ART & TECHNOLOGY PUBLISHING PTE. LTD.

Computer System Networking and Telecommunications

2023 Volume 2, Issue 1

ISSN: 2630-4635

# An Overview of False Data Injection Attack Against Cyber-Physical Power Systems

**Junhyung Bae***

School of Electronic and Electrical Engineering, Catholic University of Daegu, Republic of Korea

***Corresponding author:** Junhyung Bae, baejh80@cu.ac.kr



## Abstract

With the advancement of technology, cyber-physical systems (CPSs) are constantly being upgraded, and it resulted in more types of cyber-attacks being discovered. There are many forms of cyber-attack, and all cyber-attacks are made to manipulate the target systems. A representative system among CPSs is a cyber-physical power system (CPPS), that is, a smart grid. Smart grid is a new type of power system that provides reliable, safe, and efficient energy transmission and distribution. This paper discusses false data injection attacks targeting state estimation and energy distribution in the smart grid, along with protective strategies and dynamic monitoring for detection.

## Keywords

Cyber-physical system
Smart grid
State estimation
Bad data detection
False data injection attack

## 1. Introduction

The design of a cyber-physical system (CPS) refers to the integration of computing and communication capabilities with the monitoring and control of entities in the physical world [1]. Unlike traditional embedded systems, CPSs are physical systems that are integrated, monitored, and controlled by an intelligent computing core. Many CPSs, including smart grids, process control systems, and transportation systems, are expected to be developed using advanced computing and communication technologies. The smart grid is a typical electricity-based CPS that integrates the physical power transmission system with cyber processes in computing and communication networks.

False data injection attacks are a new and powerful class of attacks on the safety and security of CPSs [2]. The goal of the attacker is to inject false input data into the system to cause the system to make incorrect decisions without attacking the system itself. Attackers typically implement cyberattacks against sensors that measure the parameters of the physical plant of a CPS.

**Figure 1** shows a model of a CPS from a control

perspective, that is, as a control loop that controls and manages the plant. Different types of network attacks can be launched against all components and connections in the loop, but as shown in **Figure 1**, a false data injection attack is launched against the sensor. Injecting false data into the system as measurements can cause the system to make incorrect decisions and take misleading actions.

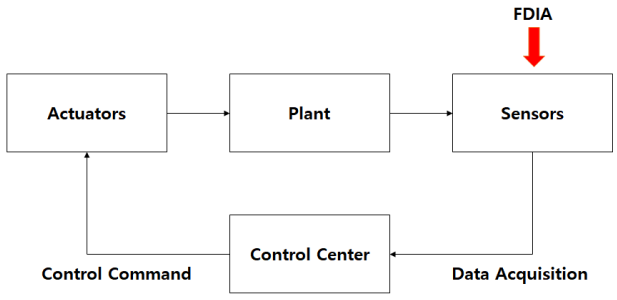


**Figure 1.** Control block diagram of CPS

Given that sensors are typically located in the field, a false data injection attack is detrimental. Attackers can gain control over sensors, and in various scenarios, particularly in measurements like temperature, pressure, and chemical concentrations, they can manipulate the interfaces connecting these sensors to the physical plant with a high degree of precision. For example, in a CPS with temperature sensors distributed at different locations, an attacker could capture temperature measurements over a specific duration and subsequently inject these recorded values into the sensor. With certain manipulations, this process could lead to a significant alteration of the actual temperature readings. In fact, the physical system could be completely out of control. A well-known attack, the Stuxnet attack, manipulated a uranium enrichment centrifuge to increase its rotational speed, but the rotational speed appeared to be within acceptable limits to the operator. As a result, the centrifuge was destroyed, which caused serious problems in plant operations for a long period of time [3,4].

This paper focuses on false data injection attacks, a type of intelligent cyberattack on state estimation in smart grids, which are cyber-physical power systems. False data injection attacks cannot be detected by traditional bad data detection algorithms and manipulate state estimation results in an arbitrary and predictable way by cooperatively modifying selected measurements. With knowledge of the system topology, an attacker can easily construct a false data injection attack by modifying only a few measurements.

This paper is organized as follows: **Section 2** introduces traditional state estimation theory; and **Section 3** describes how to construct a false data injection attack. **Section 4** describes the protection strategies and dynamic monitoring techniques for detecting false data injection attacks. Finally, conclusions and future research are presented.

## 2. Condition estimation

Since the operating conditions of the power system vary from day to day, the operators in the local control center ensure that the system remains in a safe and normal condition. Achieving this goal requires continuous monitoring of the system's condition and taking necessary preventive measures if the system is found to be unsafe. Monitoring the state of the system in real time is the first and most important step. The deployment of SCADA systems in power systems today allows control centers to collect all kinds of analog measurements and circuit breaker status information. However, the information provided by SCADA systems is not always reliable due to measurement errors, telemetry errors, communication noise, etc. and the corresponding operating state of the system cannot be directly extracted from the collected measurements [5].

The above-mentioned concerns have been addressed by a technique called state estimation of power systems. State estimation is a mathematical procedure for calculating the best estimate of the state variables of a power system. State estimation eliminates the effects of bad data and produces reliable state estimates. The state estimator output, consisting of transmission line active and reactive currents calculated from all bus voltage magnitudes and phase angles, and

busloads and generation calculated from the currents, is the basis for economic dispatch programs and what-if analysis programs.

## 2.1. Weighted least squares (WLS) state estimation

The WLS state estimation is the most widely used method due to its simplicity and low computational requirements. This state estimator adheres to the WLS criterion, which aims to minimize the weighted sum of squared measurement residuals. Consider a nonlinear measurement model $z = h(x) + e$, where $z = (z_1, z_2,...,z_m)^T$ is the measurement vector, $x = (x_1, x_2,...,x_n)^T$ is the state vector, $h(\cdot)$ is a nonlinear function relating the measurement to the state, and $e$ is the measurement error vector. Assume that the measurement error follows an independent zero-mean Gaussian distribution, i.e., $e_i$ –$N$(0,) for measurement $i$. There are $m$ measurements and $n$ states, with $n < m$. WLS state estimation can be formulated mathematically as the optimization problem $\min_x J(x) = [z - h(x)^T W[z - h(x)]$, where $J(x)$ is the weighted sum of the measurement residuals and $W$ is the covariance inverse of the measurement errors [5].

For alternating current state estimation in power systems, the state vector $x$ includes all bus voltage magnitudes and bus voltage phase angles except the reference bus. Since the relationship between state $x$ and measurement $z$ is nonlinear, an iterative technique is adopted to minimize $J(x)$. A commonly used technique is to calculate the slope of $J(x)$ and then use the Newton method to force $J(x)$ to zero.

For DC state estimation, the linear measurement model can be expressed as $z = Hx + e$, where $H$ is the measurement Jacobian matrix. The $H$-matrix is determined by the system topology and the resistance of the transmission line. At the optimal solution $= (H^T WH)^{-1} H^T Wz$, the slope of $J(x)$ vanishes, and an estimate of the measurement residual $r = z - Hx$ is given by $= [I - H(H^T WH)^{-1} H^T W]e$. In the DC tidal model, the voltage magnitude is assumed to be constant and 1 p.u. on all buses, and reactive power is completely ignored. Therefore, the state variables consist only of the voltage phase angle on all buses except the reference bus.

## 2.2. Detecting and identifying bad data

Once the state estimates and are determined, the presence of bad data can be identified by checking whether these estimates are correctly associated with the standard deviation. The $J()$ test is commonly used to detect the presence of bad measurements. It assumes that the random variable $J(x)$ follows a chi-square distribution with degrees of freedom $k = m - n$. If $J()$ is greater than a detection threshold with a determined significance level, there is sufficient reason to suspect the presence of a bad measurement. If the presence of bad data is detected, the $r^n$ test is used to identify the bad data, where $r^n$ is the vector of normalized residuals. This test is based on the fact that bad measurements produce the largest normalized residuals. Identifying and discarding bad data improves the accuracy of state estimation.

Bad data can be broadly categorized into single bad data and multiple bad data. Multiple bad data can appear in measurements where the residuals are strongly or weakly correlated. The $J(x)$ and $r^n$ tests are very effective in situations involving single bad data, multiple non-interaction bad data, and multiple interaction bad data when the bad data are non-conforming bad data that contradict each other. However, for conforming bad data with multiple interacting and non-contradictory bad data, a good measurement may have the largest normalized residual, while a bad measurement may have a small normalized residual or no residual at all. A method for distinguishing between good and bad measurements is the combinatorial optimization identification (COI) method [6]. This method is based on the fact that the Euclidean norm of the multiple normalized residuals corresponding to the bad data set is maximal. Assuming

that all meters are equally reliable, it is optimal to identify the minimum number of bad measurements. Another way to deal with multiple interactions and bad-fit data is the hypothesis testing identification (HTI) method [5]. The method first selects a set of suspect bad measurements based on their normalized residuals, assuming that the extra measurements are free of error. Hypothesis testing is then used to eliminate the list of suspicious measurements. The effectiveness of the method therefore depends on the initial selection of the set of suspicious measurements.

## 3. False data injection attacks

Smart grid state estimation plays a critical role in maintaining the reliable and economical operation of power systems. Existing state estimation approaches traditionally assume that random bad measurements can be detected. However, they have recently been shown to be vulnerable to intentional false data injection attacks. These attacks cooperatively modify measurements taken from multiple meters to skew state estimation results without being detected. As SCADA/EMS systems are increasingly connected to control center LANs, they are potentially accessible over the Internet. In addition, measurement data is often transmitted without encryption over heterogeneous SCADA communication networks consisting of fiber optic, satellite, and microwave connections. Therefore, it is clear that false data injection attacks on state estimation pose a potential security threat.

The main idea of Liu *et al.* [2] is that if the attack vector $a$ is a linear combination of the column vectors of the Jacobian matrix $H$, i.e., if $a = Hc$, then the false data injection attack cannot be detected at all. Here, $c$ can be any nonzero vector. By considering the injected attack as an addition to the measurement error, the estimated measurement residual due to the attack can be expressed as $_a = + [I – H(H^TWH)^{-1}H^TW]Hc =$ , which is exactly the same as the original measurement. Since all existing bad data detection techniques are based on measurement residuals, they cannot detect false data injection attacks at all. The state estimation solution under attack is $_a = + (H^TWH)^{-1}H^TWHc = + c$. Since $c$ can be a nonzero vector, a spurious data injection attack can manipulate the state estimation results in arbitrary and predictable ways. Furthermore, if an attacker has access to information about the power network configuration and transmission line parameters (i.e., the $H$-matrix), it is easy to construct a false data injection attack. Besides, as pointed out in Kosut *et al*.'s [7] paper, due to the sparsity of matrices in power systems, a false data injection attack only needs to modify the data by a few meters.

Indeed, fundamental limitations on the ability of state estimation to deal with cooperative bad data have long been recognized. As pointed out by Dan *et al.* [8], a spurious data injection attack can be viewed as a complete set of interacting bad data, which leads the estimated state from to $+ c$ without changing the measurement residuals. Another explanation for the success of spurious data injection attacks is provided in another paper by Kosut *et al* [9]. If is the true network state and both and $+ c$ are valid network states, then the attacker's injection vector $a = Hc$ will cause the control center to believe that the true network state is $+ c$. Since no detector can distinguish between and $+ c$, this attack vector $a$ is called an unobservable attack. Constructing a false data injection attack is equivalent to removing some meters from the network, making the network unobservable.

Liu *et al.* [2] investigated how an adversary can systematically and efficiently construct attack vectors under two realistic attack scenarios, where the attacker is either limited to a specific meter or limited in the resources required to compromise the meter. The physical limitations of the power system were not considered in the construction of the false data injection attack described above.

If an attacker fails to launch a false data injection attack due to resource limitations, they can construct an incomplete false data injection attack with a low probability of detection. Such attacks are categorized

as weak attack schemes.

Most research on false data injection attacks is based on DC state estimation. In Teixeira et al.'s [10] study, covert deception attacks were attempted on linear and nonlinear state estimators. The study of false data injection attacks on more realistic AC state estimation is much more challenging and still open to exploration.

## 5. False data injection attack protection strategy and dynamic monitoring

Ideally, the power system should be fully protected so that false data injection attacks are impossible. To achieve full protection, the operator needs to protect $n$ measurements, which are chosen so that the submatrix of $H$ over these measurements is full rank. Mathematically, the Jacobian matrix $H$ has $H_s c = 0$ if and only if $c = 0$ for the $n \times n$ non-specific matrix $H_s$. This means that if the measurements according to the submatrix $H_s$ are protected, an attack vector cannot be constructed because $H_s c = 0$ cannot be achieved. These measurements are referred to as the basic measurements, which are the minimum set of measurements required to ensure the observability of the power system. However, because the number of state variables in a system is typically large, complete protection is impractical. To address this problem, effective incomplete protection strategies have been proposed. In Dan *et al.*'s paper [8], cryptographic devices were assigned to measurements with low-security indices associated with sparse attack vectors to increase the security level of the entire power system. They also proposed an algorithm to find the lowest-cost stealth attack in a model of attack and protection costs and two greedy algorithms based on the maximum minimum attack cost and maximum average attack cost models to provide imperfect protection against false data injection attacks. Kim *et al.* [11] proposed a greedy algorithm that strategically identifies the measurements to be protected. This strategy only considers the number of measurements that are subjected to a false data injection attack and does not consider the impact of the attack on the entire power system.

Despite robust power system design, field operations are subject to breakdowns and failures due to unexpected conditions. Therefore, real-time monitoring in the field is necessary to ensure continuous safe operation through early detection of problems and rapid recovery. The designed monitor should detect all types of parameter changes and faults to cover all types of attacks and failures. Especially in the case of false data injection attacks, the monitor must observe incoming parameter measurements and detect deviations from normal operation. The wide range of parameter changes and failures due to sensor and plant faults presents a need to develop robust computational tools and methods. Existing methods for condition monitoring, a common technique in plant monitoring for such parameter changes, are limited, and early detection and diagnosis of all types of failures is virtually impossible. A good way to develop real-time monitors for robust and secure cyber-physical power systems is to extend existing condition monitors to detect cyber-attacks, especially false data injection and early failure types.

Computational tools and methods for condition monitoring of critical facilities can be categorized into two approaches: model-based and model-free [12]. Model-based approaches utilize prior information about the dynamics of the monitored system under fault-free conditions. This information is embedded in a model, such as a state space representation or equivalence formula. Model-free approaches process the raw data and represent it in the form of a non-parametric approximation, such as a neural network.

To detect parameter changes that indicate a cyberattack or failure, it is necessary to detect deviations between non-attack/faulty system behavior and behavior when there is a problem. This difference is detected through deviations that exceed a threshold. This threshold is very important to set for effective early detection as well as to avoid false alarms. The reliability of a method for detecting and isolating

attacks and faults is determined by three main factors: the accuracy of the model representing the no-attack/no-fault operation of the monitoring system, the accuracy of the estimation method for the output of the monitoring system under no-attack/no-fault conditions, and the accuracy of the statistical decision procedure and the fault thresholds used to infer the presence of faults. In most model-based fault diagnosis approaches, where linear and nonlinear observers or filters are used, the accuracy of the model is important for reliability. In model-free fault diagnosis methods, the accuracy of the model extracted from the raw data is an indication of reliability. This accuracy can be measured using model validation methods. The system may be fault-free, but the values of the parameters may be different from those used in the model.

To determine whether a model is valid, the output of a fault-free system under new operating conditions is compared to the expected output provided by the model. This is done by calculating the residuals and analyzing them statistically, i.e., by evaluating the model validity, one can identify whether these residuals exceed the predefined threshold and thus determine the need for model updating.

The accuracy of the estimation method is important in the attack/failure diagnosis process. The estimator of an attack-free/fault-free system should have minimum variance so that the effects of measurement noise are removed and the estimated system output is close to the true value. There are many different state observers or filters, but among them, the Kalman filter is known to provide the minimum variance estimate and is widely used in practice [13,14].

Compared to other state observers or filters, the Kalman filter performs better in terms of computational speed and achieves fast convergence, which makes it applicable to real-time fault diagnosis of dynamic systems. Moreover, Kalman filters can be redesigned to ensure robustness against measurement noise and model errors. Therefore, the Kalman filter can be used as a suitable estimation method for CPSs because it guarantees the optimality of the minimum variance estimate and outperforms other state observers or filters. The optimal choice of thresholds for detecting attacks and faults is important for preemptive attacks, including early failures, as well as for fault diagnosis and false alarm rates. Similar to model validation, the residual sequence is used to determine the random variable for a statistical test to detect an attack or failure. It can be seen that the elements of the residual sequence follow a zero-mean Gaussian distribution and that the sum of squares of the residual vector weighted by the reciprocal of the covariance matrix follows a chi-square distribution. The confidence interval of this distribution can be used to detect deviations between the attack-free/failure-free model and the monitored system behavior.

A simple and effective way to use a Kalman filter-based health monitor is to use a Kalman filter as a virtual sensor to identify deviations between the results of a virtual sensor that mimics the plant's sensor operation in a no-fault mode and the actual sensor measurements. Deviations outside of a threshold are considered an attack or failure. Combined with statistical decision criteria, this method is used to detect attacks on smart grid sensors. The Kalman filter is used as a virtual sensor that mimics the operation of a grid sensor in a no-fault mode, and its output is compared to the output of a real sensor to generate a residual sequence vector. The squares of this residual vector, weighted by the inverse of the covariance matrix, are composed of random variables that follow a chi-square distribution. Therefore, this variable utilizes the properties of a chi-square distribution, and a confidence interval approach can be used to define the threshold for this statistical test. If the output of the statistical test exceeds the threshold, an alarm is raised because it indicates that the sensor behavior is outside the acceptable range. Most importantly, the statistical test can be applied to a group of sensors, allowing the identification of vulnerable segments within the smart grid. Furthermore, this test can be applied to individual

sensors to pinpoint any compromised sensors.

## 5. Conclusion

In this paper, we described a false data injection attack on a representative energy-based CPS, the smart grid. Since false data injection attacks are one of the major attacks in the field of CPSs, there is a lot of research on this topic. However, most of the research is mainly focused on sensor networks and smart grids. In the future, it is necessary to study false data injection attacks in various safety-critical control systems.

## Disclosure statement

The author declares no conflict of interest

# References

[1] Son S, Park T, Won M, 2014, An Overview of Cyber Physical Systems. Telecommunications Review, 24(4): 450–459. https://www.doi.org/10.1007/978-3-030-43494-6_1

[2] Liu Y, Ning P, Reiter MK, 2009, False Data Injection Attacks against State Estimation in Electric Power Grids. ACM Transactions on Information and System Security, 14(1): 13. https://www.doi.org/10.1145/1952982.1952995

[3] Liang G, Zhao J, Luo F, et al., 2017, A Review of False Data Injection Attacks against Modern Power Systems. IEEE Trans. Smart Grid, 8(4): 1630–1638. https://www.doi.org/10.1109/TSG.2015.2495133

[4] Wang Q, Tai W, Tang Y, et al., 2018, Review of the False Data Injection Attack Against the Cyber-Physical Power System. IET Cyber-Physical Systems: Theory & Applications, 4(2): 101–107, https://www.doi.org/10.1109/TSG.2015.2495133

[5] Abur A, Exposito AG, 2004, Power System State Estimation: Theory and Implementation, CRC Press.

[6] Monticelli A, 1999, State Estimation in Electric Power Systems: A Generalized Approach, Springer.

[7] Kosut O, Jia L, Thomas R, et al., 2010, Limiting False Data Attacks on Power System State Estimation. Proceedings of the 44th Annual Conf. Information Sciences and Systems.

[8] Dan G, Sandberg H, 2010, Stealth Attacks and Protection Schemes for State Estimators in Power Systems. Proceedings of the IEEE Conf. Smart Grid Comm, 214–219.

[9] Kosut O, Jia L, Thomas R, et al., 2011, Malicious Data Attacks on the Smart Grid. IEEE Trans. Smart Grid, 2(4): 645–658, https://www.doi.org/10.1109/TSG.2011.2163807

[10] Teixeira A, Amin S, Sandberg H, et al., 2010, Cyber Security Analysis of State Estimators in Electric Power Systems. Proceedings of the 49th IEEE Conference on Decision and Control, 5991–5998.

[11] Kim TT, Poor HV, 2011, Strategic Protection Against Data Injection Attacks on Power Grids. IEEE Trans. Smart Grid, 2(2): 326–333. https://www.doi.org/10.1109/TSG.2011.2119336

[12] Rigatos G, Serpanos D, Zervos N, 2017, Detection of Attacks Against Power Grid Sensors Using Kalman Filter and Statistical Decision Making. IEEE Sensors Journal, 17(23): 7641–7648. https://www.doi.org/10.1109/JSEN.2017.2661247

[13] Rawat DB, Bajracharya C, 2015, Detection of False Data Injection Attacks in Smart Grid Communication Systems. IEEE Signal Processing Letters, 22(10): 1652–1656. https://www.doi.org/10.1109/MCOM.2015.7045410

[14] Sargolzaei A, Yazdani K, Abbaspour A, et al., 2020, Detection and Mitigation of False Data Injection Attacks in Networked Control Systems. IEEE Trans. Industrial Informatics, 16(6): 4281–4292. https://www.doi.org/10.1109/IISR.2018.8535978