Research on FDI Attacks in Edge Computing Environment

Qing Li¹, Aidong Xu², Haojie Lin¹, Wenjing Hou¹, Yixin Jiang², Yunan Zhang², Zhongqi Mao¹ and Hong Wen¹

¹School of Aeronautics and Astronautics, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China
²Electric Power Research Institute, China Southern Power Grid Co., Ltd. Guangzhou, China

Email: sunlike@uestc.edu.cn; turbolq1004@163.com

Abstract. The False Data Injection (FDI) attack in the smart grid can bypass the bad data detection mechanism and make the control center make an incorrect estimate of the system status. The smart grid under the edge computing architecture is more likely to be maliciously attacked by the terminal because it is close to the terminal. For this reason, this paper proposes a method combining PCA and ReliefF algorithm to achieve the purpose of dimensionality reduction. It can solve the overfitting problem caused by the high dimensionality of traditional machine learning attack signal detection, and it is suitable for terminals with low computing power.

1. Introduction

The smart grid under the edge computing (EC) architecture has the advantages of low latency, lightweight, high security and high reliability, and can provide computing and storage resources. Using advanced digital information and communication technology, online monitoring and real-time information control of important operating parameters of all links of the power grid became available under enables that the big data analysis can be integrated to achieve a more environmentally friendly, safer and more efficient Power management [1], but the system is also facing potential network attack risks [2]. Due to nearby terminals, malicious attacks on terminals will be easier in the EC system. This new network architecture brings new security protection requirements. False data injection attacks can tamper with the measurement information collected by the supervisory control and data acquisition (SCADA) system, so it can also tamper with the measurement data collected by the smart grid. If the attacker knows the topology of the system, he can construct a hidden FDI attack vector without changing the measurement residuals to attack the system. The state estimation has an impact. The traditional bad data detection (BDD) method cannot detect hidden and false data intrusion attacks. This will affect the important decisions of the power grid and pose a security threat to the state estimation of the smart grid.

The machine learning methods have made some progress in detecting hidden FDI attacks. But they are all obtained on the premise that the training set and the test set are highly similar. These methods are likely to have poor learning effects [3]. Moreover, the power system is often highly complex, and the dimensionality of its historical data is often hundreds or even thousands of dimensions. Therefore, in the use of machine learning methods to detect hidden FDI attacks, it is particularly important to avoid overfitting of training results through dimensionality reduction and reduce model training time[4]. Through the dimensionality reduction method, the measurement data is pre-processed by dimensionality reduction, and then used to train the neural network to obtain a suitable detection model, which is of great significance to realize real-time and efficient FDI attack detection [5].
innovation of this paper is that the PCA and ReliefF algorithms are used to reduce the dimensionality of the measured data. Also the EC provides the resources to process the PCA and ReliefF algorithms.

2. FDI Attacks Scheme for the Power System

In power system state estimation, the types of measurement data include: node injected active and reactive power, branch active and reactive power, and node voltage amplitude [6]. Due to the existence of measurement noise, there will be a small amount of bad data in the measurement data, which will reduce the accuracy of state estimation. When there are many bad data, it may even cause non-convergence of state estimation. Therefore, in practice, the identification of bad data will be added on the basis of state estimation. The commonly used bad data identification method is mainly based on residual detection [7]. To ensure the success of the attack, it is necessary to make the attack vector constructed based on nonlinear state estimation able to avoid the identification of bad data. Ideally, the attacker initiates a false data injection attack, inputs the attack vector into the original measurement vector, and obtains a new measurement vector. The measurement vector received by the state estimation program is actually the new measurement vector after the attack. The measurement of [8] has already deviated from the actual measurement, and the state estimation value obtained from this will also deviate [8]. After the attack, as long as the residual error is less than or equal to the bad data detection threshold, the false data injection attack can avoid the bad data identification link and achieve the purpose of successfully tampering with the state estimation result [9].

The empirical rule of using part of the network information is to construct the FDIA injection vector to satisfy Kirchhoff’s circuit law, and only the selected voltage angle difference is given [10]. Specifically, the voltage angle difference of the line connecting the damaged bus and the non-damaged bus is used to calculate the feasible power flow of the respective lines, so the power injection of the damaged bus is obtained. For other buses, it can be calculated by the algebraic sum of all connected buses. Using the above ideas, the following attack vectors can be constructed.

1) Initialize the system state vector $[V_0^{\theta}] = [V_0^P, V_0^Q, \theta_0]$, where $V_0$ is the initial attack voltage distribution.

2) Use the current $[V\theta]^{T}$ to calculate the attack vector $[PQpq]^{T}$ (bus real/reactive power injection and line flow).

3) Check the constructed attack vector based on the real power bus injection and the upper and lower bounds of real/reactive power flow.

If all the boundaries are satisfied, $[V\theta]^{T}$ is used as the attack vector. Otherwise continue.

4) Calculate the incremental state vector $[\Delta V\Delta \theta]^{T}$ by solving the optimization problem

$$\text{minimize} \sum_{i=1}^{10} l_i^{T} S_i$$

Subordinate to

$$
\begin{bmatrix}
\Delta P \\
\Delta Q \\
\Delta P \\
\Delta Q \\
\Delta P \\
\Delta Q \\
\Delta V \\
\Delta \theta
\end{bmatrix} = 
\begin{bmatrix}
\partial P/\partial V & \partial P/\partial \theta \\
\partial Q/\partial V & \partial Q/\partial \theta \\
\partial P/\partial \theta & \partial p/\partial \theta \\
\partial Q/\partial \theta & \partial q/\partial \theta \\
0 & 1
\end{bmatrix} 
\begin{bmatrix}
\Delta V \\
\Delta \theta
\end{bmatrix}
$$

(2)

5) Update the attack vector $[V\theta]^{T} \leftarrow [V\theta]^{T} + [\Delta V\Delta \theta]^{T}$ and go to step 2.

By repeating this process, can construct an attack vector for AC state estimation [11], but with the continuous expansion of the power grid, the dimensionality of the measurement data has also doubled,
which in turn leads to the challenge of dimensionality disasters in machine learning detection methods. So that the training results are at risk of over-fitting, so dimensionality reduction is required in advance [12].

3. Performances
Firstly, the voltage phase angle attack is simulated, and the ordinary attack method is used to destroy the data based on the time seed random number, so that its authenticity is lost, which can affect the system state estimation [13].

Then attack the voltage phase angle according to the hidden attack in the technical route, and get a comparison chart with the original voltage phase angle, and finally verify that the ordinary residual detection method cannot detect the hidden data attack. From the perspective of the BDD detection mechanism, if the attacker knows the topology of the system, he can construct a hidden FDI attack vector without changing the measurement residuals, which will affect the system state estimation. Therefore, the traditional BDD detection method cannot detect hidden falsehoods. Data intrusion attacks, experimental results verify this.

**Figure 1.** The model of FDI attack under the EC

**Figure 2.** Contrast diagram of concealed attack and original voltage phase angle.
Aiming at the problems of long training time and high computational complexity for high-dimensional data, literature [14] proposed a combination of PCA and ReliefF algorithm (PCA-RF) to reduce the dimensionality of the data, but the PCA algorithm is an unsupervised reduction. In the dimensionality algorithm, there is a risk of deleting features that are beneficial to classification in the process of dimensionality reduction, thereby reducing the classification recognition rate. Therefore, a principal component analysis algorithm based on multi-resolution analysis and ReliefF (MulRF-PCA) [15] can be used. First, use multi-resolution analysis to extract the features of the current model, and then use ReliefF algorithm to assign weights to each feature. Delete the unfavorable features for classification, and finally perform dimensionality reduction and decorrelation through the PCA algorithm to achieve better recognition performance under lower complexity [16].

4. Conclusion
Taking edge computing as the application background, this paper conducts preliminary research on false data attacks, which is a new idea in the field of edge computing. Attack signal detection methods that can reduce traditional machine learning have over-fitting problems caused by high dimensionality, and can also be applied to smart terminals with low computing power. Later, the identification algorithm can be verified in the real power grid system.

5. Acknowledgments
This work is supported by the National major R & D program (No. 2018YFB0904900, 2018YFB0904905).

6. References
[1] Hossain E, Khan I, Un-Noor F, Sikander S S and Sunny M S H 2019 Application of big data and machine learning in smart grid, and associated security concerns: a review IEEE Access vol 7 p 13960-13988
[2] Zhao J H, Liang G Q, Wen F S and Dong Y C 2016 Lessons learnt from Ukrainian blackout: protecting power grids against false data injection attacks Automation of Electric Power Systems vol 40 p 149-151
[3] Liao R F, Wen H, Wu J, Song H, Pan F and Dong L 2018 The rayleigh fading channel prediction via deep learning Wireless Communications and Mobile Computing 2018
[4] Wen H, Ho P H, Qi C and Gong G 2010 Physical layer assisted authentication for distributed ad hoc wireless sensor networks IET information security vol 4 p 390-396
[5] Wen H, Ho P H and Wu B 2014 Achieving secure communications over wiretap channels via security codes from resilient functions IEEE Wireless Communications Letters vol 3 p 273-276
[6] Chen S, Wen H, Wu J, Lei W, Hou W, Liu W and Jiang Y 2019 Internet of things based smart grids supported by intelligent edge computing IEEE Access vol 7 p 74089-74102
[7] Ji B, Li Y, Zhou B, Li C, Song K and Wen H 2019 Performance analysis of UAV relay assisted IoT communication network enhanced with energy harvesting IEEE Access vol 7 p 38738-38747
[8] Wen H, Tang J, Wu J, Song H, Wu T, Wu B and Sun L M 2014 A cross-layer secure communication model based on Discrete Fractional Fourier Transform (DFRFT) IEEE Transactions on emerging topics in computing vol 3 p 119-126
[9] Hu L, Wen H, Wu B, Tang J, Pan F and Liao R F 2017 Cooperative-jamming-aided secrecy enhancement in wireless networks with passive eavesdroppers IEEE Transactions on Vehicular Technology vol 67 p 2108-2117
[10] Liu X and Li Z 2016 False data attacks against AC state estimation with incomplete network information IEEE Transactions on smart grid vol 8 p 2239-2248
[11] Wen H, Wang Y, Zhu X, Li J and Zhou L 2013 Physical layer assist authentication technique for smart meter system IET Communications vol 7 p 189-197
[12] Xie Y, Wen H, Wu B, Jiang Y and Meng J 2015 A modified hierarchical attribute-based encryption access control method for mobile cloud computing IEEE Transactions on Cloud Computing vol 7 p 383-391
[13] Xiao J, Wen H, Wu B, Jiang X, Ho P H and Zhang L 2013 Joint design on DCN placement and survivable cloud service provision over all-optical mesh networks IEEE Transactions on Communications vol 62 p 235-245
[14] Jiang Y J, Wang X D, Wang W J and Bi K 2010 New feature selection approach by PCA and ReliefF Computer Engineering and Applications vol 46 p 170-172
[15] Li Y S 2019 Research on access authentication based on radio fingerprint characteristics University of Electronic Science and Technology of China
[16] Daengduang S and Vateekul P 2016 Enhancing accuracy of multi-label classification by applying one-vs-one support vector machine International Joint Conference on Computer Science and Software Engineering (JCSSE) p 1-6