Multiact Dynamic Game Strategy for Jamming Attack in Electricity Market

Jinghuan Ma, Yuting Liu, Lingyang Song, *Senior Member, IEEE*, and Zhu Han, *Fellow, IEEE*

***Abstract*— As the current power grid system is upgrading to the smart grid, it becomes more vulnerable to security attacks on its communication subsystem such as the denial-of-service attack. Jamming, as a kind of denial-of service attack, can be applied to interfere the real-time communication in smart grid. In this paper, we analyze the scenario in which the attacker can jam a reduced number of signal channels carrying measurement information in order to manipulate the locational marginal price and create the opportunity for gaining profit, and the defender is able to guarantee a limited number of channels in information delivery. Based on the electricity marketing model, we propose a multiact dynamic game between the attacker and defender, in which the optimal strategies are taken by the two sides to maximize their own profits. We study the gaming process and discuss the prosperities of the outcome. Simulation results present the affect of jamming attack on the electricity prices and the gained profits of the two sides. Moreover, they confirm the optimality of the proposed scheme in pursuing profit.**

***Index Terms*—Electrical market, game theory, security, smart grids.**

I. Introduction

SMART GRID is an emerging cyber-physical system integrating power infrastructures with communication technologies [1]. The well-deployed sensor network in the smart grid provides observations to identify the current operating state such as the transmission line loadings and bus voltage, and strongly supports the online monitoring and state estimation [2] by control center [such as supervisory control and data acquisition (SCADA) center] to guarantee a reliable operation of the power system. However, attacks on the cyber-physical system can cause malfunctions of the electricity market or the power system [3]. As for the physical side, a novel electrical topological model based on weighted undirected graph is proposed for structural vulnerability analysis of power grids in [4]. In [5], a hybrid approach on complex networks has been proposed to analyze the structural vulnerability of power transmission networks. Meanwhile, a number of researches have been conducted over cyber security for smart grid [6]–[9]. Liu *et al*. [6] presented an undetectable attack method based on the Jacobian matrix. The design method of information security protection architecture in U.S. smart grid and information security protection requirements of China smart grid were analyzed in [7]. In [8], based on the hierarchical information and communication model, the information security risks and information security protection demands of smart grid were studied. In [9], a novel criterion of reliable strategies for defending power systems was derived and two allocation algorithms were developed to seek reliable strategies.

Specifically, the denial-of-service (DoS) attack on communication infrastructure in the smart grid is a severe threat, in which the attacker tries to prevent the remote sensors from sending measurement information to the control center, causing the instability of the power system or even regional blackout. One of the DoS attacks is the jamming on the physical layer of the grid's communication networks [10], [11]. Until now, many works have been done over jamming attacks in wireless sensor networks (WSNs) [12]–[16], but few have paid attention to the jamming attack in smart grid. During the attack, the jammer emits undesired signals to the communication channel to interfere the ongoing measurement data transmission, resulting in the incompleteness of the received real-time measurement information in the control center. Due to the jamming, the online monitoring and state estimation may fail to reflect the actual operating state of the system, and the corresponding electricity price will be calculated in error [17], [18]. Common consumers and the power supplier may suffer an economic loss from deploying the false electricity prices, while the jammer can profit from the price gap in the electrical power market. Hence, it is critical to ensure the grid system's robustness against the jamming attack.

The jamming attack, in general case, always lasts for a long term. Once the attack is launched, the detection module equipped with sensor nodes is triggered to inform the control center for countermeasures. However, when the control center responds to take action after the detection, the attacker can further change the jamming targets to continue its attack. Instead of reacting to the detected jamming, with the dense and well-organized WSN in smart grid, the control center can take preset measures to guarantee the transmission of measurement

Manuscript received July 2, 2014; revised December 6, 2014 and January 17, 2015; accepted January 28, 2015. Date of publication February 26, 2015; date of current version August 19, 2015. This work was supported in part by the National 973 Project under Grant 2013CB336700; in part by the National Nature Science Foundation of China under Grant 61222104 and Grant U1301255; in part by the Ph.D. Programs Foundation of the Ministry of Education of China under Grant 20110001110102; and in part by the U.S. National Science Foundation (NSF) under Grant CMMI-1434789, Grant CNS-1443917, Grant ECCS-1405121, and Grant CNS-0953377. Paper no. TSG-00670-2014.

J. Ma, Y. Liu, and L. Song are with the State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China.

Z. Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA (e-mail: zhan2@uh.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2015.2400215

information, considering the fact that an jammer with limited attacking choices intends to attack the a bus with higher reward.

Thus, the attacker and the control center constitute a game, where the participators choose a limit number of buses out of the bus system to attack/guarantee the transmissions of measurement information. To analyze the attacker and defender's strategies, the game theoretic approaches have been applied in smart grid to simulate the optimization of the strategy choices during the attack [19]–[21]. In this paper, we divide the whole attack and defense process into time slots, each of which is defined as an independent level, where both attacker and defender decide their strategies based on their observation and prediction to attain optimal profit. The backward induction algorithm for finite dynamic game provides optimal strategies for the players in all stages during a long-term gaming, and is adopted in this paper to tackle the jamming attack and defense problem. Our contributions are summarized as below.

1) We study the impact of the jamming attack on the electricity market and propose countermeasures to antagonize attacks for security in smart grid.
2) We adopt the multiact dynamic game and investigate the strategy equilibrium between the attacker and defender. We propose a backward induction-based algorithm to find the saddle-point solutions in all the levels and thus achieve the Nash equilibrium solution of the dynamic game.
3) The simulation results confirm the effectiveness of the proposed algorithm over the PJM five-bus test system.

The remainder of this paper is organized as follows. The system model is provided in Section II. The elements of the proposed game are defined in Section III, and the Nash equilibrium is analyzed in Section IV. The numerical results is provided in Section V and the conclusion is provided in Section VI.

II. System Model

In this section, we study the power state estimation in transmission system, which provides the real-time information of power demand and generation. Then, we investigate the pricing mechanism optimal power flow (OPF) and the locational marginal price (LMP) that have been applied in the electricity market.

A. Power System State Estimation

In the state estimation, the control center obtains the observation of m real-time measurements from n sensors among the network with phase angles ϕ_i. Since the voltage phase (ϕ_i) of a reference bus is fixed and known, we only have to estimate $(n-1)$ left unknown. We define the state vector as $\phi = [\phi_1, \dots, \phi_n]^T$ and the observed vector $\mathbf{P}$ for m active power measurements [22], related with the active power, which can be described as follows [23]:

$$\mathbf{P} = \mathbf{p}(\phi) + \epsilon \tag{1}$$

where $\mathbf{P} = [P_1, \dots, P_m]^T$ denotes the vector of measured active power in transmission lines, $\mathbf{p}(\cdot)$ is the nonlinear relation between measurements, ϕ denotes the vector of n bus phase angles ϕ_i, and $\epsilon = [\epsilon_1, \dots, \epsilon_m]^T$ is the Gaussian measurement noise vector with covariant matrix Σ_ϵ. The Jacobian matrix $\mathbf{H} \in \mathbb{R}^m$ is defined as

$$\mathbf{H} = \frac{\partial \mathbf{p}(\phi)}{\partial \phi} \,|_{\phi=\mathbf{0}}. \tag{2}$$

Since the phase difference $(\phi_i - \phi_j)$ is small, (1) can be reduced to the following linear approximation:

$$\mathbf{P} = \mathbf{H}\phi + \epsilon. \tag{3}$$

The bad data can be injected to $\mathbf{P}$ to impact the state estimation of ϕ. Given the power flow measurements $\mathbf{P}$, the estimated state vector $\hat{\phi}$ can be computed as

$$\hat{\phi} = \left(\mathbf{H}^T \Sigma_\epsilon^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^T \Sigma_\epsilon^{-1} \mathbf{P} = \mathbf{BP} \tag{4}$$

where $\mathbf{B} = \mathbf{H}\left(\mathbf{H}^T \Sigma_\epsilon^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^T \Sigma_\epsilon^{-1}$.

The false data detection can be performed with the residue vector $\mathbf{r}$ computed from the difference between measured vector and the estimated value based on the endogenous parameters: $\mathbf{r} = \mathbf{P} - \mathbf{H}\hat{\phi}$. With the given threshold value for test whether the cyber has been attacked, the hypothesis of not being attacked should satisfy [24]

$$\max_i |r_i| \le \gamma. \tag{5}$$

B. DC OPF and LMP

OPF is adopted to provide the constraints of optimization of electricity allocation in power systems. The locational marginal pricing methodology has been the primary approach in electricity markets to set electricity prices and deal with transmission congestion. On the basis of the OPF model, LMPs are classified into tow types: 1) day-ahead LMP; and 2) real-time market.

1) Day-Ahead LMP: The linear form of dc OPF to predict the day-ahead electricity price in the market is proved to be effective in generation prescheduling with static parameters [25]–[27]. Then, LMP at each bus of the power network is decided by the linear programming solution of the problems described as

$$\min_{\mathbf{G_i}}. \quad \sum_{i=1}^{N} C_i \times G_i \tag{6}$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^N G_i - \sum_{i=1}^N D_i = 0 \\ \sum_{i=1}^N \mathrm{GSF}_{k-i} \times (G_i - D_i) \le \mathrm{Limit}_k^{\max}, \ k \in \kappa \\ G_i^{\min} \le G_i \le G_i^{\max}, \ i \in \Im \end{cases}$$

where N denotes the number of buses, C_i denotes the generation cost at bus i in (\$/MWh), G_i is the generation dispatch at bus i in (MWh), GSF_{k-i} denotes the generation shift factor from bus i to line k, κ is the set of all lines in the grid, $\Im$ is the set of all generators, and $\mathrm{Limit}_k^{\max}$ denotes the transmission limit for line k. In particular, D_i is the demand for the electricity, which is a one-variable function of the measurement $\mathbf{P}$.

In the day-ahead market, the general formulation of LMP at bus i (LMP_i) is consist of three components, including locational marginal energy price ($\mathrm{LMP}_i^{\mathrm{energy}}$), locational marginal congestion price ($\mathrm{LMP}_i^{\mathrm{cong}}$), and locational marginal loss price ($\mathrm{LMP}_i^{\mathrm{loss}}$)

$$\mathrm{LMP}_i = \mathrm{LMP}_i^{\mathrm{energy}} + \mathrm{LMP}_i^{\mathrm{cong}} + \mathrm{LMP}_i^{\mathrm{loss}} \tag{7}$$

$$\mathrm{LMP}^{\mathrm{energy}} = \lambda \tag{8}$$

$$\mathrm{LMP}_i^{\mathrm{cong}} = \sum_{k=1}^{L} \mathrm{GSF}_{k-i} \times \mu_k \tag{9}$$

$$\mathrm{LMP}_i^{\mathrm{loss}} = \lambda \times (\mathrm{DF}_i - 1) \tag{10}$$

where L denotes the number of lines, λ denotes the Lagrangian multiplier of the equality constraint, μ_k denotes the Lagrangian multiplier of the kth transmission constraint, and DF_i denotes the delivery factor at bus i. In order to emphasize the main part of LMP_i, we assume that the optimization model in (7) ignores losses, and we have $\mathrm{DF}_i = 1$ and $\mathrm{LMP}_i^{\mathrm{loss}} = 0$ in (10). Thus, LMP_i can be described as

$$\mathrm{LMP}_i = \lambda + \sum_{k=1}^{L} \mathrm{GSF}_{k-i} \times \mu_k. \tag{11}$$

2) Real-Time LMP: In the real-time market, an ex-post market model is based on the run time data. The real-time LMP in the real-time market is deduced from the dc Optimal Power Flow (DCOPF) model with the change of value in power flow on each bus for the real-time electricity dispatch, which satisfies the following incremental linear programming [28]:

$$\min_{\Delta \mathbf{G_i}}. \quad \sum_{i=1}^{N} C_i \times \Delta G_i \tag{12}$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^{N} \Delta G_i - \sum_{i=1}^{N} \Delta D_i = 0 \\ \sum_{i=1}^{N} \mathrm{GSF}_{k-i} \times (\Delta G_i - \Delta D_i) \le 0, \ k \in \mathcal{C} \\ \Delta G_i^{\min} \le \Delta G_i \le \Delta G_i^{\max}, \ i \in \Im \end{cases}$$

where $\mathcal{C}$ is the set of estimated congestion lines, which are defined as

$$\mathcal{C} = \left\{ l: \sum_{i=1}^{N} \mathrm{GSF}_{l-i} \times (G_i - D_i) \ge \mathrm{Limit}_l^{\max} \right\}. \tag{13}$$

In practice, the upper and lower bound of ΔG_i is set as 0.1–2.0 MWh. With the assumption of $\mathrm{DF}_i = 1$, then LMPs in the real-time market can be depicted as

$$\hat{\mathrm{LMP}}_i = \hat{\lambda} + \sum_{k=1}^{L} \mathrm{GSF}_{k-i} \times \hat{\mu}_k \tag{14}$$

where $\hat{\lambda}$ denotes the Lagrangian multiplier of the equality constraint in the incremental linear programming and $\hat{\mu}_k$ denotes the Lagrangian multiplier of the kth transmission constraint in the set of congestion lines.

III. Jamming Attack and Defense

A. Jamming Attack Procedure

Manipulating the prices in electricity market is the incentive for the jammer to launch the attack. The pricing mechanism depends on the state estimation from the sensors. However, when being jammed, the measurements from state estimators are unavailable to the control center [29]. We adopt a discrete-time model of jamming attacks, in which time is divided into time slots. We note that the jammer will only attack a limited number of sensors out of the whole WSN in smart grid, mainly because of the following.

1) An excessive jamming attack can cause power blackout, resulting in failure to manipulate the price.
2) A wide-area jamming attack will seriously increase the risk of being detected.

The procedure of jamming attack is given below.

1) At the beginning of a time slot, the attacker jams specific channels in the network to cause measurements unavailable, leaving real-time prices at corresponding buses ($\hat{LMP}_i^{RT}$) undecided.
2) The control center will use default values to substitute lost measurements for the dc OPF model.
3) The attacker keeps monitoring the power market and jamming the insecure measurements during a whole time slot.
4) With the access to real-time measurements, the attacker can predict real-time prices after ceasing the jamming.
5) Comparing the real-time during and after jamming, the attacker will buy electricity at lower price and sell electricity at higher price to profit from the difference between two prices.

B. Jamming Attack Strategies

With online monitoring of power systems, the transmitted power load on the transmission lines can be depicted in a linear model as

$$\hat{p}_{\mathrm{ij}} = \frac{\phi_i - \phi_j}{X_{\mathrm{ij}}} = \frac{(B_i - B_j)^T}{X_{\mathrm{ij}}} \mathbf{P} = \mathbf{M}^T \mathbf{P} \tag{15}$$

where $\mathbf{M}^T = (B_i - B_j)^T / X_{\mathrm{ij}}$, B_i, and B_j are the ith and jth components of the vector denoted in (4), respectively. Based on (13), we can find the linear relation between $\hat{p}_{\mathrm{ij}}$ and $\mathbf{P}$, which is an N-dimensional vector reflecting the measurement of the voltage angle on different transmission lines.

1) During Jamming: With no state estimation received from equipments during the jamming attack, the control center substitutes the default value $\mathbf{P}_{\mathrm{def}}$ for sensors jammed by the attacker. Once the control center detects the lost signals, the default values are required for DCOPF to price the electricity in the market. Consequently, the lost demand datas in (12) ΔD_i are replaced by the predetermined values $\Delta D_i(\mathbf{P}_{\mathrm{def}})$.

The optimal result with the constraints with default values can be deduced from the DCOPF model, given LMP_i^{jam}

$$\hat{LMP}_i^{\mathrm{jam}} = \hat{\lambda}^{\mathrm{jam}} + \sum_{k=1}^{L} \mathrm{GSF}_{k-i} \times \hat{\mu}_k^{\mathrm{jam}}. \tag{16}$$

It is proved that default values of measurements jammed can directly impact assumed values of transmitted power. The optimal dispatch strategy based on the incorrect assumption leads to the deviation of the electricity price in the market during jamming attacks.

2) After Jamming: At the end of a time slot, the attacker ceases jamming, so the control center receives the real-time estimation again. The DCOPF program decides the real-time price when sensors report the changes in measurements among the grid. Then, the price during the jamming is obviously different from the real-time price. The inequality condition after jamming is altered in the form as

$$\sum_{i=1}^{N} \mathrm{GSF}_{k-i} \times (G_i - D_i(\Delta\mathbf{P})) \leq 0 \tag{17}$$

where $\Delta\mathbf{P}$ is an N-dimensional measurement that the control center obtains from the monitoring system. Given the DCOPF program, the price at bus i, denoted as LMP_i^{AJ}, is given as

$$L\hat{MP}_i^{AJ} = \lambda^{\hat{A}J} + \sum_{k=1}^{L} \mathrm{GSF}_{k-i} \times \hat{\mu}^{AJ}. \tag{18}$$

Given the definition of two prices LMP_i^{jam} and LMP_i^{AJ}, we can clearly define the profit that the attacker gains from the attack during one time slot. We assume that the attacker will gain the whole difference between two prices at every bus i

$$\Delta\mathbf{L} = \sum_i \left| L\hat{M}P_i^{\mathrm{jam}} - L\hat{M}P_i^{AJ} \right| \tag{19}$$

where $\Delta\mathbf{L}$ is the gross profit for per unit of electricity that the attacker can gain from the whole difference of LMP at every bus i during and after the jamming attack.

C. Defense Strategies

We investigate the existing countermeasures against jamming attack in WSN and study the techniques applicable for WSNs in smart grid. Li *et al.* [29] proposed an anti narrowband jamming technique where the remote sensor can utilize multiple channels to deliver information and avoid the jamming interference. Cagalj *et al.* [30] traded-off the network robustness with its complexity and cost, and assigned a portion of pairs of sensor nodes, one of which is out of the jammed area, to create wormhole communication links to pass the information out of a jammed area. Xu *et al.* [31] utilized channel surfing method involving on-demand frequency hopping to defend jamming attack, and studied two different approaches to channel surfing. The coordinated channel switching requires the entire sensor network to adjust its channel while the spectral multiplexing assigns the nodes in a jammed region to switch channels and nodes on the boundary of a jammed region as radio relays between different spectral zones.

As smart grid manages to develop a large-scare WSNs [32], it is feasible to deploy alternative sensors to monitor the state of a bus and create multiple paths to deliver the measurement information. The techniques of wireless power transfer and energy harvesting will enable the sensors in smart grid to carry a long-term monitoring service [33]. Hence, the control center is able to adopt the available anti-jamming techniques in the WSNs in smart grid. For considerations on energy saving, in each monitoring round the control center will only assign a limit number of bus to utilize the defense strategies in data transmissions.

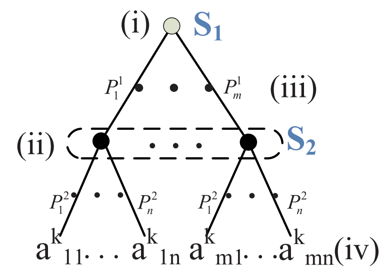


Fig. 1. Elements of a single-act game in extensive form. (i) Starting node: $\mathbf{S}_1$'s action stage. (ii) Enumerated action nodes: independent action stage of $\mathbf{S}_2$. (iii) Strategy/action branches: $\Gamma^1 = \{P_1^1, \ldots, P_m^1\}$, set of player's possible actions. (iv) Outcome terminals: final payoffs denoted by $\mathbf{a}_{\mathrm{mn}}^k$.

IV. Attacker and Defender Gaming

In this section, we firstly introduce the single-act game solutions in both pure strategy and mixed strategy. Then we turn to the dynamic game with multiple stages and specifically construct the recursive algorithm to analyze behaviors in the electricity market.

A. Single-Act Games in Extensive Form

The multiact dynamic game of complete information consists of a series of single-act games of complete information in the time order, where a single-act game is also named a sub-game [34]. We note that the symbols for defining a single-act game will contain subscripts and superscripts, for which we regard it as a sub-game of the dynamic game based on which to further introduce the entire dynamic gaming process. A single-act two-person zero sum game is defined as $\mathcal{G}^k$,[1] in which the two players, denoted, respectively by $\mathbf{S}_1$ and $\mathbf{S}_2$, compete with each other for more profit given the zero sum of gains [35]. As the extensive form is convenient to display the process of a multistage game, we also adopt this form to introduce the single-act game besides matrix form.

A single-act two-person zero-sum game is structured as a finite tree, as shown in Fig. 1. The basic structure in the tree is a node indicating the action stage of a player with branches representing every possible strategy of the player. The branch may point to another node representing the action stage of the other player, or to a payoff value indicating the end of the game. Let P_j^i denote the jth strategy of player i and Γ^i denote the strategy set of i. $\Gamma^1 = \{P_1^1, \ldots, P_m^1\}$, i.e., $\mathbf{S}_1$ has totally m strategies, and $\Gamma^2 = \{P_1^2, \ldots, P_n^1\}$. The game allows players to act successively or simultaneously. In this scenario, we assume that the attacker and defender choose their strategies simultaneously. $\mathbf{S}_2$ does not know $\mathbf{S}_1$'s strategy when choosing its own strategy. We model this by encircling all the stage nodes of $\mathbf{S}_2$ that correspond to $\mathbf{S}_1$'s possible actions with the dashed lines. The payoff function maps a strategy pair $\{P_m^1, P_n^2\}$ uniquely to a payoff vector $\mathbf{a}_{\mathrm{mn}}^k$ represents the gained profits of the players. The matrix form of the game is given in Table I. Each entry of the matrix a_{mn} corresponds to an end point assigned by a particular pair of strategies $\{P_m^1, P_n^2\}$ taken by both players, all of which represent the payoffs of the game $\mathcal{G}_k$.

[1] $\mathcal{G}^k$ also represents the kth level sub-game of the dynamic game.

TABLE I
kTH SINGLE-ACT GAME $\mathcal{G}^k$ IN MATRIX FORM

Def. \ Att.	P_1^2	...	P_t^2
P_1^1	a_{11}^k	...	a_{1t}^k
...	...	a_{mn}^k	...
P_s^1	a_{s1}^k	...	a_{st}^k

Let $J(\gamma^1, \gamma^2) \geq 0$ denote the attacker's profit by successful attack [$-J(\gamma^1, \gamma^2)$ is the defender's loss] where $\gamma^1 \in \Gamma^1$ and $\gamma^2 \in \Gamma^2$, so that $a_{\rm mn} = J(\gamma^1 = P_m^1, \gamma^2 = P_n^2)$. Here, we do not include the cost of launching jamming attack in the game, so that $J(\gamma^1, \gamma^2)$ is nonnegative. A pair of strategies $\{\gamma^{1*} \in \Gamma^1, \gamma^{2*} \in \Gamma^2\}$ is in saddle-point equilibrium if the following set of the inequalities is satisfied $\forall \gamma^1 \in \Gamma^1, \gamma^2 \in \Gamma^2$:

$$J\left(\gamma^{1*}, \gamma^2\right) \leq J\left(\gamma^{1*}, \gamma^{2*}\right) \leq J\left(\gamma^1, \gamma^{2*}\right) \tag{20}$$

where $J(\gamma^{1*}, \gamma^{2*})$ is the saddle-point value of the zero-sum game. To find the saddle-point(s) in a single-act game, we have to enumerate all the possible outcomes of the game, i.e., calculate all the payoffs at the end points of the game tree or every entry of the game matrix. In all the possible outcomes, a saddle point $(\gamma^{1*}, \gamma^{2*})^k$ of game $\mathcal{G}^k$ has to satisfy (20).

The satisfaction of the condition to be a saddle point can be expressed in two different ways depending on the given form of the strategies.

1) Saddle-Point of Pure Strategy: Given a $(s \times t)$ matrix game $\mathcal{G}^k = \{a_{\rm mn}\}^k$, $(\{\text{row } m^*, \text{column } n^*\})^k$ constitutes a saddle-point equilibrium of pure strategy when the inequality below is satisfied for all $a_{\rm mn}^k \in \mathcal{G}^k$

$$a_{m^*n}^k \leq a_{m^*n^*}^k \leq a_{mn^*}^k \tag{21}$$

and $a_{m^*n^*}^k$ is the value of the saddle-point. In this case, no players have the incentive to betray the equilibrium, such that the game is running stably under the same strategy choices.

2) Saddle-Point of Mixed Strategy: In this case, the strategy of a player is a probability distribution on its strategy space. For example, an allowable strategy for the defender is to choose P_1^1 *w.p.* y_1, P_2^1 *w.p.* $y_2, \ldots, P_s^1$ *w.p.* y_s, where $\sum_{i=1}^s y_i = 1$ and likewise, the attacker is allowed to choose P_1^2 *w.p.* x_1, P_2^2 *w.p.* $x_2, \ldots, P_s^2$ *w.p.* x_s, where $\sum_{i=1}^s x_i = 1$. Let y and x, respectively denote the probability distribution vectors as $\mathbf{y} = (y_1, \ldots, y_s)', \mathbf{x} = (x_1, \ldots, x_t)$, and s-dimensional simplex Y and t-dimensional simplex X, respectively denote the two players' strategy spaces. Hence, the average value of the outcome of the game is expressed as

$$J(\mathbf{y}, \mathbf{x})^k = \sum_{m=1}^{s} \sum_{n=1}^{t} y_m a_{\rm mn}^k x_n = \mathbf{y}' \mathcal{G}^k \mathbf{x}. \tag{22}$$

In a $(s \times t)$ matrix game $\mathcal{G}^k$, the defender and attacker try to minimize and maximize the value of $J(\mathbf{y}, \mathbf{x})$, respectively, by the appropriate choice of the probability distribution vector $\mathbf{y} \in Y$ and $\mathbf{x} \in X$. In any matrix game, the average security levels of the players in mixed strategies coincide, that is

$$\overline{V}_{\mathcal{B}}(\mathcal{G}^k) = \min_Y \max_X \mathbf{y}' \mathcal{G}^k \mathbf{x} = \max_X \min_Y \mathbf{y}' \mathcal{G}^k \mathbf{x} = \underline{V}_{\mathcal{B}}(\mathcal{G}^k) \tag{23}$$

where $\overline{V}_{\mathcal{B}}$ is the average security level of the defender (equivalently the average upper value of the game) and $\underline{V}_{\mathcal{B}}$ is the average security level of the attacker (equivalently the average lower value of the game). Hence, as for an $(m \times n)$ matrix game, a saddle point of mixed strategy is comprised the mixed security strategies for both players, in the form of $\{\gamma_{1*}, \gamma_{2*}\} = \{\mathbf{y}^*, \mathbf{x}^*\}$ which satisfies (23). Hence, the mixed-strategy equilibrium is uniquely given by

$$V_{\mathcal{B}}(\mathcal{G}^k) = \overline{V}_{\mathcal{B}}(\mathcal{G}^k) = \underline{V}_{\mathcal{B}}(\mathcal{G}^k). \tag{24}$$

B. Dynamic Games Between Attacker and Defender

While attacking, the attacker may constantly change the target sensor for mainly two reasons.

1) The defender may succeed in avoiding the effect of jamming when it has exactly protected the information transmission under attack.
2) A static jamming will increase the risk of being detected and punished.

The defender also has to adjust its defense actions facing a dynamic attacker. Hence, both the attacker and defender should adjust their strategies according to the observation of both players' past choices in different levels. Their behaviors can be modeled by a multiact zero-sum game.

Let $\mathcal{N}$ denote the set of all K-level strategy profiles of the players. We define $\mathcal{A} = (K, (\Gamma^i)_{i \in \mathcal{R}}, (U_i)_{i \in \mathcal{N}})$ as a game in which, the defender and attacker compete to compromise and defend the insecure measurements in set $\mathcal{N}$ within K levels. The aims of attacker and defender are to increase and decrease the change in LMP, respectively. The game is described as follows.

1) *Players Set:* $\mathcal{R} = \{1, 2\}$ the defender (the No. 1 player) and the attacker (the No. 2 player).
2) *Strategies:* Attacker chooses measurements of a bus to attack at different levels in order to get the maximum profit $V(\mathcal{A})$. Defender choose measurements of a bus to protect at different levels in order to minimize the profit $V(\mathcal{A})$.
3) *Strategy Set Γ^i:* The set of available strategies for player i, $\Gamma^1 = \{P_1^1, \ldots, P_m^1\}$, and $\Gamma^2 = \{P_1^2, \ldots, P_n^2\}$, where m and n are the maximum number of strategies of all insecure measurements and their profiles that the attacker and defender can choose from.
4) *Utility:* $U_2 = \sum_{k=1}^K h^k \cdot \Delta L^k((\gamma^1, \gamma^2)^k)$ and $U_1 = -U_2$ for the attacker and defender, respectively. h^k is the profited electricity amount at level k and $\Delta L^k(\cdot)$ is the per unit price difference at level k.

We assume that the defense and attack process lasts for K levels. The behaviors of the defender and attacker are modeled by a multiact discrete-time game tree illustrated in Fig. 2. A typical strategy profile of a player is composed of K components as $(\gamma_1^1, \ldots, \gamma_K^1)$ is for the defender and $(\gamma_1^2, \ldots, \gamma_K^2)$ for the attacker, where $\gamma_i^1 \in \Gamma_i^1$ and $\gamma_j^2 \in \Gamma_j^2$ are the strategy pair at the jth level. We denote the set of all strategies of $\mathbf{S}_i$'s at the jth level of play by Γ_j^i.

1) Properties of the Dynamic Game: Before illustrating the dynamic game theoretic algorithm, we introduce the properties of dynamic game solutions, which guarantee the optimality

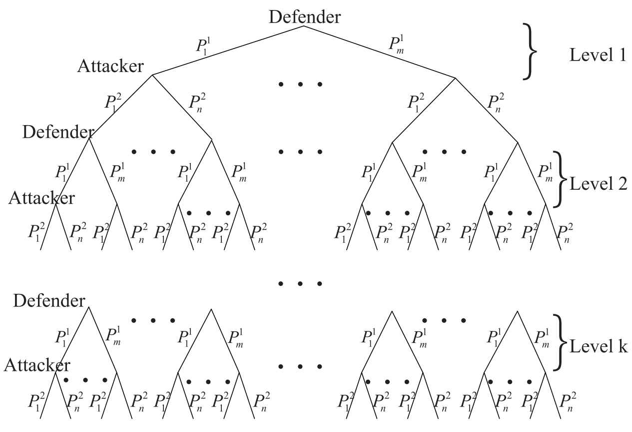


Fig. 2. Two-person zero-sum game in extensive form with K levels.

TABLE II
RECURSIVE ALGORITHM FOR OPTIMIZATION IN MULTIACT GAME

1. Solve each single-act sub-game $\mathcal{G}_i^K$ assigned by $\mathbf{S_2}$'s information set division at Kth level by the enumeration operations.
2. Find every $\{\gamma_K^{1*}, \gamma_K^{2*}\}$ satisfying (20) for all $\mathcal{G}_i^K$.
3. Compute corresponding outcomes $\mathrm{V}(\mathcal{G}_i^K)$ in pure strategies or $V_{\mathcal{B}}(\mathcal{G}_i^K)$ in mixed strategies when two players take the strategy $(\gamma_K^{1*}, \gamma_K^{2*})_i$.
4. Combine all these pairs at level K to establish the final solution at Kth level, $\{(\gamma_K^{1*}, \gamma_K^{2*})_i\}$.
5. Assign outcomes $\mathrm{V}(\mathcal{G}_i^K)$ to every vertex corresponding to each single-act game $\mathcal{G}_i^K$ and then the K-level game $\mathcal{A}$ is simplified into a $(K-1)$-level game.
6. Repeat step 2, 3, 4 and 5 until $\mathcal{A}$ is simplified into a single-act game and record every saddle-point solution at each level i, $\{(\gamma_i^{1*}, \gamma_i^{2*}); i = 1, ..., K\}$.

of the outcome assigned by the saddle-point-strategy profiles.

1) *Backward Induction:* The algorithm with backward induction begins with finding the optimal strategies in the final level K for every possible historical action branches in level $K-1$. Then, it backs to level $K-1$ and determines the optimal choices in that level. The algorithm proceeds this operation to the initiate level and constructs a strategy profile which is a Nash equilibrium. It guarantees that every player's actions are optimal at every possible history.
2) *Subgame Perfection:* A subgame is a subset of a dynamic game. When the game is carried out in divided time slots, we need to establish a multiact game model. A strategy profile of dynamic game is in the subgame perfect Nash equilibrium, if it constitutes a Nash equilibrium in each of the subgames.
3) *Dynamic Programming:* Dynamic programming in multiact games requires the value reassignment through the backward induction. The vertexes of all single-act subgames at current level will be considered as the end nodes of the previous level. Then, the recursive procedure is required to simplify the multiact feedback games from K-level games to $(K-1)$-level games, and ultimately the single-act ones with the optimal outcome.

Compared with the repeated game, players in the dynamic game do not have to keep their stable strategies. With complete information at the current level and that of all the past choices, they are able to evaluate every possible outcome combining the effect of the other player's behavior at the current stage and try to find an optimal strategy for the current stage. With the above properties of the dynamic game, both players can reach the optimal solutions of the dynamic game. The multiact game is divided by time slots and the decisions made in different time slots do not interact with each other, so the payoffs of the attacker at each level will be independently addictive.

2) Algorithm for the Defense and Attack Game: Next, we introduce the proposed algorithm for the defense and attack process to find the equilibrium solution of the multiact game, as shown in Table II. Based on decomposition of the multiact game, the algorithm begins with determining the optimal choices for all sub-game in level K and operate recursively till the first level is reached, satisfying the property of backward induction. For every recursion, the operation of dynamic programming simplifies the game structure. Step-by-step, the saddle point equilibrium strategy pairs, $\{\gamma_1^{1*}, \gamma_1^{2*}\}, \ldots, \{\gamma_K^{1*}, \gamma_K^{2*}\}$ at each level can be computed. We prove that all these pairs constitute a entire Nash equilibrium strategy profile of the dynamic game. This method provides the players with a systematic procedure to find the most effective strategy to maximize their payoffs.

Proposition 1: The saddle-point solutions of the dynamic game $\mathcal{A}$ satisfy recursively the following set of K pairs of inequalities for all $\gamma_j^i \in \Gamma_j^i$, $i = 1, 2; j = 1, \ldots, K$:

$$\left.\begin{array}{c}
J(\gamma_1^1, \ldots, \gamma_{K-1}^1 \gamma_K^{1*}; \gamma_1^2, \ldots, \gamma_K^2) \le J(\gamma_1^1, \ldots, \gamma_K^{1*}; \\
\gamma_1^2, \ldots, \gamma_{K-1}^2, \gamma_K^{2*}) \le J(\gamma_1^1, \ldots, \gamma_K^1; \gamma_1^2, \ldots, \gamma_{K-1}^2, \gamma_K^{2*}) \\
J(\gamma_1^1, \ldots, \gamma_{K-2}^1, \gamma_{K-1}^{1*}, \gamma_K^{1*}; \gamma_1^2, \ldots, \gamma_{K-1}^2, \gamma_K^{2*}) \\
\le J(\gamma_1^1, \ldots, \gamma_{K-2}^1, \gamma_{K-1}^{1*}, \gamma_K^{1*}; \gamma_1^2, \ldots, \gamma_{K-2}^2, \gamma_{K-1}^{2*}, \gamma_K^{2*}) \\
\le J(\gamma_1^1, \ldots, \gamma_{K-1}^1, \gamma_K^{1*}; \gamma_1^2, \ldots, \gamma_{K-2}^2, \gamma_{K-1}^{2*}, \gamma_K^{2*}), \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
J(\gamma_1^{1*}, \ldots, \gamma_K^{1*}; \gamma_1^2, \gamma_2^{2*}, \ldots, \gamma_K^{2*}) \le J(\gamma_1^{1*}, \ldots, \gamma_K^{1*}; \\
\gamma_1^{2*}, \ldots, \gamma_K^{2*}) \le J(\gamma_1^1, \gamma_2^{1*}, \ldots, \gamma_K^{1*}; \gamma_1^{2*}, \ldots, \gamma_K^{2*}).
\end{array}\right\} \tag{25}$$

Proof: According to the specific characteristic of the established model, the function of the value of outcomes in the whole game, $J(\gamma_1^1, \gamma_2^1, \ldots, \gamma_K^1; \gamma_1^2, \gamma_2^2, \ldots, \gamma_K^2)$ is independently addictive. That is, the value of the multiact game is added by the value of the every single-act sub-game in different levels together. It can be described as

$$J\left(\gamma_1^1, \ldots, \gamma_{K-1}^1, \gamma_K^1; \gamma_1^2, \ldots, \gamma_{K-2}^2, \gamma_{K-1}^2, \gamma_K^2\right) = J\left(\gamma_1^1, \gamma_1^2\right) + \cdots + J\left(\gamma_K^1, \gamma_K^2\right). \tag{26}$$

We denote $J(\gamma_j^1, \gamma_j^2)_i$ as the ith possible profit the attacker can gain at level j. Based on the algorithm introduced before, every saddle-point solution at level j, $J(\gamma_j^{1*}, \gamma_j^{2*})_i$ for corresponding $\mathcal{G}_i^j$ satisfies the inequality below

$$\begin{aligned}J&\left(\gamma_1^1,\gamma_2^1,\ldots,\gamma_{j-1}^1,\left(\gamma_j^{1*}\right)_i;\gamma_1^2,\gamma_2^2,\ldots,\left(\gamma_j^2\right)_i\right)\\&\le J\left(\gamma_1^1,\gamma_2^1,\ldots,\gamma_{j-1}^1,\left(\gamma_j^{1*}\right)_i;\gamma_1^2,\gamma_2^2,\ldots,\gamma_{j-1}^2,\left(\gamma_j^{2*}\right)_i\right)\\&\le J\left(\gamma_1^1,\gamma_2^1,\ldots,\left(\gamma_j^1\right)_i;\gamma_1^2,\gamma_2^2,\ldots,\gamma_{j-1}^2,\left(\gamma_j^{2*}\right)_i\right).\end{aligned}\tag{27}$$

Then, we minus $J(\gamma_1^1,\gamma_2^1,\ldots,\gamma_{j-1}^1;\gamma_1^2,\gamma_2^2,\ldots,\gamma_{j-1}^2)$ on both sides of the inequality sign. Together with the additivity, (27) will be altered into the inequality below for all $\mathcal{G}_i^j$

$$J\left(\gamma_j^{1*},\gamma_j^2\right)_i\le J\left(\gamma_j^{1*},\gamma_j^{2*}\right)_i\le J\left(\gamma_j^1,\gamma_j^{2*}\right)_i.\tag{28}$$

As for the final strategy pair at level j, $\{\gamma_j^{1*},\gamma_j^{2*}\}$ is the combination of all $(\gamma_j^{1*},\gamma_j^{2*})_i$. Then, it obviously leads to the conclusion

$$J\left(\gamma_i^{1*},\gamma_i^2\right)\le J\left(\gamma_i^{1*},\gamma_i^{2*}\right)\le J\left(\gamma_i^1,\gamma_i^{2*}\right).\tag{29}$$

Thus, with the same mathematical tricks to add the same value to the items on both sides of the inequality sign, we can deduce the set of inequalities (25) by

$$\begin{aligned}&\sum_{i=1}^{j}J\left(\gamma_i^1,\gamma_i^2\right)+J\left(\gamma_{j+1}^{1*},\gamma_{j+1}^2\right)+\sum_{i=j+2}^{K}J\left(\gamma_i^{1*},\gamma_i^{2*}\right)\\&\le\sum_{i=1}^{j}J\left(\gamma_i^1,\gamma_i^2\right)+J\left(\gamma_{j+1}^{1*},\gamma_{j+1}^{2*}\right)+\sum_{i=j+2}^{K}J\left(\gamma_i^{1*},\gamma_i^{2*}\right)\\&\le\sum_{i=1}^{j}J\left(\gamma_i^1,\gamma_i^2\right)+J\left(\gamma_{j+1}^1,\gamma_{j+1}^{2*}\right)+\sum_{i=j+2}^{K}J\left(\gamma_i^{1*},\gamma_i^{2*}\right).\end{aligned}\tag{30}$$

According to (26), for all $j=1,2,\ldots,(K-2)$, (30) can be changed into

$$\begin{aligned}J&\left(\gamma_1^1,\ldots,\gamma_j^1,\gamma_{j+1}^{1*},\ldots,\gamma_K^{1*};\gamma_1^2,\ldots,\gamma_{j+1}^2,\gamma_{j+2}^{2*},\ldots,\gamma_K^{2*}\right)\\&\le J\left(\gamma_1^1,\ldots,\gamma_j^1,\gamma_{j+1}^{1*},\ldots,\gamma_K^{1*};\gamma_1^2,\ldots,\gamma_j^2,\gamma_{j+1}^{2*},\ldots,\gamma_K^{2*}\right)\\&\le J\left(\gamma_1^1,\ldots,\gamma_{j+1}^1,\gamma_{j+2}^{1*},\ldots,\gamma_K^{1*};\gamma_1^2,\ldots,\gamma_j^2,\gamma_{j+1}^{2*},\ldots,\gamma_K^{2*}\right).\end{aligned}\tag{31}$$

■

C. Discussions on the Algorithm

In this subsection, we compare the proposed algorithm with the simply repeated algorithm and discuss the scalability of the algorithm.

1) Algorithm Comparison: The multiact game can be solved with different algorithms one of which is the simply repeated algorithm. Different from the dynamic programming algorithm, simply repeated games ignore the information evolution, with mere consideration of the best outcome in the current level. The players take simply repeated algorithm will repeat their choices during the whole process without any strategy evolution based on the observation of the history. In this case, the maximization will be processed for one time at the first level, followed with the repetition in the rest levels. Then, backward induction is not required in simply repeated algorithm. Both players in simply repeated games will repeat their strategies at the start in the following levels, so they fix their strategy choice on $\{\gamma_1^{1*},\gamma_1^{2*}\}$. Then, two players' strategy sequences can be denoted as $(\gamma_1^{1*},\ldots,\gamma_1^{1*};\gamma_1^{2*},\ldots,\gamma_1^{2*})$. The optimality proved in (31) shows that simply repeated cannot perform as well as the dynamic programming algorithm, since for $\forall i\in\{1,2,\ldots,K\}$

$$J\left(\gamma_1^{1*},\gamma_1^{2*}\right)\le J\left(\gamma_i^{1*},\gamma_i^{2*}\right)\tag{32}$$

is satisfied for all two-person zero-sum dynamic games with information evolution.

2) Discussion on Scalability: As the proposed algorithm is based on backward induction, it begins with the last level and has to enumerate all the possible strategy pairs at that level. Let N_1 and N_2, respectively denote the number of strategies held by players 1 and 2, and K denote the level of dynamic game. The time complexity of the proposed algorithm is $O\left((N_1N_2)^{K+1}\right)$. If we assume that the defender can choose n_1 out of m_1 buses to defense and the attacker can attack n_2 out of m_2 buses, we have $N_1=\binom{m_1}{n_1}$ and $N_2=\binom{m_2}{n_2}$. From the above, we can see that the increase of game level will effectively increase the time complexity as the work of evaluations has largely increased. As for n_1 and n_2, we see that with higher capabilities to attack and defense, the players will have more possible choices of the strategy combinations, causing the complexity to increase.

TABLE III
LINE REACTANCE AND THERMAL LIMIT FOR FIVE-BUS TEST SYSTEM

Line	L_{12}	L_{14}	L_{15}	L_{23}	L_{34}	L_{45}
X (%)	2.81	3.04	0.64	1.08	2.97	2.97
$Limit_k^{max}(MW)$	999	999	999	999	999	240

TABLE IV
GENERATION SHIFT FACTORS OF LINES IN FIVE-BUS TEST SYSTEM

Line \ Bus	B_1	B_2	B_3	B_4	B_5
L_{1-2}	0.1939	-0.476	-0.349	0	0.1595
L_{1-4}	0.4376	0.258	0.1895	0	0.36
L_{1-5}	0.3685	0.2176	0.1595	0	-0.5195
L_{2-3}	0.1939	0.5241	-0.349	0	0.1595
L_{3-4}	0.1939	0.5241	0.6510	0	0.1595
L_{5-4}	0.3685	0.2176	0.1595	0	0.4805

V. NUMERICAL RESULTS

A. Parameters

We analyze the effect of attack on the PJM five-bus test system in [36] with some slightly modifications. Transmission lines' parameters are given in Tables III and IV, generators' and loads' parameters (including $G_i^{\max}$, C_i, and D_i) in Fig. 3. The default values of the measurements are shown in Table V. These default values are utilized to substitute corresponding insecure measurements, when the real-time measurements have been jammed. Fig. 4 demonstrates the effect of jamming attack on the LMPs when the measurement P_5 has been attacked within a 5 min operation. The gross profit for per unit of electricity gained by the attack $\Delta\mathbf{L}=25$ (\$/MWh).

TABLE V
DEFAULT VALUES OF MEASUREMENTS

MEASUREMENT	P_1	P_2	P_3	P_4	P_5	P_6
VALUE(MW)	250	340	-180	170	500	370
MEASUREMENT	P_7	P_8	P_9	P_{10}	P_{11}	
VALUE(MW)	300	-80	220	300	-300	

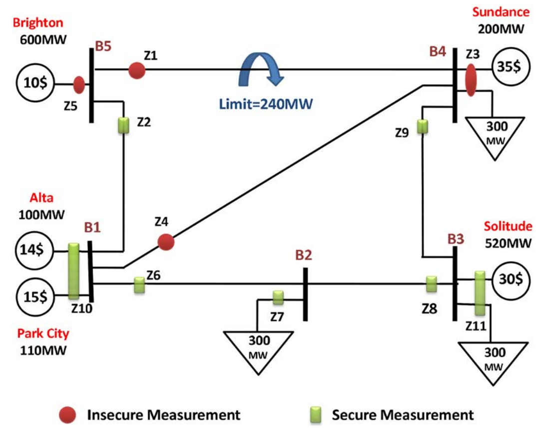


Fig. 3. Measurement configuration in PJM five-bus test system.

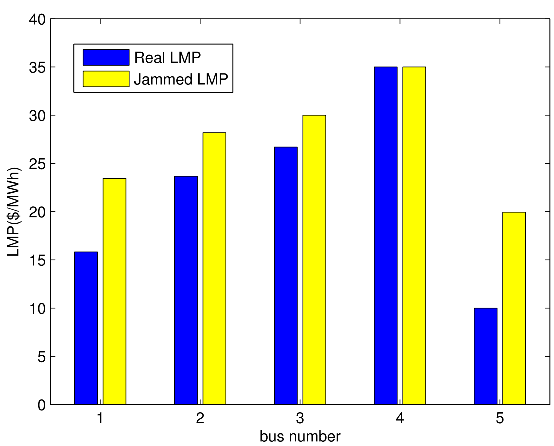


Fig. 4. Comparison of LMPs with and without attack.

If the attacker buys 10 MWh electricity at every bus before attacking and sells it when successfully jamming, it can gain \$250 in total.

B. Two-Person Zero-Sum Dynamic Games

In the real electricity power systems, we suppose that there are three insecure measurements $\{\mathcal{P}_1, \mathcal{P}_3, \mathcal{P}_5\}$, only one of which can be compromised by the attacker at each level. Once the previous transmitted data is jammed, the system will be aware of it. Different from static games, the attacker and the defender can adapt their choices with the observation of the past strategy sequences, but their choices at the current level will not be aware of by each other.

It is assumed that only when the attacker chooses the same $\mathcal{P}_i$ with the defender, it will not be compromised through jamming; otherwise, the jamming is successfully achieved with the change in profit. In Table VI, the change in LMP when the attacker successfully compromises one of measurements in $\mathcal{S}$ at one level is provided.

TABLE VI
CHANGES IN PROFIT FROM DIFFERENT MEASUREMENTS

MEASUREMENT	$\mathcal{P}_1$	$\mathcal{P}_3$	$\mathcal{P}_5$
ΔLMP(\$ /MWh)	17.17	3.18	25.00

TABLE VII
FEEDBACK GAMES IN MATRIX FORM WITH TWO LEVELS

$\mathbf{S}_1$ \ $\mathbf{S}_2$	P_1P_1	P_1P_3	P_1P_5	P_3P_1	P_3P_3	P_3P_5	P_5P_1	P_5P_3	P_5P_5
P_1P_1	0	3.18	25	3.18	6.36	28.18	25	28.18	50
P_1P_3	17.17	0	25	20.35	3.18	28.18	42.17	25	50
P_1P_5	17.17	3.18	0	20.35	6.36	3.18	42.17	28.18	25
P_3P_1	17.17	20.35	42.17	0	3.18	25	25	28.18	50
P_3P_3	34.34	17.17	42.17	17.17	0	25	42.17	25	50
P_3P_5	34.34	20.35	17.17	17.17	3.18	0	42.17	28.18	25
P_5P_1	17.17	20.35	42.17	3.18	6.36	28.18	0	3.18	25
P_5P_3	34.34	17.17	42.17	20.35	3.18	28.18	17.17	0	25
P_5P_5	34.34	20.35	17.17	20.35	6.36	3.18	17.17	3.18	0

TABLE VIII
VALUES OF J CORRESPONDING TO DEFENDER'S INFORMATION SETS AT SECOND LEVEL

NO.	1	2	3	4	5	6	7	8	9
J^*(\$/MWh)	3.18	6.36	28.18	20.35	3.18	28.18	20.35	6.36	3.18

Here, we assume a two-level attack, in which all values assigned to terminal nodes in Fig. 1 is provided in Table VII, in which the possible outcomes are decided by defender's two choices in order represented in rows and the attacker's two choices in columns. In each level, the attacker will buy 10 MHw at every bus before attacking and sell the power when jamming. Then, we will show how the attacker optimizes its profit in this two-level attack. Obviously, more levels involved will only require for more steps to repeat the same optimization procedure in our analysis.

C. Results of Dynamic Recursive Algorithm

Applying the proposed algorithm, we start at the second level (the last level). The recursive procedure requires nine single-act saddle-point solutions corresponding to the defender's nine information sets at this level. The outcomes belonging to different single-act games is calculated from Table VII. After the integration of all mixed strategies $\hat{\gamma}_2^{1*}$ and $\hat{\gamma}_2^{2*}$ which satisfies (20), we have

$$\hat{\gamma}_2^{1*} = \begin{cases} \left.\begin{array}{l} P_1, \mathbf{w.p.}0.33 \\ P_5, \mathbf{w.p.}0.67 \end{array}\right\} \text{if } \gamma_1^1 = \gamma_1^2 \\ \left.\begin{array}{l} P_3, \mathbf{w.p.}0.32 \\ P_5, \mathbf{w.p.}0.68 \end{array}\right\} \text{if } \gamma_1^1 \neq \gamma_1^2,\ \gamma_1^2 = P_1 \\ \left.\begin{array}{l} P_1, \mathbf{w.p.}0.29 \\ P_5, \mathbf{w.p.}0.71 \end{array}\right\} \text{if } \gamma_1^1 \neq \gamma_1^2,\ \gamma_1^2 = P_3 \\ \left.\begin{array}{l} P_1, \mathbf{w.p.}0.38 \\ P_5, \mathbf{w.p.}0.62 \end{array}\right\} \text{if } \gamma_1^1 \neq \gamma_1^2,\ \gamma_1^2 = P_5, \end{cases} \quad \hat{\gamma}_2^{2*} = P_3. \tag{33}$$

Then, with the optimal strategy in the second level given, we can simplify the original game into the single-act one with its terminal points. Moreover, all values of $\{J_1^*, J_2^*, \ldots, J_9^*\}$ are shown in Table VIII. Similarly, through the same procedure,

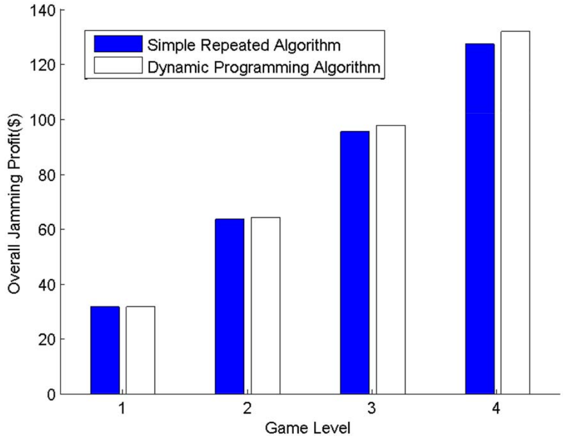


Fig. 5. Dynamic optimality between two algorithms.

the saddle-point equilibrium of the first level denoted by $(\gamma_1^{1*}, \gamma_1^{2*})$ is calculated as

$$\hat{\gamma}_1^{1*} = \begin{cases} P_1, \mathbf{w.p.}0.29 \\ P_5, \mathbf{w.p.}0.71, \end{cases} \qquad \hat{\gamma}_1^{2*} = P_3. \tag{34}$$

Then, we can solve the final value $\Delta\mathbf{L} = 6.36$. Thus, the benefit achieved by the attacker is $10 \times 6.36 = \$63.6$.

To sum-up, the attacker choosing the optimal strategy $\{\gamma_1^{1*}, \gamma_2^{1*}; \gamma_1^{2*}, \gamma_2^{2*}\}$ to jam the bus on which the electricity transmitted will be paid for \$6.36 per unit to the attacker. Finally, the dynamic solution can provide the players with the information evolution to optimize their strategy with the strategy selection sequences. Both players can take the advantage of the information available to maximize their profit and reach an equilibrium at each level.

We can find that the saddle-point equilibrium at each level is in the mixed strategy, all of which altogether constitutes the optimal outcome in the long term. Besides, the dynamic programming and backward induction are necessary for information evolution, during which both attacker and defender's optimal choices at the different levels are not static. We can see the difference between the dynamic game with multiple levels and the simple repetitive game, which lies in the availability to past choices of the other player participating the game. From the saddle-point solution in the second level, we can find that their strategies are based on their observation of the past choice sequences. Every time they decide what to choose at the beginning of the game at each level, they will base their choices with the consideration of what the other one has chosen before.

D. Dynamic Optimality Comparison With Simple Repeated Algorithm

The players in simple repeated games are unaware of the past strategy sequences chosen by the other participants in the game. They repeat their choices constantly at each level, regardless of what happened in the previous stage. In such situation, the outcome of the multiact game will be exactly linear in the amount of the profit paid to the attacker at each level. Then, we can find that the attacker's profit from two different algorithms are given in Fig. 5. As for single-act games, two algorithms will be quite the same with each other. Obviously, with the number of the game level increasing, the algorithm involved with the dynamic evolution better improves the attacker's final outcome.

VI. CONCLUSION

In this paper, we introduced the pricing mechanism and the method attackers utilize to change the congestion and the electricity price. Then, we formulated the optimization problem of maximizing attacker's profit from the most effective strategy choices with the context of the theory about multi-act two-person zero-sum game with the extensive forms. We constructed the detailed algorithm to solve the problem step by step and give the further demonstration of its optimality. In simulation, we gave the specific example of a PJM five-bus test system, in which we provided the detailed procedure shown in Table II to find the saddle-point equilibrium of the game at each level, all of which altogether were combined to build the final solution in a multiact game.

REFERENCES

[1] T. F. Garrity, "Getting smart," *IEEE Power Energy Mag.*, vol. 6, no. 2, pp. 38–45, Mar./Apr. 2008.
[2] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.
[3] H. Li and Z. Han, "Manipulating the electricity power market via jamming the price signaling in smart grid," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, TX, USA, Dec. 2011, pp. 1168–1172.
[4] G. Chen *et al.*, "Attack structural vulnerability of power grids: A hybrid approach based on complex networks," *Phys. A*, vol. 389, no. 3, pp. 595–603, Feb. 2010.
[5] G. Chen *et al.*, "An improved model for structural vulnerability analysis of power networks," *Phys. A*, vol. 388, no. 19, pp. 4259–4266, Oct. 2009.
[6] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, Chicago, IL, USA, Nov. 2009, pp. 21–32.
[7] T. Zhang *et al.*, "The design of information security protection framework to support smart grid," in *Proc. Int. Conf. Power Syst. Technol. (POWERCON)*, Hangzhou, China, Oct. 2010, pp. 1–5.
[8] Y. Wang, B. Zhang, W. Lin, and T. Zhang, "Smart grid information security—A research on standards," in *Proc. Int. Conf. Adv. Power Syst. Autom. Protect. (APAP)*, Beijing, China, Oct. 2011, pp. 1188–1194.
[9] G. Chen, Z. Y. Dong, D. J. Hill, and Y. S. Xue, "Exploring reliable strategies for defending power systems against targeted attacks," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1000–1009, Aug. 2011.
[10] E. Altman, K. Avrachenkov, and A. Garnaev, "Jamming game with incomplete information about the jammer," in *Proc. ICST/ACM Int. Workshop Game Theory Commun. Netw.*, Pisa, Italy, Dec. 2009.
[11] G. T. Amariucai and S. Wei, "Mixed anti-jamming strategies in fixedrate wireless systems over fast fading channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jun. 2009, pp. 2522–2526.
[12] H. M. Sun, S. P. Hsu, and C. M. Chen, "Mobile jamming attack and its countermeasure in wireless sensor networks," in *Proc. 21st Int. Conf. Adv. Inf. Netw. Appl. Workshops (AINAW)*, Niagara Falls, ON, Canada, May 2007, pp. 457–462.
[13] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A survey on jamming attacks and countermeasures in WSNs," *IEEE Commun. Surv. Tuts.*, vol. 11, no. 4, pp. 42–56, Dec. 2009.
[14] Z. Lu, W. Wang, and C. Wang, "From jammer to gambler: Modeling and detection of jamming attacks against time-critical traffic," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 1871–1879.
[15] M. Li, I. Koutsopoulos, and R. Poovendran, "Optimal jamming attack strategies and network defense policies in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 8, pp. 1119–1133, Aug. 2010.

[16] B. Wang, Y. Wu, K. Liu, and T. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 877–889, Apr. 2011.

[17] T. Orfanogianni and G. Gross, "A general formulation for LMP evaluation," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1163–1173, Aug. 2007.

[18] R. Frowd and A. Papalexopoulos, "Market simulation for LMP forcasting," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Calgary, AB, Canada, Jul. 2009, pp. 1–6.

[19] A. Holmgren, E. Jenelius, and J. Westin, "Evaluating strategies for defending electric power networks against antagonistic attacks," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 76–84, Feb. 2007.

[20] Z. M. Fadlullah, Y. Nozaki, A. Takeuchi, and N. Kato, "A survey of game theoretic approaches in smart grid," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Nov. 2011, pp. 1–4.

[21] P. H. Nguyen, W. L. Kling, and P. F. Ribeiro, "A game theory strategy to integrate distributed agent-based functions in smart grids," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 568–576, Mar. 2013.

[22] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "On malicious data attacks on power system state estimation," in *Proc. 45th Int. Univ. Power Eng. Conf. (UPEC)*, Cardiff, U.K., Aug./Sep. 2011, pp. 1–6.

[23] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. Hoboken, NJ, USA: Wiley, 1996.

[24] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York, NY, USA: Marcel Dekker, 2004.

[25] F. Li and R. Bo, "DCOPF-based LMP simulation: Algorithm, comparison with ACOPF, and sensitivity," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1475–1485, Nov. 2007.

[26] F. Li, J. Pan, and H. Chao, "Marginal loss calculation in competitive electrical energy markets," in *Proc. IEEE Int. Conf. Elect. Util. Deregul. Restruct. Power Technol. (DRPT)*, Hong Kong, Apr. 2004, pp. 205–209.

[27] E. Litvinov, T. Zheng, G. Rosenwald, and P. Shamsollahi, "Marginal loss modeling in LMP calculation," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 880–888, May 2004.

[28] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 528–534, May 2003.

[29] H. Li, L. Lai, and R. C. Qiu, "A denial-of-service jamming game for remote state monitoring in smart grid," in *Proc. 45th Annu. Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 2011, pp. 1–6.

[30] M. Cagalj, S. Capkun, and J. Hubaux, "Wormhole-based anti-jamming techniques in sensor networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 1, pp. 1–15, Jan. 2007.

[31] W. Xu, T. Wood, W. Trappe, and Y. Zhang, "Channel surfing: Defending wireless sensor networks from interference," in *Proc. IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Cambridge, MA, USA, Apr. 2007, pp. 499–508.

[32] A. Nasipuri, L. Zel, and R. McKosky, "Design considerations for a large-scale wireless sensor network for substation monitoring," in *Proc. IEEE 35th Conf. Local Comput. Netw. (LCN)*, Denver, CO, USA, Oct. 2010, pp. 866–873.

[33] M. Erol-Kantarci and H. Mouftah, "Suresense: Sustainable wireless rechargeable sensor networks for the smart grid," *IEEE Wireless Comm.*, vol. 19, no. 3, pp. 30–36, Jun. 2012.

[34] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.

[35] *Overview of the Smart Grid: Policies, Initiatives and Needs*, ISO New England Inc., 2009.

[36] F. Li and R. Bo, "Small test systems for power system economic studies," in *Proc. Power Energy Soc. Gen. Meeting*, Minneapolis, MN, USA, Jul. 2010, pp. 1–4.

Jinghuan Ma received the B.S. degree in electronic engineering from Peking University, Beijing, China, in 2013, where he is currently pursuing the Ph.D. degree in signal and information processing.

His current research interests include optimization, game theory, 5G communications, and smart grid communications.

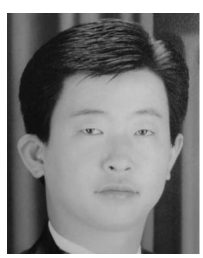
Yuting Liu received the B.S. degree in electronic engineering from Peking University, Beijing, China, in 2014. She is currently pursuing the M.S. degree in electrical and computer engineering from the University of Wisconsin–Madison, Madison, WI, USA.

Her current research interests include computer system and communication network.

Lingyang Song (S'03–M'06–SM'12) received the Ph.D. degree in electronics engineering from the University of York, York, U.K., in 2007.

He was a Post-Doctoral Research Fellow at the University of Oslo, Oslo, Norway, and Harvard University, Cambridge, MA, USA, until rejoining Philips Research, Cambridge, U.K., in 2008. In 2009, he was a Full Professor at the School of Electronics Engineering and Computer Science, Peking University, Beijing, China. His current research interests include multiple-input multiple-output, orthogonal frequency-division multiplexing, cooperative communications, cognitive radio, physical layer security, game theory, and wireless *ad-hoc*/sensor networks. He has authored/co-authored over 100 journal and conference papers, and holds a number of patents (standard contributions).

Dr. Song was the recipient of the 2012 IEEE Asia Pacific Young Researcher Award; the 2012 National Science Foundation of China Outstanding Young Investigator Award; the Best Paper Award at the IEEE International Conference on Wireless Communications, Networking, and Mobile Computing in 2007; the Best Paper Award at the First IEEE International Conference on Communications in China in 2012; the Best Student Paper Award at the 7th International Conference on Communications and Networking in China in 2012; the Best Paper Award at the IEEE Wireless Communication and Networking Conference in 2012; and the K. M. Stott Prize for his excellent research. Since 2012, he has been an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.



Zhu Han (S'01–M'04–SM'09–F'14) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, USA, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer at JDS Uniphase, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an Assistant Professor at Boise State University, Boise, ID, USA. He is currently an Associate Professor with the Electrical and Computer Engineering Department, University of Houston, TX, USA. His current research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communication.

Prof. Han was the recipient of the IEEE Fred W. Ellersick Prize in 2011 and an National Science Foundation CAREER Award in 2010. Since 2010, he has been an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is currently an IEEE Distinguished Lecturer.