On False Data-Injection Attacks against Power System State Estimation: Modeling and Countermeasures

Qingyu Yang, *Member*, *IEEE*, Jie Yang, Wei Yu, Dou An, Nan Zhang, and Wei Zhao, *Fellow*, *IEEE*

Abstract—It is critical for a power system to estimate its operation state based on meter measurements in the field and the configuration of power grid networks. Recent studies show that the adversary can bypass the existing bad data detection schemes, posing dangerous threats to the operation of power grid systems. Nevertheless, two critical issues remain open: 1) how can an adversary choose the meters to compromise to cause the most significant deviation of the system state estimation, and 2) how can a system operator defend against such attacks? To address these issues, we first study the problem of finding the optimal attack strategy—i.e., a data-injection attacking strategy that selects a set of meters to manipulate so as to cause the maximum damage. We formalize the problem and develop efficient algorithms to identify the optimal meter set. We implement and test our attack strategy on various IEEE standard bus systems, and demonstrate its superiority over a baseline strategy of random selections. To defend against false data-injection attacks, we propose a protection-based defense and a detection-based defense, respectively. For the protection-based defense, we identify and protect critical sensors and make the system more resilient to attacks. For the detection-based defense, we develop the spatial-based and temporal-based detection schemes to accurately identify data-injection attacks.

Index Terms—Cyber-physical systems, power grid, state estimation, cyber security

1 Introduction

As a typical cyber-physical system (CPS), the power grid integrates a physical power transmission system with the cyber computation and communication core [1]. It supplies electric power from generators through power transmission and distribution networks to large geographical areas. The development of a next-generation power grid, i.e., the smart grid, has received significant attention recently [2].

State estimation has been widely used by Energy Management Systems (EMS) at the control center to ensure that the power grid is running in desired states. It provides the estimation of system states in real time based on meter measurements in the field. The meter measurements are collected by the *Supervisory Control and Data Acquisition (SCADA) Systems* and processed by a state estimator to filter the measurement noise and to detect gross errors. The results of state estimation are then used by applications at the control center, for purposes such as contingency analysis, optimal power flow, economic dispatch, and others [3]. One can see that state estimation plays a critical role in the stability of power grid systems.

Meter measurements collected via the SCADA system contain not only measurement noise due to the finite accuracy of meters and communication media, but also errors caused by various issues—for example, meters with faulty connection and calibration. To reduce the impact of noise and errors, power system researchers have developed numerous methods to process meter measurements after the state estimation process [4], [5], [6]. The essential goal of these methods is to leverage the redundancy of multiple measurements to identify and remove anomalies.

While most existing techniques for protecting power grid systems were designed to ensure system reliability (i.e., against random failures), recently there have been growing concerns in smart grid initiatives on the protection against malicious cyber attacks [7], [8], [9]. There are growing concerns in the smart grid on protection against malicious cyber threats and the operation and control of smart grid depend on a complex cyberspace of computers, software, and communication technologies [10], [11]. Because the measurement component supported by smart equipment (e.g., smart meters and sensors) plays an important role, it can be a target for attacks. As those measuring devices may be connected through open network interfaces and lacking tamper-resistance hardware increases the possibility of being compromised by the adversary. For example, *Stuxnet*

- *Q. Yang, J. Yang, and D. An are with SKLMSE Lab, School of Electronics & Information Engineering, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an, Shaanxi 710049, People's Republic of China. E-mail: yangqingyu@mail.xjtu.edu.cn, yangjie.1987@stu.xjtu.edu.cn, adkaka.an@gmail.com.*
- *W. Yu is with the Department of Computer and Information Sciences, Towson University, 7800 York Road, Towson, Maryland, MD 21252. E-mail: wyu@towson.edu.*
- *N. Zhang is with the Department of Computer Science, The George Washington University, 725 23rd Street NW, Washington, DC 20052. E-mail: nzhang10@gwu.edu.*
- *W. Zhao is with the Department of Computer Science, University of Macau, Rector's Office, Av. Padre Tomás Pereira Taipa, Macau, China, and UMacau Zhuhai Research Institute, Zhuhai, Guangdong, China. E-mail: weizhao@umac.mo.*

Manuscript received 16 June 2012; revised 27 Feb. 2013; accepted 10 Mar. 2013; published online 28 Mar. 2013.
Recommended for acceptance by W. Lou.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDS-2012-06-0574.
Digital Object Identifier no. 10.1109/TPDS.2013.92.

is the first publicly known malware to exploit vulnerabilities in SCADA systems [12]. With compromised devices in the grid, the adversary can inject false measurement reports to the controller. This causes the controller to estimate wrong system states, posing dangerous threats to the operation of the power grid system.

False data-injection attack were proposed by Liu et al. [8], [13]. In such attacks, the adversary could attack the state of a power grid system. Their proposed approach could bypass existing bad data detection schemes and arbitrarily manipulate the state estimation of grid systems, posing damage on the operation of power grid. As such, in this attack, the goal of the adversary is to mislead power operation, the methodology used by the adversary is to compromise meters and inject data, and the consequence of such attacks is to let malicious data pass traditional bad data detection, posing the negative impact on many applications based on state estimation. In addition to the attack against static state estimation, false data-injection attacks can be used to attack other key functional modules in the smart grid such as energy price, distributed energy distribution, and others.

In this paper, we study a novel problem of defending against false data-injection attacks from the system operator's point of view. Because most adversaries are limited in the amount of resources they possess, we first consider a least-effort attack model—i.e., the objective of the adversary is to identify the minimum number of meters that one has to manipulate to change a predetermined number of state variables (so as to launch a false data-injection attack accordingly). We prove the NP-hardness of this problem by reduction from the *minimum subadditive join problem*.

To address this problem in a practical setting, we develop a linear transformation-based approach, which finds the optimal solution through the matrix transformation. Nevertheless, the computation complexity of the matrix transformation grows exponentially with the size of the power network. To address this issue, we develop a heuristic yet extremely efficient approach. Specifically, through the analysis of the $\mathbf{H}$ matrix, for a set of bus state variables, the adversary needs to compromise less meters when the buses are connected to one another with the largest degrees and connected to the least number of buses beyond its area. Based on this insight, we divide the network into a number of overlapping areas. The linear transformation or brute-force search (BF) can be used to identify the optimal set of meters for individual small areas and then derive the set of meters for the whole network. We have implemented our proposed heuristic-based approach on power system state manipulation on various IEEE standard buses. Our extensive experimental data validate the feasibility and effectiveness of the developed approach.

Since the false data-injection attacks may successfully change the state variables by manipulating sensor measurements based on the knowledge of the system topological information, there are two ways for the system operator to develop countermeasures. One is to make some of the critical sensors more resilient to attacks. In practice, when a power grid system contains thousands of meters and covers a large area, the cost of protecting all meters can be very high. For the protection-based defense, we identify the critical meters to protect based on the false data-injection model. Our experimental data show the superiority of our approach over the random selection of meters to protect.

The other is to develop the anomaly detection mechanisms to recognize false data-injection attacks. To continually manipulate the power grid state estimation causing extensive damage such as disrupting the contingency analysis and others, the adversary will need to manipulate measurements from a set of sensors over time. Hence, we design spatial and temporal-based detection algorithms in the control center to identify stealthy attacks. Our developed spatial-based detection algorithm is designed by leveraging the fact that when we view all the measurements received at a certain time as a unity spatially, the accumulated deviation of all the marginally compromised measurements will be significant. To detect false data-injection attacks that manipulate sensor measurements over time, we develop the temporal-based detection algorithm based on the nonparametric *cumulative sum* (cusum) change detection technique. We implement our proposed detection algorithms and test their performance using the IEEE 14-bus system. Our experimental data show that the spatial-based detection algorithm is able to recognize at least 95 percent of the false data-injection attacks once the attack changes more than 6 percent of the state variable values. Our experimental data also show that the temporal-based detection algorithm can identify the compromised meters that send manipulated measurements quickly. In comparison with the conference version of this work [14] that is only five pages, this extended journal version contains substantial new materials, including new attack and countermeasure methods and evaluation, theoretical analytical results, and others.

The remainder of this paper is organized as follows: We introduce the background of state estimation and false data-injection attacks in Section 2. In Section 3, we model the least-effort attack, which effectively computes the minimal set of meter measurements given the number of state variables to manipulate and present the hierarchical false data-injection attack strategy. In Section 4, we develop countermeasures against false data-injection attacks. In Section 5, we present our experimental results of the least-effort attack and countermeasures. In Section 6, we review the related work. Finally, we conclude the paper in Section 7.

2 Preliminaries

In this section, we first discuss the state estimation of power grid systems and then review the state-of-the-art false data-injection attack proposed in [8]. The system we focus on in this paper is actually the power transmission system—i.e., a network or power grid consisting of generators, transformers, and transmission lines and devices. The state estimator in the control center [15] uses the steady system model and data reported from a SCADA system to estimate the system state—for example, the phasor voltages at all buses—over the time. Here, "buses" can be generators, loads or transformers in a power grid system. Each meter is supposed to collect sensing measurements in the field and report them to the control center. The control center then

uses the collected measurements to estimate the state variables of power grid networks. Generally speaking, the state estimation can be formalized by

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}. \tag{1}$$

Here, $\mathbf{z} = (z_1, z_2, \ldots, z_m)^T$ represents the measurement vector. Examples of $z_i \in \mathbb{R}$ $(i = 1, 2, \ldots, m)$ include bus voltages V_i, bus active and reactive power flows P_i, Q_i, and branch active and reactive power flows P_{ij}, Q_{ij} of the power system. $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ represents the state vector, $x_j \in \mathbb{R}$ $(j = 1, 2, \ldots, n)$ can be the bus voltage magnitudes V_i and phase angles θ_i of the system buses. $\mathbf{e} = (e_1, e_2, \ldots, e_m)^T$ represents the measurement error vector, $(e_k \in \mathbb{R}$ $(k = 1, 2, \ldots, m)$, which is a zero-mean variable in Gaussian distribution). Let $cov(e_i) = \sigma_i^2, cov(\mathbf{e}) = \mathbf{R}$.

We notice that $\mathbf{h}(\mathbf{x})$ is a nonlinear vector function derived from the network topology. If the sensor measurements are based on the branch and bus power flows, the function can be derived by $P_{ij} = \frac{v_i v_j}{X_{ij}} sin(\theta_i - \theta_j) + e$, and $P_i = \sum_{j \in \aleph_i} P_{ij} + e$, where P_{ij} is the active power flow from bus i to bus j, P_i is the active power injection at bus i, θ_i and θ_j represent the phase angles of buses i and j, respectively, v_i is the bus i's voltage magnitude, X_{ij} denotes the reactance, and $\aleph_i$ represents the set of buses connected to bus i.

In this paper, we focus on the DC-state estimation model, in which the bus voltage magnitudes are already known and is equal to 1 per unit, and all shunt elements, bus and branch and reactive power flow can be ignored. Note that there is 1 meter on each bus measuring its real power injection and 2 meters on each transmission line measuring the line real power flows [3]. Then, above equations can be linearized and rewritten by

$$P_{ij} = \frac{\theta_i - \theta_j}{X_{ij}}. \tag{2}$$

Hence, for a DC-state estimation model, (1) can be simplified by a linear model and we have $\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}$, where the measurement matrix $\mathbf{H}^{\mathbf{m}\times\mathbf{n}}$ is a constant *Jacobian* matrix. For a system with N buses, if we choose bus 1 as the reference bus, i.e., $\theta_1 = 0$, the state vector in DC model becomes $\mathbf{x} = (\theta_2, \theta_3, \ldots, \theta_N)^T$.

In a DC-state estimation procedure, the number of sensor measurements is normally larger than that of the state variables $(m > n)$. Redundant measurements are used to improve the accuracy of state estimation. In particular, the objective of weighted least square (WLS) state estimation is to find $\hat{\mathbf{x}}$ and minimize $\mathbf{J}(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\hat{\mathbf{x}})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}})$. Then, $\hat{\mathbf{x}}$ can be derived by $\hat{\mathbf{x}} = (\mathbf{H}^{\mathbf{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathbf{T}}\mathbf{R}^{-1}\mathbf{z}$.

In false data-injection attacks, the adversary who has the knowledge of the network configuration of power grid (e.g., topology) can inject some of meter readings and manipulate the state variables arbitrarily. Different from the classic "unintentional" bad data case, this type of malicious attacks can effectively bypass the bad data detection. The main idea of false data-injection attack is to add a nonzero attack vector $\mathbf{a} = (a_1, a_2, \ldots, a_m)^T$ to the original sensor measurements vector $\mathbf{z}$. The observed sensor measurements vector becomes $\mathbf{z_a} = \mathbf{z} + \mathbf{a}$. The control center receives this manipulated vector and uses it to derive false estimates $\hat{\mathbf{x}}_{\mathbf{bad}}$. Let $\hat{\mathbf{x}}_{\mathbf{bad}} = \hat{\mathbf{x}} + \mathbf{c}$, $\hat{\mathbf{x}}$ be the original estimates and let $\mathbf{c}$ be the malicious errors added to the original estimates. Liu et al. [8] investigated an attack strategy, which can bypass the existing bad data test based on the principle, in which $\mathbf{a} \in \mathrm{Im}(\mathbf{H})$, i.e., $\mathbf{a} = \mathbf{H}\mathbf{c}$. The following analysis shows the feasibility of this false data-injection attack:

$$\begin{aligned} \|\mathbf{z_a} - \mathbf{H}\hat{\mathbf{x}}_{\mathbf{bad}}\|^{\mathbf{2}}_{\mathbf{R}^{-1}} &= \|\mathbf{z} + \mathbf{a} - \mathbf{H}(\hat{\mathbf{x}} + \mathbf{c})\|^{\mathbf{2}}_{\mathbf{R}^{-1}} \\ &= \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}} + (\mathbf{a} - \mathbf{H}\mathbf{c})\|^{\mathbf{2}}_{\mathbf{R}^{-1}} \\ &= \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\|^{\mathbf{2}}_{\mathbf{R}^{-1}}, \end{aligned} \tag{3}$$

where $\mathbf{a} = \mathbf{H}\mathbf{c}$. Please refer to [8] for the detailed description of this attack.

3 Least-Effort False Data-Injection Attack

In this section, we first present the formalization of the least-effort attack, which enables the adversary to find the sparsest attack vector $\mathbf{a}$ so that the attack can be conducted with the least effort by compromising the minimal necessary number of sensors. We then describe a linear transformation-based approach, which is capable of generating the optimal solution, albeit only computationally feasible for small networks. Finally, we present our heuristic-based approach for large grid networks in an efficient fashion.

3.1 Problem Formulation

We assume that the number of manipulated state variables is limited to $k(k < n)$. The adversary needs to find the sparsest attack vector $\mathbf{a}$ that satisfies the relation $\mathbf{a} = \mathbf{H}\mathbf{c}$. The least-effort attack can be formulated by

$$\begin{aligned} &\min_{\mathbf{c}} \|\mathbf{H}\mathbf{c}\|_0, \\ &\text{s.t.} \|\mathbf{c}\|_0 = k. \end{aligned} \tag{4}$$

Liu et al. [8] obtained an approximate boundary of $\|\mathbf{H}\mathbf{c}\|_0$. In particular, they conducted experiments on IEEE standard bus systems. They first randomly chose $k(1 \leq k \leq 10)$ state variables and injected malicious data (e.g., 100 times larger than the real estimates of the state variables) into each of them and then examined the number meters that need to be compromised. For a given k, they conducted the above experiment 1,000 times to examine the distribution of $\|\mathbf{H}\mathbf{c}\|_0$. As an example, when injecting the malicious data into 10 state variables, [8, Fig. 6] shows that the adversary needs to compromise 60-140 meter measurements for the IEEE 118-bus system, and 50-140 meter measurements for the IEEE 300-bus system. As we can see, one issue remains open: how can an adversary choose the meters to compromise to cause the most significant deviation of the system state estimation? We address this issue and study the false data-injection attack with the least effort.

Given k state variables to manipulate, let $\Gamma = \{i_1, i_2, \ldots, i_k\}$, g_Γ be the number of sensor measurements to compromise, the computation of g_Γ is showed in Theorem 1.

Theorem 1. *Assume that the attack vector* $\mathbf{a} = \mathbf{H}\mathbf{c}$, $c_{i_1} = c_{i_2} = \cdots = c_{i_k} \neq 0$, $\Gamma = \{i_1, i_2, \ldots, i_k\}$. *The number of nonzero elements in* $\mathbf{a}$ *can be derived by*

$$g_\Gamma = k + 3\sum_{i=i_1}^{i_k}|Q_i| - \sum_{i=i_1}^{i_{k-1}} r_i - q, \tag{5}$$

where $\aleph_i$ (the buses adjacent to bus i) is divided into two parts: one is P_i which satisfies $P_i \subseteq \Gamma$ and the other is Q_i—i.e., buses, which does not belong to Γ. r_i represents the number of buses connected to bus i and $j(j \in \Gamma, j > i)$ at the same time and q represents the number of buses that $|Q_i| = 0$ in Γ.

Proof. The measurement matrix $\mathbf{H}$ is a constant *Jacobian* matrix, which depends on (1), that is, $H_{ij} = \frac{\partial z_i}{\partial x_j}$.

In the measurement vector $\mathbf{z}$, the first n elements is denoted as the active power flows at buses $1, 2, \ldots, n$, and the rest $m - n$ elements are active power flows from bus i to bus j. We know that every row sum of $\mathbf{H}$ is 0, as well as its column sum. Thus, for the buses $j \in \aleph_i$ and $k \notin \aleph_i$, we have H_{ij}, which is equal to

$$\begin{cases} \dfrac{\partial P_i}{\partial \theta_i} = \displaystyle\sum_{j\in\aleph_i} \frac{1}{X_{ij}}, \\ \dfrac{\partial P_i}{\partial \theta_j} = -\dfrac{1}{X_{ij}}, \\ \dfrac{\partial P_i}{\partial \theta_k} = 0, \end{cases} \tag{6}$$

and

$$\begin{cases} \dfrac{\partial P_{ij}}{\partial \theta_i} = \dfrac{1}{X_{ij}}, \\ \dfrac{\partial P_{ij}}{\partial \theta_j} = -\dfrac{1}{X_{ij}}, \\ \dfrac{\partial P_{ij}}{\partial \theta_k} = 0. \end{cases} \tag{7}$$

Let $\mathbf{H} = (\mathbf{h_1}, \mathbf{h_2}, \ldots, \mathbf{h_m})^{\mathbf{T}}$, where $\mathbf{h_i}$ is denoted as the ith row vector of $\mathbf{H}$. Equations (6) and (7) indicate that $\mathbf{h_i}(i = 1, 2, \ldots, n)$ has $|\aleph_i| + 1$ nonzero elements, which are H_{ii} and H_{ij}. We note that $h_i(i = n + 1, 2, \ldots, m)$ has two nonzero elements: H_{ij} and H_{ji}.

Based on [8, Theorem 1], we know that the malicious metering measurements can bypass the bad data detection if $\mathbf{a} = \mathbf{Hc}$. We define $\mathbf{c} = (\cdots c_{i_1} \cdots c_{i_2} \cdots c_{i_k} \cdots)^T$, $\Gamma = \{i_1, i_2, \ldots, i_k\}$, the attack vector $\mathbf{a}$ is equivalent to

$$\begin{bmatrix} \vdots \\ a_{i_1} \\ \vdots \\ a_{i_{g_\Gamma}} \\ \vdots \end{bmatrix} = \begin{bmatrix} H_{11} & \ldots & \ldots & H_{1n} \\ H_{21} & \ddots & & H_{2n} \\ \vdots & & \ddots & \vdots \\ H_{m1} & \ldots & \ldots & H_{mn} \end{bmatrix} \begin{bmatrix} \vdots \\ c_{i_1} \\ \vdots \\ c_{i_k} \\ \vdots \end{bmatrix}. \tag{8}$$

We can see for a power grid network, $\mathbf{H}$ is a very sparse matrix, and each of its last $(m - n)$ rows only has two opposite numbers. This is because the last $(m - n)$ rows are determined by (7). For example, if bus i connects to bus j, P_{ij} and P_{ji} are measured by z_k and z_l, respectively, and $\mathbf{a} = \mathbf{Hc}$, the ith and jth elements of both $\mathbf{h}_k$ and $\mathbf{h}_l$ will be two opposite numbers. Hence, if we set $c_i = c_j$, we have $a_k = a_l = 0$. In the same way, if bus i also connects to bus t and $c_i = c_t$, two more elements of the attack vector $\mathbf{a}$ will be 0. That is to say, for the special $\mathbf{H}$ matrix of a power grid system, if the adversary can choose the most appropriate state variable to attack, setting the nonzero elements of $\mathbf{c}$ to the same value will make a sparse attack vector $\mathbf{a}$. Consider these factors, we believe $c_{i_1} = c_{i_2} = \cdots = c_{i_k} \neq 0$ represents an effective attack strategy. Hence, in the paper, we focus on finding the least-effort attack vector with $c_{i_1} = c_{i_2} = \cdots = c_{i_k} \neq 0$.

We first divide $\aleph_i$ into two parts: one is P_i that $P_i \subseteq \Gamma$ and the other is Q_i, in which buses do not belong to Γ. When $c_{i_1} = c_{i_2} = \cdots = c_{i_k} \neq 0$, in $a_1, a_2, \ldots, a_n$, there are $k + \sum_{i=i_1}^{i_k}|Q_i| - \sum_{i=i_1}^{i_{k-1}} r_i - q$ nonzero elements, where q denotes the number of buses that $|Q_i| = 0$ in Γ. In $a_{n+1}, a_{n+2}, \ldots, a_m$, there are $2\sum_{i=i_1}^{i_k}|Q_i|$ nonzero elements.

Hence, the nonzero elements in $\mathbf{a}$ is

$$g_\Gamma = k + 3\sum_{i=i_1}^{i_k}|Q_i| - \sum_{i=i_1}^{i_{k-1}} r_i - q. \tag{9}$$

□

From the above analysis, we know that the least-effort attack is to find an optimal set of meter measurements $\Psi = \{z_{i_1}, z_{i_2}, \ldots, z_{i_{g_\Gamma}}\}$, where $g^*_\Gamma = \min_\Gamma\{g_\Gamma\}$.

The complexity of obtaining the optimal solution is shown in Theorem 2.

Theorem 2. *When $c_{i_1} = c_{i_2} = \cdots = c_{i_k} \neq 0$, the problem of finding optimal sensor measurement for the least-effort attack in the DC-state estimation is NP-hard.*

Proof. Notice that the problem for the least-effort attack is to compute a $\mathbf{c}$ with k nonzero elements so that $\mathbf{a}$ has the least number of nonzero elements for a given $\mathbf{H}$. In other words, we need to find k columns from $\mathbf{H}$, which minimizes the number of nonzero elements in $\mathbf{a}$. We can transform the above problem to the *Minimum Subadditive Join problem* [16], which is NP-hard.

Generally speaking, the Minimum Subadditive Join problem can be described as follows: let $(L, *)$ be a semilattice and let $c : L \to [0, \infty)$ be monotone and increase over L, $c(a * b) \leq c(a) + c(b)$. Given the size n subcollection X of L and an integer k with $1 \leq k \leq n$, the Minimum Subadditive Join problem is to find a size k subcollection $(x_1^{'}, x_2^{'}, \ldots, x_k^{'})$ of X that minimizes $c(x_1^{'} * x_2^{'} * \cdots * x_k^{'})$.

The transformation works in the following way: we first choose a subcollection $X = (1, 2, \ldots, n)$ of L. Then, based on Theorem 1, the problem of finding optimal sensor measurement to attack is to find a size of k subcollection $\Gamma = (i_1, i_2, \ldots, i_k)$ of X, which minimize $c(i_1 * i_2 * \cdots * i_k) = k + 3\sum_{i=i_1}^{i_k}|Q_i| - \sum_{i=i_1}^{i_{k-1}} r_i - q$. From (5), we can see if the adversary tends to manipulate more state variables, Γ will contain more elements. Hence, we conclude that $c(a * b) \leq c(a) + c(b)$.

By now, the least-effort attack problem in Theorem 1 has been transformed to the Minimum Subadditive Join problem. The Minimum Subadditive Join problem is essentially at least as hard to approximate as the well-known Maximum Balanced Complete Bipartite Subgraph problem, which is NP-hard. Hence, we can say the MSJ problem is NP-hard and details can be found in [16]. We know that the least-effort attack problem is NP-hard as well. □

From the above results, we conclude that to obtain the optimal g_Γ^* for k state variables, the adversary should search $\binom{n}{k}$ combinations of state variables. Equation (5) can be used to compute g_Γ for each fixed set of state variables. Ultimately, the global optimal one g_Γ^* should be the minimum value of all combinations.

3.2 Linear Transformation-Based Approach

The least-effort attack stated in Section 3.1 is to find an attack vector $\mathbf{a}$ that contains the least number of nonzero elements. Recall that $\mathbf{a} \in \mathrm{Im}(\mathbf{H})$—i.e., $\mathbf{a}$ is a linear combination of the column vectors of $\mathbf{H}$. We now develop a linear transformation-based approach to construct the optimal attack vector, the main idea of which is illustrated in the following Theorem.

Theorem 3. *After each column exchange of $\mathbf{H}^T$, there will be a new matrix. We can obtain the new matrix's row simplest form[1] by elementary matrix transformation. The optimal solution of the problem (denoted as $\mathbf{a}^*$) is in one of the row simplest form.*

Proof. We assume the optimal solution $\mathbf{a}^*$ exists and it has d nonzero elements. Because $\mathbf{a}^*$ is also a linear combination of the row vectors of $\mathbf{H}^T$, $\mathbf{H}^T$ could be transformed to the following form (denoted as $(\mathbf{H}^T)'$) by row transformation:

$$(\mathbf{H}^T)' = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ (\mathbf{a}^*)^T \end{bmatrix}, \tag{10}$$

where $\alpha_1, \ldots, \alpha_{n-1}$ are m-dimensional row vectors. We then make some column exchanges for $(\mathbf{H}^T)'$ to make all the nonzero elements in $(\mathbf{a}^*)^T$ in the last d columns. The transformed $(\mathbf{H}^T)'$ is

$$(\mathbf{H}^T)'' = \begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_{n-1}' \\ 0, \ldots 0, a_1^*, \ldots, a_d^* \end{bmatrix}. \tag{11}$$

We now transform the matrix composed of first $(n-1)$ vectors of $(\mathbf{H}^T)''$ to row simplest form through elementary row transformation, and then use Reduction and Reduction to Absurdity to prove that $(\mathbf{H}^T)''$ is the row simplest form. We first assume that $(\mathbf{H}^T)''$ is not row simplest form. Because if $(\mathbf{H}^T)''$ is not row simplest form while the matrix composed of first $(n-1)$ vectors of $(\mathbf{H}^T)''$ is row simplest form, there must exit a vector in $\alpha_1', \ldots, \alpha_{n-1}'$, the column subscript of the first nonzero elements of which is larger than that of $(\mathbf{a}^*)^T$. Then, $\mathbf{a}^*$ could not be the vector, which contains the least number of nonzero elements. This is contrary to our previous assumption: $(\mathbf{H}^T)''$ is not the row simplest form. Hence, $(\mathbf{H}^T)''$ is the row simplest form.

We now conclude that when we make some special column exchanges for $\mathbf{H}^T$, we obtain the row simplest form of the transformed matrix. Notice that the optimal solution $\mathbf{a}^*$ must contain the same elements with the last row of the row simplest form. Because the column subscripts of $\mathbf{a}^*$ and the last row of its corresponding simplest form are different, we can obtain the real $\mathbf{a}^*$ by tracing historical column exchanges. □

From the above analysis, we can see that when we make all possible column exchanges of $\mathbf{H}^T$, the corresponding row simplest form can be obtained. Because $\mathbf{H}$ is a $m \times n (m > n)$ matrix, we should make $n!$ times column exchanges to obtain all the row simplest forms. There will be a solution in each row simplest form. Then, one of the solutions which has the least number of nonzero elements will be chosen as the optimal solution ultimately.

3.3 A Heuristic Approach

For a small-size network, we can use a straightforward brute-force search or the above-described linear-programming approach to find the optimal set of meter measurements to compromise. Notice that our experimental results in Section 5 show that the linear transformation-based approach obtains the same results as brute-force search (when $k = 1$). For example, it requires 0.013 and 0.561 seconds for the MATLAB program to search the whole space of IEEE 9- and 14-bus systems, respectively. However, when the size of the network becomes larger, the overhead of both the brute-force and linear transformation-based approach increase exponentially. For example, for a 30-bus power network, the brute-force method needs more than one day to completely search the space and the linear transformation-based approach needs to make $29! = 8.84 \times 10^{30}$ times column exchanges to calculate the optimal solution. A detailed comparison of the computation cost of different approaches will be presented at the end of this section.

To deal with the large network, we now develop a heuristic approach. In particular, we first investigate g_Γ given a set of state variables. From (5), we can see that for a given set of buses Γ, the largest factor to g_Γ is $(3\sum_{i=i_1}^{i_k}|Q_i|)$, which is from the buses connected to Γ. Hence, for a given k, when Γ represents a set of buses connected to each other with the largest degree and connected to the least number of buses beyond Γ, the value of g_Γ will be smaller. This insight indicates that the optimal set of sensor measurements should consider buses that are geographically close to each other.

Based on the above observation, we propose our heuristic-based approach that searches the problem space in a hierarchical way. Specifically, we divide the large network into regions and then obtain g_Γ^* through multiple levels. In this way, the complexity of the algorithm can be largely reduced. Notice that to cover a more complete search space, these regions should be overlapping. Our heuristic algorithm consists of two levels. In the first level, the large network is divided into N small overlapping areas. In each area, the brute-force search method can be used to obtain a subsolution $g_{N_i}^*$. In the second level, g_Γ^* is determined by combining all subsolutions for the large network.

1. Notice that the row simplest form has the characteristics that the first nonzero element of every nonzero row is 1, and in the column it belongs to, other elements are all 0.

TABLE 1
Computation Complexity Comparison

LT	BF	LT+H	BF+H
$n!$	$o\left(2^n\right)$	$\left(\frac{n}{m}\right)!$	$o\left(2^{\frac{n}{m}}-1\right)$

To demonstrate the efficiency of our heuristic-based approach (denoted as BF+H, where BF presents brute force for region and H presents the hierarchical regions), we compare its computational complexity (BF+H) with that of pure linear transformation algorithm(LT), brute-force search method, and hierarchical linear transformation algorithm (LT+H). All results are shown in Table 1. We assume that vector $\mathbf{c}$ is n-dimensional. Because the pure linear transformation algorithm needs to conduct $n!$ times matrix transformation to obtain an optimal solution, its computational complexity is $n!$. The method in [8] essentially belongs to the random search algorithm. Notice that only if we traverse all possibilities, the optimal solution can be derived. The complexity of the algorithm is $o(2^n)$. Recall that, they first randomly conducted 1,000 times experiments by choosing $k(1 \leq k \leq 10)$ state variables and generating malicious data for each of them. Then, the solution of the problem is selected based on 1,000 times experiments. In a small-size network, the method may achieve a good result in a short period of time. However, for a large-size network, such as IEEE 300-bus system, when $k = 3$ the algorithm could only go through 0.66 percent of all the possibilities. It is obvious that when the size of the network becomes larger, the complexity of both the pure linear transformation algorithm and the brute-force method will become too large to obtain the solution efficiently. Differently, our proposed heuristic-based approach divides the state variable vector $\mathbf{c}$ into several regions, adopts the brute-force search method to traverse all the possibilities and calculates the suboptimal solution of each region independently. In this way, we can achieve a local optimum solution quickly. Suppose that vector $\mathbf{c}$ is n-dimensional, and partitioned into N parts, then the computational complexity of our algorithm is $o(2^{\frac{n}{N}}-1)$. In the same way, if we use the linear transformation algorithm in each region instead, the complexity will be $(\frac{n}{N})!$. In summary, we conclude our heuristic-based algorithm derives solution with a low computation complexity.

4 Countermeasures

In this section, we develop countermeasures against the false data-injection attacks. We design the defensive countermeasures from the following perspectives: 1) protection-based defense, and 2) anomaly detection-based defense.

4.1 Protection

Because a power grid usually covers a large geographical area—for example, >1 million kilometer^2, it is practical for the system operators to choose some "important meters" from all the meters to protect, for example, by a combination of encryption, continuous monitoring, separation from the Internet [17]. The meters in power system DC state estimation belong to two kinds: bus real power injections and line real power flows. As mentioned above, there is 1 meter on each bus measuring its real power injection and 2 meters on each transmission line measuring the line real power flows.

We now illustrate our proposed protection strategy using an example based on IEEE 30-bus system shown in Fig. 3. From this figure, bus 9 is connected to buses 6,10, and 11. We know that if the adversary wants to change the state variable of bus 9, he/she needs to compromise the meters of power injections P_9, P_6, P_{10}, and P_{11}, and power flows P_{96}, P_{69}, $P_{9,10}$, $P_{10,9}$, $P_{9,11}$, and $P_{11,9}$. On the other hand, if the system operator protects the meter of bus power injection P_9, the adversary will have no chance to manipulate the state variables of buses 9, 6, 10, and 11. If the meter on the transmission line (e.g., P_{96}) is protected, only the state variables of buses 9 and 6 cannot be manipulated.

From the above example, we can see the meters that measure bus injection powers play a more important role in meter protection than the meters that measure the transmission line power flows. We can explain this from the essential of state estimation, the measurements of bus power injections play a critical role in determining a specific state variable, while the measurements of line power flows are redundant to improve the accuracy of state estimation. If the injection power flow of bus i is protected, the search space for the adversary will be decreased from $\binom{n}{k}$ to $\binom{n-|\aleph_i|-1}{k}$. Hence, protecting meters that measure the injection power flows of buses connected to the largest number of buses will increase the difficulty of conducting false data-injection attacks when it impractical to protect all buses in large power grid systems.

4.2 Spatial-Based Detection

Recall that in stealthy attacks, the adversary may change measurements from multiple sensors marginally, so that individual compromised measurement will not be detected by the statistical anomaly detection discussed above. To detect such stealthy attacks, we introduce a novel spatial-based detection scheme based on hypothesis testing.

Once the measurement vector has been collected in the control center, a decision should be made on whether the measurement vector is associated with an attack. Based on the hypothesis test [18], we use the distribution of measurements as the detection feature. To detect false data-injection attacks, we consider two hypotheses: H_0 and H_1. H_0 is the null hypothesis, where the measurement is valid; and H_1 is the alternative hypothesis, where the measurement is under attack. With the nonzero elements in the attack vector $\mathbf{a}$, the two hypothesis can be described by

$$\begin{aligned} H_0 &: \|\mathbf{a}\|_0 = 0, \\ H_1 &: \|\mathbf{a}\|_0 > 0. \end{aligned} \tag{12}$$

As we can see in (3), the measurement residual in state estimation can be divided into two parts: $(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}})$ and $(\mathbf{a} - \mathbf{H}\mathbf{c})$, which are derived from the measurement vector and the attack vector. Recall that the false data-injection attacks cannot be detected by LNR test because the false measurement makes $\mathbf{a} - \mathbf{H}\mathbf{c}$ to be zero when the adversary knows the system topology (i.e., $\mathbf{H}$) and bad detection algorithm.

We assume a Bayesian model, where the measurements are random with a multivariate Gaussian distribution,

which can be estimated based on historical data. Notice that Gaussian distribution has been widely used to describe quantitative phenomena in the natural and behavioral sciences. The use of the Gaussian distribution can be theoretically justified by assuming that many small independent effects are additively contributing to each observation. However, when false data-injection attacks exist, it must change some specific measurements marginally and the combination of those measurements will lead to the state variables derived far from their true values. We would like to point out that one compromised measurement's deviation from its mean value ($|z_i - Z_i|$) may be too small and recognized as the random noise. However, when we accumulate the deviations of all the measurements in a vector from their means, the accumulated deviation ($\sum_{i=1}^{m} \frac{(z_i - Z_i)^2}{\Sigma_i^2}$) can be stood out. Our detection scheme is based on this insight to recognize attacks.

We assume that the measurement vector $\mathbf{z}$ follows multivariate Gaussian distributed and z_i $(i = 1, 2, \ldots, m)$ is independent of each other, denoted as $\mathbf{z} \sim N_m(\mathbf{Z}, \mathbf{\Sigma})$. $\mathbf{Z}$ is the mean vector and $\mathbf{\Sigma}$ is a diagonal covariance matrix. Let $J(\mathbf{z}) = (\mathbf{z} - \mathbf{Z})^T \mathbf{\Sigma^{-1}} (\mathbf{z} - \mathbf{Z})$, we have

$$J(\mathbf{z}) \sim \chi^2(m). \tag{13}$$

Based on this, the hypothesis test is given by

$$J(\mathbf{z}) \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{14}$$

where $\tau = \chi_\alpha^2(m)$ is a threshold determined by considering the null hypothesis at a certain false positive rate α.

To evaluate the effectiveness of the spatial-based detection based on hypothesis test, we choose the detection rate (same as the true positive) and false positive rate as the evaluation metrics. In particular, we define detection rate P_D as the probability that the false data-injection attacks is correctly recognized. False positive rate, P_F is the probability that a normal measurement vector is misclassified as attacks. The more effective detection algorithms are the ones that can achieve higher detection rates and lower false positive rates. In our performance evaluation in Section 5, we will use *receiver operating characteristic* (ROC) curve to show the relationship between P_D and P_F.

4.3 Temporal-Based Detection

To detect attacks marginally manipulating data over time, we adopt the temporal-based detection using online nonparametric cusum change detection mechanism. In particular, given the two hypotheses in (12), the cusum algorithm assumes the observation $y(i)$ starts under H_0 (with a probability distribution p_0) and at a time k_s it changes to hypothesis H_1 (with a probability distribution p_1). The goal of such a detection scheme is to recognize this change within the shortest time [19]. Given a false positive rate, the temporal-based detection tends to minimize the time N ($N \geq k_s$) of detection. The classic cusum algorithm is based on computing the following statistics:

$$S(k) = max\left\{0, \log\frac{p_1(y(k))}{p_0(y(k))} + S(k-1)\right\}, S(0) = 0. \tag{15}$$

Hence, the detection time can be derived by $N = \inf_n\{n : S(n) \geq \tau\}$, where τ is a threshold selected based on the false positive rate.

Considering the case where the probability distributions of measurements and the attack are not known, we use nonparametric statistics. Let $z_i(k)$ be the measurement of the ith meter at the time k. We define the observation $y_i(i)$ as following, $y_i(k) = |z_i(k) - Q_i(\gamma)|$, where $Q_i(\gamma)$ is the γ-percentiles $\gamma \in (0, 1)$ derived from the historical data. The nonparametric cusum statistics for the ith measurement becomes $S_i(k) = max\{0, S_i(k-1) + |z_i(k) - Q_i(\gamma)|\}$, $S_i(0) = 0$.

Then, a decision rule can be made by

$$S_i(k) \underset{H_0}{\overset{H_1}{\gtrless}} \tau_i, \tag{16}$$

where τ_i is a threshold selected based on the false positive rate for the ith measurement. Notice the selection of threshold is a tradeoff between the detection time and the false positives.

To evaluate the effectiveness of the spatial-based detection-based online nonparametric cusum change, we consider two metrics: false positive rate and detection time. False positive rate is defined as the ratio, at which the actual value is H_0 but it is determined to be H_1. Detection time is the average time that it takes to detect the change or attack. As we can see, the most useful temporal detection scheme is the one that has low values of both metrics. In Section 5, we will evaluate the effectiveness of temporal-based detection on these two metrics.

5 Performance Evaluation

In this section, we conduct experiments to investigate the effectiveness of the least-effort attack and corresponding countermeasures.

Simulation setup. To validate the least-effort attacks and countermeasures studied in Sections 3 and 4, we simulate our approaches using MATLAB 7.9.0. All the parameters used in the test cases, including the real value of state variables, sensor measurements, and the measurement matrix $\mathbf{H}$, are based on the MATLAB package MATPOWER [20]. We first evaluate the least-effort attacks investigated in Section 3, we use the brute-force search and linear transformation-based approach to inject attacks into small networks (i.e., the IEEE 9-bus and 14-bus systems), and the hierarchical search approach into large networks (i.e., IEEE 30-bus, 118-bus, and 300-bus systems), respectively. We then evaluate countermeasures investigated in Section 4 based on the IEEE 14-bus network.

Least-effort false data-injection attacks. For small-size networks, the optimal set of meters to compromise can be obtained by using a straightforward brute-force or linear transformation-based approach, while for large-size networks, it is obtained by our proposed heuristic-based approach. We first simulate the least-effort attacks using brute-force approach on small networks, including IEEE 9-bus and 14-bus systems. In our experiments, we set the number of state variables to be manipulated as

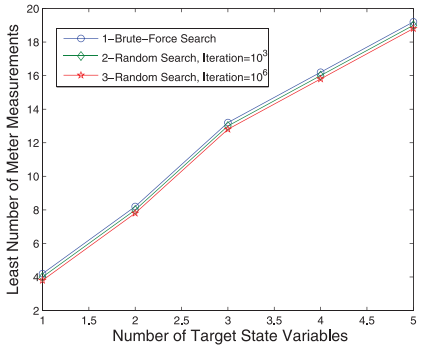


Fig. 1. Brute-force search for IEEE 9-bus.

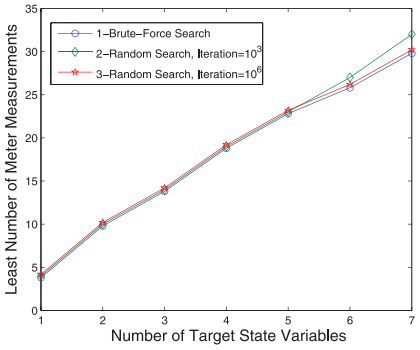


Fig. 2. Brute-force search for IEEE 14-bus.

$k = 1, 2, \ldots, 5$ for the IEEE 9-bus system ($n = 8$) and $k = 1, 2, \ldots, 7$ for the IEEE 14-bus system ($n = 13$), respectively. We set the false data $c_i (i = i_1, i_2, \ldots, i_k)$ as 0.5 (10 percent of the true state variable value in the IEEE 14-bus system). We use (5) and the brute-force approach to compute the number of nonzero elements of all the $\binom{n}{k}$ attack vector $\mathbf{a}$, and examine the combination of error vector $\mathbf{c}$, which requires the least effort. It takes only 0.22 and 0.55 seconds for the brute-force approach to find the least number of meters to compromise for the two systems, respectively.

We compare the brute-force approach with the random search approach in [8]. To be specific, we randomly choose k state variables and inject false data into them, by conducting such experiments 10^3 times like [8]. To thoroughly test the efficiency of the random search approach, we conduct the same experiments for 10^6 times. The results of the least number of meter measurements to be compromised for each scenario are demonstrated in Figs. 1 and 2.

For the IEEE 9-bus system, we can see that the brute-force approach and the random search approach achieve almost the same results because of the small search space for such a small-scale system. Nonetheless, for the larger IEEE 14-bus system, the results of random search approach with 10^3 rounds are larger than those of the brute-force approach when $k = 6$ and 7, while the results of random search approach with 10^6 rounds achieve the similar results as the brute-force one. This is because 10^3 times random experiments cannot cover the whole space while $\binom{13}{6} = 1{,}716$. Hence, we can conclude that for a small network, the random search approach is almost as efficient as the brute-force search method.

We simulate the least-effort attacks based on the linear transformation-based approach on the IEEE 9-bus system. We make $8! = 40{,}320$ times column exchanges for $\mathbf{H}^T$, and obtain the row simplest form for each changed matrix. We calculate the number of nonzero elements in the last row of every row simplest form. The optimal solution $\mathbf{a}^*$ is the one which has the minimum number of nonzero elements. We show the simplest form that contains the optimal solution in Fig. 4. We can see the last row of this row simplest form contains four nonzero elements, which are in the columns of 13, 22, 26, and 27. By tracing the way we make column exchange, we find this row simplest form belongs to a matrix which is obtained by exchanging the third column with the 27th column, and the sixth column with the 26th column of $\mathbf{H}^T$. So the nonzero elements in the real $\mathbf{a}^*$ is a_3^*, a_6^*, a_{13}^*, and a_{22}^*. That is to say, the least-effort attack needs to compromise four measurements (i.e., z_3, z_6, z_{13}, and z_{22}). As we can see, the linear transformation-based approach obtains the optimal solution, which is the same as the brute-force search (when $k = 1$).

We also simulate the false data-injection attacks against large networks, including the IEEE 30-bus, 118-bus, and 300-bus systems, using the heuristic-based approach. In our experiments, we set the number of targeted state variables $k = 1, 2, \ldots, 15$. For example, we divided the IEEE 30-bus systems into $m = 3$ areas and each area contains $n_i (i = 1, 2, 3)$ buses, as shown in Fig. 3. In each area, we inject the false data into all the $\binom{n_i}{k}$ possible state vectors, and used (5) to examine the attack vector $\mathbf{a}$, which contains the minimum number of nonzero elements. We compare the suboptimal sets of each area and obtain the optimal set of meters to compromise. As we can see in Fig. 5, to manipulate four state variables, we need to compromise at least 10-meter measurements.

We also compare the random search approach with the hierarchical-based search approach. Fig. 5 shows that when k is larger than 8, the hierarchical-based approach always works better than the other two approaches. From the perspective of probability theory, when the network becomes large, the search space and search time become larger, leading to the random search approach inefficient.

We conducted the same experiments on networks with the IEEE 118-bus and 300-bus. In particular, we divide the IEEE 118-bus system into $N = 12$, and the IEEE 300-bus system into $N = 30$ areas, respectively. Figs. 6 and 7 show the results of the two tested networks. As we can see, the

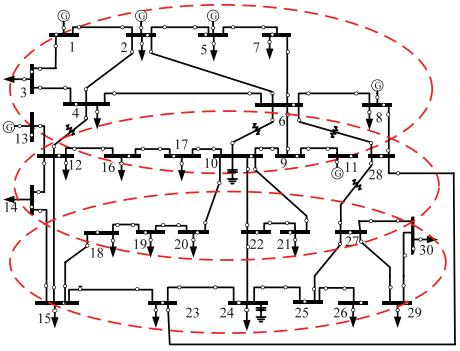


Fig. 3. Example of IEEE 30-bus with measurements.

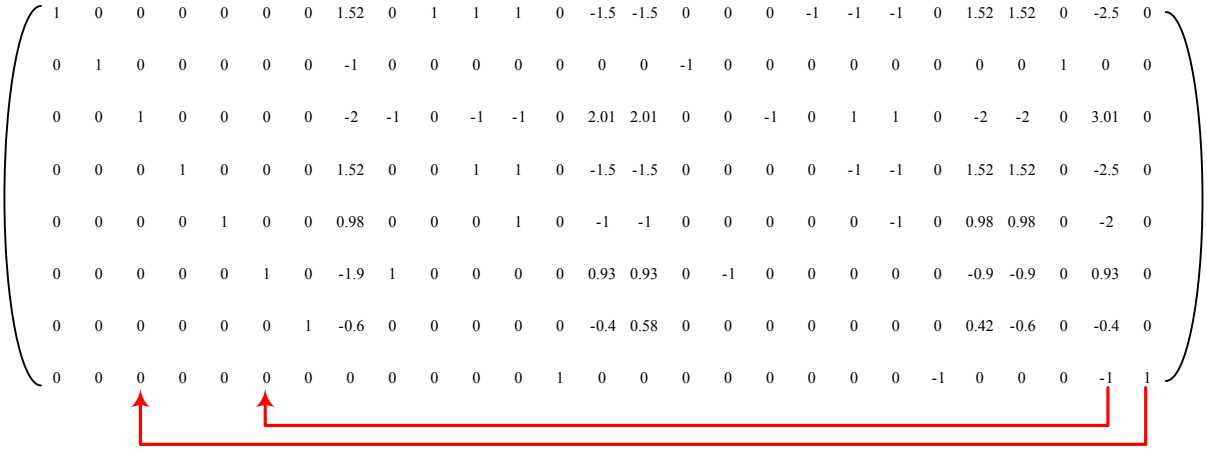


Fig. 4. Linear transformation approach for IEEE 9-bus.

hierarchical-based approach can always find more accurate results than the random search approach quickly, especially when the size of the system is large. Notice that in Figs. 5, 6, and 7, there are bumpy curves for two reasons. First, the results of random search rely heavily on the number of times the search is conducted. When the number of searches is not sufficiently large, the random search process only covers part of the search space. This makes the results look random and bumpy. Second, in the heuristic approach, we first divide the network into a number of overlapping areas, use the brute-force search or linear programming to identify the optimal set of meters for individual small areas, and then derive the set of meters for the whole network. Although the heuristic-based approach could significantly reduce the complexity of the algorithm and yield much better results than the random search approach, it could not cover the whole search space and may, therefore, miss certain optimal solutions, leading to the curves of results occasionally having bumpy points.

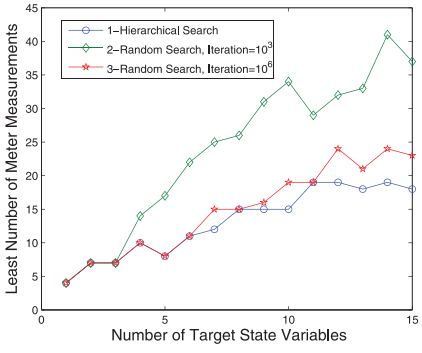


Fig. 5. Hierarchical search for IEEE 30-bus.

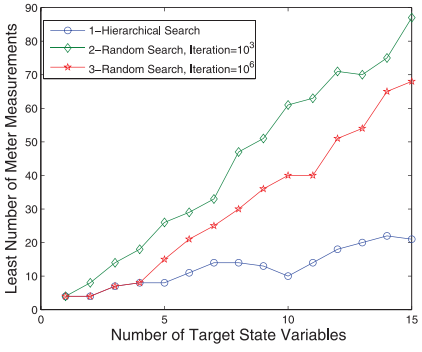


Fig. 6. Hierarchical search for IEEE 118-bus.

We now show the overhead of the computation complexity of our hierarchical-based approach. Table 2 shows the complexity of brute-force search (BF), the complexity of heuristic-based search (BF+H), the random search time (RS) given the iteration of 10^6, and the time for BF+H for the IEEE 30-bus, 118-bus, and 300-bus, respectively. As we can see, our hierarchical-based approach can not only find more accurate attacking vector but also compute in shorter time than the random search approach with low computation complexity.

Protecting critical meters. Fig. 8 shows the topology of IEEE 14-bus system. In this figure, we choose to protect bus 4 because it connects to the largest number of buses. As we can see, bus 4 is connected to five other buses, 2, 3, 5, 7, and 9. We know that if the meter measuring P_4 is unknown to the adversary or cannot be compromised, he cannot change $\theta_4, \theta_2, \theta_3, \theta_5, \theta_7$, and θ_9. This is because if he does inject false data into one of the above state variables, he must compromise P_4 at the same time. Obviously, this leads to

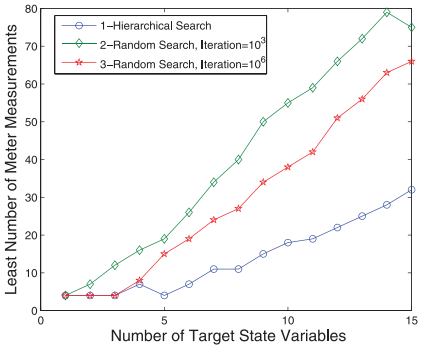


Fig. 7. Hierarchical search for IEEE 300-bus.

TABLE 2
Attack Overhead

System	BF	BF+H	RS(s)	BF+H(s)
30-bus	1.07×10^9	3.07×10^3	437.22	24.43
118-bus	3.32×10^{35}	4.92×10^5	830.90	156.13
300-bus	2.04×10^{90}	9.83×10^5	6940.31	787.66

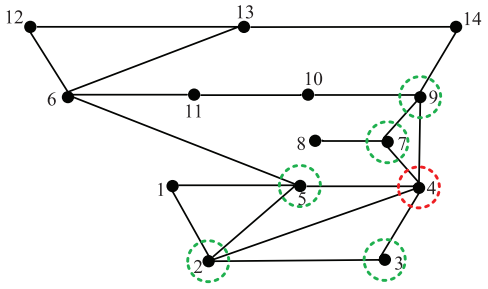


Fig. 8. Topology of IEEE 14-bus system.

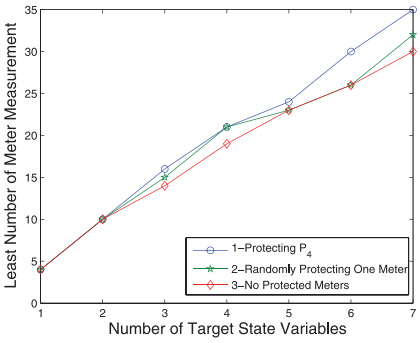


Fig. 9. Results of protection-based defense.

a smaller search space (e.g., higher cost) to conduct the least-effort attack than the scenario when the complete topology information is known.

Fig. 9 shows the least number of meters to compromise when P_4 is protected and cannot be compromised. We randomly choose one sensor measurement to be protected and compute the least number to compromise. We repeated the experiments 100 times and obtained the mean value to compare with the scenario, where P_4 is protected. As we can see, if the adversary wants to manipulate $k(k > 2)$ state variables, he will have to compromise more meters when the important measurements are protected than the scenarios, where the meters are randomly protected.

If we make a set of active power flows of buses whose adjacent buses can cover all the buses in the power system unknown to the adversary, the search space of meters to compromise becomes 0, so it is impossible for the adversary to deploy false data-injection attacks.

Results of spatial-based detection. Notice that the SCADA system collects the real-time information of power field and reports the data to the control center every 2 to 4 seconds [21]. In our evaluation, we run simulations for 1 hour without attacks to collect enough samples and derive the mean values and covariance of all measurements in the IEEE 14-bus system. We then set the detection threshold τ of chi-square distribution by choosing the false positive rate $\alpha = 0.1$. We exhaustively search all the bad measurement vectors based on the attack model $\mathbf{a} = \mathbf{Hc}$. We use the detection rule defined in (14) to calculate the number of times that the $J(\mathbf{z}) > \tau$, and derive the detection rate by $P_D = \frac{\#J(\mathbf{z})>\tau}{\#all\ trails}$. We collect 1,000 normal measurement vectors, which only contain random noise, use the detection rule defined in (14) to compute the total number of times when the $J(\mathbf{z}) > \tau$ and derive the false positive rate by $P_F = \frac{\#J(\mathbf{z})>\tau}{1{,}000}$. Based on the (P_F, P_D) given a threshold, we draw the ROC, which represents the detection rate versus false positive rate.

Fig. 10 shows the ROC curves for the IEEE 14-bus system in terms of different magnitudes, which determine the strength of data to be manipulated. Similar to the signal-to-noise ratio (SNR), we define signal-to-attack ratio (*SAR*) to quantify the attack's "power" or strength, that is, $SAR = 10\log_{10}\frac{x_i}{c_i}$. We find that when we increase the threshold, both the detection rate and the false positive rate increase. Hence, the selection of the threshold τ is a tradeoff between the detection rate and false positive rate.

We can see that the efficiency of the detection algorithm is improved compared with the case when $SAR = 17$ dB and the detection rate approaches 88 percent. While $SAR = 11$ dB, the detection algorithm can recognize 98.8-98.9 percent attacks. As we can see, the detection rate becomes higher when the attack's power/strength rises. This is because the deviation of the observed measurements from their means is more significant when the attack becomes stronger.

Results of temporal-based detection. We implemented the proposed temporal-based detection algorithm on the IEEE 14-bus system. In the experiments, we run simulations for 1 hour without attacks and collected the meters of z_3, z_4, z_{20}, which measure the power injections of P_3, P_4, and power flow of P_{34}, respectively. We set $\gamma = 0.9$ and then obtain the $\gamma-$percentiles of P_3, P_4, and P_{34} as $Q_3(\gamma) = 94.19$, $Q_4(\gamma) = 47.81$, $Q_{20}(\gamma) = -24.24$.

We run simulations for 1,000 times without attacks and compute the total number of false positives for different values of τ. The false positive rate, P_F can be defined as $P_F = \frac{\#false\ alarms}{1{,}000}$. Fig. 11 shows the results for z_3, z_4, z_{20}. We can see for z_{20}, the false positive rate will become very low when we set $\tau > 50$.

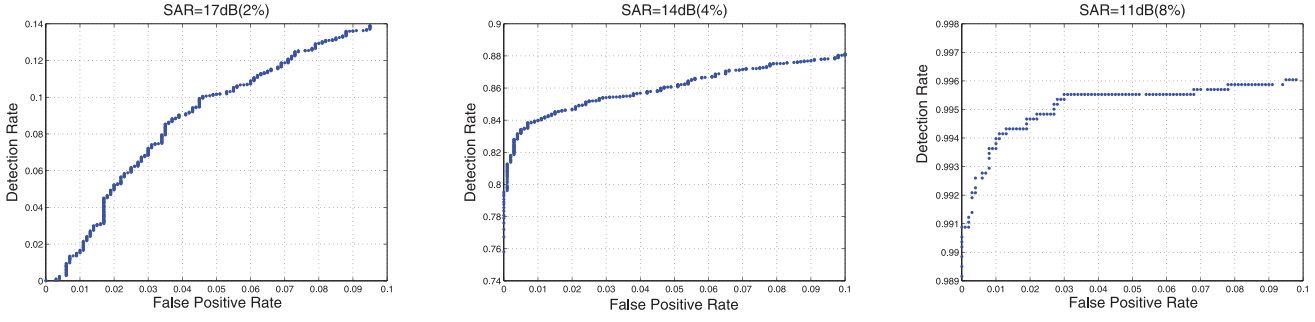


Fig. 10. Detection rate versus false positive rate.

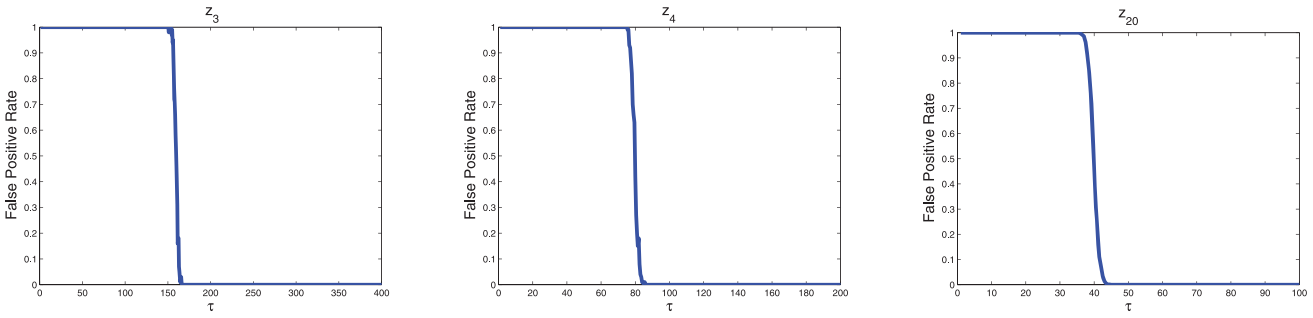


Fig. 11. False positive rate versus τ for z_3, z_4, and z_{20}.

With the attack model $\mathbf{a} = \mathbf{Hc}$, Fig. 12 shows the average detection time by conducting 1,000 times experiments based on our proposed *nonparametric cusum* algorithm in terms of the thresholds: z_3, z_4, and z_{20}, respectively. As we can see, the detection time increases linearly as the threshold τ increases. This is because τ is selected based on the false positive rate, which is a tradeoff between the detection time and the false positives. We can also observe that the more measurement's value is changed by the attack, the earlier the false data-injection attacks will be detected by our proposed detection algorithm. From Fig. 11, we can see selecting τ as high as possible for each sensor can reduce the false positives. Nevertheless, increasing τ will require more time to detect attacks.

6 Related Work

In this section, we briefly review the related work and illustrate the main novelties of our proposed research. After [8], a number of efforts have been paid to false data-injection attacks against power system state estimation [13], [17], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32]. For example, based on the work of [8], Liu et al. further studied the false data-injection attacks and proposed more generalized attack models in [13]. To defend against the false data-injection attacks, Bobba et al. [25] developed schemes to protect a selected set of meter measurements. Kosut et al. [21] developed a scheme to detect the attack. Pasqualetti et al. [28] proposed a distributed state estimation method with a group of connected control centers and developed a distributed algorithm to detect false data injection attacks.

The study of false data-injection attacks has been extended [17], [19], [30], [31], [33], [34], [35], [36], [37]. For example, Mo and Sinopoli [30] studied false data injection attacks and defensive method in the models of using Kalman Filter and LQG controller. Lin et al. [35] investigated the vulnerability of distributed energy distribution and investigate novel false data injection attacks against the energy distribution. Mohsenian-Rad and Leon-Garcia [36] studied attacks against the consumption sector, by investigating load altering attacks and proposed a cost-efficient load protection strategy. Kim and Poor [17] developed a unified framework for constructing attack vectors, considering the constraints on measurements and limited sources associated with the adversary. Based on this formulation, they proposed a defensive mechanism of choosing a subset of measurements immune to the attacks based on a greedy algorithm. Their work is similar to our study of giving the attack partial knowledge of $\mathbf{H}$ matrix. Nevertheless, our method is more general based on the graph theory. Hossain et al. [37] studied the state-of-the-art theory, key strategies, protocols, applications, deployment aspects, and experimental studies of communication and networking technologies for the smart grid.

Different from the existing research efforts, to the best of our knowledge, our research is among the first of efforts toward finding the minimal number of meter measurements to compromise in a systematic way, developing efficient algorithms to obtain the vector for attacks, and developing novel protection and detection mechanisms to effectively reduce the risk by false data injection attacks and detect attacks accurately and quickly.

7 Conclusion

In this paper, we studied the false data-injection attacks and corresponding countermeasures. In particular, we first presented the least-effort attack model, which enables an adversary to find the optimal set of meter measurements

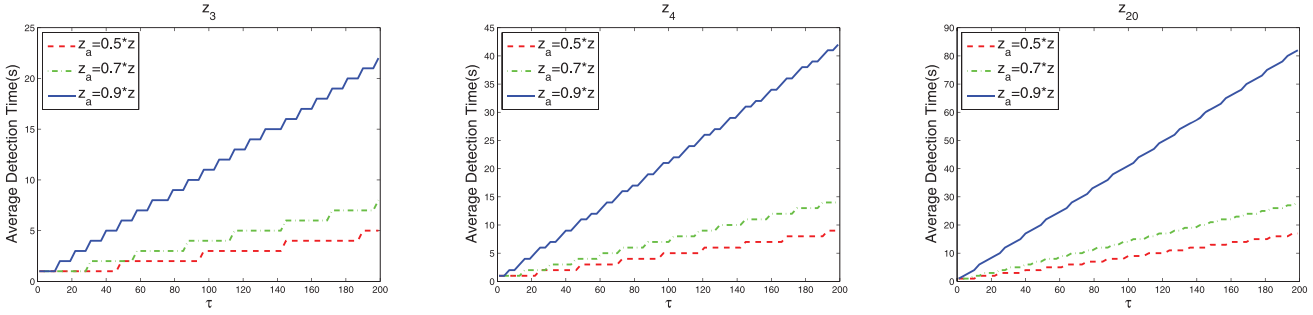


Fig. 12. Detection time versus τ for z_3, z_4, and z_{20}.

given the number of state variables to manipulate in a power grid system. We formalized the problem and developed efficient algorithms to find the set of meters. We implemented and tested our attacking approach on various IEEE standard bus systems, and demonstrated their superiority over the random search approach. To defend against false data-injection attacks, we investigated the protection-based defense and detection-based defense. For the protection-based defense, we identified and protected critical sensor and make the system more resilient to attacks. For the detection-based defense, we developed spatial-based and temporal-based detection schemes. Our experimental data show that our spatial-based detection can detect at least 95 percent of attacks when the adversary changes up to 6 percent of the magnitude values of state variables and the temporal-based algorithm can identify compromised meters quickly given the historical data from those meters.

ACKNOWLEDGMENTS

The work was supported in part by the following funding agencies in China: National 973 Basic Research Program of China under grant No. 2011CB302801, the Fundamental Research Funds for the Central Universities (xjj2011078), and Xi'an industrial applied technology research project (CXY1017(4)). This research was also supported in part by the US National Science Foundation under grants 0852673, 0852674, 0915834, 1117175, and by the George Washington University under a Research Enhancement Fund. Any opinions, findings, conclusions, and/or recommendations expressed in this material, either expressed or implied are those of the authors and do not necessarily reflect the views of the sponsor listed above. Wei Yu is the corresponding author of this paper.

REFERENCES

[1] T. Morris, A.K. Srivastava, B. Reaves, K. Pavurapu, S. Abdelwahed, R. Vaughn, W. McGrew, and Y. Dandass, "Engineering Future Cyber-Physical Energy Systems: Challenges, Research Needs, and Roadmap," *Proc. North Am. Power Symp. (NAPS),* Oct. 2009.

[2] U.D. of Energy Smart Grid System Report, http://energy.gov/oe/technology-development/smart-grid, 2009.

[3] A. Albur and A.G. Exposito, *Power System State Estimation: Theory and Implementation.* CRC Press, 2004.

[4] A. Monticelli, F.F. Wu, and M.Y. Multiple, "Multiple Bad Data Identification for State Estimation by Combinatorial Optimization," *IEEE Trans. Power Delivery,* vol. PD-1, no. 3, pp. 361-369, July 1986.

[5] M.M.G.P. Granelli, "Identification of Interacting Bad Data in the Framework of the Weighted Least Square Method," *Electric Power System Research,* vol. 78, no. 5, pp. 806-814, May 2008.

[6] J. Khwanramand and P. Damrongkulkamjorn, "Multiple Bad Data Identification in Power System State Estimation Using Particle Swarm Optimization," *Proc. Sixth Int'l Conf. Electrical Eng./Electronics, Computer, Telecomm. and Information Technology,* pp. 3-6, May 2009.

[7] A.A. Cardenas, S. Amin, and S. Sastry, "Secure Control: Towards Survivable Cyber-Physical Systems," *Proc. the First Int'l Workshop Cyber-Physical Systems,* pp. 495-500, June 2008.

[8] Y. Liu, M.K. Reiter, and P. Ning, "False Data Injection Attacks against State Estimation in Electric Power Grids," *Proc. the 16th ACM Conf. Computer and Comm. Security,* Nov. 2009.

[9] NIST, *Guidelines for Smart Grid Cyber Security,* http://csrc.nist.gov/publications/PubsNISTIRs.html#NIST-IR-7628, 2010.

[10] A.A. Cardenas, S. Amin, and S. Sastry, "Research Challenges for the Security of Control Systems," *Proc. Third USENIX Workshop Hot Topics in Security (HotSec),* July 2008.

[11] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S.S. Sastry, "Challenges for Securing Cyber Physical Systems," *Proc. Workshop Future Directions in Cyber-physical Systems Security,* July 2009.

[12] J. Vijayan, "Stuxnet Renews Power Grid Security Concerns," http://www.computerworld.com/s/article/9179689/Stuxnet_renews_power_grid_security_concerns, July 2010.

[13] Y. Liu, P. Ning, and M.K. Reiter, "Generalized False Data Injection Attacks against State Estimation in Electric Power Grids," *ACM Trans. Information and System Security,* vol. 14, no. 1, pp. 13:1-13:32, May 2011.

[14] Q. Yang, J. Yang, W. Yu, N. Zhang, and W. Zhao, "On a Hierarchical False Data Injection Attack on Power System State Estimation," *Proc. IEEE Globecom,* Dec. 2011.

[15] F.C. Schweppe, J. Wildes, and D.B. Rom, "Power System Static State Estimation. Parts 1, 2, 3," *IEEE Trans. Power Apparatus and Systems,* vol. PAS-89, no. 1, pp. 120-135, Jan. 1970.

[16] S.A. Vinterbo, "A Stab at Approximating Minimum Subadditive Join," *Proc. 10th Int'l Conf. Algorithms and Data Structures (WADS '07),* pp. 214-225, 2007.

[17] T. Kim and H. Poor, "Strategic Protection against Data Injection Attacks on Power Grids," *IEEE Trans. Smart Grid,* vol. 2, no. 2, pp. 326-333, June 2011.

[18] P.J. Bickel, *Mathematical Statistics.* Holden Day, 1991.

[19] A.A. Cardenas, S. Amin, Z.S. Lin, Y.L. Huang, C.Y. Huang, and S. Sastry, "Attacks against Process Control Systems: Risk Assessment, Detection, and Response," *Proc. ACM Symp. Information, Computer and Comm. Security (AsiaCCS '11),* Mar. 2011.

[20] R.D. Zimmerman, C.E. Murillo-Sanchez, and D. Gan, "Matpower, a Matlab Power System Simulation Package," http://www.pserc.cornell.edu/matpower/manul.pdf, 2007.

[21] O. Kosut, L. Jia, R.J. Thomas, and L. Tong, "On Malicious Data Attacks on Power System State Estimation," *Proc. 45th Int'l Univ. Power Eng. Conf. (UPEC '10),* Aug. 2010.

[22] O. Kosut, L. Jia, R.J. Thomas, and L. Tong, "Malicious Data Attacks on the Smart Grid," *IEEE Trans. Smart Grid,* vol. 2, no. 4, pp. 645-658, Dec. 2011.

[23] H. Sandberg, A. Teixeira, and K.H. Johansson, "On Security Indices for State Estimators in Power Networks," *Proc. Preprints of the First Workshop Secure Control Systems (CPSWEEK '10),* 2010.

[24] G. Dan and H. Sandberg, "Stealth Attacks and Protection Schemes for State Estimators in Power Systems," *Proc. First IEEE Int'l Conf. Smart Grid Comm.,* Oct. 2010.

[25] R.B. Bobba, K.M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T.J. Overbye, "Detecting False Data Injection Attacks on DC State Estimation," *Proc. Preprints of the First Workshop Secure Control Systems (CPSWEEK '10),* 2010.

[26] O. Kosut, L. Jia, R.J. Thomas, and L. Tong, "Limiting False Data Attacks on Power System State Estimation," *Proc. Conf. Information Sciences and Systems,* Mar. 2010.

[27] O. Kosut, L. Jia, R.J. Thomas, and L. Tong, "Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures," *Proc. First IEEE Int'l Conf. Smart Grid Comm.,* Oct. 2010.

[28] F. Pasqualetti, R. Carli, and F. Bullo, "A Distributed Method for State Estimation and False Data Detection in Power Networks," *Proc. IEEE Int'l Conf. Smart Grid Comm. (SmartGridComm),* Oct. 2011.

[29] A. Teixeira, G. Dan, H. Sandberg, and K.H. Johansson, "A Cyber Security Study of a SCADA Energy Management System: Stealthy Deception Attacks on the State Estimator," *Proc. 18th IFAC World Congress,* 2011.

[30] Y.L. Mo and B. Sinopoli, "False Data Injection Attacks in Control Systems," *Proc. Preprints of the first Workshop Secure Control Systems,* 2010.

[31] L. Xie, Y.L. Mo, and B. Sinopoli, "False Data Injection Attacks in Electricity Markets," *Proc. First IEEE Int'l Conf. Smart Grid Comm.,* Oct. 2010.

[32] S. Cui, Z. Han, S. Kar, T.T. Kim, H.V. Poor, and A. Tajer, "Coordinated Data-Injection Attack and Detection in the Smart Grid: A Detailed Look at Enriching Detection Solutions," *IEEE Signal Processing Magazine,* vol. 29, no. 5, pp. 106-115, Sept. 2012.

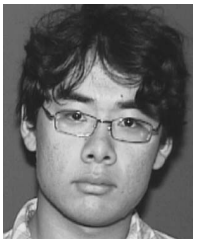
[33] A. Teixeira, S. Amin, H. Sandberg, K.H. Johansson, and S.S. Sastry, "Cyber Security Analysis of State Estimators in Electric Power Systems," *Proc. 49th IEEE Conf. Decision and Control,* Dec. 2010.
[34] A. Tajer, S. Kar, H.V. Poor, and S. Cui, "Distributed Joint Cyber Attack Detection and State Recovery in Smart Grids," *Proc. IEEE Int'l Conf. Smart Grid Comm. (SmartGridComm),* Oct. 2011.
[35] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao, "On False Data Injection Attacks against Distributed Energy Routing in Smart Grid," *Proc. ACM/IEEE Third Int'l Conf. Cyber-Physical Systems (ICCPS),* Apr. 2012.
[36] A.H. Mohsenian-Rad and A. Leon-Garcia, "Distributed Internet-Based Load Altering Attacks against Smart Power Grids," *IEEE Trans. Smart Grid,* vol. 2, no. 4, pp. 667-674, Dec. 2011.
[37] E. Hossain, Z. Han, and V. Poor, *Smart Grid Communications and Networking*. Cambridge Univ. Press, 2012.

Qingyu Yang received the BS and MS degrees both in mechatronics engineering from Xi'an Jiaotong University, China, in 1996 and 1999, respectively, the PhD degree in control science and engineering from Xi'an Jiaotong University, China, in 2003. He is an associate professor in School of Electronics & Information Engineering at Xi'an Jiaotong University. He is also with the State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University. His current research interests include cyber-physical systems, power grid security, control and diagnosis of power system, and intelligent control of industrial process. He is a member of the IEEE.

Jie Yang received the BS degree in detection guidance and control technology from Northwestern Polytechnical University, China, in 2009, the MS degree in control engineering and control theory from Xi'an Jiaotong University, in 2012. She is also with the State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University. Her research interests include cyber-physical systems and power grid security.

Wei Yu received the BS degree in electrical engineering from Nanjing University of Technology, Nanjing, China, in 1992, the MS degree in electrical engineering from Tongji University, Shanghai, China, in 1995, and the PhD degree in computer engineering from Texas A&M University, College Station, in 2008. He is currently an assistant professor with the Department of Computer and Information Sciences, Towson University, Towson, MD. Before joining Towson, he was with Cisco Systems Inc. for almost nine years. His research interests include cyberspace security, computer networks, cyber-physical systems, and distributed computing.

Dou An received the BS degree in applied mathematics from Northwestern Polytechnical University, China, in 2011. Currently, he is working toward the PhD degree in the Department of Automation Science and Technology, Xi'an Jiaotong University. He is also with the State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University. His research interests include cyber-physical systems, power grid security.



Nan Zhang received the BS degree from Peking University in 2001 and the PhD degree from Texas A&M University in 2006, both in computer science. He is an associate professor of computer science at the George Washington University. His current research interests include security and privacy issues in databases, data mining, and computer networks, in particular privacy and anonymity in data collection, publishing, and sharing, privacy-preserving data mining, and wireless network security and privacy.

Wei Zhao completed his undergraduate program in physics at Shaanxi Normal University, Xian, China, in 1977. He received the MS and PhD degrees in computer and information sciences from the University of Massachusetts at Amherst in 1983 and 1986, respectively. He is currently the rector of the University of Macau. Before joining the University of Macau, he served as the dean of the School of Science at Rensselaer Polytechnic Institute. Between 2005 and 2006, he served as the director for the Division of Computer and Network Systems in the US National Science Foundation when he was on leave from Texas A&M University, where he served as a senior associate vice president for Research and a professor of computer science. He was the founding director of the Texas A&M Center for Information Security and Assurance, which has been recognized as a Center of Academic Excellence in Information Assurance Education by the National Security Agency. Since then, he has served as a faculty member at Amherst College, the University of Adelaide, and Texas A&M University. As an elected IEEE fellow, he has made significant contributions in distributed computing, real-time systems, computer networks, and cyber space security.