2022 American Control Conference (ACC)
Atlanta, USA, June 8-10, 2022

Security Resource Investment Optimization for Critical Infrastructure Systems: A Game-Theoretic Approach

Joe Clanin[1] and Sourabh Bhattacharya[1,2]

Abstract— **Motivated by the problem of optimal security resource deployment in critical infrastructure systems, we study a non-zero-sum security game over the substations of a power system in which the player payoffs depend upon the maturity of the security resources at each substation according to NERC-CIP standards. Extending previous work, we give a structural characterization of the possible types of Nash equilibria in our non-zero-sum additive security game model, present feasibility conditions for equilibria of each type, and propose a novel algorithm to compute an equilibrium. Utilizing our characterization of the possible equilibria in additive security games, we propose a method to obtain a suboptimal solution to the problem of maximizing the expected outcome to the system operator by varying the maturity of security resources deployed at each substation and demonstrate the method by simulation.**

I. INTRODUCTION

The pervasive introduction of cyber systems into critical infrastructure domains has rendered the task of defense against intrusion by malicious actors extraordinarily complex. System operators must decide how to allocate their limited security resources across elaborate cyber-physical systems to reduce the impact of increasingly frequent and larger scale intrusions into both the cyber and physical layers. Recent high profile cyberattacks have caused major disruptions to critical infrastructure in the oil and gas [18], meat processing [2], and chemical distribution [3] industries. Increased concern among experts about the scope of the vulnerability of power grids in the U.S. to large-scale cyber attacks [14] has highlighted the urgency with which new security measures must be developed and deployed.

A *security game* is a type of resource allocation game between attackers and defenders over a set of targets. Applications of game theory to security scenarios have been extensively studied (e.g [5], [13], [15]) and successfully deployed in the real world. For example, the United States Federal Air Marshals Service, Coast Guard, and Transportation Security Administration have implemented game-theoretic models to defend critical infrastructure [9], [10] by determining optimal security resource deployment strategies such as assignment of a limited number of air marshals to flights or law enforcement personnel to patrol routes. A zero-sum security game model is proposed in [8] for contingency analysis in power grids. Real world attack and defense scenarios, however, are in general not necessarily zero-sum games [17] due to a variety of factors such as the diverse incentives of attackers and defenders [19]. The present work concerns simultaneous move one-shot two-player security games of complete information between a single defender (representing a system operator) and a single attacker. Under the assumption that player utilities are additive over attacked targets, [11] proposed an algorithm to compute a Nash equilibrium in such a game in the case of a single attacker resource and this work was later generalized to the case of multiple attacker resources in [12]. The special case of zero-sum additive security games under a restricted payoff model was studied in [7] in which a structural characterization of the Nash equilibria is given along with an algorithm to compute the equilibrium strategies in such games in linear time. Extending this approach to non-zero-sum security games, [6] generalizes this structural characterization to non-zero-sum games under the same restricted payoff model and gives a quadratic time algorithm for equilibrium computation. In this work, we further generalize the structural approach in [6], [7] to the general payoff model studied in [12].

The contributions of our current work are as follows: (1) The structural analysis [6], [7] of Nash equilibria in additive security games is extended to a more general payoff structure given in [12] and a characterization of the possible types of equilibria in such games is presented along with feasibility conditions for equilibria of each type. (2) Utilizing our characterization of the possible types of equilibria, we present a novel algorithm for computing an equilibrium and its type. (3) Motivated by the non-zero-sum nature of real security scenarios, we formulate a non-zero-sum additive security game over the substations of a power system and propose a method to vary the maturity of security resource deployment at each substation to address the problem of optimizing defender expected outcome at equilibrium.

The rest of the paper is organized as follows. In section II, we present the problem formulation. In section III, we give a structural characterization of possible types equilibria in our additive security game model, and a suboptimal solution to the optimization problem. Section IV presents simulation of our approach. Finally, Section V presents our conclusions and proposes several directions for future work.

II. PRELIMINARIES

A. Additive Security Games

We define a two player game between an attacker and defender over a target set $T = \{1, \ldots, m\}$. The attacker will attack $1 \leq k_a < m$ targets and the defender will defend

This material is based upon work supported by the National Science Foundation under Grant No. #1739969

[1]Department of Computer Science, [2] Department of Mechanical Engineering, Iowa State University, Ames, IA.
{jsc, sbhattac}@iastate.edu

$1 \leq k_d < m$ targets. We assume that at most one resource is allocated to a target, an assumption introduced in [12]. If a target is defended, we say the target is 'covered', and if a target is not defended we shall say this target is 'uncovered'. Let $a_t^u > 0$ denote the payoff to the attacker when he attacks the uncovered target t, and let $a_t^c > 0$ denote his payoff when he attacks the covered target t. Similarly, let $b_t^u > 0$ and $b_t^c > 0$ denote the payoffs to the defender when target t is attacked and is uncovered and covered, respectively. We assume the payoffs satisfy $\Delta_a(t) = a_t^u - a_t^c > 0, \Delta_d(t) = b_t^c - b_t^u > 0$. That is, when a target is attacked, it is better for the defender that the target is defended, and better for the attacker that it is exposed. Let $s_1^a, \ldots, s_{\binom{m}{k_a}}^a$ be the pure strategies of the attacker and let $s_1^d, \ldots, s_{\binom{m}{k_d}}^d$ be the pure strategies of the defender. We assume that the payoff matrices A and B, for the attacker and defender respectively, satisfy the following additive property:

$$A_{ij} = \sum_{t \in s_i^a \cap s_j^d} a_t^c + \sum_{t \in s_i^a \setminus s_j^d} a_t^u, \quad B_{ij} = \sum_{t \in s_i^a \cap s_j^d} b_t^c + \sum_{t \in s_i^a \setminus s_j^d} b_t^u. \quad (1)$$

This security game formulation coincides with the security games considered in [12]. By convention we assume the attacker is the row player. Suppose p and q are mixed strategies for the attacker and defender respectively. Using (1), we can write the expected payoffs as

$$\begin{aligned} v_a &= p^T A q = \sum_{t=1}^{m} [\alpha_t a_t^c \beta_t + \alpha_t a_t^u (1 - \beta_t)] \\ v_d &= p^T B q = \sum_{t=1}^{m} [\alpha_t b_t^c \beta_t + \alpha_t b_t^u (1 - \beta_t)] \end{aligned} \quad (2)$$

Where $\alpha = [\alpha_1, \ldots, \alpha_m]^T$ and $\beta = [\beta_1, \ldots, \beta_m]^T$ are the attack and defense probability vectors (respectively) given by $\alpha_i = \sum_{i \in S_t^a} p_t$ and $\beta_i = \sum_{i \in S_t^d} q_t$ We consider α and β to be mixed-strategy actions of the players and arrive at the following formulation of Nash equilibrium:

Definition 1. *A pair (α^*, β^*) is a Nash Equilibrium of a security game $(\mathcal{X}, \mathcal{Y}, A, B)$ if and only if any feasible deviation from α^* (β^*) does not lead to a better payoff for the attacker (defender). That is $v_a \leq v_a^*$ ($v_d \leq v_d^*$).*

Note that $\alpha_i, \beta_i \in [0, 1]$ (p, q are probability vectors) and

$$\sum_{i=1}^{m} \alpha_i = k_a \quad \text{and} \quad \sum_{i=1}^{m} \beta_i = k_d \quad (3)$$

We have that $v_a, v_d : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ are linear functions of α and β. We shall use the notation $\nabla_\alpha v_a|_{\alpha^*, \beta^*}$ to denote the partial derivative of v_a with respect to α computed at (α^*, β^*). We denote the standard basis for $\mathbb{R}^m$ by $e_1, \ldots, e_m$.

B. Problem Formulation: Optimization of Defender Expected Utility

We shall examine the following problem: Consider a power network with m substations $1, \ldots, m$. The system operator will prevent the simultaneous loss of k_d substations, and an attacker will choose to attack k_a substations.

TABLE I: Definition of the sets $I_1, \ldots, I_9$.

β_i^* \ α_i^*	$= 0$	$\in (0, 1)$	$= 1$
$= 0$	$i \in I_1$	$i \in I_2$	$i \in I_3$
$\in (0, 1)$	$i \in I_4$	$i \in I_5$	$i \in I_6$
$= 1$	$i \in I_7$	$i \in I_8$	$i \in I_9$

Associated to each substation are d *security domains* given by the NERC-CIP standards [1] and the Department of Energy Cybersecurity Capability Maturity Model (C2M2). Each domain at a substation is scored from 0 to 3 according to the maturity of the cybersecurity resources deployed at the substation. The summation of the scores from each domain at a substation give the Maturity Indicator Level (MIL) of the substation. Examples of security domains selected from the NERC-CIP standards include Personnel and Training, Password Protection, Physical Security, Recovery Plans for Cyber Systems, Information Protection, and Systems Security Management [1][16]. We formulate a non-zero-sum security game over the set of substations of the power system in which the payoffs to players for attacks on uncovered targets are given by $a_i^u = I_i - \text{MIL}_i$ and $b_i^u = I_i + \text{MIL}_i$. where I_i is a quantity representing the impact to the grid of the loss of substation i (without loss of generality, we assume $I_i > 3d$). Intuitively, the payoff to the defender when an uncovered target is attacked increases as the maturity of the cybersecurity resources is increased and the payoff to the attacker in such an instance is decreased. We formulate the following optimization problem:

$$\begin{aligned} \textbf{maximize} \quad & v_2^*(a^c, a^u, b^c, b^u) \\ \textbf{subject to} \quad & a_i^u \in [a_i^{lb}, a_i^{ub}], b_i^u \in [b_i^{lb}, b_i^{ub}] \end{aligned}$$

for given bounds $a_i^{lb} \leq a^{ub}$, $b_i^{lb} \leq b_i^{ub}$.

III. RESULTS

In this section, we extend the structural analysis of equilibria in [7], [6] to the more general utility framework presented in the previous section by giving a characterization of the equilibria in terms of necessary structural properties. We classify the possible equilibria into two types and present an algorithm to compute an equilibrium in terms of target attack/defense probability, its type, and the expected outcome of the game at equilibrium. Finally, we present a procedure to obtain a suboptimal solution to the problem of optimizing the defender expected outcome at equilibrium.

A. Structural Properties of Equilibria

For any (α^*, β^*), let $I_1, \ldots, I_9$ denote the partition of T defined by Table I. Note that for any $i, j \in T$ we have

$$\begin{aligned} (e_i - e_j)^T \nabla_\alpha v_a|_{(\alpha^*, \beta^*)} &= a_i^u - \beta_i^* \Delta_a(i) - (a_j^u - \beta_j^* \Delta_a(j)), \\ (e_i - e_j)^T \nabla_\beta v_d|_{(\alpha^*, \beta^*)} &= \alpha_i^* \Delta_d(i) - \alpha_j^* \Delta_d(j). \end{aligned}$$

Our analysis is based upon the observation that if (α^*, β^*) is an equilibrium and the attacker (resp. defender) can deviate

by shifting probability from target j to target i, then $(e_i - e_j)^T \nabla_\alpha v_a|_{(\alpha^*,\beta^*)} \leq 0$ (resp. $(e_i - e_j)^T \nabla_\beta v_d|_{(\alpha^*,\beta^*)} \leq 0$).

Lemma 1. *If $I_4 \cup I_7 \neq \emptyset$ then $I_2 \cup I_3 \cup I_5 \cup I_6 = \emptyset$.*

Proof. Suppose $i \in I_4 \cup I_7$. Then $\alpha_i^* = 0$ and $\beta_i^* \in (0, 1]$. As $k_a \geq 1$ there exists j so that $\alpha_j > 0$. We must have $\beta_j = 1$ for all j such that $\alpha_j > 0$ (else the defender has a feasible deviation increasing their payoff). □

This yields the following theorem.

Theorem 1. *An equilibrium (α^*, β^*) is of precisely one of the following two types:*

- *Type I: $\forall i \in T$, $i \in T \setminus (I_4 \cup I_7)$*
- *Type II: $\forall i \in T$, $i \in I_1 \cup I_4 \cup I_7 \cup I_8 \cup I_9$, $I_4 \cup I_7 \neq \emptyset$*

Lemma 2. *Suppose (α^*, β^*) is an equilibrium.*

(i) There exists a constant c_1 so that $\forall i \in I_2 \cup I_5 \cup I_8$ $a_i^u - \beta_i^* \Delta_a(i) = c_1$. Furthermore, for all $i \in I_2$ we have $c_1 = a_i^u$ and for all $i \in I_8$ we have $c_1 = a_i^c$.

(ii) There exists a constant c_2 so that $\forall i \in I_5 \cup I_6$ $\alpha_i^* \Delta_d(i) = c_2$. Furthermore, for $i \in I_6$ we have $c_2 = \Delta_d(i)$.

Proof. For $i, j \in I_2 \cup I_5 \cup I_8$ we must have $(e_i - e_j)^T \nabla_\alpha v_a|_{(\alpha^*,\beta^*)} = 0$ which implies $a_i^u - \beta_i^* \Delta_a(i) = a_j^u - \beta_j^* \Delta_a(j)$. The remaining assertions follow from similar arguments and the definitions of the sets $I_1, \ldots, I_9$. □

We make the following assumption:

Assumption 1. *The parameters $a_1^c, \ldots, a_m^c, a_1^u, \ldots, a_m^u$ are distinct and the parameters $\Delta_d(1), \ldots, \Delta_d(m)$ are distinct.*

Under this assumption, we obtain the following structural results regarding equilibria of Type I.

Lemma 3. *Suppose (α^*, β^*) is an equilibrium.*

(i) $\Delta_d(i) < \Delta_d(j) < \Delta_d(k)$ for $i \in I_3$, $j \in I_6$, $k \in I_9$

(ii) $\alpha_i^* \Delta_d(i) \leq \alpha_j^* \Delta_d(j) \leq \alpha_k^* \Delta_d(k)$ for $i \in I_2$, $j \in I_5$, $k \in I_8$

(iii) $\Delta_d(i) > \Delta_d(j)$ for $i \in I_5 \cup I_8$, $j \in I_6$ and $\Delta_d(k) < \Delta_d(\ell)$ for $k \in I_3, \ell \in I_8$.

(iv) $a_i^u < a_j^u < a_k^u$ and $a_j^u < a_\ell^u$ for $i \in I_1$, $j \in I_2$, $k \in I_3$, $\ell \in I_5 \cup I_6 \cup I_9$.

(v) $a_i^u - \beta_i^* \Delta_a(i) \leq a_j^u - \beta_j^* \Delta_a(j)$ for $(i, j) \in I_2 \times I_3, I_5 \times I_6, I_8 \times I_9$ with strict inequality for $j \in I_3 \cup I_9$

(vi) $a_i^u < a_j^u$ for $i \in I_1$ and $j \in I_2 \cup I_3 \cup I_5 \cup I_6 \cup I_8 \cup I_9$

(vii) $a_i^c < a_j^c$ for $i \in I_2 \cup I_5 \cup I_8$, $j \in I_9$ and $a_k^c < a_\ell^c$ for $k \in I_5$ $\ell \in I_8$.

Proof. We prove the left inequality in (i). Every other assertion follows from similar arguments and the fact that $\Delta_a(i), \Delta_d(i) > 0$ for all $i \in T$. Suppose that $i \in I_3$ and $j \in I_6$. We must have $(e_i - e_j) \nabla_\beta v_d|_{(\alpha^*,\beta^*)} = \Delta_d(i) - \Delta_d(j) \leq 0$ and therefore $\Delta_d(i) < \Delta_d(j)$ by Assumption 1. □

B. Equilibrium Computation

Under Assumption 1, by Lemma 2 it follows that at most one of I_2 and I_8 is nonempty, and I_2, I_6, I_8 contain at most one element. When these sets are nonempty, write $I_2 = \{j^{[2]}\}, I_6 = \{j^{[6]}\}, I_8 = \{j^{[8]}\}$. Label the possible cases as follows: When $I_6 = \emptyset$, the equilibrium is of type I.A and otherwise is of type I.B. In both cases, define subtypes i,ii, and iii corresponding to the cases in which $I_2, I_8 = \emptyset$; $I_2 \neq \emptyset$, $I_8 = \emptyset$; and $I_2 = \emptyset$, $I_8 \neq \emptyset$ respectively. Order the set of targets such that $a_i^u < a_j^u$ for $i < j$. Define parameters $r = |I_1|$, $s = |I_3|$, $t = |I_9|$. For each possible value of the parameters, we check the feasibility of an equilibrium constructed according to the necessary structural properties given in the previous subsection. We now give the details of the construction for each subtype of type I equilibrium in terms of a, b, k_a, k_d, r, s, t. By Lemma 3 (vi), in every case and for any r we set $I_1 = \{1, \ldots, r\}$.

Type I.A.i: By Lemma 2(ii), Lemma 3(i), we have that $I_3 = \{j_1, \ldots, j_s\}$ where $j_1, \ldots, j_s$ are the targets of least $\Delta_d(i)$ in $T \setminus I_1$. From Lemma 3(vii), we have $I_9 = \{\ell_1, \ldots, \ell_t\}$ where $\ell_1, \ldots, \ell_t$ are the t targets of greatest a_j^c in $T \setminus (I_1 \cup I_3)$. Note $t \leq k_a - s$. Finally, we set $I_5 = T \setminus (I_1 \cup I_3 \cup I_9)$. We have $\alpha_i = c_2/\Delta_d(i)$ and $\beta_i = (a_i^u - c_1)/\Delta_a(i)$ for all $i \in I_5$ where

$$c_1 = \frac{t - k_d + \sum_{j \in I_5} \frac{a_j^u}{\Delta_a(j)}}{\sum_{j \in I_5} \frac{1}{\Delta_a(j)}} \text{ and } c_2 = \frac{k_a - s - t}{\sum_{j \in I_5} \frac{1}{\Delta_d(j)}} \quad (4)$$

Type I.A.ii: By Lemma 3(iv) we set $I_2 = \{r + 1\} = \{j^{[2]}\}$. Then, by Lemma 2(ii), Lemma 3(i) we set $I_3 = \{j_1, \ldots, j_s\}$ where $j_1, \ldots j_s$ are the s targets of least $\Delta_d(i)$ in $T \setminus (I_1 \cup I_2)$. By Lemma 3(viii) we set $I_9 = \{\ell_1, \ldots, \ell_t\}$ where $\ell_1, \ldots, \ell_t$ are the t targets of largest a_j^c in $T \setminus (I_1 \cup I_2 \cup I_3)$. Note $t \leq k_a - s$. Finally $I_5 = T \setminus (I_1 \cup I_2 \cup I_3 \cup I_9)$. Now, for each $i \in I_5$ we have

$$\alpha_i = \frac{k_a - s - t - \alpha_{j^{[2]}}}{\Delta_d(i) \sum\limits_{j \in I_5} \frac{1}{\Delta_d(i)}} \text{ and } \beta_i = \frac{a_i^u - a_{j^{[2]}}^u}{\Delta_a(i)}. \quad (5)$$

Type I.A.iii: By Lemma 3(i)(iii) we set $I_3 = \{j_1, \ldots, j_s\}$ where $j_1, \ldots, j_s$ are the s targets of least $\Delta_d(i)$ in $T \setminus I_1$. By Lemma 3(vii) we set $I_9 = \{\ell_1, \ldots, \ell_t\}$ where $\ell_1, \ldots, \ell_t$ are the t targets of greatest a_j^c in $T \setminus (I_1 \cup I_3)$ and $I_8 = \{j^{[8]}\}$ where $j^{[8]}$ is the target of greatest a_j^c in $T \setminus (I_1 \cup I_3 \cup I_9)$. Finally $I_5 = T \setminus (I_1 \cup I_3 \cup I_8 \cup I_9)$. Now, for each $i \in I_5$ we have

$$\alpha_i = \frac{k_a - s - t - \alpha_{j^{[8]}}}{\Delta_d(i) \sum\limits_{j \in I_5} \frac{1}{\Delta_d(j)}} \text{ and } \beta_i = \frac{a_i^u - a_{j^{[8]}}^c}{\Delta_a(i)}. \quad (6)$$

The constructions for Type I.B are similar to those just outlined for type I.A, but take into account the inequalities in Lemma 3 involving I_6. For reasons of space, we have omitted these calculations here. We now derive feasibility conditions for each subtype of type I equilibrium. As noted in [12], (α^*, β^*) is a Nash Equilibrium if and only if there exist constants c_1 and c_2 such that for all $i \in T$, $\alpha_i^* > 0 \implies \beta_i^* a_i^c + (1 - \beta_i^*) a_i^u \leq c_1$ and $\alpha_i^* < 1 \implies \beta_i^* a_i^c + (1 - \beta_i^*) a_i^u \geq c_1$ as well as $\beta_i^* > 0 \implies \alpha_i^* \Delta_d(i) \geq c_2$ and $\beta_i^* < 1 \implies \alpha_i^* \Delta_d(i) \leq c_2$ (i.e. both players are playing a best response). Based upon this observation, for any Type I

TABLE II: Feasibility conditions for Type I equilibria

I.A.i	• $\forall i \in I_5\ 0 < \alpha_i, \beta_i < 1$
I.A.ii	• $\exists 0 < \alpha_{j^{[2]}} < \min_{i\in I_5}\{\alpha_i\Delta_d(i)/\Delta_d(j^{[2]})\}$ such that $\forall i \in I_5\ 0 < \alpha_i, \beta_i < 1$ • $\forall i \in I_5 \cup I_9\ a^c_{j^{[2]}} < a^c_i$
I.A.iii	• $\exists \max_{i\in I_5}\{\alpha_i\Delta_d(i)/\Delta_d(j^{[8]})\} < \alpha_{j^{[8]}} < 1$ such that $\forall i \in I_5\ 0 < \alpha_i, \beta_i < 1$
I.B.i	• $\exists 0 < \beta_{j^{[6]}} < 1$ so that $\forall i \in I_5$, $\beta_{j^{[6]}} \leq \frac{a^u_{j^{[6]}} - a^u_i + \beta_i\Delta_a(i)}{\Delta_a(j^{[6]})}$ and $0 < \alpha_i, \beta_i < 1$
I.B.ii	• For all $i \in I_5$, $0 < \alpha_i, \beta_i < 1$ • $0 < \alpha_{j^{[2]}} \leq \min_{i\in I_5}\{\alpha_i\Delta_d(i)/\Delta_d(j^{[2]})\}$ • $0 < \beta_{j^{[6]}} \leq \min_{i\in I_5}\left\{\frac{a^u_{j^{[6]}} - a^u_i + \beta_i\Delta_a(i)}{\Delta_a(j^{[6]})}\right\}$
I.B.iii	• For all $i \in I_5\ 0 < \alpha_i, \beta_i < 1$ • $0 < \beta_{j^{[6]}} \leq \min_{i\in I_5}\left\{\frac{a^u_{j^{[6]}} - a^u_i + \beta_i\Delta_a(i)}{\Delta_a(j^{[6]})}\right\}$ • $\max_{i\in I_5}\{\alpha_i\Delta_d(i)/\Delta_d(j^{[8]})\} \leq \alpha_{j^{[8]}} < 1$

equilibrium we have

$$\begin{aligned} &c_1 \geq \max\left\{\max_{i\in I_3} a^u_i, \max_{i\in I_9} a^c_i, \beta_{j^{[6]}}a^c_{j^{[6]}} + (1-\beta_{j^{[6]}})a^u_{j^{[6]}}\right\}, \\ &c_1 \leq \min_{i\in I_1} a^u_i, \end{aligned} \tag{7}$$

$$\begin{aligned} &c_2 \geq \max\left\{\max_{i\in I_3}\Delta_d(i), \alpha_{j^{[2]}}\Delta_d(j^{[2]})\right\}, \\ &c_2 \leq \min\left\{\min_{i\in I_9}\Delta_d(i), \alpha_{j^{[8]}}\Delta_d(j^{[8]})\right\}. \end{aligned} \tag{8}$$

A given $\Gamma(r, s, t, \texttt{type})$ is feasible if it satisfies (7) and (8) as well as the additional conditions required by Lemma 3 given in Table II. A Type I equilibrium (when it exists) can be computed by simply checking the feasibility of $\Gamma(r, s, t, \texttt{type})$ for all possible values of the parameters.

We now consider how to compute a Type II equilibrium. First note that by arguments exactly similar to those in the proof of Lemma 3 we have $a^c_i < a^c_j$ for $i \in I_1 \cup I_4 \cup I_7$ and $j \in I_9$. By Lemma 2, Assumption 1 and (3), $I_8 = \emptyset$. By the aforementioned observation about c_1, there must exist a constant c_1 such that $a^c_i \leq c_1 \leq a^c_j$ for $i \in I_9$ and $j \in I_7$. By Assumption 1, we have $I_7 = \emptyset$. Thus, in a Type II equilibrium, all targets belong to I_1, I_4 and I_9. By (3) and the fact that $\alpha_i > 0$ implies $i \in I_9$ and thus $\alpha_i = 1$ in a Type II equilibrium, we set $I_9 = \{m - k_a + 1, \ldots, m\}$. Note that a Type II equilibrium is feasible only when $k_d > k_a$, as I_4 must be nonempty. Now, we must have $\max_{i\in I_9} a^c_i \leq c_1 \leq \beta_j a^c_j + (1-\beta_j)a^u_j$ for all $j \in I_4$. For each $i = 1, \ldots, m - k_a$, let $\epsilon_i = \min\{1, (a^u_j - \max_{i\in I_9} a^c_i)/\Delta_a(i)\}$. Note that we must have $\sum_{i\in I_4}\beta_i = k_d - k_a$. There exists a feasible assignment of the remaining β_i, $i = 1, \ldots, m - k_d$ if $\sum_{1\leq i\leq m-k_d, \epsilon_i>0}\epsilon_i \geq k_d - k_a$. When this condition is satisfied, a Type II equilibrium exists and can be computed using this construction.

C. Optimization of Defender Expected Utility

We now outline a procedure to, given an initial state as input, increase v_d for each subtype of Type I equilibrium while preserving the type. Note that in a Type II equilibrium there are no successful attacks and thus the defender expected utility is optimized by maximal increase of b^c_i for $i \in I_9$. For a given r, s, t and $\texttt{type}$ let $v^5_d(r, s, t, \texttt{type})$ denote the quantity $c_2\sum_{i\in I_5}\left[\frac{b^u_i}{\Delta_d(i)} + \frac{a^u_i - c_1}{\Delta_a(i)}\right]$ where c_1, c_2, I_5 are constructed with parameters r, s, t and the construction in the previous subsection for $\texttt{type}$. Similarly, let $v^{3,9}_d(r, s, t, \texttt{type})$ denote $\sum_{i\in I_3} b^u_i + \sum_{i\in I_9} b^c_i$ where I_3 and I_9 are constructed according to $r, s, t, \texttt{type}$. By Lemma 3, the expression (2) for v_d, and the definitions of the sets $I_1, \ldots, I_9$, we have that $v^5(r, s, t, \texttt{type})$ is the contribution to v_d by targets in I_5 and $v^{3,9}_d(r, s, t, \texttt{type})$ is the contribution to v_d by targets in I_3 and I_9 for Type I equilibria. When $r, s, t, \texttt{type}$ are clear from context, abbreviate these quantities by v^5_d and $v^{3,9}_d$.

By Lemma 3, a perturbation of $a^u_i, b^u_i, i = 1, \ldots, m$ which does not change the parameters r, s, t or subtype of Type I equilibrium must not violate the following inequalities:

(a) $\max_{i\in I_3}\Delta_d(i) < \min_{i\in I_6}\Delta_d(i)$ (b) $\max_{i\in I_6}\Delta_d(i) < \min_{i\in I_9}\Delta_d(i)$
(c) $\alpha_{j^{[2]}}\Delta_d(j^{[2]}) \leq \min_{i\in I_5}\alpha_i\Delta_d(i)$
(d) $\max_{i\in I_5}\alpha_i\Delta_d(i) \leq \alpha_{j^{[8]}}\Delta_d(j^{[8]})$
(e) $\Delta_d(j^{[6]}) < \min_{i\in I_5\cup I_8}\Delta_d(i)$ (f) $\max_{i\in I_3}\Delta_d(i) < \Delta_d(j^{[8]})$
(g) $\max_{i\in I_1} a^u_i < a^u_{j^{[2]}} < \min_{i\in I_3} a^u_i$ (h) $a^u_{j^{[2]}} < \min_{i\in I_5\cup I_6\cup I_9} a^u_i$
(i) $\max_{i\in I}(a^u_i - \beta_i\Delta_a(i)) \leq \min_{j\in J}(a^u_i - \beta_i\Delta_a(i))$ for $(I, J) \in \{(I_2, I_3), (I_5, I_6), (I_8, I_9)\}$.
(j) $\max_{i\in I_1} a^u_i < \min_{i\in I_5\cup I_6\cup I_8\cup I_9} a^u_i$
(k) $\max_{i\in I_2\cup I_5\cup I_8} a^c_i < \min_{i\in I_9} a^c_i$ (l) $\max_{i\in I_5} a^c_i < \min_{i\in I_8} a^c_i$

Perturbations must also preserve (7) and (8), and we shall require the preservation of $\Delta_a(t), \Delta_d(t) > 0$ and Assumption 1. For each type of equilibrium, we give a procedure to perturb the b^u_i such that the sets $I_1, \ldots, I_9$ do not change. For each type the procedure will consist of a preprocessing phase in which b^u_i for $i \in T \setminus I_5$ are perturbed followed by a second phase in which we perturb the b^u_i for $i \in I_5$.

In a Type I.A.i, we have that c_1 and c_2 are given by (4) and $v^{I.A.i}_d = v^{3,9}_d + v^5_d$. Note that c_2 is an increasing function of $\Delta_d(i)$ for each $i \in I_5$, and we have

$$\frac{\partial c_1}{\partial a^u_i} = \frac{c_1 - a^c_i}{(\Delta_a(i))^2\sum_{j\in I_5}\frac{1}{\Delta_a(j)}} > 0, \tag{9}$$

for $i \in I_5$ so c_1 is a decreasing function of b_i^u for $i \in I_5$. We have that $v_d^5(r,s,t,I.A.i)$ is given by

$$\frac{k_a - s - t}{\sum_{j \in I_5} \frac{1}{\Delta_d(j)}} \left(k_d - t + \sum_{j \in I_5} \frac{b_j^u}{\Delta_d(j)} \right). \quad (10)$$

Note that for $i \in I_5$, $\partial v_d^5 / \partial b_i^u$ is given by $\frac{c_2}{(\Delta_d(i))^2} \left(b_i^c - \frac{v_d^5}{k_a - s - t} \right)$, so v_d^5 is an increasing function of any b_i^u such that $b_i^c > v_d^5/(k_a - s - t)$ (refer to such a b_i^u as type a) and a decreasing function of any b_i^u such that $b_i^c < v_d^5/(k_a - s - t)$ (refer to such a b_i^u as type b).

In a Type I.A.ii equilibrium, c_1 is simply $a_{j^{[2]}}^u$, $v_d^{I.A.ii} = v_d^{3,9} + v_d^5 + \alpha_{j^{[2]}} b_{j^{[2]}}^u$, and by (5) we see c_2 is an increasing function of $\Delta_d(i)$ for $i \in I_5$. We have that $v_d^5(r,s,t,I.A.ii)$ is given by

$$\frac{k_a - s - t - \alpha_{j^{[2]}}}{\sum_{j \in I_5} \frac{1}{\Delta_d(j)}} \left(k_d - t + \sum_{i \in I_5} \frac{b_i^u}{\Delta_d(i)} \right). \quad (11)$$

For any $i \in I_5$, inspection of $\partial v_d^5 / \partial b_i^u$ for gives that v_d^5 is an increasing function of b_i^u for any i such that $b_i^c > v_d^5/(k_a - s - t - \alpha_{j^{[2]}})$ (type a), and a decreasing function o b_i^u such that $b_i^c < v_d^5/(k_a - s - t - \alpha_{j^{[}[2]}$ (type b).

In a Type I.A.iii equilibrium, $c_1 = a_{j^{[8]}}^c$, $v_d^{I.A.iii} = v_d^{3,9} + v_d^5 + \alpha_{j^{[8]}} b_{j^{[8]}}^c$ and by (6) c_2 is again an increasing function of $\Delta_d(i)$ for $i \in I_5$. Now, $v_d^5(r,s,t,I.A.iii)$ is given by

$$\frac{k_a - s - t - \alpha_{j^{[8]}}}{\sum_{j \in I_5} \frac{1}{\Delta_d(j)}} \left(k_d - t - 1 + \sum_{i \in I_5} \frac{b_i^u}{\Delta_d(i)} \right). \quad (12)$$

For each $i \in I_5$, we again inspect $\partial v_d^5 / \partial b_i^u$ to determine that v_d^5 is an increasing function of b_i^u for any i such that $b_i^c > v_d^5/(k_a - s - t - \alpha_{j^{[8]}})$ (type a), and a decreasing function o b_i^u such that $b_i^c < v_d^5/(k_a - s - t - \alpha_{j^{[}[8]})$ (type b). We execute the following procedure for a Type I.A equilibrium:

I.A Phase 1: To accommodate the perturbation of b_i^u of both types a and b in Phase 2, we seek to both reduce the lower bound and increase the upper bound of (7) and (8). Note that in a Type I.A.ii, we have $\max_{i \in I_3} a_i^u \le c_1 = a_{j^{[2]}}^u$ by (7) and $a_{j^{[2]}}^u < \min_{i \in I_3} a_i^u$ by (g), so $s = 0$. Similarly, in a Type I.A.iii, $\max_{i \in I_3} a_i^u \le a_{j^{[8]}}^c$ and $a_{j^{[8]}}^c < a_i^u$ for $i \in I_3$ so $s = 0$. In Type I.A.i we select the least feasible $\alpha_{j^{[2]}}$ and increase $b_{j^{[2]}}^u$ to a maximal extent. In Type I.A.iii we select the greatest feasible $\alpha_{j^{[8]}}$ and increase $b_{j^{[2]}}$ to a maximal feasible extent. In any type I.A, we decrease b_i^u for $i \in I_1$ and I_9 to a maximal feasible extent. For a Type $I.A.i$, we may have $s > 0$ and in this case we also increase b_i^u for $i \in I_3$ maximally.

I.A Phase 2: For each i of type a we compute the maximal feasible increase of b_i^u (if any) and the corresponding value of v_d^5 resulting from the increase. For each i of type b, we compute the maximal feasible decrease of b_i^u (if any) and the corresponding value of v_d^5 resulting from the decrease. We select the i resulting in the greatest increase of v_d^5 and execute the corresponding perturbation. We repeat this process for as long as there exist feasible perturbations among the targets of type a and b. As each perturbation increases v_d^5, the procedure terminates after a finite number of steps.

For any Type I.B equilibrium, we have $c_2 = \Delta_d(j^{[6]})$. Note that (9) also holds in a Type I.B.i equilibrium, so c_1 is a decreasing function of b_i^u for $i \in I_5$. For a Type $I.B.i$, we have $v_d^{I.B.i} = v_d^{3,9} + v_d^5 + b_{j^{[6]}}^u - \beta_{j^{[6]}} \Delta_d(j^{[6]})$.We have that $v_d^5(r,s,t,I.B.i)$ is given by

$$\Delta_d(j^{[6]}) \left(k_d - t - \beta_{j^{[6]}} + \sum_{j \in I_5} \frac{b_j^u}{\Delta_d(j)} \right). \quad (13)$$

In a Type I.B.ii equilibrium, $c_1 = a_{j^{[2]}}^u$ and so $v_d^5(r,s,t,I.B.ii)$ is given by

$$\Delta_d(j^{[6]}) \left(k_d - t - \beta_{j^{[6]}} + \sum_{i \in I_5} \frac{b_i^u}{\Delta_d(i)} \right). \quad (14)$$

We have $v_d^{I.B.ii} = v_d^{3,9} + v_d^5 + \alpha_{j^{[2]}} b_{j^{[2]}}^u + b_{j^{[6]}}^u - \beta_{j^{[6]}} \Delta_d(j^{[6]})$. Similarly, in a Type I.B.iii we have $c_1 = a_{j^{[8]}}^c$ and thus $v_d^5(r,s,t,I.B.iii)$ is given by

$$\Delta_d(j^{[6]}) \left(k_d - t - 1 - \beta_{j^{[6]}} + \sum_{i \in I_5} \frac{b_i^u}{\Delta_d(i)} \right). \quad (15)$$

We have $v_d^{I.B.iii} = v_d^{3,9} + v_d^5 + \alpha_{j^{[8]}} b_{j^{[8]}}^c + b_{j^{[6]}}^u - \beta_{j^{[6]}} \Delta_d(j^{[6]})$.

Clearly, for any Type I.B equilibrium, v_d^5 is an increasing function of b_i^u for $i \in I_5$. We execute the following procedure for a Type I.B equilibrium:

I.B Phase 1: As c_1 and v_d^5 are decreasing and increasing functions of b_i^u for $i \in I_5$ respectively, we seek to reduce the lower bound of (7) as much as possible. Accordingly, for type I.B.i, we choose the least feasible $\beta_{j^{[6]}}$ and increase b_i^u for $i \in I_3$ as much as possible without violating (a)-(l). As v_d^5 increases with $c_2 = \Delta_d(j^{[6]})$, we then increase the upper bound of (8) by decreasing b_i^u for $i \in I_9$ maximally without violating (a)-(l) and then decrease $b_{j^{[6]}}^u$ to the maximum feasible extent.

I.B Phase 2: For each $i \in I_5$, compute the maximal feasible increase of b_i^u and the corresponding value of v_d^5 when b_i^u is increased by this increment. Select the i corresponding to the greatest increase in v_d^5 and execute the maximal increment of b_i. Repeat until no increment of b_i^u for $i \in I_5$ is feasible.

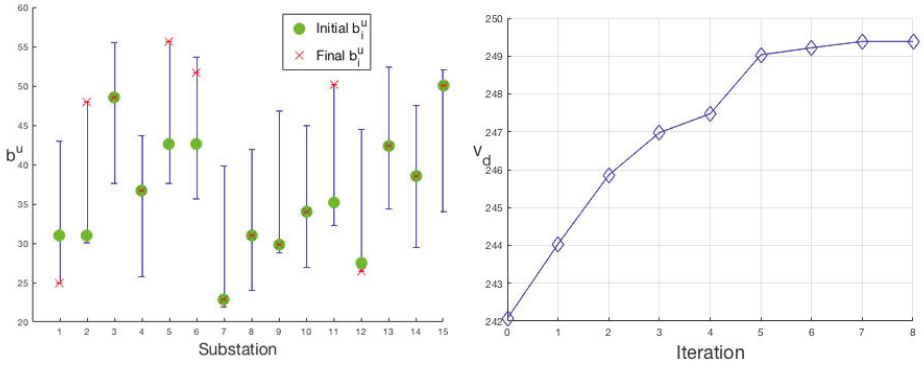
IV. Simulation and Discussion

In this section we present an example case study on a 15-substation power system. The initial state of the system used in our simulation is specified by Table III. We assume that each substation in the system has six security domains that the initial maturity indicator level of each substation is nonzero. To demonstrate our algorithm, we have generated values for the impact associated with the loss of each substation as well as payoffs to the players when an attack is launched on a covered target. In practice, such quantities are computed by taking into account physical parameters of the grid, such as the impact of loss-of-substation induced cascades [4]. Figure 1 depicts the result of the algorithm

TABLE III: 15-bus system with 6 security domains exhibiting a Type I.A.i equilibrium with k_a=4, k_d=10; $v_d \approx 242.07$

Substation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$I \approx$	24.993	30.008	37.580	25.721	37.599	35.668	21.879	23.976	28.809	26.972	32.218	26.507	34.399	29.516	34.011
MIL^{initial}	6	1	11	11	5	7	1	7	1	7	3	1	8	9	16
$a^c \approx$	3.0130	1.8306	5.9220	0.9003	6.5156	8.7267	2.6624	2.9456	3.8905	8.8047	10.1046	8.3042	2.6379	4.1097	1.4826
$b^c \approx$	55.058	78.030	90.750	86.859	100.277	98.641	78.238	67.144	76.697	69.017	72.235	50.586	78.045	72.114	68.364

Fig. 1: Initial and final values of b_i^u for the example given in Table III (left) and the path of the defender expected outcome v_d when the our method is applied to the example in Table III (right). In the final state, $v_d \approx 249.38$



proposed in the previous section applied to the case study. The procedure identifies a Type I.A.i equilibrium in the game specified by the initial system state and perturbs the values of b_i^u from their initial positions shown in green in Figure 1 to the final values shown in red.

V. Conclusion and Future Work

In this work, we have formulated the interaction between a resource-constrained attacker and defender of a power system as a non-zero-sum additive security game. In this model, we investigated the problem of maximizing the expected outcome to the defender at equilibrium by varying the maturity of the security resources deployed in the system. We propose the following directions for future work: (a) Extension of our structural approach to scenarios of incomplete information, dynamic attack scenarios, and non-additive games. (b) Further study of how the expected outcome to the defender varies as payoff perturbation violates the feasibility conditions of each equilibrium type. (c) Application of our approach to additional critical infrastructure domains for which security resource maturity standards have been developed.

References

[1] North american electric reliability corporation - critical infrastructure protection (nerc-cip) standards. https://www.nerc.com/pa/Stand/Pages/CIPStandards.aspx.

[2] Fabiana Batista, Michael Hirtzer, and Mike Dorning. All of jbs's u.s. beef plants were forced shut by cyberattack, May 2021.

[3] Melody M Bombgardner. Siegfried, brenntag, and symrise hit by cyberattacks, May 2021.

[4] Jacques Delport. *Critical Substation Risk Assessment and Mitigation*. PhD thesis, Virginia Tech, 2018.

[5] Cuong T Do, Nguyen H Tran, Choongseon Hong, Charles A Kamhoua, Kevin A Kwiat, Erik Blasch, Shaolei Ren, Niki Pissinou, and Sundaraja Sitharama Iyengar. Game theory for cyber security and privacy. *ACM Computing Surveys (CSUR)*, 50(2):1–37, 2017.

[6] Hamid Emadi. *A game-theoretic framework for contingency analysis in power systems*. PhD thesis, Ames IA, 2021.

[7] Hamid Emadi and Sourabh Bhattacharya. On the characterization of saddle point equilibrium for security games with additive utility. In *Decision and Game Theory for Security - 11th International Conference,GameSec 2020, College Park, MD, USA, October 28-30, 2020, Proceedings*, volume 12513 of *Lecture Notes in Computer Science*, pages 349–364. Springer, 2020.

[8] Hamid Emadi, Joe Clanin, Burhan Hyder, Kush Khanna, Manimaran Govindarasu, and Sourabh Bhattacharya. An efficient computational strategy for cyber-physical contingency analysis in smart grids. In *IEEE PES General Meeting*, 2021.

[9] Manish Jain, Bo An, and Milind Tambe. Security games applied to real-world: Research contributions and challenges. In *Moving Target Defense II*, pages 15–39. Springer, 2013.

[10] Manish Jain, Jason Tsai, James Pita, Christopher Kiekintveld, Shyamsunder Rathi, Milind Tambe, and Fernando Ordónez. Software assistants for randomized patrol planning for the lax airport police and the federal air marshal service. *Interfaces*, 40(4):267–290, 2010.

[11] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordóñez, and Milind Tambe. Computing optimal randomized resource allocations for massive security games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[12] Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr. Security games with multiple attacker resources. In *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[13] Xiannuan Liang and Yang Xiao. Game theory for network security. *IEEE Communications Surveys & Tutorials*, 15(1):472–486, 2012.

[14] Jim Magill. Experts say cyberattacks likely to result in blackouts in u.s., Jul 2021.

[15] Mohammad Hossein Manshaei, Quanyan Zhu, Tansu Alpcan, Tamer Bacşar, and Jean-Pierre Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys (CSUR)*, 45(3):25, 2013.

[16] U. S. Department of Energy. Electricity subsector cybersecurity capability maturity model (es-c2m2), version 2.0. https://www.energy.gov/ceser/cybersecurity-capability-maturity-model-c2m2.

[17] Robert Powell. Defending against terrorist attacks with limited resources. *American Political Science Review*, 101(3):527–541, 2007.

[18] David E. Sanger and Nicole Perlroth. Pipeline attack yields urgent lessons about u.s. cybersecurity, May 2021.

[19] Milind Tambe. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge university press, 2011.