Network intrusion detection model based on improved BP algorithm

Luo Qun[1]*, Liu Zhen-dong[1]

[1] *Chongqing Creation Vocational College, YongChuan, Chongqing, China, 402160*

Received 1 November 2014, www.cmnt.lv

Abstract

With the rapid development of network, the performance of intrusion detection system that ensures the security of network information has been paid more and more attention. In order to overcome the disadvantages of RBF neural network, a novel RBF scheme based on improved particle swarm optimization is proposed, which can overcome the disadvantage of premature convergence. The experiment result shows that the proposed algorithm has better detection rate and false positive rate than traditional algorithms based on RBF, and it can provide important reference for network intrusion detection system in practice.

Keywords: network intrusion detection; RBF neural network; particle swarm optimization.

1 Introduction

With the rapid development of Internet and computer, the application of the network also increases, but the systems and network information are being threatened by the invasion, which is becoming more and more serious, coupled with the mature knowledge, hackers technical level unceasing enhancement, and the diversified attacker means, the traditional security techniques have been unable to keep up with the pace of the network security needs [1]. The network information and system security problems are becoming more and more serious.

Intrusion detection system has gradually evolved into one kind of active network security defense technologies, which effectively makes up for the deficiency of the traditional security technologies [2]. Especially in recent years, IDS (Intrusion Detection System, IDS) has made greater progress. An distributed intrusion detection system model based on multi-agent was proposed by M.Chang-lou[3]. Detection feature selection based on improved quantum genetic algorithm was proposed by Liu Jun[4]. Design of intrusion detection system based on neural network was proposed by Z.Ping-hui [5]. Intrusion detection system design in wireless LANs based on optimized BP algorithm was proposed by L.Feng-chun, Z.Hao and Z.Bao-hua[6]. Fault detection and diagnosis of the gearbox in marine propulsion system based on bi-spectrum analysis and artificial neural networks was proposed by Z.Li[7]. Support vector machine based on genetic algorithm for network intrusion prediction was proposed by Xie Zhiqiang [8]. Intrusion detection method using neural networks based on the reduction of characteristics was proposed by Iren Lorenzo-Fonseca [9]. Network intrusion detection system based on expert system and neural network was proposed by ZHANG Ren-shang [10]. A neural network-based intrusion detector to recognize novel attacks was proposed by Lee, S.C [11]. Intrusion detection system based on fuzzy technology was proposed by Zhixin SUN [12]. Data integration system for IDS based multi-agent systems was proposed by E. C.Claudino[14]. Utilizing fuzzy logic and trend analysis for effective intrusion detection was proposed by MART INB [15]. But network attack means constantly change, the shortcomings of the traditional intrusion detection algorithm are exposed gradually [16], especially in the face of unknown attack types, it cannot adapt to the environment, and it cannot extend its performance. Artificial neural network with the self-learning and adaptive ability is a good way to solve the above problems.

The paper is organized as follows. In the next section, model of RBF neural network is proposed. In Section 3, RBF scheme based on improved particle swarm optimization is proposed. In Section 4, in order to test the performance of proposed algorithm, experiment is carried out to test its efficiency and the proposed algorithm is compared with other algorithms. Finally, section 5 gives some conclusions.

2 RBF neural networks

RBF neural network as shown in figure 1 is a kind of three-layer feed-forward neural network, which includes input layer, hidden layer and output layer. RBF neural network is a local approximation network, for each input and output data, only a small amount of weight need to be adjusted, and it has advantages of fast learning speed, global approximation and the best approximation performance [9]. A RBF neural network consists of n number of input nodes, m number of hidden layer nodes and one output node. Hidden layer node is RBF function, which can be represented as

* ***Corresponding author's*** e-mail: luozhiqiong121@163.com

$$h = \exp\left[-\frac{\|x-c\|}{2\sigma^2}\right]. \tag{1}$$

σ represents width of RBF hidden layer node, c represents the centre of the i-th RBF hidden layer and w represents output weight. In order to obtain RBF neural network of high performance, the best σ, w, and c should be obtained.

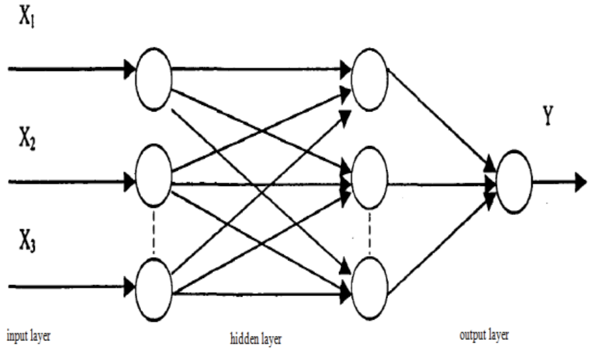

FIGURE 1 RBF neural network structure

Optimization methods mainly include gradient descent algorithm, genetic algorithm (GA) and particle swarm optimization algorithm (Particle Swarm Optimization, PSO). Gradient descent algorithm has fast search speed, and computational complexity is high. Traditional GA and PSO only choose RBF neural network parameters and intrusion characteristic, ignoring the connection between network intrusion characteristics and parameters of RBF neural network, which is difficult to obtain the parameters of the model to predict the overall optimal performance.

3 Production-distribution schemes based on genetic harmony algorithm RBF

The improved PSO algorithm takes each particle as a charged particle in space, and each particle charge is determined by fitness value of the objective function to be optimized. The charge also determines the attraction or repulsion of this particle to other particles. If fitness is optimal, attraction is stronger. If fitness is inferior, attraction is weaker. Then we use the charge provide acceleration for each particle to correct velocity updating formula of the algorithm. The charge of particle i is

$$q_i = \exp\left(-n\frac{f(x_i)-f(p_g)}{\sum_{k=1}^{m}(f(x_k)-f(p_g))}\right) \tag{2}$$

$f(x_i)$ represents fitness value of particle i, and $f(p_g)$ represents the optimal fitness value at present. Supposing particle 1 is better than particle 2, and particle 3 is inferior to particle 1, so that particle 2 has attractive force F_{21} to particle 1, and particle 3 will have a repulsive force F_{31} to the particle 1. The two forces is superimposed to produce joint force, which can make particle move to a better direction, as is shown in figure 2.

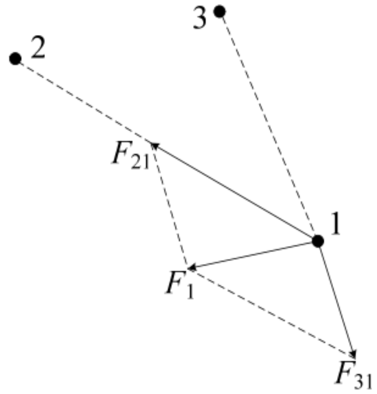

FIGURE 2 Particle move direction

After comparing fitness value of object function of two particles, the direction of the force between two particles is determined. The component force of other particles is calculated by

$$F_i^k = (x_k - x_i)\frac{q_i q_k}{\|x_i - x_k\|_2},\ f(x_k) < f(x_i) \tag{3}$$

$x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ represents current position of the i-th particle and its speed is $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})$. Particle d, which is farthest from the optimal particle, has a disturbance process compared with force of the other particles. When particles ignore some searching area, premature convergence will occur. In order to effectively avoid premature, component force of particle d, which is the farthest from the current optimal particle, is given certain disturbance, which can be represents by

$$F_d^k = (x_k - x_d)\frac{r q_d q_k}{\|x_d - x_k\|_2},\ f(x_k) < f(x_d). \tag{4}$$

r is a random number between 0 and 1. For two particles, particles with better fitness value will attract another particle, attraction force between particles lead particles move to a better area. On the other hand, particles with inferior fitness value will reject another particle, and repulsive force pushed particles to unsearched area. The object of calculating superimposed force is to make particle move to the area that is better than itself and move it far away from the area that is worse than itself. Speed update formula and position update formula are

$$\begin{aligned} v_{ij} &= w \cdot v_{ij} + c_1 r_1 (pbest_{ij} - x_{ij}) + c_2 r_2 (gbest_j - x_{ij}) + a_{ij} \\ x_{ij}(k+1) &= x_{ij}(k) + v_{ij}(k+1) \end{aligned} \tag{5}$$

$a_{ij} = r_3 \frac{F_{ij}}{\|F_{ij}\|_2}$ And r_3 is a random number between 0 and 1. In the proposed algorithm, the particles should be encoded, which includes c, σ, w, v and fitness f_i, c, σ, and w correspond to the position of particle. Supposing there are k number of centres, one output node and each centre is a vector of n dimension. The encoding structure is $c_{11}c_{12}\cdots c_{1n}\sigma_1 w_1 \cdots c_{21}c_{22}\cdots c_{2n}\sigma_2 w_2 \cdots c_{k1}c_{k2}\cdots c_{kn}\sigma_k w_k$, $v_1 v_2 \cdots v_{(k\cdot(n+1)+k)}$, f_i.

$$f_i = \frac{1}{2}\sum_{i=1}^{N}(y(k) - y_m(k))^2 . \tag{6}$$

f_i represents objective function and N represents the number of samples. $y(k)$ represents the expected output value and $y_m(k)$ represents the actual output value. The proposed RBF neural network based on improved particle swarm optimization is as follows.

Step1.The sample data is collected.

Step2.Determine the number of hidden layer neuron and calculate σ, w, according to the method proposed by Sun Dan.

Step3.Initialize the position and speed of particles, the individual optimal value $p_i = (p_{i1}, p_{i2}, \dots, p_{in})$, global optimal value $p_g = (p_{g1}, p_{g2}, \dots, p_{gn})$ and superimposed force F_i.

Step4. For each particle i, carry out the following operation.

Calculate its superimposed force and component force.

Update the particle speed and position.

Calculate the fitness value $f(x_i)$.

If $f(x_i)$ is better than fitness value of p_i, p_i is the current position of x_i.

If $f(x_i)$ is better than fitness value of p_g, p_g is the current position of x_i.

Step5. If it meets stopping condition of improved PSO, turn to step 6. Otherwise turn to step 4.

Step6. Decode the global extreme value to obtain the corresponding parameter of RBF neural network. If it meets stopping condition of RBF, the algorithm stops. Otherwise turn to step 3.

4 Production-distribution schemes based on genetic harmony algorithm

The experimental data set comes from KDDCUP99 dataset of MIT Lincoln laboratory data and attack types include four categories. Four groups of sample data are randomly chosen from the data set. Each data set contains 10000 normal data and 100 abnormal data, a total of 10100 data records. In order to test the attack detection effect of different kinds of algorithms, the number of specific types of attacks of four sample data sets is not the same. In the experiment, under the same test environment, the selected sample data sets are used to test the performance of proposed scheme, PSO_RBF and RBF respectively. Table 2 is intrusion detection effect of three algorithms.

It can be seen that the proposed algorithm has the best detection performance. For test set 1, detection rate and false positive rate of the algorithm are respectively 92.61% and 2.85%, and the corresponding detection rate and false positive rate of the PSO-RBF algorithm are respectively 85.45% and 5.56%, rate corresponding to RBF are 84.61% and 6.54%. From the above data, we can see that the proposed algorithm can significantly improve detection rate, and also has significantly reduced false positive rate. Similarly, it can also be concluded that the rest of the three sets of test data also satisfy the above laws.

TABLE1. Data sample set

Test set	Normal	DOS	R2L	U2R	Probing
1	10000	30	20	25	25
2	10000	35	15	20	30
3	10000	40	10	15	35
4	10000	45	5	10	40

TABLE2. Intrusion detection effect of three algorithms

Test set	Proposed scheme		PSO-RBF		RBF	
	DR	FPR	DR	FPR	DR	FPR
1	92.61%	2.85%	85.45%	5.56%	84.61%	6.54%
2	95.16%	2.21%	87.27%	4.44%	86.26%	5.51%
3	92.36%	2.86%	92.47%	3.36%	89.34%	3.69%
4	97.65%	2.76%	94.38%	2.98%	90.58%	3.23%

5 Conclusions

With the rapid development of network, how to ensure.
The security of network information becomes more and more important. As an active security technology, intrusion detection has been paid more and more attention. First of all, the principle of RBF neural network is discussed. Then a novel intrusion detection algorithm based on RBF and improved PSO is proposed, which is a kind of intrusion detection algorithm using the method of improved PSO to optimize the parameters of BP neural network algorithm. The experiment result shows that the proposed algorithm has better detection rate and false positive rate, which can provide reference for actual network intrusion detection system.

References

[1] X. Zhao, R. Jing and M. Gu 2002 Adaptive intrusion detection algorithm based on rough sets, *J T singhua Univ (Sci & Tech)*, 1(48), 1165-1168

[2] Yong Hou,Xue feng Zheng 2010 Quantum Growing Hierarchical Self Organized Map-based Intrusion Detection.*International Conference on System Science Engineering Design and Manufacturing Information* 110-115

[3] M.Chang-lou and L.Yong-qing 2009 An Distributed Intrusion Detection System Model Based on Multi-agent *Computer & Digital Engineering* 37(6), 102-106

[4] Liu Jun, Di Wenhui.Intrusion 2011 Detection Feature Selection Based on Improved Quantum Genetic Algorithm *Computer Measurement & Control* 04-09

[5] Z.Ping-hui 2009 Design of Intrusion Detection System Based on Neural Network *Microelectronics &Computer* 26(8), 240-242

[6] L.Feng-chun, Z.Hao and Z. Bao-hua 2010 Intrusion detection system design in wireless LANs based on optimized BP algorithm *Journal of University of Science and Technology of China* 40(10) 1096-1100

[7] Z. Li, X. Yan, C. Yuan, J. Zhao and Z. Peng 2011 Fault detection and diagnosis of the gearbox in marine propulsion system based on bispectrum analysis and artificial neural networks, *Journal of Marine Science and Application* 17-24

[8] Xie Zhiqiang 2010 Support vector machines based on genetic algorithm for network intrusion prediction, *Journal of Computer Simulation*, 27, 110-113

[9] IrenLorenzo-Fonseca, Francisco Macia-Perez, Francisco Jose Mora-Gimeno 2009 Intrusion detection method using neural networks based on the Reduction of Characteristics.*Proceedings of the 10th International Work-Conference on Artificial Neural Networks* 1296-1303

[10]ZHANG Ren-shang 2012 Network Intrusion Detection System Based on Expert System and Neural Network *Computer Simulation* 29(9), 162-165

[11]Lee, S.C. and D.V. Heinbuch 2001 Training a neural network-based intrusion detector to recognize novel attacks. *IEEE Trans. Systems, Man and Cybernetics Part A Systems and Humans* 294-299

[12]Zhixin SUN, Hongxia XU 2006 Research of Fuzzy technology in Intrusion Detection System *Journal of Nanjing University of Posts and Telecommunications* 26(4), 73-78

[13]Sun Dan, li-ming wan, Yanfeng Sun 2010 An improved RBF neural network hybrid learning algorithm *Journal of Ji Lin university* 48(5), 817-822

[14]E.C.Claudino, Z.Abdelouahab, M.M.eixeira 2006 Management and integration of information in intrusion detection system: Data integration system for IDS based multi-agent systems *Pro. 2006 IEEE/WIC/ACM Inter. Conf*, 49-52

[15]MART INB, ROSSOUW S 2013 Utilizing fuzzy logic and trend analysis for effective intrusion detection *Computers and Security* 22(5), 423-424

[16]Li jie YANG, BoYANG 2004 Neural Network Applied to Intrusion Detection *Computers and Security* 18(4), 69-71

Author

< Luo Qun >, <19811.02>,< Chongqing, China>

Current position, grades: the college lecturer of Department of Information Engineering, Chongqing Creation Vocational College, China.
University studies: received her Master's degree in Computer application technology from Chongqing University in China.
Scientific interest: Her research interest fields include Computer Network Technology, Semantic Network, etc.
Publications: more than 12 papers published in various journals.
Experience: She has teaching experience of 10 years, and has completed three scientific research projects.

< Liu zhen-dong >, <1975.10>, < Chongqing, China>

Current position, grades: college lecturer of Department of Information Engineering, Chongqing Creation Vocational College, China.
University studies: received his Bachelor's degree in Computer application technology from JISHOU University in China.
Scientific interest: His research interest fields include Computer Network Technology.
Publications: more than 8 papers published in various journals.
Experience: He has teaching experience of 6 years, has completed three scientific research projects.