Research on Fake Data Injection Attack Detection in Intelligent Ship Power System

Ganlong Wang 1, Jun Zhu1 *
1 CSIC Information Technology Co., Ltd, Lianyungang, Jiangsu, 222000, China
*Corresponding author’s e-mail: huzi881210@foxmail.com

Abstract. The application of computers in marine power systems is becoming more and more extensive, and the network security of marine power systems has become an important part. This paper studies the false data injection attack (FDIA) of the ship's power information physical system. First, it analyzes the objectives and principles of the FDIA attack from the perspective of the attacker and the possible consequences on the power grid. From the technical point of view, it studies two aspects of FDIA attacks, analyzes and summarizes various methods in detection and defense. The FDIA defense model based on deep learning is proposed, and on this basis, the measurement points that need to be protected are confirmed according to the importance index $I_p$. And its index is applied to the established fake data injection attack detection model based on deep learning, and the accuracy of the two traditional fake data injection attacks is compared. Experiments show that the accuracy of this model is increased to more than 90% compared with traditional algorithms, which can effectively enhance the defense capability of the ship's power system against FDIA.

1. Introduction
FDIA destroys the integrity of power grid information by tampering with state estimates. It will jeopardize the normal safe and economic operation of the ship's power system, and is more likely to cause the disintegration of the entire ship's power system and the entire ship into darkness. This attack method was first proposed by Yao Liu [1]. This attack exploits the loopholes of the traditional bad data detection mechanism of the control center to induce the state estimation of the control center to produce an erroneous estimation of the system state, thereby causing the EMS to generate wrong instructions, which will affect the operation of the power grid. Effective detection and protection measures against fake data injection attacks are major challenges for power systems to maintain their security and stability. In recent years, with the deepening of research, attackers have made good progress in the method of completing vicious data injection attacks with little or completely unknown system topology and line parameter information[2]. Therefore, this paper analyzes the state attack in the fake data injection attack, and establishes an online defense model that can effectively detect such attacks. In response to fake data injection attacks, due to the large number of smart meters in the power grid, the cost of achieving complete security protection is quite high, so the realistic approach is to use offline defense to protect a part of the key measurement.

Literature discusses the minimum number of measuring instruments that an attacker needs to invade in order to complete a malicious data injection attack, and then gives the optimal configuration strategy of PMU required to replace the instrument [3]. Literature proposes A set of strategies for dynamically reconstructing microgrids makes it impossible to inject vicious data attacks[4]. Literature uses graph theory to select a set of protected quantity measurements to curb the occurrence of vicious
data injection attacks[5]. On the other hand, online defense makes analysis based on differences in meter data. Reference uses a set of safe and reliable PMU measurement and Huber robust estimation to detect the consistency of the measurement data to detect attacks[6]. Reference uses wavelet transform and deep neural network to obtain the system through machine learning State dynamic spatio-temporal characteristics[7], and then judge the difference in state quantity at the current moment. Literature classifies potential vulnerable nodes in the power grid by optimizing clustering algorithm [8]. Detects the malignant data with the help of the traditional autoregressive state prediction results. Literature uses the KL divergence method to obtain the quantity measurement variation probability distribution difference to determine whether the current system is under attack [9-11].

At present, the development of ship power system automation technology and sensor networks is rapid. The traditional ship power grid is gradually transformed into an intelligent ship power grid with independent coordination of information systems and physical systems with independent decision-making capabilities. As an inevitable result of ship development, the smart ship power grid has stronger self-repair and safety. Therefore, research on virus attacks targeting FDIA detection technology and defense technology is of great significance.

2. Traditional detection technology for FDIA

2.1. Detection technology based on state estimation
If the network attack encountered has a comprehensive understanding of grid information and power protection optimization algorithms, it can build an optimization algorithm that can avoid the identification and monitoring of current abnormal information based on the least square method state estimation. According to the characteristics of this type of information forgery, the initial state estimation optimization algorithm can be improved to improve the level of identifying others as forged information. The more professional monitoring technology includes the method of measuring mutation, the correlation and residual monitoring of the measurement. In addition, it is noted that the monitoring threshold is greatly disturbed by the scale of the system. For a very large system, the global system can be divided into multiple blocks, and the threshold can be different according to the specific parameters of each subsystem.

The advantage of this type of method is that it uses a perfect optimization algorithm, monitors faster, and can better reflect the characteristics of the power system. However, in specific applications, different monitoring thresholds will obtain different monitoring precision, which is very easy to cause missed detection and false detection.

2.2. Detection technology based on trajectory prediction
The monitoring method of state estimation is selected for static analysis, and the aggressive behavior at a certain point in time is monitored. In the continuous dynamic operation of the power system, various state quantities have a strong spatial relationship with each other. This type of method estimates the decentralized rules for state quantities based on the operating rules of the system state and historical data tables. It is paired according to the operation trajectory and can efficiently monitor various types of forged information. This type of estimation method is very complicated and the monitoring rate is very slow, especially not suitable for complex systems inside our power grid.

Fake information injection attacks generally rely on the flaws in the monitoring and identification methods based on the number of anomalies in the residuals. Then it can falsify the measured values in the power system, in order to achieve the goal of obtaining economic profits and affecting the stable operation of the power system. To this end, an effective FDIA defense method is to conceive a better monitoring strategy. According to the defense measures of the power grid, we first define two indicators, including: indicator $V_k$ and indicator $R_k$. Consider the level of defense against FDIA of the power system from the perspective of part of the power system to the overall perspective. Subsequently, we constructed the FDIA defensive model with indicators $V_k$ and $R_k$. On this basis, the
quantity control points that must be protected are checked according to the criticality monogram $l_k$. According to the test system of IEEE30 and IEEE117 nodes, the defensive measures are simulated and tested, and the experiment proves that the method is efficient.

3. Formatting the title, authors and affiliations

In this paper, the deep convolutional neural network inception-v3 model is used to establish a false data injection fault detection model structure to complete the task of false data injection fault detection. The Inception-v3 model was proposed and announced by Google in 2015, and this model uses this model for fault detection. The Inception-v3 model implements a new convolution module that diversifies the width of the network. In each module, multiple branches are used to extract the feature map of the input. The depth, susceptor and pooling parameters of each branch are different. The final output is the fusion of the features of each branch before. Features rich in information at different scales. The input feature map is fused in the third dimension after a variety of convolution calculations, as shown in Figure 1, where the $1 \times 1$ convolution is used for the fusion of different channels of the feature map.

Convolutional neural network (CNN) is composed of many convolutional layers and is the core component of the entire network. Each convolutional layer has its own local connection and weight sharing function. The convolution kernel is mainly used to perform convolution operations, and the activation feature is used to obtain the text feature map. First represent the ith data vector in the text according to the vector from the previous step, as shown in Equation 1:

$$X_{i, n} = x_1 \oplus x_2 \oplus \cdots \oplus x_n$$

Where $\oplus$ is a concatenation operator, a sentence can be expressed in the form of a matrix in this way, which is convenient for convolution to obtain local feature vectors through calculation. On this basis, we assume a convolution kernel, which has $h \times k$ parameters to be estimated, only when our
The convolution kernel has a sequence of words from \( x_i \) to \( x_i + h \). The product operation will output \( o_i \) as shown in Equation 2:

\[
o_i = w \cdot X[i : i + h - 1]
\] (2)

Where \( i=1,...,s-h+1 \) represents the calculation of the dot product of the convolution kernel and sub-matrix. Finally, using an activation function, a characteristic formula can be calculated, as shown in Equation 3:

\[
c_i = f(o_i + b)
\] (3)

We use ReLU as our activation function. It can be very effective to improve the learning efficiency of the entire network and can improve our convergence speed in network learning. Then it can reduce the number of iterations. When we normalize the data we collected, we list each situation, and perform a convolution operation through \( \{x_{ih}, x_{2h}, \cdots, x_{s-h+i}\} \) to obtain a feature map \( c \in \mathbb{R}^{c-h+1} \), as shown in Equation 4:

\[
c = [c_1, c_2, \cdots, c_{s-h+1}]
\] (4)

4. Comparative analysis of experimental results

4.1. Experiment preparation
The specific environment of the experiment is: 2.5GHz i5-7300HQ, 16G memory, 256G SSD, the operating system uses the Linux system, MatPower is used to generate the network topology and measurement data of the standard nodes, and the programming environment uses Visual Studio code 2019 Python 3.7.5. The TensorFlow framework and Pandas open source library are used to extract the abnormal scores.

4.2. Solving model parameters
We use MatPower to generate standard node systems and measurement data. Then we simulate the false data injection attack under non-complete topology information by injecting attack vectors. Finally, training data samples and test data samples containing attack data samples and normal measurement data samples are generated. At the beginning, we need to normalize the generated data. Then, we use the inception-v3 model to extract data features in layers. Since there are so many parameters, we first perform a dimension reduction process on the data. Finally, we began to use the model to train data samples non-stop. Constantly adjust the model parameters to get good prediction results.

In general, the parameters of the system mainly include: frame parameters and classification algorithm parameters. Among them, the important parameters in the framework are convolution kernel size, step size, number of iterations, learning rate, etc. In the SVM classification algorithm, there are two important parameters: C penalty coefficient and gamma. Gamma is the RBF function selected as a parameter attached to the kernel. C is the weight of adjusting the interval between the two indicators in the optimization direction, which is a preference for classification accuracy, that is, its tolerance for errors. The higher the C, the higher the tolerance for errors, and the easier it is to overfit; the smaller the C, the easier it is to underfit. If C is too large or too small, the generalization ability becomes poor. After mapping to a new feature space, Gamma implicitly determines the distribution of the data. The larger the gamma, the fewer the support vectors. The smaller the gamma value, the more the support vectors. The number of support vectors will directly affect the speed of training and prediction. The selection of these parameters will affect the final test results.
We verify the accuracy of training and the correctness of the model. We divide all the data into two parts, one part of which takes up 90% of the total data and is dedicated to training. The remaining 10% is used to verify the correctness of the model to test the accuracy of our system.

We first perform simple preprocessing on the collected field data, and then reduce the dimension of the data. Finally, we use the inception-V3 model structure to extract the features of the data we selected. The parameters of the model are shown in Table 1.

| Network Type | Convolution kernel size / step |
|--------------|--------------------------------|
| Conv         | 3*3/2                          |
| Conv         | 3*3/1                          |
| Conv padded  | 3*3/1                          |
| Pool         | 3*3/2                          |
| Conv         | 3*3/1                          |
| Conv         | 3*3/2                          |
| Conv         | 3*3/1                          |
| 3*Inception  | -                              |
| 5*Inception  | -                              |
| 2*Inception  | -                              |
| Pool         | 8*8                            |
| Linear       | logits                         |
| SVM          | classifier                     |

The selection of SVM parameters, the changes in prediction results and prediction time caused by different C values are shown in Figure 2.

![Figure 2. C value change accuracy](image)

During our training, the number of iterations is 4000, the initial learning rate is 0.01, and the learning rate per 100 steps becomes 96% before. Each iteration randomly extracts data from the
training set for training, and outputs accuracy and average loss after 10 iterations. When the model is trained from 0 to 200 times, the value of the loss function quickly decreases and reaches approximately. As the number of iterations increases, it gradually decreases. It is basically stable at 0.01, and training can be stopped at this time, and the accuracy is the highest.

We used three methods for comparative analysis, including two traditional methods for false data injection failure detection. And the fault detection model of fake data injection based on deep learning in this paper. Three algorithms are used for comparative analysis of detection accuracy. The results are shown in Figure 3.

![Figure 3. Prediction accuracy](chart.png)

It can be seen from the above figure that the state estimation method has the lowest accuracy and is not suitable for the detection of false data injection failures. The accuracy of the trajectory prediction method is more than 70%, which can meet the initial detection function, but it has not achieved the desired effect. Driven by artificial intelligence algorithms, we tried deep learning algorithms with an accuracy rate of about 95%. This can fully meet our current industrial error requirements, further improve the detection accuracy of the system, and is of great significance to the stability of the entire marine power grid system.

5. Conclusion
Based on previous research, this paper proposes a deep data-based fake data injection fault detection and defense model. First we preprocess the data, then use Inception-v3 to design the model, and then use R2 for classification. It is proved by experiments that the accuracy of this model is improved to more than 90% compared with traditional algorithms. It can effectively enhance the ability of the ship's power system to defend against FDIA.

Acknowledgments
This paper was supported by Shipbuilding industry built-in information security function Industrial control equipment promotion and application (NO. YZF18.008), network security solution provider for industrial enterprises, industrial Internet platform enterprises (NO.TC190H3XG), and industrial enterprise network security comprehensive protection platform (NO.TC190H3WQ).
References

[1] Liu, Y., Reiter, M. K., & Ning, P. (2009). False Data Injection Attacks Against State Estimation in Electric Power Grids. Proceedings of the 2009 ACM Conference on Computer and Communications Security, CCS 2009, Chicago, Illinois, USA, November 9-13, 2009. ACM.

[2] Yao, Liu, Peng, Ning, Michael, & K. (2011). False data injection attacks against state estimation in electric power grids. ACM Transactions on Information and System Security.

[3] André Teixeira, Amin, S., Sandberg, H., Johansson, K. H., & Sastry, S. S. (2011). Cyber security analysis of state estimators in electric power systems. 49th IEEE Conference on Decision and Control (CDC). IEEE.

[4] Hug, G., & Giampapa, J. A. (2012). Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks. IEEE Transactions on Smart Grid, 3(3), 1362-1370.

[5] M.A. Rahman, & H. Mohsenian-Rad. (2013). False data injection attacks against nonlinear state estimation in smart power grids. Power & Energy Society General Meeting. IEEE.

[6] Liang, J., Kosut, O., & Sankar, L. (2014). Cyber attacks on AC state estimation: Unobservability and physical consequences. power and energy society general meeting.

[7] Chin, W., Lee, C., & Jiang, T. (2018). Blind False Data Attacks Against AC State Estimation Based on Geometric Approach in Smart Grid Communications. IEEE Transactions on Smart Grid, 9(6), 6298-6306.

[8] Kim, J., Tong, L., & Thomas, R. J. (2013). Data framing attack on state estimation with unknown network parameters. asilomar conference on signals, systems and computers. , 32(7):1460-1470.

[9] Zhang, C., Ren, Z., Zhang, A., Zhang, Y., & Geng, Y. (2013). Malicious data injection attack against power system state estimation based on orthogonal matching pursuit. asian control conference., 1-6.

[10] Ozay, M., Esnaola, I., Vural, F. T., Kulkarni, S. R., & Poor, H. V. (2013). Sparse Attack Construction and State Estimation in the Smart Grid: Centralized and Distributed Models. IEEE Journal on Selected Areas in Communications, 31(7), 1306-1318.

[11] Yamaguchi, Y., Ogawa, A., Takeda, A., & Iwata, S. (2014). Cyber security analysis of power networks by hypergraph cut algorithms. international conference on smart grid communications. 6(5):2189-2199.