International Journal of Security and Its Applications
Vol. 10, No. 11 (2016), pp.303-316
http://dx.doi.org/10.14257/ijsia.2016.10.11.25

Research on an Improved Intrusion Detection Algorithm

Yue Liu [1] and Mei-shan Li [2,*]

[1,2]*College of Information Science &Electronic Technology Jiamusi University, Jiamusi 154007, china*
[1]*friend9023@sina.com and* [2]*limeishanz@163.com*
[2,]**Corresponding author*

Abstract

First of all, the principle of neural network is discussed, and the traditional BP network learning algorithm and BP neural network adaptive learning algorithm are researched. Combining the advantages of two algorithms, the distributed neural network self-learning algorithm is proposed, which is a kind of intrusion detection algorithm using the method of distributed learning to optimize the BP neural network algorithm. Using this algorithm to study and test the network intrusion data, it solves the problem that directly using BP learning caused by the training sample size too large and difficult to convergence. At the same time, the sample training time is shortened, and the BP neural network classification accuracy is improved. Secondly, based on the research of the improved algorithm, this paper gives the specific steps of the algorithm, and uses the improved algorithm to establish mathematical model which is used to analyzing and forecasting. Compared with the traditional BP network learning algorithm and BP neural network adaptive learning algorithm, verify the effectiveness and feasibility of the improved algorithm. Finally, the algorithm is applied to intrusion detection. Through appropriate test method, use the sample data of this paper adopted to verify the example. Through the results of the testing data, it verifies the performance of the distributed neural network self-learning algorithm, and comes to the conclusion.

Keywords*: network security, intrusion detection, back propagation neural network, distributed neural network intrusion algorithm*

1. Introduction

Nowadays, the network has developed rapidly, which makes the network structure more complex, network security and intrusion detection are facing a difficult task. This makes artificial neural network widely used in network security intrusion detection. BP neural network is one of the common neural network technologies. It has the advantages of fast speed, strong analysis ability and so on. Not only application of pattern recognition is very good, but also to promote the development of network intrusion detection technology. However, BP algorithm also has its own shortcomings, such as time-consuming, easy to fall into the local minimum value, *etc*. This will affect its own development. In order to improve the performance of BP algorithm, the research on the performance of the algorithm is also a lot of scholars have been trying to do.

In recent years, most of the domestic and foreign scholars focus on the research of intrusion detection technology based on neural network algorithm. The theory can be used to follow the model of biological information processing, and obtain a theory of intelligent information processing function. It can be used in any test. The emphasis here is to establish a model for the detection of anomaly detection. It can learn some behavior characteristics, performance and some information resources of the object.

Through analyzing and processing these data and using neural network to realize this function. At the same time, the neural network has the ability to learn and identify the

ISSN: 1738-9976 IJSIA

unknown attacks. This greatly improves the performance of IDS. Because it can acquire the latest attack behavior through self learning ability, and can collect their data, update their own sample library, so as to improve the ability to attack cognitive unknown behavior.

For known attacks, the expert system is able to detect, and the unknown attack can do nothing. Other methods can be used to predict the likelihood of attack, but cannot detect the complex form. So the mode is likely to be limited by continuous events. In order to overcome the above problems, the neural network is proposed, which can improve the detection efficiency by simulating the human brain. It is collecting, processing, storing and processing information by simulated human brain, and there are many characteristics of neural network, including learning ability, abstract ability, adaptive ability and parallel computing. The application of neural networks with these features to IDS will make IDS a great advantage:

(1) Neural network by calling their learning ability given a large number of actual attack behavior, let the neural network to learn and get predictive ability, it is able to automatically grasp the inherent relationship of the system. No need to analyze the details of the data distribution of neural network.

(2) In order to train the neural network to reflect the behavior of some new attacks. We need to apply the new attack instances to the neural network, making the IDS enhanced adaptive capability.

(3) After learning the normal mode of operation, it can make the neural network learn to deviate from the normal pattern of the behavior of the system to make response measures, found that the new attack behavior.

(4) Neural network the neural network training by a large number of examples of training, the model can be converted to the judgment and numerical calculation. This can improve the processing efficiency of the system, and it is suitable for real time processing.

Neural network algorithm need to be trained to learn, then to deal with the actual problem. The input variable of neural network module is given, including the learning sample library and network training to achieve the goal. Then some training algorithms are used to study and deal with the sample in the sample library, and finally reach the required training target. After a period of training, the neural network is a form of memory weight value to distinguish aggressive behavior relevant knowledge. Thus, the model is built up in the neural network to identify the attack behavior. When the input of the unknown sample is obtained, it is necessary to analyze and process the sample to achieve the purpose of distinguishing. However, the neural network model to limit their learning ability. Moreover, this kind of technology can judge whether the behavior is abnormal, it is very difficult to identify the type of invasion, and there are many problems in the actual implementation process.

Many scholars put forward a lot of algorithms, such as voting, weighted voting, feature matching, and so on, in order to merge knowledge. These methods reduce the error of the fusion stage by weighting the rules of the merging sub concentration. This is because there is a certain error in the sub concentration of rule learning, which is the error of the existence of a single sub set as a subset of the additional information, in order to improve the performance. However, if the specific subset of this information is too small and the weight is still too small, it will produce a large error, but it does not affect the results of this study. In addition, the decision tree algorithm is a typical hybrid learning method, merging, pruning sub set of rule, make it become the global knowledge. Due to re learning all the regular subsets, small samples are easily overlooked and the problem can be solved, but restructuring knowledge will consume a large amount of computing resources.

A neural network intrusion detection algorithm based on distributed learning is proposed in this paper, intrusion detection algorithm applied to the behavior of distributed network learning. The goal of the algorithm is to effectively study the arbitrary

segmentation of the network data. So it can reduce the requirement of load balancing module.

The neural network has the following characteristics:

1. Learning ability. As mentioned above, the input and output conditions of self-learning is to initialize the neural network. After a given continuing learning and training process, in order to modify the connection path weights between the neurons system rule. As the basis for judging the input conditions, the neural network system can be formed as a benchmark in the basic framework of the network. This process is to correct the weights of the neural network, without user intervention, so the neural network itself has a better learning ability, and does not require user interference.

2. Applicability. The neural network is used in the computer system, because the neural network can realize the logical thinking of the brain. Therefore, the neural network as the human brain, has the characteristics of strong applicability. At the same time, the problem of visual thinking can be solved by neural network, and the problem of visual thinking includes two aspects. The first is the classification of large amounts of data. The second is a complex nonlinear mapping learning process. According to the content of these two aspects, it is difficult to use the traditional artificial intelligence technology to achieve a new breakthrough.

Along with the development of modern technology, the application of neural network has been spread all over the area. For example, the production engineering system modeling will be used in neural networks, fuzzy control on the speech, vision, *etc*. can also be used in neural networks.

3. Generalization ability. The neural network that is trained to end is a common ability, when the following conditions, even if the input conditions have a slight deviation, it will change and do not give the corresponding response. On this issue, the neural network of the high accuracy algorithm caused by the high accuracy algorithm will become defective, and low precision arithmetic problems can be converted to a good neural network for fault tolerance. The above reflects the neural network of the two functions, that is, noise and incomplete function.

4. Distributed storage information. In the neural network, once the neurons are damaged in a node and lose their function, distributed storage can make the normal operation of the neural network system. The system will not be paralyzed because of damage to a node. For a part of the network structure, this feature will be destroyed after learning knowledge led to the destruction of knowledge before the emergence of unstable changes, so that the whole also appears unstable phenomenon.

Topology structure and neural network learning algorithm are two important aspects of neural network theory. Advantages of neural networks are as follows:

1 with computational power, it is closer to a higher level of the human brain.

2. With self-learning, self-organization and comprehensive description of the three skills.

3. Be able to identify objects based on the object feature, the recognition ability is very strong, has a strong robustness

1. The main problem of neural networks is that the system itself cannot explain the process of self - learning and reasoning.

2. In operation, the neural network is to analyze whether the input information is saved and the numerical information is formed. For example, the logic of the configuration information may exist to calculate the inevitable loss.

3. After the system is initialized, the neural network must be self - learning, that is, the learning and training of the network.

The disadvantage of the neural network is that the system cannot ask the necessary questions like the user.

The system will be a number of loopholes, the neural network will not work, when the training data accumulation is not sufficient or not comprehensive.

4. Further improve the algorithm and theoretical knowledge of the support neural network.

The concept of intrusion detection was proposed in 1980; then in 1987 the abstract model of intrusion detection system was firstly proposed [1-4]. The first network intrusion algorithm was a detection method based on pattern matching, which has been developing continuously to be used as network intrusion detection system [5-9]. However, the efficiency of the algorithm performed not ideally in the course of matching data packet. Then Boyer Moore algorithm was introduced by some researchers to the network intrusion detection system. Due to that, the detection system's effect was greatly enhanced. Even some scholars on Boyer Moore algorithm did some improvements to it. Unfortunately with increasing network scale, the single-mode network intrusion detection algorithm becomes difficult to meet development requirements of computer network [10-14].

Artificial intelligence technology (AIT) develops rapidly nowadays [15-16]. Lots of scholars introduced intrusion detection system to RBF neural network. For the traditional RBF neural network algorithm, RBF gradient descent algorithm is a training approach based on temporal network [17-20]. But it needs adjust error weight and threshold, together with back propagation network and learning speed not ideal, and network easily falling into local minimum condition. Those constrain its application for network intrusion detection. During the recognition of intrusion detection, some scholars employed BP neural network approach to do and process data with BP neural network; but time complexity increased along [21].

Here it presents distributive neural network intrusion algorithm. It is a kind of technique with the use of distributive intrusion algorithm for optimizing BP neural network algorithm. It can improve better the efficiency and performance of intrusion detection.

2. Principle of Intrusion Detection Algorithm

Intrusion detection is defined as a process during which an intrusion action is identified, through collecting, analyzing and processing some key information. Such kinds of information exist earlier in computer network or system. From it, discover unreliable, suspicious cases and breach of security rules and perceive any sign which hints any possible attack. Hence the completeness of security structure can be guaranteed. Intrusion detection does not affect network performance even during work. It is a technique which is devised and configured as to enable computer network to work in a secure and dependable environment and report any abnormalities in time. Without it, network performance can hardly be warrantied. Intrusion detection technology can effectively monitor network and improve real-time protection effect of external, internal or false operations by acquiring, processing, fetching information and result handling.

Intrusion detection technology is the key to the intrusion detection system data analysis. Its efficiency is a standard to evaluate the quality of an intrusion detection system. Intrusion detection technology is to analyze the characteristics of the extraction, the combination of reasoning. Anomaly and misuse of two kinds of intrusion detection technologies are included in the classification of intrusion detection technology.

Classification of anomaly intrusion detection technology:

1. The mature detection technology of intrusion detection system is statistical analysis, through the self-learning system combined with statistical principles, to learn the user's habits of learning.

Therefore, the system aims at self-learning, record some normal activities other than study. These activities are not illegal activities, the advantage is that the detection speed, availability is strong. Data mining is a combination of network security management and intrusion detection technology. To study the accuracy of IDS has become the focus of the

discussion. Intrusion detection models built in most fields are based on data mining technology. The theory used in the process of dealing with large data is pattern recognition and artificial intelligence. At the same time, the results of processing data and the technology are combined, after a detailed analysis, can effectively reduce the false alarm rate.

Random process in detection technology, refers to the quantitative description of the dynamic relationship between a series of random events. In the field of computer, usually used to create the corresponding data model. The algorithm of the Markov process of the intrusion detection system is generally combined with the combination of the Markov model, the system uses a conversion matrix,

The classification of the events described in order to change the state of the system represents the state transition and the results. In contrast, the presence of infiltration in the system will result in a smaller state transition probability matrix to determine the intrusion.

Connection model is a mathematical algorithm model, neural network is used to imitate the behavior of animals. It is the basis of neural network detection. Parallel information distributed processing, neural network model based on the neural network model is called.

The neural network system is a nonlinear system with a large number of interconnected nodes. The system is characterized by its dependence on the system to adjust its internal nodes. Realized information processing of the relationship between the methods of information by means of the complexity.

3. Description of BP Algorithm

The basic theory of BP algorithm is feedback guiding, confirming errors of other layers with the help of input and output of each layer and inferring gradually from back to front, and so on. Guiding anywhere is a feature of any activation function. For forward propagation, sample data can pass to output layer through hidden layer, making certain the number of nodes on input same with sample data. If output value in ideal sense is not identical to the feedback result of actual output layer, the system will implement automatically back propagation, returning reversely the acquired error signal by forward-propagating path and modifying weighted value of the path with gradient descent method as possibly to reduce the difference between actual output value and learning target. If output value is beyond target range, repeat the process several times.

4. Deficiency of BP Algorithm and its Improvement

In practice, the utilization of BP algorithm can well solve problems [22]. The method has following drawbacks and weakness:

1. Global optimal solution can't be reached due to existence of local minimum;

2. Training more times reduces remarkably learning efficiency and convergence speed becomes slow;

3. For the lack of theoretical foundations, the number of nodes on hidden layers can only be determined according to experimental guide;

4. During new sample learning, it interfere old samples and those tend to be forgotten.

With regards to learning efficiency and convergence, based on the above points, lots of scholars home and abroad put forward these solutions:

1. Since vibration occurs during the training, which leads to slow convergence, so it needs to change gradient direction and increase momentum items to avoid that;

2. Extend or shorten the step length of BP algorithm, meanwhile adjust adaptively learning rate as per environment changes;

3. After entrance to flat area, introduce new variable gradient factor; attempt to compress the net input of neurons to change the shape of error function as thus to separate from the flat area.

5. Improved BP Neural Network Model

5.1. Selection of Initial Weight

On the basis of randomness thinking, we use principle of statistics to select the algorithm's initial weight to generate randomly a huge number of initial nodes and iterate each time as to choose the best initial weight. Although error function E can gain the optimal solution through initial weight as to raise the possibility of random production, unavoidably it has blindness and randomness. Through BP algorithm, with S-shape function $f(x)=\frac{1}{1+e^{-x}}$, the actual obtained output value ranges [0,1]. So choose initial weight in [-1,1], which can't be too big. By distributive parallel detection mechanism, with stepwise searching method, the problem of selecting initial weight blindly is solved. Divide initial value region H into N equal parts; next choose the least error function E which is relative to the region and divide to N equal parts; at last repeat former steps till error function E won't reduce; end iteration to get the best node. Therefore as long as little enough local behavior appears in the region, it can't be effectively avoided. It's full manifestation of the feature of a parallel mechanism.

5.2 Determination of the Number of Nodes on Input Layers

Determining the dimension of input and output layer is always key to studies on traditional BP network modeling issue. For example when we model complicated system of social or economical system, the existing theoretical knowledge can't interpret completely and reasonably the complicated associations among every factors in such system, and also can't specify what factors closely correlated with variables to our attention. To avoid missing option of those important factors, it's a common way to select some independent variables through qualitative analysis method. If those independent invariables affect fairly dependent variables and are of large quantity, then the model of system is created. If BP network inputs too many independent invariables, it will increase considerably network complexity and time for computer operation, weakening computer performance and affecting computer precision.

5.3. Determination of the Number of both Network Hidden Layers and Hidden Nodes

Input node and BP neural network's topological structure decides its own output nodes, but key factor is hidden nodes' hidden layers and topological structure of those layers. A lot of scholars made in-depth investigations from theoretical aspect and they stated as long as there are two hidden layer nodes, it's possible to divide them arbitrarily. Many years later, Robert Hecht Nielson *et al*. pointed out that as long as there are enough hidden nodes of neural network, it's likely to make nonlinear function approximate any precision. Relatively, it's very difficult to select hidden nodes. On the one hand, the learning process can't converge too few hidden nodes; on the other hand, network performance declines and node redundancy arises because of too big number of hidden nodes. To find suitable hidden nodes, during network learning, in order to get an appropriate neural network model, it's necessary to adjust self-structure and learn adaptively in accordance to environmental requirements.

Suppose O_{pi} is output, the output value generated by the hidden node i when it's learning the pth sample; O_{pi} stands for output, the output value generated by hidden node j when it's learning the pth sample; N is learning amount of samples; so:

$$\begin{aligned} \bar{O}_i &= \frac{1}{N}\sum_{p=1}^{N} O_{pi} \\ \bar{O}_j &= \frac{1}{N}\sum_{p=1}^{N} O_{pj} \end{aligned} \tag{1}$$

Suppose

$$\begin{aligned} x_p &= O_{pi} - \frac{1}{N}\sum_{p=1}^{N} O_{pi} = O_{pi} - \bar{O}_i \\ y_p &= O_{pj} - \frac{1}{N}\sum_{p=1}^{N} O_{pj} = O_{pj} - \bar{O}_j \end{aligned} \tag{2}$$

Then the correlation coefficient between O_{pi} and O_{pj} is:

$$R_{ij} = \frac{\sum_{P=1}^{N} x_p y_p}{\sum_{P=1}^{N} x_p^2 . \sum_{P=1}^{N} y_p^2} \tag{3}$$

If the R_{ij} is more close to ± 1, it indicates that the linear correlation between O_{pi} and O_{pj} is greater than that of the two sequences. The degree of dispersion of each other is the smaller. In contrast, the linear correlation between the two sequences is smaller, and the degree of dispersion of each other is greater.

6. Distributive Neural Network Intrusion Algorithm

6.1. Relevant Definitions and Rules

If the study on data which is divided by learning result is finished, some rather complete network actions are included in data of every partition and also some show incompleteness. In this case, we need learn the result of sample space, which is again mentioned in the middle part, in order to form complete network behavior knowledge. In the course, features of some network behaviors are formed may because knowledge can't be partitioned. The problem has been solved but in the meantime it decreases the generalization ability of learning achievement. So it's better to use pruning algorithm to realize structure risk minimization principle, with every trained nerve cell as candidate for pruning and similar nerve cells merged to one or more neurons.

Definition 1: relative coefficient of hidden node i and j on the same layer is:

$$\rho_{ij} = \mid R_{ij} \mid = \frac{\sum_{P=1}^{N} x_p y_p}{\sum_{P=1}^{N} x_p^2 . \sum_{P=1}^{N} y_p^2} \tag{4}$$

The degree of correlation of hidden nodes i and j expressed by ρ_{ij}, if ρ_{ij} is too large, then repeat the node i and j function, compress and merge.

Definition 2: The divergence degree of the sample is S_i, and the definition is as follows:

$$S_i = \frac{1}{N}\sum_{P=1}^{N} O_{pi}^2 - \bar{O}_i^2 \tag{5}$$

If the S_i is too small, it shows that there is little change in the output value of the hidden node i, there is no network training, you can delete it.

To sum up, the relevant rules of dynamic merging and deleting nodes are as follows:

Rule 1: If $|\rho_{ij}| \geq C_1$ and $S_i, S_j \geq C_2$, i and j can be made one

Rule 2: When the $S_i < C_2$ condition is satisfied, node i can be merged with the threshold, delete the node i.

6.2. Steps of the Algorithm

Distributive neural network intrusion algorithm steps are as follows:

1. Based on characteristics of BP neural network model, construct an initial structure, *i.e*. one input layer, one hidden layer, one output layer and many enough hidden nodes;

2. Through studies on real problems, use adaptive method to select variables which affect dependent variables the most to determine the number of output nodes;

3. With reference to practical problems, confirm the number of output layer nodes; use pruning algorithm to initialize variables; $\in$is learning precision; M_0 is iterative step; R is upper limit of hidden node quantity; η is initial value of learning parameters; a is momentum coefficient; other constants are respectively C_1 and C_2;

4. According to distributive parallel principle, divide tolerance region H of initial weight into N equal partitions, put as $\beta_1, \beta, ..., \beta_N$;

5. Input given learning sample to make sample parameter value in [0,1];

6. Based on above steps, produce randomly in selected little region $\beta_i (\beta_i \subset H)$ the required initial weight;

7. Do learning of network by combining BP algorithm with distributive principle;

8. Judge iterative steps and detect if iterative steps surpass specified steps or learning precision meets requirement; if yes, go to 9; or turn back to 7;

9. Adopt parallel principle to calculate relevant coefficient and divergence degree between hidden nodes;

10. In the end, confirm if learning precision lives up to requirement or iterative step is over preset step; if yes, the algorithm terminates; or return to 6.

7. Modeling of Neural Network Intrusion Algorithm on the Distributed Basis

7.1. Utilize Adaptive Method to Choose Output Variables of BP Neural Network

We choose a few key factors which have impacts on consumption level of Chengdu city to establish the relationship model between them and total retail sales of consumer goods of Jiamusi City. We collect data from January 2014 to September 2015 as sample interval. The model of intrusion algorithm based on distributed neural network. It is shown in Table 1.

Table 1. List of Total Retail Sales of Social Consumer Goods and the Related Factors of Heilongjiang

Vari / Date	y	p	M_0	V1	V2	V3	V4	V5	V6
2014.1	30.6	7.48	12.5	102.9	-0.98	825.3	10.6	7.2	14.6
2014.2	30.65	7.23	12.5	104.2	0.47	827.3	10.5	18.3	17.5
2014.3	30.69	7.23	13.5	103.7	1.11	827.6	9.6	8.6	18.5
2014.4	29.68	7.23	17.5	103.2	0	827.0	11.5	8.8	20.5
2014.5	32.64	7.23	17.7	103.8	0.5	821.2	10.1	6.5	19.2
2014.6	33.15	7.23	19	102.5	-0.26	827.3	9.5	7.6	19.3
2014.7	33.57	7.23	16.8	101.6	0.87	827.8	12.5	9.8	17.6
2014.8	32.77	7.23	14.6	101.8	1.28	827.6	12.6	11.3	17.6
2014.9	33.45	7.23	12.4	100	0.54	825.9	12.1	10.5	16.5
2014.10	35.24	5.68	12.8	99.8	-0.65	824.6	10.8	9.8	16.7
2014.11	31.89	5.68	14.5	101.5	0.45	825.6	9.8	9.1	16.5
2014.12	39.54	5.68	15.4	100.1	0.35	825.6	10.5	8.3	16.7
2015.1	36.1	5.68	13.4	100.5	0.031	826.7	10.5	8.7	17.6
2015.2	35.87	5.68	11.8	99.6	0.14	826.5	10.3	24.8	17.5
2015.3	33.75	5.22	9.9	100.5	-0.5	825.4	10.5	8.5	15.8
2015.4	32.4	5.22	9.7	99.8	-0.04	831.2	15.5	8.6	14.8
2015.5	33.58	5.22	10.5	99.9	-0.74	830.2	15.8	-1.5	15.8
2015.6	34.5	5.22	6.5	98.5	2.08	827.6	16.5	1.8	15.4
2015.7	33.87	4.55	8.5	98.5	-0.06	826.8	22.8	3.6	15.8
2015.8	34.58	4.77	9.8	98.5	2.05	824.6	26.8	-2.8	15.8
2015.9	35.28	4.77	9.7	98.6	2.08	822.5	27.6	-1.8	14.9

With computer and adaptive principle and method, we select out input variables of BP neural network. Such variables are those relating with total retail sales of consumer goods of Jiamusi City.

Input variables:

y: Total retail sales of social consumer goods in Heilongjiang (100 million yuan)

p the one-year deposit rate (%)

M_0 : cash flow

V1: national consumer price index (%)

V2 : national budget deficit (100 million yuan)

V3 : RMB against the U.S. dollar ($one hundred)

V4 : fixed asset investment growth (100 million yuan)

V5: national export growth (%)

V6 : broad money supply

We choose totally 17 pairs of data from May 2014 to December 2015; then do simulation of them to optimize; get six independent variables respectively v1, v2, v3, v4, v5 and v6, as well as relevant model:

$$y = a_0 + \sum_{i=1}^{6} a_i v_i \qquad (6)$$

Where,

$a_0 = 470.5688$, $a_1 = -0.6044$, $a_2 = 0.6856$, $a_3 = -0.4489$, $a_4 = -0.3580$, $a_5 = -0.0007$ $a_6 = -0.5967$

7.2. Establish the Model with the Use of Distributed Neural Network Intrusion Algorithm

Input variables of BP neural network are screened out by adaptive method, which are totally 6; so the number of input layer node of BP neural network is 6.

The total retail sales of consumer goods of heilongjiang city is output variable of BP neural network; so the number of output layer node is 1 and the number of hidden layer node is 15; the relative coefficient and divergence of hidden node during adjustment are shown in Table 2.

Table 2. The Correlation Coefficient and Divergence When the Hidden Nodes Are Adjusted

i / j	4	6	9	11	12	13	15	s_i
4	1.0	0.24	0.92	0.98	0.38	0.03	0.48	0.05
6	0.25	1.0	0.24	0.37	0.52	0.44	0.18	0.03
9	0.93	0.24	1.0	0.93	0.32	0.25	0.07	0.04
11	0.92	0.38	0.93	1.0	0.52	0.38	0.45	0.02
12	0.38	0.52	0.32	0.51	1.0	0.65	0.99	0.06
13	0.04	0.44	0.23	0.38	0.67	1.0	0.54	0.05
15	0.48	0.12	0.07	0.49	0.01	0.54	1.0	0.07

8. Experiment Design and Discussion

So far, the intrusion detection dataset used mostly is security audit dataset KDD CUP99. The dataset is considered as fundamental data of many research and paper achievements, with obvious validated effect. However apparent differences exist between analyzing and utilizing the dataset. Here we use KDD CUP99 to exemplify the model of the proposed algorithm and analyze experimental result before making according conclusions.

8.1. Input Data and Output Data

8.1.1. Input Data

Input data means intrusion detection data, which are generally given safe audit dataset used as processed object.

1. All:training data set: kddcup.data.gz,18M

: test data set: kddcup.testdata.unlabeled.gz，11.2M

2. 10% data set: kddcup.data_10_percent.gz，2.1M

: test data set: kddcup.newtestdata.unlabeled_10_percent.gz,1.4M

3 corrected.gz With labeled attacks, researchers can detect the results of their algorithms and compare the results of the data set.

The testing dataset in the paper is part of dataset, only 10% of it as training data; use labels of testing dataset as testing data. Here we make it 521028 records including 70552 pieces of normal behavior record, and 351446 pieces of intrusion record, the dataset containing 42 kinds of intrusion.

8.1.2. Output Data

Output data is detection result of intrusion data. In the testing dataset, there are 42 kinds of intrusion behaviors. Firstly we classify them to different broad categories:

Probe：{portsweep，mscan，saint，satan，ipsweep，nmap}

DoS：{udpstorm，smurf，pod，land，processtable，warezmaster，apache2，mailbomb，Neptune，back，teardrop}

U2R：{httptunnel，tip_write，sqlattack，xterm，multihop，buffer_overflow，perl，loadmodule，rootkit，ps}

R2L：{guess_passwd，phf，snmpguess，named，imap，snmpgetattack，xlock，sendmail，xsnoop，worm}

Then, submitted to the training system for testing. Test data is the output data. It is shown in Table 3.

Table 3. The Test Results Based On Test Data Set

		Test results					
		Normal	Probe	DoS	U2R	R2L	TR(%)
Intrusion behavior	Normal	58267	367	251	801	920	94.8
	Probe	336	3418	304	4	205	81.4
	DoS	5890	590	232867	783	580	97.6
	U2R	180	24	20	26	3	9.3
	R2L	14321	5	6	10	141	0.3
	TR(%)	71.9	75.2	99.8	1.5	7.4	

8.2. Data Processing

Data KDD CUP99 owns 42 attributes; except the last attribute, it has totally 7 kinds of symbolic variables and 34 kinds of continuous variables. For instance, flag attribute is symbol variable, with 11 kinds of identity, which are respectively SH-1, RSTR-2，OTH-3，S0-4，REJ-5，S2-6，SF-7，S1-8，S3-9，RSTOS0-10，RSTO-11

We use matrix A to express variable value of symbol; A_i is symbol variable which contains g kinds of identity, defined like:

$$A_i = \{A_i^1, A_i^2, ..., A_i^g\} \tag{7}$$

$b_{1\times n}$ is continuous variable, including n continuous variables; X is input variable, defined as:

$$X = \{A_1, ..., A_m, b_1, ..., b_{n-m}\} \tag{8}$$

$|x_i - y_i|$ is continuous variable in Minkowski; normalize it and then compute it; here we use a method similar to Hamming distance to calculate it; y_i does not mean a quantity value but a mark, *i.e*. the relative symbolic attribute A_i of synaptic weight (neuron) N; if $x_i = A_i^k (k \in 1,.., g)$, then:

$$|x_i - y_i| = 1 - \frac{c_k}{C} \tag{9}$$

In short, according to the above formula, calculate the value of symbolic variable Minkowski measure; if A_i^k is identity value of one training sample's symbolic variable A_i, A_i then for more times the neuron N learns A_i^k, the more similar it looks with the sample. In the meantime, normalization method is realized.

With given input data, we compare traditional BP network intrusion algorithm and improved distributed neural network intrusion algorithm. The false alarm rate, detection rate and missing report rate is listed in Table 4.

Detection model of evaluation criteria for the false alarm rate, detection rate and false negative rate, which is defined as follows:

False positive rate = the number of normal samples and normal samples of the normal samples.

Detection rate = the number of detected samples / normal samples

Table 4. The Comparison of Algorithms Performance

Performance / Algorithm	False alarm rate (%)	Detection rate (%)	The false negative rate (%)
Traditional BP network learning algorithm	2.6	89.5	10.5
Distributed neural network intrusion algorithm	0.4	91.6	7.5

Comparison of the execution time of the traditional BP network learning algorithm and the distributed neural network intrusion algorithm. It is shown in Figure 1.

8.2.1. Analysis of Experimental Results

The above results reveal that traditional BP neural network algorithm has very small computational amount but very slow operating speed; besides running time takes longer and output results vibrate heavily and fluctuate obviously; while the improved algorithm has huge computational amount, with complicated input parameters, due to improvements, the distributed neural network intrusion algorithm takes far shorter running time than BP algorithm and it's more stable with little vibration.

8.2.2. Experimental Conclusions

Through analysis of experimental results, we proved that the improved algorithm advanced greatly in learning speed, with shorter training time and higher model precision. For the same dataset, we compared it with traditional BP algorithm. It confirmed the effectiveness and feasibility of the distributed neural network intrusion algorithm. Our findings suggest that the method here realized higher detection efficiency and lower false alarm rate, fulfilling the expected objective.

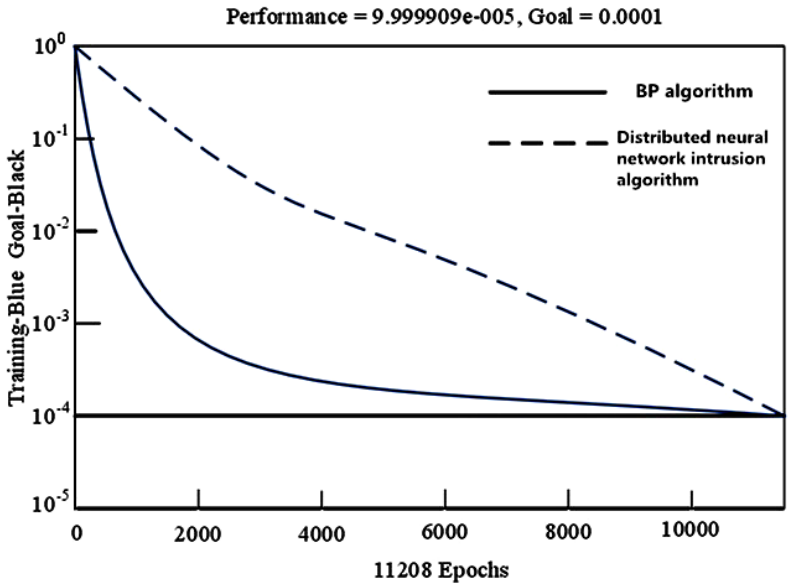

Figure 1. The Time of the Executions of Two Algorithms

9. Conclusion

The paper introduced neural network technique and generation and principle of such kind of network. Based on the analysis of existing BP algorithm, it proposed neural network intrusion detection algorithm based on distributed learning, *i.e*. distributed neural network intrusion detection algorithm, which was improved on the basis of traditional BP algorithm. The algorithm's time complexity was deduced and efficiency was enhanced. A good foundation was laid to the following example verification with improved algorithm to create model and analyze characteristics of the model. With given intrusion detection testing data, it validated and compared with traditional BP algorithm. The improved method advanced a lot the learning speed, effectively compressing training time. To a large degree, it increased the stability and convergence of data detection. The method achieved higher detection efficiency and lower false alarm rate.

Acknowledgement

This work was supported by the key scientific and research projects of jiamusi university (12Z1201520) and surface project on scientific and research projects of jiamusi university (13Z1201576)

References

[1] Liu yanheng, Tian Dah, Yu Xuegang, Wang Jian. Large networks of distributed learning based intrusion detection algorithm . Journal of software, 2008,04:993-1003.
[2] Ji Genlin, the revitalization of the Chinese, Xue Yun Cheng. Neural network integration of distributed intrusion detection method. Journal of Nanjing University of Aeronautics and Astronautics, 2007,02:231-235.
[3] Li Ming. Xianyi. Based on Improved BP neural network intelligent agent distributed intrusion detection system . Computer applications and software. 2009,01:105-107.
[4] Fu Yadan, Yang Geng. Based on optical fiber sensing of intrusion detection signal extraction and recognition algorithm . Computer technology and development, 2014,06:161-165.
[5] Wu Chunqiong. Neural network intrusion detection method based on improved ant colony algorithm . Journal of Fuzhou University (NATURAL SCIENCE EDITION), 2013,05:845-849.
[6] Fan Ying. Improved ant colony algorithm combined with BP network for intrusion detection. Journal of Liaoning Technical University (NATURAL SCIENCE EDITION), 2010,05:966-969.
[7] Liu Yun. Research on Intrusion Detection Technology in cloud computing environment. Shandong Normal University, 2015
[8] Ouyang. Research and implementation of network intrusion detection system based on neural network BP algorithm . Southeast University, 2006
[9] Zhang Ying. Research on Intrusion Detection Technology Based on genetic algorithm and neural network in TDCS network. LanZhou JiaoTong University, 2012
[10] Yan Hao. Research on Intrusion Detection Algorithm Based on BP neural network. South-Central University For Nationalities, 2013
[11] Jiang Changdong. Research on the localization algorithm of continuous distributed optical fiber sensing security system. Dalian Maritime University, 2013
[12] Yu Xiaowang. Multi target recognition. Xiangtan University distributed optical fiber sensor intrusion signal of perimeter security, 2014
[13] Lai Jibao. Network security situation awareness heterogeneous sensors on some key technologies. Harbin Engineering University, 2009
[14] Luolin. Research on Key Technologies of distributed denial of service attack defense. Northeastern University, 2009
[15] Zhou guiwang. A distributed intrusion detection model based on BP neural network was used to construct . Shanxi University, 2011
[16] Xie Yuxin. Distributed ensemble learning in Intrusion Detection Research on the application. Jilin University, 2012
[17] Liu Yantao. Intrusion detection model based on hybrid neural network technology. Northeast Normal University, 2010
[18] Lu Rong. Research on Distributed Intrusion Detection Algorithm . Jilin University, 2011
[19] Xu Zhenhua. Improved algorithm of Distributed Intrusion Detection Model Based on BP neural network . Network security technology and application, 2016,02:77-78.

[20] Guo Dechao. Intrusion detection system based on genetic algorithm and wavelet neural network. Jinan University, 2008
[21] Ju Jin Ju, Xu Hui. Distributed intrusion detection system based on neural network in wireless sensor networks . instrument technology, 2011,08:67-70.
[22] Hu Jinbin, Tang Xuqing. BP algorithm of artificial neural network and its application. Information technology, 2004,28 (4):1-4

Authors

Yue Liu, She received her B.S degree from Daqing Petroleum Institute and received her M.S degree from Beijing University of Posts and Telecommunications. She is a Lecturer in College of Information Science &Electronic Technology Jiamusi University. Her research interests include network intelligence and applications.

Mei-shan Li, She received her M.S degree from Harbin normal university. She is a lecturer in College of Information Science &Electronic Technology Jiamusi University. Her research interests include computer vision and image processing.

