HOKKAIDO UNIVERSITY

Title	Sensor Scheduling-Based Detection of False Data Injection Attacks in Power System State Estimation
Author(s)	OBATA, Sho; KOBAYASHI, Koichi; YAMASHITA, Yuh
Citation	IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E105.A(6), 1015-1019 https://doi.org/10.1587/transfun.2021EAL2098
Issue Date	2022-06-01
Doc URL	http://hdl.handle.net/2115/90183
Rights	copyright©2022 IEICE
Type	article
File Information	Sensor Scheduling-Based Detection of False Data Injection Attacks in Power System State Estimation.pdf

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

LETTER

Sensor Scheduling-Based Detection of False Data Injection Attacks in Power System State Estimation

Sho OBATA[†], *Nonmember*, **Koichi KOBAYASHI**[†a)], *and* **Yuh YAMASHITA**[†], *Members*

SUMMARY In the state estimation of steady-state power networks, a cyber attack that cannot be detected from the residual (i.e., the estimation error) is called a false data injection (FDI) attack. In this letter, to enforce the security of power networks, we propose a method of detecting an FDI attack. In the proposed method, an FDI attack is detected by randomly choosing sensors used in the state estimation. The effectiveness of the proposed method is presented by two examples including the IEEE 14-bus system.

key words: *power networks, state estimation, false data injection (FDI) attacks, random sensor scheduling*

1. Introduction

A cyber attack in control systems has attracted much attention (see, e.g., [2]–[6]). In power networks, a false data injection (FDI) attack is well known as one of the typical attacks [1]. In the state estimation of steady-state power networks, a cyber attack that cannot be detected from the residual (i.e., the estimation error) is called an FDI attack. An FDI attack can be realized by attacking multiple sensors simultaneously under the assumption that an attacker knows the structure (i.e., the sensor placement) of a given power network. There have been many previous studies (see, e.g., [7]). In previous studies, many techniques such as machine learning have been used.

In this letter, we propose a simpler method of detecting an FDI attack. In the proposed method, an FDI attack can be detected by randomly choosing sensors used in the state estimation. In other words, the sensor placement is randomly changed. As a result, an FDI attack can be detected. Thus, comparing the proposed method with the existing methods, the proposed method is simpler under the assumption that secure communications can be used for only some signals. The effectiveness of the proposed method is presented by two examples including the IEEE 14-bus system.

Notation: Let $\mathcal{R}$ denote the set of real numbers. Let I_n denote the n dimensional identity matrix. For the vector x, let $\|x\|$ denote the Euclidean norm of x.

2. Preliminaries

In this section, as preliminaries, the state estimation of power networks and FDI attacks are explained.

First, the state estimation of steady-state power networks with $n+1$ buses is explained. The bus phase angles are denoted by δ_i, $i = 1, 2, \ldots, n+1$. One (arbitrary) bus phase angle is fixed as a reference angle. In this letter, δ_1 is fixed by $\delta_1 := 0$, and therefore, only n angles have to be estimated. Let z_i, $i = 1, 2, \ldots, m$ denote the active power flow measured by the sensor i.

Assume that the phase differences $\delta_i - \delta_j$ in the power network are small. Then, since a linear approximation is accurate, we can obtain

$$z = Hx + e, \tag{1}$$

where $z = [z_1, z_2, \ldots, z_m]^\top \in \mathcal{R}^m$, $x = [\delta_2, \delta_3, \ldots, \delta_{n+1}]^\top \in \mathcal{R}^n$, and $e = [e_1, e_2, \ldots, e_m]^\top \in \mathcal{R}^m$ is the measurement noise. The noise e_i has a Gaussian distribution with zero mean and variance σ_i^2. The matrix H can be derived from a given power network (see, e.g., [3]).

For example, in the power network shown in Fig. 1, the matrix H in (1) can be obtained as

$$H = \begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{2}$$

Here, we assume that $\mathrm{rank}(H) = n$. Then, from (1), the estimated value $\hat{x}$ of x can be derived as

$$\hat{x} = (H^\top W H)^{-1} H^\top W z,$$

where $W \in \mathcal{R}^{m \times m}$ is the diagonal matrix in which diagonal elements are given by $1/\sigma_1^2, 1/\sigma_2^2, \ldots, 1/\sigma_m^2$.

Next, as a detection method of false data, we consider calculating the residual (i.e., the difference between the observed value z and the estimated value $H\hat{x}$). It is said that false data is injected if the following condition holds:

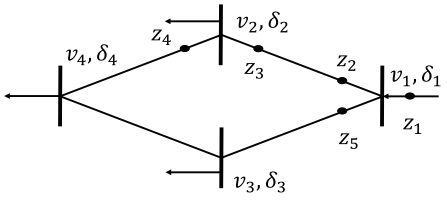

Fig. 1 Example of power networks.

Manuscript received November 8, 2021.
Manuscript publicized December 13, 2021.

[†]The authors are with the Graduate School of Information Science and Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.
a) E-mail: k-kobaya@ssi.ist.hokudai.ac.jp

DOI: 10.1587/transfun.2021EAL2098

Copyright © 2022 The Institute of Electronics, Information and Communication Engineers

$$r := \|z - H\hat{x}\| > \tau, \tag{3}$$

where τ is a given threshold.

A cyber attack that cannot be detected from the residual r is called an FDI attack. Based on [1], [3], [7], we suppose the following abilities for an attacker:

(i) An attacker knows the matrix H (i.e., the structure (the sensor placement) of a power network).
(ii) An attacker can tamper with the measured value z over a communication network, that is, an attacker can change the measured value z to $z_a = z + a$, where a is called an attack vector.

From (i), in an FDI attack, we assume that the attack vector a is generated by $a = Hc$, where c is an arbitrary vector. Then, the estimated value $\hat{x}_a$ after attack can be calculated by

$$\begin{aligned}\hat{x}_a &= (H^\top WH)^{-1}H^\top Wz_a \\ &= \hat{x} + (H^\top WH)^{-1}H^\top WHc \\ &= \hat{x} + c.\end{aligned}$$

In other words, $\hat{x}$ can be changed to $\hat{x} + c$ if an attacker utilizes $a = Hc$. However, in this case, the residual r is not changed as follows:

$$\begin{aligned}\|z_a - H\hat{x}_a\| &= \|z + Hc - H(\hat{x} + c)\| \\ &= \|z - H\hat{x}\|.\end{aligned}$$

Thus, by attacking multiple sensors simultaneously, an attacker can achieve tampering that the controller that the state estimation is performed cannot detect.

3. Proposed Detection Method

In this section, based on random sensor scheduling, we propose a method for detecting an FDI attack. The key idea of the proposed method is to randomly choose sensors used in the state estimation performed by the controller. In other words, the matrix H in (1) is changed randomly. As a result, we can detect an FDI attack. In this paper, we suppose the following defense abilities:

(i) For the information on sensors used in the state estimation, secure communications can be utilized.
(ii) For a part of the measured values, secure communications can be utilized.

We remark that since secure communications are generally expensive, minimum use of these is desirable.

3.1 Outline

First, we explain the outline of the proposed method. Figure 2 shows the proposed control system. A random binary sequence satisfying a certain condition is sent from the controller to each sensor. In Fig. 2, the binary value for each sensor is included in $D(k)$. See the next subsection for further details. The sensor i determines if the measured value z_i is sent to the controller, based on the received binary value. If the sensor i does not send the measured value z_i, then $z_i = 0$ is assigned. The controller calculates the estimated state using the received data. In addition, the controller determines if an FDI attack occurs. In the proposed method, FDI attacks can be detected from the elements of z that are not used in the state estimation.

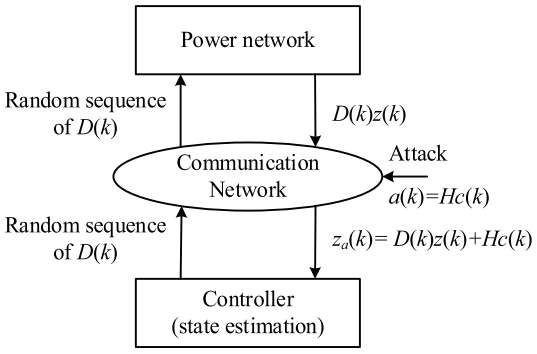

Fig. 2 Proposed control system.

3.2 Method

Now, we explain the details of the proposed method. We introduce a binary diagonal matrix $D \in \mathcal{R}^{m\times m}$ to determine sensors to be used. If z_i (i.e., the measured value of the sensor i) is used in the state estimation, then the (i,i)-th element of D is 1, otherwise it is 0. For example, consider the power network shown in Fig. 1. If we use z_1, z_3, z_4 in the state estimation, then the matrix D is given by

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A binary diagonal matrix is generated for each combination of measured values used in the state estimation. Let $D_i, i = 1, 2, \ldots, L$ denote the binary diagonal matrix representing the measured values to be used, where L is determined depending on a given power network. The following assumption is made for D_i.

Assumption 1: $\mathrm{rank}(D_iH) = n$ holds.

This assumption implies that the state can be estimated by using D_iH, instead of H. We set $\mathcal{D} := \{D_1, D_2, \ldots, D_L\}$.

Next, we propose an on-line procedure for detecting an FDI attack. At each discrete time, one is randomly chosen from the set $\mathcal{D}$. Let $D(k)$ denote the binary diagonal matrix chosen from $\mathcal{D}$ at discrete time k. Let $d_i(k)$ denote the (i,i)-th element of $D(k)$, which implies a random binary sequence generated by $D(k)$. In addition, since z, z_a, and c change in time, these are denoted by $z(k)$, $z_a(k)$, and $c(k)$, respectively. Using $D(k)$, $z_a(= z + Hc)$ and H are replaced with $z_a(k) = D(k)z(k) + Hc(k)$ and $D(k)H$, respectively. Since $D(k)$ depends on the discrete time, other variables

except for H and W also depend on the discrete time. Then, the estimated state $\tilde{x}_a(k)$ at time k can be derived by

$$\begin{aligned}\tilde{x}_a(k) &= (H^\top D(k)WD(k)H)^{-1}H^\top D(k)Wz_a(k)\\ &= (H^\top D(k)WH)^{-1}H^\top D(k)W\\ &\quad\times(D(k)z(k)+Hc(k))\\ &= \tilde{x}(k)+c(k),\\ \tilde{x}(k) &= (H^\top D(k)WH)^{-1}H^\top D(k)Wz(k).\end{aligned}$$

Based on the residual, we introduce the attack detection parameter s. The attack detection parameter s at discrete time k is defined by

$$\begin{aligned}s(k) &:= \|z_a(k)-D(k)H\tilde{x}_a(k)\|\\ &= \|D(k)z(k)+Hc(k)-D(k)H(\tilde{x}(k)+c(k))\|\\ &= \|D(k)(z(k)-H\tilde{x}(k))+(I_m-D(k))Hc(k)\|.\end{aligned} \tag{4}$$

When $D(k)=I_m$ holds, $s(k)$ is the same as r in (3), i.e., the normal residual. In (4), $D(k)(z(k)-H\tilde{x}(k))$ is the residual occurred by the sensors used. The second term $(I_m-D(k))(z(k)+Hc(k))$ is the value other than the residual, and implies the information obtained from sensors, which are not used in the state estimation. In the normal case where the power network is not attacked, this term is equal to zero. When the power network is attacked, this term is changed significantly. As a result, using $s(k)$, we can detect an FDI attack.

Finally, the proposed procedure for detecting a false data injection attack is summarized as follows.

Detection Procedure:

Step 0: In the controller, generate the set of binary diagonal matrices, $\mathcal{D}=\{D_1,D_2,\ldots,D_L\}$, where D_i satisfies Assumption 1.

Step 1: Set $p:=0$ and $t:=0$.

Step 2: In the controller, generate the random sequence of binary diagonal matrices, $D(k)$, $k=pT,pT+1,\ldots,pT+T-1$, where $D(k)\in\mathcal{D}$ and T is a given large positive integer.

Step 3: Send a random binary sequence $d_i(k)$, $k=pT,pT+1,\ldots,pT+T-1$ from the controller to the sensor i by using secure communications.

Step 4: Send the measured value $z_i(t)$ from the sensor i to the controller if $d_i(t)=1$ holds.

Step 5: Calculate $s(t)$ of (4) in the controller. Determine that an FDI attack occurs, if the following condition holds:

$$s(t)>\tau, \tag{5}$$

where τ is a given threshold.

Step 6: Update $t:=t+1$. If $t=pT$, then update $p:=p+1$ and go to Step 2, otherwise go to Step 3.

The threshold τ in Step 5 is set depending on both the noise and the value of $c(k)$ that we want to detect. Using this procedure, persistent surveillance of FDI attacks can be performed. In this procedure, secure communications are required at each T. We remark that a part of z must be sent to the controller by using secure communications. See Sect. 3.3 and Sect. 4 for further details.

Remark 1: All elements in $\mathcal{D}$ must be included in the random sequence $D(k)$. We must check if this condition is satisfied. Instead of random sequences, we may use periodic sequences. In this case, we may sometimes change periodic patterns from the viewpoint of security.

Remark 2: In general, FDI attacks can be detected by using z itself. However, z may be changed by operations. It must be determined if the change of z is caused by the operation or the attack. Furthermore, FDI attacks may not be detected by using only some elements of z (see also Sect. 4). In the proposed method, we use the elements of z that are set to zero. These elements are randomly chosen. By random sensor scheduling, we can realize detection of FDI attacks.

3.3 Relation to Security Indices

For the matrix H in (1) and the attack vector $a=Hc$, the security indices α_i, $i=1,2,\ldots,m$ are defined by $\alpha_i := \min_c \|Hc\|_0$ such that $\sum_j H_{ij}c_j=1$, where $\|Hc\|_0$ and H_{ij} denote the number of non-zero elements in the vector Hc and the (i,j)-th element of H, respectively (see, e.g., [3], [6]). As an example, consider H of (2). The security indices can be obtained as $\alpha_1=2$, $\alpha_2=3$, $\alpha_3=3$, $\alpha_4=1$, and $\alpha_5=2$. For example, $\alpha_1=2$ implies that in order to tamper z_1 with z_1+1, two elements (i.e., z_1 and other one element of z) must be tampered. From $\alpha_4=1$, we see that the measured value z_4 is critical from the viewpoint of security.

Let $\bar{i}$ denote i satisfying $\alpha_i=1$. Let $\bar{D}$ denote the diagonal matrix such that the $(\bar{i},\bar{i})$-th element is 0 and other diagonal elements are 1. Then, $z_{\bar{i}}$ must be necessarily used in the state estimation. Because $\mathrm{Ker}(\bar{D}H)\neq\{0\}$ holds, that is, $\mathrm{rank}(\bar{D}H)<n$ holds, where $\mathrm{Ker}(\bar{D}H)$ denotes the null space of $\bar{D}H$. Hence, Assumption 1 does not hold. As a result, the attack to $z_{\bar{i}}$ cannot be detected by the proposed method. For $z_{\bar{i}}$, secure communications are required. From the above discussion, we see that if $\alpha_i\geq 2$, $i=1,2,\ldots,m$ hold, any FDI attack can be detected without using secure communications for the measured value.

4. Examples

4.1 Example 1

First, we explain the problem setting. Consider the power network shown in Fig. 1. The number of the candidates of $D(k)$ is four (i.e., $\mathcal{D}=\{D_1,D_2,D_3,D_4\}$ and $L=4$). For each D_i, only one sensor is not used in the state estimation as follows:

- D_1: only z_1 is not used (i.e., z_2,z_3,z_4,z_5 are used).
- D_2: only z_2 is not used.
- D_3: only z_3 is not used.
- D_4: only z_5 is not used.

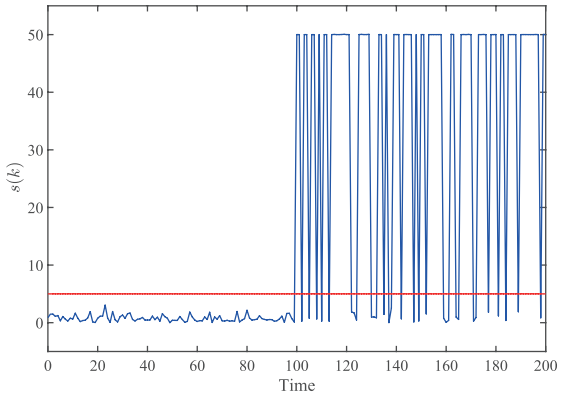

Fig. 3 Attack detection parameter $s(k)$.

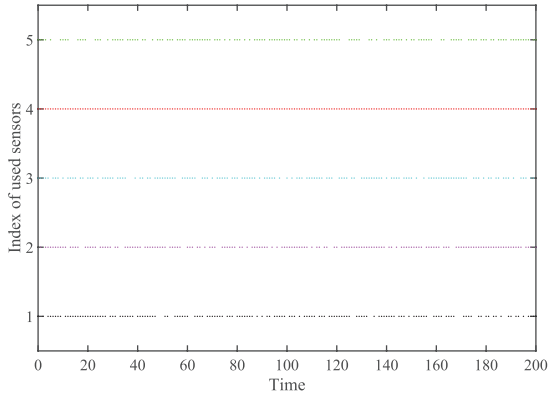

Fig. 4 Used sensors.

From discussion in Sect. 3.3, since $\alpha_4 = 1$ holds, z_4 must be necessarily used in the state estimation. Hence, it is necessary to utilize secure communication for z_4. In the detection procedure, T and τ are given by $T = 200$ and $\tau = 5$, respectively.

In the simulation, we suppose that x is fixed, and is given by $x = [-30, -20, -40]^\top$. In addition, W is given by $W = I_5$. We also suppose that the attack starts at $k = 100$, that is, the attack vector $c(k)$ is given by

$$c(k) = \begin{cases} [0,0,0]^\top & \text{if } k \le 99, \\ [50,0,50]^\top & \text{if } k \ge 100. \end{cases}$$

Next, we explain the computation result. Figures 3 and 4 show one sample using random sensor scheduling. In this simulation, when $k = 100$, the FDI attack was detected. Thus, the FDI attack can be detected quickly. From Fig. 4, we see that the used sensors are chosen randomly.

Finally, we further discuss the proposed method. In this simulation, $Hc(k)$, $k \ge 100$ is given by $Hc(k) = [-50, -50, 50, 0, 0]^\top$. Hence, $s(k)$ takes on either about 0 or about 50, except for the noise. Furthermore, if the set $\mathcal{D}$ is given by $\mathcal{D} = \{D_4\}$, then the attack cannot be detected. Because the fifth element of $Hc(k)$ is always zero. This implies that even if a part of measurements is directly utilized, then the attack may not be detected. Detection of FDI attacks can be realized using sensor scheduling.

4.2 Example 2

In the second example, we consider the IEEE 14-bus system [8]. In this system, n and m are given by $n = 13$ and $m = 23$, respectively. The matrix H is given by

$$H = \begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
5 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.$$

In this case, the security indices are obtained as $\alpha_1 = \alpha_4 = 6$, $\alpha_2 = \alpha_3 = \alpha_5 = \alpha_6 = \alpha_{17} = \alpha_{20} = \alpha_{21} = \alpha_{22} = 5$, $\alpha_7 = \alpha_{13} = \alpha_{19} = 3$, and $\alpha_8 = \alpha_9 = \alpha_{10} = \alpha_{11} = \alpha_{12} = \alpha_{14} = \alpha_{15} = \alpha_{16} = \alpha_{18} = \alpha_{23} = 2$. Since $\alpha_i \ge 2$, $i = 1, 2, \ldots, 23$ hold, any FDI attack can be detected by the proposed method with appropriate T and τ. For example, we may set $\mathcal{D}$ as $\mathcal{D} = \{D_1, D_2, \ldots, D_{23}\}$ ($L = 23$), where D_i is the diagonal matrix that the (i,i)-th element is 0, and other diagonal elements are 1. It is important to reduce the number of the candidate of $D(k)$. This is one of the future efforts.

5. Conclusion

In this letter, we proposed a new method to detect an FDI attack using random sensor scheduling. In the proposed method, it is necessary to send the information on sensor scheduling from the controller to the power network in a certain interval by using secure communications. Under this situation and the assumption on the security indices, the proposed method can detect any FDI attack.

In future work, it is important to apply the proposed method to distributed state estimation in large-scale power networks. It is also important to study the relationship between the randomness of switching frequency and the attack frequency. Developing a method to find an appropriate $\mathcal{D}$ is also significant.

This work was partly supported by JSPS KAKENHI Grant Numbers JP17K06486, JP19H02157, JP19H02158.

References

[1] Y. Liu, P. Ning, and M. Reiter, "False data injection attacks against

state estimation in electric power grids," Proc. 16th ACM Conf. on Computer and Communications Security, pp.21–32, 2009.
[2] F. Pasqualetti, F. Dörfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," IEEE Control Syst. Mag., vol.35, no.1, pp.110–127, 2015.
[3] H. Sandberg, A. Teixeira, and K.H. Johansson, "On security indices for state estimators in power networks," First Workshop on Secure Control Systems, 2010.
[4] T. Shinohara and T. Namerikawa, "On the vulnerabilities due to manipulative zero-stealthy attacks in cyber-physical systems," SICE Journal of Control, Measurement, and System Integration, vol.10, no.6, pp.563–570, 2017.
[5] T. Shinohara, T. Namerikawa, and Z. Qu, "Resilient reinforcement in secure state estimation against sensor attacks with a priori information," IEEE Trans. Autom. Control, vol.64, no.12, pp.5024–5038, 2019.
[6] A. Teixeira, K.C. Sou, H. Sandberg, and K.H. Johansson, "Secure control systems: A quantitative risk management approach," IEEE Control Syst. Mag., vol.35, no.1, pp.24–45, 2015.
[7] D.B. Unsal, T.S. Ustun, S.M.S. Hussain, and A. Onen, "Enhancing cybersecurity in smart grids: False data injection and its mitigation," Energies, vol.14, no.9, article ID 2657, 2021.
[8] https://icseg.iti.illinois.edu/ieee-14-bus-system/