Stochastic Games for Power Grid Protection Against Coordinated Cyber-Physical Attacks

Longfei Wei, *Student Member, IEEE,* Arif I. Sarwat, *Member, IEEE,* Walid Saad, *Senior Member, IEEE,* and Saroj Biswas, *Member, IEEE*

***Abstract*—Owing to the critical nature of the power grid, coordinated cyber-physical attacks on its critical infrastructure can lead to disastrous human and economic losses. In this paper, a stochastic game-theoretic approach is proposed to analyze the optimal strategies that a power grid defender can adopt to protect the grid against coordinated attacks. First, an optimal load shedding technology is devised to quantify the physical impacts of coordinated attacks. Taking these quantified impacts as input parameters, the interactions between a malicious attacker and the defender are modeled using a resource allocation stochastic game. The game is shown to admit a Nash equilibrium and a novel learning algorithm is introduced to enable the two players to reach such equilibrium strategies while maximizing their respective minimum rewards in a sequence of stages. The convergence of the proposed algorithm to a Nash equilibrium point is proved and its properties are studied. Simulation results of the stochastic game model on the WSCC 9-bus system and the IEEE 118-bus system are contrasted with those of static games, and show that different defense resources owned lead to different defense strategies.**

***Index Terms*—Coordinated attacks, optimal load shedding, power grid security, stochastic game theory.**

I. Introduction

The power grid is one of our nation's most critical technological infrastructures which, in turn, renders it susceptible to a range of physical and cyber attacks [1]. Physical attacks may disrupt the power plants, transmission lines, and substations of the power grid. For instance, the California electrical power substation attack [2] in April 2013 by unidentified gunmen led to not only power outages, but also underscored the vulnerability of the grid. If an attacker has enough expertise in power grid operation, monitoring and control, it can launch an attack that can result in long-term power outages over a large territory of any country. In this context, protecting the power grid against a variety of attacks is extremely challenging. The increasing growth of the grid in scale and complexity makes it impossible both financially and logistically to protect the entire infrastructure [3]. Moreover, the emergence of the smart grid, in which new communication and information technologies are integrated in the power grid has led to new cyber security concerns and new points of entry for attackers. Cyber attacks may take advantage of accessibility through the supervisory control and data acquisition (SCADA) or advanced metering infrastructure (AMI) components to attempt to remotely access, compromise, or control electronic resources. These increasing threats have particularly culminated with the discovery of the Stuxnet worm (see [4] and [5]) that infected numerous industrial control systems in 2010.

Essentially, a cyber-physical system [6]–[8] is any system that is composed of cyber elements, such as communication nodes, and physical elements that follow the laws of physics, such as any control system. In particular, the power grid is a complex cyber-physical system whose cyber elements, such as SCADA components, depend on the power supply of the physical system, such as generators, while the physical system relies on the cyber system for operation, monitoring and control. Compared with single cyber or physical attacks, highly motivated, sophisticated groups are capable of implementing coordinated cyber-physical attacks on both cyber and physical components of the power grid [1]. For example, a cyber attack can seek to disable the power grid failure detection system in order to facilitate a physical attack against a critical component of the grid. Coordinated cyber-physical attacks can have a compounded effect greater than the sum of its individual attacks. As a result, a comprehensive approach for analyzing optimal defense strategies against such coordinated attacks must utilize cyber-physical system interactions to quantify attack impacts and evaluate corresponding defense mechanisms.

Recently, a wealth of research [9]–[16] based on optimization theory, Markov decision processes (MDPs) and game theory has been proposed for defending the power grid against various attacks. An optimization-theoretic approach to protect power grid state estimators against false data injection attacks was proposed in [9], which considered the attack vectors to be linearized measurement models. This work in [9] demonstrated that in the case in which a small subset of measurements are immune to attacks, it is possible to defend against such cyber attacks. In [10], power grid attack graphs generated by off-the-shelf model checking technology were interpreted as MDPs, and the value iteration algorithm was introduced to compute the probabilities of successful attacks. However, these works ignore the decision making process of the attacker, and the solutions of these works only optimize the defender's actions using the expected rewards of these actions.

More recently, the application of game-theoretic approaches for the power grid security, such as in [11]–[16], has attracted attention due to the inherent ability of such approaches to capture the multi-faced decision making process involved in power grid security problems. A zero-sum static game model was proposed in [11] to compute optimal defense strategies that seek to protect physical infrastructures of the power grid

This work was supported by the National Science Foundation under Grant CNS-1446570, Grant CNS-1446621, and Grant CNS-1446574. Paper no. TSG-00919-2015.

L. Wei and A. I. Sarwat are with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL, USA (e-mail: lwei004@fiu.edu, asarwat@fiu.edu).

W. Saad is with Wireless@VT, the Bradley Department of Electrical and Computer Engineering, Virginia Tech, VA, USA (e-mail: walids@vt.edu). He is also International Scholar at the Department of Computer Science and Engineering, Kyung Hee University, Seoul, South Korea.

S. Biswas is with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, USA (e-mail: saroj.biswas@temple.edu).

against physical attacks. In order to protect the communication network of the power grid, a game equilibrium obtained from a zero-sum static game model between an intentional attacker and a fusion-based defender was introduced and studied in [12]. In [13], a zero-sum game-theoretic framework was formulated to investigate the interactive decision-making process between a sensor node and an attacker who can launch denial-of-service (DoS) attacks. For defending against false data injection (FDI) attacks on power grid state estimation, in [14], the least-budget defense strategy in the game equilibrium was proposed to render power grids immune to FDI attacks. In [15], a general-sum game-theoretic framework was proposed to explore and evaluate optimal defense strategies for the power grid operator to protect the grid against a combination of cyber and physical attacks.

However, the works in [11]–[15] rely on a static game formulation in which the dynamics of the power grid are ignored and the interactions between the attacker and the defender are assumed to be one-shot events. In [16], a zero-sum stochastic game was proposed for modeling single transmission line attack-defense scenarios while focusing on deriving the probabilistic strategies of the involved players. While interesting, the stochastic game model of [16] focused only on a single attack (e.g. physical or cyber). Traditionally, power grid planning techniques have accommodated $N-1$ contingency in their scope. Even if the attacker successfully compromises part of the cyber-physical power grid system, it is quite possible that no load shedding will be caused. However, great power failures could be triggered if the attacker launches coordinated attacks to compromise multiple parts or functions of the power grid in cyber and physical aspects. Compared with single attacks, coordinated attacks, when smartly structured, can not only have severe physical impacts, but can also potentially nullify the effect of system redundancy and other defense mechanisms. In a recent report by North American Electric Reliability Corporation (NERC), the coordinated attack was identified as one of the three representative high-impact, low-frequency (HILF) threats [17]. Indeed, devising new defense mechanisms against such coordinated attacks is both challenging and desirable.

The main contribution of this paper is to develop a new framework that enables one to model and analyze how a power grid can dynamically respond to coordinated cyber-physical attacks. To address this problem, we formulate a stochastic game-theoretic mechanism to characterize and defend against coordinated attacks. Compared with related game-theoretic works such as in [11]–[16], we provide several novel contributions. We identify and classify practical cyber-physical attacks faced by the power grid, and introduce a new optimal load shedding technology to quantify the amount of shed load under different types of attacks representing their physical impacts on the power grid. Using these quantified physical impacts, we develop defense mechanisms based on stochastic game models for dynamically optimizing the security and resiliency of the power grid, especially against coordinated attacks. A novel learning algorithm for the proposed stochastic game models is devised to obtain the defender's Nash equilibrium strategies that provide realistic guidelines on how to deploy limited resources for protecting critical elements of the power grid. Simulation results using the WSCC 9-bus system and the IEEE 118-bus system are presented to illustrate the proposed approach and deriving different defense resource allocation strategies under different resource limitations.

II. System Model and Problem Formulation

Consider a power grid system with $N_{\mathcal{V}}$ buses and $N_{\mathcal{E}}$ branches. This system can be modeled using an undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ with $N_{\mathcal{V}}$ vertices and $N_{\mathcal{E}}$ edges. The set of vertices $\mathcal{V} = \{v_1, v_2, ..., v_{N_{\mathcal{V}}}\}$ represents $N_{\mathcal{V}}$ nodes in the graph that can include generation plants, transformers, substation devices, and customers. The set $\mathcal{E} = \{e_1, e_2, ..., e_{N_{\mathcal{E}}}\}$ of edges encompasses $N_{\mathcal{E}}$ edges that correspond to transmission lines. Thus, the total number of elements (vertices and edges) of the system that must be protected against cyber-physical attacks is $N_{\mathcal{G}} = N_{\mathcal{V}} + N_{\mathcal{E}}$.

Consider an attacker that seeks to disrupt the system by distributing its finite attack resources over one or more elements of the graph to maximize the physical impact on the system. The resources owned by the attacker can include a) personnel or hackers that are assigned to the attack, b) technological resources such as advanced tools or malwares to disrupt the power grid, and c) economic resources among others. Due to the resource limitation, we define $C_{\mathcal{A}}$ as the maximum number of attacks that can be carried out at a time, each of which corresponds to a specific attack. For example, the attacker may launch a DoS attack over poorly protected wireless channels to block the communication between power grid sensors and remote estimators [18]–[21]. Alternatively, the attacker may plan to launch a physical attack on a high voltage (HV) transformer. Thus, the action space $\mathcal{A}$ of the attacker contains all possible methods of allocating $C_{\mathcal{A}}$ attacks over the $N_{\mathcal{G}}$ elements of the graph.

In order to maximize the system resiliency against such attacks, the defender needs to allocate its limited defense resources over the $N_{\mathcal{G}}$ elements of the graph to reinforce operational elements or to repair disabled elements. Similarly, the defense resources can include a) personnel such as users, administrators, and support personnel, b) technological resources such as advanced tools or softwares to reinforce operational elements or to repair broken elements of the power grid, and c) economic resources such as investments in new infrastructures or nodes, among others. Let $C_{\mathcal{D}}$ be the maximum number of defense mechanisms that can be implemented at a time, each of which is dependent on the type of the attack and the element disrupted. For instance, for false data injection attacks on the automatic generation control system, the defense mechanism can be to implement saturation filters. Alternatively, the defender needs to build some barriers and fortification for preventing physical attacks on critical infrastructures. Briefly, the defender should make sensible decisions about how to allocate finite resources over elements of the graph. Let $\mathcal{D}$ be the action space of the defender, then, it conditions all possible methods to distribute $C_{\mathcal{D}}$ defense mechanisms over the $N_{\mathcal{G}}$ elements of the graph.

In the literature, the physical impact of an attack on the system is measured by the cost of shed load following the failure of elements. In order to analyze the physical impacts of various attacks, attacks on the system can be classified into two categories: *isolated* and *coordinated*. The former can only destroy one element of the graph at a time, while the latter can target two or more elements. A coordinated attack that can collapse a combination of elements will naturally have a more

detrimental impact, as opposed to a single, isolated attack. For the system impacted by coordinated attacks, load shedding must be performed in order to regain stability.

In order to exactly quantify the cost of shed load, we need to solve an optimal load shedding problem which determines where and how many load must be shed under successful attacks. For the system composed of $N_{\mathcal{V}}$ buses, we assume that N_g are generation buses and N_l are load buses. As is common in [22]–[24], this optimal load shedding problem can be formulated as a constrained optimization problem, under the physical constraints of stable power flows. In practice, the load buses of the power grid may be of different importance. For example, load buses serving critical hospitals normally have higher importance than load buses stemming from electric vehicle charging stations. Thus, the weight vector $\boldsymbol{w}_l = [w_{l1}, w_{l2}, ..., w_{lN_l}]^T$ is introduced to represent the relative importance of different kinds of load buses. The optimal load shedding problem is therefore formulated to minimize the total cost of shed load at N_l load buses:

$$\begin{aligned} \min_{z,\theta} \quad & L = \sum_{j=1}^{N_l} w_{lj} u_{lj} z_{lj}, \\ \text{s.t.} \quad & \boldsymbol{\Gamma}^T \boldsymbol{B} \sin(\boldsymbol{\Gamma\theta}) - (\boldsymbol{p} + \boldsymbol{z}) = 0, \\ & \boldsymbol{p}_{\text{gmin}} \le \boldsymbol{p}_g + \boldsymbol{z}_g \le \boldsymbol{p}_{\text{gmax}}, \\ & \boldsymbol{z}_{\text{gmin}} \le \boldsymbol{z}_g \le \boldsymbol{z}_{\text{gmax}}, \\ & \boldsymbol{p}_l \le \boldsymbol{p}_l + \boldsymbol{z}_l \le \boldsymbol{p}_{\text{lmax}}, \\ & \boldsymbol{\theta}_{\min} \le \boldsymbol{\Gamma\theta} \le \boldsymbol{\theta}_{\max}, \\ & \boldsymbol{c}_{\min} \le \boldsymbol{B} \sin(\boldsymbol{\Gamma\theta}) \le \boldsymbol{c}_{\max}, \end{aligned} \tag{1}$$

where $\boldsymbol{u}_l = [u_{l1}, ..., u_{lN_l}]^T$ (\$/kW) is the load cost vector whose element represents the independent load values or costs of the corresponding bus. $\boldsymbol{p} = [\boldsymbol{p}_g; \boldsymbol{p}_l]$ represents the power distribution over $N_{\mathcal{V}}$ buses, in which $\boldsymbol{p}_g = [p_{g1}, ..., p_{gN_g}]^T \ge \boldsymbol{0}$ refers to power generation over N_g generation buses while $\boldsymbol{p}_l = [p_{l1}, ..., p_{lN_l}]^T \le \boldsymbol{0}$ refers to load distribution over N_l load buses. $\boldsymbol{z} = [\boldsymbol{z}_g; \boldsymbol{z}_l]$ represents the changes in power assignment for $N_{\mathcal{V}}$ buses due to element failures and corresponding load shedding, in which $\boldsymbol{z}_g = [z_{g1}, ..., z_{gN_g}]^T$ refers to re-dispatched power at given generation buses while $\boldsymbol{z}_l = [z_{l1}, ..., z_{lN_l}]^T$ refers to load to be shed at given load buses. $\boldsymbol{p}_{\text{gmin}} = [p_{\text{gmin}1}, ..., p_{\text{gmin}N_g}]^T \ge \boldsymbol{0}$ and $\boldsymbol{p}_{\text{gmax}} = [p_{\text{gmax}1}, ..., p_{\text{gmax}N_g}]^T \ge \boldsymbol{0}$ represent, respectively, the minimum and maximum outputs at given generation buses. $\boldsymbol{z}_{\text{gmin}} = [z_{\text{gmin}1}, ..., z_{\text{gmin}N_g}]^T \le \boldsymbol{0}$ represents the maximum power can be reduced at given generation buses for a time step, and $\boldsymbol{z}_{\text{gmax}} = [z_{\text{gmax}1}, ..., z_{\text{gmax}N_g}]^T \ge \boldsymbol{0}$ represents the maximum power can be increased at given generation buses for a time step. $\boldsymbol{p}_{\text{lmax}} = [p_{\text{lmax}1}, ..., p_{\text{lmax}N_l}]^T \le \boldsymbol{0}$ refers to important load that cannot be shed at given load buses. $\boldsymbol{\theta} = [\theta_1, ..., \theta_{N_{\mathcal{V}}}]^T$ is the phase angle at each bus. $\boldsymbol{\theta}_{\min} = [\theta_{\min 1}, ..., \theta_{\min N_{\mathcal{V}}}]^T$ and $\boldsymbol{\theta}_{\max} = [\theta_{\max 1}, ..., \theta_{\max N_{\mathcal{V}}}]^T$ represent the minimized and maximized limitations of the phase angle at each bus, respectively. $\boldsymbol{\Gamma} \in R^{N_{\mathcal{V}} \times N_{\mathcal{E}}}$ is the incidence matrix for the topology of the power grid, and $\boldsymbol{B} \in R^{N_{\mathcal{E}} \times N_{\mathcal{E}}}$ is the diagonal matrix whose diagonal entries correspond to line admittances. $\boldsymbol{c}_{\min} = [c_{\min 1}, ..., c_{N_{\min \mathcal{E}}}]^T$ and $\boldsymbol{c}_{\max} = [c_{\max 1}, ..., c_{\max N_{\mathcal{E}}}]^T$ independently represent the minimized and maximized power limitations of each branch.

The physical interpretations of the optimization problem constraints are explained as follows. The first constraint corresponds to the physical power flow equation that must be satisfied during the load shedding. The second constraint relates to the minimum and maximum output limitations for generation buses. The third constraint represents the ramping capacity for generation buses. In the fourth constraint, $\boldsymbol{p}_l + \boldsymbol{z}_l \le \boldsymbol{p}_{\text{lmax}}$ guarantees that some important load can not be shed at any time for load buses, and $\boldsymbol{p}_l \le \boldsymbol{p}_l + \boldsymbol{z}_l$ guarantees that the load at given load buses can only be shed, not added. The fifth indicates that, in order to keep the power system in a stable state, each phase angle of $N_{\mathcal{V}}$ buses should be in such an interval. Similarly, the capacity limits of transmission lines are given in the sixth constraint.

The first and sixth constraints including trigonometric sine functions make the load shedding problem to be a non-convex optimization problem. However, as discussed in [22], the phase angle differences $\boldsymbol{\Gamma\theta} \approx \boldsymbol{0}$ under normal operations of the power grid, the limits $\boldsymbol{\theta}_{\min} = -\boldsymbol{\pi/2}$ and $\boldsymbol{\theta}_{\max} = \boldsymbol{\pi/2}$ are required to guarantee steady state stability. Thus, the first constraint can be linearized as $\boldsymbol{\Gamma}^T \boldsymbol{B\Gamma\theta} - (\boldsymbol{p} + \boldsymbol{z}) = 0$, and the sixth constraint becomes $\boldsymbol{c}_{\min} \le \boldsymbol{B\Gamma\theta} \le \boldsymbol{c}_{\max}$ according to $\sin(\boldsymbol{\Gamma\theta}) \approx \boldsymbol{\Gamma\theta}$ for $\boldsymbol{\Gamma\theta} \approx \boldsymbol{0}$. The non-convex optimization problem (1) is therefore linearized as a linear programming problem that can be solved by efficient techniques including simplex methods [25] and interior-point methods [26] for optimal solutions. A coordinated attack could lead to a combination of element failures, then, the incidence matrix $\boldsymbol{\Gamma}$ representing the topology of the power grid should be updated according to the targets of the attack. Taking the updated incidence matrix $\boldsymbol{\Gamma}$ as an input parameter, the cost of shed load under the attack, denoted by L, will then be equal to $\sum_{j=1}^{N_l} w_{lj} u_{lj} z_{lj}$ derived by (1). Using this optimal load shedding technology, a game-theoretic approach for analyzing the interactions between the attacker and the defender is proposed next.

III. Game Model for Attack-Defense Interactions

We now mathematically analyze and identify the interactions between the attacker and the defender using the advanced tools of stochastic, noncooperative game theory [27]–[31]. In particular, we formulate a two-player stochastic game in normal form, $\Xi = \langle \mathcal{S}, \mathcal{A}, \mathcal{D}, R^{\mathcal{A}}, R^{\mathcal{D}} \rangle$, in which the players are the attacker and the defender. This game is played over a finite state space and each player has a finite number of actions to choose from. The main components of the game include:

- $\mathcal{S} := \{s_1, s_2, ..., s_{N_{\mathcal{S}}}\}$ which represents the power grid's state space;
- $\mathcal{A} := \{a_1, a_2, ..., a_{N_{\mathcal{A}}}\}$ which represents the attacker's action space;
- $\mathcal{D} := \{d_1, d_2, ..., d_{N_{\mathcal{D}}}\}$ which represents the defender's action space;
- $R^{\mathcal{A}}(s) := [R^{\mathcal{A}}(a, d, s)]_{N_{\mathcal{A}} \times N_{\mathcal{D}}}$ $(\mathcal{S} \times \mathcal{A} \times \mathcal{D} \to \mathcal{R})$ which represents the attacker's expected reward function corresponding to attack action $a \in \mathcal{A}$ against defense action $d \in \mathcal{D}$ in state $s \in \mathcal{S}$;
- $R^{\mathcal{D}}(s) := [R^{\mathcal{D}}(a, d, s)]_{N_{\mathcal{A}} \times N_{\mathcal{D}}}$ as the defender's expected reward function corresponding to defense action $d \in \mathcal{D}$ against attack action $a \in \mathcal{A}$ in state $s \in \mathcal{S}$.

In this game, each state $s \in \mathcal{S}$ of the power grid is associated with the amount of load, P_{shed} that is shed in this state. For

instance, if the total number of load in the grid is P_{all}, two states can be defined as follows:

- state $s_1 : P_{\text{shed}} = 0$;
- state $s_2 : 0 < P_{\text{shed}} \leq P_{\text{all}}$.

Let $\boldsymbol{m}^{\mathcal{S}}(t) := [\Pr\{s(t) = s_1\}, ..., \Pr\{s(t) = s_{N_{\mathcal{S}}}\}]^T$ be the probability distribution over the state space $\mathcal{S}$ at time t, then the state probability distribution at time $t+1$ will be $\boldsymbol{m}^{\mathcal{S}}[t+1] = \boldsymbol{T}(a,d) \times \boldsymbol{m}^{\mathcal{S}}[t]$, where $\boldsymbol{T}(a,d) := [T_{s_i,s_j}(a,d)]_{N_{\mathcal{S}} \times N_{\mathcal{S}}}$ is the state transition matrix parameterized by attack action $a \in \mathcal{A}$ and defend action $d \in \mathcal{D}$. The transition matrix entry $T_{s_i,s_j}(a,d)$ represents the probability of state s_i moving to state s_j under attack action a and defense action d. For the pair of players' actions (a,d), let $p_i^{\text{fail}}(a,d)$ be the probability that the normal element i of the power grid fails in one time step, and $p_i^{\text{rec}}(a,d)$ be the probability of the failed element j of the power grid recovering in one time step. The entry $T_{s_i,s_j}(a,d)$ can be derived based on two corresponding probabilities $p_i^{\text{fail}}(a,d)$ and $p_i^{\text{rec}}(a,d)$. Assuming that element i is disrupted by $c_i^{\mathcal{A}}$ attacks and protected of $c_i^{\mathcal{D}}$ defense mechanisms, the failure probability $p_i^{\text{fail}}(a,d)$ not only depends on the types of $c_i^{\mathcal{A}}$ attacks, but also on the types of $c_i^{\mathcal{D}}$ defense mechanisms. Accordingly, the recovery probability $p_j^{\text{rec}}(a,d)$ will be also determined by both the types of $c_j^{\mathcal{A}}$ attacks and $c_j^{\mathcal{D}}$ defense mechanisms implemented on element j. Therefore, a pair of players' actions (a,d) can cause the topology and structure of the power grid to be changed, and the grid transitions from the current state to another. If we take the number of power grid's elements $N_{\mathcal{G}} = n$, the probability of the power grid transitions from state s_1 to state s_2 with the transition probability: $T_{s_1,s_2}(a,d) = p_1^{\text{fail}}(a,d) + p_2^{\text{fail}}(a,d) + \cdots + p_n^{\text{fail}}(a,d)$.

A. Attacker and Defender's Strategies

For the power grid composed of $N_{\mathcal{G}}$ elements to be protected, the actions of both the attacker and the defender are constrained by a finite amount of resources. Thus, the attacker and the defender can only implement $C_{\mathcal{A}}$ attacks and $C_{\mathcal{D}}$ defense mechanisms, respectively, at a given time. We define each attack action vector $\boldsymbol{a}_i \in \mathcal{A}$, $i = 1, ..., N_{\mathcal{A}}$, as a method of allocating an attacker's finite resources over $N_{\mathcal{G}}$ elements:

$$\boldsymbol{a}_i = [c_{i,1}^{\mathcal{A}}, c_{i,2}^{\mathcal{A}}, ..., c_{i,N_{\mathcal{G}}}^{\mathcal{A}}]^T, \tag{2}$$

$$\sum_{j=1}^{N_{\mathcal{G}}} c_{i,j}^{\mathcal{A}} = C_{\mathcal{A}}, \tag{3}$$

where $0 \leq c_{i,j}^{\mathcal{A}} \leq C_{\mathcal{A}}$, $j = 1, 2..., N_{\mathcal{G}}$, represents the number of attacks related to action $\boldsymbol{a}_i$ that target element j of the grid. Similarly, each defense action vector $\boldsymbol{d}_i \in \mathcal{D}$, $i = 1, ..., N_{\mathcal{D}}$, conditions one method to distribute its limited defense resources over $N_{\mathcal{G}}$ elements:

$$\boldsymbol{d}_i = [c_{i,1}^{\mathcal{D}}, c_{i,2}^{\mathcal{D}}, ..., c_{i,N_{\mathcal{G}}}^{\mathcal{D}}]^T, \tag{4}$$

$$\sum_{j=1}^{N_{\mathcal{G}}} c_{i,j}^{\mathcal{D}} = C_{\mathcal{D}}, \tag{5}$$

where $0 \leq c_{i,j}^{\mathcal{D}} \leq C_{\mathcal{D}}$, $j = 1, 2..., N_{\mathcal{G}}$, denotes the number of defense mechanisms in action $\boldsymbol{d}_i$ that the defender plans to commit to element j of the grid.

For this game, we consider mixed strategies [28]–[31], in which the attacker and the defender's strategies are defined as probability distributions over their action spaces $\mathcal{A}$ and $\mathcal{D}$. Thus, the attacker's mixed strategy for a given state s will be:

$$\boldsymbol{\pi}_{\mathcal{A}}(s) := [\Pr\{a(s) = a_1\}, ..., \Pr\{a(s) = a_{N_{\mathcal{A}}}\}]^T, \tag{6}$$

$$\sum_{i=1}^{N_{\mathcal{A}}} \Pr\{a(s) = a_i\} = 1, \tag{7}$$

where $\Pr\{a(s) = a_i\}$ represents the probability of choosing attack action a_i in state $s \in \mathcal{S}$, and $\boldsymbol{\pi}_{\mathcal{A}}(s)$ yields the probability distribution over the attacker's action space $\mathcal{A}$. Correspondingly, the defender's mixed strategy in state s is given by:

$$\boldsymbol{\pi}_{\mathcal{D}}(s) := [\Pr\{d(s) = d_1\}, ..., \Pr\{d(s) = d_{N_{\mathcal{D}}}\}]^T, \tag{8}$$

$$\sum_{i=1}^{N_{\mathcal{D}}} \Pr\{d(s) = d_i\} = 1, \tag{9}$$

where $\Pr\{d(s) = d_i\}$ indicates the likelihood of selecting defense action d_i in state $s \in \mathcal{S}$, and $\boldsymbol{\pi}_{\mathcal{D}}(s)$ derives the probability distribution over the defender's action space $\mathcal{D}$.

B. Nash Equilibrium Strategies

In this game, for a given state $s \in \mathcal{S}$, a pair of players' actions (a,d) will lead to an immediate expected reward for both players. For the attacker, the expected reward, denoted by $R^{\mathcal{A}}(a,d,s)$, is measured by the expected cost of shed load due to the element failures of the power grid. $R^{\mathcal{A}}(a,d,s)$ will then be equal to $\sum_{s' \in \mathcal{S}}[T_{s,s'}(a,d) \times \sum_{j=1}^{N_l} w_{lj} u_{lj} z_{lj}]$, where the cost of shed load $\sum_{j=1}^{N_l} w_{lj} u_{lj} z_{lj}$ is derived by (1). The attacker intends to maximize its expected reward, while the defender aims to minimize it. Thus, the defender's expected reward is just the negative of the attacker's expected reward, denoted by $R^{\mathcal{D}}(a,d,s) = -R^{\mathcal{A}}(a,d,s)$. The proposed stochastic game Ξ is therefore a *zero-sum stochastic game*.

Moreover, this game can be considered as a collection of static games at each time step; the attacker and the defender repeatedly play games from this collection, and the particular game played at any given time depends probabilistically on the previous game played and on the actions taken by all players in that game. For example, given state $s \in S$, the attacker and the defender independently choose actions $a \in \mathcal{A}$ and $d \in \mathcal{D}$, and receive immediate expected rewards $R^{\mathcal{A}}(a,d,s)$ and $R^{\mathcal{D}}(a,d,s)$. The state then transits to the next state s' based on the fixed transition probability $T_{s,s'}(a,d)$. New expected rewards $R^{\mathcal{A}}(a,d,s')$ and $R^{\mathcal{D}}(a,d,s')$ will be obtained in the new state. We have specified the immediate rewards of the attacker and the defender at each stage game, but not how these rewards are aggregated into an overall payoff. To solve this problem, the most commonly used aggregation method is the discounted-sum reward [32]. For an attack action a and a defense action d, the discounted-sum reward of the attacker is the discounted sum of expected rewards at each time step t, with a discount factor $\gamma \in [0,1)$:

$$Q := \sum_{t=0}^{\infty} \gamma^t R^{\mathcal{A}}(a,d,s(t)), \tag{10}$$

where γ^t represents the weight of the immediate reward at the time step t, given by $R^{\mathcal{A}}(a,d,s(t))$, which denotes the

relative importance of the immediate reward in the overall payoff. Small values of γ emphasize near-term gains while large values emphasize future rewards. Correspondingly, the defender's discounted-sum reward is the negative of the number.

In this game, the attacker aims to maximize the discounted sum of expected rewards Q, while facing the defender who intends to minimize it. In order to solve for the two players' optimal strategies of a stochastic game in normal form such as Ξ, one popular solution is that of a closed-loop *Nash equilibrium* [27], [28]. A Nash equilibrium is a state of the game such that no player can increase its reward by *unilaterally* deviating from this equilibrium state. Formally, the Nash equilibrium of the proposed stochastic game Ξ is defined as follows:

Definition 1: Consider the proposed stochastic game $\Xi = \langle \mathcal{S}, \mathcal{A}, \mathcal{D}, R^{\mathcal{A}}, R^{\mathcal{D}} \rangle$, where expected rewards $R^{\mathcal{A}}$ and $R^{\mathcal{D}}$ are derived by solving the optimal load shedding problem (1), a *Nash equilibrium* solution of the proposed game is a two-tuple of mixed strategies $\{\boldsymbol{\pi}^*_{\mathcal{A}}, \boldsymbol{\pi}^*_{\mathcal{D}}\}$, where $\boldsymbol{\pi}^*_{\mathcal{A}} = \{\boldsymbol{\pi}^*_{\mathcal{A}}(s) | s \in \mathcal{S}\}$ and $\boldsymbol{\pi}^*_{\mathcal{D}} = \{\boldsymbol{\pi}^*_{\mathcal{D}}(s) | s \in \mathcal{S}\}$, such that, for all attack mixed strategies $\boldsymbol{\pi}_{\mathcal{A}}(s)$ and defense mixed strategies $\boldsymbol{\pi}_{\mathcal{D}}(s)$, $s \in \mathcal{S}$, it satisfies the following set of inequalities in state $s_i \in \mathcal{S}$, $i = 1, ..., N_{\mathcal{S}}$:

$$\begin{aligned} & Q(\boldsymbol{\pi}_{\mathcal{A}}(s_1), ..., \boldsymbol{\pi}^*_{\mathcal{A}}(s_i), ..., \boldsymbol{\pi}_{\mathcal{A}}(s_{N_{\mathcal{S}}}), \\ & \quad \boldsymbol{\pi}_{\mathcal{D}}(s_1), ..., \boldsymbol{\pi}_{\mathcal{D}}(s_i), ..., \boldsymbol{\pi}_{\mathcal{D}}(s_{N_{\mathcal{S}}})) \\ \geq\ & Q(\boldsymbol{\pi}_{\mathcal{A}}(s_1), ..., \boldsymbol{\pi}^*_{\mathcal{A}}(s_i), ..., \boldsymbol{\pi}_{\mathcal{A}}(s_{N_{\mathcal{S}}}), \\ & \quad \boldsymbol{\pi}_{\mathcal{D}}(s_1), ..., \boldsymbol{\pi}^*_{\mathcal{D}}(s_i), ..., \boldsymbol{\pi}_{\mathcal{D}}(s_{N_{\mathcal{S}}})) \\ \geq\ & Q(\boldsymbol{\pi}_{\mathcal{A}}(s_1), ..., \boldsymbol{\pi}_{\mathcal{A}}(s_i), ..., \boldsymbol{\pi}_{\mathcal{A}}(s_{N_{\mathcal{S}}}), \\ & \quad \boldsymbol{\pi}_{\mathcal{D}}(s_1), ..., \boldsymbol{\pi}^*_{\mathcal{D}}(s_i), ..., \boldsymbol{\pi}_{\mathcal{D}}(s_{N_{\mathcal{S}}})). \end{aligned} \tag{11}$$

However, the existence of Nash equilibrium in stochastic games is not immediate: for static games, the existence of Nash equilibrium is guaranteed by Nash's theorem [27], but in case of stochastic games, the possible number of strategies is infinite [33]. The existence of Nash equilibrium is known only in very special cases of stochastic games. Fortunately, if we limit our study to stationary optimal strategies by solving two players' mixed strategies in each state instead of each time step, where the attacker's mixed strategy $\boldsymbol{\pi}_{\mathcal{A}}(s) = \boldsymbol{\pi}_{\mathcal{A}}(s(t))$, $\forall t$ and the defender's mixed strategy $\boldsymbol{\pi}_{\mathcal{D}}(s) = \boldsymbol{\pi}_{\mathcal{D}}(s(t))$, $\forall t$ are optimal, it has been proved in [34] that there exists a unique Nash equilibrium in stationary strategies for two-player zero-sum discounted stochastic games. Therefore, there always exists a Nash equilibrium for guiding the attacker and the defender with the stationary strategy selection in the proposed stochastic game Ξ.

IV. Game Solution

We now solve the proposed stochastic game Ξ. Our objective is to characterize the attacker and the defender's Nash equilibrium strategies for each state $s \in \mathcal{S}$, where one player's Nash equilibrium strategy is the optimal strategy to maximize the minimum discounted sum of expected rewards under the opponent's optimal strategy. Due to the zero-sum nature of the game, it is sufficient to describe the solution of optimal strategies for only one player. Therefore, hereinafter, we focus on the defender's side of the game as the attacker's solution will be analogous.

Using the minimax Q-learning algorithm in [27], the defender's Nash equilibrium strategies can be derived recursively through the following dynamic programming approach. At time step t, the optimal discounted sum of expected rewards Q^* in state s under a pair of player actions (a, d) can be devised iteratively by following recursions:

$$Q_{t+1}(s, a, d) = R^{\mathcal{A}}(a, d, s) + \gamma \sum_{s' \in \mathcal{S}} T_{s,s'}(a, d) V(s'), \tag{12}$$

$$V(s') = \min_{\boldsymbol{\pi}_{\mathcal{D}}} \max_{a} \boldsymbol{\pi}^T_{\mathcal{D}}(s') Q_t(s', a, d), \tag{13}$$

for a given initial condition Q_0. The defender's mixed strategy $\boldsymbol{\pi}^*_{\mathcal{D}}(s)$, $\forall s \in \mathcal{S}$ calculated by (13) is the Nash equilibrium strategy. Unfortunately, every iteration of this algorithm just depends on rewards derived at the current time step while ignoring those derived before. Therefore, the computational complexity of such algorithm grows exponentially with the size of the power grid, making it impractical for grids of reasonable sizes. Inspired by the improved linear programming algorithm for MDPs in [31], we introduce a changeable learning rate $\alpha_t = 1/(t+1)^{\omega}$ for each time step t, for $\omega \in (1/2, 1]$, into the minimax Q-learning algorithm. Using the learning rate α_t, two new recursions are defined for computing the optimal discounted sum of expected rewards Q^* at time step t shown as follows:

$$\begin{aligned} Q_{t+1}(s, a, d) = (1 - \alpha_t) Q_t(s, a, d) + \alpha_t (R^{\mathcal{A}}(a, d, s) \\ + \gamma \sum_{s' \in \mathcal{S}} T_{s,s'}(a, d) \times V(s')), \end{aligned} \tag{14}$$

$$V(s') = \min_{\boldsymbol{\pi}_{\mathcal{D}}} \max_{a} \boldsymbol{\pi}^T_{\mathcal{D}}(s') Q_t(s', a, d), \tag{15}$$

for a given initial condition Q_0. (15) can be formulated as a linear constrained optimization problem:

$$\begin{aligned} \min_{\boldsymbol{\pi}_{\mathcal{D}}} \quad & V(s'), \\ \text{s.t.} \quad & \boldsymbol{\pi}^T_{\mathcal{D}}(s') Q_t(s', a, d) \leq V(s'), \forall a \in \mathcal{A}. \end{aligned} \tag{16}$$

The defender's mixed strategy $\boldsymbol{\pi}^*_{\mathcal{D}}(s)$, $\forall s \in \mathcal{S}$ calculated by (16) is the Nash equilibrium strategy. The fixed points of (14) and (15), V^* and Q^*, lead to the optimal minimax solution for the defender. Correspondingly, the attacker's Nash equilibrium strategy $\boldsymbol{\pi}^*_{\mathcal{A}}(s)$, $\forall s \in \mathcal{S}$ can be obtained by solving the dual of the linear constraint optimization (16):

$$\begin{aligned} \max_{\boldsymbol{\pi}_{\mathcal{A}}} \quad & V_{\text{dual}}(s'), \\ \text{s.t.} \quad & \boldsymbol{\pi}^T_{\mathcal{A}}(s') Q_t(s', a, d) \geq V_{\text{dual}}(s'), \forall d \in \mathcal{D}. \end{aligned} \tag{17}$$

For zero-sum stochastic games, the *strong max-min property* in [26] proves that strong duality applies and $V_{\text{dual}}(s')$ is equal to $V(s')$. Therefore, the tuple of Nash equilibrium strategies $(\boldsymbol{\pi}^*_{\mathcal{D}}(s), \boldsymbol{\pi}^*_{\mathcal{A}}(s))$, $\forall s \in \mathcal{S}$ obtained by (16) and (17) is the *Nash equilibrium* that we are looking for each state of the power grid. The procedure of computing the Nash equilibrium of the proposed stochastic game Ξ is detailed presented in Table I.

In the proposed algorithm, the whole state space $\mathcal{S}$ must be evaluated for deriving the Nash equilibrium of the proposed stochastic game Ξ. However, in practice, the players may focus on a small subset of $\mathcal{S}$, which corresponds to the current power grid conditions [35]. In such a case, we define $\mathcal{I}$ as the set of states that the two players are interested in. Then, to reduce

TABLE I
PROPOSED ALGORITHM

Phase 1 - Security Game Formulation:

a) Define the power grid state space $\mathcal{S}$ and the attacker's action space $\mathcal{A}$ based on potential threats on the grid.
b) Determine the defender's action space $\mathcal{D}$ corresponding to $\mathcal{A}$.
for $\forall s \in \mathcal{S}$, $a \in \mathcal{A}$ and $d \in \mathcal{D}$ **do**
c) Derive state transition matrix $T(a,d)$, $a \in \mathcal{A}$ and $d \in \mathcal{D}$ using probabilities $p^{\text{fail}}(a,d)$ and $p^{\text{rec}}(a,d)$ proposed in Section III.
d) Derive optimal cost $L(s,a,d)$ of shed load according to (1).
e) Derive attacker and defender's expected reward $R^{\mathcal{A}}(s,a,d)$ and $R^{\mathcal{D}}(s,a,d)$ according to $\sum_{s' \in \mathcal{S}} T_{s,s'}(a,d) L(s,a,d)$.
end for

Phase 2 - Security Game Solution:

Next, solve the *Nash equilibrium* of the formulated security game $\Xi = \langle \mathcal{S}, \mathcal{A}, \mathcal{D}, R^{\mathcal{A}}, R^{\mathcal{D}} \rangle$.
a) Set initial $Q_0(s,a,d)$ and $V(s)$, $\forall s \in \mathcal{S}$, $a \in \mathcal{A}$ and $d \in \mathcal{D}$.
b) Define the learning rate $\alpha_t = 1/(t+1)^{\omega}$, for $\omega \in (1/2, 1]$.
repeat
for $\forall s \in \mathcal{S}$ **do**
for $\forall a \in \mathcal{A}$ and $\forall d \in \mathcal{D}$ **do**
c) Update $Q_{t+1}(s,a,d)$ according to (14).
end for
d) Update linear programming $V(s)$ according to (16).
e) Derive optimal defense strategy by solving $V(s)$.
f) Update dual problem $V_{\text{dual}}(s)$ according to (17).
g) Derive optimal attack strategy by solving $V_{\text{dual}}(s)$.
end for
until reward sequence $\{Q_t\}$ derived converges to the equilibrium Q^*.

the computational complexity, we can eliminate the states that are not likely to be reached by the players. To decide if a state $s' \notin \mathcal{I}$ needs to be evaluated, we define a closeness value, $\beta_{s'}$, which bounds the highest probability of transitioning to state s' from state $s \in \mathcal{I}$, where the transition probabilities are considered. We only evaluate and update the value of a state when its closeness value exceeds a threshold parameter, given as $\delta \in (0,1]$. Otherwise, the state is not considered. For all states $s \in \mathcal{I}$, we have $\beta_s = 1$, and $\beta_{s'} = 0$ for all the other states $s' \notin \mathcal{I}$ initially. Then, for computing the value $V(s)$ of a state s (with $\beta_s \geq \delta$), we assign $\beta_{s'}$ as $\max\{\beta_{s'}, \beta_s \times T_{s,s'}(a,d)\}$ for state s'. Intuitively, the step means that we always use the highest closeness value of a state $s' \notin \mathcal{I}$ (whether it connects to a state $s \in \mathcal{I}$ directly or through other uninteresting but important states), to decide whether to include the state in the computation. We only update $V(s)$ (and compute the corresponding optimal player strategies) for all s where $\beta_s \geq \delta$. Here, δ controls the tradeoff between the efficiency and accuracy of solutions.

Next, we would like to prove the convergence of the proposed algorithm, where the reward sequence $\{Q_t\}_{t \to \infty}$ derived by the algorithm converges to the optimal point Q^*, given by $R^{\mathcal{A}}(s,a,d) + \gamma \sum_{s' \in \mathcal{S}} T_{s,s'}(a,d) \times V(s', \boldsymbol{\pi}^*_{\mathcal{D}}(s'), \boldsymbol{\pi}^*_{\mathcal{A}}(s'))$. Let $\mathcal{Q}$ be the space of all reward functions, and define $P_t : \mathcal{Q} \to \mathcal{Q}$ as a mapping on the reward space $\mathcal{Q}$ into the reward space $\mathcal{Q}$, where

$$P_t Q = R^{\mathcal{A}}(s,a,d) + \gamma \sum_{s' \in \mathcal{S}} [T_{s,s'}(a,d) \times \min_{\boldsymbol{\pi}_{\mathcal{D}}} \max_{a} \boldsymbol{\pi}^T_{\mathcal{D}}(s') Q_t(s', a, d)]. \quad (18)$$

Then, the convergence of the proposed algorithm follows from the application of the following Lemma in [31], which establishes convergence given three conditions.

Lemma 1: Assume that the learning rate $\alpha_t \in [0,1]$ satisfies the following condition that, for all state $s \in \mathcal{S}$, $a \in \mathcal{A}$, $d \in \mathcal{D}$, and time t, such that

$$\sum_{t=0}^{\infty} \alpha_t(s,a,d) = \infty, \sum_{t=0}^{\infty} [\alpha_t(s,a,d)]^2 < \infty. \quad (19)$$

And the mapping $P_t : \mathcal{Q} \to \mathcal{Q}$ satisfies the following condition: there exists a number $\gamma \in (0,1)$ and a sequence $\{\lambda_t\} \geq 0$ converging to zero with probability 1 such that

$$\|P_t Q - P_t Q^*\| \leq \gamma \|Q - Q^*\|^2 + \lambda_t, \quad (20)$$

for all $Q \in \mathcal{Q}$ and $Q^* = E[P_t Q^*]$, then the iteration defined by

$$Q_{t+1} = (1 - \alpha_t) Q_t + \alpha_t [P_t Q_t], \quad (21)$$

converges to Q^* with probability 1.

In order for the condition in Lemma 1 that $\|P_t Q - P_t Q^*\| \leq \gamma \|Q - Q^*\|^2 + \lambda_t$ to hold in the proposed algorithm, we have to restrict the domain of the reward functions in $\mathcal{Q}$. Our restriction focuses on the stage games $(Q_t(s))$ encountered during learning with special types of Nash equilibrium points: *global optima*, and *saddles*. Therefore, a following assumption about these stage games is proposed for the convergence proof of the proposed algorithm.

Assumption 1: Every stage game $(Q_t(s))$, $\forall t$ and $\forall s \in \mathcal{S}$, has a global optimal point or a saddle point, and players' expected rewards in this equilibrium are used for updating the discounted sum of expected rewards Q.

Now, we can present the convergence of the algorithm.

Theorem 1: Under Lemma 1 and Assumption 1, the reward sequence $\{Q_t\}_{t \to \infty}$ derived by the proposed algorithm in Table I converges to the optimal point Q^*.

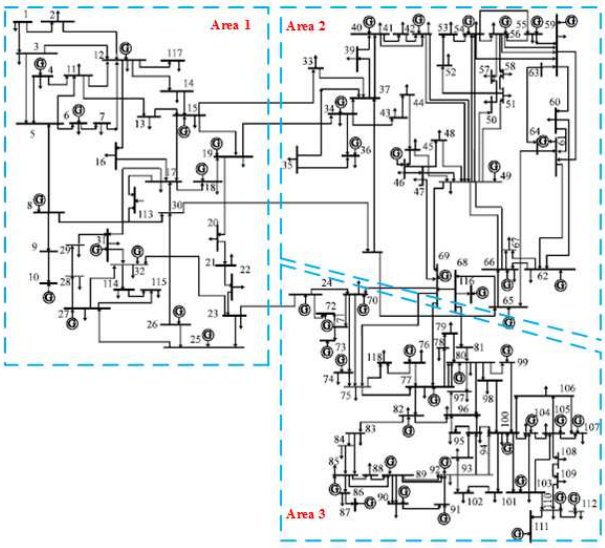
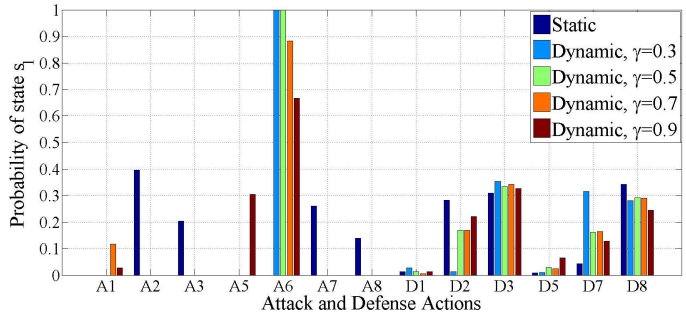


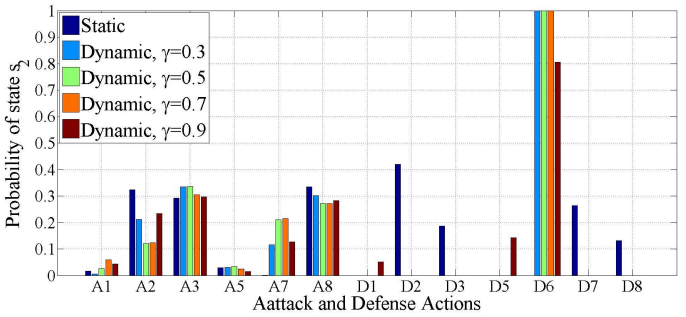
Fig. 1. The IEEE 118-bus system contains 19 generators, 177 transmission lines, 9 transformers, and 91 loads, in which all 118 buses are separated into three areas.

V. SIMULATION RESULTS AND ANALYSIS

We present simulation results on the WSCC 9-bus system [36] and the IEEE 118-bus system [37] represented by Fig. 1 to illustrate the solutions of the proposed stochastic game and evaluate the performance of the proposed algorithm. For simulating the game, we consider both physical attacks and denial-of-service (DoS) attacks targeted at disrupting the transmission lines of the system. The defender must implement

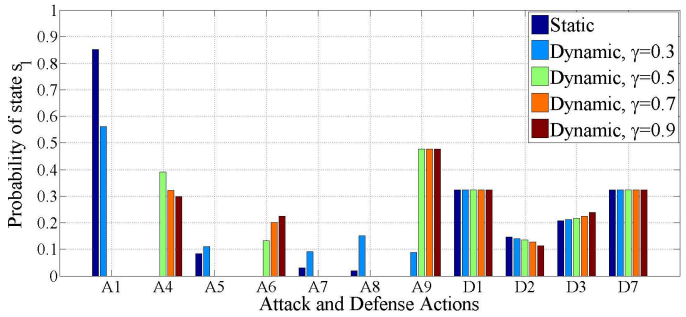


(a) state s_1

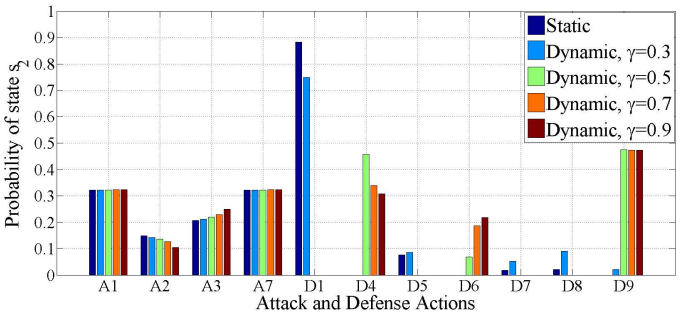


(b) state s_2

Fig. 2. The attacker and the defender's *Nash equilibrium* strategies in two states of the proposed game with various discounted factors γ for the WSCC 9-bus system. (a) state s_1. (b) state s_2.



(a) state s_1



(b) state s_2

Fig. 3. The attacker and the defender's *Nash equilibrium* strategies in two states of the proposed game with various discounted factors γ for the IEEE 118-bus system. (a) state s_1. (b) state s_2.

proper defense mechanisms such as building barriers and implementing filters, to reinforce normal transmission lines and repair broken lines. Due to finite attack resources, here, we assume that the attacker can only implement 6 attacks at a time, each of which corresponds to a specific attack for one transmission line. Similarly, we assume 5, 10, 20, and 40 defense mechanisms that can be carried out at a time for protecting transmission lines, respectively. Additionally, we intend to derive the *Nash equilibrium* of the proposed game Ξ in stationary strategies for each state of the power grid. Here, two power grid's states are considered including state s_1: $P_{\text{shed}} = 0$ that no load has to be shed, and state s_2: $P_{\text{shed}} > 0$ that load must be shed to balance the load with generation.

In the WSCC 9-bus system, 6 types of *isolated attacks* and 15 types of *coordinated attacks* are investigated. The amount of shed load following each of successful attacks is listed in Table II. Table II shows that coordinated attacks can lead to more load to be shed than isolated attacks, and attack 7-9, 13-15, 17 and 19-21 are ten coordinated attacks that can cause more physical damages on the system than other coordinated attacks. Thus, the attack action space $\mathcal{A}$ contains above ten coordinated attacks, where a_1 for Line 1 and 2, a_2 for Line 1 and 3, a_3 for Line 1 and 4, a_4 for Line 2 and 4, a_5 for Line 2 and 5, a_6 for Line 2 and 6, a_7 for Line 3 and 5, a_8 for Line 4 and 5, a_9 for Line 4 and 6, and a_{10} for Line 5 and 6. Similarly, the defense action space $\mathcal{D}$ includes ten corresponding defense actions.

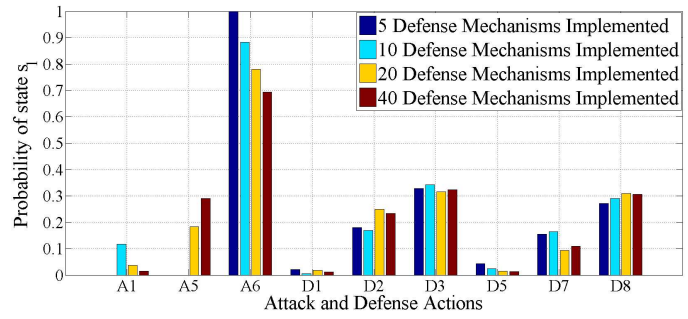
Fig. 2 presents the attacker and the defender's *Nash equilibrium* strategies in states s_1 and s_2 for the WSCC 9-bus system with various discount factors γ. In this figure, the y-axis shows the probability with which a certain attack or defense action will be chosen. Here, we assume 10 defense mechanisms that can be implemented, and define the fail probability as $p_i^{\text{fail}} = [c_i^A/(1+c_i^A)] \times [1/(1+c_i^D)]$ and the recovery probability

TABLE II
SHED LOAD DUE TO ATTACKS IN THE WSCC 9-BUS SYSTEM

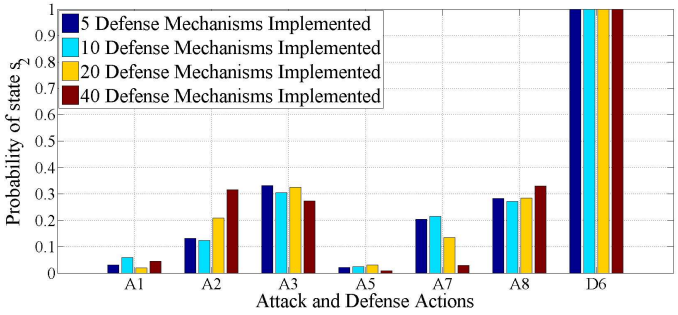
Attack No.	Attack Target	Shed Load (MW)	Attack No.	Attack Target	Shed Load (MW)
1	Line 1	0	12	Line 2 and 3	0
2	Line 2	0	13(a_4)	Line 2 and 4	352
3	Line 3	0	14(a_5)	Line 2 and 5	132
4	Line 4	0	15(a_6)	Line 2 and 6	27
5	Line 5	0	16	Line 3 and 4	0
6	Line 6	0	17(a_7)	Line 3 and 5	361
7(a_1)	Line 1 and 2	120	18	Line 3 and 6	0
8(a_2)	Line 1 and 3	376	19(a_8)	Line 4 and 5	220
9(a_3)	Line 1 and 4	232	20(a_9)	Line 4 and 6	328
10	Line 1 and 5	0	21(a_{10})	Line 5 and 6	135
11	Line 1 and 6	0			

as $p_i^{\text{rec}} = [c_i^D/(1+c_i^D)] \times [1/(1+c_i^A)]$, where $c_i^{a(d)}$ represents the number of attacks (defense mechanisms) related to action a (d) that targeted at the transmission line i of the system. In Fig. 2, we can see that the proposed game derives different Nash equilibrium strategies as we vary the discount factor γ from 0 (static game) to 0.9. For instance, Fig. 2(a) shows that in state s_1 of the static game, the attacker focuses on taking attack actions a_2, a_3, a_7 and a_8 that could lead to more load to be shed when the attack succeeds. In contrast, the attacker shifts its focus to attack action a_6 in the stochastic game, in which less impacts are caused. This observation can be explained according to the defender's strategy selection. In both the static game and the stochastic game, the defender focuses on taking defense actions d_2, d_3, d_7 and d_8. Although attack actions a_2, a_3, a_7 and a_8 could lead to higher physical damages on the system, they are difficult to be successful under a thorough protection of the defender. Thus, the attacker shifts its focus to the easier target.

For a bulk power grid, the failure of an element in the grid due to cyber-physical attacks may cascade to other parts of

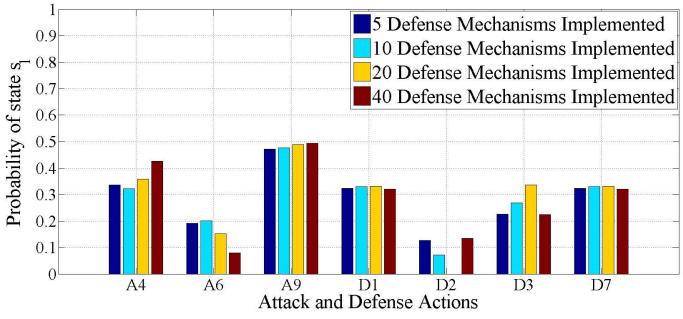


(a) state s_1

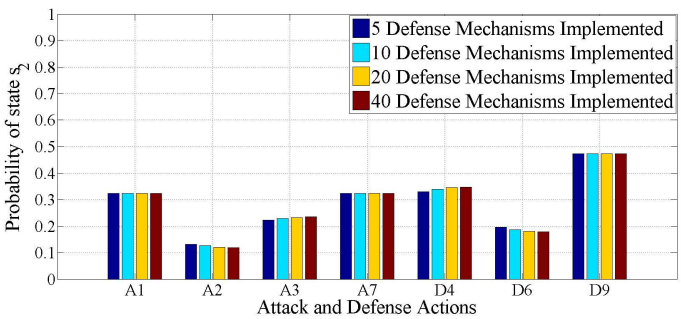


(b) state s_2

Fig. 4. The attacker and the defender's *Nash equilibrium* strategies in two states of the proposed stochastic game for the WSCC 9-bus system with various defense mechanisms can be implemented at a time. (a) state s_1. (b) state s_2.



(a) state s_1



(b) state s_2

Fig. 5. The attacker and the defender's *Nash equilibrium* strategies in two states of the proposed stochastic game for the IEEE 118-bus system with various defense mechanisms can be implemented at a time. (a) state s_1. (b) state s_2.

the grid and cause the failure of other interdependent elements. This process can be cascaded back and forth between multiple interdependent elements, resulting in a catastrophic failure. Controlled islanding is the last countermeasure for a bulk power grid when it suffers from severe cascading contingencies, in which the objective of controlled islanding is to maintain the stability of each island and to keep the total loss of loads of the whole system to a minimum. Based on controlled islanding strategies in [38], [39], the IEEE 118-bus system is separated into three areas for analysis, and the coherent stability is guaranteed within an area under *coordinated attacks* targeted at two elements. In the IEEE 118-bus system, 40 types of *coordinated attacks* are investigated in 4 scenarios, respectively. In scenario 1-3, ten attacks are explored for each of three areas of the system, where attack 1-10 for Area 1, attack 11-20 for Area 2, and attack 21-30 for Area 3. In scenario 4, attack 31-35 investigated are five attacks occurred between Area 1 and 2, while attack 36-40 are carried out between Area 2 and 3. The amount of shed load following each of successful attacks is listed in Table III. From Table III, we can find that attack 1, 11-15, 21, 24 and 27 are nine coordinated attacks with higher physical impacts on the system than others. Therefore, the attack action space $\mathcal{A}$ contains above nine coordinated attacks, where a_1 for Line 1 and 2, a_2 for Line 59 and 60, a_3 for Line 59 and 61, a_4 for Line 59 and 62, a_5 for Line 59 and 63, a_6 for Line 60 and 61, a_7 for Line 121 and 122, a_8 for Line 121 and 125, and a_9 for Line 122 and 125. Similarly, the defense action space $\mathcal{D}$ contains nine corresponding defense actions.

Given the same setting as the WSCC 9-bus system, Fig. 3 presents the attacker and the defender's *Nash equilibrium* strategies in states s_1 and s_2 for the IEEE 118-bus system with various discount factors γ. In Fig. 3, we can also find that the Nash equilibrium derived by the proposed stochastic game varies with different discount factors γ. However, Fig. 3(a) shows in state s_1, the defender's Nash equilibrium strategy varies less with different discount factors γ than the WSCC 9-bus system. Compared with the WSCC 9-bus system, the IEEE 118-bus system has a larger scale and the attack targets are less connected, the transmission lines targeted to be attacks are more independent between each other. Moreover, as state s_1 is the safe state without load to be shed, the defender's objective is just to reinforce the transmission lines against potential attacks. Thus, the defender's Nash equilibrium strategy varies little of different game models. Similarly, Fig. 3(b) shows the attacker's Nash equilibrium strategy varies little of different game models in the danger state s_2.

TABLE III
SHED LOAD DUE TO ATTACKS IN THE IEEE 118-BUS SYSTEM

Attack No.	Attack Target	Shed Load (MW)	Attack No.	Attack Target	Shed Load (MW)
1(a_1)	Line 1 and 2	196	21(a_7)	Line 121 and 122	196
2	Line 1 and 3	0	22	Line 121 and 123	0
3	Line 1 and 4	0	23	Line 121 and 124	0
4	Line 1 and 5	0	24(a_8)	Line 121 and 125	122
5	Line 2 and 3	0	25	Line 122 and 123	0
6	Line 2 and 4	0	26	Line 122 and 124	0
7	Line 2 and 5	0	27(a_9)	Line 122 and 125	35
8	Line 3 and 4	0	28	Line 123 and 124	0
9	Line 3 and 5	0	29	Line 123 and 125	0
10	Line 4 and 5	0	30	Line 124 and 125	0
11(a_2)	Line 59 and 60	194	31	Line 1 and 59	0
12(a_3)	Line 59 and 61	197	32	Line 1 and 60	0
13(a_4)	Line 59 and 62	34	33	Line 2 and 59	0
14(a_5)	Line 59 and 63	122	34	Line 2 and 60	0
15(a_6)	Line 60 and 61	36	35	Line 2 and 61	0
16	Line 60 and 62	0	36	Line 108 and 116	0
17	Line 60 and 63	0	37	Line 108 and 118	0
18	Line 61 and 62	0	38	Line 116 and 118	0
19	Line 61 and 63	0	39	Line 116 and 126	0
20	Line 62 and 63	0	40	Line 118 and 126	0

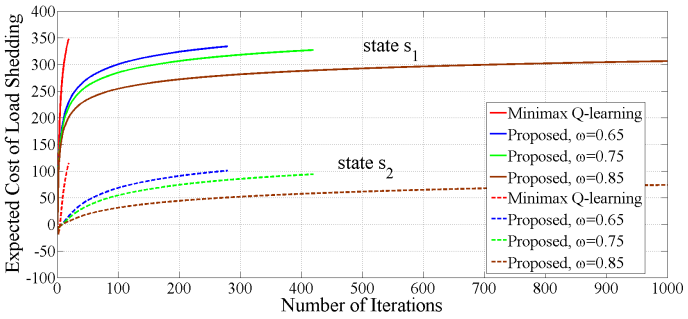


Fig. 6. The expected costs of shed load resulting from the stochastic game approach using various algorithms for two states of the WSCC 9-bus system.

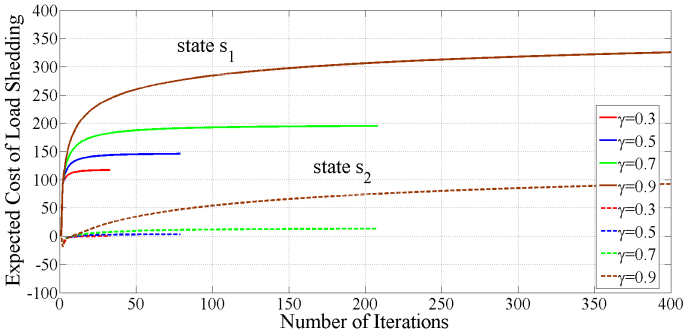


Fig. 7. The expected costs of shed load resulting from the stochastic game approach using the proposed algorithm for two states of the WSCC 9-bus system with various discount factors.

Fig. 4 presents the attacker and the defender's *Nash equilibrium* strategies in states s_1 and s_2 of the proposed stochastic game for the WSCC 9-bus system with various defense mechanisms that can be implemented at a time. Given the discount number to $\gamma = 0.7$, here, we assume 5, 10, 20, and 40 defense mechanisms that can be carried out at a time for protecting transmission lines, respectively. In Fig. 4, we can see that different defense resources owned could lead to two players' different Nash equilibrium strategies. For instance, Fig. 4(a) shows that when 5 defense mechanisms can be implemented, the attacker just focuses on attack action a_6 targeting the transmission lines 2 and 6. However, with 40 defense mechanisms implemented, the attacker shifts some part of its focus to attack action a_5, where it also adds the the transmission line 5. Similarly, Fig. 5 shows two players' Nash equilibrium strategies in two states of the proposed stochastic game for the IEEE 118-bus system with defense mechanisms can be implemented at a time.

Fig. 6 presents the expected cost of shed load resulting from the stochastic game approach for states s_1 and s_2 of the WSCC 9-bus system. Here, we choose 10 defense mechanisms that can be implemented, and we let $\gamma = 0.9$. As we define the same threshold $\Delta = e^{-3}$ for the termination of the minimax Q-learning algorithm and the proposed algorithm, in Fig. 6, we can see that the curve of the minimax Q-learning algorithm vanishes earlier than the curves of the proposed algorithm, which illustrates that the minimax Q-learning algorithm requires less iterations for converging to the Nash equilibrium compared with the proposed algorithm. However, compared with the minimax Q-learning algorithm, the proposed algorithm yields lower expected costs of shed load for the defender. In particular, the proposed algorithm yields an expected cost reduction ranging between 7.14% (for $\omega = 0.65$) and 14.29% (for $\omega = 0.85$) in state s_1 of the system relative to the minimax Q-learning algorithm. Fig. 6 also shows that, as ω increases, the proposed algorithm reaches a lower expected costs of shed load due to the decrease of the learning rate $\alpha_t = 1/(t+1)^{\omega}$.

Fig. 7 presents the expected cost of shed load resulting from the stochastic game approach using the proposed algorithm for states s_1 and s_2 of the WSCC 9-bus system with various discount factors. The performance of the proposed algorithm is compared for various discount factors γ. In this figure, as we define the same threshold $\Delta = e^{-3}$ for the termination of all proposed algorithms of various discount factors γ, we can find, as the discount factor increases, the curve of the proposed algorithm continues more iterations, which shows that more iterations are needed for the proposed algorithm to converge with the increase of the discount factor. However, the algorithm results in a higher expected costs of shed load with the increase of the discount factor. For instance, the proposed algorithm with the discount factor $\gamma = 0.9$ converges to an expected cost 166.67% higher than the cost of $\gamma = 0.3$ in state s_1 of the system.

VI. Conclusions

In this paper, we have presented a novel approach for analysis of the complex interactions between the attacker and the defender in the power grid, while considering the finite resources owned by the attacker and the defender that will have an important influence on their strategy selections. We have formulated a two-player zero-sum stochastic game between the attacker and the defender in which each player seeks to maximize its respective minimum rewards under the opponent's optimal strategy. In order to quantify their rewards, the optimal load shedding technology is introduced to determine the minimum cost of shed load. Using these quantified rewards as input, the attacker and the defender's *Nash equilibrium* strategies are derived by the proposed algorithm. The convergence of the algorithm and its properties are studied. The WSCC 9-bus system and IEEE 118-bus system are used as test models to illustrate solutions of the proposed game theoretic framework which can explore and evaluate the defense strategies for protecting the power grid against coordinated attacks. Simulation results have shown that the attacker and the defender should take different strategies corresponding to the resources owned.

Appendix

A. Convergence Proof of Theorem 1

Proof: Similar with Lemma 1, our proof establishes convergence given three conditions. Since the learning rate α_t defined by $1/(t+1)^{\omega}$ in the proposed algorithm satisfies $\sum_{t=0}^{\infty} \alpha_t(s,a,d) = \infty$ and $\sum_{t=0}^{\infty}[\alpha_t(s,a,d)]^2 < \infty$, two other conditions for convergence need to be proved. First, we prove that the proposed algorithm satisfies $Q^* = E[P_t Q^*]$.

Based on (12) and (13), we have

$$\begin{aligned} & Q^*(s,a,d) \\ =& \sum_{s' \in \mathcal{S}} T_{s,s'}(a,d) \cdot [R^{\mathcal{A}}(a,d,s) + \gamma \sum_{s' \in \mathcal{S}} (T_{s,s'}(a,d) \min_{\boldsymbol{\pi}^*_{\mathcal{D}}} \max_{a} \\ & \boldsymbol{\pi}^{*T}_{\mathcal{D}}(s') Q^*(s,a,d))] \\ =& E[P_t Q^*]. \end{aligned} \tag{22}$$

Next, if $\|P_tQ - P_tQ'\| \leq \gamma\|Q - Q'\|$, $\forall Q, Q' \in \mathcal{Q}$, is guaranteed, the convergence of the proposed algorithm is proved.

$$\begin{aligned} &\|P_tQ - P_tQ'\| \\ =&\max_{s\in\mathcal{S}} |\gamma \boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \gamma \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s')| \\ \leq& \gamma|\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s')|. \end{aligned} \tag{23}$$

We proceed to prove that

$$|\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s')| \leq \gamma\|Q - Q'\|. \tag{24}$$

Suppose that $\boldsymbol{\pi}_{\mathcal{D}}(s)$ and $\boldsymbol{\pi}'_{\mathcal{D}}(s)$ satisfies Assumption 1, which indicates that they are global optimal points or saddle points. If $\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s') \geq 0$, we have

$$\begin{aligned} &|\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s')| \\ =&\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s') \\ \leq& \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s') \\ \leq& \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s)\|Q(s) - Q(s')\| \\ =&\|Q(s) - Q(s')\|. \end{aligned} \tag{25}$$

Similarly, if $\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s') < 0$, we have

$$\begin{aligned} &|\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s')| \\ =&\boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s') - \boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) \\ \leq& \boldsymbol{\pi}_{\mathcal{D}}^T(s')Q(s') - \boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) \\ \leq& \boldsymbol{\pi}_{\mathcal{D}}^T(s)\|Q(s) - Q(s')\| \\ =&\|Q(s) - Q(s')\|. \end{aligned} \tag{26}$$

Thus, we can get, $\forall Q, Q' \in \mathcal{Q}$,

$$\begin{aligned} \|P_tQ - P_tQ'\| &\leq \gamma|\boldsymbol{\pi}_{\mathcal{D}}^T(s)Q(s) - \boldsymbol{\pi}_{\mathcal{D}}^{'T}(s')Q(s')| \\ &\leq \gamma\|Q - Q'\|. \end{aligned} \tag{27}$$

Therefore, the reward sequence $\{Q_t\}_{t\to\infty}$ derived by the proposed algorithm converges to the optimal point Q^*. ■

REFERENCES

[1] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 86–105, Sep. 2012.

[2] R. A. Serrano and E. Halper, "Sophisticated but low-tech power grid attack baffles authorities," *Los Angeles Times*, Feb. 2014.

[3] B. Ram and D. N. Vishwakarma, *Power system protection and switchgear*. New York: McGraw-Hill, 2007.

[4] S. Greengard, "The new face of war," *Commun. ACM*, vol. 53, no. 12, pp. 20–522, Dec. 2010.

[5] S.Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," in *Proc. 37th IEEE Conf. Ind. Electron. Soc.*, Melbourne, Australia, Nov. 2011.

[6] S. Backhaus, R. Bent, J. Bono, R. Lee, B. Tracey, D. Wolpert, D. Xie, and Y. Yildiz, "Cyber-physical security: A game theory model of humans interacting over control systems," *IEEE Trans. Smart Grid*, Dec. 2013.

[7] X. Cao, P. Cheng, J. Chen, and Y. Sun, "An online optimization approach for control and communication codesign in networked cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, Feb. 2013.

[8] X. Cao, P. Cheng, J. Chen, S. S. Ge, Y. Cheng, and Y. Sun, "Cognitive radio based state estimation in cyber-physical systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, Mar. 2014.

[9] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, Jun. 2011.

[10] S. Jha, O. Sheyner, and J. Wing, "Two formal analysis of attack graphs," in *Proc. IEEE Workshop Computer Security Foundations*, 2002.

[11] A. J. Holmgren, E. Jenelius, and J. Westin, "Evaluating strategies for defending electrical power networks against antagonistic attacks," *IEEE Trans. Power Syst.*, vol. 22, no. 1, Feb. 2007.

[12] P.-Y. Chen, S. M. Cheng, and K.-C. Chen, "Smart attacks in smart grid communication networks," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 24–29, 2012.

[13] Y. Li, L. Shi, P. Cheng, J. Chen, and D. E. Quevedo, "Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2831–2836, Oct. 2015.

[14] R. Deng, G. Xiao, and R. Lu, "Defending against false data injection attacks on power system state estimation," *IEEE Trans. Ind. Informat.*, 2015.

[15] L. Wei, A. H. Moghadasi, A. Sundararajan, and A. Sarwat, "Defending mechanisms for protecting power systems against intelligent attacks," in *Proc. IEEE 10th SoSE Conf.*, San Antonio, the United States, May 2015.

[16] C. Y. T. Ma, D. K. Y. Yau, X. Lou, and N. S. V. Rao, "Markov game analysis for attack-defense of power networks under possible misinformation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1676–1686, May 2013.

[17] NERC, "The highimpact, low-frequency event risk to the North American Bulk Power System," 2009.

[18] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal denial-of-service attack scheduling with energy constraint," *IEEE Trans. Autom. Control*, Mar. 2015.

[19] ——, "Optimal denial-of-service attack scheduling against linear quadratic gaussian control," in *Proc. 2014 American Control Conf.*, Portland, USA, Jun. 2014.

[20] ——, "Optimal dos attack scheduling in wireless networked control system," *IEEE Trans. Control Sys. Technol.*, Aug. 2015.

[21] ——, "Optimal DoS Attack Policy Against Remote State Estimationl," in *Proc. IEEE 52nd Annu. Conf. Decision Control*, Dec. 2013.

[22] A. Pinar, J. Meza, V. Donde, and B. Lessieutre, "Optimization strategies for the vulnerability analysis of the electric power grid," *SIAM J. Optimiz.*, vol. 20, no. 4, pp. 1786–1810, 2010.

[23] B. Otomega and T. V. Cutsem, "Undervoltage load shedding using distributed controllers," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1898–1907, 2007.

[24] V. C. Nikolaidis, C. D. Vournas, G. A. Fotopoulos, G. P. Christoforidis, E. Kalfaoglou, and A. Koronides, "Automatic load shedding schemes against voltage stability in the hellenic system," in *Proc. IEEE PES Gen. Meet.*, Tampa, FL, Jun. 2007.

[25] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer Press, 2006.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[27] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, Aug. 2012.

[28] A. P. Maitra and W. D. Sudderth, *Discrete Gambling and Stochastic Games*. New York: Sringer-Verlag, 1996.

[29] T. E. S. Raghaven, T. S. Ferguson, T. Parthasarathy, and O. Vrieze, *Stochastic Games and Related Topics: In Honor of Professor L. S. Shapley*. New York: Springer Netherlands, 1991.

[30] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM Series in Classics in Applied Mathematics, 1999.

[31] A. Neyman and S. Sorin, *Stochastic Games and Applications*. New York: Kluwer Academic, Jul. 1999.

[32] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[33] K. Chatterjee and A. Tarlecki, *Computer Science Logic*. Springer Berlin Heidelberg Press, 2004.

[34] L. Shapley, "Stochastic games," in *Proc. Nat. Acad. Sci. USA*, vol. 39, 1953, pp. 1095–1100.

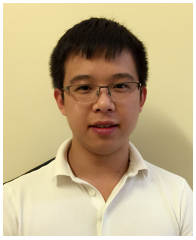
[35] C. Y. T. Ma, D. K. Y. Yau, and N. S. V. Rao, "Scalable solutions of markov games for smart-grid infrastructure protection," *IEEE Trans. Smart Grid*, vol. 4, no. 1, Mar. 2013.

[36] P. M. Anderson and A. A. Fouad, *Power system control and stability*. Delhi, India: Galgotia, 1981.

[37] C. Canizares and F. Alvarado, "Point of collapse and continuation methods for large AC/DC systems," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 1–8, 1993.

[38] A. Peiravi and R. Ildarabadi, "A fast algorithm for intentional islanding of power systems using the multilevel kernel k-means approach," *J Appl Sci*, vol. 9, no. 12, pp. 2247–2255, 2009.

[39] H. Song, J. Wu, and K. Wu, "A wide-area measurement systems-based adaptive strategy for controlled islanding in bulk power systems," *Energies*, vol. 7, no. 4, pp. 2631–2657, 2014.



Longfei Wei (S'15) received his B.S. and M.S. degrees in applied mathematics from Hebei University of Technology, Tianjin, China, in 2011 and 2014, respectively. He is currently pursing the Ph.D. degree with the Department of Electrical and Computer Engineering, Florida International University, under the supervision of Prof. A. Sarwat.

His current research focuses on numerical optimization, game theory, and cyber-physical security of smart grids.

Arif I. Sarwat (M'08) received his M.S. degree in electrical and computer engineering from University of Florida, Gainesville. In 2010 Dr. Sarwat received his Ph.D. degree in electrical engineering from the University of South Florida.

He worked in the industry (SIEMENS) for nine years executing many critical projects. Currently, he is an Assistant Professor at the Department of Electrical and Computer Engineering at Florida International University (FIU), where he leads the Energy, Power and Stainability (EPS) group. Before joining Florida International University, he was Assistant Professor of Electrical Engineering at the University at Buffalo, the State University of New York (SUNY). His significant work in energy storage, microgrid and DSM is demonstrated by Sustainable Electric Energy Delivery Systems in Florida. His research areas are smart grids, high penetration renewable systems, cyber-physical systems, power system reliability, large scale distributed generation integration, large scale data analysis, cyber security, and vehicular technology. Dr. Sarwat is the recipient of the NSF CAREER award in 2015.

Walid Saad (S'07, M'10, SM'15) received his Ph.D. degree from the University of Oslo in 2010. Currently, he is an Assistant Professor and the Steven O. Lane Junior Faculty Fellow at the Department of Electrical and Computer Engineering at Virginia Tech, where he leads the Network Science, Wireless, and Security (NetSciWiS) laboratory, within the Wireless@VT research group. His research interests include wireless networks, game theory, cyber security, and cyber-physical systems. Dr. Saad is the recipient of the NSF CAREER award in 2013, the AFOSR summer faculty fellowship in 2014, and the Young Investigator Award from the Office of Naval Research (ONR) in 2015. He was the author/co-author of five conference best paper awards at WiOpt in 2009, ICIMP in 2010, IEEE WCNC in 2012, IEEE PIMRC in 2015, and IEEE SmartGridComm in 2015. He is the recipient of the 2015 Fred W. Ellersick Prize from the IEEE Communications Society. Dr. Saad serves as an editor for the IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, and IEEE Transactions on Information Forensics and Security.

Saroj Biswas (M'86) received the Ph.D. degree in Electrical Engineering from University of Ottawa in 1986, and BSEE and MSEE degrees from Bangladesh University of Engineering and Technology, in 1975 and 1977, respectively. He is a Professor of Electrical and Computer Engineering at Temple University, Philadelphia, specializing in control and optimization of dynamic systems, power systems, and distributed parameter systems. His current research focuses on security of cyber-physical systems with applications to power grid based on multiagent framework, and the development of an intelligent virtual laboratory for electrical machines. He has also developed a control theoretic framework for modeling and control of magnetic signatures. Dr. Biswas is a member of IEEE, ASEE, and Sigma Xi.