An Information Theoretic Metric for Measurement Vulnerability to Data Integrity Attacks on Smart Grids

Xiuzhen Ye, Iñaki Esnaola, Samir M Perlaza, Robert F Harrison

An Information Theoretic Metric for Measurement Vulnerability to Data Integrity Attacks on Smart Grids

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

[1] *Dept. of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK*
[2] *Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA*
[3] *INRIA, Sophia Antipolis 06902, France*
* *E-mail: Xye15@sheffield.ac.uk*

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.

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 characterize 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) analysis Information Technology (IT) system [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 categorized 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 [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 [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 characterize 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 characterize the amount of information acquired by the measurements in the system in fundamental terms.

The rest of the paper is organized as follows: In Section 2, we introduce a Bayesian framework with linearized 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 characterize 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 follow: (1) A notion of vulnerability for the measurements in the system is proposed. The proposed notion is characterized 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 minimization 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 characterizations 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 $\mathrm{diag}(\mathbf{A})$ 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 i-th entry. Random variables are denoted by capital letters and their realizations by the corresponding lower case, e.g., x is

a realization of the random variable X. Vectors of n random variables are denoted by a superscript, e.g., $X^n = (X_1, \ldots, X_n)^\mathsf{T}$ with corresponding realizations 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 linearized 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,

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

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 (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 $\boldsymbol{\Sigma}_{AA} \sim S_{++}^m$, 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 minimized [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 (2), and therefore, A^m is Gaussian. In view of this, in the following, we assume that

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

where $\mathbf{0} = (0, 0, \ldots, 0)$ and $\boldsymbol{\Sigma}_{AA} \in \mathcal{S}_+^m$ are the mean vector and the covariance matrix of the random attack vector A^m. The assumption in (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 $\boldsymbol{\Sigma}_{Y_AY_A} \in \mathcal{S}_+^m$, that is,

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

with

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

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 minimize the mutual information between the vector of state variables X^n in (3) and the vector of compromised measurements Y_A^m in (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 (6) and the distribution P_{Y^m} in (1), that is, $D(P_{Y_A^m} \| P_{Y^m})$. For the observation model and attack setting described in Section 2, and assuming optimal detection, the Chernoff-Stein Lemma [23] states that the minimization of KL divergence leads to the minimization of the asymptotic detection probability.

The following propositions characterize 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* (3) *and Y_A^m in* (8) *is*

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

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

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

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

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

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

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

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

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

where $\lambda \in \mathbb{R}_+$ is the weighting parameter that determines the trade-off between mutual information and KL divergence. Note that the optimization domain in (13) is the set of m-dimensional Gaussian multivariate distributions. The optimal Gaussian attack for $\lambda \geq 1$ as

a solution to (13) is given by [24]

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

Note that the attack realizations from (14) are nonzero with probability one, that is, $\mathbb{P}[|\text{supp}(A^m)| = m] = 1$, where

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

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 optimization domain over the attack vector A^m in (6) to the distributions with cardinality of the support satisfying $|\text{supp}(A^m)| = k \leq m$, that is,

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

The resulting sparse attack construction is [18]

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

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* (17) *for the case* $k = 1$ *is*

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

where

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

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

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

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 idependent, that is,

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

where A_i is the i-th 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 (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,

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

Let the set of measurements that have not been compromised be

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

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

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

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

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

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

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

where the inequality in (28) holds from plugging (10) and (12) into (27); the equality in (29) follows from cancelling $|\mathbf{\Sigma}_{XX}|$ in the first term [25, Sec. 14.17] and noting that $\mathbf{\Sigma}_{Y_AY_A} = \mathbf{\Sigma}_{YY} + \mathbf{\Sigma}_{AA}$ in (9); and the equality in (30) holds from plugging (25) into (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 [24] to characterize 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* (24)*, defines the vulnerability of measurement* i *in the following form:*

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

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

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

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

where $\mathcal{K}_o$ is defined in (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 the following holds

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

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

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 $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ defined in (31). For the uncompromised system case, the optimization problem in (32) can be solved in closed form expression. The following theorem provides the solution.

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

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

where $\mathbf{\Sigma}_{YY}$ *is in* (5)*.*

Proof: We start by noting that (33) establishes that the vulnerability of measurement i in (31) is $\Delta(\mathbf{0}, \lambda, v, i)$. From the equality in (30), the function $f(\mathbf{0}, \lambda, 0, i)$ is constant with respect to i. Hence, for $\mathbf{\Sigma} = \mathbf{0}$, the optimization problem in (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 (34). Recall that $\lambda \in \mathbb{R}_+$ and $v \in \mathbb{R}_+$. From (30), the resulting problem in (36) is equivalent to the following optimization problem:

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

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

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

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

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

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

Note that (40) increases monotonically with t. Therefore, the cost function in (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 (26), and the value of the variance v. That is, it only depends on the system topology and parameters denoted by $\mathbf{\Sigma}_{YY}$ defined in (5). This result coincides with Theorem 1 in the sense that in the attack construction for $k = 1$, the most vulnerable measurement is characterized in (19), which is independent of the value of λ. The following corollary formalizes this observation.

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

$$\mathbf{s} \triangleq (s_1, s_2, \ldots, s_m) \tag{41}$$

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

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

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 generalize 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* (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* (24) *and*

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

Let the vulnerability ranking

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

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 optimization problem in (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 (1);
σ^2 in (2);
$\mathbf{\Sigma}_{XX}$ in (3);
$\mathbf{\Sigma} \in \mathcal{S}_k$ in (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 (24)
2: **for** $i \in \mathcal{K}_o$ **do**
3: Compute $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ in (31)
4: **end for**
5: Sort $\Delta(\mathbf{\Sigma}, \lambda, v, i)$ in ascending order
6: Set $\mathbf{r} = (r_1, r_2, \ldots, r_{|\mathcal{K}_o|})$
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,

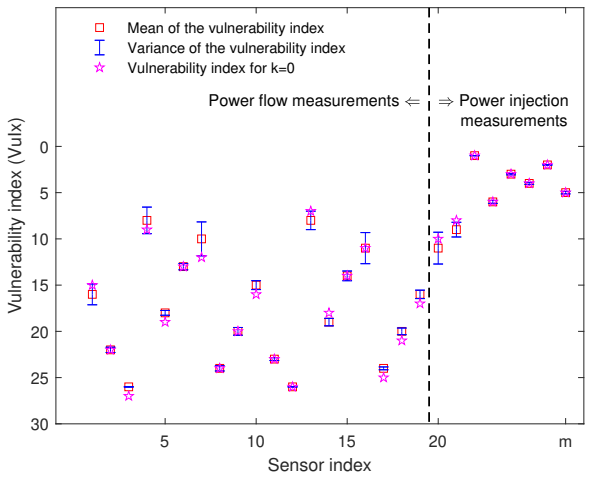

Fig. 1: Vulnerability index (VuIx) when $k = 1$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

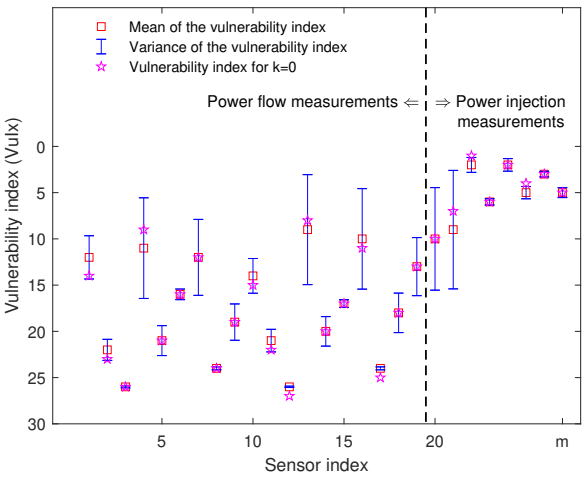

Fig. 3: Vulnerability index (VuIx) when $k = 1$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

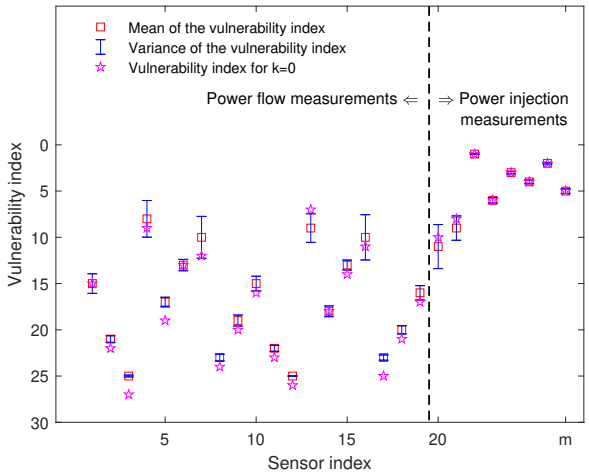

Fig. 2: Vulnerability index (VuIx) when $k = 2$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

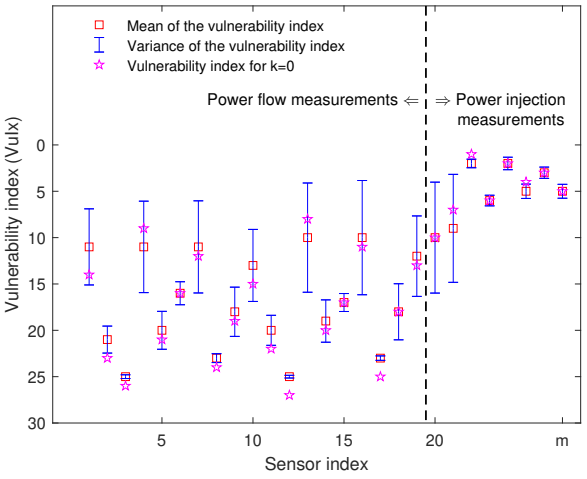

Fig. 4: Vulnerability index (VuIx) when $k = 2$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 9-bus system.

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 (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 characterized by the signal to noise ratio (SNR) defined as

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

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

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

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

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

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

$$\Delta(\tilde{\mathbf{\Sigma}}, \lambda, 1, i), \tag{49}$$

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

6.1 Assessment of vulnerability index (VuIx)

Fig. 1 and Fig. 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. Fig.1-6 depict a prevalence of

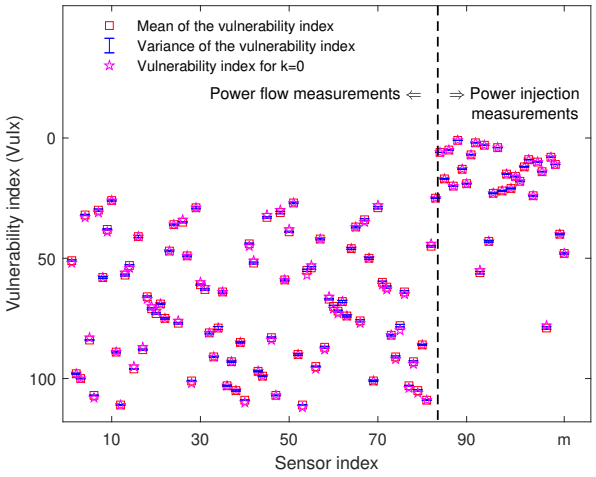

Fig. 5: Vulnerability index (VuIx) when $k = 1$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

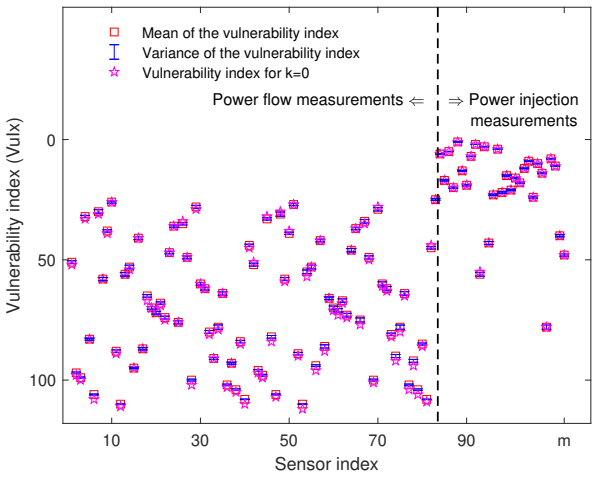

Fig. 6: Vulnerability index (VuIx) when $k = 2$, SNR = 10 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

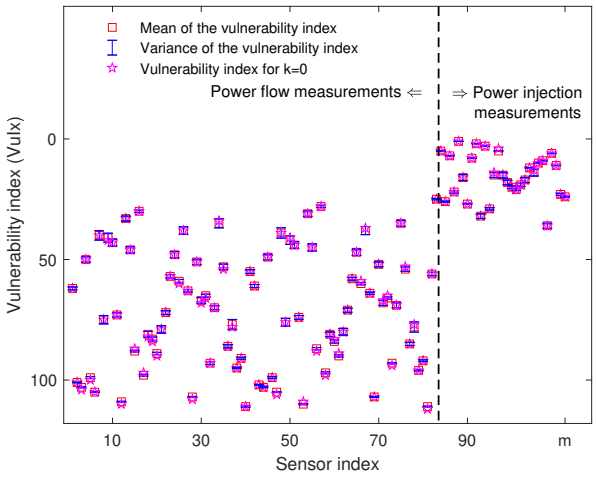

Fig. 7: Vulnerability index (VuIx) when $k = 1$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

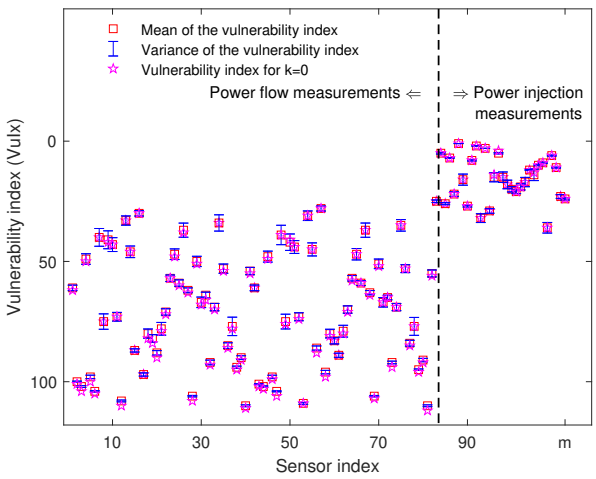

Fig. 8: Vulnerability index (VuIx) when $k = 2$, SNR = 30 dB, $\lambda = 2$ and $\rho = 0.1$ on the IEEE 30-bus system.

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.

Fig. 3 and Fig. 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. Similarly 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 Fig. 1 and Fig. 2 with SNR = 10 dB. Fig. 5 and Fig. 6 depict the results on IEEE 30-bus systems with the same setting as in Fig. 1 and Fig. 2, respectively. Fig. 7 and Fig. 8 depict the results on IEEE 30-bus systems with the same setting as in Fig. 3 and Fig. 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, Fig. 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, Fig. 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.

It is observed that in both systems, small VuIx are more likely to correspond to power injection measurements than to power flow

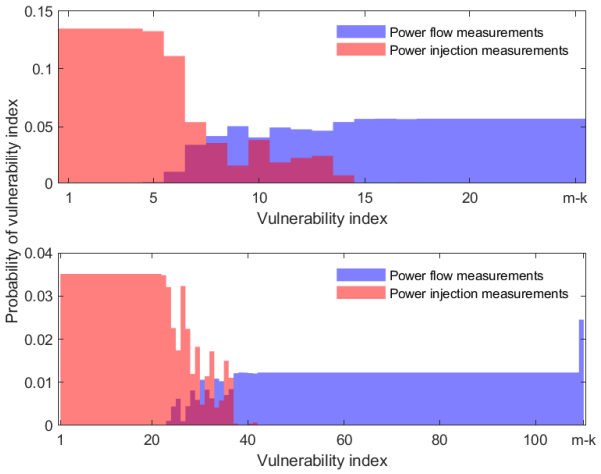

Fig. 9: Probability mass function of Vulnerability index (VuIx) for power injection measurements and power flow measurements when $\lambda = 2$, $k = 2$, SNR = 30 dB and $\rho = 0.1$ on IEEE 9-bus and 30-bus systems, respectively.

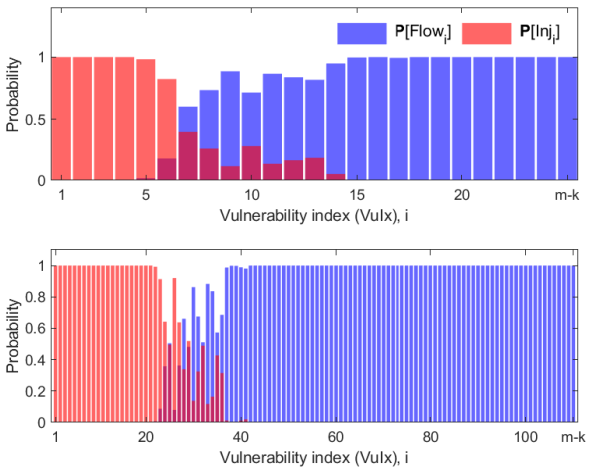

Fig. 10: Probability of Vulnerability index (VuIx) corresponds to power injection measurements and power flow measurements when $\lambda = 2$, $k = 2$, SNR = 30 dB and $\rho = 0.1$ on IEEE 9-bus and 30-bus systems, respectively.

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.

Fig. 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, Fig. 10 depicts the probability mass function of the following events:

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

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.

7 Conclusion

In this paper, we have proposed, from a fundamental perspective, a novel security metric referred to as vulnerability index (VuIx) that characterizes 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.

8 References

1 Grainger JJ, Stevenson WD. Power system analysis. McGraw-Hill; 1994.
2 Abur A, Exposito AG. Power system state estimation: Theory and implementation. CRC press; 2004.
3 Wang W, Lu Z. Cyber security in the smart grid: Survey and challenges. Computer networks. 2013 Jan;57(5):1344–1371.
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 Security & Privacy. 2006;4(6):85–89.
6 Pallitteri VY, Brewer TL. Guidelines for Smart Grid Cybersecurity. NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology; 2014. Available from: https://doi.org/10.6028/NIST.IR.7628r1.
7 Pendleton M, Garcia-Lebron R, Cho JH, Xu S. A survey on systems security metrics. ACM Computing Surveys. 2017 Dec;49(4):1–35.
8 Venkataramanan V, Hahn A, Srivastava A. CP−SAM: Cyber-physical security assessment metric for monitoring microgrid resiliency. IEEE Trans Smart Grid. 2022 Mar;11(2):1055–1065.
9 Liu Y, Ning P, Reiter MK. False data injection attacks against state estimation in electric power grids. ACM Trans Info Syst Sec. 2011 May;14(1):1–33.
10 Cui S, Han Z, Kar S, Kim TT, Poor HV, Tajer A. Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions. IEEE Signal Process Mag. 2012 Aug;29(5):106–115.
11 Ozay M, Esnaola I, Vural FTY, Kulkarni SR, Poor HV. Sparse attack construction and state estimation in the smart grid: Centralized and distributed models. IEEE J Sel Areas Commun. 2013 Jul;31(7):1306–1318.
12 Esnaola I, Perlaza SM, Poor HV, Kosut O. Maximum distortion attacks in electricity grids. IEEE Trans Smart Grid. 2016 Jul;7(4):2007–2015.
13 Ozay M, Esnaola I, Vural FTY, Kulkarni SR, Poor HV. Machine learning methods for attack detection in the smart grid. IEEE Trans on neural networks and learning systems. 2015;27(8):1773–1786.

14 Bretas A, Bretas N, London Jr JB, Carvalho B. Cyber-physical power systems state estimation. Elsevier; 2021.

15 Sun K, Esnaola I, Tulino AM, Poor HV. Learning requirements for stealth attacks. In: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing. Brighton, UK; 2019. p. 8102–8106.

16 Sun K, Esnaola I, Tulino AM, Poor HV. Asymptotic learning requirements for stealth attacks. IEEE Trans on Smart Grid. 2023;14(4):3189 – 3200.

17 Ye X, Esnaola I, Perlaza SM, Harrison RF. Information theoretic data injection attacks with sparsity constraints. In: Proc. IEEE Int. Conf. on Smart Grid Comm. Tempe, AZ, USA: IEEE; 2020. p. 1–6.

18 Ye X, Esnaola I, Perlaza SM, Harrison RF. Stealth data injection attacks with sparsity constraints. IEEE Trans on Smart Grid. 2023 Jul;14(4):3201–3209.

19 Genes C, Esnaola I, Perlaza SM, Ochoa LF, Coca D. Robust recovery of missing data in electricity distribution systems. IEEE Trans Smart Grid. 2018 Jun;10(4):4057–4067.

20 Shomorony I, Avestimehr AS. Worst-case additive noise in wireless networks. IEEE Trans Inf Theory. 2013 Jun;59(6):3833–3847.

21 Lévy P. Propriétés asymptotiques des sommes de variables aléatoires enchaînées. J Math Pures Appl. 1935;14:109–128.

22 Cramér H. Über eine Eigenschaft der normalen Verteilungsfunktion. Math Z. 1936;41:405–414.

23 Cover TM, Thomas JA. Elements of information theory. John Wiley & Sons; 1999.

24 Sun K, Esnaola I, Perlaza SM, Poor HV. Stealth attacks on the smart grid. IEEE Trans Smart Grid. 2019 Aug;11(2):1276–1285.

25 Seber GA. A matrix handbook for statisticians. vol. 15. John Wiley & Sons; 2008.

26 of Washington U. Power Systems Test Case Archive;. Available from: `https://labs.ece.uw.edu/pstca/`.

27 Zimmerman RD, Murillo-Sánchez CE, Thomas RJ. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE Trans Power Syst. 2010 Feb;26(1):12–19.