International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 09 - Issue 04 || April 2023 || PP. 06-10

Defense method for malicious data injection attacks in power Internet of Things

Cao Na[1], Fu Yufei[2], Qiu Rixuan[1], Dang Fangfang[3], Yan Lijing[3]
[1](*Information and Communication Branch of State Grid Jiangxi Electric Power Co., Ltd., China)*
[2](*Xi'an Jiaotong-Liverpool University, China)*
[3](*Information and Communication Branch of State Grid Henan Electric Power Co., Ltd.(Data Center), China)*

Abstract: The Power Internet of Things (PIoT) is vulnerable to malicious data attacks that can stealthily manipulate the data of power systems. Studying how to defend against such attacks is crucial for ensuring the security and reliability of power system operations. In this paper, we propose a defense method. Our method leverages both malicious and historical normal data to jointly train the generator and discriminator networks, fine-tune the network parameters, and restore the original data. We conduct simulation experiments on the IEEE system to demonstrate the effectiveness and performance of our model .
Keywords: Power Internet of Things; malicious data attack; defense model.

I. INTRODUCTION

The power Internet of Things (PIoT) is an emerging paradigm for Industry 4.0 that involves the infrastructure construction and application research of diversified power supply demands, open network structures, information technology development and integration, and new energy sources incorporation. However, these factors also increase the complexity and security risks of the system, making it more vulnerable to malicious attacks and unauthorized intrusions. In recent years, researchers have investigated various forms and methods of network attacks on the power grid and proposed some novel attack patterns that can evade traditional detection mechanisms. These attacks are more diverse, stealthy, and resource-efficient than before, posing great challenges for prevention and control. Liu et al. [1] introduced a false data injection attack (FDIA) that can bypass bad data detection for DC state estimation in the power system. Reference [2] presented an adaptive Markov strategy (AMS) that can provide strong protection against unpredictable attacks on smart grids. Moreover, they designed a greedy algorithm to ensure the security of each data subset obtained from measurements. Reference [3] explored how to effectively cope with the correlation problem of FDIA for AC state estimation in power systems and proposed a low-cost defense strategy.

We proposes a defense model against malicious data attacks based on MisGAN. The model trains a generator with the internal network structure and parameter settings learned from a historical database. When a malicious data attack is detected by the detection module, the attacked measurement data is fed into the generator, which produces recovered data that approximates normal data[4-6]. Then, the recovered data is sent to the state estimation module as if it were not attacked by malicious data, and the subsequent processes are completed. The performance of the MisGAN defense model is validated using standard IEEE-14 and IEEE-118 node test systems [7-9]. Experimental results demonstrate that the recovered data generated by the MisGAN defense model closely resembles normal data, and the effect of different iteration numbers on the defense model performance is also verified.

II. PROBLEM DESCRIPTION

This paragraph presents the mathematical principle of residual detection. The residual vector is the difference between the actual measurement and the estimated measurement：

$$r = z - \hat{z} = z - H\hat{x} \quad (1)$$

In the absence of an attack, it holds that：

$$r \le \tau \quad (2)$$

τ is the threshold for detecting bad data.

FDIA is a stealthy attack that injects an attack vector to manipulate the original measurement vector, leading to a wrong state estimation , such that the manipulated and still satisfy equations (1) and (2).

and，

$$a = [a_1, a_2, \cdots, a_m]^T \quad (3)$$
$$z_a = z + a \quad (4)$$
$$\hat{x}_a = \hat{x} + c \quad (5)$$

Therefore, we derive the residual computation formula：

$$\begin{aligned}\|r\| &= \|z_a - H\hat{x}_a\| \\ &= \|z + a - H(\hat{x} + c)\| \\ &= \|z - H\hat{x} + a - Hc\| \end{aligned} \quad (6)$$

According to equation (2), the injected attack vector $a = Hc$can ensure $r \leq \tau$.

Generative Adversarial Networks (GANs) [10] are a deep learning-based model that consists of two modules: a generator and a discriminator. The generator and the discriminator learn from each other through a game-theoretic process, resulting in better outputs. GANs require complete normal measurement data for training, but when the power system is subjected to malicious data attacks, the dispatch center can only obtain incomplete data. Therefore, using GANs directly to solve this problem leads to a gap between the recovered data and the normal data.

The framework of MisGAN has two main steps: 1) In addition to using a complete data generator, it also uses a masking generator to explicitly model the data loss process, which allows us to see clearly the missing parts of the incomplete dataset and use standard GANs to estimate the data distribution; 2) It uses complete data to adversarially train the generator, masks its output with f_τ using the masking generator , and compares it with f_τ masked incomplete dataset , where the latter is similar to the actual incomplete dataset. MisGAN uses two pairs of generators and discriminators: masking pair (G_m, D_m) and data pair (G_x, D_x), where each generator is independent and has its own noise distribution p_z and p_ε.The loss functions for each generator are defined as follows:

$$L_m(D_m, G_m) = E_{(x,m)\sim p_D}[D_m(m)] - E_{\varepsilon\sim p_\varepsilon}[D_m(G_m(\varepsilon)] \quad (7)$$

$$L_x(D_x, G_x, G_m) = E_{(x,m)\sim p_D}[D_x(f_\tau(x,m)] - E_{\varepsilon\sim p_\varepsilon, z\sim p_z}[D_x(f_\tau(G_x(z), G_m(\varepsilon)))] \quad (8)$$

The objective functions for optimizing the generator and the discriminator are:

$$\min_{G_x} \max_{D_x \in F_x} L_x(D_x, G_x, G_m) \quad (9)$$

$$\min_{G_m} \max_{D_m \in F_m} L_m(D_m, G_m) + \alpha(D_x, G_x, G_m) \quad (10)$$

Where $\alpha$is an optimization coefficient that can be set to 0. Experiments show that setting it between 0.1~0.2 can yield relatively optimal solutions.

III. A DEFENSE MODEL FOR MALICIOUS DATA ATTACKS

Construct two pairs of generator and discriminator, namely masking pair(G_m, D_m) and data pair (G_x, D_x), where the data pair discriminator is trained with normal measurement data, which ensures that the data pair generator has smaller error and is closer to the normal measurement data. The model loss functions and optimization functions are as follows (8)、(9)、(10)、(11).

Random noise vectors z are fed into generator G_x to produce $\tilde{x}$ random noise vectors ε are fed into generator G_m to produce $\tilde{m}$,based on $\tilde{x}, \tilde{m}$, $f_\tau(\tilde{x}, \tilde{m})$ is computed, based on the normal measurement data and abnormal measurement data in the dataset, the true $m$can be obtained, and then $f_\tau(x, m)$ is computed[11-14]. The outputs of generator G_x, G_m ,which are indirectly derived from $f_\tau(\tilde{x}, \tilde{m})$ along with $f_\tau(x, m)$ are fed into discriminator D_x for training. Through loss functions (8) and (9), objective optimization functions (10) and (11) provide feedback to generators G_x and G_m.The final outputs of generators (G_x, D_x) and (G_m, D_m) are obtained after convergence. After the model training is completed, use generator G_x to perform the final data imputation task. which is transmitted to dispatch center as normal measurement data for normal state estimation.

IV. EXPERIMENT

We perform experiments on IEEE-14 and IEEE-118 systems using real power consumption data from a certain location. The data includes active load and reactive load, as well as phase angle values and voltage amplitude that need to be measured. We use the Mat power toolkit to solve these data with the IEEE systems. The original data covers 90 days with a sampling frequency of every 5 minutes. Each 5-minute sample is considered as one time slice, resulting in a total of 25920 time slices in this dataset. For training purposes, we balance the normal measurement data and the attacked measurement data with a ratio of 1:1. We evaluate our method by comparing the mean square error (MSE) of node voltage amplitude and node phase angle value between normal measurement data and Mis GAN model recovered measurement data under different attack

strengths. Figures 1 to 4 compare voltage amplitude and phase angle under normal and various attack scenarios for IEEE standards-based systems with either a) fourteen or b) one hundred eighteen buses.

System

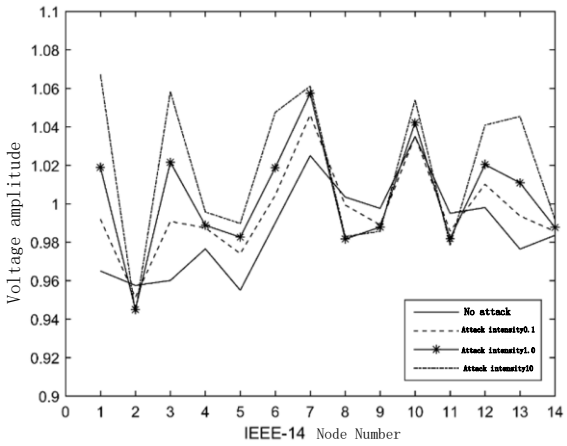

Figure 1 IEEE-14 Voltage amplitude under different attack intensities in IEEE-14 bus standard test

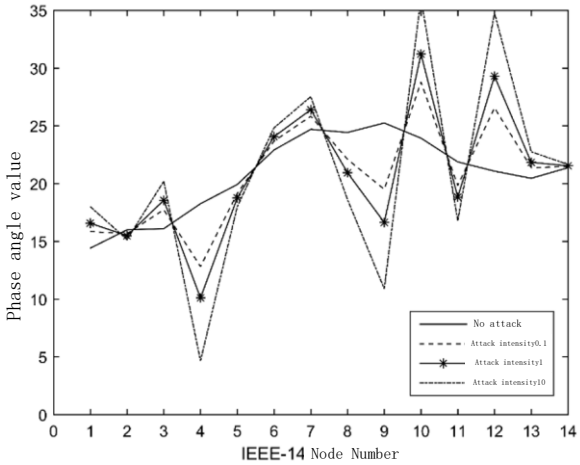

Figure 2 Phase angle under different attack intensities in IEEE 14 bus standard test system

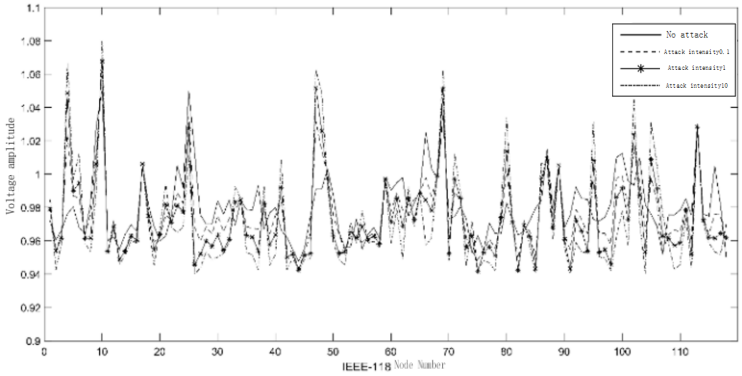

Figure 3 Voltage amplitude under different attack intensities in IEEE 118 bus standard test system

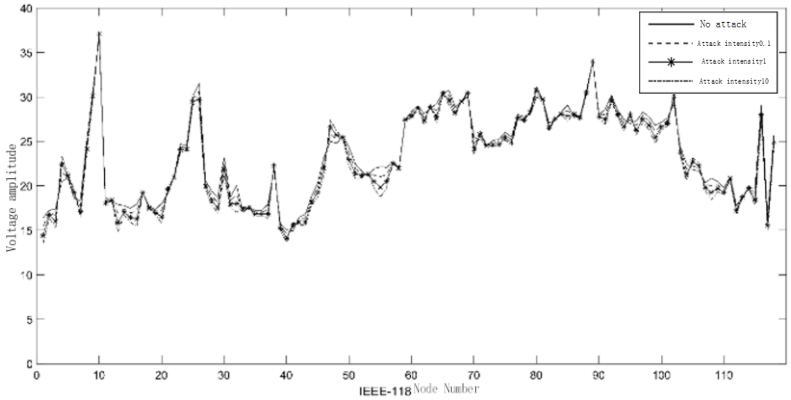

Figure 4 Phase angle under different attack intensities in IEEE 118 bus standard test system

Figures show that even if an attacker targets only one or a few buses, other buses are also impacted because of how they are electrically connected (topology). This lets an attacker manipulate overall data so that it passes state estimation and power flow checks while avoiding bad data detection mechanisms. The more complex a grid's topology is, the more vulnerable it is to targeted attacks with high intensity on its weak points.

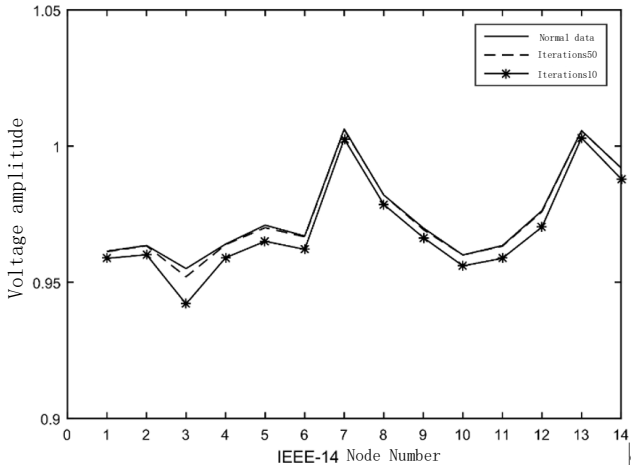

Figure 5 Comparison of voltage amplitude recovery by generator trained with different number of iterations in IEEE-14 bus standard test system

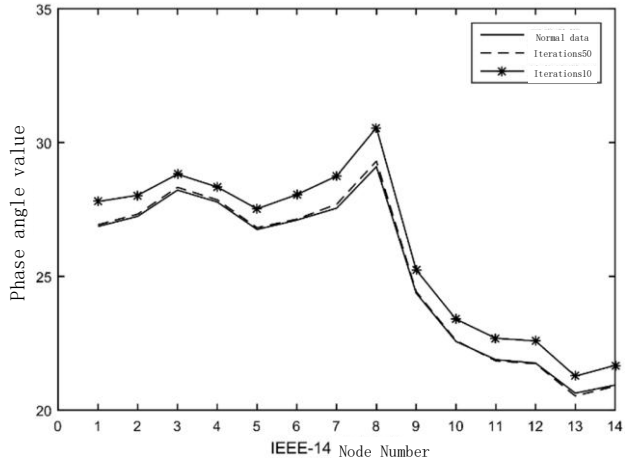

Figure 6 Comparison of phase angle recovery by generator trained with different number of iterations in IEEE-14 bus standard test system

Figures 5 to 8 show the comparison of normal and recovered data for voltage magnitude and angle under the IEEE-14 and IEEE-118 bus standard test systems123. The IEEE-14 bus system is a simple approximation of the American Electric Power system with 14 buses, 5 generators, and 11 loads2, while the IEEE-118 bus system

is a simple approximation of the same power system as of December 1962 with 118 buses, 19 generators, 35 synchronous condensers, 177 lines, 9 transformers, and 91 loads1. We use the same topology and normal data as the benchmark for both systems. We train the Mis GAN model with different numbers of iterations. The more iterations we use, the higher the model performance and the closer the recovered data to the normal data. The curves in Figures 5 and 6 confirm this hypothesis. Figures 7 and 8 present the experimental results for the IEEE-118 bus standard test system1, which also verify the same conclusion. All these experimental results validate the effectiveness of our proposed defense model.

V. CONCLUSION

After a power grid system suffers from a malicious data attack, in order to enable the dispatch center to continue normal operations such as state estimation, optimal power flow, economic dispatch, and issuing dispatch instructions, we first design a malicious data attack defense model. This defense model trains the internal network structure and sets the parameters based on the historical database, and finally obtains a generator. When the detection module detects a malicious data attack, it takes the attacked measurement data as the input of the generator and outputs recovered data that is close to normal data. Then it sends the recovered data as normal data that has not been attacked by malicious data to the state estimation module for subsequent processes. We then propose an overall architecture for defending against malicious data attacks in power internet of things. Finally, we apply our proposed model on IEEE-14 and IEEE-118 bus standard test systems123, and conduct related verification work. We find that the recovered data output by the generator of our proposed Mis GAN defense model is close to normal data. We also verify the impact of different numbers of iterations on the performance of our defense model.

VI. ACKNOWLEDGEMENTS

This work was supported by the State Grid Jiangxi Electric Power Corporation Science and Technology Project “Research on Active Defense Technology for Advanced Sustainable Network Attacks Based on Dynamic Obfuscation” under Grant 521835220003.

REFERENCE

[1]. Liu Y, Ning P, Reiter M K. False Data Injection Attacks against State Estimation in Electric Power Grids[J]. Acm Transactions on Information and System Security, 2011, 14(1).
[2]. Hao J, Kang E, Sun J, et al. An Adaptive Markov Strategy for Defending Smart Grid False Data Injection from Malicious Attackers [J]. IEEE Transactions on Smart Grid, 2018, 9(4): 2398-2408.
[3]. Deng R L, Xiao G X, Lu R X. Defending Against False Data Injection Attacks on Power System State Estimation[J]. Ieee Transactions on Industrial Informatics, 2017, 13(1): 198-207.
[4]. Liang G Q, Weller S R, Luo F J, et al. Distributed Blockchain-Based Data Protection Framework for Modern Power Systems against Cyber Attacks [J]. Ieee Transactions on Smart Grid, 2019, 10(3):3162-3173.
[5]. Tian J, Tan R, Guan X H, et al. Enhanced Hidden Moving Target Defense in Smart Grids [J]. Ieee Transactions on Smart Grid, 2019, 10(2):2208-2223.
[6]. Higgins M, Teng F, Parisini T. Stealthy MTD against Unsupervised Learning-Based Blind FDI Attacks in Power Systems [J]. IEEE Transactions on Information Forensics and Security, 2021, 16:1275-1287.
[7]. Wang J Y, Shi D Y, Li Y H, et al. Realistic Measurement Protection Schemes Against False Data Injection Attacks on State Estimators[J]. 2017 Ieee Power & Energy Society General Meeting, 2017.
[8]. Hao J P, Piechocki R J, Kaleshi D, et al. Sparse Malicious False Data Injection Attacks and Defense Mechanisms in Smart Grids[J]. Ieee Transactions on Industrial Informatics, 2015, 11(5):1198-1209.
[9]. Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets [J]. MIT Press, 2014.
[10]. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [J], 2017.
[11]. Gulrajani I, Ahmed F, Arjovsky M, et al. Improved Training of Wasserstein GANs. arXiv e-prints. 2017.
[12]. Ahmadi M, Nest T, Abdelnaim M, et al. Reproducing Ambient GAN: Generative models from lossy measurements. arXiv e-prints. 2018.
[13]. Cheng-Xian Li S, Jiang B, Marlin B. Mis GAN: Learning from Incomplete Data with Generative Adversarial Networks. arXiv e-prints. 2019.
[14]. Li Y, Wang Y, Hu S. Online Generative Adversary Network Based Measurement Recovery in False Data Injection Attacks: A Cyber-Physical Approach [J]. IEEE Transactions on Industrial Informatics, 2019:1-1.