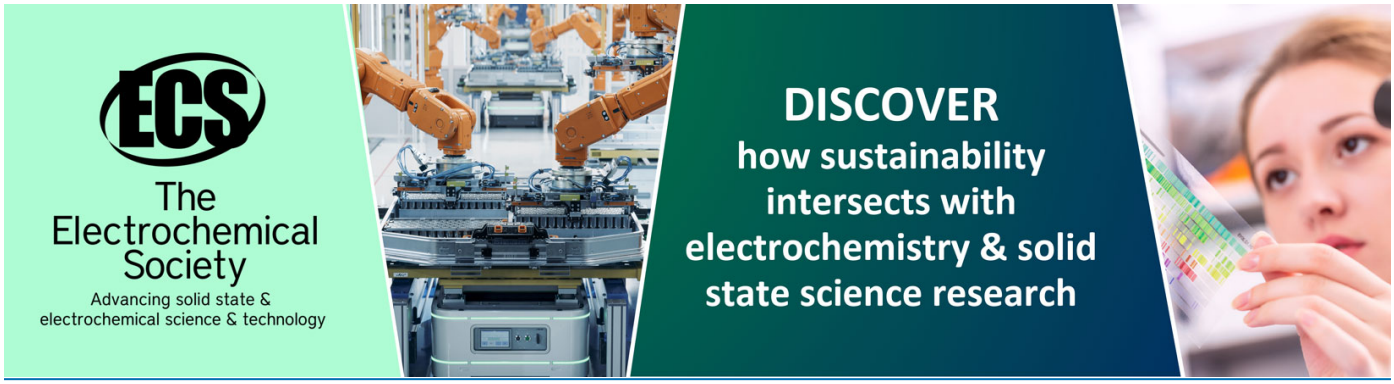
Intrusion Detection Method of Electric Power Information Network in Cloud Computing Environment



Intrusion Detection Method of Electric Power Information Network in Cloud Computing Environment

Jiaqi Zhang [1*], Guoping Feng [2] , Dexi Zhou[2] and Mingjiu Li [2]

[1] Guangzhou Power Supply Bureau of Guangdong Power Grid

[2] China Energy Engineering Group Guangdong Electric Power Design Institute Co., Ltd.

*Corresponding author's e-mail: zhangjiaqi@gzpsc.csg.cn

Abstract. With the widespread application of power grid systems, the information security problems faced by power grids have become more obvious. Various internal and external intrusion attacks that occur frequently have become an important issue affecting the normal operation of power generation and operations. The purpose of this paper is to study the intrusion detection method of electric power information(PI) network in the cloud computing environment. With the help of the cloud platform's ability to process big data, and based on the analysis of the PI network structure, a DBN optimized BP network algorithm is proposed, and the optimized BP neural network is used as a runtime classification program. Experimental results show that MR-DBN-BP has a detection rate of 96.7% for intrusion detection of PI networks, which can effectively detect intrusions and effectively protect the power dispatch system network.

1. Introduction

With the rapid development of Internet applications and distributed computing technology [1-2], cloud computing has become a mature network technology and is widely used in PI networks [3-4]. As the PI system is subject to various external and internal attacks, the information security problem of the PI system has become more obvious, which has seriously affected the normal production and stable operation of the power system [5-6]. With the continuous improvement of the integrated information network of the power supply network, the security threats from the Internet have become more complex and changeable. These threats include the information network, the collection of energy user information, smart power consumption, power trading, power distribution network automation, chargers, etc. [7-8].

In the research of intrusion detection methods for PI networks based on cloud computing environments, many scholars have studied them and achieved good results. For example, Suzuki S, Hiraoka M, Shiraishi T and others proposed the first profile based on user behavior characteristics. A real-time intrusion detection system model [9]. Wang D, Bing-Yu G E, Xuan J X and others have proposed a GPU-based high-speed network intrusion detection system for the high underreporting problem faced by traditional intrusion detection under high-speed traffic conditions [10].

Based on the analysis of the PI network structure and its security partitions, this paper proposes a PI network intrusion detection framework based on cloud computing technology in view of the intrusion attack problems in the PI network.

2. PI Network Intrusion Detection Method Under Cloud Computing Environment

2.1. PI Network Structure

When constructing a secure power system network, it is necessary to comprehensively consider how to improve the security of the entire power system network. Therefore, it is necessary to study the security between the real-time system and the local area network to ensure the most important power production real-time system security, to study the security between the real-time systems to ensure the independence between systems, and to avoid mutual interference; to study the relationship between the local area network and the Internet To ensure the safety of power application systems such as MIS, OA, and power management systems.

2.2. Structure of PI Network Intrusion Detection System

Nowadays, the emergence of high-speed networks and the continuous update of attack methods make it more and more difficult to update and maintain the knowledge base by manual methods. As a method of obtaining knowledge from massive data, cloud computing technology was first introduced by WenkeLee et al. to intrusion detection. The field subsequently developed vigorously. The system adopts misuse intrusion detection, that is, judges whether the data in the network has intrusions based on known intrusion patterns. The intrusion detection system must increase the speed of rule generation, increase the accuracy of the rules, and reduce the system's false negative and false positive rates; it must detect intrusions quickly and accurately; the rule base must be continuously updated to accurately discover new ones in a timely manner Intrusion attacks. Due to the huge number of power information networks and unknown sources, it is very important to strengthen the maintenance of their safety performance. Applying cloud computing to the security detection of power information networks, by establishing a database, it can effectively detect intrusions in the power information network, maintain the security of the power information network, and ensure its normal operation.

This article will obtain network data information flowing through the power dispatch automation system. After preprocessing the data, use data mining methods to discover the attack mode of the system data information, thereby detecting illegal behaviors and intrusion attacks.

2.3. BP Neural Network Optimization Based on DBN

Because the BP network relies heavily on the initial weight, it is easy to fall into the local minimum, and the convergence speed is slow, and the training time is long. Many improved BP algorithms have been proposed. In order to solve the problem of BP's sensitivity to initial weights, a global search strategy in the range space using a global search algorithm is proposed. Starting from the BP algorithm itself, an adaptive variable step size for the momentum item is proposed, based on LM (Levenberg-Marquardt) Improved algorithm of algorithm. The above method can effectively solve the problem of BP network in solving the global minimum and the number of iterations, but the first strategy is prone to oscillations, and the second method is prone to overfitting.

Combined with the deep belief network, a BP improved algorithm based on DBN optimization (DBN-BP) is proposed, because the multiple Boltzmann machines of DBN can obtain better model parameter values in the process of unsupervised learning, and then pass BP. The algorithm is fine-tuned, and it can quickly converge to the global minimum. For a given RMB state vector (v, h), define the energy of RBM as:

$$E(v,h|\theta) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{n} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} v_i w_{ij} h_j \quad (1)$$

Equation (1) is the parameter of RBM, and the joint probability of the visible layer and the hidden layer is:

$$P(v,h|\theta) = \frac{1}{Z(\theta)}\exp(-E(v,h|\theta)) \quad (2)$$

Where Z(θ) is the normalization factor, Z(θ)=∑_(v,h)▒〖exp (-E(v,h|θ))〗. The distribution of observation data defined by RBM is P(v|θ) and P(h|θ):

$$P(v|\theta) = \frac{1}{Z(\theta)}\sum_v exp(-E(v,h|\theta)) \quad (3)$$

$$P(h|\theta) = \frac{1}{Z(\theta)}\sum_h exp(-E(v,h|\theta)) \quad (4)$$

3. Experimental Research on PI Network Intrusion Detection Method Under Cloud Computing Environment

3.1. Description of Test Data Set

The data set used in this article is the network traffic audit log data of a power company. In order to increase the data of various intrusion behaviors, a part of the security audit data set is added to the data set after analysis and processing to form an intrusion detection data set (KD). A total of 2 million network traffic is collected as training data, and a test data set is a total of 1 million data sets. It includes four types of intrusions: Dos attacks, port scanning attacks, unauthorized access by remote hosts, and unauthorized access by local users. Each record has 53-dimensional attributes, and the last attribute is its category.

3.2. Data Processing Method

Because the data contains two types of discrete and continuous data. In order to fit the input of neurons and eliminate the situation that big numbers eat decimals, the data needs to be standardized and normalized.

(1) Data standardization

The input of the neural network requires numeric input, and uniformly encodes items whose attribute fields are of character type. The encoding method uses lexicographic sorting to assign ordinal numbers to character fields in turn. Other character fields are standardized in the same way.

(2) Data normalization

The size range of the attribute is reduced to the interval of [0,1]. This article uses the formula a=(a-min)/(max-min) to normalize the data, where a represents the value of the attribute field.

4. Data Analysis of PI Network Intrusion Detection Method Under Cloud Computing Environment

4.1. MR-DBN-BP Algorithm Classification Performance Test Under Cloud Computing Applications

Test the extraction performance of the DBN-BP algorithm under different data volumes. KD dataset data is randomly generated into multiple groups of datasets, KD1 to KD8. Among them, KD1 contains 10,000 pieces of sample data, KD2 contains 30,000 pieces of data, KD3 contains 50,000 pieces of data, KD4 contains 70,000 pieces of sample data, KD5 contains 100,000 pieces of sample data, KD6 contains 130,000 pieces of sample data, and KD7 contains 16 pieces of data. There are 10,000 pieces of sample data, and KD8 has 200,000 pieces of sample data.

The training time of DBN-BP algorithm on single machine is compared with that of MR-DBN-BP in cluster environment. The comparison results are shown in Table 1.

Table 1. Single-machine DBN-BP and 4-node cloud cluster MR-DBN-BP data

Sample data (ten thousand)	Stand-alone DBN-BP (training time S)	Cluster MR-DBN-BP (training time S)
1	218	320
3	345	369
5	420	384
7	645	418
10	972	458
13	1341	489
16	1867	550

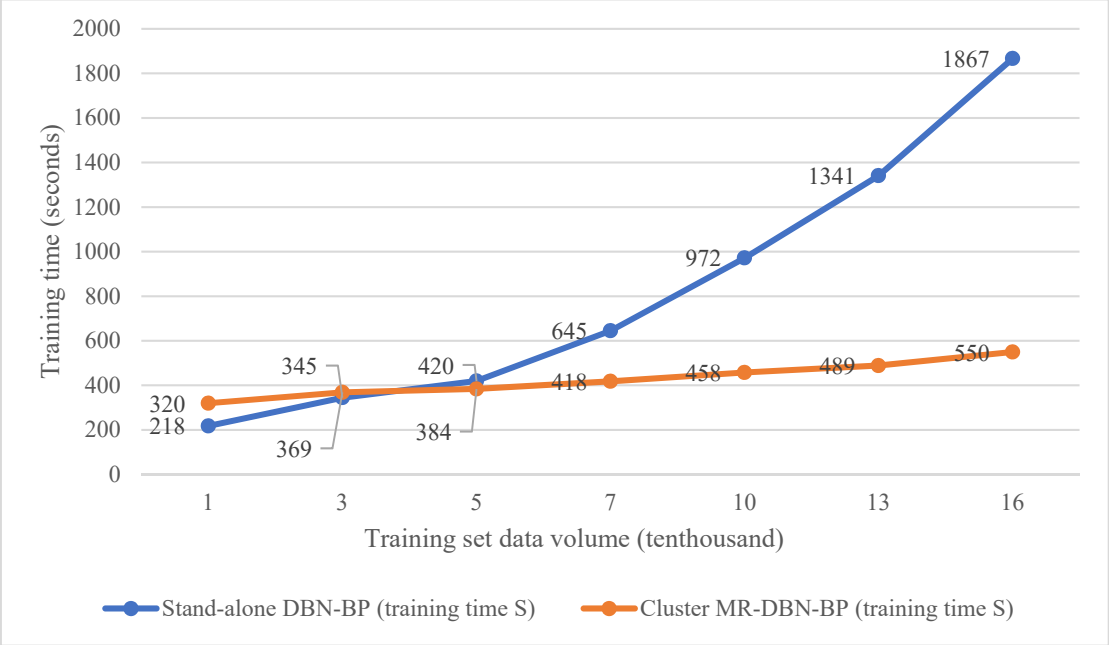


Figure 1. Single-machine DBN-BP and 4-node cloud cluster MR-DBN-BP

As shown in Figure 1, with the expansion of the training set size, the execution time of the MR-DBN-BP algorithm on a 4-node cloud computing cluster has little change. In a stand-alone environment, with the expansion of the training set size, DBN -BP algorithm execution time is also increasing. Through the analysis of the experimental results, it can also be concluded that in a 4-node cloud computing cluster environment, the training speed of the MR-DBN-BP algorithm is more than 3 times that of the single-machine DBN-BP algorithm. In practical application, the increase of the number of cloud cluster nodes and the growth of power information network connection record sample data set are conducive to the operation of MR-DBN-BP algorithm for cluster computing.

4.2. Parallel Performance Analysis Of MR-DBN-BP Algorithm in Cloud Computing Applications

In order to compare the method in this paper with the BP method, PSO-BP, and GA-BP method that introduces the steepness factor λ, now set the network parameters: set the number of RBMs constituting the DBN network to two, and the number of nodes to be 53-22 -12, 1000 iterations; BP neural network has a three-layer structure, the number of nodes is 12-6-3, 500 iterations, the learning rate is 0.01, and the error rate is 0.01. Table 2 is the comparison between MR-DBN-BP and λBP, PSO-BP, GA-BP in three aspects: detection rate, false alarm rate and false alarm rate. It can be seen that compared with λBP, MR-DBN-BP, PSO-BP and GA-BP have the highest detection rate of 96.7%.

Table 2. Comparison of BP network detection rate with other optimization methods

Detection method	Normal record	Attack record	Detection rate (%)			Time-consuming detection (S)
	850	500	R_C	R_w	R_l	
λ BP	784	466	91.3	4.51	10.8	18.7
PSO-BP	815	488	94.5	2.13	7.6	21.4
GA-BP	826	472	95.2	1.25	7.4	20.7
MR-DBN-BP	831	491	96.7	1.09	6.5	21.7

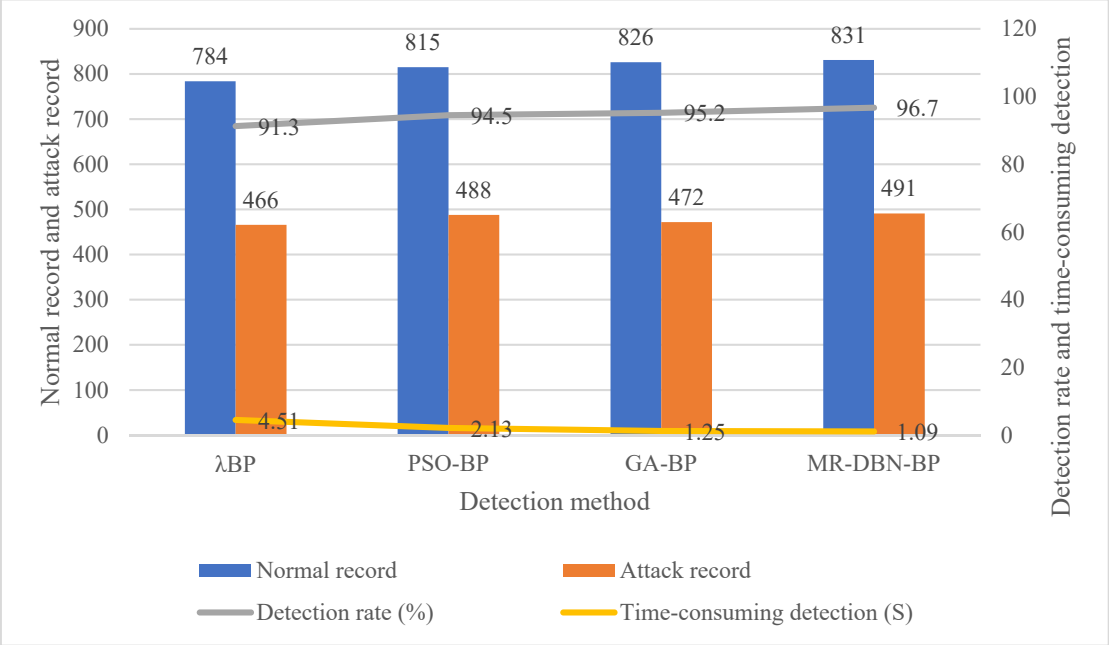


Figure 2. Comparison of network detection rates

According to Figure 2, the detection time of the DBN-optimized BP network in a single machine is slightly longer than other methods, which is 21.7 seconds. However, because this method can select the more essential features of the data during training, the detection accuracy is Compared with other methods, it also effectively reduces the false detection rate and missed detection rate.

5. Conclusions

The rapid development of big data and cloud computing has further ensured the safety and maintenance of power systems. Through the introduction of cloud computing, the detection and prevention of intrusion behaviors in the power system information network is of great significance for effectively improving power information security and ensuring the normal operation of the power system. However, it is undeniable that due to the limitations of time, energy and knowledge, the intrusion detection experiments carried out in this research have certain limitations. In order to obtain a more scientific power information network intrusion detection method, further research is needed.

References

[1] Li S , Wang X P , Wang Q D , et al. Research on intrusion detection based on SMDP reinforcement learning in electric PI network[J]. Electric Power Automation Equipment, 2006, 26(12):75-78.

[2] Li C , Kang H , Chen S . Research on the Service Innovation Path for Information Platform in the Cloud Computing Environment[J]. International Journal of Grid & Distributed Computing, 2016, 9(10):129-140.

[3] Wang T , Zhang G , Yang X , et al. A Trusted and Energy Efficient Approach for Cluster-Based Wireless Sensor Networks[J]. International Journal of Distributed Sensor Networks,2016,(2016-4-10), 2016, 2016:1-13.

[4] Xinhui D , Shuai W , Juan Z . Research on Marine Photovoltaic Power Forecasting Based on Wavelet Transform and Echo State Network[J]. Nephron Clinical Practice, 2017, 24(s2):53-59.

[5] Kb A , Bbg A . Hypothesis Test for Low-rate DDoS Attack Detection in Cloud Computing Environment - ScienceDirect[J]. Procedia Computer Science, 2018, 132:947-955.

[6] Chuang C L , Chiu W Y , Chuang Y C . Dynamic Multiobjective Approach for Power and

Spectrum Allocation in Cognitive Radio Networks[J]. IEEE Systems Journal, 2021, PP(99):1-12.
[7] Chen Y , Li L , Shen Y , et al. Group decision-making framework for site selection of coastal nuclear power plants in a linguistic environment: a sustainability perspective[J]. International Journal of Green Energy, 2021:1-12.
[8] Gulmkhan O , El-Saadany E , Youssef A , et al. Cyber Security of Market-based Congestion Management Methods in Power Distribution Systems[J]. IEEE Transactions on Industrial Informatics, 2021, PP(99):1-1.
[9] Suzuki S , Hiraoka M , Shiraishi T , et al. Power Prediction for Sustainable HPC[J]. Journal of Information Processing, 2021, 29:283-294.
[10] Wang D , Bing-Yu G E , Xuan J X , et al. Research on Operation Test Technology in PI System[J]. Electric PI and Communication Technology, 2017(6):50-55.