Detection of False Data Injection Attacks in Power System State Estimation Using Sensor Encoding

Rodrigo D. Trevizan
Energy Storage Technology & Systems
Sandia National Laboratories
Albuquerque, NM, USA
rdtrevi@sandia.gov

Matthew Reno
Electric Power Systems Research
Sandia National Laboratories
Albuquerque, NM, USA
mjreno@sandia.gov

***Abstract*—In this paper, we present a sensor encoding technique for the detection of stealthy false data injection attacks in static power system state estimation. This method implements low-cost verification of the integrity of measurement data, allowing for the detection of stealthy additive attack vectors. It is considered that these attacks are crafted by malicious actors with knowledge of the system models and capable of tampering with any number of measurements. The solution involves encoding all vulnerable measurements. The effectiveness of the method was demonstrated through a simulation where a stealthy attack on an encoded measurement vector generates large residuals that trigger a chi-squared anomaly detector (e.g. χ^2). Following a defense in-depth approach, this method could be used with other security features such as communications encryption to provide an additional line of defense against cyberattacks.**

***Index Terms*—bad data detection, cybersecurity, false data injection, sensor encoding, state estimation.**

I. Introduction

With the adoption of communications and data processing capabilities, power grids have increased their operational flexibility. Analysis of this new paradigm of communications-enabled power grids has drawn attention to cybersecurity-related risks. Recent cyberattacks on power grids (e.g. [1]) have confirmed the soundness of these concerns. In North America, regulators have worked to enforce cybersecurity regulations on bulk power systems [2].

In academia, several works have been dedicated to investigating cybersecurity risks and methods to make power systems more robust and resilient to cyberattacks. Power systems applications that have been identified as vulnerable to cyberattacks include power system state estimators (PSSEs), automatic generation control, voltage control, and energy markets [3]. Cyberattacks targeting the integrity of data have been identified as a serious threat and they add up to challenges PSSE faces in practical applications, such as model errors [4]–[6]. One such attack named False data injection (FDI) attack has received considerable attention from researchers. FDI is the modification of data, such as power system measurements or control signals, by a threat actor. FDI attacks targeting PSSEs could be implemented as an integrity attack to modify the values of measurements, circuit breaker statuses, and other critical data. These attacks can be perpetrated at the meter or communication system levels, and their goal is to harm the situational awareness of power system operators and to induce errors in the operation of applications that rely on state estimates. In [7], the authors have introduced models for stealthy cyberattacks on PSSEs. The measurements used in the PSSE could be manipulated by an additive FDI vector capable of defeating traditional residual-based approaches for bad data detection (BDD), therefore called stealthy. To achieve this goal, a malicious actor needs knowledge of the system model and the capability to manipulate several measurements. The system representations of the static state estimator used in the paper are linear, which limits the scope of attacks. This approach was extended for nonlinear state estimators in [8]. In [9], the authors have designed a method to quantify the cost of attacks as a function of the number of measurements that need to be compromised to obtain a stealthy FDI attack vector. Additionally, a method to obtain the largest minimum cost attack is proposed as a criterion to determine the best set of measurements to protect under constraints of the number of protected meters.

Several methods based on statistical tests have been proposed. To detect stealthy FDI attacks on supervisory control and data acquisition (SCADA) measurements, an online anomaly detection algorithm that leverages load forecasts, generation schedules and synchrophasor data was proposed by [10]. It is assumed that remote terminal unit (RTU) measurements communicated over SCADA systems are vulnerable to attacks while it requires a significantly larger effort to compromise Synchrophasor measurements, generation schedules and load forecasts. The algorithm obtains an estimate of the states based on the protected data and compares it to the results of the SCADA-fed PSSE. In [11] the authors have proposed a modified χ^2 test for the detection of gross errors in power

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Number 34226. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2022-XXXX C

978-1-6654-6591-5/22/$31.00 ©2022 IEEE

system measurements that combines residuals from a weighted least-squares state estimator (WLSSE) with a score obtained by the Mahalanobis distance of a Reed-Xiaoli (RX) Anomaly Detector [12]. Methods for FDI detection in dynamic state estimation have also been proposed. In [13], the χ^2 and the Euclidean detectors were applied to the pre-fit residuals of the Kalman Filter (KF) used for dynamic PSSE.

Machine learning approaches have also been proposed as an alternative to detect FDI attacks on PSSE. In [14], a mixture Gaussian distribution learning method was applied to FDI detection in linear PSSE and has shown higher detection scores than support-vector machines, multi-layer perceptron neural networks, and a semi-supervised anomaly detection method.

Outside of the research on FDI attacks targeting static PSSEs, researchers have proposed other solutions for detecting integrity cyberattacks in dynamical systems. One of such techniques, known as measurement encoding, involves modifying the values of sensor readings to harm the capability of the attacker to design a stealthy attack sequence in measurements. Conditions for designing undetectable cyberattacks capable of introducing estimation errors on KF have been designed in [15]. A measurement encoding strategy was developed in [16] to increase the residuals obtained by the KF applied to the attack sequence previously defined in [15]. Those methods, however, are tailored for dynamic systems and their applicability to static PSSE is limited.

In this paper, we propose two sensor encoding techniques for the detection of stealthy FDI attacks in static PSSE. These methods implement low-cost verification of the integrity of measurement data, allowing for the detection of stealthy additive attack vectors. It is considered that these attacks are crafted by malicious actors with knowledge of system models and capable of tampering with any number of measurements. The solution involves encoding the measurements, such that if an attacker tries to add bias and remain undetected, its attack should generate large residuals that will trigger an anomaly detector (e.g. χ^2). Because this method is based on the numerical transformation of measurements values, it can be applied to any communications protocol. While practical implementation aspects are not the focus of the paper, conceptually, this approach could be implemented in the measurement device level in a way similar to calibration.

These methods can be combined with other information security features, such as data encryption and network segmentation. Compared to common encryption techniques, the proposed methods have some advantages:

1) The encoding and decoding of measurements are faster when compared to data encryption methods. For a real-time PSSE, the encryption overhead could be significant.
2) An attacker would notice that the data is being encrypted, and either they can crack it to inject data or they cannot. Here the attacker might not even notice that the data had been encoded, which means you have better ability to flag the FDI.

The remainder of the paper is organized as follows. The problem is stated in more detail in Section II. Section III presents the details of the proposed solution. The application of the solution to a simulated problem is presented in Section IV. Finally, the conclusions are presented in Section V.

II. Problem Description

Let's consider the system shown in Fig. 1. In this framework, a malicious actor attacks a power system. The defense of this power system is performed by PSSE and the RTUs. The PSSE is used to estimate the states of the power system based on knowledge of the physical system (topology, parameters, etc) and measurements obtained from this system, $\mathbf{z}$. These measurements are collected from several points in the power grid by RTUs and transmitted over an insecure network connection to the state estimator. It is assumed that the RTUs are secure, i.e., it is impossible for an attacker to compromise them, and they are capable of encoding data before transmission. The PSSE is capable of decoding measurement data and is also considered secure. Both sensor encoding and decoding functions have access to a secret encoding vector, $\mathbf{w}$.

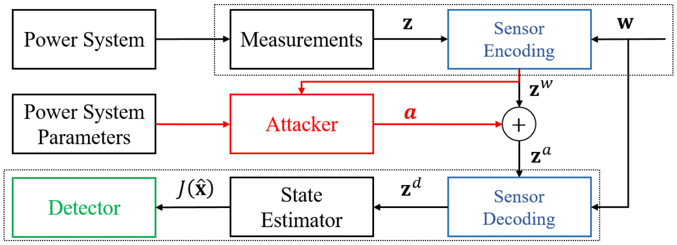

Fig. 1. Problem depiction. The attacker has access to system parameters and encoded measurements and it can change any measurement.

It is considered that a malicious actor is capable of circumventing any data protection schemes implemented within the communications system (e.g. packet encryption). Consequently, the attacker can read data-in-flight between the meters and the PSSE. Furthermore, the attacker can craft measurement data packets and impersonate the meters gathering power and voltage data from the power system. With these capabilities, the attacker performs a man-in-the-middle attack on the system. Furthermore, the attacker knows in detail the physical model of the power system. The goal of the attacker is to manipulate arbitrarily the state estimate obtained by the PSSE while remaining undetected. The attacker uses PSSE-based methods such as [8] to obtain a stealthy attack vector from manipulated meter data. It is considered that the bias the attacker will want to introduce in the state estimates remains unknown by the defender. In this paper, the goal of the defender is to detect a FDI attack to its measurements. Any response procedures following the attack detection are out of the scope of the paper.

A. Weighted Least-Squares State Estimation

In this paper, it is assumed that both attacker and defender utilize WLSSE algorithm. The problem of state estimation of steady-state condition of power systems using noisy measurements is given by a set of algebraic equations described by

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

where $\mathbf{z} \in \mathbb{R}^m$ is the vector of measurements, $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{h}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ is the vector of measurement functions and $\mathbf{e} \in \mathbb{R}^m$, is the vector of zero-mean ($\mathbb{E}[\mathbf{e}] = \mathbf{0}$), uncorrelated measurement errors with known diagonal covariance matrix ($\text{Cov}[\mathbf{e}] = \mathbf{R} \in \mathbb{R}^{m \times m}$). The states are bus voltage angles (θ) and magnitudes ($\mathbf{V}$), while the measurements are composed of real and reactive branch power flows, real and reactive bus power injections, and bus voltage magnitudes.

The solution of the WLSSE is achieved by obtaining the vector of states $\mathbf{x}^*$ that minimizes the weighted least-squares problem described by (2).

$$\min_{\mathbf{x}} J(\mathbf{x}) = \frac{1}{2}[\mathbf{z} - \mathbf{h}(\mathbf{x})]^\mathsf{T}\mathbf{W}[\mathbf{z} - \mathbf{h}(\mathbf{x})] \tag{2}$$

where the weight matrix is defined as $\mathbf{W} = \mathbf{R}^{-1}$. The optimal solution is found at the point where the first-order optimality condition, $\nabla_\mathbf{x} J(\mathbf{x}) = \mathbf{0}$, is observed. The gradient of the objective function can be defined as (3).

$$\nabla_\mathbf{x} J(\mathbf{x}) = -\mathbf{H}(\mathbf{x})^\mathsf{T}\mathbf{W}[\mathbf{z} - \mathbf{h}(\mathbf{x})] \tag{3}$$

where $\mathbf{H}(\mathbf{x}) = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$.

Due to the nonlinear relationships between system states and the measurement models of power flows and injections, the Newton-Raphson algorithm is used to obtain a solution. Starting from an initial guess $\hat{\mathbf{x}}_0$, the algorithm iteratively approaches the optimal solution $\mathbf{x}^*$ by solving successive linear approximations, (4) - (6), of the original problem (3).

$$\mathbf{h}(\hat{\mathbf{x}}_k) \approx \mathbf{h}(\hat{\mathbf{x}}_{k-1}) + \mathbf{H}_{k-1}\Delta\hat{\mathbf{x}}_{k-1} \tag{4}$$

$$\Delta\mathbf{z}_k = \mathbf{z} - \mathbf{h}(\hat{\mathbf{x}}_k) \tag{5}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \Delta\hat{\mathbf{x}}_{k-1} \tag{6}$$

where k is the iteration number and $\mathbf{H}_{k-1}$ is a shorthand notation for $\mathbf{H}(\mathbf{x}_{k-1})$. The estimate for the k-th state update vector, $\Delta\hat{\mathbf{x}}_k$, is obtained by

$$\Delta\hat{\mathbf{x}}_k = (\mathbf{H}_k^\mathsf{T}\mathbf{W}\mathbf{H}_k)^{-1}\mathbf{H}_k^\mathsf{T}\mathbf{W}\Delta\mathbf{z}_k. \tag{7}$$

The solution of the problem is obtained when the gradient (3) is sufficiently close to zero, which means that, in practice, the first-order optimality condition has been achieved. Alternatively, a very small $\Delta\mathbf{x}_k$ indicates that $\mathbf{x}^*$ has been found.

B. Bad Data Analysis

The detection of bad data in the measurement set can be achieved by the χ^2 (chi-squared) test, shown in (8). It is assumed that both attacker and defender can perform this test. Given that $2 \cdot J(\mathbf{x}^*) \sim \chi^2_{\nu,\alpha}$ where ν is the number of degrees of freedom and α is the significance level. In this case $\nu = m - n$. This statistical test aims at determining the goodness of fit of data with respect to a model. To perform the χ^2 test, it is first necessary to obtain a threshold C for the significance level.

$$P(J(\mathbf{x}) \geq C) = \alpha \tag{8}$$

α can be determined based on an acceptable level of false positives, for example. If $J(\mathbf{x}) \geq C$ is true, then it is considered that the data ($\mathbf{z}$) poorly fits the model ($\mathbf{h}(\mathbf{x}^*)$). In the context of this paper, this could signify that an attack targeting the integrity of data has been performed. Furthermore, if the data fits well the model, it is expected that the vector of residuals $\mathbf{r} \in \mathbb{R}^m$ obtained by (9) has a small magnitude.

$$\mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x}^*) \tag{9}$$

C. Stealthy False-Data Injection Attacks

The goal of the FDI attack is to introduce a bias, $\mathbf{c}$, in $\mathbf{x}^*$:

$$\mathbf{x}_c = \mathbf{x}^* + \mathbf{c}. \tag{10}$$

To achieve that, an attack vector $\mathbf{a}$ is devised to bias the measurements used as inputs to the PSSE.

$$\mathbf{z}_a = \mathbf{z} + \mathbf{a} \tag{11}$$

A stealthy FDI attack can be achieved by crafting $\mathbf{a}$ to produce a vector $\mathbf{r}$ with a magnitude smaller or equal than the one that would be obtained by solving the WLSSE for the original data $\mathbf{z}$ [8]. Such vector can be calculated by:

$$\mathbf{a} = -\mathbf{K}_c\mathbf{r} + \mathbf{h}(\mathbf{x}_c) - \mathbf{h}(\mathbf{x}^*), \tag{12}$$

$$\mathbf{K}_c = \mathbf{H}_c(\mathbf{H}_c^\mathsf{T}\mathbf{W}\mathbf{H}_c)^{-1}\mathbf{H}_c^\mathsf{T}\mathbf{W}. \tag{13}$$

where $\mathbf{H}_c = \mathbf{H}(\mathbf{x}_c)$. Since $\mathbf{K}_c$ is a projection vector, we can decompose it in two parts: a part that is orthogonal to it, $\mathbf{r}_a$ and a component that is in the column space of $\mathbf{K}_c$, $\mathbf{r}_c = \mathbf{K}_c\mathbf{r}$.

$$\mathbf{r} = \mathbf{r}_a + \mathbf{r}_c \tag{14}$$

If the measurement vector $\mathbf{z}$ has passed the χ^2 test, then we have that $\frac{1}{2}\|\mathbf{r}\|^2_\mathbf{R} < C$. Consequently, we have

$$\|\mathbf{r}_\mathbf{a}\|^2_\mathbf{R} = \|\mathbf{r}\|^2_\mathbf{R} - \|\mathbf{r}_c\|^2_\mathbf{R} \leq \|\mathbf{r}\|^2_\mathbf{R} < 2C. \tag{15}$$

which means that $\mathbf{z}_a$ it is not detected by the χ^2 test.

III. Sensor encoding

The proposed solution is inspired by symmetric cryptography, where the same encryption key is used to encrypt and decrypt plaintext. In this case, the solution involves encoding all vulnerable measurements using a secret encoding vector, $\mathbf{w} \in \mathbb{R}^m$, such that if an attacker tries to add bias and remain undetected, its attack should generate residuals large enough to trigger the anomaly detector (8).

Given $\mathbf{w}$, the encoding function $\mathbf{f}(\mathbf{z}, \mathbf{w})$ should produce an encoded vector $\mathbf{z}^w$, which can be used to recover the original measurement vector $\mathbf{z}^d = \mathbf{z}$ following the application of a decoding function $\mathbf{g}(\mathbf{z}^\mathbf{w}, \mathbf{w})$.

$$\mathbf{z}^w = \mathbf{f}(\mathbf{z}, \mathbf{w}) \tag{16a}$$

$$\mathbf{z}^d = \mathbf{g}(\mathbf{z}^w, \mathbf{w}) \tag{16b}$$

A. Naive sensor encoding (NSE)

For this application, an ideal encoding function should have the following characteristics: i) support distributed encoding of data, ii) be hard to identify, iii) induce triggering of the anomaly detector, and iv) not be easily identifiable. In order to allow distributed encoding, a simple pair of encoding and decoding functions could be implemented by additive scalar encoding and decoding pair

$$\mathbf{z}^{w} = \mathbf{z} + \mathbf{w}, \tag{17a}$$

$$\mathbf{z}^{d} = \mathbf{z}^{w} - \mathbf{w}, \tag{17b}$$

where $\mathbf{z}^d = \mathbf{z}$ must be true at the end of the decoding process.

The magnitude of $\mathbf{w}$ should be chosen such that it can trigger an anomaly detector but still generate values that are within an acceptable range and reasonably close to the original values so the attacker does not suspect an encoding method is being used. The additive encoding vector can move the solution of (2) to a point where the Jacobian matrices assume a value that is significantly different from the original solution. In that way, the encoding vector can introduce errors in $\mathbf{H}_c$ that propagate to $\mathbf{z}^a$ to generate large residuals when the $\mathbf{z}^d$ is processed. Furthermore, the encoding vector should have high entropy to reduce the chances of brute-force attacks. For these reasons the additive encoding vector was chosen as a sequence of numbers from a uniform distribution between -0.05 and 0.05 p.u., i.e, $w_i \sim \mathcal{U}(-0.05, 0.05)$, $i \in \{1, 2, \ldots, m\}$. In practice, this encoding step could be performed as part of a calibration of an instrument or by a feature of the equipment.

B. Undetectable sensor encoding (USE)

Even though the NSE approach described above could be designed to evade some data integrity checks on PSSE inputs, an attacker utilizing any BDD method would identify that there is a large mismatch between the system model and measurements. That could lead the attacker to suspect the defender is using some sort of encoding method and to devise ways to circumvent this defense. To avoid detection of encoding by the attacker, the defender can utilize an additive encoding function that uses the same strategy used to design a stealthy cyberattack [8], as defined by (10)–(13).

$$\mathbf{x}_u = \mathbf{x}^* + \mathbf{u} \tag{18a}$$

$$\mathbf{w}^u = -\mathbf{K}_u\mathbf{r} + \mathbf{h}(\mathbf{x}_u) - \mathbf{h}(\mathbf{x}^*) \tag{18b}$$

$$\mathbf{K}_u = \mathbf{H}_u(\mathbf{H}_u^{\mathsf{T}}\mathbf{W}\mathbf{H}_u)^{-1}\mathbf{H}_u^{\mathsf{T}}\mathbf{W} \tag{18c}$$

where $\mathbf{H}_u = \mathbf{H}(\mathbf{x}_u)$.

Unlike in the case where the attacker aims at manipulating the state estimate in an undetected manner, the defense approach for USE does not have the goal to obtain a predefined state estimate by the adversary. Therefore, the state bias vector $\mathbf{u}$ used to calculate the encoding vector can be chosen as a random percent deviation of the estimated state such as $u_i = x_i^* \cdot w_i^u$ where $w_i^u \sim \mathcal{U}(-0.1, 0.1)$, $i \in \{1, 2, \ldots, n\}$.

The process of obtaining an USE vector can be summarized by the following steps:

1) Receive telemetered vector of measurements $\mathbf{z}$;
2) Solve state estimation problem with $\mathbf{z}$ to obtain $\mathbf{x}^*$;
3) Add random encoding vector to $\mathbf{u}$ (18a);
4) Obtain encoding vector $\mathbf{w}^{\mathbf{u}}$ (18b);
5) Apply encoding vector to measurements.

It is important to note that, depending on how the encoding method is applied, Step 5 is unlikely to happen before a new set of measurements is obtained. Consequently, the state of the system changes due to fluctuations in load and generation, so this USE vector might become detectable at some point in the future.

IV. Case Study

The effectiveness of the proposed strategy for FDI attack detection was validated numerically. The simulations were performed on the IEEE 14-bus power flow test case using the MATLAB package MATPOWER [17]. The measurement set included all real power flows, all real and reactive power injections and all voltage magnitudes, resulting in 82 measurements and a global redundancy index (ratio between number of measurements and number of states) of 3.03. All measurements are corrupted by Gaussian noise with standard deviation of 0.01 p.u. and 0.001 p.u. for power and voltage magnitude measurements, respectively. The FDI detection statistical tests are performed with a confidence level of 99% (threshold of $\chi^2_{55,99\%}$ = 82.29).

A. Unmitigated FDI attack

Suppose the attacker wants to increase the voltage magnitude at bus 1 by 0.1 p.u. In order to do that, the attacker can obtain $\mathbf{x}^*$ by solving the state estimation problem from the original vector of measurements $\mathbf{z}$. Then, $\mathbf{x}^c$ is calculated by adding 0.1 p.u. to the state of $\mathbf{x}^*$ that corresponds to the voltage magnitude of bus 1. Then, the attack vector $\mathbf{a}$ is crafted using (12). In this example, the attack vector obtained by this method is shown in Fig. 2. It is necessary to tamper with only 11 out of 82 measurements to obtain the desired attack.

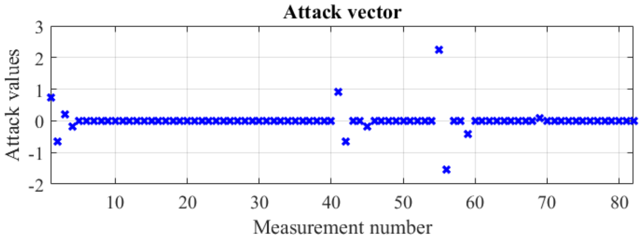

Fig. 2. Stealthy FDI additive attack vector.

The $J(\mathbf{x})$ score of the attacked PSSE is shown in Fig. 3. The state estimation process converges after 5 iterations and the $J(\mathbf{x})$ is lower than the detection threshold. Therefore, the estimator accepts this result. As a consequence, the error of the state estimator incorporates the 0.1 p.u. error introduced by the attacker, as shown in Fig. 4.

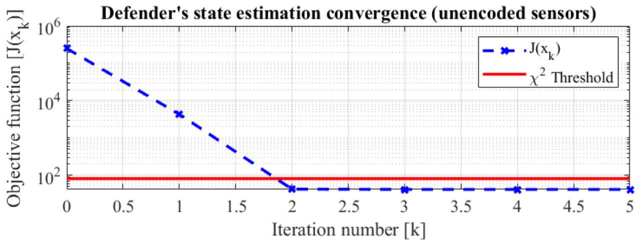

Fig. 3. Convergence of state estimation under stealthy FDI.

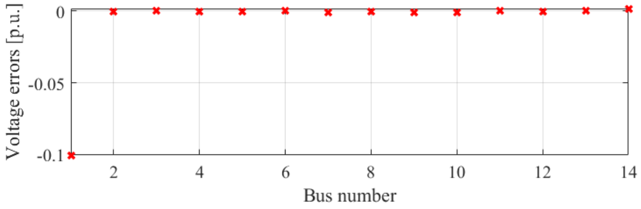

Fig. 4. Errors of state estimation when FDI is unmitigated.

B. Detecting FDI Attack with Naive Sensor Encoding

By applying the encoding method described in Section III-A we obtain the NSE vector shown in Fig. 5. When the attacker uses the encoded vectors $\mathbf{z}^w$ to generate its attack vector, the WLSSE of the attacker converges to a solution that would be rejected by the BDD approach, as shown in Fig. 6. If the attacker proceeds to use this solution to try to produce a stealthy attack, the defender's state estimator converges to a solution with high $J(\mathbf{x})$ (Fig. 7), triggering the BDD. Consequently, the PSSE solution is rejected, and the attack fails.

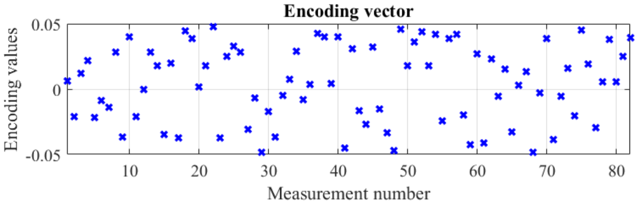

Fig. 5. The NSE vector is full with small magnitude relative to other measurements.

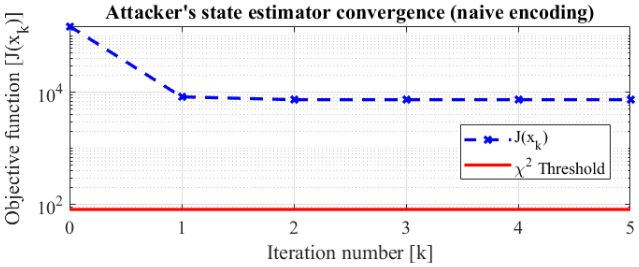

Fig. 6. The attacker's state estimator can detect an NSE vector method has been used.

C. Detecting FDI Attack with Undetectable Sensor Encoding

The encoding method described in Section III-B produces the encoding vector shown in Fig. 8. After the attacker runs its PSSE with the encoded measurement vectors it converges to a solution that is accepted by the BDD approach, as shown in Fig. 9. Unaware that the measurements are encoded, the attacker produces a stealthy attack vector and adds it to the encoded measurement vector. The defender's state estimator takes the biased and encoded measurements produced by the attacker, decodes it, and runs its PSSE. The estimator converges to a solution with high $J(\mathbf{x})$ as shown in Fig. 10, which is detected by the BDD method. Consequently, the PSSE solution is rejected, and the attack fails.

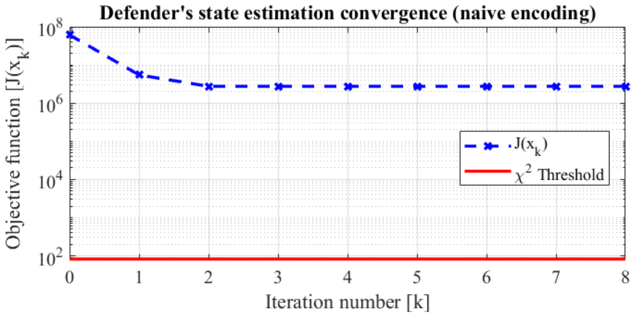

Fig. 7. The defender's state estimator can detect FDI attack when NSE is used.

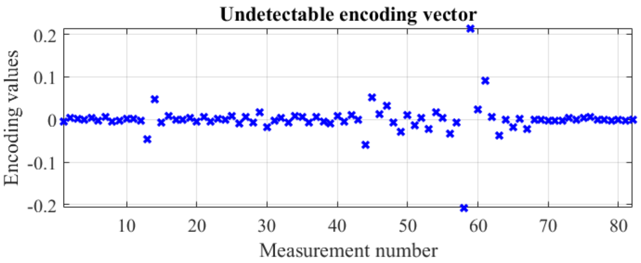

Fig. 8. The USE vector is full.

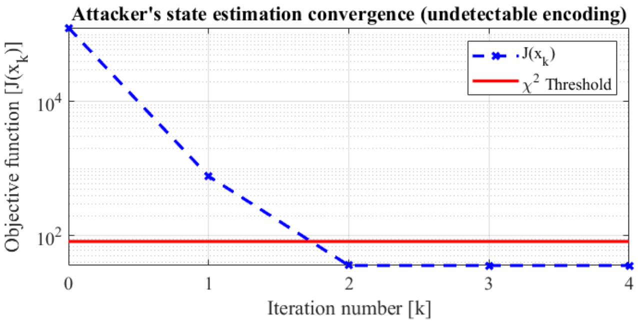

Fig. 9. Attacker's state estimator converges to an acceptable $J(\mathbf{x})$ score when USE is applied.

1) Continued Effectiveness of Undetectable Sensor Encoding: Because the USE method is dependent on the current state estimate of the system $\mathbf{x}^*$, it is not known if after several measurement scans the same result will be achieved. To determine if the effectiveness of the USE approach fades over time, 1001 sequential simulations were performed. In the first time step, the meters are not yet attacked and the defender's PSSE obtains its USE vector $\mathbf{w}^u$ and encodes all sensors. In

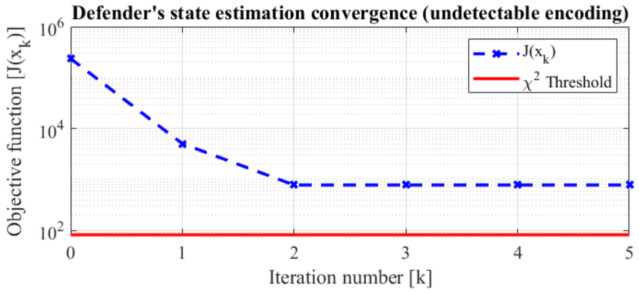

Fig. 10. The defender can detect a FDI attack when USE is used.

the following time steps, the attacker injects a new stealthy FDI attack into all sensors and the defender tries to detect the FDI using the decoding procedure and BDD method. The encoding vector obtained at the first step remains the same for all PSSE runs. The load changes over time following a random walk with a standard deviation of 20% for both active and reactive power injections in all load buses (Fig. 12).

The results shown in Fig. 11 show that the $J(\mathbf{x})$ score of the defender remains high and above the threshold, which means that it is capable of detecting the FDI attacks all the time, while the attacker's $J(\mathbf{x})$ score remains below the detectability threshold, meaning that it remains unaware of the measurement encoding defense.

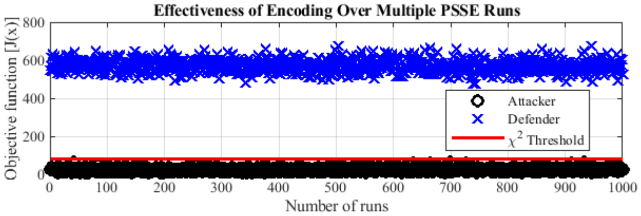

Fig. 11. The USE method is effective after 1000 PSSE runs.

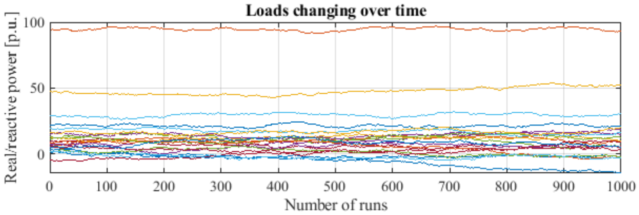

Fig. 12. Loads' real and reactive power follow a random walk process.

V. Conclusion

In this paper, we have presented two simple methods for defense against FDI attacks on PSSE using sensor encoding. The case study has demonstrated their effectiveness by inducing an attacked measurement vector to trigger the BDD method used for data integrity verification on the defender's state estimator. The proposed method has detected attacks that could circumvent traditional residual-based BDD approaches. This low-cost method could be applied to PSSEs with minimal intervention. Following a defense in depth strategy for defending the PSSE application, this method could be used with other security features such as communications encryption to provide an additional line of defense against cyberattacks.

In future work we plan to extend this approach to incorporate considerations to practical implementation and to enable its application in dynamic PSSE. Also, we plan to provide analytical demonstrations for conditions in which the encoding approaches are effective and undetectable. Furthermore, we want to extend the approach to incorporate constraints in the number of encoded measurements.

Acknowledgment

The authors would like to thank Dr. Alvaro Furlani Bastos for his insightful contributions to the work.

References

[1] A. Greenberg, "Crash override malware took down Ukraine's power grid last december," Jun 2017. [Online]. Available: https://www.wired.com/story/crash-override-malware/

[2] "CIP-002-5.1a—Cyber Security —BES Cyber System Categorization," North American Electric Reliability Corporation, Standard, Dec 2016.

[3] K. Chatterjee, V. Padmini, and S. A. Khaparde, "Review of cyber attacks on power system operations," in *2017 IEEE Region 10 Symp. (TENSYMP)*, July 2017, pp. 1–6.

[4] L. Blakely, M. J. Reno, and J. Peppanen, "Identifying common errors in distribution system models," in *2019 IEEE 46th Photovoltaic Specialists Conference (PVSC)*, 2019, pp. 3132–3139.

[5] R. D. Trevizan and M. J. Reno, "Distribution system state estimation sensitivity to errors in phase connections," in *2021 IEEE 48th Photovoltaic Specialists Conf. (PVSC)*, 2021, pp. 1049–1056.

[6] C. Francis, V. Rao, R. D. Trevizan, and M. J. Reno, "Topology identification of power distribution systems using time series of voltage measurements," in *2021 IEEE Power and Energy Conf. at Illinois (PECI)*, 2021, pp. 1–7.

[7] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. of the 16th ACM Conf. on Computer and Comm. Security*, ser. CCS '09, 2009, pp. 21–32.

[8] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks against nonlinear state estimation in smart power grids," in *2013 IEEE Power Energy Society General Meeting*, 2013, pp. 1–5.

[9] G. Dan and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *2010 First IEEE Int. Conf. on Smart Grid Communications*, Oct 2010, pp. 214–219.

[10] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1636–1646, May 2018.

[11] R. D. Trevizan, C. Ruben, K. Nagaraj, L. L. Ibukun, A. C. Starke, A. S. Bretas, J. McNair, and A. Zare, "Data-driven physics-based solution for false data injection diagnosis in smart grids," in *2019 IEEE Power Energy Society General Meeting (PESGM)*, 2019, pp. 1–5.

[12] H. Kwon and N. Nasrabadi, "Kernel rx-algorithm: a nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 388–397, 2005.

[13] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter," *IEEE Trans. Control of Network Systems*, vol. 1, no. 4, pp. 370–379, Dec 2014.

[14] S. A. Foroutan and F. R. Salmasi, "Detection of false data injection attacks against state estimation in smart grids based on a mixture gaussian distribution learning method," *IET Cyber-Physical Systems: Theory Applications*, vol. 2, no. 4, pp. 161–171, 2017.

[15] C. Kwon, W. Liu, and I. Hwang, "Security analysis for cyber-physical systems against stealthy deception attacks," in *2013 American Control Conf.*, June 2013, pp. 3344–3349.

[16] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding sensor outputs for injection attacks detection," in *53rd IEEE Conf. on Decision and Control*, Dec 2014, pp. 5776–5781.

[17] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.