2018 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS): BigSecurity 18: The Fourth International Workshop on Security and Privacy in Big Data

A Stochastic Game Approach to Cyber-Physical Security with Applications to Smart Grid

Yuanxiong Guo*, Yanmin Gong*, Laurent L. Njilla†, and Charles A. Kamhoua‡
*School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078
†Air Force Research Laboratory, Rome, NY 13441
‡Army Research Laboratory, Adelphi, MD 20783
Email: *{richard.guo, yanmin.gong}@okstate.edu, †laurent.njilla@us.af.mil, ‡charles.a.kamhoua.civ@mail.mil,

***Abstract*—This paper proposes a game-theoretic approach to analyze the interactions between an attacker and a defender in a cyber-physical system (CPS) and develops effective defense strategies. In a CPS, the attacker launches cyber attacks on a number of nodes in the cyber layer, trying to maximize the potential damage to the underlying physical system while the system operator seeks to defend several nodes in the cyber layer to minimize the physical damage. Given that CPS attacking and defending is often a continual process, a zero-sum Markov game is proposed in this paper to model these interactions subject to underlying uncertainties of real-world events and actions. A novel model is also proposed in this paper to characterize the interdependence between the cyber layer and the physical layer in a CPS and quantify the impact of the cyber attack on the physical damage in the proposed game. To find the Nash equilibrium of the Markov game, we design an efficient algorithm based on value iteration. The proposed general approach is then applied to study the wide-area monitoring and protection issue in smart grid. Extensive simulations are conducted based on real-world data, and results show the effectiveness of the defending strategies derived from the proposed approach.**

I. Introduction

Cyber-physical systems (CPSs) are engineered systems that are built from, and depend upon, the seamless integration of computational algorithms and distributed physical components. They are expected to drive innovation and competition in sectors such as agriculture, energy, transportation, building design and automation, healthcare, and manufacturing. However, CPSs are subject to threats due to their increasing reliance on information and communication technologies. Recent real-world events [1], [2] have clearly demonstrated the vulnerability of a CPS to various malicious attacks and the destructive effects that such attacks can have in practice. Therefore, it is imperative for us to ensure the security of current and emerging CPSs.

Given its great importance, CPS security has attracted increasing attention recently. Some general frameworks to model and analyze CPS security have been proposed in [3], [4], [5], [6]. The CPS operational objectives under potential security threats are characterized in [3]. Pasqualetti et al. [4] propose a mathematical framework to detect and identify various attacks in a CPS. A hierarchical architecture for CPS security is developed in [5] to design a cross-layer security solution. In [6], resilient control schemes are developed to mitigate the impact of denial-of-service (DoS) attacks for a control system. All of these work are from a control perspective and consider a general CPS, which can be applied to several applications including smart grid. There also exist several studies that solely focus on smart grid security. In particular, false data injection attack in smart grid [7], [8], [9] has been studied extensively. Surveys about smart grid security can be found in [10], [11]. Although these work provide interesting insights to CPS and smart grid security, few of them consider the strategic interactions between attackers and defenders in CPSs. In a CPS, the attacker often tries to maximize the potential damage to the system while the defender seeks to minimize the damage. Therefore, the actions and objectives of attackers and defenders in a CPS are closely related, which makes the consideration of their strategic interactions necessary and motivates a game-theoretic approach.

Game theory has been applied to study CPS security (see [12] for a recent survey on related work). However, there are two aspects that are often ignored in existing literature. First, CPS attacking and protecting is a continual process, where attackers and defenders continue to interact to produce dynamic states that reflect their respective best actions. Second, there is an interdependence between the cyber layer and the physical layer in a CPS, which means that the cyber attackers could propagate to the physical layer and cause physical damage. However, none of the existing work [13], [14], [15], [16] on applying game theory to CPS or smart grid security fully address both of these aspects.

In this paper, we propose a game-theoretic approach to analyze the interactions between an attacker and a defender in a CPS and develop effective defense strategies. Given a general CPS model, the CPS security problem is formulated as a zero-sum Markov game between the attacker and defender that are continually interacting in the cyber layer of the CPS. In this game, the attacker launches cyber attackers against a set of nodes in the cyber layer of the CPS and aims to maximize the potential physical damage to the underlying physical system. On the other hand, the defender selects a set of nodes in the cyber layer to protect such that the physical damage to the underlying physical system is minimized. The

DISTRIBUTION A. Approved for public release: distribution unlimited. Case Number: 88ABW-2017-5830. Dated 17 Nov 2017.

978-1-5386-5979-3/18/$31.00 ©2018 IEEE

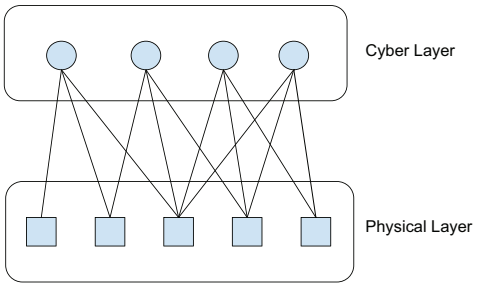


Fig. 1. A cyber-physical system model.

impact of cyber attacks on the underlying physical system of a CPS is characterized by a conditional probability that indicates the failure probability of a physical node when a cyber node has been attacked successfully. The system state will evolve based on the actions made by both the attacker and defender as well as the underlying uncertainties of real-world events. We then propose an algorithm to solve the proposed Markov game and, as a case study, apply it to the wide-area monitoring and protection problem in smart grid. In summary, the contributions of this paper are as follows:

- We develop a novel framework to analyze the attacker-defender interactions of a general CPS and propose a Markov game to model it considering the continual nature of the interactions.
- We design an algorithm based on value iteration to find the Nash equilibrium of the proposed Markov game and derive the optimal defending strategy under an intelligent attacker.
- We apply our general framework to a wide-area monitoring and protection scenario in smart grid. Experimental results verify the effectiveness of our proposed approach.

The rest of the paper is organized as follows. Section II describes our general CPS security model and the proposed Markov game. Section III presents the designed algorithm to compute the Nash equilibrium of the proposed game. Experimental results are discussed in Section IV based on a real-world power system, and the conclusion is made in Section V.

II. System Modeling and Problem Formulation

Consider a cyber-physical system consisting of n nodes in the cyber layer and m nodes in the physical layer as shown in Fig. 1. The cyber layer and the physical layer are interconnected, and security breaches can propagate from the cyber layer to the physical layer. Both the defender and the attacker operate on the cyber layer of the system and care about the effect on the physical layer. In the following, we first describe the game model and then illustrate how the interdependence between cyber layer and physical layer can be incorporated in the proposed game model.

A. Game Formulation

We define our Markov game as follows. There are two players in the game: a defender and an attacker. For the defender, its action $d := \{d_1, \ldots, d_n\}$ is the cyber nodes that it chooses to defend, where d_i is equal to 1 if node i is chosen or 0 otherwise. For the attacker, its action $a := \{a_1, \ldots, a_n\}$ is the cyber nodes that it chooses to attack, where a_i is equal to 1 if node i is chosen or 0 otherwise. Both players have limited budget such that in each time slot, the attacker can choose a limited number of cyber nodes n_a to attack, and the defender can choose a limited number of cyber nodes n_d to defend. Therefore, the action sets of the defender $\mathcal{D}$ and the attacker $\mathcal{A}$ are defined as

$$\mathcal{D} := \left\{ d \mid \sum_{i=1}^{n} d_i \leq n_d, d_i \in \{0,1\}, i = 1, \ldots, n. \right\}, \quad (1)$$

$$\mathcal{A} := \left\{ a \mid \sum_{i=1}^{n} a_i \leq n_a, a_i \in \{0,1\}, i = 1, \ldots, n. \right\}. \quad (2)$$

The state of the game denotes the set of normal cyber nodes that are currently working correctly in the cyber-physical systems. Let x_i be a binary variable indicating whether the cyber node i is working correctly (i.e., $x_i = 1$) or not (i.e., $x_i = 0$). Then each state s of the game can be represented as an n-dimensional vector, where each element represents whether the cyber node i is working or not and has a binary value, and the state space is represented as $\mathcal{S}$.

The game repeats in discrete time slots. Assume the failure or recover of each node in the cyber layer is independent. In general, the operating status of a cyber node at the current time slot depends on its previous status and the current decisions of both the defender and the attacker. In each time slot, the two players choose a pair of actions which may cause state transition in a Markov manner characterized as follows. Let μ be the successful attacking probability if the attacker attacks a cyber node, and λ be the successful defending probability if the defender defends a cyber node. For simplicity of presentation, we assume that these events are independent. Note that our game-theoretic framework can easily accommodate more-complicated scenarios as long as the state transition is in a Markov manner. An abnormal cyber node will restore to the normal state if it has been defended successfully. Similarly, a normal cyber node will move to the abnormal state if it has been attacked successfully. Then for each cyber node i at each time slot t, we have the following transition probabilities:

$$\Pr(x_i^t = 0 | x_i^{t-1} = 1, d_i^t, a_i^t) = \begin{cases} \mu(1-\lambda) & \text{if } d_i^t = 1, a_i^t = 1 \\ \mu & \text{if } d_i^t = 0, a_i^t = 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$\Pr(x_i^t = 1 | x_i^{t-1} = 0, d_i^t, a_i^t) = \begin{cases} \lambda(1-\mu) & \text{if } d_i^t = 1, a_i^t = 1 \\ \lambda & \text{if } d_i^t = 1, a_i^t = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, we have $\Pr(x_i^t = 1 | x_i^{t-1} = 1, d_i^t, a_i^t) = 1 - \Pr(x_i^t = 0 | x_i^{t-1} = 1, d_i^t, a_i^t), \forall a_i^t, d_i^t$ and $\Pr(x_i^t = 0 | x_i^{t-1} = 0, d_i^t, a_i^t) = 1 - \Pr(x_i^t = 1 | x_i^{t-1} = 0, d_i^t, a_i^t), \forall a_i^t, d_i^t$.

Each player in the game has a payoff function. In a Markov game, a player's stationary policy is a function that given a

state, returns a probability distribution (i.e., a mixed strategy) over the set of actions that the player may perform. Let $\pi_{\text{att}} := [\pi_a(s), \forall a, s]$ and $\pi_{\text{def}} := [\pi_d(s), \forall d, s]$ be policies of the attacker and the defender, respectively. Obviously, we have $\sum_{a\in\mathcal{A}} \pi_a(s) = 1, \forall s$ and $\sum_{d\in\mathcal{D}} \pi_d(s) = 1, \forall s$.

The defender aims to minimize the cost of the system by choosing the defending policy while the attacker aims to maximize it by choosing the attacking policy. Therefore we have a zero-sum game. Denote by $Q(s, a, d)$ the expected long-term system cost when the attacker takes action $a \in \mathcal{A}$ and the defender takes action $d \in \mathcal{D}$ in state $s \in \mathcal{S}$. In the Markov game, the value of the attacker at state $s \in \mathcal{S}$ is

$$V_{\text{att}}(s) = \max_{\pi_{\text{att}}} \min_{d} \sum_{a\in\mathcal{A}} Q(s, a, d)\pi_a(s). \tag{3}$$

Here $Q(s, a, d)$ can be represented as

$$Q(s, a, d) = c(s, a, d) + \gamma \sum_{s'\in\mathcal{S}} T(s, a, d, s')V_{\text{att}}(s'), \tag{4}$$

where $c(s, a, d)$ is the immediate system cost when the attacker takes action a and the defender takes action d in state s, $T(s, a, d, s')$ is the probability of state transition from s to s' under the actions a and d of the attacker and defender, respectively, and γ is a discount factor satisfying $0 \le \gamma < 1$. The state transition matrix $\mathbf{T} := [T(s, \cdot, \cdot, s'), \forall s, s']$ is calculated based on the probabilities λ and μ as defined before.

Similarly, the value of the defender at state $s \in \mathcal{S}$ is

$$V_{\text{def}}(s) = \min_{\pi_{\text{def}}} \max_{a} \sum_{d\in\mathcal{D}} Q(s, a, d)\pi_d(s). \tag{5}$$

As our game is zero-sum, we have $V_{\text{att}}(s) = V_{\text{def}}(s)$ and hence can use $V(s)$ to denote the value of state s without ambiguity. Therefore, the payoff functions of the attacker and the defender are $V(s)$ and $-V(s)$, respectively.

B. Impact of Cyber Attacks on Physical Domain

The immediate system cost $c(s, a, d)$ in (4) depends on the interdependence between the cyber layer and the physical layer in a cyber-physical system and varies for different applications. In this section, we adopt a model proposed in [16] to characterize such independence. In practice, the control law that governs the operations of the physical system relies on the remote data collected by cyber nodes. Such remote data is first sent via communication channels to the control center which then computes the optimal control commands and sends them back to the cyber nodes. After that, the cyber nodes initiate control actions over the physical nodes. Considering such a scenario, let $\mathbf{B} := [b_{ij}, \forall i, j]$ be the interdependence matrix of the cyber-physical system, where $b_{ij} \in [0, 1]$ is the weight that captures the effect of the data sent by cyber node i on the control action over physical node j and represents the conditional probability that physical node j temporarily fails given the corrupt data sent by cyber node i. It is obvious that $\sum_{i=1}^{n} b_{ij} = 1, \forall j = 1, \ldots, m$.

At each time slot t, let $p_i(t)$ be the probability that cyber node i fails at time t, and $q_j(t)$ be the probability that physical node j fails at time t. We have

$$q_j(t) = \sum_{i=1}^{n} b_{ij}p_i(t), \forall j, t. \tag{6}$$

Note that in our system $p_i(t)$ is uniquely determined by the system state s as the cyber nodes that fails in state s have $p_i = 1$, and can thus be rewritten as $p_i(s)$. Similarly, $q_j(t)$ can be rewritten as $q_j(s)$. Assume that each physical node j is associated with a cost of failure ρ_j to the system cost. Then the expected immediate system cost can be represented as

$$c(s, a, d) = \sum_{s'\in\mathcal{S}} T(s, a, d, s') \sum_{j=1}^{m} \rho_j q_j(s'). \tag{7}$$

III. Solution of the Game

In this section, we solve the Markov game defined before. The solution concept we use here is Nash equilibrium (NE) [17]. A NE solution gives the defender an idea of the attacker's policy and a plan for what to do in each state in the event of an attack. It has been proved in [18] that every Markov game has a stationary NE. Our goal is to find this stationary NE, which gives us the optimal policies for both players at equilibrium. Note that the optimal solutions computed individually by both players in (3) and (5) are best responses to each other. Since $V(s) = V_{\text{att}}(s) = V_{\text{def}}(s)$ for all states $s \in \mathcal{S}$, based on the definition of NE, these optimal solutions are also in NE. Therefore, we only need to solve (3) and (5) to obtain the NE in our game.

The optimal policy π^*_{att} of the attacker can be obtained by solving (3), which can be reformulated as the following optimization problem:

$$\max_{\pi_{\text{att}}(s)} \quad V(s) \tag{8a}$$

$$\text{s.t.} \quad V(s) \le \sum_{a\in\mathcal{A}} Q(s, a, d)\pi_a(s), \forall d \in \mathcal{D} \tag{8b}$$

$$\sum_{a\in\mathcal{A}} \pi_a(s) = 1, \tag{8c}$$

$$\pi_a(s) \ge 0, \forall a \in \mathcal{A}. \tag{8d}$$

Similarly, the optimal policy π^*_{def} of the defender can be obtained from (5) as follows:

$$\min_{\pi_{\text{def}}(s)} \quad V(s) \tag{9a}$$

$$\text{s.t.} \quad V(s) \ge \sum_{d\in\mathcal{D}} Q(s, a, d)\pi_d(s), \forall a \in \mathcal{A} \tag{9b}$$

$$\sum_{d\in\mathcal{D}} \pi_d(s) = 1, \tag{9c}$$

$$\pi_d(s) \ge 0, \forall d \in \mathcal{D}. \tag{9d}$$

The key challenge to solve (8) and (9) is the absence of the explicit forms for V and Q. To tackle this challenge, we propose to use the value iteration algorithm [19] to compute the optimal Q and V for given state s, attacker action a, and

defense action d. The value iteration algorithm is described in Algorithm 1. The proposed algorithm iteratively estimates the values of V and Q, and will converge to their correct values. Note a linear program is solved in each iteration (Line 7) to obtain a mixed strategy of the attacker, which will converge to the optimal one. Then we use the converged values of Q to obtain the optimal policy of the defender (Line 13).

Algorithm 1 Value iteration algorithm to calculate NE

1: Initialization: set $V^0(s) \leftarrow 0, \forall s \in \mathcal{S}$ and $k \leftarrow 0$;
2: **repeat**
3: **for all** $s \in \mathcal{S}$, $a \in \mathcal{A}$, $d \in \mathcal{D}$ **do**
4: update $Q^{k+1}(s, a, d)$ according to (4) with $V(s') = V^k(s')$;
5: **end for**
6: **for all** $s \in \mathcal{S}$ **do**
7: update $V^{k+1}(s)$ and $\pi_{\text{att}}^{k+1}(s)$ as the optimal objective function value and solution to (8) with $Q(s, a, d) = Q^{k+1}(s, a, d)$, respectively;
8: **end for**
9: $k \leftarrow k + 1$;
10: **until** convergence criteria is satisfied;
11: set $\pi_{\text{att}}^*(s) = \pi_{\text{att}}^k(s), \forall s \in \mathcal{S}$;
12: **for all** $s \in \mathcal{S}$ **do**
13: set $\pi_{\text{def}}^*(s)$ to be the optimal solution to (9) with $Q(s, a, d) = Q^k(s, a, d)$;
14: **end for**

IV. Evaluation Results

In this section, as a case study, we apply the proposed framework to the wide-area monitoring and protection in smart grid.

A. Experimental Setting

We consider the scenario where system-wide information sent from a collection of cyber nodes is used to generate protective actions affecting the connectivity of the system's physical components to prevent the propagation of large disturbances [20]. We use the same setting as [16]. The test system we consider here is a 5-bus system from PJM as shown in Fig. 2. The data related to this system is available in [21]. In the figure, there are 12 cyber nodes: $c_1, c_2, \ldots, c_{12}$. These cyber nodes collect real-time data from their attached transmission lines and send it to the supervisory control and data acquisition (SCADA) for processing. If possible disturbances are detected by the SCADA, some transmission lines might be temporarily disconnected through the flexible alternating current transmission system (FACTS) to stop the propagation of the disturbance. Here the transmission lines $p_1, p_2, \ldots, p_6$ are the physical nodes. In this system, the malicious attacker may inject false data into the cyber nodes to mislead the SCADA to send false disconnection commands.

To characterize the independence matrix $\mathbf{B}$, observe that locally collected data often gives a better indication of the real-time operation state of a transmission line and should have a larger impact on the decision of disconnecting that line. Here we use the following setting in the 5-bus PJM system:

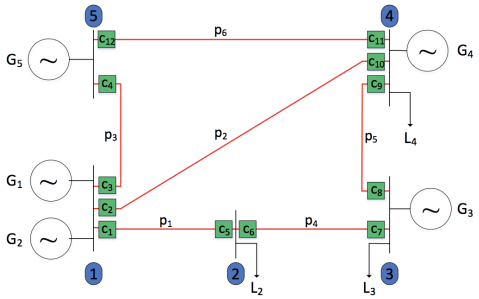


Fig. 2. PJM 5-bus system [16]

$$b_{ij} = \begin{cases} 0.25 & \text{if } c_i \text{ is locally connect to } p_j \\ 0.05 & \text{otherwise.} \end{cases}$$

To model the cost of loss of each physical node, we use the increased economic dispatch cost due to the temporary disconnection of a transmission line. Let δ_0 be the minimum dispatch cost without loss of any transmission line and δ_j be the minimum dispatch cost with the loss of transmission line j. Furthermore, denote by T the length of a time slot. Then we have

$$\rho_j = (\delta_0 - \delta_j)T. \tag{10}$$

In our setting, we set T to be 1 hour. We can use MATPOWER [22] to run seven optimal power flow to obtain the following values for all ρ_j: $\rho_1 = \$700$, $\rho_2 = \$1,000$, $\rho_3 = \$600$, $\rho_4 = \$1,200$, $\rho_5 = \$800$, and $\rho_6 = \$900$.

For simplicity, as with [16], we assume that the attacker is interested in putting down a single transmission line by compromising the two cyber nodes that have the most effect on this line. Therefore, the action spaces for both the attacker and the defender reduce to the set of the pairs of cyber nodes that are connected to the same transmission line, i.e., $(c_1, c_5), (c_2, c_{10}), (c_3, c_4), (c_6, c_7), (c_8, c_9), (c_{11}, c_{12})$, which is equivalent to the set of the transmission lines in the system. Therefore, the state of the system becomes the set of transmission lines whose pairs of cyber nodes are working correctly. The successful attacking and defending probabilities μ and λ for each pair of cyber nodes are set to be 1 and 0.5, respectively. The attacking and defending budgets for each time slot are both set to be 1, meaning that the attacker/defender can at most attack/defend one pair of cyber nodes in a time slot.

B. Numerical Results

First, we show the player strategies in a static game that does not consider rewards in future time periods as comparison. This is equivalent to set $\gamma = 0$. Table I shows the payoff matrix $Q(s, a, d)$ of the static game for state $s = [1, 1, 1, 1, 1, 1]$ corresponding to the case that all cyber nodes are working correctly. Note that the matrix represents the payoff to the attacker, and therefore, the attacker prefers an action that

returns a larger number while the defender prefers an action that returns a smaller number. The rows of the table correspond to the actions taken by the attacker and the columns of the table correspond to the actions taken by the defender. The optimal player strategies for the static game at this state are $\pi_a = [0.2207, 0.1919, 0, 0.1766, 0.2102, 0.2006]$ and $\pi_b = [0.0137, 0.2728, 0, 0, 0.4110, 0.1083, 0.1943]$. However, if we consider future rewards, the payoff matrix for this state will evolve during the iterative process and finally converge to a composite payoff matrix. For instance, if we set $\gamma = 0.3$, the payoff matrix at this state will finally evolve into the matrix shown in Table II. Note that for any state s, the optimal strategies of the players in the Markov game is the same as the optimal strategies of the players in the static game with the composite matrix as payoffs [17].

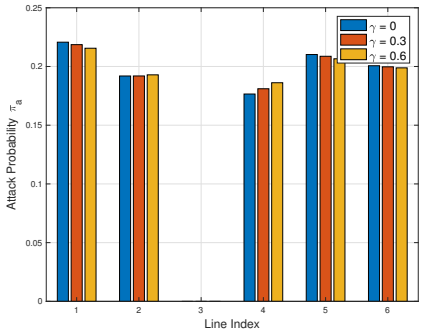
TABLE I
PAYOFF MATRIX $Q(s, a, d)$ FOR A STATIC GAME WITHOUT CONSIDERING FUTURE REWARDS UNDER STATE $s = [1, 1, 1, 1, 1, 1]$

400	800	800	800	800	800
920	460	920	920	920	920
760	760	380	760	760	760
1000	1000	1000	500	1000	1000
840	840	840	840	420	840
880	880	880	880	880	440

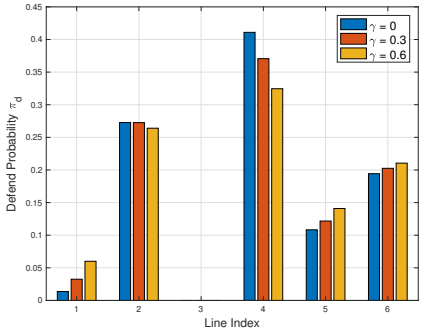
TABLE II
PAYOFF MATRIX $Q(s, a, d)$ FOR A MARKOV GAME WITH $\gamma = 0.3$ UNDER STATE $s = [1, 1, 1, 1, 1, 1]$

956.3	1475.1	1475.1	1475.1	1475.1	1475.1
1619.3	1028.4	1619.3	1619.3	1619.3	1619.3
1425.8	1425.8	931.6	1425.8	1425.8	1425.8
1690.4	1690.4	1690.4	1063.9	1690.4	1690.4
1524.4	1524.4	1524.4	1524.4	980.9	1524.4
1573.2	1573.2	1573.2	1573.2	1573.2	1005.4

Next, we show the optimal player strategies of the Markov game in the dynamic case. Fig. 3–5 show the optimal player strategies of the Markov game in state $s = [1, 1, 1, 1, 1, 1]$, $s = [0, 1, 1, 1, 1, 1]$ and $s = [1, 0, 0, 0, 0, 0]$ for $\gamma = \{0, 0.3, 0.6\}$, respectively. Obviously, the results show that the optimal strategies of the players change a lot as the value of the penalty factor γ varies. Specifically, the figures show that in general, as γ increases, the attacker will progressively shift its focus from attacking the cyber nodes on some lines to attacking other cyber nodes, while the defender will also shift its focus to minimize the total cost. Note that in our simulation setting, the budgets of the attack and the defend are both 1. Therefore, depending on the specific system status, as the attacker shift its focus to attack the cyber nodes on a specific transmission line, the defender may use this opportunity to restore the operation of other damaged cyber nodes if they are more important (e.g., the results shown in Fig. 4) or defend against such attacks directly if the damage of the specific line is high (e.g., the results shown in Fig. 5) to minimize the cost.



(a)



(b)

Fig. 3. Optimal player strategies in state $s = [1, 1, 1, 1, 1, 1]$ of the Markov game with different values of γ: (a) attacker strategy; (b) defender strategy.

V. CONCLUSION

In this paper, we have investigated the interactions between the attacker and the defender in a CPS and developed a general game-theoretic framework to analyze them. A probabilistic interdependence model has been introduced to characterize the impact of cyber attacks on the underlying physical system. A Markov game has been formulated to model the CPS security problem, considering both the continual nature of the interaction process and the interdependence model. We have applied our framework to study the security problem in the wide-area monitoring and protection of smart grid. Numerical results show the effectiveness of the derived optimal defending strategies from our proposed approach.

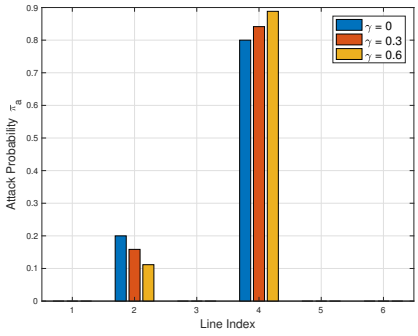
REFERENCES

[1] The future of security: A combination of cyber and physical defense, Sep. 2016. [Online]. Available: http://www.networkworld.com/article/3125476/security/the-future-of-security-a-combination-of-cyber-and-physical-defense.html

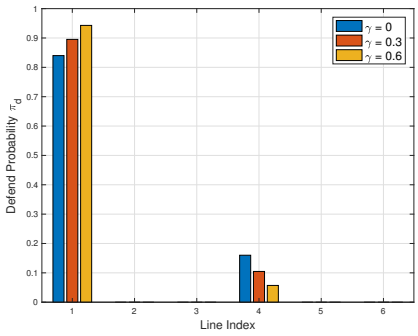
[2] When Cybersecurity Meets Physical Security, Jan. 2017. [Online]. Available: https://www.forbes.com/sites/kalevleetaru/2017/01/13/when-cybersecurity-meets-physical-security/

[3] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *Distributed Computing Systems Workshops, 2008. ICDCS'08. 28th International Conference on*. IEEE, 2008, pp. 495–500.

[4] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

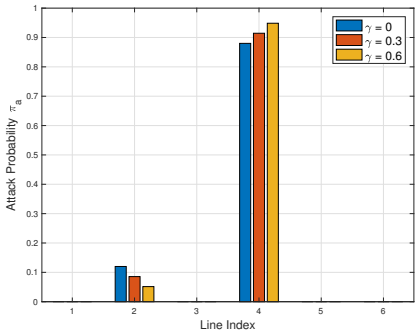


(a)

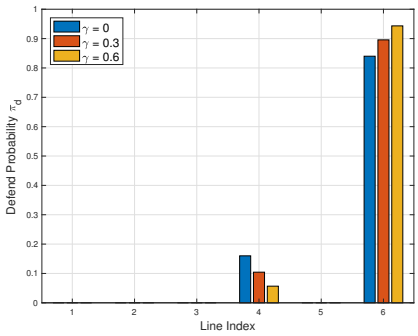


(b)

Fig. 4. Optimal player strategies in state $s = [0, 1, 1, 1, 1, 1]$ of the Markov game with different values of γ: (a) attacker strategy; (b) defender strategy.



(a)



(b)

Fig. 5. Optimal player strategies in state $s = [1, 1, 1, 1, 1, 0]$ of the Markov game with different values of γ: (a) attacker strategy; (b) defender strategy.

[5] Q. Zhu, C. Rieger, and T. Başar, "A hierarchical security architecture for cyber-physical systems," in *Resilient Control Systems (ISRCS), 2011 4th International Symposium on*. IEEE, 2011, pp. 15–20.

[6] Y. Yuan, Q. Zhu, F. Sun, Q. Wang, and T. Başar, "Resilient control of cyber-physical systems against denial-of-service attacks," in *Resilient Control Systems (ISRCS), 2013 6th International Symposium on*. IEEE, 2013, pp. 54–59.

[7] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[8] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[9] S. Lakshminarayana, T. Z. Teng, D. K. Yau, and R. Tan, "Optimal attack against cyber-physical control systems with reactive attack mitigation," in *Proceedings of the Eighth International Conference on Future Energy Systems*. ACM, 2017, pp. 179–190.

[10] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012.

[11] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Computer Networks*, vol. 57, no. 5, pp. 1344–1371, 2013.

[12] C. T. Do, N. H. Tran, C. Hong, C. A. Kamhoua, K. A. Kwiat, E. Blasch, S. Ren, N. Pissinou, and S. S. Iyengar, "Game theory for cyber security and privacy," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, p. 30, 2017.

[13] Y. W. Law, T. Alpcan, and M. Palaniswami, "Security games for voltage control in smart grid," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*. IEEE, 2012, pp. 212–219.

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

[15] J. Hao, E. Kang, D. Jackson, and J. Sun, "Adaptive defending strategy for smart grid attacks," in *Proceedings of the 2nd Workshop on Smart Energy Grid Security*. ACM, 2014, pp. 23–30.

[16] A. Sanjab and W. Saad, "On bounded rationality in cyber-physical systems security: Game-theoretic analysis with application to smart grid protection," in *Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG), Joint Workshop on*. IEEE, 2016.

[17] R. B. Myerson, *Game Theory*. Harvard University Press, 1991.

[18] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *International Conference on Machine Learning*, 1994, pp. 157–163.

[19] R. Bellman, "A markovian decision process," *Journal of Mathematics and Mechanics*, pp. 679–684, 1957.

[20] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*. John Wiley & Sons, 2012.

[21] F. Li and R. Bo, "Small test systems for power system economic studies," in *IEEE Power and Energy Society General Meeting*, 2010.

[22] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2011.