
Article:
Ye, X., Esnaola, I., Perlaza, S.M. et al. (2024) An information theoretic metric for measurement vulnerability to data integrity attacks on smart grids. IET Smart Grid, 7 (5). pp. 583-592. ISSN: 2515-2947

https://doi.org/10.1049/stg2.12163

Received: 4 November 2022 | Revised: 31 July 2023 | Accepted: 20 February 2024

IET Smart Grid

DOI: 10.1049/stg2.12163

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

An information theoretic metric for measurement vulnerability to data integrity attacks on smart grids

Xiuzhen Ye[1] | Iñaki Esnaola[1,2] | Samir M. Perlaza[2,3] | Robert F. Harrison[1]

[1]Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK

[2]Department of Electrical Engineering, Princeton University, Princeton, New Jersey, USA

[3]INRIA, Sophia Antipolis, Sophia Antipolis, France

Correspondence

Xiuzhen Ye.
Email: Xye15@sheffield.ac.uk

Funding information

The China Scholarship Council, Grant/Award Number: 201906150124; The European Commission through the H2020-MSCA-RISE-2019 program, Grant/Award Number: 872172

Abstract

A novel metric that describes the vulnerability of the measurements in power systems to data integrity attacks is proposed. The new metric, coined vulnerability index (VuIx), leverages information theoretic measures to assess the attack effect in terms of the fundamental limits of the disruption and detection tradeoff. The result of computing the VuIx of the measurements in the system yields an ordering of their vulnerability based on the degree of exposure to data integrity attacks. This new framework is used to assess the measurement vulnerability of IEEE 9-bus and 30-bus test systems and it is observed that power injection measurements are significantly more vulnerable to data integrity attacks than power flow measurements. A detailed numerical evaluation of the VuIx values for IEEE test systems is provided.

KEYWORDS

power system cyber-security and privacy, power system security

1 | INTRODUCTION

Supervisory Control and Data Acquisition (SCADA) systems and more recently advanced communication systems facilitate efficient, economic and reliable operation of power systems [1]. For instance, the communication system transmits the measurements to a state estimator that evaluates the operational status of the system accurately [2]. However, the integration between the physical layer and the cyber layer exposes the system to cybersecurity threats. Cyber incidents highlight the vulnerability of power systems to sophisticated attacks. To ensure the security and reliability of power system operation, it is essential to quantitatively characterise the vulnerabilities of the system in order to set up appropriate security mechanisms [3]. To that end, security metrics provide operationally meaningful vulnerability descriptors and identify the impact that security threats pose to the system. Moreover, security metrics enable operators to assess the defence mechanisms requirements to be embedded into cybersecurity policies, processes, and technology [4]. For example, the Common Vulnerability Scoring System (CVSS) is one of the typical systems that provides security metrics [5]. Typical security metrics for power systems focus on integrity, availability, and confidentiality as envisioned by the cybersecurity working group in the NIST Smart Grid interoperability panel [6]. System security objectives are categorised into system vulnerability, defence power, attack severity, and situations to develop security metrics in a systematic manner [7]. A cyberphysical security assessment metric (CP-SAM) based on quantitative factors is proposed to assess the specific security challenges of microgrid systems in ref. [8].

This fragmented landscape showcases a wide variety of metrics available that depend on the security services, threat characteristics, and system parameters. Remarkably, there is a lack of general data integrity vulnerability metrics for power systems. For instance, the impact of data injection attacks (DIAs) [9] can be assessed with a wide variety of criteria that depend on the objectives of the attackers [10–13]. A large body of literature addresses DIAs that compromise both the confidentiality and integrity of the information contained by the system measurements [14]. With the unprecedented data acquisition capabilities available in cyberphysical systems, attackers can learn the statistical structure of the system and incorporate the underlying stochastic process to launch the attacks [15, 16]. DIAs that

operate within a Bayesian framework by leveraging stochastic models of the system are studied in refs. [17, 18]. From the perspective of the operator, the introduction of stochastic descriptors opens the door to information theoretic quantification of the measurement vulnerability.

In this paper, we propose a novel information theoretic metric to assess the vulnerability of measurements in power systems to data integrity attacks. Specifically, we characterise the fundamental information loss induced by data integrity attacks via mutual information and the stealthiness of the attack via Kullback–Leibler divergence. Our aim is to provide a metric that is grounded on fundamental principles, and therefore, informs the vulnerabilities of the measurements in the system to a wide range of threats. This is enabled by the use of information theoretic measures which characterise the amount of information acquired by the measurements in the system in fundamental terms.

The rest of the paper is organised as follows: In Section 2, we introduce a Bayesian framework with linearised dynamics for DIAs. Information theoretic attacks are presented in Section 3. The vulnerability metric on information theoretic attacks is proposed in Section 4. In Section 5, we characterise the vulnerability of measurements in uncompromised systems and propose an algorithm to evaluate the vulnerability of measurements. The vulnerability of measurements of the IEEE test systems is presented in Section 6. The paper concludes in Section 7.

The main contributions of this paper is as follows: (1) A notion of vulnerability for the measurements in the system is proposed. The proposed notion is characterised by the information theoretic cost induced by random attacks. Specifically, mutual information and KL divergence are used to construct a quantitative measure of vulnerability. (2) The vulnerability assessment of the measurements is posed as a minimisation problem and closed-form expressions are obtained for the case in which the initial state of the system is uncompromised. (3) An algorithm that computes the proposed vulnerability indices for general state estimators in power systems is proposed. (4) The proposed framework is numerically evaluated in IEEE 9-bus and 30-bus test systems to obtain qualitative characterisations of the vulnerability of the measurements in the systems.

Notation We denote the number of state variables on a given system by n and the number of the measurements by m. The set of positive semidefinite matrices of size $n \times n$ is denoted by S_+^n. The n-dimensional identity matrix is denoted as $\mathbf{I}_n$. For a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we denote by $(\mathbf{A})_{ij}$ the entry in row i and column j and diag($\mathbf{A}$) denotes the vector formed by the diagonal entries of $\mathbf{A}$. The elementary vector $\mathbf{e}_i \in \mathbb{R}^n$ is a vector of zeros with a one in the ith entry. Random variables are denoted by capital letters and their realisations by the corresponding lower case, for example, x is a realisation of the random variable X. Vectors of n random variables are denoted by a superscript, for example, $X^n = (X_1, \ldots, X_n)^\mathsf{T}$ with corresponding realisations denoted by $\mathbf{x}$. Given an n-dimensional vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and a matrix $\boldsymbol{\Sigma} \in S_+^n$, we denote by $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ the multivariate Gaussian distribution of dimension n with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The mutual information between random variables X and Y is denoted by $I(X; Y)$ and the Kullback-Leibler (KL) divergence between the distributions P and Q is denoted by $D(P\|Q)$.

2 | SYSTEM MODEL

2.1 | Observation model

In a power system, the state vector $\mathbf{x} \in \mathbb{R}^n$ that contains the voltages and phase angles at all the buses describes the operational state of the system. State vector $\mathbf{x}$ is observed by the acquisition function $F : \mathbb{R}^n \to \mathbb{R}^m$. A linearised observation model is considered for state estimation, which yields the observation model

$$Y^m = \mathbf{H}\mathbf{x} + Z^m, \tag{1}$$

where $\mathbf{H} \in \mathbb{R}^{m\times n}$ is the Jacobian of the function F at a given operating point and is determined by the parameters and topology of the system. The vector of measurements Y^m is corrupted by additive white Gaussian noise introduced by the sensors [1, 2]. The noise vector Z^m follows a multivariate Gaussian distribution, that is,

$$Z^m \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}_m\right), \tag{2}$$

where σ^2 is the noise variance.

In a Bayesian estimation framework, the state variables are described by a vector of random variables X^n with a given distribution. In this study, we assume X^n follows a multivariable Gaussian distribution [19] with zero mean and covariance matrix $\boldsymbol{\Sigma}_{XX} \in S_+^n$, that is,

$$X^n \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{XX}). \tag{3}$$

From Equation (1), it follows that the vector of measurements is with zero mean and covariance matrix $\boldsymbol{\Sigma}_{YY} \in S_+^m$, that is,

$$Y^m \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{YY}), \tag{4}$$

where

$$\boldsymbol{\Sigma}_{YY} \triangleq \mathbf{H}\boldsymbol{\Sigma}_{XX}\mathbf{H}^\mathsf{T} + \sigma^2\mathbf{I}_m. \tag{5}$$

2.2 | Attack setting

Let us denote the measurements corrupted by the malicious attack given by the random vector A^m taking values in $\mathbb{R}^m$, that is,

$$Y_A^m = \mathbf{H}X^n + Z^m + A^m, \tag{6}$$

where $Y_A^m \in \mathbb{R}^m$ random vector of measurements. With a fixed covariance matrix $\mathbf{\Sigma}_{AA} \sim S_{++}^m$, when the additive disturbance to the system, that is, $Z^m + A^m$ follows a multivariate Gaussian distribution, the mutual information between the state variables X^n and the compromised measurements Y_A^m denoted by $I(X^n; Y_A^m)$ is minimised [20]. Hence, from the Lévy-Cramér decomposition theorem [21, 22], it holds that the sum $Z^m + A^m$ is Gaussian, given that Z^m satisfies Equation (2), and therefore, A^m is Gaussian. In view of this, in the following, we assume that

$$A^m \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{AA}), \tag{7}$$

where $\mathbf{0} = (0, 0, \ldots, 0)$ and $\mathbf{\Sigma}_{AA} \in \mathcal{S}_+^m$ are the mean vector and the covariance matrix of the random attack vector A^m. The assumption in Equation (7) is further discussed in Section 3. Consequently, the vector of compromised measurements Y_A^m follows a multivariate Gaussian distribution with zero mean and covariance matrix $\mathbf{\Sigma}_{Y_AY_A} \in \mathcal{S}_+^m$, that is,

$$Y_A^m \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{Y_AY_A}), \tag{8}$$

with

$$\mathbf{\Sigma}_{Y_AY_A} \triangleq \mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^\top + \sigma^2\mathbf{I}_m + \mathbf{\Sigma}_{AA}. \tag{9}$$

3 | INFORMATION THEORETIC ATTACKS

The aim of the attack is twofold. Firstly, the attack aims to disrupt the state estimation procedure. Secondly, it aims to stay undetected. For the first objective, we minimise the mutual information between the vector of state variables X^n in Equation (3) and the vector of compromised measurements Y_A^m in Equation (6), that is, $I(X^n; Y_A^m)$. In other words, the attack yields less information about the state variables contained by the compromised measurements. The stealth constraint in the second objective is captured by the Kullback–Leibler (KL) divergence between the distribution $P_{Y_A^m}$ in Equation (6) and the distribution P_{Y^m} in Equation (1), that is, $D\left(P_{Y_A^m} \| P_{Y^m}\right)$. For the observation model and attack setting described in Section 2, and assuming optimal detection, the Chernoff-Stein Lemma [23] states that the minimisation of KL divergence leads to the minimisation of the asymptotic detection probability.

The following propositions characterise mutual information and KL divergence with Gaussian state variables and attacks, respectively [24], Prop. 1, 2.

Proposition 1 *The mutual information between the random vectors X^n in Equation (3) and Y_A^m in Equation (8) is*

$$I(X^n; Y_A^m) = \frac{1}{2}\log\frac{|\mathbf{\Sigma}_{XX}||\mathbf{\Sigma}_{Y_AY_A}|}{|\mathbf{\Sigma}|}, \tag{10}$$

where the matrices $\mathbf{\Sigma}_{X\,X}$ and $\mathbf{\Sigma}_{Y_AY_A}$ are in Equations (3) and (9), respectively; and the matrix $\mathbf{\Sigma}$ is the covariance matrix of the joint distribution of X^n and Y_A^m, that is, $(X^n; Y_A^m) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ with

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{XX} & \mathbf{\Sigma}_{XX}\mathbf{H}^\top \\ \mathbf{H}\mathbf{\Sigma}_{XX} & \mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^\top + \sigma^2\mathbf{I}_m + \mathbf{\Sigma}_{AA} \end{pmatrix}, \tag{11}$$

where $\sigma \in \mathbb{R}_+$ is in Equation (2); and matrices $\mathbf{H}$ and $\mathbf{\Sigma}_{AA}$ are in Equations (1) and (7), respectively.

Proposition 2 *The KL divergence between the distribution of random vector Y_A^m in Equation (8) and the distribution of random vector Y^m in Equation (4) is as follows:*

$$D\left(P_{Y_A^m} \| P_{Y^m}\right) = \frac{1}{2}\left(\log\frac{|\mathbf{\Sigma}_{YY}|}{|\mathbf{\Sigma}_{Y_AY_A}|} - m + \mathrm{tr}\left(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{Y_AY_A}\right)\right), \tag{12}$$

where the matrices $\mathbf{\Sigma}_{YY}$ and $\mathbf{\Sigma}_{AA}$ are in Equations (5) and (7), respectively.

The information theoretic attack construction is proposed in the following optimisation problem [17, 24]:

$$\min_{P_{A^m}} I(X^n; Y_A^m) + \lambda D\left(P_{Y_A^m} \| P_{Y^m}\right), \tag{13}$$

where $\lambda \in \mathbb{R}_+$ is the weighting parameter that determines the tradeoff between mutual information and KL divergence. Note that the optimisation domain in Equation (13) is the set of m-dimensional Gaussian multivariate distributions. The optimal Gaussian attack for $\lambda \geq 1$ as a solution to Equation (13) is given by the following [24]:

$$A^m \sim \mathcal{N}\left(\mathbf{0}, \lambda^{-1/2}\mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^\top\right). \tag{14}$$

Note that the attack realisations from Equation (14) are non-zero with probability one, that is, $\mathbb{P}[|\mathrm{supp}(A^m)| = m] = 1$, where

$$\mathrm{supp}(A^m) \triangleq \{i : \mathbb{P}[A_i = 0] = 0\}. \tag{15}$$

The attack implementation requires access to the sensing infrastructure of the industrial control system (ICS) operating the power systems. For that reason, the attack construction incorporates a sparsity constraint that limits the optimisation domain over the attack vector A^m in Equation (6) to the distributions with cardinality of the support satisfying $|\mathrm{supp}(A^m)| = k \leq m$, that is,

$$\mathcal{P}_k \triangleq \bigcup_{i=1}^{k} \{A^m \sim \mathcal{N}(\mathbf{0}, \overline{\mathbf{\Sigma}}) : |\mathrm{supp}(A^m)| = i\}. \tag{16}$$

The resulting sparse attack construction is [18]

$$\min_{\mathcal{P}_k} I\left(X^n; Y_A^m\right) + \lambda D\left(P_{Y_A^m} \| P_{Y^m}\right). \tag{17}$$

The following theorem provides the optimal single sensor attack construction.

Theorem 1 *[17], Th. 1 The solution to the sparse stealth attack construction problem in Equation (17) for the case $k = 1$ is*

$$\overline{\mathbf{\Sigma}}^* = v\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}, \tag{18}$$

where

$$i = \underset{j\in\{1,2,\ldots,m\}}{\arg\min} \left\{\left(\mathbf{\Sigma}_{YY}^{-1}\right)_{jj}\right\}, \tag{19}$$

$$v = -\frac{\sigma^2}{2} + \frac{1}{2}\left(\sigma^4 - \frac{4\left(\underline{w}\,\sigma^2 - 1\right)}{\lambda \underline{w}^2}\right)^{\frac{1}{2}}, \tag{20}$$

with $\underline{w} \triangleq \left(\mathbf{\Sigma}_{YY}^{-1}\right)_{ii}$.

4 | VULNERABILITY METRIC FOR INFORMATION THEORETIC ATTACKS

4.1 | Attack structure with sequential measurement selection

To assess the impact of the attacks to different measurements, we model the entries of the random attack vector A^m as independent, that is,

$$P_{A^m} = \prod_{i=1}^{m} P_{A_i}, \tag{21}$$

where A_i is the ith entry of A^m and for all $i \in \{1, 2, \ldots, m\}$, the distribution P_{A_i} is Gaussian with zero mean and variance $v \in \mathbb{R}_+$, that is, $A_i \sim \mathcal{N}(0, v)$. Consider that k sensors have been attacked with $k \in \{0, 1, 2, \ldots, m-1\}$ and let the covariance matrix of the corresponding attack vector A^m in Equation (6) be

$$\mathbf{\Sigma} \in \mathcal{S}_k, \tag{22}$$

where $\mathcal{S}_k$ is the set of m-dimensional positive semidefinite matrix with k positive entries in the diagonal, that is,

$$\mathcal{S}_k \triangleq \left\{\mathbf{S} \in \mathcal{S}_+^m : \|\mathrm{diag}(\mathbf{S})\|_0 = k\right\}. \tag{23}$$

Let the set of measurements that have not been compromised be

$$\mathcal{K}_o \triangleq \left\{i \in \{1, 2, \ldots, m\} : (\mathbf{\Sigma})_{ii} = 0\right\}, \tag{24}$$

where $(\mathbf{\Sigma})_{ii}$ is the entry of $\mathbf{\Sigma}$ in row i and column i. The sequential measurement selection imposes the following structure in the covariance matrix of the attack vector in Equation (7):

$$\mathbf{\Sigma}_{AA} = \mathbf{\Sigma} + v\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}, \tag{25}$$

where $i \in \mathcal{K}_o$ and $v \in \mathbb{R}_+$. From Equation (25), the cost function $f : \mathcal{S}_k \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{K}_o \to \mathbb{R}_+$ defined by adding Equations (10) and (12) is as follows:

$$f(\mathbf{\Sigma}, \lambda, v, i) \tag{26}$$

$$\triangleq I\left(X^n; Y_A^m\right) + \lambda D\left(P_{Y_A^m} \| P_{Y^m}\right) \tag{27}$$

$$= \frac{1}{2}\log\frac{|\mathbf{\Sigma}_{XX}||\mathbf{\Sigma}_{Y_AY_A}|}{|\mathbf{\Sigma}|} + \frac{1}{2}\lambda\left(\log\frac{|\mathbf{\Sigma}_{YY}|}{|\mathbf{\Sigma}_{Y_AY_A}|} - m + \mathrm{tr}\left(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{Y_AY_A}\right)\right) \tag{28}$$

$$= \frac{1}{2}\log\frac{|\mathbf{\Sigma}_{Y_AY_A}|}{|\sigma^2\mathbf{I}_m + \mathbf{\Sigma}_{AA}|} + \frac{1}{2}\lambda\left(\log\frac{|\mathbf{\Sigma}_{YY}|}{|\mathbf{\Sigma}_{Y_AY_A}|} + \mathrm{tr}\left(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{AA}\right)\right), \tag{29}$$

$$= \frac{1}{2}(1-\lambda)\log\left|\mathbf{\Sigma}_{YY} + \mathbf{\Sigma} + v\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right| - \frac{1}{2}\log\left|\mathbf{\Sigma} + v\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}} + \sigma^2\mathbf{I}_m\right|$$

$$+ \frac{1}{2}\lambda\left(\mathrm{tr}\left(\mathbf{\Sigma}_{YY}^{-1}\left(\mathbf{\Sigma} + v\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right)\right) + \log|\mathbf{\Sigma}_{YY}|\right), \tag{30}$$

where the inequality in Equation (28) holds from plugging Equations (10) and (12) into Equation (27); the equality in Equation (29) follows from cancelling $|\mathbf{\Sigma}_{XX}|$ in the first term [25], [Section 14.17] and noting that $\mathbf{\Sigma}_{Y_AY_A} = \mathbf{\Sigma}_{YY} + \mathbf{\Sigma}_{AA}$ in Equation (9); and the equality in Equation (30) holds from plugging Equation (25) into Equation (29).

4.2 | Information theoretic vulnerability of a measurement

We propose a notion of vulnerability that is linked to the information theoretic cost function proposed in ref. [24] to characterise the disruption and detection tradeoff incurred by the attacks. Taking the state of the system with k compromised measurements as the baseline, we quantify the vulnerability of measurement $i \in \mathcal{K}_o$ in terms of the cost decrease that i induces. In the following, we define the vulnerability of a measurement according to this idea.

Definition 1 The function $\Delta : \mathcal{S}_+^m \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{K}_o \to \mathbb{R}_+$, where $\mathcal{K}_o$ is in Equation (24), defines the vulnerability of measurement i in the following form:

$$\Delta(\boldsymbol{\Sigma}, \lambda, v, i) \triangleq f(\boldsymbol{\Sigma}, \lambda, v, i) - f(\boldsymbol{\Sigma}, \lambda, 0, i), \tag{31}$$

where the function f is defined in Equation (26).

Note that the attacker aims to minimise Equation (26) by choosing an index i and a variance v, and therefore, the definition above implies that given that k measurements in $\{1, 2, \ldots, m\} \backslash \mathcal{K}_o$ are already under attack in the system, the most vulnerable measurement is obtained by solving the following minimisation problem

$$\min_{i \in \mathcal{K}_o} \Delta(\boldsymbol{\Sigma}, \lambda, v, i), \tag{32}$$

where $\mathcal{K}_o$ is defined in Equation (24).

5 | VULNERABILITY OF MEASUREMENTS

5.1 | Vulnerability analysis of uncompromised systems

We first consider the case in which no measurements are under attacks, that is, $k = 0$, for which the following holds

$$\boldsymbol{\Sigma} = \mathbf{0}, \tag{33}$$

$$\mathcal{K}_o = \{1, 2, \ldots, m\}. \tag{34}$$

The attacker selects a single measurement with a given variance budget $v \leq v_0$. We quantify the vulnerability of measurement i in terms of $\Delta(\boldsymbol{\Sigma}, \lambda, v, i)$ defined in Equation (31). For the uncompromised system case, the optimisation problem in Equation (32) can be solved in closed form expression. The following theorem provides the solution.

Theorem 2 *The solution to the problem in Equation (32), with* $\mathcal{K}_o = \{1, 2, \ldots, m\}$, *is*

$$i = \arg\min_{j \in \{1, 2, \ldots, m\}} \left\{ \left(\boldsymbol{\Sigma}_{YY}^{-1} \right)_{jj} \right\}, \tag{35}$$

where $\boldsymbol{\Sigma}_{Y\,Y}$ is in Equation (5).

Proof We start by noting that Equation (33) establishes that the vulnerability of measurement i in Equation (31) is $\Delta(\mathbf{0}, \lambda, v, i)$. From the equality in Equation (30), the function $f(\mathbf{0}, \lambda, 0, i)$ is constant with respect to i. Hence, for $\boldsymbol{\Sigma} = \mathbf{0}$, the optimisation problem in Equation (32) is equivalent to

$$\min_{i \in \mathcal{K}_o} f(\mathbf{0}, \lambda, v, i), \tag{36}$$

where $\mathcal{K}_o$ is defined in Equation (34). Recall that $\lambda \in \mathbb{R}_+$ and $v \in \mathbb{R}_+$. From Equation (30), the resulting problem in Equation (36) is equivalent to the following optimisation problem:

$$\begin{aligned} \min_{i \in \{1, 2, \ldots, m\}} & (1 - \lambda) \log \left| \boldsymbol{\Sigma}_{YY} + v \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right| - \log \left| v \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} + \sigma^2 \mathbf{I}_m \right| \\ & + \lambda v \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right) \end{aligned} \tag{37}$$

$$\begin{aligned} = \min_{i \in \{1, 2, \ldots, m\}} & (1 - \lambda) \log \left| \mathbf{I}_m + v \boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right| - \log\left(v + \sigma^2 \right) \\ & + \lambda v \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right) \end{aligned} \tag{38}$$

$$\begin{aligned} = \min_{i \in \{1, 2, \ldots, m\}} & (1 - \lambda) \log\left(1 + v \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right) \right) \\ & + \lambda v \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right), \end{aligned} \tag{39}$$

where the equivalence in Equation (37) holds from plugging $\boldsymbol{\Sigma} = \mathbf{0}$ into the equality in Equation (30); the equality in Equation (38) follows from removing a constant $(1 - \lambda) \log |\boldsymbol{\Sigma}_{YY}|$ from the first term; and the equality in Equation (39) follows from the fact that $\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}}$ is a matrix with non-zero entries in the ith column and all the other entries are zero.

We now proceed by defining $t \triangleq v \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \right)$, with $t \in \mathbb{R}_+$, and rewriting the equality in Equation (39) as

$$\min_{t \in \mathbb{R}_+} (1 - \lambda) \log(1 + t) + \lambda t. \tag{40}$$

Note that Equation (40) increases monotonically with t. Therefore, the cost function in Equation (39) is monotonically increasing with t. This completes the proof. □

From Theorem 2, it follows that the identification of the most vulnerable measurement is independent of λ, introduced in Equation (26), and the value of the variance v. That is, it only depends on the system topology and parameters denoted by $\boldsymbol{\Sigma}_{Y\,Y}$ defined in Equation (5). This result coincides with Theorem 1 in the sense that in the attack construction for $k = 1$, the most vulnerable measurement is characterised in Equation (19), which is independent of the value of λ. The following corollary formalises this observation.

Corollary 1 *Let* $\boldsymbol{\Sigma} = \mathbf{0}$. *The vulnerability ranking for measurement indices*

$$\mathbf{s} = ^{\Delta}(s_1, s_2, \ldots, s_m) \tag{41}$$

is such that for all measurement index i, *with* $i \in \{1, 2, \ldots, m\}$, $s_i \in \{1, 2, \ldots, m\}$ *and*

$$\operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_{s_1} \mathbf{e}_{s_1}^{\mathsf{T}} \right) \leq \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_{s_2} \mathbf{e}_{s_2}^{\mathsf{T}} \right) \leq \ldots \leq \operatorname{tr}\left(\boldsymbol{\Sigma}_{YY}^{-1} \mathbf{e}_{s_m} \mathbf{e}_{s_m}^{\mathsf{T}} \right). \tag{42}$$

For all $i \in \{1, 2, \ldots, m\}$, *the* i*th most vulnerable measurement index is* s_i.

5.2 | Vulnerability index (VuIx)

The vulnerability analysis of uncompromised systems in Section 5.1 is constrained to $k = 0$. To generalise the vulnerability analysis to systems compromised with $k > 0$, in the following we propose a novel metric, coined *vulnerability index*.

Definition 2 For $k \in \{1, 2, \ldots, m-1\}$ and $\mathcal{S}_k$ in Equation (23), let the parameters be $\mathbf{\Sigma} \in \mathcal{S}_k$, $v \in \mathbb{R}_+$, $\lambda \in \mathbb{R}_+$. Consider the set $\{(i, \Delta) : i \in \mathcal{K}_o\}$, with $\mathcal{K}_o$ in Equation (24) and

$$\Delta_i \triangleq \Delta(\mathbf{\Sigma}, \lambda, v, i). \tag{43}$$

Let the vulnerability ranking

$$\mathbf{r} = \left(r_1, r_2, \ldots, r_{|\mathcal{K}_o|}\right) \tag{44}$$

be such that for all $i \in \{1, 2, \ldots, |\mathcal{K}_o|\}$, $r_i \in \mathcal{K}_o$ and moreover,

$$\Delta_{r_1} \leq \Delta_{r_2} \leq \ldots \leq \Delta_{r_{|\mathcal{K}_o|}}. \tag{45}$$

The vulnerability index (VuIx) of measurement $r_j \in \mathcal{K}_o$ is j, that is, $\text{VuIx}(r_j) = j$.

Note that the measurement with the smallest VuIx is the most vulnerable measurement and corresponds to the solution of the optimisation problem in Equation (32). The proposed VuIx for $i \in \mathcal{K}_o$ is obtained by Algorithm 1.

Algorithm 1 Computation of Vulnerability Index (VuIx)

Input: $\mathbf{H}$ in Equation (1);
σ^2 in Equation (2);
$\mathbf{\Sigma}_{X\,X}$ in Equation (3);
$\mathbf{\Sigma} \in \mathcal{S}_k$ in Equation (22);
$\lambda \in \mathbb{R}_+$ and $v \in \mathbb{R}_+$.
Output: the VuIx for all $i \in \mathcal{K}_o$.
1: Set $\mathcal{K}_o$ in Equation (24)
2: **for** $i \in \mathcal{K}_o$ **do**
3: Compute $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ in Equation (31)
4: **end for**
5: Sort $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ in ascending order
6: Set $\mathbf{r} = \left(r_1, r_2, \ldots, r_{|\mathcal{K}_o|}\right)$
7: Set the VuIx of measurement $r_j \in \mathcal{K}_o$ as j.

6 | NUMERICAL RESULTS

In this section, we numerically evaluate the VuIx of the measurements on a direct current (DC) setting for the IEEE Test systems [26]. The voltage magnitudes are set to 1.0 per unit, that is, the measurements of the systems are active power flow between the buses that are physically connected and active power injection to all the buses. The Jacobian matrix $\mathbf{H}$ in Equation (1) determined by the topology of the system and the physical parameters of the branches is generated by MATPOWER [27]. We adopt a Toeplitz model for the covariance matrix $\mathbf{\Sigma}_{XX}$ that arises in a wide range of practical settings, such as autoregressive stationary processes. Specifically, we model the correlation between state variable X_i and X_j with an exponential decay parameter $\rho \in \mathbb{R}_+$, which results in the entries of the matrix $(\mathbf{\Sigma}_{XX})_{ij} = \rho^{|i-j|}$ with $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$. In this setting, the VuIx of the measurements is also a function of the correlation parameter ρ, the noise variance σ^2, and the Jacobian matrix $\mathbf{H}$. The noise regime in the observation model is characterised by the signal to noise ratio (SNR) defined as follows:

$$\text{SNR} \triangleq 10\log_{10}\left(\frac{\text{tr}\left(\mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathsf{T}}\right)}{m\sigma^2}\right). \tag{46}$$

For all $\lambda \in \mathbb{R}_+$ and $v \in \mathbb{R}_+$, we generate a realisation of k attacked indices $\mathcal{K}_a \subseteq \{1, 2, \ldots, m\}$ that is uniformly sampled from the set of sets given by the following:

$$\widetilde{\mathcal{K}} = \{\mathcal{A} \subseteq \{1, 2, \ldots, m\} : |\mathcal{A}| = k\}. \tag{47}$$

We then construct a random covariance matrix describing the existing attacks on the system as follows:

$$\widetilde{\mathbf{\Sigma}} = \sum_{i \in \mathcal{K}_a} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}}, \tag{48}$$

with $\mathcal{K}_a \in \widetilde{\mathcal{K}}$. In the numerical simulation, we obtain the vulnerability of measurement i by computing

$$\Delta\left(\widetilde{\mathbf{\Sigma}}, \lambda, 1, i\right), \tag{49}$$

where $i \in \mathcal{K}_o$ is in Equation (24) and Δ is defined in Equation (31).

6.1 | Assessment of vulnerability index (VuIx)

Figures 1 and 2 depict the mean and variance of the VuIx obtained by Algorithm 1 for all the measurements with SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system when $k = 1$ and $k = 2$, respectively. Therein, it is observed that in general power injection measurements take higher vulnerability indices. Note that the vulnerability index captures the threat posed by an attack on sensor i expressed in terms of the vulnerability of the measurement as described by $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ in Algorithm 1. A larger value of $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ indicates a larger potential for an stealthy data integrity disruption induced by an attacker. Figures 1–6 depict a prevalence of higher vulnerability indices assigned to power injection

measurements for different system settings. This implies that corrupting the sensor data of power injection measurements is linked to larger information losses about the state of the grid, regardless of the attack construction used by the malicious attacker. Most power injection measurements correspond to higher ranked vulnerability indices but there are instances of power flow measurements with a higher ranked VuIx than that of power injection measurements. Interestingly, the power injection measurements with lower vulnerability indices correspond to the buses that are more isolated in the system, that is, the buses with a lower number of connections. On the other hand, the power flow measurements with higher ranked vulnerability indices correspond to the branches with higher admittance. The VuIx for $k = 0$ obtained in Corollary 1 is depicted for the purpose of serving as a reference to assess the deviation when $k > 0$. In this setting, the VuIx of most measurements does not change substantially for different values of k, which suggests that the VuIx is insensitive to the state of the system.

Figures 3 and 4 depict the mean and variance of the VuIx from Algorithm 1 for all the measurements with SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system when $k = 1$ and $k = 2$, respectively. Similar to what is observed above, the mean of the VuIx for most of the measurements does not deviate significantly from the case when $k = 0$. However, most of the variance values deviate significantly in comparison with the cases in Figures 1 and 2 with SNR = 10 dB. Figures 5 and 6 depict the results on IEEE 30-bus systems with the same setting as in Figures 1 and 2, respectively. Figures 7 and 8 depict the results on IEEE 30-bus systems with the same setting as in Figures 3 and 4, respectively. Surprisingly, the mean of the VuIx in larger systems coincides with that obtained for the case $k = 0$, which suggests that the VuIx is a robust security metric for large systems. In line with the previous observation, the power injection measurements corresponding to the least connected buses decrease in the VuIx when SNR = 10 dB.

6.2 | Comparative vulnerability assessment of power flow and power injection measurements

In Section 6.1, we have established that power injection measurements and power flow measurements are qualitatively different in terms of the VuIx. To provide a quantitative description of this difference, Figure 9 depicts the probability of a given VuIx $i \in \{1, 2, \ldots, m - |\mathcal{K}_a|\}$ being taken by a power injection measurement or a power flow measurement for the IEEE 9-bus and 30-bus systems when $\lambda = 2$, $k = 2$, SNR = 30 dB and $\rho = 0.1$. Specifically, Figure 9 depicts the probability of the following events:

Flow_i : VuIx i corresponds to a power flow measurement,
Inj_i : VuIx i corresponds to a power injection measurement.

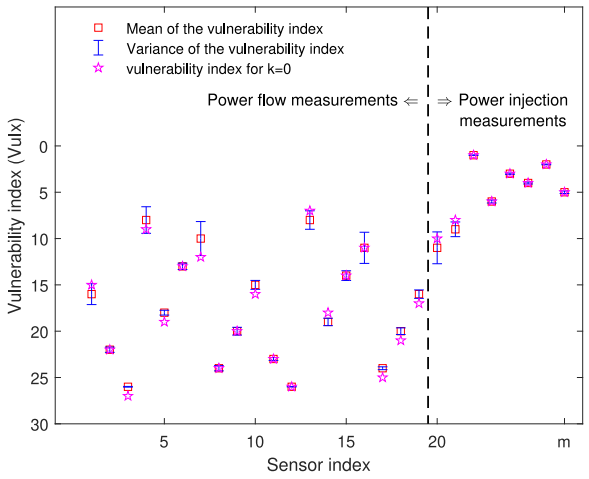

FIGURE 1 Vulnerability index (VuIx) when $k = 1$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

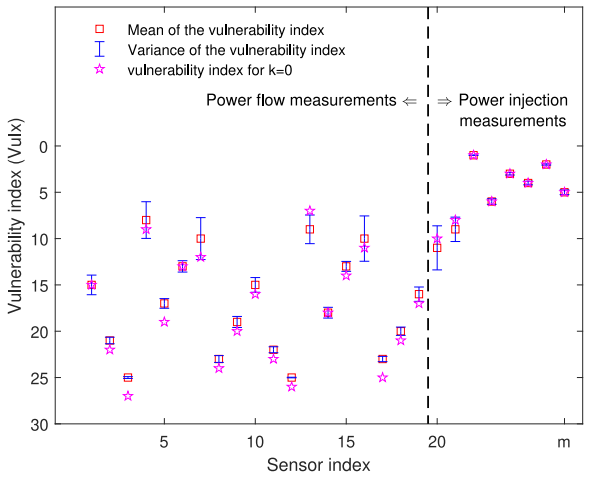

FIGURE 2 Vulnerability index (VuIx) when $k = 2$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

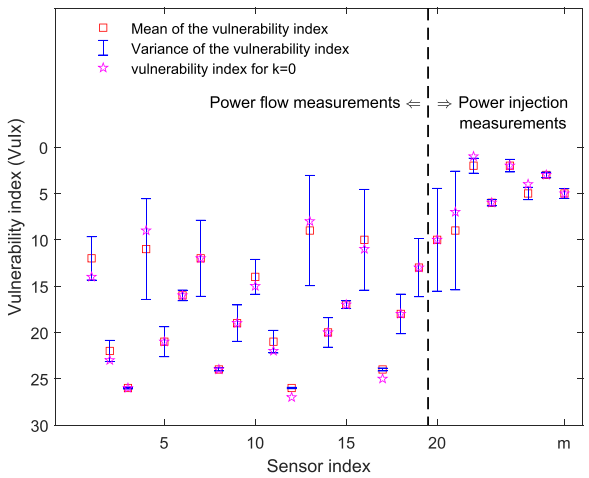

FIGURE 3 Vulnerability index (VuIx) when $k = 1$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

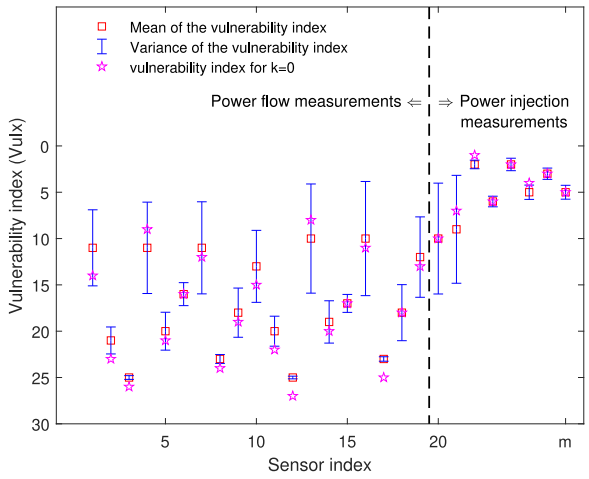

FIGURE 4 Vulnerability index (VuIx) when $k = 2$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

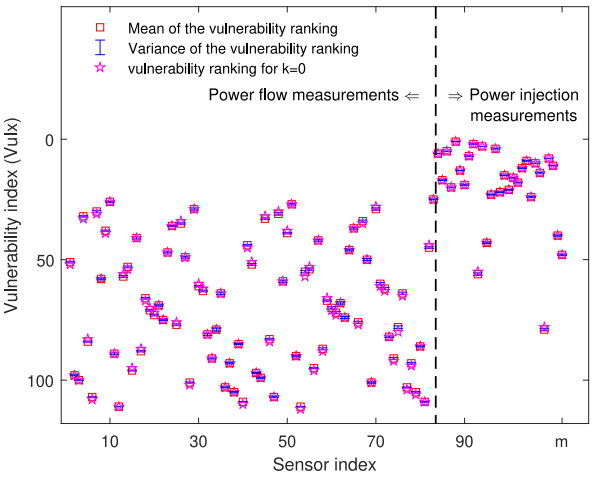

FIGURE 5 Vulnerability index (VuIx) when $k = 1$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

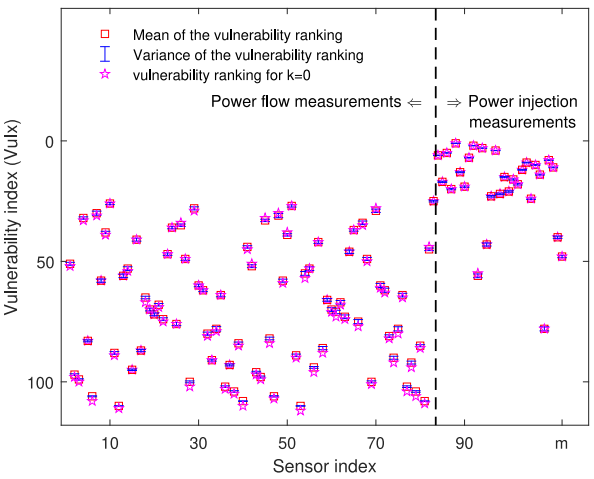

FIGURE 6 Vulnerability index (VuIx) when $k = 2$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

It is observed that in both systems, small VuIx are more likely to correspond to power injection measurements than to power flow measurements, that is, $\mathbb{P}[\mathsf{Inj}_i] > \mathbb{P}[\mathsf{Flow}_i]$ for small values of i. Conversely, it holds that $\mathbb{P}[\mathsf{Inj}_i] < \mathbb{P}[\mathsf{Flow}_i]$ for large values of i. In fact, small VuIx correspond to power injection measurements with probability one, which suggests that the most vulnerable measurements in the system tend to be power injection measurements. Conversely, the larger VuIx values correspond to power flow measurements with probability one, which indicates that the least vulnerable measurements tend to be power flow measurements. Interestingly, there is a clear demarcation for each system for which $\mathbb{P}[\mathsf{Inj}_i]$ and $\mathbb{P}[\mathsf{Flow}_i]$ change rapidly with the VuIx value, which points to a phase transition type phenomenon for measurement vulnerability.

The probability of VuIx taken by power injection measurements concentrates higher probability mass for higher priority vulnerability indices. One the other hand, power flow measurements with higher probability mass coincide with low ranked VuIx values. Precisely, the probability of the vulnerability indices with higher priority taken by power injection measurements is one in both IEEE 9-bus and 30-bus systems. Meanwhile, the probability of the lower ranked vulnerability indices taken by power flow measurements is one. Note that the probability of mid-ranked vulnerability indices taken by power injection measurements drops significantly, which indicates that there are some power flow measurements that are equally as vulnerable as power injection measurements. We observe that these power flow measurements correspond to the branches with higher admittance. The power injection measurements with lower vulnerability indices correspond with the buses that are isolated in the systems.

Figure 10 depicts the distribution of VuIx for power injection measurements and power flow measurements on the IEEE 9-bus and 30-bus systems when $\lambda = 2$, $k = 2$, SNR = 30 dB and $\rho = 0.1$. Specifically, Figure 10 depicts the probability mass function of the following events:

$\mathsf{VuIx}(\mathsf{Flow}) = i$: VuIx for power flow measurements is i,
$\mathsf{VuIx}(\mathsf{Inj}) = i$: VuIx for power injection measurements is i.

Power injection measurements have a higher probability with high ranked VuIx, whereas power flow measurements have much higher probability with low ranked VuIx. It is worth noting that the probability mass functions are close to uniform for high and low vulnerability index ranges. This suggests that the most vulnerable measurements in the system are contained with high probability in a subset of the power injection measurements. Conversely, the least vulnerable measurements comprise the majority of the power flow measurements with no apparent preference over the majority. Surprisingly, in the 30-bus system, the probability of lowest ranked VuIx for power flow measurements experiences a sharp increase.

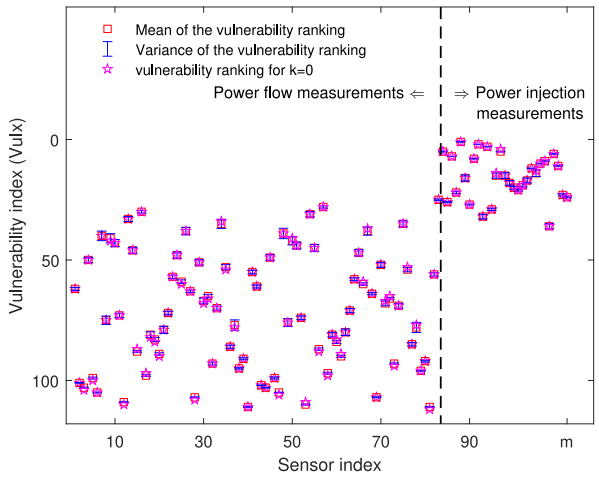

FIGURE 7 Vulnerability index (VuIx) when $k = 1$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

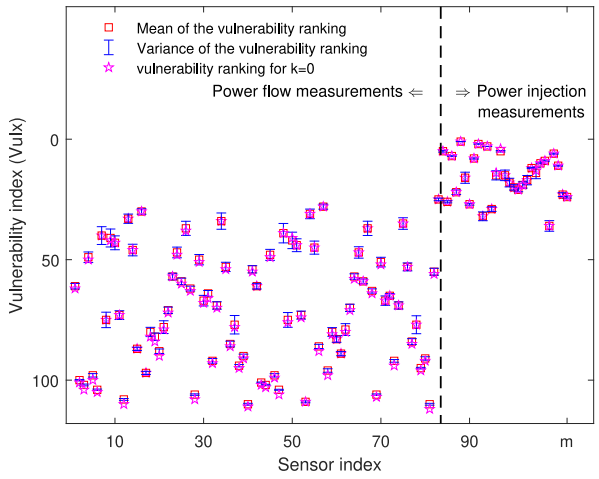

FIGURE 8 Vulnerability index (VuIx) when $k = 2$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

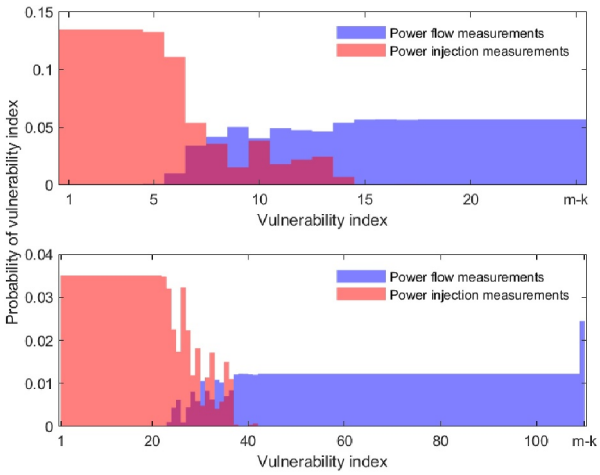

FIGURE 9 Probability mass function of Vulnerability index (VuIx) for power injection measurements and power flow measurements when $\lambda = 2$, $k = 2$, SNR = 30 dB and $\rho = 0.1$ on IEEE 9-bus and 30-bus systems, respectively.

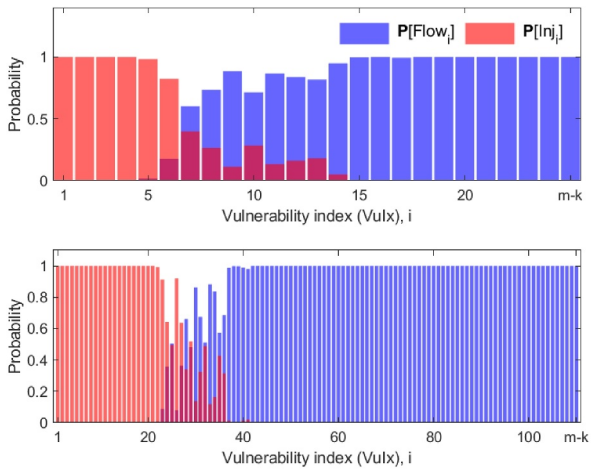

FIGURE 10 Probability of Vulnerability index (VuIx) corresponds to power injection measurements and power flow measurements when $\lambda = 2$, $k = 2$, SNR = 30 dB and $\rho = 0.1$ on IEEE 9-bus and 30-bus systems, respectively.

7 | CONCLUSION

In this paper, we have proposed, from a fundamental perspective, a novel security metric referred to as vulnerability index (VuIx) that characterises the vulnerability of power system measurements to data integrity attacks. We have achieved this by embedding information theoretic measures into the metric definition. The resulting VuIx framework evaluates the vulnerability of all the measurements in the systems and enables the operator to identify those that are more exposed to data integrity threats. We have tested the framework for IEEE test systems and concluded that power injection measurements are more vulnerable to data integrity attacks than power flow measurements.

AUTHOR CONTRIBUTIONS

Xiuzhen Ye: Proposed the original novel idea and refined the idea with co-authors; designed the framework and the test; wrote and revised the manuscript. Iñaki Esnaola: Worked with the first author and gave advice on the ideas and helped refined the idea; revised the paper. Samir M. Perlaza: Worked with the first author and gave advice on the ideas and helped refined the idea. Robert F. Harrison: Gave high level advice on the ideas and provided broader vision on the field and the application on data security in general.

ACKNOWLEDGEMENTS

This study was funded by the China Scholarship Council, grant number 201906150124 and the European Commission through the H2020-MSCA-RISE-2019 program, grant number 872172.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable – no new data generated.

ORCID

Xiuzhen Ye https://orcid.org/0000-0002-7859-713X

REFERENCES

1. Grainger, J.J., Stevenson, W.D.: Power System Analysis. McGraw-Hill (1994)
2. Abur, A., Exposito, A.G.: Power System State Estimation: Theory and Implementation. CRC press (2004)
3. Wang, W., Lu, Z.: Cyber security in the smart grid: survey and challenges. Comput. Network. 57(5), 1344–1371 (2013). https://doi.org/10.1016/j.comnet.2012.12.017
4. Jaquith, A.: Security Metrics: Replacing Fear, Uncertainty, and Doubt. Pearson Education (2007)
5. Mell, P., Scarfone, K., Romanosky, S.: Common vulnerability scoring system. IEEE Secur. Priv. 4(6), 85–89 (2006). https://doi.org/10.1109/msp.2006.145
6. Pallitteri, V.Y., Brewer, T.L.: Guidelines for smart grid cybersecurity. In: NIST Interagency/Internal Report (NISTIR). National Institute of Standards and Technology (2014). https://doi.org/10.6028/NIST.IR.7628r1
7. Pendleton, M., et al.: A survey on systems security metrics. ACM Comput. Surv. 49(4), 1–35 (2017). https://doi.org/10.1145/3005714
8. Venkataramanan, V., Hahn, A., Srivastava, A.: CP-SAM: cyber-physical security assessment metric for monitoring microgrid resiliency. IEEE Trans. Smart Grid 11(2), 1055–1065 (2022). https://doi.org/10.1109/tsg.2019.2930241
9. Liu, Y., Ning, P., Reiter, M.K.: False data injection attacks against state estimation in electric power grids. ACM Trans. Inf. Syst. Secur. 14(1), 1–33 (2011). https://doi.org/10.1145/1952982.1952995
10. Cui, S., et al.: Coordinated data-injection attack and detection in the smart grid: a detailed look at enriching detection solutions. IEEE Signal Process. Mag. 29(5), 106–115 (2012). https://doi.org/10.1109/msp.2012.2185911
11. Ozay, M., et al.: Sparse attack construction and state estimation in the smart grid: centralized and distributed models. IEEE J. Sel. Area. Commun. 31(7), 1306–1318 (2013). https://doi.org/10.1109/jsac.2013.130713
12. Esnaola, I., et al.: Maximum distortion attacks in electricity grids. IEEE Trans. Smart Grid 7(4), 2007–2015 (2016). https://doi.org/10.1109/tsg.2016.2550420
13. Ozay, M., et al.: Machine learning methods for attack detection in the smart grid. IEEE Trans. Neural Networks Learn. Syst. 27(8), 1773–1786 (2015). https://doi.org/10.1109/tnnls.2015.2404803
14. Bretas, A., et al.: Cyber-physical Power Systems State Estimation. Elsevier (2021)
15. Sun, K., et al.: Learning requirements for stealth attacks. In: Proc. IEEE Int. Conf. On Acoustics, Speech and Signal Processing, pp. 8102–8106 (2019)
16. Sun, K., et al.: Asymptotic learning requirements for stealth attacks. IEEE Trans. Smart Grid 14(4), 3189–3200 (2023). https://doi.org/10.1109/tsg.2023.3236785
17. Ye, X., et al.: Information theoretic data injection attacks with sparsity constraints. In: Proc. IEEE Int. Conf. On Smart Grid Comm, pp. 1–6. IEEE, Tempe (2020)
18. Ye, X., et al.: Stealth data injection attacks with sparsity constraints. IEEE Trans. Smart Grid 14(4), 3201–3209 (2023). https://doi.org/10.1109/tsg.2023.3238913
19. Genes, C., et al.: Robust recovery of missing data in electricity distribution systems. IEEE Trans. Smart Grid 10(4), 4057–4067 (2018). https://doi.org/10.1109/tsg.2018.2848935
20. Shomorony, I., Avestimehr, A.S.: Worst-case additive noise in wireless networks. IEEE Trans. Inf. Theor. 59(6), 3833–3847 (2013). https://doi.org/10.1109/tit.2013.2248875
21. Lévy, P.: Propriétés asymptotiques des sommes de variables aléatoires enchaînées. J. Math. Pure Appl. 14, 109–128 (1935)
22. Cramér, H.: Über eine Eigenschaft der normalen Verteilungsfunktion. Math. Z. 41(1), 405–414 (1936). https://doi.org/10.1007/bf01180430
23. Cover, T.M., Thomas, J.A.: Elements of Information Theory. John Wiley and Sons (1999)
24. Sun, K., et al.: Stealth attacks on the smart grid. IEEE Trans. Smart Grid 11(2), 1276–1285 (2019). https://doi.org/10.1109/tsg.2019.2935353
25. Seber, G.A.: A Matrix Handbook for Statisticians, vol. 15. John Wiley and Sons (2008). https://doi.org/10.1002/9780470226797
26. UM: Power Systems Test Case Archive. https://labs.ece.uw.edu/pstca/
27. Zimmerman, R.D., Murillo-Sánchez, C.E., Thomas, R.J.: MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education. IEEE Trans. Power Syst. 26(1), 12–19 (2010). https://doi.org/10.1109/tpwrs.2010.2051168

How to cite this article: Ye, X., et al.: An information theoretic metric for measurement vulnerability to data integrity attacks on smart grids. IET Smart Grid. 1–10 (2024). https://doi.org/10.1049/stg2.12163