2021 8th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS) | 978-1-6654-0245-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCSS53909.2021.9721954

Stealthy False Data Injection Attacks against Extended Kalman Filter Detection in Power Grids

Yifa Liu, Wenchao Xue, Shuping He, and Long Cheng*

***Abstract*—The power grid is a kind of national critical infrastructure directly affiliated to human daily life. Because of the vital functions and potentially significant losses, the power grid becomes an excellent target for many malicious attacks. Due to the special nonlinear measurements, many detection methods do not match the grid very well. The extended Kalman filter based detection is one of the few methods suitable for nonlinear system detection, and therefore can be used in power system to spot malicious attacks. However, the reliability and effectiveness of the extended Kalman filter based detection have not been fully studied and adequately guaranteed. By proposing a two-step false data injection attack strategy, this paper gives a stealthy way to inject false data of increasing magnitude into the power grid, which can eventually cause a certain degree of deviation of the grid state without being detected. In the simulation, the method proposed in this paper caused a voltage deviation of more than 25% before being discovered in the power system.**

***Index Terms*—False data injection, state estimation, extended Kalman filter, attack sequence.**

I. Introduction

Integrating modern computation, control and communication technologies, the modern power systems realize the interaction with the power transmission and distribution process through the human-computer interaction interface, and provide a way to control a power grid entity in a remote, real-time, and collaborative manner through networked space [1].

However, while improving convenience and efficiency, the power systems also introduce security problems. Remote communications and connections give opportunities for malicious attacks [2]–[4]. Due to the significant functions and instant loss upon damage, the power grid becomes a brilliant target of malicious attacks for enemies and hostile forces [5], [6]. And in reality, there are many examples of attacks on power systems [7]–[9]. Especially for the false data injection attacks, the smart grid relies on a large number of distributed meters in the field to collect measurements, which are vulnerable to attacks on the physical layer [10]–[12]. By replacing or distorting sensor information, the attacker can launch a false data injection attack against the measurements to disrupt the system operation [13]–[16], which causes a serious threat to the security of the power supply.

However, it is noted that most studies about false data injection attacks, including attack methods and detection methods mainly focus on the linear system model [17]–[22], especially those studies about attacks on the general cyber-physical systems. Even the false data injection attacks designed for the power system are mainly aimed at the linearized power system [6], [21], [22]. In actual grid operation, the (active and reactive) powers are the most important indexes and measurements for evaluating operating conditions, which are nonlinear. And on the other hand, in order to counter attacks, the power grids also adopt supervisory control and data acquisition system for anomaly monitoring and attack detection. Concretely speaking, in power system monitoring, the extended Kalman filter (EKF) is widely used for detecting malfunctions, faults, and attacks, and is placed high hopes [23]–[25]. Inappropriate attacks can be detected by the control center and then targeted measures are taken to cut off the impact of the attacks. Those facts mean that the existing attack methods cannot fully match the actual power system and may be found before causing substantial damage to the power grid.

Though the EKF raises the difficulty of attacks significantly and suppresses the effect of attacks theoretically, this paper indicates that neither the EKF nor any filter variants under the Kalman filter framework can distinguish between small false data and noises. Based on that fact, this paper first proposes a stealthy attack strategy to inject a small amount of false data, which can bypass the EKF based detection. Furthermore, when this step plan has been implemented, the false data injected continuously present in subsequent system states and estimations. Then those injected data become a springboard, with which the attacker can launch succeeding and fiercer attacks to offset the reflection of false data in the detection. Under this two-step strategy, the constructed attack sequence has increasing destructive force, and eventually interferes with the power grid and causes voltage deviations.

The main contributions are summarized as follows:

1) This paper points out the flaws of the EKF based detection and proves that the EKF can still be hijacked by false data to cause the grid operation to be disrupted.
2) This paper proposes a two-step stealthy false data injection attack strategy against the power grid with the EKF based detection, which can inject considerable amounts

Y. Liu and L. Cheng are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and are also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China.

W. Xue is with the Key Laboratory of System and Control, Academy of Mathematics and System Science, Chinese Academy of Sciences, Beijing 100190, China.

S. He is with the School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China.

All correspondences should be addressed to Prof. Long Cheng (Email: long.cheng@ia.ac.cn).

This work was supported in part by the National Natural Science Foundation of China (Grants 61633016 61873268 62025307 U1913209) and the Beijing Natural Science Foundation (Grant JQ19020).

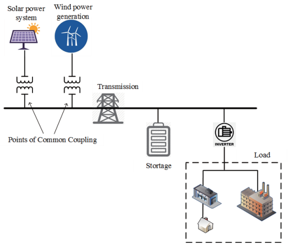


Fig. 1. Schematic diagram of the power transmission in the smart grid.

of false data into the power system to cause harm before being discovered. Compared with

II. Preliminaries

A. Power Grid and Voltage Control Model

The modern smart grid increases integration of large-scale renewable energy systems by combining these distributed generation resources. Figure 1 shows a schematic diagram of power transmission in the modern power system. Each distributed energy generator is connected into the main grid network at the point of common coupling, and distribution test feeders are interfaced to the local load through converter [26]. The main control objective is to keep the voltages at the points of common couplings V to be the reference values V_d. Therefore, the system state is defined as the deviation of the voltages from the reference value, *i.e.*, $x = V - V_d$ [27]. The desired state value x_d is the origin. Then the mathematical models of power operation and measurement can be derived as follows

$$\begin{aligned} x[k+1] &= Fx[k] + Bu[k] \\ z[k] &= h(x[k]) + v[k], \end{aligned} \tag{1}$$

where $x[i] \in \mathbb{R}^n$ is the state vector at i-th time instant. $u[k] \in \mathbb{R}^p$ is the input vector. $F \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times p}$, $(F,\ B)$ is controllable, and usually F is unstable at the origin; $z[i] \in \mathbb{R}^m, m > n$ is the measurements vector at i-th time instant. $v[i] = [v_1[k]\ v_2[k]\ \cdots\ v_m[k]]^T \sim N_m(0, R)$ is the *i. i. d* measurement noise vector. Due to the physical isolation of the sensors, the measurement noises $v_1,\ v_2\ \cdots\ v_m$ are assumed to be independent, *i.e.*, $R = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_m^2)$, $v_i \sim N(0, \sigma_i^2)$. $h(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$ is the nonlinear measurement function.

In the high-voltage power transmission systems, it is difficult and not economical to directly measure all states in the power grid. Therefore, distributed meters in the field are used to collect power flow measurements and send those data to the control center. Then the control center estimates the state through the power flow model and measurement data. Transmitted active power from the i-th bus to the j-th bus P_{ij} and the active power injection at the i-th bus P_i can be written as follows,

$$\begin{aligned} P_{ij} &= V_i^2 G_{ij} - V_i V_j [G_{ij}\cos(\theta_i - \theta_j) + B_{ij}\sin(\theta_i - \theta_j)], \\ P_i &= \sum_{j\in\mathcal{N}_i} V_i V_j [G_{ij}\cos(\theta_i - \theta_j) + B_{ij}\sin(\theta_i - \theta_j)], \end{aligned}$$

where V_i and θ_i are the voltage magnitude and phase of the i-th bus, respectively; G_{ij} and B_{ij} are the conductance and the susceptance of the transmission line between bus i and bus j, respectively; $\mathcal{N}_i$ is the set of adjacent buses of the i-th bus.

It is noted that each power flow measurement can be written as the quadratic form of voltage as follows

$$\begin{aligned} P_{ij} &= V^T H_{ij} V = (V_d + x)^T H_{ij} (V_d + x), \\ P_i &= V^T H_{ii} V = (V_d + x)^T H_{ii} (V_d + x), \end{aligned} \tag{2}$$

where$\{H_{ij}\}$are constant matrices. Then $h(\cdot)$ is a second-order function and can be written as the sum of finite Taylor series,

$$\begin{aligned} h(x) &= [I \otimes (V_d + x)^T]\mathcal{H}(V_d + x) = C + J_d x + (I \otimes x^T)\mathcal{H}x, \\ \frac{\partial h(x)}{\partial x} &= J_d + 2(I \otimes x^T)\mathcal{H}, \end{aligned} \tag{3}$$

where $J_d = \frac{\partial}{\partial x} h(x)|_{x=x_d} \in \mathbb{R}^{m\times n}$ is the Jacobian matrix of $h(\cdot)$ at the stable operation point x_d and $\mathcal{H} = [\mathcal{H}_1^T\ \mathcal{H}_2^T\ \cdots\ \mathcal{H}_m^T]^T \in \mathbb{R}^{mn\times n}$ where $\mathcal{H}_i \in \{H_{ij}\}, i = 1, \cdots, m$. $\otimes$ denotes the Kronecker product.

In the operation of power system, the feedback control adopts estimation of system state

$$u = K\hat{x}. \tag{4}$$

B. Extended Kalman Filter Based State Estimation

To acquire the estimated state $\hat{x}$ by measurements z, the EKF based state estimation methods are used. Denote $\hat{X}[t_1|t_2]$ as the state estimation at time t_1 by using measurements up to time t_2, and denote $P[t_1|t_2]$ as the covariance of the estimation $\hat{X}[t_1|t_2]$. Then the iterations of EKF based state estimation process is described as follows.

From (2), the state equation of the power grid is linear, and therefore, the classic linear Kalman filter can be applied to predict the state at the next instant. At time k, the estimated state and its covariance matrix are

$$\begin{aligned} \hat{x}[k+1|k] &= F\hat{x}[k|k] + BK\hat{x}[k|k] \\ P[k+1|k] &= (F + BK)P[k|k](F + BK)^T \end{aligned} \tag{5}$$

If the prediction process is correct, the measurement vector at time $k+1$ should be, namely the estimated measurement is

$$\begin{aligned} \hat{z}[k+1|k] &= h(\hat{x}[k+1|k]) \\ &= h(F\hat{x}[k|k] + BK\hat{x}[k|k]). \end{aligned} \tag{6}$$

Then at time $k+1$, the truly measurement vector $z[k+1]$ is collected and sent to the control center. By adding the innovation $(z[k+1] - \hat{z}[k+1|k])$, the posteriori estimation and its covariance can be derived by

$$\begin{aligned} \hat{x}[k+1|k+1] &= \hat{x}[k+1|k] + W[k+1](z[k+1] - \hat{z}[k+1|k]) \\ P[k+1|k+1] &= (I - W[k+1]\frac{\partial h(x)}{\partial x}|_{\hat{x}[k+1|k]})P[k+1|k] \\ W[k+1] &= P[k+1|k](\frac{\partial h(x)}{\partial x}|_{\hat{x}[k+1|k]})^T[(\frac{\partial h(x)}{\partial x}|_{\hat{x}[k+1|k]}) \\ &\quad \times P[k+1|k](\frac{\partial h(x)}{\partial x}|_{\hat{x}[k+1|k]})^T + R]^{-1}, \end{aligned} \tag{7}$$

where $W[i] \in \mathbb{R}^{n\times m}$ is the Kalman gain at time i.

C. False Data Injection Attack and Extended Kalman Filter Based Detection Description

When the attacker launches a false data injection attack against measurements, some measured data are tampered with the mix of false data as follows

$$z^a[k] = z[k] + A[k] = h(x[k]) + v[k] + A[k], \tag{8}$$

where $z^a[k]$ is the measurement vector after tampering at the k-th sampling instant, $A[k] \in \mathbb{R}^m$ is the malicious data added to the original measurements. Then the control center receives the false measurement data and uses them to adjust system operation incorrectly.

By replacing the original measurements $z[k]$ with tampered measurements $z^a[k]$ and taking the partial derivative of $h(\cdot)$ as a variable H, the system estimation model of the grid state and the corresponding residual after attacks are

$$\begin{aligned} x[k+1] &= Fx[k] + BK\hat{x}[k|k] \\ z^a[k] &= h(x[k]) + v[k] + A[k] \\ \hat{x}[k+1|k] &= (F+BK)\hat{x}[k|k] \\ P[k+1|k] &= (F+BK)P[k|k](F+BK)^T \\ \hat{x}[k+1|k+1] &= \hat{x}[k+1|k] + W[k+1](z[k+1]-\hat{z}[k+1|k]) \\ P[k+1|k+1] &= P[k+1|k] - W[k+1]H[k+1]P[k+1|k] \\ W[k+1] &= P[k+1|k]H^T[k+1] \\ &\quad \times (H[k+1]P[k+1|k]H^T[k+1] + R)^{-1} \\ H[k+1] &= \frac{\partial h(x)}{\partial x}|_{\hat{x}[k+1|k]} \end{aligned} \tag{9}$$

However, thanks to attack detection, the attackers cannot attack at will. When the power grid works normally, the measurement should be close to the estimation, *i.e.* $\hat{z}[k] \approx z[k]$. The residual between the measurement and the estimation is

$$\begin{aligned} r[k] &= z[k] - \hat{z}[k] \\ &= h(x[k]) + v[k] - h(\hat{x}[k|k-1]). \end{aligned} \tag{10}$$

In this paper, Euclidean norm is selected for analysis and discussion, and the results can be extended to other common norms. If the attacker tampers with the measurement as (8), the norm of the residual after the attack may become larger,

$$\begin{aligned} r^a[k] &= z^a[k] - \hat{z}[k] \\ &= z[k] + A[k] - \hat{z}[k] \\ &= r[k] + A[k], \end{aligned} \tag{11}$$

where r^a is the residual after attacks.

Then the control center can monitor the norm of the residual and compare it with a threshold τ to detect attacks. If the norm of the residual exceeds the threshold, it can be thought that an exception has occurred, *i.e.*

$$\|r^a[k]\|_2 > \tau. \tag{12}$$

Due to the existence of noises, the threshold τ cannot be set too low, otherwise, false alarms may be triggered frequently. The mechanism of the EKF based detection is shown as Fig. 2.

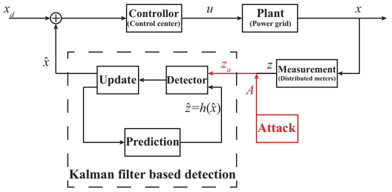


Fig. 2. An information flow diagrams of Kalman filter framework based detection.

The primary target of the attacker is to disrupt grid operation and force the control center to make the wrong schedule by tampering with the measurement data. However, modern power grids have already been equipped with anomaly detection. This fact demands the attacker to execute a stealthy attack that does not trigger the EKF based detection.

In order to achieve two objectives, the attacker should investigate the parameters and mechanism of the EKF based detection. Therefore, the following assumption is adopted:

- (A1) The attacker knows the precise parameters of the system x_d, F, B, K, $h(\cdot)$, R.
- (A2) The noise intensity is not large enough to affect the smooth operation of the system.
- (A3) When the system works in the steady state (before the attack), each random vector approximately obeys Gaussian distribution.

III. Attack against Extended Kalman Filter

To solve the main problem: that is the conflict between being detected and insufficient power, this section adopts two kinds of strategies. The one is to carry out attacks that are not easily detected, and the other is the strategy of increasing strength through continuous attacks. And the two are combined through the analysis of the system with the EKF operation mechanism. However, before launching attacks, there are things need to be prepared in advance.

A. Preparation before Attacks

In order to implement the stealthy attack, the attacker needs to know the states of the EKF, which is infeasible. Therefore, asking for second place, the attacker should at least approximate the parameters before launching attacks. In this subsection, two Kalman filter parameter estimation methods are proposed.

Most of the time, the power grid is operating near the stable point x_d in a peaceful situation, therefore, the EKF is also operating near that point. This fact reveals that the states of the EKF are stable before the attack, *i.e.*

$$\begin{aligned} P[t|t] &\xrightarrow{p} P_\infty, \\ P[t+1|t] &\xrightarrow{p} \tilde{P}_\infty, \\ W[t] &\xrightarrow{p} W_\infty,\ t \to \infty. \end{aligned} \tag{13}$$

Then the attacker can use the following two methods to get an approximate value of these parameters.

1) Combining (9) and (13) leads to the following equations

$$\begin{cases} P_\infty \approx & \tilde{P}_\infty - W_\infty J_d \tilde{P}_\infty, \\ \tilde{P}_\infty \approx & (F+BK)P_\infty(F+BK)^T, \\ W_\infty \approx & \tilde{P}_\infty J_d^T (J_d \tilde{P}_\infty J_d^T + R)^{-1}, \end{cases}$$

By solving (14), the estimated parameters can be gained.

2) Obviously, it is hard to get the solution to (14). Since the EKF is achieved through iteration, its states converge iteratively. Then the attacker can use the following iterative algorithm to acquire those parameters.

Algorithm 1: Speculation over EKF states information

```
**Initialization**: Initialize $P_0^a = \mathbf{0}$.
**while** the stopping criteria is not satisfied **do**
    $\tilde{P}_{t+1}^a = (F+BK)P_t^a(F+BK)^T$.
    $W_{t+1}^a = \tilde{P}_{t+1}^a J_d^T (J_d \tilde{P}_{t+1}^a J_d^T + R)^{-1}$.
    $P_{t+1}^a = \tilde{P}_{t+1}^a - W_{t+1}^a J_d \tilde{P}_{t+1}^a$.
    $t \leftarrow t+1$.
**end while**
Acquire $P_\infty^a = P_t^a,\ \tilde{P}_\infty^a = \tilde{P}_t^a,\ W_\infty^a = W_t^a$
 as the recurrence of $P_\infty,\ \tilde{P}_\infty,\ W_\infty$, respectively.
```

B. Pre-attack Strategy

The design of the first attack is hard, which faces the detection head-on. Therefore, the goal of the first attack is not to cause sufficient damage immediately. Since the basic operating principle of the power system is known, this subsection first provides a specific attack method which is suitable for power system measurement.

Denote s as a large time when the system has already reached the steady state, then this subsection discusses the process after time s. Supposing the attacks start at time $s+1$, *i.e.* $A[t] \equiv 0,\ t \leq s$. From (10), the residual is mainly determined by noises. By applying (3) into (11), it can be obtained that

$$\begin{aligned} & r^a[s+1] \\ = & [I \otimes (V_d + Fx[s] + BK\hat{x}[s|s] + w[s])^T]\mathcal{H} \\ & \times (V_d + Fx[s] + BK\hat{x}[s|s] + w[s]) + v[s+1] + A[s+1] \\ & -[I \otimes (V_d + (F+BK)\hat{x}[s|s])^T]\mathcal{H}(V_d + (F+BK)\hat{x}[s|s]). \end{aligned}$$

As the expectation of the quadratic form of noise is not zero, the expectation of residual can be derived by

$$\begin{aligned} & E(r^a[s+1]) \\ = & E\{[I \otimes (V_d + Fx[s] + BK\hat{x}[s|s])^T]\mathcal{H}(V_d + Fx[s] + BK\hat{x}[s|s])\} \\ & -E\{[I \otimes (V_d + (F+BK)\hat{x}[s|s])^T]\mathcal{H}(V_d + (F+BK)\hat{x}[s|s])\} \\ & +A[s+1] \\ = & E\{[I \otimes (Fx[s] - F\hat{x}[s|s])^T]\mathcal{H}(V_d + (F+BK)\hat{x}[s|s])\} \\ & +E\{[I \otimes (V_d + (F+BK)\hat{x}[s|s])^T]\mathcal{H}(Fx[s] - F\hat{x}[s|s])\} \\ & +E\{[I \otimes (Fx[s] - F\hat{x}[s|s])^T]\mathcal{H}(Fx[s] - F\hat{x}[s|s])\} + A[s+1] \\ \approx & E\{[I \otimes (Fx[s] - F\hat{x}[s|s])^T]\mathcal{H}(Fx[s] - F\hat{x}[s|s])\} + A[s+1] \end{aligned} \quad (14)$$

Therefore, the first attack vector should be designed as

$$\begin{aligned} A[s+1] &= -E\{(Fx[s] - F\hat{x}[s|s])\mathcal{H}(Fx[s] - F\hat{x}[s|s])\} \\ &= tr(FP_\infty F)[\mathcal{H}_1\ \mathcal{H}_2\ \cdots\ \mathcal{H}_m]^T. \end{aligned} \quad (15)$$

C. Post-attack Strategy

After the grid and its EKF suffer attacks, the false data are injected and continuously affect grid operation. The mismatch between false data content in the estimation and that in the measurement gives the attacker the opportunity to launch more and more fierce attacks. In order to reduce the change of the innovation caused by the attacks, the attacker needs the current state of the grid. Therefore, the attacker should estimate the system state based on the system evolutionary equation and previous attack data to design stealthy attacks.

Denote X^r and $\hat{X}^r$ as the reconstruction of system state X^a and estimation $\hat{X}^a$ during attacks by the attacker, respectively. According to system (9), X^r and $\hat{X}^r$ should obey the following dynamics

$$X^r[s+k+1] = FX^r[s+k] + BK\hat{X}^r[s+k] \quad (16)$$

$$\begin{aligned} \hat{X}^r[s+k+1] = & (F+BK)\hat{X}^r[s+k] + W^r[s+k+1] \\ & \times [h(FX^r[s+k] + BK\hat{X}^r[s+k]) \\ & + A[s+k+1] - h((F+BK)\hat{X}^r[s+k])] \end{aligned} \quad (17)$$

$$H^r[s+k+1] = \frac{\partial h(x)}{\partial x}|_{(F+BK)\hat{X}^r[s+k]} \quad (18)$$

$$\begin{aligned} P^r[s+k+1] = & (I - W^r[s+k+1]H^r[s+k+1]) \\ & \times [(F+BK)P^r[s+k](F+BK)^T] \end{aligned} \quad (19)$$

$$\begin{aligned} W^r[s+k+1] = & (F+BK)P^r[s+k](F+BK)^T H^{rT}[s+k+1] \\ & \times [H^r[s+k+1](F+BK)P^r[s+k](F+BK)^T \\ & \times H^{rT}[s+k+1] + R]^{-1}, \end{aligned} \quad (20)$$

where $W^a,\ P^a,\ H^a$ are the attacker's recurrence of $W,\ P,\ H$, respectively. Because it is impossible for the attacker to get the noise information, and it is also nonsense to simulate noise, therefore, there is no noise term $V[t]$ in (20). The initial values are

$$\begin{aligned} P^r[s+1] = P_\infty, \qquad H^r[s+1] = J_d, \\ W^r[s+1] = P_\infty J_d (J_d P_\infty J_d^T + R)^{-1}. \end{aligned} \quad (21)$$

Since the initial value of the system $X[s]$ is close to $\mathbf{0}$, the initial values are also set as $X^r[s] = \mathbf{0},\ \hat{X}^r[s] = \mathbf{0}$.

To achieve the final goal which is making the attacks stealthy to avoid detection, the attacker should ensure the residual between measurement and estimation does not exceed the threshold value. As long as the attacker maintains that the residuals after attacks mainly come from noises, it can make the system look like it is not compromised, *i.e.*

$$E(r^a|_{\{A[s+k]\}} - r|_{\{A[s+k]\equiv 0\}}) \approx 0. \quad (22)$$

The impact of the attacks on the residual can be derived by

$$\begin{aligned} & E(r^a[s+k] - r[s+k]) \\ = & E[h(x[s+k]) + v[s+k] + A[s+k] - h(\hat{x}[s+k|s+k-1])] \\ & -E[h(x[s+k]|_{X^r[s+k]=\mathbf{0}}) + v[s+k] \\ & -h(\hat{x}[s+k|s+k-1]|_{\hat{X}^r[s+k-1]=\mathbf{0}})] \\ \approx & h(X^r[s+k]) + A[s+k] \\ & -h((F+BK)(\hat{X}^r[s+k-1])). \end{aligned} \quad (23)$$

To realize (22) as far as possible, the attacker should use accessible information to satisfy the following conditions

$$\mathbf{0} = h(X^r[s{+}k]){+}A[s{+}k]{-}h((F{+}BK)(\hat{X}^r[s{+}k{-}1])).$$

And finally, the attacker should adopt the following strategy,

$$A[s{+}k] = h((F{+}BK)\hat{X}^r[s{+}k{-}1]) - h(X^r[s{+}k]).(24)$$

In conclusion, the procedure to attack extended Kalman filter based estimation and detection can be summarized into the following algorithm,

Algorithm 2: Implement stealthy attacks on the power system with EKF for the attacker

```
Estimate the parameters of EKF during steady-state
 operation as Algorithm 1.
Carry out the first attack according to (15).
The attacker simulates the state values of the system
 after the first attack, namely $X^r$, $\hat{X}^r$, $H^a$, $P^a$, and $W^a$,
 and finishes the initialization as (21).
Calculate the deviations of the estimation at the time
 the first attack occurred $\hat{X}^r[s{+}1] = W^r[s{+}1]A[s{+}1]$.
$k = 1$.
while(1) do
    Update the system equation parameters $H^r[s{+}k{+}1]$,
     $P^r[s{+}k{+}1]$ and $W^r[s{+}k{+}1]$, according to the
     current system state as (18)-(20).
    Calculate the deviations of the estimation and the
     system state $X^r$, $\hat{X}^r$ as (16) and (17), respectively.
    Design subsequent attack $A[s{+}k{+}1]$ as (24).
    $k \leftarrow k{+}1$.
end while
```

IV. Numerical Simulation

This section verifies the effectiveness of the proposed false data injection attack strategy against EKF based state estimation, namely, whether the attacks can bypass the detection of EKF while disrupting the grid operation.

Consider a 4-bus model of the distribution test feeders proposed by the IEEE distribution test feeder working group [28], which is shown in Fig. 3. The admittance matrix Y of this power network is given in (25).

Each distributed energy resource is connected into the smart grid at the point of common coupling, and it needs to control the voltages ($V = [V_1 \; V_2 \; V_3 \; V_4]^T$) of those points of common coupling to keep them at the reference values $V_d = [12.470 \;\; 7.123 \;\; 2.258 \;\; 1.987]^T$(KV). The power grid adopts voltage regulation control, and the phase angles are set fixed $\theta = [0 \;\; -0.333 \;\; -3.600 \;\; -8.833]^{T\circ}$. From [27], the dynamics of the system state can be written by

$$\dot{x} = fx + bu, \qquad (26)$$

where the state transition matrix f and input matrix b are shown as follows

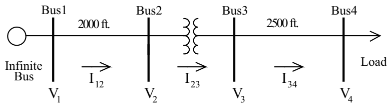


Fig. 3. An illustration of the IEEE 4-bus test feeder model.

$$f = \begin{bmatrix} 0.1759 & 0.1768 & 0.5110 & 1.0360 \\ -0.3500 & 0 & 0 & 0 \\ -0.5442 & -0.4748 & -0.4088 & -0.8288 \\ -0.1197 & -0.5546 & -0.9688 & -1.0775 \end{bmatrix} \times 10^3,$$

$$b = \begin{bmatrix} 0.0008 & 0.3342 & 0.5251 & -1.0360 \\ -0.3500 & 0 & 0 & 0 \\ -0.0693 & -0.0661 & -0.4201 & -0.8288 \\ -0.4349 & -0.4142 & -0.1087 & -1.0775 \end{bmatrix} \times 10^3.$$

By applying the Euler formula, the sampled-data based forml can be expressed by the following linear state-space model

$$x[k{+}1] = Fx[k] + Bu[k],$$

where $F{=}\exp(f\triangle t) \approx (I{+}f\triangle t), \; B = \int_0^{\triangle t} \exp(f\tau) b d\tau \approx b \triangle t$. The sampling period $\triangle t$ is set to be 10ms. A linear feedback control is adopted as $u[k] = -K\hat{x}[k]$, where

$$K = \begin{bmatrix} 1.0000 & -1.4286 & -0.0000 & -0.0000 \\ -0.8052 & 1.5995 & 2.4960 & 0.7234 \\ 1.5276 & 0.2175 & -0.6126 & 1.6858 \\ -0.1371 & 0.4545 & 0.0014 & 0.0878 \end{bmatrix}.$$

The measurements are based on the sensors in branch and bus power flows which are the same with (2), the measurement vector is selected as $z = [P_1 \; P_2 \; P_3 \; P_4 \; P_{12} \; P_{23} \; P_{34}]^T = h(x) + E$, where $E = [e_1 \; e_2 \; e_3 \; e_4 \; e_{12} \; e_{23} \; e_{34}]^T \sim N_7\big(0, \; diag \; (1.7050, 5.2910, 2.2502, 6.2500, 2.0610, 4.1305, 6.2500)\big)$ is the noise vector.

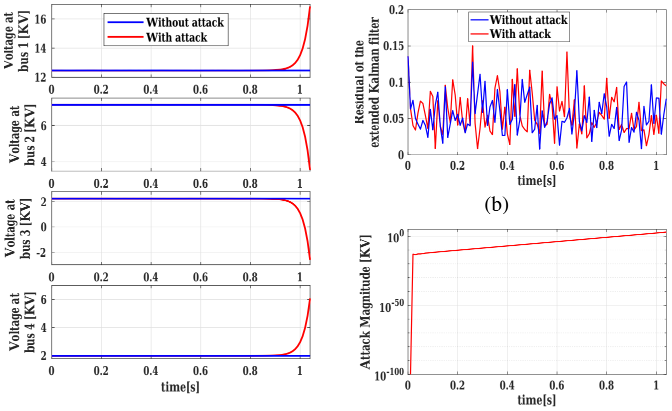
When the attacker launches a false data injection attack proposed in this paper, the first attack is designed as $A[1] = [1\,1\,1\,1\,1\,1\,1] \times 10^{-99}$, which is certainly small enough to be not spotted. Then the subsequent attacks are obtained by Algorithm 2.

Figure 4 shows the evolutions of the states of the power grid with the EKF based detection in cases of the normal operation and the attacked situation. From Fig. 4(c), the attack magnitude soars rapidly. While from Fig. 4(b), the residual given by EKF is quite normal and has almost no difference from the case without attack. However, this increasing attack sequence can eventually cause the bus voltages to significantly deviate from the stable operating point as shown in Fig. 4(a). This experiment can demonstrate that the attack strategy proposed in this paper can cheat the EKF based detector in the power grid successfully.

V. Conclusion

This paper focuses on a kind of typical cyber-physical system with nonlinear measurements: the power systems, and proposes a specific stealthy attack sequence construction

$$Y(s)=\begin{bmatrix} \frac{1}{0.1750+0.0005s} & -\frac{1}{0.1750+0.0005s} & 0 & 0 \\ -\frac{1}{0.1750+0.0005s} & \frac{1}{0.1750+0.0005s}+\frac{1}{0.1667+0.0004s} & -\frac{1}{0.1667+0.0004s} & 0 \\ 0 & -\frac{1}{0.1667+0.0004s} & \frac{1}{0.1667+0.0004s}+\frac{1}{0.2187+0.0006s} & -\frac{1}{0.2187+0.0006s} \\ 0 & 0 & -\frac{1}{0.2187+0.0006s} & \frac{1}{0.2187+0.0006s}+\frac{1}{12.3413+0.0148s} \end{bmatrix}. \tag{25}$$



(a)

(b)

(c)

Fig. 4. The attack strategy proposed in this paper cheats the extended Kalman filter based detection: (a) voltage on each bus fluctuates sharply; (b) there is no significant difference in extended Kalman filter detection indicators with and without attack; (c) attack magnitude is growing rapidly.

method. Based on this attack strategy, the attacker can launch a stealthy false data injection attack to cause the grid voltage to jump, while bypassing the detection of the extended Kalman filter. It is proved that only an extended Kalman filter cannot guarantee the safety of power systems.

REFERENCES

[1] G. Lu, D. De, and W. Song, "SmartGridLab: a laboratory-based smart grid testbed," in *Proceedings of the First IEEE International Conference on Smart Grid Communications*, Gaithersburg, USA, 2010, pp. 143-148.

[2] F. Li and Y. Tang, "False data injection attack for cyber-physical systems with resource constraint," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 729-738, 2020.

[3] Y. Mo and B. Sinopoli, "False data injection attacks in control systems," *Preprints of the 1st Workshop on Secure Control Systems*, 2010.

[4] A. Lu and G. Yang, "Malicious attacks on state estimation against distributed control systems," *IEEE Transactions on Automatic Control* vol. 14, no. 8, pp. 1-8, 2015.

[5] M. Jin, J. Lavaei, and K. H. Johansson, "Power grid AC-based state estimation: vulnerability analysis against cyber attacks," *IEEE Transactions on Automatic Control*, vol. 64, no. 5, pp. 1784-1799, 2019.

[6] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 13, 2011.

[7] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317-3318, 2017.

[8] J. Wang and L. Rong, "Cascade-based attack vulnerability on the US power grid," *Safety Science*, vol. 47, no. 10, pp.1332-1336, 2009.

[9] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," *Workshop on Future Directions in Cyber-physical Systems Security*, 2009.

[10] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645-658, 2011.

[11] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362-1370, 2012.

[12] R. Deng, P. Zhuang, and H. Liang, "False data injection attacks against state estimation in power distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2871-2881, 2019.

[13] G. Wu, J. Sun, and J. Chen, "Optimal data injection attacks in cyber-physical systems," *IEEE Transactions on Cybernetics*, vol. 48, no. 12, pp. 3302-3312, 2018.

[14] L. Guo, H. Yu, and F. Hao, "Optimal allocation of false data injection attacks for networked control systems with two communication channels," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 2-14, 2021.

[15] Z. Guo, D. Shi, K. H. Johansson, and L. Shi, "Worst-case stealthy innovation-based linear attack on remote state estimation," *Automatica*, vol. 89, pp. 117-124, 2018.

[16] C. Zhao, J. He, P. Cheng, and J. Chen, "Analysis of consensus-based distributed economic dispatch under stealthy attacks," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 5107-5117, 2017.

[17] Z. Pang, L. Fan, J. Sun, K. Liu, adn G. Liu, "Detection of stealthy false data injection attacks against networked control systems via active data modification," *Information Sciences*, vol. 546, pp. 192-205, 2021.

[18] M. Ghaderi, K. Gheitasi, and W. Lucia, "A blended active detection strategy for false data injection attacks in cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 168-176, 2021.

[19] A. Sargolzaei, K. Yazdani, A. Abbaspour, C. D. Crane III and W. E. Dixon, "Detection and mitigation of false data injection attacks in networked control systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4281-4292, 2020.

[20] A. Chattopadhyay and U. Mitra, "Security against false data-injection attack in cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 2, pp. 1015-1027, 2020.

[21] Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao, "On false data-injection attacks against power system state estimation: modeling and countermeasures," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 717-729, 2014.

[22] R. Deng and H. Liang, "False data injection attacks with limited susceptance information and new countermeasures in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1619-1628, 2019.

[23] 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 Transactions on Control of Network Systems*, vol. 1, no. 4, pp. 370-379, 2014.

[24] C. Bai, V. Gupta, and F. Pasqualetti, "On Kalman filtering with compromised sensors: attack stealthiness and performance bounds," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6641-6648, 2017.

[25] J. Sawodny, O. Riedel, and T. Namerikawa, "Detection of attacks in smart grids via extended Kalman filter and correlation analysis," in *Proceedings of 2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan*, Thailand, 2020, pp. 663-669.

[26] H. Li, F. Li, Y. Xu, D. T. Rizy, and J. Kueck, "Adaptive voltage control with distributed energy resources: algorithm, theoretical analysis, simulation, and field test verification," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp.1638-1647, 2010.

[27] H. Li, L. Lai, and H. V. Poor, "Multicast routing for decentralized control of cyber physical systems with an application in smart grid," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1097-1107, 2012.

[28] IEEE PES AMPS DSAS Test Feeder Working Group, Available: https://site.ieee.org/pes-testfeeders/.