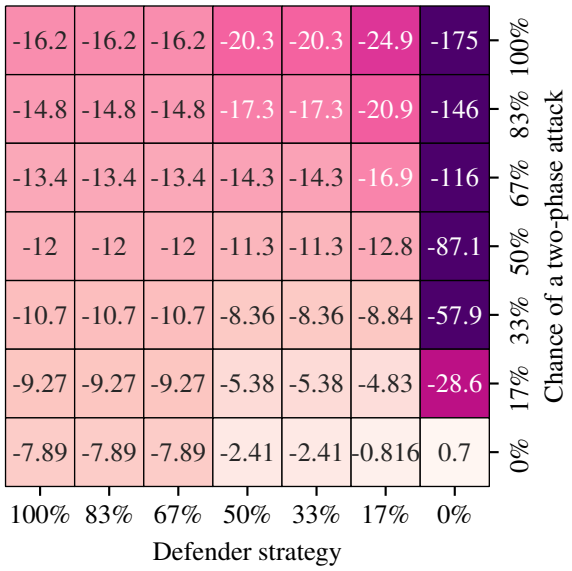

Figure 2: Expected defender payoff when playing a strategy from Figure 1 against a given chance of a two-phase attack. As we can see in the last column, the loss incurred by playing a strategy that ignores the possibility of a two-phase attack is an order of magnitude larger than over-cautious protection against such attacks.

standard one-phase model, while $t = 100\%$ describes a pure two-phase attack.

Figure 1 shows that presence of two-phase attackers significantly alters the Stackelberg equilibrium of the game. For example, for 33% probability of a two-phase attack (with 53% chance of a single-phase low-profile attack and 13% chance of a single-phase high-profile attack, keeping the $4:1$ low- to high-profile ratio), the optimal defender strategy becomes

$$(x_{S_1S_2}, x_{S_1S_3}, x_{S_1S_4}, x_{S_2S_3}, x_{S_2S_4}, x_{S_3S_4}) = (12\%, 15\%, 17\%, 17\%, 18\%, 21\%).$$

As we see in Figure 1, two-phase Stackelberg equilibria are much more robust against changes of attacker profiles.

Figure 2 shows how defender payoffs change against different compositions of attacker groups. For example, the expected payoff of the defender $\mathrm{E}(R) = \mathbf{0.7}$ against a single-phase attack drops to $-\mathbf{175}$ when single-phase strategy is pitted against a two-phase attacker.

In order to fix this flaw, we propose a new model which allows for considering one-phase and two-phase attackers simultaneously. With our security model, the expected payoff against coordinated attackers jumps from **-175** to **-16.2** (the defender is still at a disadvantage). The optimal strategy:

$$(x_{S_1S_2}, x_{S_1S_3}, x_{S_1S_4}, x_{S_2S_3}, x_{S_2S_4}, x_{S_3S_4}) = (8.5\%, 11\%, 12\%, 20\%, 25\%, 23\%)$$

forces the low-profile attacker to attack S_1 and the high-profile attacker to back off if S_1 was not patrolled. Note that this comes at a cost: for the uncoordinated (one-phase) attack, when low- and high-profile attackers act independently, this strategy brings payoff **-7.89** to the defender (a drop from **0.7**).

3 PRELIMINARIES

In the Bayesian Stackelberg game, the defender plays against a group of attackers of n distinct types. In each round, the defender plays against a single attacker and encounters the attacker of type $1 \leq t \leq n$ randomly, with probability p_t. Attackers may have different sets of moves at their disposal that inflict different damage to the defender.

Let I denote the set of defender's moves. In the Bayesian Stackelberg game, the defender picks his mixed strategy x first. Here $x = \{x_i\}_{i \in I}$ is a probability measure on I, which we denote by $x \in \mathrm{Prob}(I)$ with $\mathrm{Prob}(I) = \{x \colon I \to \mathbb{R} \colon \sum_{i \in I} x_i = 1, x \geq 0\}$. Strategy x does not depend on t as the defender doesn't know the type of attacker he will encounter. Let J_t denote the set of moves

of attacker of type t. Attacker t picks his strategy $y^t = y^t(x) \in \mathrm{Prob}(J_t)$ second, with the knowledge of the defender's strategy x. In each round of the game, both players move independently, according to strategies x and $y^t(x)$ they picked prior. Let $r_{i,t,j}$ denote the defender's payoff if he played move $i \in I$ against the attacker of type $1 \leq t \leq n$ who played a move $j \in J_t$. Let $c_{i,t,j}$ denote attacker's payoff (which may be different from $-r_{i,t,j}$). Attacker t picks an optimal strategy $\overline{y}^t = \overline{y}^t(x)$ that depends on strategy x known by him and that maximizes his expected payoff $\overline{c} = \sum_{i \in I} \sum_{j \in J_t} x_i \overline{y}^t_j c_{i,t,j}$. This payoff is maximized by a pure strategy, i.e., $\overline{y}^t$ is optimal if and only if $\overline{c} \geq \sum_{i \in I} x_i c_{i,t,j}$ for each $j \in J_t$. The defender acts to maximize his expected payoff against the optimal strategies of the attackers, i.e. he picks an optimal strategy $\overline{x}$ that maximizes his expected payoff $\sum_{i \in I} \sum_{t=1}^{n} \sum_{j \in J_t} p_t x_i \overline{y}^t_j r_{i,t,j}$.

These observations coupled with a linearization technique lead to a mixed integer linear programming formulation of Bayesian Stackelberg games published in [Paruchuri et al., 2008] as the celebrated DOBSS algorithm.

4 OUR MODEL

Let us now describe our model of a two-phase security game, which is an extension of the model of Bayesian Stackelberg games specified above in Section 3.

In a **two-phase Bayesian Stackelberg game** the defender picks his mixed strategy $x \in \mathrm{Prob}\, I$, where I denotes the set of possible defender's moves. Then the attacker of type t (encountered with probability p_t) picks his first-phase mixed strategy $y^t(x) \in \mathrm{Prob}\, J_t$ with the knowledge of defender's strategy x, where J_t denotes the set of possible first-phase moves of attacker of type t. After both the defender and the attacker make their moves $i \in I$ and $j \in J_t$ independently according to probability distributions x and $y^t(x)$ the attacker learns his first-phase payoff $c_{i,t,j}$. This narrows a possible range of moves that the defender played. With this information the attacker picks his second-phase mixed strategy $z^{t,j,c_{i,t,j}}(x) \in \mathrm{Prob}(K_t)$, where K_t denotes the set of possible second-phase moves of attacker of type t and makes his second-phase move $k \in K_t$ according to this probability distribution. The outcome of the game for the defender is $r_{i,t,j} + r'_{i,t,j,k}$, where r denotes the first-phase defender's payoff and r' denotes the second-phase one. The outcome for the attacker is $c_{i,t,j} + c'_{i,t,j,k}$, where c' is the second phase payoff.

In the above scenario, we assume that the attacker is much more agile than the defender, who picked his move (e.g., patrolling routes) for a period of time. Still, the defender wishes to maximize his expected payoff $\mathrm{E}(r + r')$ even if the attacker can gain partial information about the defender's position i with a small-scale attack.

4.1 EXPECTED PAYOFFS

The set of all possible play-outs in a two-phase game is

$$\Omega = \{(i,t,j,k) \colon i \in I, 1 \leq t \leq n, j \in J_t, k \in K_t\}\,.$$

Let us introduce following random variables on Ω: X - the defender's move; T - the attacker's type; Y - the attacker's first move; Z - the attacker's second move; C - the attacker's first-phase payoff; C' - the attacker's second-phase payoff; R - the defender's first-phase payoff; R' - the defender's second-phase payoff. Note that variables are defined on Ω so, for example, R is evaluated on (i,t,j,k) but it is equal to $r_{i,t,j}$ and is independent of k. We have

$$\begin{aligned} P(X = i) = x_i, P(T = t) = p_t, \\ P(Y = j | T = t) = y^t_j(x), \\ P(Z = k | T = t, Y = j, C = c) = z^{t,j,c}_k(x,y). \end{aligned}$$

The functional dependency $y^t(x)$ of y^t on x means that y is picked with the knowledge of strategy x. Similarly for dependency $z^{t,j,c}(x,y)$ of $z^{t,j,c}$ on x and y. From now on, for simplicity, we will write y^t and $z^{t,j,c}$.

Using this notation, we can write the expected payoff of the defender:

$$\mathrm{E}(R + R') = \sum_{(i,t,j,k) \in \Omega} x_i p_t y^t_j z^{j,c_{i,t,j}}_k \left(R(i,t,j,k) + R'(i,t,j,k)\right), \tag{1}$$

as well as the expected payoff of the attacker:

$$\mathrm{E}(C + C') = \sum_{(i,t,j,k) \in \Omega} x_i p_t y^t_j z^{j,c_{i,t,j}}_k \left(C(i,t,j,k) + C'(i,t,j,k)\right). \tag{2}$$

Let $\mathscr{C}_{t,j} = \{c_{i,t,j} \colon i \in I\}$. Given the defender's strategy x, the attacker's of type t best response maximizes his payoff:

$$(\overline{y}^t, \overline{z}^{t,j,c} \colon j \in J_t, c \in \mathscr{C}_{t,j}) \in \operatorname*{arg\,max}_{y^t \in \mathrm{Prob}(J_t), z^{t,j,c} \in \mathrm{Prob}(K_t)} \{\mathrm{E}(C + C' | T = t)\}\,.$$

Note that there is a single first-phase strategy y^t for attacker of type t and multiple second-phase strategies $z^{t,j,c}$ that depend on first-phase move $j \in J_t$ and first-phase reward $c \in \mathscr{C}_{t,j}$ obtained by the attacker. Assuming perfect rationality of the attacker, the defender adjusts his strategy to maximize his own payoff:

$$\overline{x} \in \operatorname*{arg\,max}_{x \in \mathrm{Prob}(I)} \left\{ \sum_{t=1}^{n} p_t\, \mathrm{E}(R + R' | T = t) \colon (y^t, z^{t,j,c}) \in \operatorname*{arg\,max}_{y,z} \{\mathrm{E}(C + C' | T = t)\} \right\}.$$

4.2 OPTIMAL STRATEGIES

Let $I_{t,j,c} = \{i \in I \colon c_{i,t,j} = c\}$ and $\mathscr{C}_{t,j} = \{c_{i,t,j} \colon i \in I\}$.

Proposition 4.1. Assume that attacker of type $1 \le t \le n$ played a first-phase move $j \in J_t$ against defender's strategy $x \in \mathrm{Prob}(I)$ and learned his first-phase payoff $c \in \mathscr{C}_{t,j}$. His second move strategy $z^{t,j,c}$ is optimal if and only if it maximizes

$$\mathrm{E}(C'|T=t, Y=j, C=c) = \\ = \frac{1}{\sum_{i \in I_{t,j,c}} x_i} \sum_{k \in K_t} z_k^{t,j,c} \cdot \left(\sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \right).$$

Hence any strategy $z^{t,j,c}$ that distributes probability among moves $k \in K$ with maximal $\sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k}$ is optimal. There always exists an optimal *pure* strategy $z^{t,j,c}$, i. e. without a loss of generality, we may assume that an optimal attacker's strategy satisfies $z_k^{t,j,c} \in \{0,1\}$ for each $k \in K_t$.

Proof. We have

$$P(C=c|T=t, Y=j) = \sum_{i \in I_{t,j,c}} x_i,$$

$$P(X=i, Z=k, C=c|T=t, Y=j) \\ = \begin{cases} x_i z_k^{t,j,c} & \text{if } c_{i,t,j} = c \\ 0 & \text{otherwise .} \end{cases}$$

Therefore we have

$$E(C'|T=t, Y=j, C=c) = \\ \sum_{i \in I} \sum_{k \in K_t} c'_{i,t,j,k} P(X=i, Z=k|T=t, Y=j, C=c) = \\ \sum_{k \in K_t} \frac{\sum_{i \in I} c'_{i,t,j,k} P(X=i, Z=k, C=c|T=t, Y=j)}{P(C=c|T=t, Y=j)} \\ = \frac{1}{\sum_{i \in I_{t,j,c}} x_i} \sum_{k \in K_t} \sum_{i \in I_{t,j,c}} c'_{i,t,j,k} x_i z_k^{t,j,c},$$

as claimed. □

Corollary 4.1. Assume that attacker of type $1 \le t \le n$ played a fist-phase move $j \in J_t$ against defender's strategy $x \in \mathrm{Prob}(I)$. The expected defender's second-phase payoff against an optimal attacker's strategy is

$$\mathrm{E}(R'|T=t, Y=j) = \sum_{c \in \mathscr{C}_{t,j}} \sum_{i \in I_{t,j,c}} x_i r'_{i,t,j,k_{t,j,c}},$$

where

$$k_{t,j,c} \in \arg\max_{k \in K_t} \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k}.$$

Proposition 4.2. Assume that in the first-phase attacker of type $1 \le t \le n$ plays against the defender's strategy $x \in \mathrm{Prob}(I)$. His first move strategy y^t is optimal if and only if it maximizes

$$E(C + C'|T=t) = \\ \sum_{j \in J_t} y_j^t \left(\sum_{i \in I} x_i c_{i,t,j} + \sum_{c \in \mathscr{C}_{t,j}} \max_{k \in K_t} \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \right).$$

Hence any strategy y^t that distributes probability among moves $j \in J_t$ with maximal

$$\sum_{i \in I} x_i c_{i,t,j} + \sum_{c \in \mathscr{C}_{t,j}} \max_{k \in K_t} \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k}$$

is optimal. There always exists a pure optimal strategy y^t, i.e., without a loss of generality, we may assume that an optimal attacker's first-phase strategy satisfies $y_j^t \in \{0,1\}$ for each $j \in J_t$.

Proof. We have

$$\mathrm{E}(C|T=t) = \sum_{i \in I} \sum_{j \in J_t} x_i y_j^t c_{i,t,j}.$$

On the other hand

$$\mathrm{E}(C'|T=t) = \sum_{j \in J_t} \sum_{i \in I} y_j^t x_i \,\mathrm{E}(C'|T=t, Y=j, X=i) = \\ = \sum_{j \in J_t} y_j^t \left(\sum_{i \in I} x_i \left(\sum_{k \in K_t} z_k^{t,j,c_{i,t,j}} c'_{i,t,j,k} \right) \right) = \\ = \sum_{j \in J_t} y_j^t \left(\sum_{c \in \mathscr{C}_{t,j}} \sum_{i \in I_{t,j,c}} \sum_{k \in K_t} z_k^{t,j,c} x_i c'_{i,t,j,k} \right) = \\ = \sum_{j \in J_t} y_j^t \left(\sum_{c \in \mathscr{C}_{t,j}} \sum_{k \in K_t} z_k^{t,j,c} \left(\sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \right) \right).$$

Since $\mathrm{E}(C|T=t)$ does not depend on $z^{t,j,c}$, optimal strategies $z^{t,j,c}$ should be chosen as to maximize $E(C'|T=t)$. Since each value $z_k^{t,j,c}$ appears exactly once in the formula, it is enough if we set $z_k^{t,j,c} = 1$ next to the largest coefficient for each j and c. Hence with optimal attacker's response, we have

$$\mathrm{E}(C'|T=t) = \sum_{j \in J_t} y_j^t \left(\sum_{c \in \mathscr{C}_{t,j}} \max_{k \in K_t} \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \right)$$

We are done, since $\mathrm{E}(C + C'|T=t) = \mathrm{E}(C|T=t) + \mathrm{E}(C'|T=t)$. □

Corollary 4.2. Assume that the defender's strategy $x \in \mathrm{Prob}(I)$ is played against an attacker of type $1 \le t \le n$. Then the expected defender's payoff against an optimal attacker's strategy is

$$E(R+R'|T=t) = \left(\sum_{i \in I} x_i r_{i,t,j_t} \right) + \mathrm{E}(R'|T=t, Y=j_t),$$

where

$$j_t \in \arg\max_{j \in J_t} \left(\sum_{i \in I} x_i c_{i,t,j} + \sum_{c \in \mathscr{C}_{t,j}} \max_{k \in K_t} \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \right)$$

and $E(R'|T = t, Y = j_t)$ is the payoff computed in the statement of Corollary 4.1.

Remark 4.1. The defender's payoffs computed in Corollaries 4.2 and 4.1 depend on the choices of j_t's and $k_{t,j,c}$, respectively. If a multiple choices are possible, we assume, following the existing literature, a choice that maximizes the defender's payoff.

4.3 SOLVING TWO-PHASE GAMES

Using Proposition 4.2 and Proposition 4.1 we can derive the following quadratic programming solution of two-phase security games.

$$\underset{x_i, y_j^t, z_k^{t,j,c}, \gamma_{t,j,c}}{\text{maximize}} \quad \sum_{t=1}^{n} \sum_{i \in I} \sum_{j \in J_t} p_t x_i y_j^t \times \left(r_{i,t,j} + \sum_{k \in K_t} z_k^{t,j,c_{i,t,j}} r'_{i,t,j,k} \right) \tag{4a}$$

subject to

$$\sum_{i \in I} x_i = 1, \tag{4b}$$

$$\sum_{j \in J_t} y_j^t = 1 \text{ for each } 1 \le t \le n, \tag{4c}$$

$$\sum_{k \in K_t} z_k^{t,j,c} = 1 \text{ for each } 1 \le t \le n, j \in J_t, c \in \mathscr{C}_{t,j}, \tag{4d}$$

$$\gamma_{t,j,c} \ge \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \quad \text{for each } 1 \le t \le n, j \in J_t, c \in \mathscr{C}_{t,j}, k \in K_t, \tag{4e}$$

$$\sum_{k \in K_t} \sum_{i \in I} z_k^{t,j,c_{i,t,j}} x_i c'_{i,t,j,k} \ge \sum_{c \in \mathscr{C}_{t,j}} \gamma_{t,j,c} \quad \text{for each } 1 \le t \le n, j \in J_t, \tag{4f}$$

$$\sum_{j \in J_t} y_j^t \left(\sum_{i \in I} x_i c_{i,t,j} + \sum_{c \in \mathscr{C}_{t,j}} \gamma_{t,j,c} \right) \ge \sum_{i \in I} x_i c_{i,t,j} + \sum_{c \in \mathscr{C}_{t,j}} \gamma_{t,j,c} \text{ for each } 1 \le t \le n, j \in J_t, \tag{4g}$$

$$x_i, y_j^t, z_k^{t,j,c} \ge 0, \gamma_{t,j,c} \in \mathbb{R}. \tag{4h}$$

Using a linearization technique we derive a mixed integer linear programming solution (3a) listed in Figure 3. The details of the derivation of quadratic programming formulation (4a) and mixed integer linear programming formulation (3a) are published in the appendix. Note that a two-phase security game may be expressed in extensive form and technique from [Bosansky and Cermak, 2015] may be used to derive a mixed-integer linear programming formulation of similar size.

5 COMPARISON TO THE STANDARD BAYESIAN STACKELBERG GAMES

In this section, we show that it is possible to reduce a two-phase Bayesian Stackelberg game to a one-phase Bayesian Stackelberg game using a transformation that is similar to a Harsanyi normal-form transformation. However, this reduction results in an exponential explosion of the problem size. Using equations (2) and (1) and the observation that the attackers have optimal *pure* strategies, we can write the following mixed quadratic linear problem that solves two-phase Bayesian Stackelberg games.

$$\text{maximize} \quad \mathrm{E}(R + R') \tag{5a}$$

subject to

$$x \in \mathrm{Prob}(I), y^t \in \mathrm{Prob}(J_t), z^{t,j,c} \in \mathrm{Prob}(K_t) \tag{5b}$$

$$\mathrm{E}(C + C') \ge \sum_{i \in I} x_i c_{i,t,j} + \sum_{i \in I} x_i c'_{i,t,j,k_{j,c_{i,j}}} \quad \text{for each } 1 \le t \le n, j \in J_t, \{k_{j,c}\}_{c \in \mathscr{C}_{t,j}} \subset K_t \tag{5c}$$

Note the possibly exponential number of conditions of type (5c), as we have to consider all possible combinations of first-phase move j and second-phase moves k_c depending on the first-phase outcome c.

Note that the set of attacker moves in a regular (single-phase) Bayesian Stackelberg game is: $J'_t = \bigcup_{j \in J_t} \{j\} \times K_t^{\mathscr{C}_{t,j}}$, where $K_t^{\mathscr{C}_{t,j}}$ is a set of all maps from $\mathscr{C}_{t,j}$ to K_t, i. e. a choice of a move k_c from K_t for each possible first-phase outcome $c \in \mathscr{C}_{t,j}$. The size of the set of follower's moves grows exponentially, $|J'| = \sum_{j \in J} |K|^{|\mathscr{C}_j|}$, and so does the number of constraints (5c). Compare this to the MIQP formulation (4a), where we have a polynomial number $|J||K| + 2|J| = \sum_{j \in J}(|K| + 2)$ of constraints that correspond to the optimality of the follower's actions.

6 EXPERIMENTAL EVALUATION

We evaluated performance of the three algorithms considered in the paper: mixed quadratic linear program (MQLP), mixed integer linear program (MILP) and DOBSS applied to a single-phase problem transformed from a two-phase form. We also verified that observation of Section 2 (Figures 1 and 2) that a defensive strategy computed against a single-phase attack is vulnerable to a two-phase attack is universal.

$$\underset{x_i, y_j^t, z_k^{t,j,c}, \gamma_{t,j,c}, s_{i,t,j,k}, u_{t,j,c}, w_{i,t,j,k}}{\text{maximize}} \quad \sum_{\substack{1 \le t \le n, i \in I, \\ j \in J_t, k \in K_t}} p_t \left(r_{i,t,j} + r'_{i,t,j,k} \right) w_{i,t,j,k} \tag{3a}$$

subject to

$$\sum_{i \in I} x_i = 1, \tag{3b}$$

$$\sum_{j \in J_t} y_j^t = 1 \qquad \text{for each } 1 \le t \le n, \tag{3c}$$

$$\sum_{k \in K_t} z_k^{t,j,c} = 1 \qquad \text{for each } 1 \le t \le n, j \in J_t, c \in \mathscr{C}_{t,j}, \tag{3d}$$

$$\gamma_{t,j,c} \ge \sum_{i \in I_{t,j,c}} x_i c'_{i,t,j,k} \qquad \text{for each } 1 \le t \le n, j \in J_t, c \in \mathscr{C}_{t,j}, k \in K_t, \tag{3e}$$

$$\sum_{k \in K_t} \sum_{i \in I} s_{i,t,j,k} c'_{i,t,j,k} \ge \sum_{c \in \mathscr{C}_{t,j}} \gamma_{t,j,c} \qquad \text{for each } 1 \le t \le n, j \in J_t, \tag{3f}$$

$$\sum_{m \in J_t} \left(\left(\sum_{i \in I} \left(\sum_{k \in K_t} w_{i,t,m,k} \right) c_{i,t,m} \right) + \sum_{c \in \mathscr{C}_{t,m}} u_{t,m,c} \right) \ge \sum_{i \in I} x_i c_{i,t,j} + \sum_{c \in \mathscr{C}_{t,j}} \gamma_{t,j,c} \qquad \text{for each } 1 \le t \le n, j \in J_t, \tag{3g}$$

$$s_{i,t,j,k} \le z_k^{t,j,c_{i,t,j}} \qquad \text{for each } i \in I, 1 \le t \le n, j \in J_t, k \in K_t, \tag{3h}$$

$$\sum_{k \in K_t} s_{i,t,j,k} = x_i \qquad \text{for each } i \in I, 1 \le t \le n, j \in J_t, \tag{3i}$$

$$-M(1 - y_j^t) \le u_{t,j,c} - \gamma_{t,j,c} \le M(1 - y_j^t) \qquad \text{for each } 1 \le t \le n, j \in J_t, c \in \mathscr{C}_{t,j}, \tag{3j}$$

$$-My_j^t \le u_{t,j,c} \le My_j^t \qquad \text{for each } 1 \le t \le n, j \in J_t, c \in \mathscr{C}_{t,j}, \tag{3k}$$

$$\sum_{j \in J_t} \sum_{k \in K_t} w_{i,t,j,k} = x_i \qquad \text{for each } i \in I, 1 \le t \le n, \tag{3l}$$

$$w_{i,t,j,k} \le y_j^t \qquad \text{for each } i \in I, 1 \le t \le n, j \in J_t, k \in K_t, \tag{3m}$$

$$w_{i,t,j,k} \le z_k^{t,j,c_{i,t,j}} \qquad \text{for each } i \in I, 1 \le t \le n, j \in J_t, k \in K_t, \tag{3n}$$

$$x_i, s_{i,t,j,k}, w_{i,t,j,k} \ge 0, \gamma_{t,j,c}, u_{t,j,c} \in \mathbb{R}, y_j^t, z_k^{t,j,c} \in \{0,1\}. \tag{3o}$$

Figure 3: A mixed-integer linear programming formulation of two-phase security games.

	1	2	3	4	5	6	7
1	0.009	0.012	0.013	0.013	0.015	0.019	0.023
2	0.010	0.015	0.062	0.777	0.897	1.015	1.816
3	0.015	0.023	0.584	0.621	1.959	1.232	1.693
4	0.014	0.037	0.297	1.343	1.248	1.454	2.698
5	0.009	0.028	0.162	0.605	1.450	2.166	2.501
6	0.010	0.031	0.320	0.744	1.503	7.702	13.679
7	0.113	0.107	0.430	1.115	1.907	3.627	7.968

(a) Performance of the MILP formulation, in sec.

	1	2	3	4	5	6	7
1	0.008	0.009	0.009	0.010	0.012	0.015	0.019
2	0.007	0.023	0.235	1.044	2.139	3.958	6.003
3	0.013	0.072	0.968	2.845	4.910	10.309	30.679
4	0.014	0.102	1.124	6.962	14.434	-	-
5	0.008	0.076	1.099	9.569	23.458	-	-
6	0.013	0.183	1.401	34.405	-	-	-
7	0.034	0.339	1.967	40.175	-	-	-

(b) Performance of the MQLP formulation, in sec.

	1	2	3	4	5	6	7
1	0.007	0.009	0.009	0.012	0.018	0.031	0.041
2	0.006	0.012	0.055	0.352	1.627	6.524	20.422
3	0.012	0.043	1.567	-	-	-	-
4	0.006	0.173	-	-	-	-	-
5	0.007	0.743	-	-	-	-	-
6	0.007	3.024	-	-	-	-	-
7	0.024	-	-	-	-	-	-

(c) Performance of the DOBSS formulation, in sec.

Table 1: Performance of MQLP, MILP and DOBSS. Row and column headers display number of defender and attacker moves. Averaged over 4 runs. Time limit 60 seconds.

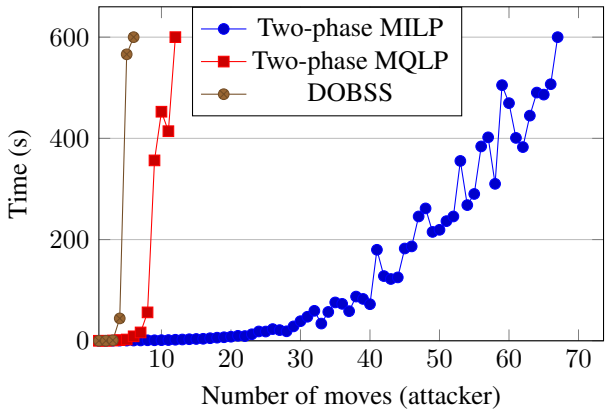

Figure 4: Running time (averaged over 20 runs) for random problems with 3 defender's moves against a number of attacker's moves marked on the x axis. Two-phase attack with 2 attacker types. Time limit 600 seconds.

	$(3,1)$	$(3,2)$	$(4,2)$	$(5,2)$
Constant	33%, 33%	50%, 0%	100%, 0%	100%, 0%
Linear	41%, 32%	50%, 30%	100%, 0%	46%, 0%
Exponential	43%, 32%	50%, 30%	100%, 0%	48%, 0%

Table 2: Chance that defender's optimal strategy is vulnerable to two-phase attack. In each column, number on the left shows chance for mixed strategy computed against one-phase attack with DOBSS. Number on the right shows chance for mixed strategy with model proposed in the paper. Lower is better.

6.1 TIME COMPLEXITY OF OPTIMAL STRATEGY COMPUTATIONS IN TWO-PHASE MODELS

Table 1 shows comparison of three algorithms that compute optimal strategies in two-phase Bayesian Stackelberg games. Table 1a shows solution times for mixed quadratic linear programming (MQLP) formulation (4a). Table 1b shows solution times for mixed integer linear programming (MILP) formulation (3a). Finally, Table 1c shows solution times for DOBSS applied to one-phase problem obtained with a transformation described in Section 5.

In each table, the row number is the number of defender's moves and the column number is the number of attacker's moves. The time is measured in seconds and was averaged over 4 independent runs. Cases where time limit of 60 seconds was reached are marked with '-'.

Table 4 shows analogous comparison for larger numbers of attacker moves and two attacker types. The computation was performed with SCIP solver on a single core of Intel Xeon 3.60GHz processor.

6.2 COMPARISON OF STRATEGIES AGAINST TWO-PHASE ATTACKS

Example from Section 2 shows that a defensive strategy computed against a single-phase attack is vulnerable to a two-phase attack is universal. Table 2 shows that this is universal (i.e. the example is not cherry-picked). In particular, we checked that this pattern (severe loss against a two-phase attack) always emerges for different value profiles (constant, linear, exponential) of the defended targets and over different ranges of random noise.

A defender's mixed strategy is *vulnerable* to two-phase attack if it permits a phase-one attack such that knowledge of the outcome guarantees that the second-phase attack will be successful. We considered border patrolling game with 3, 4 or 5 border segments and 1 or 2 patrols We considered three payoff profiles for the attacker: constant (successful attack of each segment of the border is of equal value to the

attacker), linear (the value grows linearly) and exponential. Results are averaged over 4 runs.

7 RELATED WORK

The literature on security games is vast and continuously growing (the surveys can be found in Sinha et al. [2018] and Fang and Nguyen [2016]). The first related body of works is on multi-stage Stackelberg games in which the attacker and defender interact in stages. In Luh et al. [1984], the authors analysed systems where players choose among pure strategies. In [Żychowski and Mańdziuk, 2022], an evolutionary algorithm for solving multi-stage Stackelberg games was proposed, whereas in [Guzman et al., 2022], an inspection game is formalized as a multiple-stage Stackelberg game. Two-stage (but not two-phase) Stackelberg games were considered in the literature, e.g., in [Anand et al., 2008, Gray et al., 2009, Kabul and Parlaktürk, 2019, Wang et al., 2022].

Our model can also be understood as a method to prevent a deception attack (see [Kar et al., 2015] for an example). To this end, let us assume that the attacker, aware that the defender relies on the DOBSS algorithm (against a one-phase attack), chooses an appropriate two-phase strategy for which the defender is unprepared. Now, if the defender uses our algorithm, the situation changes accordingly.

8 CONCLUSIONS

We introduced an extension of a standard Bayesian Stackelberg game that takes into account the possibility that an attack can consist of two phases. In our model, the attacker makes a preliminary strike in the first phase in order to gain extra intelligence about the defense. Next, the attacker is able to make a more informed choice of the concluding move. The model is motivated by a pattern of attacks observed on the Belarus-Ukraine border.

The usual setting of Stackelberg games assumes a large asymmetry between the defender and the attacker: on one hand, it is assumed that the attacker has the perfect knowledge of the defender's past actions; on the other hand, it is assumed that the attacker has zero knowledge of the defender's current defensive position. The model proposed in the paper reduces this asymmetry: it considers scenarios where the attacker may undertake some actions to gain knowledge about the defender's current defensive position.

For this new model, we derived a compact-form MILP formulation and we showed analytically that the reduction in problem size compared to a standard approach is exponential. Our results also revealed that using the standard approach to defend against a two-phase attack can lead to severe losses on the defender's side.

References

Krishnan Anand, Ravi Anupindi, and Yehuda Bassok. Strategic inventories in vertical contracts. *Management Science*, 54(10):1792–1804, 2008.

Branislav Bosansky and Jiri Cermak. Sequence-form algorithm for computing stackelberg equilibria in extensive-form games. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

European Monitoring Center for Drugs and Drug Addiction. Perspectives on drugs. cocaine trafficking to europe, 2016. URL `https://www.emcdda.europa.eu/system/files/attachments/2641/Cocaine%20trafficking_POD2016.pdf`.

Fei Fang and Thanh H Nguyen. Green security games: Apply game theory to addressing green security challenges. *ACM SIGecom Exchanges*, 15(1):78–83, 2016.

John V Gray, Brian Tomlin, and Aleda V Roth. Outsourcing to a powerful contract manufacturer: The effect of learning-by-doing. *Production and Operations Management*, 18(5):487–505, 2009.

Cristobal Guzman, Javiera Riffo, Claudio Telha, and Mathieu Van Vyve. A sequential stackelberg game for dynamic inspection problems. *European journal of operational research*, 302(2):727–739, 2022.

William Haskell, Debarun Kar, Fei Fang, Milind Tambe, Sam Cheung, and Elizabeth Denicola. Robust protection of fisheries with compass. In *Twenty-Sixth IAAI Conference*, 2014.

Mustafa O Kabul and Ali K Parlaktürk. The value of commitments when selling to strategic consumers: A supply chain perspective. *Management Science*, 65(10):4754–4770, 2019.

Debarun Kar, Fei Fang, Francesco Delle Fave, Nicole Sintov, and Milind Tambe. " a game of thrones" when human behavior models compete in repeated stackelberg security games. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1381–1390, 2015.

Peter B. Luh, Shi-Chung Chang, and Tsu-Shuan Chang. Solutions and properties of multi-stage stackelberg games. *Automatica*, 20(2):251–256, 1984. ISSN 0005-1098.

Praveen Paruchuri, Jonathan P Pearce, Janusz Marecki, Milind Tambe, Fernando Ordonez, and Sarit Kraus. Playing games for security: An efficient exact algorithm for solving bayesian stackelberg games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 895–902, 2008.

James Pita, Manish Jain, Fernando Ordónez, Christopher Portway, Milind Tambe, Craig Western, Praveen Paruchuri, and Sarit Kraus. Using game theory for los angeles airport security. *AI magazine*, 30(1):43–43, 2009.

V. Romanenko. Belarus uses migrants for intelligence on the border with Ukraine, 2022. URL `https://www.pravda.com.ua/eng/news/2022/12/6/7379514/`.

Eric Shieh, Bo An, Rong Yang, Milind Tambe, Craig Baldwin, Joseph DiRenzo, Ben Maule, and Garrett Meyer. Protect: A deployed game theoretic system to protect the ports of the united states. In *Proceedings of the 11th international conference on autonomous agents and multiagent systems-volume 1*, pages 13–20. Citeseer, 2012.

Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, and Milind Tambe. Stackelberg security games: Looking beyond a decade of success. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 2018.

Jason Tsai, Shyamsunder Rathi, Christopher Kiekintveld, Fernando Ordonez, and Milind Tambe. Iris-a tool for strategic security allocation in transportation networks. *AAMAS (Industry Track)*, pages 37–44, 2009.

Heinrich Von Stackelberg. *Marktform und gleichgewicht*. J. springer, 1934.

Xinyu Wang, Suresh P Sethi, and Shuhua Chang. Pollution abatement using cap-and-trade in a dynamic supply chain and its coordination. *Transportation Research Part E: Logistics and Transportation Review*, 158:102592, 2022.

Rong Yang, Benjamin J Ford, Milind Tambe, and Andrew Lemieux. Adaptive resource allocation for wildlife protection against illegal poachers. In *Aamas*, pages 453–460, 2014.

Zhengyu Yin, Albert Xin Jiang, Matthew P Johnson, Christopher Kiekintveld, Kevin Leyton-Brown, Tuomas Sandholm, Milind Tambe, and John P Sullivan. Trusts: Scheduling randomized patrols for fare inspection in transit systems. In *Twenty-Fourth IAAI Conference*, 2012.

Yunxiao Zhang and Pasquale Malacaria. Bayesian stackelberg games for cyber-security decision support. *Decision Support Systems*, 148:113599, 2021.

Adam Żychowski and Jacek Mańdziuk. Coevolutionary approach to sequential stackelberg security games. In *International Conference on Computational Science*, pages 103–117. Springer, 2022.