Cyber Risk Analysis of Combined Data Attacks Against Power System State Estimation

Pan, Kaikai; Teixeira, André; Cvetkovic, Milos; Palensky, Peter

DOI
10.1109/TSG.2018.2817387

Publication date
2018

Document Version
Final published version

Published in
IEEE Transactions on Smart Grid

Cyber Risk Analysis of Combined Data Attacks Against Power System State Estimation

Kaikai Pan, *Student Member, IEEE*, André Teixeira, *Member, IEEE*, Milos Cvetkovic, *Member, IEEE*, and Peter Palensky, *Senior Member, IEEE*

***Abstract*—Understanding smart grid cyber attacks is key for developing appropriate protection and recovery measures. Advanced attacks pursue maximized impact at minimized costs and detectability. This paper conducts risk analysis of combined data integrity and availability attacks against the power system state estimation. We compare the combined attacks with pure integrity attacks—false data injection (FDI) attacks. A security index for vulnerability assessment to these two kinds of attacks is proposed and formulated as a mixed integer linear programming problem. We show that such combined attacks can succeed with fewer resources than FDI attacks. The combined attacks with limited knowledge of the system model also expose advantages in keeping stealth against the bad data detection. Finally, the risk of combined attacks to reliable system operation is evaluated using the results from vulnerability assessment and attack impact analysis. The findings in this paper are validated and supported by a detailed case study.**

***Index Terms*—Combined integrity and availability attack, false data injection, risk analysis, power system state estimation.**

I. Introduction

THE INCREASINGLY digitized power system offers more data, details, and controls in a real-time fashion than its non-networked predecessors. One of the benefiting applications of this development is State Estimation (SE): Remote Terminal Units (RTUs) provide measurement data via Information and Communication Technology (ICT) infrastructure such as Supervisory Control and Data Acquisition (SCADA) system. The SE provides the operator with an estimate of the state of the electric power system. This state information is then used and processed by the energy management system (EMS) for optimal power flow (OPF), contingency analysis (CA), and automatic generation control (AGC). Security of supply depends on the EMS, which in turn depends on a reliable SE.

As discussed in [1], the SCADA system is vulnerable to a large number of security threats. A class of integrity data attack, known as false data injection (FDI) attack, has been studied with considerable attention. With modifying the measurement data, this attack can pass the Bad Data Detection (BDD) within SE to keep stealth [2], by tampering of RTUs, the communication links to the control center, or even the databases and IT software in the control center. However, such FDI attack needs intensive attack resources such as the knowledge of the system model and the capability to corrupt the integrity on a set of measurements. Denial-of-service (DoS) attacks [3], [4], a type of availability attack, are much "cheaper" to achieve, especially if RTUs communicate via insecure communication channels. In this paper, we focus on combined attacks where the SE is corrupted by both integrity attacks and availability attacks simultaneously. We compare combined attacks and FDI attacks under different levels of adversarial knowledge and resources.

A. State of the Art

Research in the literature has focused on FDI attacks from many aspects of risk assessment [5], e.g., vulnerability analysis, attack impact assessment and mitigation schemes development. As first shown in [2], a class of FDI attack, so-called *stealth* attack, can perturb the state estimate without triggering alarms in BDD within SE. Vulnerability of SE to *stealth FDI attacks* is usually quantified by computing attack resources needed by the attacker to alter specific measurements and keep stealth against the BDD [6]–[8].

Since state estimates are inputs of many application specific tools in EMS, the corrupted estimates can infect further control actions. The estimate errors due to FDI attacks were analyzed in [9] and [10]. The results illustrate that the errors could be significant even with a small number of measurements being compromised. The work in [11] and [12] studied the potential economic impact of FDI attacks against SE by observing the nodal price of market operation. The attacker could obtain economic gain or cause operating costs in the market. Recent work in [13] studied the physical impact of FDI attacks with the attacker's goal to cause a line overflow.

In order to defend against stealth FDI attacks, mitigation schemes have been proposed to improve the bad data detection algorithm or safeguard certain measurements from adversarial data injection. Sequential detection (or quickest detection) of FDI attacks was designed mainly based on well-known Cumulative Sum (CUSUM) algorithm in [14]. In [15], detection methods that leverage synchrophasor data and other

Manuscript received August 14, 2017; revised December 9, 2017 and February 12, 2018; accepted March 7, 2018. Date of publication March 20, 2018; date of current version April 19, 2019. Paper no. TSG-01175-2017. *(Corresponding author: Kaikai Pan.)*

K. Pan, M. Cvetkovic, and P. Palensky are with the Electrical Sustainable Energy Department, Delft University of Technology, 2600 Delft, The Netherlands (e-mail: k.pan@tudelft.nl; m.cvetkovic@tudelft.nl; p.palensky@tudelft.nl).

A. Teixeira is with the Department of Engineering Sciences, Uppsala University, 751 21 Uppsala, Sweden (e-mail: andre.teixeira@angstrom.uu.se).

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.2018.2817387

forecast information were presented. The network layer and application layer mitigation schemes, such as multi-path routing and data authentication and protection, are proved to be effective to decrease the vulnerability [16], [17].

Most of the research above assumes that the adversary has full knowledge of the system model including the power grid topology and transmission line parameters. However, the data of the system model is usually protected well and the attacks are always executed with limited adversarial knowledge. The work in [18] and [19] proposed that an FDI attack can be made with incomplete network information. The attacker can still keep stealth if it knows the local information (topology and line parameters) of the attacking region under certain conditions. Liu and Li [20] also explored how to launch a successful FDI attack against AC state estimation with incomplete knowledge. Another limited knowledge scenario is that the attacker has inaccurate network information of topology and line parameters [21]. Such FDI attacks have the possibility to be detected by the BDD while the detectability is intimately related to the detectability of topology or parameter errors [22]. For these limited knowledge cases, the adversary could also infer necessary network information based on available data using learning methods such as independent component analysis (ICA) [23] and subspace estimation technique [24].

It is worth noting that the majority of research has focused on stealth FDI attacks from a specific aspect of vulnerability or impact assessment. The work in [4] first considered adding a class of availability attacks, so-called jamming attack, to the attack scenarios against SE. Our recent paper [17] first studied the *stealth combined attacks* with different measurement routing topologies, concluding that such attacks may need less attack resources than FDI attacks. Besides, the work above still assumed that the attackers have perfect knowledge of the system model. In practice, we are more interested in the limited adversarial knowledge case that the attacker knows inaccurate network information. Such attacks are not guaranteed to be stealth. In this work we would like to explore how combined attack can differ from FDI attacks in a limited knowledge setting. Intuitively, combined attacks provide the availability attack option to block measurements that the attacker has least knowledge of. This motivates the use of attack resources and the detection probability of attacks with limited knowledge in vulnerability analysis. In addition, vulnerability and impact of attacks can be combined together in the notion of *risk*. In [25], a high-level risk assessment methodology for power system applications including SE was presented. However, risk analysis methods and tools combining vulnerability and impact assessment for data attacks are needed to implement risk assessment methodologies.

In this paper, in contrast to our previous work [17], for the first time we formulate combined attacks with limited knowledge of the system model and we conduct the risk analysis of combined attacks. In order to assess the risk, we first analyze vulnerability of SE with respect to attack resources needed by the adversary and calculate the detection probability of combined attacks. This is a necessary step in deriving the likelihood of the attack. Next, we propose attack impact metric for evaluating attack impact on load estimate. Combining the results from vulnerability and impact assessment, we present the *risk* which combined attacks bring to reliable system operation. We compare the vulnerability, impact and risk with those of FDI attacks. The simulation results show that combined attacks yield higher risk in majority of considered cases.

B. Contributions and Outline

As far as we know, our work is the first one to conduct risk analysis of combined attacks with limited adversarial knowledge. Our contributions are listed as follows:

1) The first part of vulnerability analysis is presented through the notion of security index [7], which corresponds to the minimum attack resources needed by the attacker to compromise the measurements while keeping stealth. The power system is more vulnerable to attacks with smaller security index since such attacks can be executed with less resources. We show that, the optimal solution of combined attack security index problem coincides with the optimal solution of the FDI attack security index problem.
2) Our second contribution is to address the detection probability problem of combined attacks with limited adversarial knowledge. Here we relax the full knowledge assumption which is commonly used in the literature. We show that the optimal combined attack with limited adversarial knowledge can still keep stealth under certain conditions. The empirical results also indicate that combined attacks have lower detection probability.
3) We propose risk metric to quantify the risk of combined attacks with limited adversarial knowledge. For the attacks with the same security index, the risk metric is computed by multiplying 1) the probability of the attack not to be detected, with 2) the attack impact on load estimate. We particularly consider the attack impact on load estimate because the load estimates are inputs of other applications that compute optimal control actions in EMS. Based on the analysis of risk metrics of combined attacks and FDI attacks, we show that power system operations face higher risk under combined attacks.

The outline of the paper is as follows. Section II gives an introduction of SE and stealth FDI attacks mechanism. Section III extends the attack scenario to combined attacks and proposes security index with computational method for vulnerability analysis. In Section IV, the detectability of combined attacks with limited adversarial knowledge is discussed. The risk metric is proposed to measure the risk of attacks in Section V with the analysis of the vulnerability and attack impact. Section VI presents empirical results from a power system use case. In Section VII we conclude the paper.

C. Notation

For an $m \times n$ matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$, we denote the i-th row of $\mathbf{H}$ by $\mathbf{H}(i, :)$. For a vector of m values $\mathbf{a} \in \mathbb{R}^m$, $\mathbf{a}(i)$ is the i-th entry of $\mathbf{a}$. By $diag(\mathbf{a})$, we denote an $m \times m$ diagonal matrix with the elements of vector $\mathbf{a}$ on the main diagonal.

II. Power System Model and Data Attacks

In this section, we review the state estimation and BDD techniques and the stealth data attacks problem.

A. State Estimation

The power system we consider has $n+1$ buses and n_t transmission lines. The data collected by RTUs includes line power flow and bus power injection measurements. These m measurements are denoted by $\mathbf{z} = [z_1, \ldots, z_m]^T$. The system state $\mathbf{x}$ is the vector of phase angles and voltage magnitudes at all buses except the reference bus whose phase angle is set to be zero. For the analysis of cyber security and bad data detection in SE, it is customary to describe the dependencies of measurements and system state through an approximate model called DC power flow model [8]. In the DC power flow model, all the voltage magnitudes are assumed to be constant and the reactive power is completely neglected. Thus the vector $\mathbf{z}$ refers to active power flow and injection measurements, and the state $\mathbf{x}$ refers to bus phase angles only. There are n phase angles to be estimated excluding the reference one, i.e., $\mathbf{x} = [x_1, \ldots, x_n]^T$. Hence, $\mathbf{z}$ and $\mathbf{x}$ are related by the equation

$$\mathbf{z} = \mathbf{P}\begin{bmatrix} \mathbf{W}\mathbf{B}^T \\ -\mathbf{W}\mathbf{B}^T \\ \mathbf{B}_0\mathbf{W}\mathbf{B}^T \end{bmatrix}\mathbf{x} + \mathbf{e} \coloneqq \mathbf{H}\mathbf{x} + \mathbf{e}, \tag{1}$$

where $\mathbf{e} \sim \mathcal{N}(0, \mathbf{R})$ is the measurement noise vector of independent zero-mean Gaussian variables with the covariance matrix $\mathbf{R} = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2)$, $\mathbf{H} \in \mathbb{R}^{m\times n}$ represents the system model, depending on the topology of the power network, the line parameters and the placement of RTUs. Here the topology is described by a directed incidence matrix $\mathbf{B}_0 \in \mathbb{R}^{(n+1)\times n_t}$ in which the directions of the lines can be arbitrarily specified [8]. Matrix $\mathbf{B} \in \mathbb{R}^{n\times n_t}$ is the truncated incidence matrix with the row in $\mathbf{B}_0$ corresponding to the reference bus removed. The line parameters are described by a diagonal matrix $\mathbf{W} \in \mathbb{R}^{n_t\times n_t}$ with diagonal entries being the reciprocals of transmission line reactance. Matrix $\mathbf{P} \in \mathbb{R}^{m\times(2n_t+n+1)}$ is a matrix stacked by the rows of identity matrices, indicating which power flows or bus injections are measured. Usually a large degree of redundancy of measurements is employed to make $\mathbf{H}$ full rank.

The state estimate $\hat{x}$ is obtained by the following weighted least squares (WLS) estimate:

$$\hat{\mathbf{x}} \coloneqq \arg\min_{\mathbf{x}} (\mathbf{z} - \mathbf{H}\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{x}), \tag{2}$$

which can be solved as $\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z} \coloneqq \mathbf{K}\mathbf{z}$.

The estimated state $\hat{\mathbf{x}}$ can be used to estimate the active power flows and injections by

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} = \mathbf{H}\mathbf{K}\mathbf{z} \coloneqq \mathbf{T}\mathbf{z}, \tag{3}$$

where $\mathbf{T}$ is the so-called hat matrix [26]. The BDD scheme uses such estimated measurements to identify bad data by comparing $\hat{\mathbf{z}}$ with $\mathbf{z}$, see below.

B. Bad Data Detection

Measurement data may be corrupted by random errors. Thus there is a built-in BDD scheme in EMS for bad data detection. The BDD is achieved by hypothesis tests using the statistical properties of the measurement residual:

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = (\mathbf{I} - \mathbf{T})\mathbf{z} \coloneqq \mathbf{S}\mathbf{z} = \mathbf{S}\mathbf{e}, \tag{4}$$

where $\mathbf{r} \in \mathbb{R}^m$ is the residual vector, $\mathbf{I} \in \mathbb{R}^{m\times m}$ is an identity matrix and $\mathbf{S}$ is the so-called residual sensitivity matrix [26].

We now introduce the $J(\hat{\mathbf{x}})$-test based BDD. For the measurement error $\mathbf{e} \sim \mathcal{N}(0, \mathbf{R})$, the new random variable $y = \sum_i^m R_{ii}^{-1} e_i^2$ where R_{ii} is the diagonal entry of the covariance matrix $\mathbf{R}$ has a χ^2 distribution with $m - n$ degrees of freedom. Note the quadratic cost function $J(\hat{\mathbf{x}}) = \|\mathbf{R}^{-1/2}\mathbf{r}\|_2^2 = \|\mathbf{R}^{-1/2}\mathbf{S}\mathbf{e}\|_2^2$. For the independent m measurements we have $\text{rank}(\mathbf{S}) = m - n$, which implies that $J(\hat{\mathbf{x}})$ has a so-called *generalized chi-squared distribution* with $m - n$ degrees of freedom [27]. The BDD uses the quadratic function as an approximation of y and checks if it follows the distribution χ_{m-n}^2. Defining $\alpha \in [0, 1]$ as the significance level corresponding to the false alarm rate, and $\tau(\alpha)$ such that

$$\int_0^{\tau(\alpha)} f(x)dx = 1 - \alpha, \tag{5}$$

where $f(x)$ is the probability distribution function (PDF) of χ_{m-n}^2. Hence, the BDD scheme becomes

$$\begin{cases} \text{Good data,} & \text{if } \|\mathbf{R}^{-1/2}\mathbf{r}\|_2 \le \sqrt{\tau(\alpha)}, \\ \text{Bad data,} & \text{if } \|\mathbf{R}^{-1/2}\mathbf{r}\|_2 > \sqrt{\tau(\alpha)}. \end{cases} \tag{6}$$

C. Stealth FDI Attacks

The goal of an attacker is to perturb the SE while remaining hidden from the BDD. If only data integrity attacks are considered, the attacker could inject false data on a set of measurements, modifying the measurement vector $\mathbf{z}$ into $\mathbf{z}_a \coloneqq \mathbf{z} + \mathbf{a}$. Here the *FDI attack vector* $\mathbf{a} \in \mathbb{R}^m$ is the corruption added to the original measurement $\mathbf{z}$. We have the following definition of a k_a-tuple FDI attack,

Definition 1 (k_a-Tuple FDI Attack): An attack with an FDI attack vector $\mathbf{a} \in \mathbb{R}^m$ is called a k_a-tuple FDI attack if a number of k_a measurements are injected with false data, i.e., $\|\mathbf{a}\|_0 = k_a$.

As shown in [2], an attacker with full knowledge of the system model (i.e., the matrix $\mathbf{H}$) and the capability to corrupt specific measurements can keep stealth if the FDI attack vector follows $\mathbf{a} = \mathbf{H}\mathbf{c}$ where $\mathbf{c} \in \mathbb{R}^n$ is non-zero. The corrupted measurements $\mathbf{z}_a$ becomes $\mathbf{z}_a = \mathbf{H}(\mathbf{x}+\mathbf{c})+\mathbf{e}$. This leads to the state estimate perturbed by a degree of c, while the residual for BDD checking remains the same. It has been verified that such *stealth FDI attacks* based on the DC model can be performed on a real SCADA/EMS testbed avoiding the bad data detection with full nonlinear AC power flow model [9].

To describe the vulnerability of SE to stealth FDI attacks with full knowledge of the system model, the security index is introduced as the minimum number of measurements that need to be corrupted by the attacker in order to keep stealth [7].

The security index is given by

$$\begin{aligned} \alpha_j \coloneqq \min_{\mathbf{c}} \quad & \|\mathbf{a}\|_0 \\ \text{s.t.} \quad & \mathbf{a} = \mathbf{H}\mathbf{c}, \quad \mathbf{a}(j) = \mu, \\ & \mathbf{a}(l) = 0 \quad \text{for all } l \in \Gamma, \end{aligned} \tag{7}$$

where $\mathbf{a}(j)$ denotes the injected false data on measurement j, and μ is the non-zero *attack magnitude* determined by the attacker. We add the constraint that the pseudo-measurements (in the set Γ) corresponding to zero-injection buses cannot be attacked. The result α_j is the security index that quantifies the vulnerability of measurement j to stealth FDI attacks. Here the computed α_j belongs to one of the FDI attacks with the minimum k_a ($k_a = \alpha_j$) for measurement j. It is known that this optimization problem above is NP-hard (See [28]). Teixeira *et al.* [8] proposed an approach using the big M method to directly express (7) as a mixed integer linear programming (MILP) problem which can be solved with an appropriate solver,

$$\begin{aligned} \alpha_j \coloneqq \min_{\mathbf{c},\mathbf{y}} \quad & \sum_{i=1}^{m} \mathbf{y}(i) \\ \text{s.t.} \quad & \mathbf{H}\mathbf{c} \le M\mathbf{y}, && \text{(8a)} \\ & -\mathbf{H}\mathbf{c} \le M\mathbf{y}, && \text{(8b)} \\ & \mathbf{H}(j,:)\mathbf{c} = \mu, && \text{(8c)} \\ & \mathbf{H}(l,:)\mathbf{c} = 0, \quad \text{for all } l \in \Gamma, && \text{(8d)} \\ & \mathbf{y}(i) \in \{0,1\} \quad \text{for all } i. \end{aligned}$$

In (8a), M is a constant scalar that is greater than the maximum absolute value of entries in $\mathbf{H}\mathbf{c}^*$, for some optimal solution $\mathbf{c}^*$ of (7). At optimality, for any i that $|\mathbf{H}(i,:)\mathbf{c}^*| = 0$, the corresponding $\mathbf{y}(i)$ is zero. Thus an optimal solution to (8a) is exactly the same optimal solution to (7) with $\mathbf{y}(i) = 1$ indicating that the measurement i is corrupted by an FDI attack. Here the *attack magnitude* μ is determined by the attacker and is set as a tunable parameter in the optimization problem (8). Thus, the attacker can vary the attack magnitude based on the possible constraints arising from the presence of measurement forecasts and range limitations. We denote the optimization problem (8) which computes the FDI attacks as $\mathcal{P}_a(\mathbf{H})$ where $\mathbf{H}$ corresponds to the full system model.

III. Stealth Combined Data Attacks

FDI attacks are resource-intensive since the adversary needs to coordinate integrity attacks on all targeted measurements. This usually gives the adversary more power than possible in practice [10]. In reality, an attacker would try to reduce the attack resources and would prefer data availability attacks (e.g., DoS attacks, jamming attacks) since monitoring systems are always more vulnerable to this type of attacks [29]. Thus, we focus on the scenario that the adversary would launch the combined data integrity and availability attacks.

A. Combined Data Integrity and Availability Attacks

For a large-scale SCADA system, missing data and failing RTUs are common [7]. When some of the measurements are missing, the typical solution widely employed in SE is to use the remaining data before the system becomes "unobservable". Another solution is to use pseudo measurements (e.g., previous data, forecast information), but these measurements would still lose confidence in further time intervals as long as the availability attacks continue. The combined attacks we introduce here are attacks which will not make the system unobservable or lead to non-convergence of the SE algorithm. We say that such combined attacks can still keep stealth against the BDD, with the following definition.

Definition 2 (Stealth Combined Attacks): Attacks which can launch both availability attack and FDI attack are called stealth combined attacks if no additional alerts are triggered in the current BDD.

In practice, the current BDD scheme employed in SE would not trigger alarms when some measurements are missing. Besides, even when availability attacks happen, they may be misdiagnosed as poor network conditions or physical damages to the sensors. Thus we keep the assumption in this paper that SE uses remaining data if availability attacks take place and availability attack would not trigger additional alerts in BDD. We introduce the *availability attack vector* $\mathbf{d} \in \{0,1\}^m$ for the availability attacks and $\mathbf{d}(i) = 1$ means that measurement i is unavailable. Thus the model for remaining measurements and system state can be described by

$$\mathbf{z}_d = \mathbf{H}_d x + \mathbf{e}_d, \tag{9}$$

where $\mathbf{e}_d \in \mathbb{R}^m$ and $\mathbf{z}_d \in \mathbb{R}^m$ are the noise vector and measurement vector respectively, and the entries of them are zero if the corresponding measurements are unavailable. Matrix $\mathbf{H}_d \in \mathbb{R}^{m\times n}$ denotes the model of the remaining measurements and it is obtained from $\mathbf{H}$ by replacing some rows with zero row vectors due to availability attacks on these measurements, i.e., $\mathbf{H}_d \coloneqq (\mathbf{I} - \mathrm{diag}(\mathbf{d}))\mathbf{H}$. We can further obtain the hat matrix and residual sensitivity matrix when availability attacks occur,

$$\mathbf{K}_d \coloneqq \left(\mathbf{H}_d^T \mathbf{R}^{-1} \mathbf{H}_d\right)^{-1} \mathbf{H}_d^T \mathbf{R}^{-1}, \tag{10}$$

$$\mathbf{T}_d \coloneqq \mathbf{H}_d \mathbf{K}_d, \quad \mathbf{S}_d \coloneqq \mathbf{I} - \mathbf{T}_d. \tag{11}$$

For the combined attacks, the attacker would still launch FDI attacks on the remaining measurements in concert with availability attacks, making $\mathbf{z}_d$ changed into $\mathbf{z}_{a,d} \coloneqq \mathbf{z}_d + \mathbf{a}$. Similarly, a (k_a, k_d)-tuple combined attack can be defined as.

Definition 3 ((k_a, k_d)-Tuple Combined Attack): A combined attack with an FDI attack vector $\mathbf{a} \in \mathbb{R}^m$ and an availability attack vector $\mathbf{d} \in \{0,1\}^m$ described above is called a (k_a, k_d)-tuple combined attack if $\|\mathbf{a}\|_0 = k_a$, $\|\mathbf{d}\|_0 = k_d$.

B. Security Index for Combined Attacks

Similar to the FDI attacks, if the attack vectors of a (k_a, k_d)-tuple attack satisfy $\mathbf{a} = \mathbf{H}_d\mathbf{c}$, such combined attacks can still keep stealth as the FDI attack vector $\mathbf{a}$ lies on the column space of the matrix $\mathbf{H}_d$. Using the formulation of security index in (7) for FDI attacks, we propose an intuitive security index for combined attacks as the minimum number of measurements

that need to be compromised by the attacker,

$$\begin{aligned}\beta_j \coloneqq \min_{\mathbf{c},\mathbf{d}} \quad & \|\mathbf{a}\|_0 + \|\mathbf{d}\|_0 \\ \text{s.t.} \quad & \mathbf{a} = \mathbf{H}_d\mathbf{c}, && (12a)\\ & \mathbf{H}_d = (\mathbf{I} - \text{diag}(\mathbf{d}))\mathbf{H}, && (12b)\\ & \mathbf{a}(j) = \mu, && (12c)\\ & \mathbf{a}(l) = 0 \quad \text{for all } l \in \Gamma, && (12d)\\ & \mathbf{d}(i) \in \{0,1\} \quad \text{for all } i.\end{aligned}$$

Here we also assume $\mathbf{a}(j) = \mu$ where μ is the non-zero *attack magnitude*. The result β_j is the security index that quantifies how vulnerable measurement j is to combined attacks. The computed β_j belongs to one of the combined attacks that have minimum $k_a + k_d$ ($k_a + k_d = \beta_j$) for measurement j. To solve this NP-hard problem above, we propose a computation solution which uses the big M method to formulate a MILP problem:

$$\begin{aligned}\beta_j^{'} \coloneqq \min_{\mathbf{c},\mathbf{w},\mathbf{d}} \quad & \sum_{i=1}^{m}\mathbf{w}(i) + \sum_{k=1}^{m}\mathbf{d}(k) \\ \text{s.t.} \quad & \mathbf{Hc} \le M(\mathbf{w}+\mathbf{d}), && (13a)\\ & -\mathbf{Hc} \le M(\mathbf{w}+\mathbf{d}), && (13b)\\ & \mathbf{H}(j,:)\mathbf{c} = \mu, && (13c)\\ & \mathbf{H}(l,:)\mathbf{c} = 0, \quad \text{for all } l \in \Gamma, && (13d)\\ & \mathbf{w}(i) \in \{0,1\} \quad \text{for all } i, && (13e)\\ & \mathbf{d}(k) \in \{0,1\} \quad \text{for all } k, && (13f)\end{aligned}$$

where $\mathbf{w}, \mathbf{d} \in \{0,1\}^m$ with $\mathbf{w}(i) = 1$ and $\mathbf{d}(k) = 1$ meaning FDI attack and data availability attack on measurement i and k.

The following theorem shows that the optimal solution to (12a) can be obtained from the optimal solution of (13a). Hence we denote the optimization problem (13) which computes the combined attacks as $\mathcal{P}_{a,d}(\mathbf{H})$. By solving $\mathcal{P}_a(\mathbf{H})$ from (8a) and $\mathcal{P}_{a,d}(\mathbf{H})$ from (13a), the system operators can obtain the attack vectors and further assess the risk of attacks on the measurements, which will be illustrated in Section V.

Theorem 1: For any index $j \in \{1,\dots,m\}$ and non-zero μ, let ($\mathbf{c}^*$, $\mathbf{w}^*$, $\mathbf{d}^*$) be an optimal solution to (13a). Then an optimal solution to (12a) can be computed as ($\mathbf{c}^*$, $\mathbf{d}^*$), and $\beta_j^{'} = \beta_j$.

Proof: The proof follows by re-writing (12a) as (13a). First, note that the constraint of (12a), $\mathbf{a} = (\mathbf{I} - \text{diag}(\mathbf{d}))\mathbf{Hc}$, can be formulated as a set of inequality constraints with auxiliary binary variables by using the big M method, yielding $-M\mathbf{w} \le (\mathbf{I} - \text{diag}(\mathbf{d}))\mathbf{Hc} \le M\mathbf{w}$, where $\mathbf{w} \in \{0,1\}^m$ and $\|\mathbf{a}\|_0 = \sum \mathbf{w}(i)$. Since $\mathbf{d}$ is a vector of binary variables, the pair of inequality constraints pertaining the i-th measurement can be written as $|(1 - \mathbf{d}(i))\mathbf{H}(i,:)\mathbf{c}| \le M\mathbf{w}(i)$. The latter can be read as

$$\begin{cases}\mathbf{H}(i,:)\mathbf{c} = 0, & \text{if } \mathbf{w}(i) = \mathbf{d}(i) = 0,\\ |\mathbf{H}(i,:)\mathbf{c}| \le M, & \text{if } \mathbf{w}(i) = 1 \text{ or } \mathbf{d}(i) = 1,\end{cases}$$

which can be rewritten as $|\mathbf{H}(i,:)\mathbf{c}| \le M(\mathbf{d}(i) + \mathbf{w}(i))$. Hence, recalling that $\mathbf{a}(i) = (1 - \mathbf{d}(i))\mathbf{H}(i,:)\mathbf{c}$, we conclude that the constraints of (12a) can be equivalently re-written as the constraints of (13a). The proof concludes by noting that the objective functions of both problems satisfy the equality $\|\mathbf{a}\|_0 + \|\mathbf{d}\|_0 = \sum \mathbf{w}(i) + \sum \mathbf{d}(i)$. ■

Corollary 1: For any index $j \in \{1,\dots,m\}$ and non-zero μ, let ($\mathbf{c}^*$, $\mathbf{w}^*$, $\mathbf{d}^*$) be an optimal solution to (13a). Then an optimal solution to (7) can be computed as $\mathbf{c}^*$, and $\alpha_j = \beta_j$.

Proof: The proof follows straightforwardly from Theorem 1, which establishes that an optimal solution to (12a) can be obtained from an optimal solution to (13a): comparing (13a) and (8a), we can easily see that an optimal solution to (8a) can be computed as ($\mathbf{c}^*$, $\mathbf{y}^*$) with $\mathbf{y}^* = \mathbf{w}^* + \mathbf{d}^*$, and $\alpha_j = \beta_j^{'}$. Since (8a) provides the exact solution to (7), an optimal solution to (7) can be computed as $\mathbf{c}^*$, and also $\alpha_j = \beta_j^{'} = \beta_j$. ■

Corollary 1 implies that a set of compromised measurements is an optimal solution to (12a) if and only if this set is an optimal solution to (7), and the two security indexes β_j and α_j coincide. In fact, in [30] it was shown that the set of compromised measurements in a k_a-tuple FDI attack obtained by solving (7) is a sparsest *critical tuple* containing the target measurement j. A sparsest critical tuple is characterized by the measurements that do not belong to a critical tuple of lower order. A critical tuple contains a set of measurements, where removal all of them will cause the system to be unobservable. If any subset of the critical tuple is removed, it would not lead to the loss of observability [26]. According to Corollary 1 and its proof, we can see that the set of compromised measurements of FDI attacks in this critical tuple is also an optimal solution to the security index problem (12a) of combined attacks. The interpretation of the security index problem as a critical tuple problem provides the means for comparing security indexes of attacks with full and limited adversarial knowledge; see Section IV-C for details.

The optimization problems $\mathcal{P}_a(\mathbf{H})$ and $\mathcal{P}_{a,d}(\mathbf{H})$ derived so far in (8a) and (13a) could identify the compromised measurements set of attacks but did not consider the attack costs. In what follows, we include the costs in the formulation. To simplify the discussion, we assume that the availability and integrity attacks have the costs C_A and C_I, respectively, per measurement. Thus we formulate a security index for attack resources of combined attacks as

$$\begin{aligned}\gamma_j^{a,d} \coloneqq \min_{\mathbf{c},\mathbf{w},\mathbf{d}} \quad & \sum_{i=1}^{m} C_I\mathbf{w}(i) + \sum_{k=1}^{m} C_A\mathbf{d}(k) && (14)\\ \text{s.t.} \quad & \text{(13a)}-\text{(13f)}.\end{aligned}$$

By making vector $\mathbf{d}$ in (14) to be zero, we can get the security index γ_j^a for FDI attacks. We can also see that the set of compromised measurements from the optimal solution of (14) is also the optimal solution to (12a) and (7). If $C_A = C_I$, this is the same case as the one described in Corollary 1. For C_A and C_I with different values, we have the following proposition.

Proposition 1: When $C_A < C_I$, the optimal strategy of combined attack is to inject false data on the targeted measurement j and make other measurements in the critical tuple unavailable to the SE, yielding a $(1,\beta_j - 1)$-tuple combined attack with optimal attack cost $\gamma_j^{a,d} = C_I + (\beta_j - 1)C_A$. When

$C_A > C_I$, the combined attack has the same optimal strategy as the FDI attack, i.e., injecting false data on the all measurements in the critical tuple, yielding a (β_j,0)-tuple combined attack (i.e., β_j-tuple FDI attack) with optimal attack cost $\gamma_j^{a,d} = \beta_j C_I$.

Proof: If we take the values that satisfy $C_A < C_I$, the optimal solution of $\mathbf{w}^*$ and $\mathbf{d}^*$ in (14), w.r.t. measurement j, would lead to $\sum \mathbf{w}^*(i) = 1$ and $\sum \mathbf{d}^*(k) = \beta_j - 1$. This means that the optimal combined attack in the case of $C_A < C_I$ is to corrupt one measurement with an integrity attack and make other measurements in this critical tuple unavailable. If we take the values that satisfy $C_A > C_I$, the optimal solution of $\mathbf{w}^*$ and $\mathbf{d}^*$ in (14), w.r.t. measurement j, would lead to $\sum \mathbf{w}^*(i) = \beta_j$ and $\sum \mathbf{d}^*(k) = 0$, i.e., the optimal combined attack is to inject false data on all the measurements in this critical tuple. ■

As previously indicated, availability attacks can cost less attack resources compared with integrity attacks. An intuitive example is that the attacker uses the same tool to perform a Man-In-The-Middle (MITM) attack on the exchanged measurements between substations and the control center. Thus the adversary is capable of interfering with the transmitted measurements using the MITM tool, either launching FDI or availability attacks. Unlike the FDI attack in which the attacker has to inject specific data values and repackage the packets carefully, the availability attack only needs to block the measurements or modify the data to zero or random errors [31]. Using the same MITM tool, the availability attacks become "cheaper" to achieve than FDI attacks. Of course, the true attack costs of different kinds of attacks launched by different tools are hard to quantify in practice. One possible way is to relate the attack cost to the inverse-likelihood of the attack. Likelihood assessment of attacks using attack trees or graphs also implies that availability attacks (e.g., DoS attacks, jamming attacks) have higher probability to take place considering the factors (skills, knowledge, time, etc.) [32]. Thus in the following of this paper we take the values that satisfy $C_A \leq C_I$. The above Proposition 1 for the case $C_A < C_I$ will also be validated in Section VI-A.

IV. Attacks With Limited Adversarial Knowledge

From this section we consider the scenario in which the adversary has limited knowledge of the system model and we discuss how this affects the detectability of combined attacks and FDI attacks.

A. Relaxing Assumption on Adversarial Knowledge

For the combined attacks and FDI attacks above, the adversary is assumed to have full knowledge of $\mathbf{H}$ in (1) that includes the topology of the power network, the placement of RTUs and the transmission line reactance. This system data is kept in the database of control center, which is difficult to be accessed by the attacker. We extend the previous analysis by replacing the full knowledge assumption. Hence, in what follows the attacker only has limited knowledge of the system model. In particular, the limited knowledge case that is of interest to us is the one in which the attackers have inaccurate network information. Now the system model known by the adversary gets "perturbed" that system model uncertainties exist. An attacker could acquire perturbed system model as a result of analyzing an out-dated or estimated model using power network topology data but limited information of transmission line parameters [21], [22], [33].

Looking at the problem from the attacker's perspective, without loss of generality, the perturbed system model known by the attacker can be denoted as $\tilde{\mathbf{H}}$, such that

$$\tilde{\mathbf{H}} \triangleq \mathbf{H} + \Delta\mathbf{H}, \tag{15}$$

where $\Delta\mathbf{H} \in \mathbb{R}^{m\times n}$ denotes the part of model uncertainty. We still consider that the attacker uses the same linear policies to compute attack vectors, i.e., $\mathbf{a} = \tilde{\mathbf{H}}_d\mathbf{c}$ for combined attacks and $\mathbf{a} = \tilde{\mathbf{H}}\mathbf{c}$ for FDI attacks and $\tilde{\mathbf{H}}_d := (\mathbf{I} - \text{diag}(\mathbf{d}))\tilde{\mathbf{H}}$. Correspondingly, we denote the optimization problem (8) as $\mathcal{P}_a(\tilde{\mathbf{H}})$ w.r.t $\tilde{\mathbf{H}}$ computing the FDI attacks and the optimization problem (13) as $\mathcal{P}_{a,d}(\tilde{\mathbf{H}})$ w.r.t $\tilde{\mathbf{H}}$ computing the combined attacks.

B. Detectability of Data Attacks

1) Combined Attacks: When the measurements are corrupted by a (k_a, k_d)-tuple attack, the measurement residual $\mathbf{r}_{a,d}$ can be written as

$$\mathbf{r}_{a,d} = \mathbf{S}_d\mathbf{z}_{a,d} = \mathbf{S}_d\mathbf{e}_d + \mathbf{S}_d\mathbf{a}. \tag{16}$$

As discussed in Section III-B, when the attack vectors of the combined attack satisfy $\mathbf{a} = \mathbf{H}_d\mathbf{c}$, the residual $\mathbf{r}_{a,d} = \mathbf{S}_d\mathbf{e}_d + \mathbf{S}_d\mathbf{H}_d\mathbf{c} = \mathbf{S}_d\mathbf{e}_d$ due to $\mathbf{S}_d\mathbf{H}_d = 0$, then the residual is not affected by $\mathbf{a}$ and no additional alarms are triggered; the BDD treats the measurements attacked by availability attacks as a case of missing data. However, for the attack with limited knowledge, the attack vector $\mathbf{a}$ becomes $\mathbf{a} = \tilde{\mathbf{H}}_d\mathbf{c}$ and $\mathbf{S}_d\mathbf{a}$ may be non-zero. In this case, the residual is incremented and the attack can be detected with some possibility.

Note that the quadratic cost function with the combined attack becomes $J_{a,d}(\hat{x}) = \|\mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{e}_d + \mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{a}\|_2^2$. Here the mean of $(\mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{e}_d + \mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{a})$ is non-zero $\mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{a}$ incremented by the attack. Recalling the $J(\hat{\mathbf{x}})$-test based BDD, $J_{a,d}(\hat{\mathbf{x}})$ has a *generalized non-central chi-squared distribution* with $m-n-k_d$ degrees of freedom under the combined attack. We use $J_{a,d}(\hat{\mathbf{x}})$ as an approximation of having the *non-central chi-squared distribution* $\chi^2_{m-n-k_d}(\|\mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{a}\|_2^2)$ to calculate the detection probability, where $\lambda_{a,d} = \|\mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{a}\|_2^2$ is the non-centrality parameter. Further we will validate such approximation using empirical results from Monte Carlo simulation in Section VI-B. We can further obtain

$$\int_0^{\tau_d(\alpha)} f_{\lambda_{a,d}}(x)dx = 1 - \delta_{a,d}, \tag{17}$$

where $f_{\lambda_{a,d}}(x)$ is the PDF of $\chi^2_{m-n-k_d}(\|\mathbf{R}^{-1/2}\mathbf{S}_d\mathbf{a}\|_2^2)$, $\tau_d(\alpha)$ is the threshold set in the BDD using (5) but with the PDF of $\chi^2_{m-n-k_d}$, and $\delta_{a,d}$ is the detection probability.

2) FDI Attacks: For a k_a-tuple FDI attack with limited knowledge, the quadratic function $J_a(\hat{\mathbf{x}})$ can also be approximated to have a non-central chi-squared distribution but with $m - n$ degrees of freedom, namely the distribution

$\chi^2_{m-n}(\|\mathbf{R}^{-1/2}\mathbf{Sa}\|_2^2)$. Similar to (17), the detection probability can be computed by solving

$$\int_0^{\tau(\alpha)} f_{\lambda_a}(x)dx = 1 - \delta_a, \tag{18}$$

where $\lambda_a = \|\mathbf{R}^{-1/2}\mathbf{Sa}\|_2^2$ denotes the non-centrality parameter, $\tau(\alpha)$ is the threshold set in the BDD using (5), and δ_a is the detection probability of the FDI attack.

C. Special Case: Attacks With Structured Model Uncertainty

An interesting analysis can be made to understand what the model uncertainty $\Delta\mathbf{H}$ in (15) is to the adversary. As stated in [22], the scenarios where the uncertainty is more structured are of greater interest. Here we assume that the attacker knows the exact topology of the power network, but has to estimate the line parameters. This assumption is feasible since the attacker can access the topology information by 1) collecting offline data such as topology maps and online data using attacker's own meters; 2) using market data to extract it from locational marginal prices; 3) utilizing available power flow measurements and compromised breaker status data, as summarized in [34]. However, usually the attacker has limited access to the knowledge of the exact length of the transmission line and type of the conductor being used. Even if the attacker obtains such knowledge, the values would get changed by the time of implementing the attack due to weather conditions and changes in temperature [21]. Denote the line parameters matrix with errors as $\tilde{\mathbf{W}} \triangleq \mathbf{W} + \mathbf{S}\Phi$ where $\Phi \in \mathbb{R}^{n_t \times n_t}$ is the parameter uncertainty. Thus the model with such structured uncertainty becomes

$$\tilde{\mathbf{H}} = \mathbf{P}\begin{bmatrix} (\mathbf{W}+\Phi)\mathbf{B}^T \\ -(\mathbf{W}+\Phi)\mathbf{B}^T \\ \mathbf{B}_0(\mathbf{W}+\Phi)\mathbf{B}^T \end{bmatrix} \Rightarrow \Delta\mathbf{H} = \mathbf{P}\begin{bmatrix} \Phi\mathbf{B}^T \\ -\Phi\mathbf{B}^T \\ \mathbf{B}_0\Phi\mathbf{B}^T \end{bmatrix}, \tag{19}$$

Now we consider the security index of attacks w.r.t. $\tilde{\mathbf{H}}$ in (19). As we have discussed in Section III-B, the security index problem can be interpreted as a critical tuple problem. In the remaining part of this paper we adopt the following assumption,

Assumption 1: The system with perturbed model $\tilde{\mathbf{H}}$ in (19) has the same sets of critical tuples as the system with original model $\mathbf{H}$ in (1).

Assumption 1 is expected to hold in the case that the system with $\mathbf{H}$ in (1) is topologically observable [35]. Defining the security indexes for compromised measurements set under structured uncertainty model as $\tilde{\alpha}_j$ and $\tilde{\beta}_j$, respectively, the following theorem shows that the security index remains the same although the model is perturbed with structured uncertainty.

Theorem 2: For any measurement index $j \in \{1, \ldots, m\}$ and non-zero μ, under Assumption 1, let $(\tilde{\mathbf{c}}^*, \tilde{\mathbf{w}}^*, \tilde{\mathbf{d}}^*)$ be an optimal solution to $\mathcal{P}_{a,d}(\tilde{\mathbf{H}})$ ($\tilde{\mathbf{H}}$ is from (19)). Then there exists some $\mathbf{c}^*$ such that $(\mathbf{c}^*, \mathbf{w}^*, \mathbf{d}^*)$ with $\mathbf{w}^* = \tilde{\mathbf{w}}^*$ and $\mathbf{d}^* = \tilde{\mathbf{d}}^*$ is an optimal solution to $\mathcal{P}_{a,d}(\mathbf{H})$, $(\mathbf{c}^*, \mathbf{y}^*)$ with $\mathbf{y}^* = \tilde{\mathbf{w}}^* + \tilde{\mathbf{d}}^*$ is an optimal solution to $\mathcal{P}_a(\mathbf{H})$, and $\tilde{\beta}_j = \beta_j = \alpha_j = \tilde{\alpha}_j$.

Proof: The optimal solution with $\tilde{\mathbf{w}}^*$ and $\tilde{\mathbf{d}}^*$ identifies a sparsest critical tuple containing measurement j for the perturbed model $\tilde{\mathbf{H}}$ in (19), which is also a sparsest critical tuple for the model $\mathbf{H}$ in (1) according to Assumption 1. Then the set of measurements in this critical tuple is an optimal solution to $\mathcal{P}_{a,d}(\mathbf{H})$. According to Theorem 1 and Corollary 1, the set of measurements in this critical tuple is also an optimal solution to $\mathcal{P}_a(\mathbf{H})$. ■

With respect to the security index for attack resources, let $\tilde{\gamma}_j^{a,d}$ and $\tilde{\gamma}_j^a$ be the security indexes of combined attacks and FDI attacks from (14) but w.r.t. perturbed model $\tilde{\mathbf{H}}$ in (19). We can see that the set of compromised measurements from optimal solution to (14) w.r.t. $\tilde{\mathbf{H}}$ in (19) is also the optimal solution to (13a) and (8a) according to Theorem 2. When it is the case that $C_A < C_I$, the optimal solution of $\tilde{\mathbf{w}}^*$ and $\tilde{\mathbf{d}}^*$ from (14) w.r.t. $\tilde{\mathbf{H}}$, would lead to $\sum \tilde{\mathbf{w}}^*(i) = 1$ and $\sum \tilde{\mathbf{d}}^*(k) = \tilde{\beta}_j - 1$. Such $(1,\tilde{\beta}_j - 1)$-tuple combined attack can be launched with least attack resources when $C_A < C_I$ and in the following we show that it also can achieve minimized detectability.

As discussed in Section IV-B, the detection probability would increase when attacker has limited knowledge of the system model. However, for the combined attacks, the following proposition states that the combined attacks with structured model uncertainty can still keep stealth against the BDD if the following conditions are satisfied: 1) structured model uncertainty is defined as in (19); 2) Assumption 1 holds.

Proposition 2: For any index $j \in \{1, \ldots, m\}$ and non-zero μ, under Assumption 1, let $(\tilde{\mathbf{c}}^*, \tilde{\mathbf{w}}^*, \tilde{\mathbf{d}}^*)$ with $\sum \tilde{\mathbf{w}}^*(i) = 1$ be an optimal solution to $\mathcal{P}_{a,d}(\tilde{\mathbf{H}})$ ($\tilde{\mathbf{H}}$ is from (19)). Then this $(1,\tilde{\beta}_j - 1)$-tuple combined attack from $(\tilde{\mathbf{c}}^*, \tilde{\mathbf{w}}^*, \tilde{\mathbf{d}}^*)$ is a stealth attack.

Proof: The FDI attack vector of this combined attack is $\mathbf{a} = \tilde{\mathbf{H}}_{\tilde{\mathbf{d}}^*}\tilde{\mathbf{c}}^*$. According to Theorem 2, there exists $\mathbf{c}^*$ such that $(\mathbf{c}^*, \mathbf{w}^*, \mathbf{d}^*)$ with $\mathbf{w}^* = \tilde{\mathbf{w}}^*$ and $\mathbf{d}^* = \tilde{\mathbf{d}}^*$ is an optimal solution to $\mathcal{P}_{a,d}(\mathbf{H})$. Using the attack strategy above, $k_a = \sum \tilde{\mathbf{w}}^*(i) = 1$ and the only non-zero entry of the attack vector $\mathbf{a}$ is μ while other measurements in this critical tuple are attacked by availability attacks. Thus this combined attack is with the vector $\mathbf{a} = (\mathbf{I} - \text{diag}(\tilde{\mathbf{d}}^*))\tilde{\mathbf{H}}\tilde{\mathbf{c}}^* = (\mathbf{I} - \text{diag}(\mathbf{d}^*))\mathbf{H}\mathbf{c}^* = \mathbf{H}_{\mathbf{d}^*}\mathbf{c}^*$, which can keep stealth w.r.t. $\mathbf{H}$ in (1). ■

It should be noted that, Proposition 2 is independent from the parameter uncertainty Φ. This $(1,\tilde{\beta}_j - 1)$-tuple combined attack can always keep stealth for any parameter uncertainty levels as long as the critical tuple is correctly identified by solving $\mathcal{P}_{a,d}(\tilde{\mathbf{H}})$.

V. Risk Assessment for Data Attacks

The previous sections focus on vulnerability assessment of SE to combined attacks with limited knowledge. Following the procedure of risk analysis in [25], in this section we define and analyze the *risk* brought by attacks with limited knowledge.

Usually the total *risk* of data attacks is defined as the likelihood of attack multiplied by the potential attack impact [5]. For a (k_a, k_d)-tuple combined attack, the risk metric $R(\mathbf{a}, \mathbf{d})$ can be expressed as

$$R(\mathbf{a}, \mathbf{d}) = L(\mathbf{a}, \mathbf{d}) * I(\mathbf{a}, \mathbf{d}) \tag{20}$$

where $L(\mathbf{a}, \mathbf{d})$ denotes the likelihood of the combined attack with attack vectors $\mathbf{a}$ and $\mathbf{d}$, and $I(\mathbf{a}, \mathbf{d})$ denotes the attack impact. For the attacks with larger risk metrics, they bring

more risk to reliable system operation. In the following we discuss how $L(\mathbf{a}, \mathbf{d})$ and $I(\mathbf{a}, \mathbf{d})$ are formulated.

A. Likelihood of Data Attacks

The attack likelihood relates to the vulnerability of the system. In this work, the likelihood of the attack is taken as the probability that the attack is launched and the probability that the attack can keep stealth against the detection schemes,

$$L(\mathbf{a}, \mathbf{d}) = \mathbb{P}(\mathbf{a}, \mathbf{d})\mathbb{P}(s|\,\mathbf{a}, \mathbf{d}), \tag{21}$$

where $\mathbb{P}(s|\mathbf{a}, \mathbf{d})$ denotes the conditional probability of the combined attack passing the BDD if it has been performed successfully. For the attack with limited knowledge, the detection probability $\delta_{a,d}$ can be obtained from (17), thus we have $\mathbb{P}(s|\mathbf{a}, \mathbf{d}) = 1-\delta_{a,d}$. In (21), $\mathbb{P}(\mathbf{a}, \mathbf{d})$ represents the probability that a particular adversary would perform a combined attack and successfully corrupt the data. Obtaining meaningful and realistic data for calculating $\mathbb{P}(\mathbf{a}, \mathbf{d})$ remains an unsolved and open issue for most of the established approaches [36]. The proposed security index $\tilde{\gamma}_j^{a,d}$ w.r.t. perturbed model $\tilde{\mathbf{H}}$ captures the efforts required by a combined attack and essentially can be related to the probability $\mathbb{P}(\mathbf{a}, \mathbf{d})$. We assume that if the attacks have the same security index of $\tilde{\gamma}_j^{a,d}$, they have the same probability of $\mathbb{P}(\mathbf{a}, \mathbf{d})$. In this paper, to compare the risk of attacks with the same security index, we "normalize" $\mathbb{P}(\mathbf{a}, \mathbf{d})$ to be 1, meaning that the attacks have been performed successfully. The following risk metric applies to the attacks with the same security index of $\tilde{\gamma}_j^{a,d}$,

$$R(\mathbf{a}, \mathbf{d}) = \mathbb{P}(\mathbf{a}, \mathbf{d})\mathbb{P}(s|\,\mathbf{a}, \mathbf{d})I(\mathbf{a}, \mathbf{d}) = \big(1 - \delta_{a,d}\big)I(\mathbf{a}, \mathbf{d}), \tag{22}$$

For the k_a-tuple FDI attacks with the same security index of $\tilde{\gamma}_j^{a}$, the formulation of risk metric is similar, i.e., $R(\mathbf{a}) = (1 - \delta_a)I(\mathbf{a})$ where δ_a is the detection probability from (18), $I(\mathbf{a})$ denotes the attack impact and $R(\mathbf{a})$ is the risk metric. Thus in the case of $\tilde{\gamma}_j^{a,d} = \tilde{\gamma}_j^{a}$, the risk of combined attacks and FDI attacks is comparable.

B. Attack Impact: Errors of Load Estimate

The estimated information from SE is used by further applications in EMS to compute optimal control actions. These are typically computed by minimizing network operation costs which are obtained by solving OPF algorithms. As the work in [13] and [37] shows, the OPF application uses the load estimate from SE as the inputs. In practice, the important outputs from EMS are the injection estimate and OPF results which would affect the further operations. If data attacks take place and pass the BDD, the load estimates get perturbed which influences the control actions. Therefore, we consider the impact metric as a function of the bias introduced by the attack on the load estimate.

Assuming that the actual injections are described in a vector $\mathbf{L}_{inj} \in \mathbb{R}^{n_{inj}}$ where n_{inj} is the number of buses with injections, we consider the impact on the errors of estimated power injections and actual power injections,

$$\epsilon = \hat{\mathbf{L}}_{inj,a,d} - \mathbf{L}_{inj}, \tag{23}$$

Algorithm 1 Risk Assessment for Combined Attacks

Step 1) Determine the attack magnitude μ. Compute attack vectors $\mathbf{a}$ and $\mathbf{d}$ from the optimization problem $\mathcal{P}_{a,d}(\tilde{\mathbf{H}})$.
Step 2) Solve (17) for obtaining the detection probability $\delta_{a,d}$ of the combined attack with $\mathbf{a}$ and $\mathbf{d}$.
Step 3) Calculate the attack impact metric $I(\mathbf{a}, \mathbf{d})$ according to Definition 4.
Step 4) Compute the risk metric $R(\mathbf{a}, \mathbf{d})$ for combined attack using the formulation of (22).

where $\hat{\mathbf{L}}_{inj,a,d} \in \mathbb{R}^{n_{inj}}$ is the vector of estimated injections under a (k_a, k_d)-tuple combined attack. Thus,

$$\epsilon = \mathbf{H}_{inj}\hat{\mathbf{x}}_{a,d} - \mathbf{H}_{inj}\mathbf{x}, \tag{24}$$

where $\hat{\mathbf{x}}_{a,d} = \mathbf{K}_d(\mathbf{z}_d + \mathbf{a}) = \mathbf{x} + \mathbf{K}_d\mathbf{e}_d + \mathbf{K}_d\mathbf{a}$, $\mathbf{H}_{inj} \in \mathbb{R}^{n_{inj}\times n}$ denotes the submatrix of $\mathbf{H}$ by keeping the rows corresponding to injections including loads. We can further obtain $\epsilon = \mathbf{H}_{inj}\mathbf{K}_d\mathbf{a} + \mathbf{H}_{inj}\mathbf{K}_d\mathbf{e}_d$ where the term introduced by the attacks is $\mathbf{H}_{inj}\mathbf{K}_d\mathbf{a}$. Here $\mathbf{K}_d$ is the function of the matrix $\mathbf{H}_d$ as defined in (10). The expected value of ϵ is

$$\mathbb{E}(\epsilon) = \mathbf{H}_{inj}\mathbf{K}_d\mathbf{a}. \tag{25}$$

We have the following definition of the attack impact metric for combined attacks.

Definition 4: The impact metric $I(\mathbf{a}, \mathbf{d})$ for quantifying attack impact of a combined attack with FDI attack vector $\mathbf{a}$ and availability vector $\mathbf{d}$ on load estimate is defined as the 2-norm of $\mathbf{H}_{inj}\mathbf{K}_d\mathbf{a}$, i.e., $I(\mathbf{a}, \mathbf{d}) \coloneqq \|\mathbf{H}_{inj}\mathbf{K}_d\mathbf{a}\|_2$.

Similar to the combined attacks, we define the attack impact metric $I(\mathbf{a}) = \|\mathbf{H}_{inj}\mathbf{K}\mathbf{a}\|_2$ for a k_a-tuple FDI attack with attack vector $\mathbf{a}$. We continue to adopt the linear attack policies to compute attack vectors for attacks with limited knowledge, i.e., $\mathbf{a} = \tilde{\mathbf{H}}_d\mathbf{c}$ for combined attacks and $\mathbf{a} = \tilde{\mathbf{H}}\mathbf{c}$ for FDI attacks.

Giving all the information above, the following Algorithm 1 summarizes the risk assessment procedure for combined attacks and FDI attacks. First, the system operators would solve $\mathcal{P}_a(\mathbf{H})$ and $\mathcal{P}_{a,d}(\mathbf{H})$ as a tool to compute the attack vectors from security index. Then the detection probability of attacks and the attack impact could be obtained respectively according to (17) and Definition 4, leading to the risk metric of (22). Thus in conclusion, the risk assessment presented in this paper, including the computation of attack vectors, the detection probability and the impact of attacks, provides insights at the planning stage of the grid and offline analysis of combined attacks in the limited knowledge case.

VI. Case Study

In this section we apply the analysis to the IEEE benchmark system (Figure 1). We conduct simulations on DC model for the purposes of: 1) illustrating vulnerability of SE to combined attacks; 2) providing insights into how combined attack can differ from FDI attack; 3) evaluating the risk of data attacks and giving the risk prioritization. In the performed experiments, measurements are placed on all the buses and

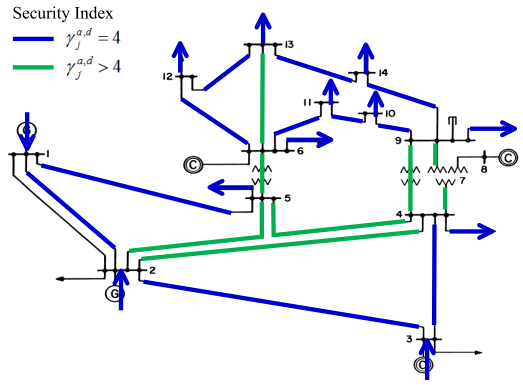

Fig. 1. The IEEE 14-bus system. The measurements are labeled different colors according to their security index $\gamma_j^{a,d}$ from Figure 2. Here the vulnerable measurements with small index ($=4$) are color coded blue. The measurements that have large index (>4) are color coded green. The pseudo-measurements (without color) on bus 7, 8 and line 7-8 can not be attacked.

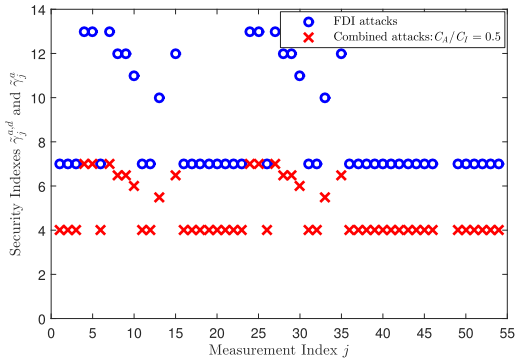

Fig. 2. The security index $\gamma_j^{a,d}$ under combined attacks and γ_j^{a} under FDI attacks are plotted versus the measurement index j. Here the cost of FDI attack on per measurement is assumed to be 1 and $C_A=0.5$ as $C_A/C_I=0.5$.

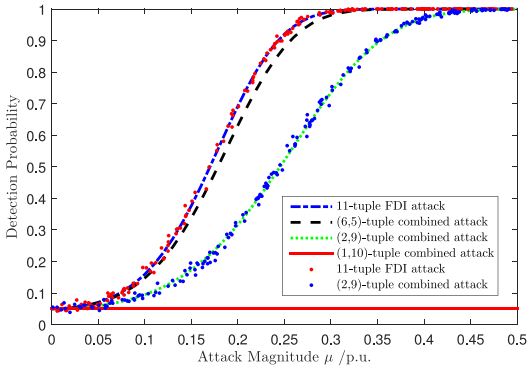

Fig. 3. The detection probability is plotted versus the attack magnitude. The attacks are all under structured uncertainty model (error on the model parameters of $\pm 20\%$) and performed in the same set of 11 measurements and the false alarm rate α is 0.05.

transmission lines to provide large redundancy. In the 14 bus system, measurements on bus 7, bus 8 and line 7-8 are pseudo-measurements for zero-injection buses and can not be attacked. The per-unit system is used and the power base is $100MW$. The measurements are generated under the DC model with Gaussian noise ($\sigma_j = 0.02$ for any measurement j). For the limited knowledge model, we assume the attacker knows the exact topology but has estimated line parameters with errors.

A. Security Index for Vulnerability Analysis

In order to expose vulnerability of SE to data attacks, we calculated the security index using the computation solutions of (13a) (according to Theorem 1) and (8a) for both combined attacks and FDI attacks. Thus the minimum number of compromised measurements and attack resources needed by the attacker to corrupt SE and pass the BDD are determined. Figure 2 shows the security indexes $\gamma_j^{a,d}$ and γ_j^{a} of combined attacks and FDI attacks in the IEEE 14 bus system. Here the cost of FDI attack on per measurement is assumed to be 1 ($C_I = 1$) and $C_A = 0.5$ as we take $C_A/C_I = 0.5$. The x-axis indicates the measurement targeted by the attacker to inject false data of $\mu = 0.1p.u.$ Note that in Figure 2 the pseudo-measurements 14, 34, 47, 48 from bus 7, 8 and line 7-8 can not be attacked and we keep their security indexes empty. The results illustrate the attack resources needed by the attacker to keep stealth. The security index of combined attacks is also showed in Figure 1 where the measurements are color coded to indicate which ones are more vulnerable. Combining Figure 2 and Figure 1, the security index can illustrate the vulnerable measurements in a power system.

The values of security index under combined attacks are smaller than the ones under FDI attacks when $C_A < C_I$ from Figure 2. For instance, in order to corrupt measurement $j = 10$, the FDI attack needs a value of 11 for attack resources (i.e., a 11-tuple FDI attack) while the combined attack only needs a value of 6 (i.e., a (1,10)-tuple combined attack). This implies that SE is more vulnerable to combined attacks with less attack resources. The results also show that $k_a = 1$ for the combined attacks and the optimal attack cost is $C_I + (\beta_j - 1)C_A$ for the case $C_A < C_I$, which is consistent with Proposition 1.

B. Detectability of Attacks With Limited Knowledge

Using the attack policy $\mathbf{a} = \tilde{\mathbf{H}}_d\mathbf{c}$ for combined attacks and $\mathbf{a} = \tilde{\mathbf{H}}\mathbf{c}$ for FDI attacks with the same given model uncertainty, the detection probability of attacks can be obtained according to (17) and (18). From Theorem 2 we see that the compromised measurements set from the optimal solutions of (14) w.r.t. $\tilde{\mathbf{H}}$ in (19) is in the same critical tuple with the one w.r.t. $\mathbf{H}$ in (1). Thus a set of 11 measurements (a critical tuple) containing measurement $j = 10$ needs to be compromised by the attacker from the security index in Figure 2. For the sake of comparison, the combined attacks and FDI attacks are performed in the same set of these 11 measurements. Figure 3 shows the detection probability of combined attacks and FDI attacks targeting these 11 measurements, with the structured model uncertainty (error on the line parameters of $\pm 20\%$). In addition to the theoretical results, the empirical detection probability results are also presented in Figure 3 for the 11-tuple FDI attack and (2,9)-tuple combined attack respectively. Figure 4 shows the detection probability of combined attacks and FDI attacks with different levels of model uncertainty (error on line parameters of $\pm 10\%$, $\pm 20\%$, $\pm 30\%$, $\pm 40\%$).

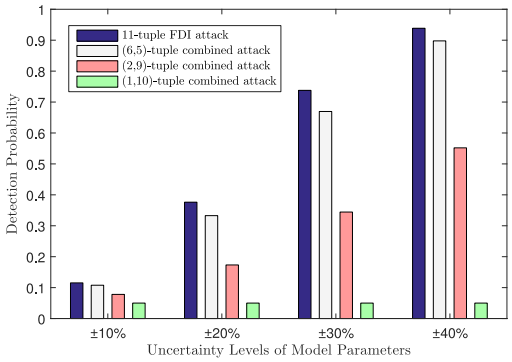

Fig. 4. The detection probability is plotted versus different levels of model uncertainty (error on the model parameters of ±10%, ±20%, ±30%, ±40%, respectively). The combined attacks and FDI attacks are performed in the same set of 11 measurements and the attack magnitudes are all $\mu = 0.15p.u.$ here. The false alarm rate α is 0.05.

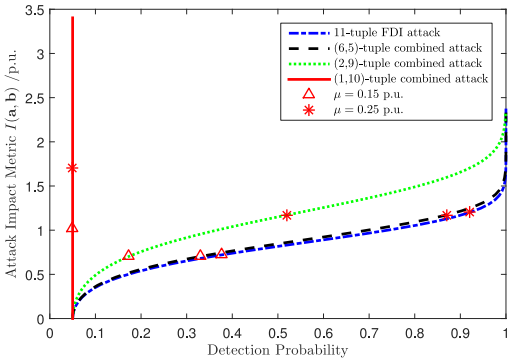

Fig. 5. The attack impact metric is plotted versus the detection probability. The attacks are all under structured uncertainty model (error on model parameters of ±20%) and performed in the same set of 11 measurements. Here we assume $C_A = C_I = 1$ and the false alarm rate α is 0.05.

To obtain the empirical detection probability in Figure 3, we use Monte Carlo simulations. Taking the (2,9)-tuple combined attack as an example, 200 different points of attack magnitude μ were taken in random from 0 to 0.5 p.u. and the corresponding attack vectors were built. For each attack vector with the taken magnitude μ, total 1000 Monte Carlo runs were executed to obtain the detection probability of such attack. In each Monte Carlo simulation, the measurements were created by the DC model with Gaussian noise and the attack vector was added to the measurements. For the attacked measurements, the SE and BDD with the false alarm rate 0.05 were executed.

From Figure 3 we can see that the empirical results of detection probability follow the theoretical one. This proves that using the approximation of the distribution of $J_{a,d}(\hat{\mathbf{x}})$ and $J_a(\hat{\mathbf{x}})$ can provide the detection probability, and it is reliable to use theoretical detection probability for risk analysis in the following. The results in Figure 3 illustrate that combined attacks can have lower detection probability comparing with FDI attacks, meaning that SE is more vulnerable to combined attacks as they have higher probability not to be discovered by the BDD. An interesting result is that with smaller k_a the combined attack also has lower probability to be detected. In the case that $k_a = 1$ and $k_d = 10$, the (1,10)-tuple combined attack can keep stealth, which is consistent with Proposition 2. The results in Figure 4 show that, for the combined or FDI attacks with different levels of model uncertainty, the detection probability of attacks would increase when the attacker has an bigger error on the transmission line parameters. This can be expected as the attacker has less knowledge to build attack vectors. Besides, combined attacks still have advantages in keeping stealth as they can have lower detection probability especially the combined attacks with smaller k_a, and the undetectability of the (1,10)-tuple combined attack is independent of parameter uncertainty as discussed in Proposition 2.

C. Risk Metrics for Attacks

We continue with the risk analysis of combined attacks. Simulations were conducted on the same scenarios as Section VI-B where the attacker manipulates the set of 11 measurements (a critical tuple). We analyze the attack impact and present the risk of the combined attacks and FDI attacks. For the risk analysis, we take the attack cost values that satisfy $C_A = C_I = 1$, thus the security indexes $\tilde{\gamma}_j^{a,d}$ and $\tilde{\gamma}_j^{a}$ w.r.t. $\tilde{\mathbf{H}}$ in (19) of these attacks are equal to each other and the probability $P(\mathbf{a}, \mathbf{d})$ can be "normalized" as discussed in Section IV-B. First, for the attacks with specific model uncertainty (error on the transmission line parameters of ±20%), the results for attack impact metrics versus detection probability are given in Figure 5, and the values of risk metrics for combined attacks and FDI attacks versus attack magnitude are shown in Figure 6. Second, we also show the risk metric values of combined attacks and FDI attacks with different levels of model uncertainty (error on line parameters of ±10%, ±20%, ±30%, ±40%) in Figure 7.

Under the perturbed model with uncertainty, the attacker has the possibility to be detected by the BDD while introducing errors on load estimate. From Figure 5, we see that combined attacks can have similar attack impact metrics with FDI attacks but lower detection probability with the same attack magnitude μ (0.15 p.u. or 0.25 p.u. as shown in Figure 5). Especially the (1,10)-tuple combined attack has larger impact metrics than attacks with limited knowledge for the both cases that attack magnitude $\mu = 0.15p.u.$ or $\mu = 0.25p.u.$

For the *risk* metrics in Figure 6, when the attack magnitude μ increases, the risk metric increases due to the low detection probability. After μ reaches certain values, the risk metric decreases since the attacks can be discovered with high probability. It's also shown that combined attacks can have larger risk metrics especially the cases of (1,10)-tuple and (2,9)-tuple combined attacks. It should be noted that though we assume $C_A = C_I$ to obtain the risk metrics, the risk prioritization of these attacks in Figure 6 would not change if $C_A < C_I$ is assumed. This is because the combined attacks can be launched with less attack resources when $C_A < C_I$, resulting in larger risk values comparing with FDI attacks. Figure 7 illustrates that with bigger errors on the model parameters, the risk metrics would decrease for most cases of attacks, meaning that the system faces less risk when the attacker has large

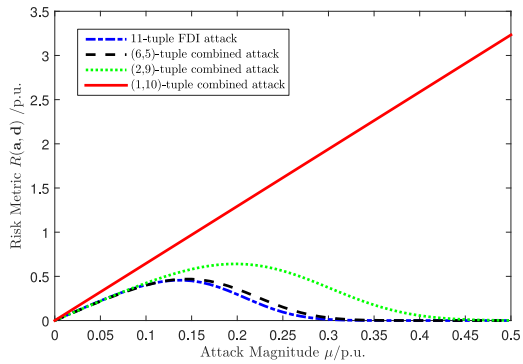

Fig. 6. The risk metric is plotted versus the attack magnitude. The attacks are all under structured uncertainty model (error on model parameters of ±20%) and performed in the same set of 11 measurements. Here we assume $C_A = C_I = 1$ and the false alarm rate α is 0.05.

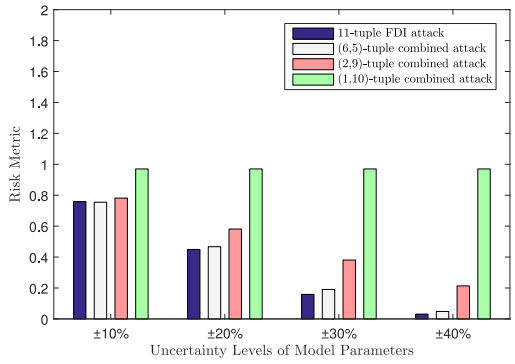

Fig. 7. The risk metric is plotted versus different levels of model uncertainty (error on the model parameters of ±10%, ±20%, ±30%, ±40%, respectively). The attacks are performed in the same set of 11 measurements and the attack magnitudes are all $\mu = 0.15p.u.$. Here we assume $C_A = C_I = 1$ and the false alarm rate α is 0.05.

model uncertainty in building attack vectors. From Figure 7 we can see, combined attacks with smaller k_a would bring more risk to the system under each level of model uncertainty and the (1,10)-tuple combined attack has the largest risk metric independently of model uncertainty. This is due to the fact that such kind of attack can always succeed in keeping stealth even with limited knowledge of the system model.

D. Discussion

1) Computation Efficiency: In this paper we use the big M method to express the security index problem as a MILP. To show the computation time of this method, we calculated security index for the IEEE 14 bus, 39 bus and 118 bus systems, all of which are with full measurements for the sake of comparison. Note that the big M method does not need the full measurements assumption. The computation time for the four IEEE benchmarks is listed in Table I. The computation was performed on a PC with 3.5 GHz CPU and 8 GB of RAM. The MILP problems were solved using the CPLEX for MATLAB where the execution time of the algorithm for calculating all the security indexes of each IEEE benchmark was recorded.

Table I shows that when the system becomes larger, the computation time increases. The MILP formulation imposes challenges for computation for large-scale power systems. However, this method could be used off-line in the assessment of the system vulnerability. Faster computation time can be achieved on the expense of accuracy using relaxations (such as 1-norm relaxation providing an overestimate of the security index [8]) or some assumptions (such as the full measurements assumption used in the min-cut algorithm [28], [30]).

TABLE I
COMPUTATION TIME OF SECURITY INDEX FOR THE IEEE BENCHMARKS

	14 bus	39 bus	118 bus
Time	4.2s	25.6s	117s

2) Existence of a Detector for Availability Attack: It should be noted that our previous results assume that the SE treats the availability attacks as missing data and no additional alerts are triggered. Although the typical BDD schemes fail to detect availability attacks, a new detector could be designed for combined attacks.

Here we propose an initial missing data detection (MDD) scheme. We assume that, under normal conditions each measurement may be missing with a given small probability. In particular, we say that the i-th measurement is missing if $\mathbf{u}_{(i)} = 1$, where $\mathbf{u}_{(i)} \in \mathbb{B}$ is a Bernoulli distributed random variable with $\mathbb{P}(\mathbf{u}_{(i)} = 1) = p_i$. The Bernoulli distributed random variables $u_{(i)i=1,\ldots,m}$ are assumed to be independent and identically distributed, with $p_i = p_0$ for all $i = 1, \ldots, m$. The missing data due to abnormal conditions can be detected based on the random variable $\mathbf{u} \in \mathbb{B}^m$. Parameterizing $\mathbf{u}_{(i)}$ as $\mathbf{u}_{(i)} \sim \mathcal{B}(p)$, we are interested in testing the hypothesis $\mathcal{H}_1$ with a null hypothesis $\mathcal{H}_0$. If $\mathcal{H}_0$ is accepted, that means there is no availability attack and alternatively availability attack exists:

- $\mathcal{H}_0$: $p \leq p_0$;
- $\mathcal{H}_1$: $p > p_0$.

In other words, we are interested in differentiating between cases of low probability of missing data, versus cases where missing data occurs with higher probability. Defining the auxiliary statistic $r_u \triangleq \sum_{i=1}^{m} \mathbf{u}_{(i)} = \mathbf{1}^\top \mathbf{u}$ which corresponds to the number of missing measurements, we know that r_u follows a binomial distribution, namely $r_u \sim B(m, p)$ with the likelihood function $\mathcal{L}(p; u) \triangleq \frac{m!}{(m-r_u)!r_u!}(1-p)^{m-r_u}p^{r_u}$. Thus the statistical test for rejection $\mathcal{H}_0$ is

$$r_u > \bar{\tau}_u,$$

where $\bar{\tau}_u$ is computed to bound the probability of false-alarm of the statistical test.

Recall the current BDD scheme in SE described in Section II-B. If the above MDD scheme is implemented in SE together with the BDD, we can obtain the detection probability of combined attacks. Note that the random variables $\mathbf{r}_{a,d}$ in (16) and r_u are not independent since the unavailable measurements will influence the degrees of freedom and the covariance matrix of the residual vector $\mathbf{r}_{a,d}$. Thus it's difficult to express the whole detection probability of combined attack under these two detectors mathematically. We use Monte Carlo simulations instead. For each taken attack magnitude, the given combined attack was implemented through 1000 Monte Carlo

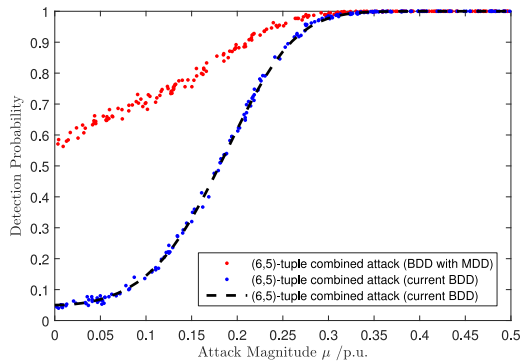

Fig. 8. The detection probability is plotted versus the attack magnitude. The same (6,5)-tuple combined attack from Figure 3 is tested with both cases: one with BDD and MDD, the other one only with BDD (without MDD). For MDD test, p_0 is assumed to be 0.06.

runs while in each run the measurements were generated with random errors. If this combined attack triggered any alert on these two detectors, we say it was detected. Here we provide Figure 8 to show the detection probability of (6,5)-tuple combined attack (from Figure 3) when the proposed MDD is equipped with the typical BDD. The results show that the MDD could help in detecting the combined attacks.

3) AC Power Flows: In this paper for the first time we look at combined attacks under limited knowledge and conduct risk analysis on these attacks. Here we are focusing on establishing the concept of risk of the combined attacks and explore this concept in the DC state estimation at the EMS of control. We hope this can be a stepping stone towards addressing risk of combined attacks in the AC power flows model.

The combined attacks explored in this paper would naturally be more complex to compute under the AC model. In the case of AC state estimation, an attacker would need to have a better knowledge of the system and its operating state. The detection probability of the combined attack constructed based on the DC model will be higher and the risk of a successful attack will be lower. Thus, the results of this paper cannot be directly extrapolated to the case with AC state estimation. However, we believe that the proposed formulation can be used to explore the AC case by replacing the DC model $\mathcal{H}$ with a linearization of the AC nonlinear power flow model at a given system state of interest.

VII. Conclusion

In this paper we see that combined attacks can succeed with less resources (if $C_A < C_I$) and lower detection probability when the adversarial knowledge is limited, bringing more risk to reliable system operation. It also should be noted that this paper assumes that the SE treats unavailable measurements due to attacks as a case of missing data, although the amount of missing data under attacks is larger than the one under normal conditions. In the discussion we also showed the potentiality of designing a detector for availability attacks. Besides, availability attacks like DoS attacks could trigger alerts on ICT-specific measures (e.g., Intrusion Detection System). These two features give the opportunities to develop better cross-domain detection schemes for availability portion of the attacks improving the overall combined attacks detection. Other research directions to explore in the future include evaluating physical impact of combined attacks and exploring the vulnerability of AC state estimation to combined attacks.

References

[1] A. Giani, S. Sastry, K. H. Johansson, and H. Sandberg, "The VIKING project: An initiative on resilient control of power networks," in *Proc. 2nd Int. Symp. Resilient Control Syst.*, Idaho Falls, ID, USA, 2009, pp. 31–35.

[2] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Security*, Chicago, IL, USA, 2009, pp. 21–32.

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

[4] D. Deka, R. Baldick, and S. Vishwanath, "Optimal data attacks on power grids: Leveraging detection & measurement jamming," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Miami, FL, USA, Nov. 2015, pp. 392–397.

[5] R. S. Ross, "Guide for conducting risk assessments," NIST, Gaithersburg, MD, USA, Rep. SP 800-30 Rev. 1, Sep. 2012.

[6] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362–1370, Sep. 2012.

[7] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *Proc. 1st Workshop Secure Control Syst. (SCS)*, Stockholm, Sweden, 2010, pp. 1–6.

[8] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, "Secure control systems: A quantitative risk management approach," *IEEE Control Syst.*, vol. 35, no. 1, pp. 24–45, Feb. 2015.

[9] A. Teixeira, G. Dán, H. Sandberg, and K. H. Johansson, "A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator," in *Proc. IFAC World Congr.*, Jan. 2011, pp. 11271–11277.

[10] 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.

[11] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, Dec. 2011.

[12] L. Jia, J. Kim, R. J. Thomas, and L. Tong, "Impact of data quality on real-time locational marginal price," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 627–636, Mar. 2014.

[13] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3864–3872, Sep. 2016.

[14] S. Li, Y. Yılmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov. 2015.

[15] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Trans. Smart Grid*, to be published.

[16] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, "Network-aware mitigation of data integrity attacks on power system state estimation," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 6, pp. 1108–1118, Jul. 2012.

[17] K. Pan, A. M. H. Teixeira, M. Cvetkovic, and P. Palensky, "Combined data integrity and availability attacks on state estimation in cyber-physical power grids," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Sydney, NSW, Australia, Nov. 2016, pp. 271–277.

[18] X. Liu and Z. Li, "Local load redistribution attacks in power systems with incomplete network information," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1665–1676, Jul. 2014.

[19] X. Liu, Z. Bao, D. Lu, and Z. Li, "Modeling of local false data injection attacks with reduced network information," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1686–1696, Jul. 2015.

[20] X. Liu and Z. Li, "False data attacks against AC state estimation with incomplete network information," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2239–2248, Sep. 2017.

[21] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Anaheim, CA, USA, 2012, pp. 3153–3158.

[22] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Proc. 49th IEEE Conf. CDC*, Atlanta, GA, USA, Dec. 2010, pp. 5991–5998.
[23] M. Esmalifalak, H. Nguyen, R. Zheng, and Z. Han, "Stealth false data injection using independent component analysis in smart grid," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Oct. 2011, pp. 244–248.
[24] J. Kim, L. Tong, and R. J. Thomas, "Data framing attack on state estimation with unknown network parameters," in *Proc. Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 1388–1392.
[25] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.
[26] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL, USA: CRC Press, 2004.
[27] D. A. Jones, "Statistical analysis of empirical models fitted by optimization," *Biometrika*, vol. 70, no. 1, pp. 67–88, 1983.
[28] J. M. Hendrickx, K. H. Johansson, R. M. Jungers, H. Sandberg, and K. C. Sou, "Efficient computations of a security index for false data attacks in power networks," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3194–3208, Dec. 2014.
[29] J. D. Markovic-Petrovic and M. D. Stojanovic, "Analysis of SCADA system vulnerabilities to DDoS attacks," in *Proc. IEEE 11th Int. Conf. TELSIKS*, vol. 2. Niš, Serbia, 2013, pp. 591–594.
[30] K. C. Sou, H. Sandberg, and K. H. Johansson, "On the exact solution to a smart grid cyber-security analysis problem," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 856–865, Jun. 2013.
[31] B. Kang, "Deliverable D4.1 high-level design documentation and deployment architecture for multi-attribute SCADA intrusion detection system," Austrian Inst. Technol., Kepler Universität Linz, Linz, Austria, Rep. 608224, 2015.
[32] M. Hutle, G. Hansch, and W. Fitzgerald, "D2.2 threat and risk assessment methodology," *Tunneling Underground Space Technol.*, vol. 24, no. 3, pp. 269–277, Sep. 2015.
[33] K. Pan, A. Teixeira, M. Cvetkovic, and P. Palensky. (2017). *Data Attacks on Power System State Estimation: Limited Adversarial Knowledge VS. Limited Attack Resources*. [Online]. Available: http://arxiv.org/pdf/1708.08355v1
[34] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.
[35] G. R. Krumpholz, K. A. Clements, and P. W. Davis, "Power system observability: A practical algorithm using network topology," *IEEE Trans. Power App. Syst.*, vol. PAS-99, no. 4, pp. 1534–1542, Jul. 1980.
[36] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proc. IEEE*, vol. 105, no. 7, pp. 1389–1407, Jul. 2017.
[37] A. Teixeira, H. Sandberg, G. Dan, and K. H. Johansson, "Optimal power flow: Closing the loop over corrupted data," in *Proc. Amer. Control Conf. (ACC)*, Montreal, QC, Canada, Jun. 2012, pp. 3534–3540.

Kaikai Pan (S'16) received the B.Eng. and M.Eng. degrees in measuring and control from Beihang University, Beijing, China, in 2012 and 2015, respectively. He is a Doctoral Researcher working toward risk assessment and mitigation of cyber attacks in smart grids with the Electrical Sustainable Energy Department, Delft University of Technology, Delft, The Netherlands. His current research interests include cyber security of intelligent power grids, risk assessment of data attacks, attack detection, cyber-physical energy systems, and co-simulation techniques.

André Teixeira (M'15) received the M.Sc. degree in electrical and computer engineering from the Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, in 2009 and the Ph.D. degree in automatic control from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2014. He is an Associate Senior Lecturer with the Division of Signals and Systems, Department of Engineering Sciences, Uppsala University, Sweden. From 2014 to 2015, he was a Post-Doctoral Researcher with the Department of Automatic Control, KTH Royal Institute of Technology. From 2015 to 2017, he was an Assistant Professor with the Department of Engineering Systems and Services, Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands. His current research interests include cyber-secure and resilient control systems, distributed fault detection and isolation, and energy systems.

Milos Cvetkovic (M'15) received the B.Sc. degree in electrical engineering from the University of Belgrade, Serbia, in 2008, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2011 and 2013, respectively. He is an Assistant Professor with the Electrical Sustainable Energy Department, Delft University of Technology, Delft, The Netherlands. His research interests include development of co-simulations for energy grids and modeling for control and optimization of the electricity grids.

Peter Palensky (M'03–SM'05) was a Principal Scientist for Complex Energy Systems with the Energy Department, Austrian Institute of Technology, Austria, where he was the Head of Business Unit "Sustainable Building Technologies," a CTO with Envidatec Corporation, Hamburg, Germany, an Associate Professor with the Department of Electrical, Electronic, and Computer Engineering, University of Pretoria, South Africa, a University Assistant with the Vienna University of Technology, Austria, and a Researcher with the Lawrence Berkeley National Laboratory, California. He is a Professor for intelligent electric power grids with TU Delft, The Netherlands. His main research field is complex energy systems. He is active in international committees like IEEE and an Associate Editor of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.