Volume 5, Issue 6, June 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Optimizing and Tuning RBF Parameters using QPSO Algorithm for Anomaly Detection in Network Intrusion Detection System

Henali Sheth* Computer Department LJIET-Ahmedabad, India
Prof. Bhavin Shah M.C.A. Programme LJIMS-Ahmedabad, India
Asst. Prof. Shruti Yagnik Computer Department LJIET-Ahmedabad, India

Abstract— Intrusion detection system (IDS) is a kind of security software which inspects all incoming and outgoing network traffic and generates alerts if any attack or unusual behaviour is found in a network. This is helpful in providing security. This paper describes RBF neural network approach for IDS. RBF is a feed forward and supervise technique of neural network. RBF approach has good classification ability but its performance depends on its parameters. Based on survey we find that RBF approach has high false alarm as one of major short comings. To overcome this problem, we need to do optimization and proper tuning of RBF parameters. Optimization of RBF parameters is done by implementing QPSO algorithm and tuning of parameters is done as mentioned in this paper. The proposed algorithm also helps to find which parameter optimizing gives better performance for NIDS. The experimental results of detection rate and false rate is improved which is 99.481% and 0.0169948 respectively. The result of using QPSO algorithm along with RBF is better than conventional RBF.

Keywords— Intrusion Detection System (IDS), Network Intrusion Detection System (NIDS), Particle swarm optimization (PSO), Quantum Behaved Particle Swarm Optimization (QPSO), Radial basis function (RBF).

I. INTRODUCTION

Internet threat are increasing nowadays, therefore automated system such as IDS is used to detect malicious activities [1] [19]. Various approaches are adopted to build IDS as described in [2] [10] [14] [21]. Out of these, neural network based model have become a promising AI approach for IDS [2] [10] [20] [21]. In this paper detection engine for IDS is trained using Radial Basis Function (RBF) Neural Network. This approach will learn the behaviour of actors (i.e. users, intruders, etc) in the system [10]. It will learn to predict next command based on some previous command by a legitimate user. So RBF can truly detect anomaly network intrusion, if the behaviour is different or deviated from actual user. RBF performance is moderate [10] and is more suitable approach for IDS in order to tackle present issues such as regular updating, detection rate, false positive, false negative, suitability and flexibility.

Moreover RBF is local approaching technique. Sometimes RBF due to local approaching stuck in the problem of local minima. This affects the performance of RBF. The reason for this is that RBF requires proper training and moreover its performance depends on its parameters. Based on literature survey as described in [14], major current challenges are 1) high false alarm rate and 2) high response time. Response time issue is tackle using parallel hierarchical Intrusion Detection System approach. To overcome high false alarm rate problem, it is necessary to properly choose RBF parameters. For this we need to optimize RBF parameters. Various optimization techniques are used but according to paper [5] QPSO has better results. So in this paper RBF parameters are optimized using QPSO (Quantum behaved particle swarm optimization) algorithm and parameter tuning is done in order to improve RBF performance for NIDS. Therefore by doing this we can achieve the objectives to have 1) less false alarm rate, 2) to improve detection rate, 3) to find which parameter optimization is important and 4) to deal with the symbolic data of kdd cup1999 [13]. The proposed algorithm approach is similar to paper [5] algorithm but there are some modifications done in it. This algorithm is described in implementation algorithm.

II. RADIAL BASIS FUNCTION

RBF was first introduced in the solution of the real multivariable interpolation problem by Broomhead & Lowe (1988) and Moody & Darken (1989) [9]. It is a kind of local approximation neural network, which has very strong approximation ability, good classification ability and rapid learning speed [7] [8]. RBF neural network is a feed forward network and it has simple structure [7] [8], as shown in Figure 1 [9]. It has three layers known as input layer (x), hidden layer which includes radial basis function, and output layer (y) which is linear summation of hidden layer [6] [8].

The neurons in the hidden layer commonly contain Gaussian transfer functions whose outputs are inversely proportional to the distance from the center of the neuron [3]. The equation of basis function can be defined as equation-1 [4].

$$\alpha_{i(x)} = exp\left[\frac{-\|x-c_i\|^2}{2\sigma_i^2}\right] \quad i = 1,2,\dots,m \qquad (1)$$

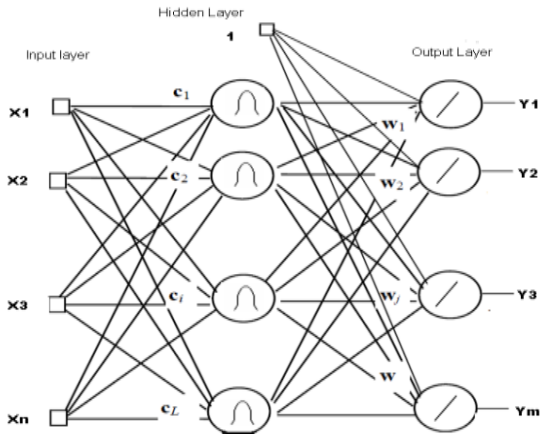

Figure 1 The structure of RBF [9]

Where αi(x) is the output of number i node of hidden layer, X = (x1, x2,.., xn)T are the input samples, Ci is the centroid, σi is the variable number which is known as width or spread parameter and m stands for total number of hidden layer nodes. The connection weight (w) between hidden layer and output layer is adjusted. And output layer realize on linear mapping [7] and can be given as defined in equation -2 [4].

$$Y_k = \sum_{i=1}^{m} \omega_{ik}\, \alpha_i(x) \quad k = 1,2,\ldots,p \qquad (2)$$

Where ωik is the weight value of network and p is the number of output layer nodes [4]. During training, the choice of the parameters has an important influence on the classification performance of RBF neural network. Following is the list of such parameters.

1. The number of neurons in the hidden layer.
2. The coordinates of the center of hidden-layer
3. The radius (spread) of each RBF function in each dimension.
4. The weights between hidden & output layer.

III. QUANTUM BEHAVED PARTICLE SWARM OPTIMIZATION

Kennedy and Eberhart developed a swarm intelligence optimization algorithm called Particle swarm optimization (PSO) in 1995 [5] [12]. Its basic idea is derived from social behavior of birds' predation. It will find optimal solution by utilizing cooperation and competition of the group particles. But as Wang and Mendel has demonstrated that PSO is not global convergent-guaranteed optimization algorithm. Therefore Sun J. et al [5] [11] proposed Quantum behaviour Particle Swarm Optimization (QPSO), whose performance is superior to PSO. It is a delta potential (quantum) model of PSO. In QPSO, state of particle is given by wave function $\psi^{(x,t)}$, instead of position and velocity. It will depict the probability of particle appearing in position x by using density function $|\psi(x,t)|^2$. The particle moves according to following iterative equation [5]:

$$x(t+1) = P \pm \beta * |mbest - x(t)| * ln\left(1/u\right) \qquad (3)$$

Where

$$mbest = \frac{1}{N} \sum_{i=1}^{N} pbest_i \qquad (4)$$

$$P = rand() * pbest_i + (1 - rand()) * gbest \qquad (5)$$

mbest (Mean Best Position) is defined as the mean value of all particles' best position; $pbest_i$ is the best individual position and gbest is the global best position for whole swarm. Here rand () and u are random numbers distributed uniformly on [0, 1] respectively. β is Contraction-Expansion Coefficient. Convergence speed of algorithm can be controlled by contraction expansion coefficient and it is the only parameter of QPSO algorithm [5] [11] [12].

IV. IMPLEMENTATION ALGORITHM

This proposed algorithm has similar approach as in paper [5] but some modifications are done in it, in. Before applying this algorithm, dataset i.e. standard benchmark kdd cup1999 [13] dataset is pre-processed. Pre-processing includes feature reduction (22 features are selected as mentioned in [16] [17] [18]) and normalization (includes symbolic to numeric conversion, scaling and lossless reduction [18]). This pre-processed dataset is given as input and then step by step following procedure is used to train RBF Neural Network.

1. RBF Neural Network is initialized. Now QPSO is used to optimize parameters during learning process, and parameter tuning is done.
2. Initialize QPSO. Here particle vector takes only one RBF parameter at a time. Means one time center is selected to optimize and similarly width and weight are selected. So QPSO initialize the population by randomly generating position vector X_i (i= 1,2,…,M). It is further adopted as $X_i = (c_1, c_{2,\ldots}, c_m)$ or $X_i = (\sigma_1, \sigma_2, \ldots, \sigma_m)$ or $X_i = (w_1, w_{2,\ldots}, w_m)$ and set $pbest_i = X_i$.
3. Structure a RBF neural network by treating the position vector of each particle as a mention in step 2. Also select value of β Contraction-Expansion Coefficient using equation (6).

$$\beta = \frac{(\beta_1 - \beta_0)*(T-t)}{T} + \beta_0 \quad (6)$$

4. Evaluate the fitness value of each particle by equation (7), update the personal best position P_i and obtain the global best position P_g across the population;

$$E = \frac{1}{2M}\sum_{p=1}^{M}\sum_{q=1}^{N}\left(d_q^p - y_q^p\right)^2 \quad (7).$$

Where d_q^p and y_q^p are the q^{th} desired target output and actual output of p^{th} input

5. If the stop criterion is met, go to step (7) or else go to step (6);
6. Update the position vector of each particle according to equation (3);
7. Output the P_g as a group of optimized parameters;
8. Save the training parameters. And thus RBF network is trained.
9. Now use testing set to validate the model.
10. Again go to step (2) and select another RBF parameter and repeat steps up to 9 until all parameters are tuned.

V. EXPERIMENT AND RESULT ANALYSIS

An experiment is conducted in order to validate the above mention algorithm to find which parameter optimization give better results for NIDS. This experiment adopt a standard KDD Cup 1999 dataset 1st developed by the Information Systems Technology Group (IST) of the MIT Lincoln Laboratory [13], to run our intrusion detection prototype. It consists of 4 different kinds of attack (DoS, PROBE, U2R, R2L). There are 22 types of attack in training dataset and 17 types of attack in testing dataset. It consists of 494021 records. Out of 41 features only 22 features are used in this paper. Then normalization of dataset (encoding, scaling and lossless reduction [18]) is done. We have implemented proposed algorithm using java programming language. Java program is also used for feature reduction and normalization. The performance of an Intrusion detection algorithm is usually measured by the detection rate (DR), false positive rate (FPR), true alarm rate (TAR), false alarm rate (FAR), true positive (TP), true negative (TN), false positive (FP), false negative (FN) and accuracy.

In this experiment, the number of hidden layer neurons is 7 as per [5]. Then one by one parameters of RBF are selected and optimized using QPSO algorithm. In order to implement QPSO algorithm, firstly its swarm size, maximum iteration, network parameter, its dimension, its range and value of β is decided. Here to implement proposed algorithm swarm size is set to 200, value of β linear varies according to equation (6) between 0.5 and 1.2 and max iteration is set to 1500 for all parameters. Moreover it will iterate certain times to calculate the fitness value function. The max iteration is also same for all parameters. The table 1 shows various performance measures results for all parameters which are optimized and tuned one by one

Table 1 Results of various performance measures

Optimizing parameters	**TP (%)**	**TN (%)**	**FP (%)**	**FN (%)**	**FAR**	**FPR (%)**	**TAR (%)**	**Detection Rate (DR) (%)**	**Accuracy (%)**
Only centroids	92.0202	1.9476	98.0622	7.9773	0.9805	19.1026	80.5179	92.0224	74.4701
Only width	99.8786	1.9954	98.0144	0.1189	0.9800	19.0933	80.8137	99.88100	74.2781
Only weight	98.2813	83.8463	16.1635	1.7162	0.1616	3.1486	96.1734	98.2837	95.4694
Weights and centroids	99.4793	98.3082	1.7016	0.5183	0.0169	0.3314	99.587	99.481	97.6474

Table 2 Results of detection rate and attack types

Parameters Optimized	**Normal (%)**	**DoS (%)**	**Probe (%)**	**R2L (%)**	**U2R (%)**
only centroids	80	95	30.36	10	5
only width	41.99	99.92	21.68	20.52	10
only weight	76.84	98.18	31.37	40.14	45.71
weights & centroids	98.31	95.42	85	95.26	85

In table 2 results of detection rate for normal and each attack type for various parameters are shown. Table 3 shows the comparison between paper [5] and proposed algorithm. It also shows that detection rate is more by optimizing centroids and weights instead of optimizing all three parameters as [5]. Moreover false rate is also less compare to paper [5]. Graphical representation of table 3 is shown in figure 2 which depicts the same scenario.

Table 3 Comparison table of proposed algorithm & paper [5]

Parameters Optimized	**Algorithms with their parameters**				
	Paper[5]	**Proposed algorithm**			
	All 3parameters	Only centroids	Only width	Only weights	Weights & centroids

Detection Rate	92.08	92.0224	99.8810	98.2837	99.481
FAR	0.0529	0.9805	0.9800	0.1616	0.0169

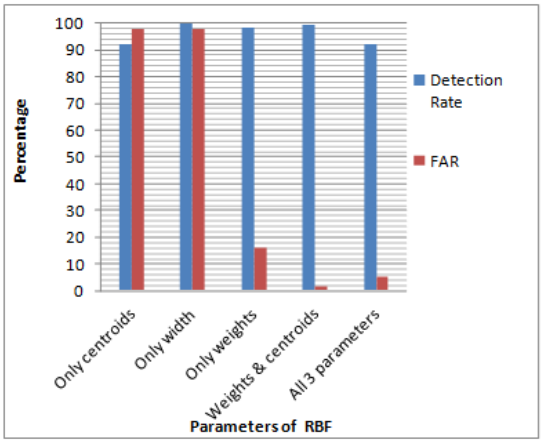

Figure 2 DR & FAR for various parameters optimized using QPSO

It shows that if we train only centroids and only width parameter of RBF then it is not enough to obtain proper performance. When we train only weight parameters at that time also it can detect attacks well but still false rate is high. So we train weight and centroids both at that time we get better accuracy and overall performance too. Thus by optimizing RBF parameters using QPSO algorithm performance of RBF for NIDS is improved. By using proposed algorithm false alarm rate is reduced and accuracy is also increased. But parameters performance depends on how it is trained, size of dataset, how many times it is train and on initial random values. Due to random values its performance differs every time. Moreover it requires more time in training of about 5-6 hours. In testing also execution time depends on dataset size.

A. Comparison Analysis

In table 4 comparisons between results of RBF-QPSO and convention RBF [15] is shown. It shows that RBF-QPSO has better results.

Table 4 Comprasion between rbf-qpso & conventional RBF

Attack types	RBF-QPSO (%)	Conventional RBF [15] (%)
Normal	98.30492	-
DoS	95.42472	85.14
Probe	85	81.71
R2L	95.2585	88.00
U2R	85	86.86

VI. CONCLUSIONS

For anomaly detection in Network Intrusion Detection System RBF-QPSO approach is used in this paper. RBF approach is used to handle issue such as regular updating, detection rate, false positive, false negative, suitability and flexibility for IDS. Moreover RBF performance depends on its parameters. To reduce false alarm rate RBF parameters need to be optimized and tuned. QPSO algorithm is used to optimize parameters. This paper concludes that optimizing centroids and weights using QPSO gives better performance than optimizing all 3 parameters. At the time of optimizing parameters centroids and weights, parameter width value is properly chosen randomly for better classification. Proposed algorithm works differently for various scenarios as initially values are randomly chosen and depends on how it is trained. Detection Rate and FAR are improved using optimization and tuning. But it requires more time in training. Moreover QPSO-RBF performance results are better than conventional RBF.

VII. FUTURE WORK

For various applications there can be different scenario in using RBF-QPSO. In future this can be extended for various applications. Moreover it is necessary to optimized RBF parameters. So in future, experiments can be carried out using various optimization techniques.

DECLARATION

The content of this paper is written by Author 1 (Henali Sheth) while Author 2 (Prof. Bhavin shah) had guided the work and Author 3 (Asst. Prof. Shruti yagnik) has reviewed this paper. Hence Author 1 is responsible for the content and issues related with plagiarism.

REFERENCES

[1] J. P. Anderson, Computer Security Threat Monitoring and Surveillance. Technical Report, Fort Washington, PA (1980).

[2] Yanwei, Fu, Zhu Yingying, and Yu Haiyang. "Study of neural network technologies in intrusion detection systems." Wireless Communications, Networking and Mobile Computing, 2009. WiCom'09. 5th International Conference on .IEEE, 2009, pp. 1-4.

[3] Yichun, Peng, Niu Yi, and Hu Qiwei. "Research on Intrusion Detection System Based on IRBF." In Computational Intelligence and Security (CIS), 2012 Eighth International Conference on. IEEE, 2012, pp. 544-548.

[4] Tian, Jingwen, Meijuan Gao, and Fan Zhang. "Network intrusion detection method based on radial basic function neural network." In E-Business and Information System Security, 2009. EBISS'09. International Conference on. IEEE, 2009, pp.1-4.

[5] Liu, Yuan. "Qpso-optimized rbf neural network for network anomaly detection."Journal of Information & Computational Science 8.9 (2011): pp. 1479-1485.

[6] Liu, Yinfeng. "An improved RBF neural network method for information security evaluation." TELKOMNIKA Indonesian Journal of Electrical Engineering 12.4 (2014): pp. 2936-2940.

[7] Chun-tao, Man, Wang Kun, and Zhang Li-yong. "A new training algorithm for RBF neural network based on PSO and simulation study." Computer Science and Information Engineering, 2009 WRI World Congress on. Vol. 4. IEEE, 2009, pp. 641-645.

[8] Bi, Jing, Kun Zhang, and Xiaojing Cheng. "Intrusion detection based on RBF neural network." In Information Engineering and Electronic Commerce, 2009. IEEC'09. International Symposium on. IEEE, 2009, pp. 357-360.

[9] Ugur Halici. Artificial Neural Network. Chapter 9. Radial basis function Network .EE543 lecture notes. Metu EEE. Ankara 139.

[10] Elfeshawy, Nawal A., and Osama S. Faragallah. "Divided two-part adaptive intrusion detection system." Wireless networks 19, no. 3 (2013): pp. 301-321 , DOI 10.1007/s11276-012-0467-7.

[11] Ma, Ruhui, et al. "Network anomaly detection using RBF neural network with hybrid QPSO."In Networking, Sensing and Control, 2008. ICNSC 2008. IEEE International Conference on. IEEE, 2008, pp. 1284-1287.

[12] Sun, Jun, Wenbo Xu, and Jing Liu. "Training RBF neural network via quantum-behaved particle swarm optimization."In Neural Information Processing. Springer Berlin Heidelberg, 2006, pp. 1156-1163.

[13] KDD dataset, 1999; http//kdd.ics.uci.edu/databases/-kddcup99/kddcup99.html

[14] Henali Sheth, Prof. Bhavin Shah, Shruti Yagni k et al. "A survey on RBF Neural Network for Intrusion Detection System " Int. Journal of Engineering Research and Applications, vol. 4, issue 12(part 4), December 2014, pp. 17-22

[15] Chen, Zhifeng, and Peide Qian. "Application of PSO-RBF neural network in network intrusion detection." Intelligent Information Technology Application, 2009. IITA 2009. Third International Symposium on. Vol. 1. IEEE, 2009, pp. 362-364.

[16] Tesfahun, Abebe, and D. Lalitha Bhaskari, "Intrusion Detection Using Random Forests Classifier with SMOTE and Feature Reduction." International Conference on Cloud & Ubiquitous Computing & Emerging Technologies (CUBE), IEEE, 2013.

[17] Shah, Bhavin, and Bhushan H. Trivedi. "Reducing Features of KDD CUP 1999 Dataset For Anomaly Detection Using Back Propagation Neural Network." In Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on, pp. 247-251. IEEE, 2015.

[18] Bhavin Shah, Bhushan Trivedi, "Dataset Normalization: For Anomaly Detection Using Back Propagation Neural Network"

[19] Shah, Bhavin, and Bhushan H. Trivedi. "Improving Performance of Mobile Agent Based Intrusion Detection System." In Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on, pp. 425-430. IEEE, 2015.

[20] Shah, Bhavin, and Bhushan H. Trivedi. "Artificial neural network based intrusion detection system: A survey." International Journal of Computer Applications 39, no. 6 (2012).

[21] Shah, Bhavin, and Bhushan H. Trivedi. "Optimizing Back Propagation Parameters for Anomaly Detection." *IEEE-International Conference on Research and Development Prospectus on Engineering and Technology (ICRDPET)*. 2013.

