International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:7, No:4, 2013

Research on Hybrid Neural Network in Intrusion Detection System

Jianhua Wang and Yan Yu

***Abstract*—**This paper presents an intrusion detection system of hybrid neural network model based on RBF and Elman. It is used for anomaly detection and misuse detection. This model has the memory function .It can detect discrete and related aggressive behavior effectively. RBF network is a real-time pattern classifier, and Elman network achieves the memory ability for former event. Based on the hybrid model intrusion detection system uses DARPA data set to do test evaluation. It uses ROC curve to display the test result intuitively. After the experiment it proves this hybrid model intrusion detection system can effectively improve the detection rate, and reduce the rate of false alarm and fail.

***Keywords*—**RBF, Elman, anomaly detection, misuse detection, hybrid neural network.

I. Introduction

ALONG with the computer network, the network security becomes a problem. In order to prevent network intrusion, intrusion detection technology becomes hot spot problems for people research. Intrusion detection test system diagram monitors and prevents possible intrusion or other harm behaviors on your system and network resources as far as possible.

Building neural network model through the authorized user's behavior characteristics so that it can be used to monitor the invasion behavior. At present, intrusion detection system research based on the neural network has made many achievements. Its application in intrusion detection system can improve the efficiency of the intrusion detection system, and enhance the system self learning ability [1, 2, 3, 4, 5, 6]. But there are also many problems which need to be solved.

At present the most studied, and the most widely applied network is a kind of multilayer feed forward neural network, but the MLP network model does not have memory function of previous events, and MLP network needs long training time. It is the nonlinear mapping global approximation [5]. In order to overcome the above problems, this paper presents a hybrid neural network structure based on generalized RBF network and Elman network. Among them, the Elman network is used to memorize the previous events. RBF network uses local index attenuation nonlinear function to do the local approximation work for nonlinear input/output. RBF network has a fast convergence speed of learning [7, 8, 9] compared with MLP network.

This paper puts forward the intrusion detection model is the combination of misuse monitoring and anomaly detection model. It uses DARPA data set to do test evaluation for system monitoring results. After the experiment it proves this model can effectively improve the detection rate, and reduce the rate of false alarm and fail.

Jianhua Wang, is with the Institute of Computer Science and Information Engineering, Harbin Normal University, China (phone: +86-451-88060537; fax: +86-451-88060602; e-mail: wjh@163.com).

Yan Yu, was with the Institute of Computer Science and Information Engineering, Harbin Normal University, China (corresponding author to provide phone:+86-451-88060537; fax:+86-451-88060602; e-mail: yuyan9999@vip.qq.com).

II. The Application of the Neural Network in Intrusion Detection

A. The Determination of Input Vector

Neural network has a wide range of applications in intrusion detection. Compared with the previous proposed "the user behavior", this article distinguishes between normal and abnormal behavior through the establishment of "software behavior" feature. In this paper the "software behavior" refers to the system call sequences.

In order to use neural network to do intrusion detection work, it is a must to solve five main questions: how to determine the neural network's input data? What network technology will be used? How to train the network? How to say intrusion behavior? How to deal with the neural network output data.

In order to solve the above-mentioned problems, first of all we need to understand BSM audit data format. A BSM record in audit file includes seven fields at most. The seven fields information can be used to distinguish between normal behavior and abnormal behavior [10]. Only these information will be converted to vector so that it can be accepted by the neural network [11,12,13]. Conversion steps are as follows:

a. acquisition of system call sequences.
b. use length L sliding window to intercept system call sequences.
c. vectorization of captured system call sequences.
d. get N "model series" after the vector quantization.
e. other vector projects to the N dimensional space, namely the N "model series" vector distance respectively, and get a group of N vector as the neural network input vector.

B. General RBF Network

RBF network belongs to the multilayer feed forward neural network. It is a kind of three layer forward network. The input layer is composed by the source node. The second is hidden layer. Hidden layer cell transformation function is center radial

symmetrical and attenuated nonnegative nonlinear functions; the third layer is output layer. It makes response to input mode [7, 8].

RBF is based on the theory of Cover theorem of model separability. The complicated pattern classification problem projects into a high dimensional space nonlinearly, and it is more likely to be linear separable to project to low dimensional space. The basic idea of the RBF neural network is: radial basis function (RBF) is used as the hidden layer unit "base", a hidden space. Hidden layer transforms input vector. The low dimensional model input data changes to high dimension space. It makes the linear inseparable problem in low dimensional space become linear separable within high dimensional space. RBF network structure, as shown in Fig. 1.

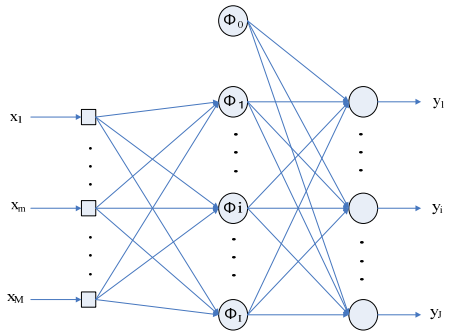

Fig. 1 RBF neural network model

From the chart that the input layer has M neurons, hidden layer has I neuron, among which the i neuron incentive output is as the base function, including $t_i = [t_{i1}, t_{i2}, ...t_{iM}]$ (i= 1, 2,…. I) are as the center of the basis function; Output layer has J neurons. Weights between Hidden layer and output layer is shown by w_{ij}. In the general RBF network ,a hidden layer neuron G_0 output constant is as "1", and the output unit connected weights is(j = 1, 2…. J).

Assume that the training sample collection is $X = [X_1, X_2, ...X_k, ..., X_N]^T$, including any training sample $X_k = [x_{k1}, x_{k2}, ...x_{km}, ...x_{km}]^T$ (k = 1, 2,…, N) corresponding actual output is $Y_k = [y_{k1}, y_{k2}, ...y_{kj}, ...y_{kJ}]^T$ (k = 1, 2,…, N), expected output is $d_k = [d_{k1}, d_{k2}, ...d_{kj}, ...d_{kJ}]^T$ (k = 1, 2,…, N).

When the network inputs training samples X_k, the j output neurons in RBF neural network actual output is

$$y_{kj}(X_k) = w_{0j} + \sum_{i=1}^{I} w_{ij}\phi(X_k, t_i),\ j = 1, 2, ..., J \tag{1}$$

In this paper the selected base function is Gaussian function, so can be said as follows

$$\phi(X_k, t_i) = G(\| X_k - t_i \|) = \exp\left(-\frac{1}{2\sigma_i^2}\| X_k - t_i \|\right) = \exp\left(-\frac{1}{2\sigma_i^2}\sum_{m=1}^{M}(x_{km} - t_{im})^2\right) \tag{2}$$

$t_i = [t_{i1}, t_{i2}, ...t_{iM}]$ which is the center of the Gaussian function, σ_i is the variance of Gaussian function.

The RBF neural network has three parameters to learn: the center of the basis function, variance and weights. Based on radial basis function center to select different method. The RBF network has a variety of learning methods. The most commonly used four kinds of study method are: random selection center method, self organization selection center method, supervision selection center method and the orthogonal least squares method. This paper uses self-organizing selection center method. This method has two- phase compositions: one is the self organization learning stages, namely, study of the center of hidden layer basis function and variance. In this stage, this paper uses the enhancement k-means clustering algorithm. This method is based on the variation of the weighted measure clustering. It can make the algorithm converge in an optimal results or approximately optimal results, and it has nothing to do with the initial position of the center. The second is supervised learning stages, namely, study the output layer weights stage. In this stage, the study of the right value can use LMS method.

C. Elman Network

According to the front narration, intrusion detection system detects discrete and related attack. It must have memory ability to recent events, Elman network can realize the memory ability.

Elman network is a famous circulation network technology [9]. It is put forward by Jeffrey Elman. A Elman network has a group of context related nodes. Every context related node get an input from a hidden layer nodes. The context related node sends its output to the hidden layer relative current node. Because these contexts related nodes only rely on the front input hidden layer nodes to activate, so context related nodes reserve two input state information [7]. This paper Elman neural network is shown as in fig. 2. Elman network uses BP algorithm for training. Elman network activation Function is Sigmoid Function.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

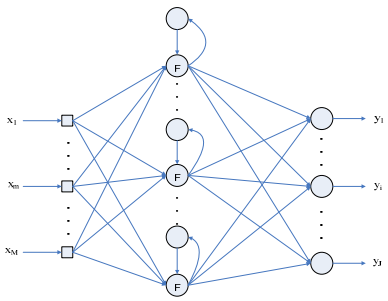

Fig. 2 Elman neural network model

D.RBF/Elman Blend Neural Network

Pure RBF network technology character has a big limitation. This feature is any input resulted input is not relevant from the front output and this feature may be suitable for those who need independent process information as input task. But it is not the best choice for continuous flow of data as input task. For example, detect DOS attack and port scanning can't depend on the current events separately .It also needs the original event information as a basis for judgment of invasion. Therefore, we put forward a kind of hybrid RBF/Elman neural network model.

1. The Structure of the Hybrid Model

RBF/Elman hybrid neural network model is as shown in fig. 3. This hybrid model is structured by a RBF neural network and several Elman neural network. RBF network realizes function of the input data real-time analysis classification. Each Elman network realizes recent some attacks memory function. In this model, a neuron output of the RBF network output layer is as a Elman network's input, so the number of Elman network equals RBF network output neuron number, and one to one correspondence. An RBF network output neuron represents unknown abnormal behavior. Every other output unit represents one kind of known attack type. Each output of RBF network can be memorized by an Elman network. When the RBF network classifies input, output classification results spread forward and memorized by the output unit connecting Elman network.

RBF/Elman hybrid neural network model, it adapts to the needs of the following three conditions:

(1) Some attack cannot be tested according to the current isolated input information. Testing this kind of attack not only needs the input information but also needs recent network input as the basis for judgment. DOS attack detection and port scanning can't depend on the current events separately. It also needs recent event information as a basis for judgment of invasion.
(2) In most cases invasion has cluster character, i.e., continuous abnormal behavior appears in a period.

The differences between comprehensive Limited system knowledge base, software running environment and version lead to a normal behavior .It may be mistaken for abnormal behavior, and this kind of situation happens occasionally.

The hybrid neural network model has memory ability. It can detect cooperated attacks. Hybrid model uses intrusion events to show cluster. It is sensitive to continuously happened abnormal behavior, and oversights occasional individual abnormal behavior in normal behavior sequence. This not only improves the generalization ability of the system, improves the efficiency of system detection rates and reduces the system error rate.

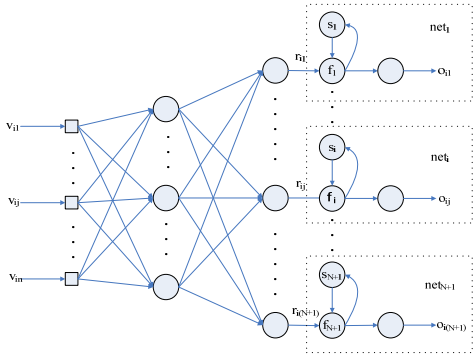

Fig. 3 The hybrid RBF/Elman neural network model

2. Hybrid Model Work Steps

The hybrid RBF/Elman neural network work step includes the following steps:

(1) Read each connection weight matrix, and generate network according to the defined hybrid neural network topology.
(2) The context related neurons value of each Elman network undertaking layer in the hybrid neural network t initializes to zero, namely $S^t = (s_1^t, s_2^t, ...s_i^t ....s_{N+1}^t)^T$ each element is zero. Each neuron activation function of the system's each Elman network undertaking layer is discrete function.
(3) In time, hybrid neural network reads an input vector. The input vector passes through the calculation of RBF neural network to get RBF network output vector. This vector passes through the system default amplifier .Signal processes to get vector $C^{(t)} = (c_1^t, c_2^t, ..., c_i^t, ..., c_{N+1}^t)^T$,i.e. reset element which is lower than the value, reserve the element which is higher than the value. Amplifier is used to make the Elman network memorize the abnormal behavior better.
(4) Each element of the processed vector of RBF neural network output vector connects forward to the Elman network. The i Elman hidden layer neurons is commonly activated by s_i and c_i^t , if $c_i^t = 0$ so $s_i^{t+1} = s_i^t - \Delta s$, otherwise $s_i^{t+1} = s_i^t + \Delta s^{'}$.
(5) mCalculate the Elman network output vector. Compare the vector with alarm vector. if it is above the corresponding element of alarm vector, the attack type occurs.
(6) In t + 1 hour, repeat the process.

From the above process, it is known that the Elman network context related neurons value becomes smaller gradually over time. When abnormal behavior happens we should increase the value to realize the memory of recent events. When RBF neural network detects an abnormal behavior, the output node connected Elman network increases a larger quantity for the value in context related neurons. Set an alarm value vector. When an Elman network output value is more than the corresponding elements of the value vector, it means one of this types of intrusion events happens.

Hybrid Model Algorithm Description

Neural network input set after pretreatment is $V=\{v_1,v_2...v_i...v_k\}$. Take any one of the input vector $v_i=(v_{i1},v_{i2},...v_{ij}...v_{in})^T$, the "n" is RBF network input layer node number. Vi passes treatment of RBF neural network to get output vector $R_i=(r_{i1},r_{i2},...r_{ij}...r_{i(N+1)})^T$, including "N + 1" is RBF network output layer node number. Model "N + 1" Elman network expresses as vector $Net=(net_1,net_2,...,net_i,...,net_{N+1})^T$. Weight matrix connected RBF network and Elman network is $W=(w_1,w_2,...w_i...w_{N+1})^T$. The connection weight matrix between Elman network undertaking layer and the hidden layer is $W^{'}=(w_1^{'},w_2^{'},...w_i^{'}...w_{N+1}^{'})^T$. Context relevant neurons value in the Elman network undertaking layer is $S=(s_1,s_2,...s_i....s_{N+1})^T$. Filter filtration value is $P=(p_1,p_2,...,p_i,...,p_{N+1})^T$. Every kind of attack alarm value is $T=(t_1,t_2,...,t_i,...,t_{N+1})^T$. Through the above data, we can get Elman network output vector $O_i=(o_{i1},o_{i2},...o_{ij}...o_{i(N+1)})^T$. RBF/Elman hybrid neural network uses the following formula to compute output vector:.

$$o_{ij}=r_{ij}*w_j+s_i*w_j^{'}$$

Input: The BSM audit data
Output: Intrusion alert
Algorithm:
process the BSM audit data to retrieve set V, where each item is a vector
build a hybrid network, whose outputs number is N+1, one used for anomaly detection and the others used for misuse detections.

```
for Vi∈V do
  RBF process Vi to get a result Ri
   for vij∈Vi do
      if rij<= pj then sj -=Δs
      else then sj+= rij *w_j
    end if
    o_ij = r_ij * w_j + s_i * w_j'
    set net1 used for anomaly detection
    set net2~netN+1 used for misuse detection
    for oij∈Oi do
       if oij >tj
         if j==1 then output an anomaly alert
         else output a misuse alert, type is which "j" mapped
         end if
       end if
    end for
  end for
end for
```

III. The Experimental Results

Anomaly and misuse detection system use the same data set to detect. The first cycle and the third cycle training data of data set are used to train hybrid network. In order to facilitate the determination of the accuracy of system output results, we use the second cycle training data to test the detecting ability of system. The second cycle data has 43 attack sessions.

By changing the value of output node can adjust the sensitivity of the system. If cycle weights value is 1, then all anomalies are ignored before this anomaly. If the cycle weights value is zero, then all the previous anomaly emerged in the process are considered cumulatively. In the test process, the cycle weights value defines between 0 and 1.

It needs to analyze the relationship between system detection rate and false positives to decide the intrusion detection system detection capacity. The ROC curve is used in this paper to analyze the system's detection ability. Different circulation weight value corresponds to the different ROC curve. After determining the cycle weights, each value of the output node can be sure of one group of detection rate and false positives. The group detection rate and false positives can be mapped to a point of ROC curve. System is in anomaly detection, and two different cycle weights get two different ROC curve, which is shown as in fig. 4. When circulation weight value is 0.75, the best operating point of system, the detection rate was 91.4%, and the misinformation rate was 3.1%. When circulation weight value is 0.25, the best operating point in the system, the detection rate is 93%, and the misinformation rate is 2.3%.

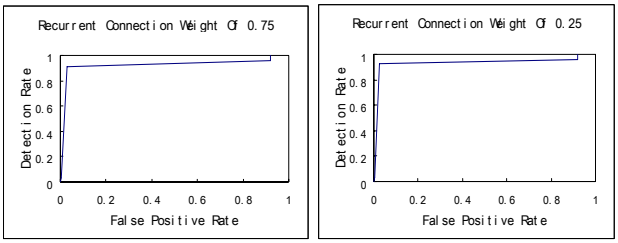

Fig. 4 Anomaly detection results when circular weights are 0.75 and 0.25, respectively

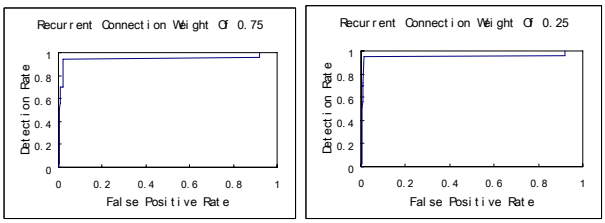

Fig. 5 Misuse detection results when circular weights are 0.75 and 0.25, respectively

TABLE I
THREE DIFFERENT NETWORK MODEL DETECTION RATE AND FALSE POSITIVES

Network Model	Anomaly Detection		Misuse Detection	
	Detection Rate	Misstatement Rate	Detection Rate	Misstatement Rate
MLP / leaky bucket	77.3%	2.2%	90.9%	18.8%
Elman/ leaky bucket	77.3%	0%	/	/
RBF/Elman	93%	2.3%	95.3%	1.4%

System is in the misuse detection, two different cycle weights get two different ROC curves, as shown in Fig. 5. When circulation weight value is 0.75, the best operating point in the system, the detection rate is 94.1%,and the misinformation rate is 2.3%. When circulation weight value is 0.25, the best operating point in the system, the detection rate is 95.7%, the misinformation rate is 1.4%.

Through the above experimental data, it can compare with network model presented in [3], as is shown in table I. Analysis shows that the proposed hybrid model anomaly detection is stronger than misuse detection ability. The reason is the shortage of the data set for training neural network misuse attack behavior. If use enough abnormal behaviors to train neural network, this system‘s misuse detection ability will be improved significantly.

IV. CONCLUSION

In this paper, a hybrid intrusion detection system model based on RBF/Elman network. This model has the following three prominent features:

(1) The model achieves a classifier which has memory ability of recent events .The RBF network realizes the function of real-time classifier, and Elman network realizes memory function.
(2) The model is a pure neural network model, and complete intrusion detection without any other algorithm auxiliary.
(3) The sensitivity of the system is easy to be configured, the end user can define the tolerance of the system, and neural network is transparent to it.

This system model has memory ability and adaptive ability. It can effectively enhance the detection rates of intrusion detection system and reduce false positives and fail through the experiment.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (No. 41071262), and supported by grants from Intellectual Education and Information Engineering in Heilongjiang Key Laboratories (081203) and Computer Applications Technology in Heilongjiang Key Subjects.

This work was supported by Youth Academic Fund in Harbin Normal University: Key technologies of intelligent tutoring system based on multi- AGENT (No. 11XQXG23).

REFERENCES

[1] Bivens A, Palagiri C, Smith R, Szymanski B. et al. Network-based Intrusion Detection using Neural Networks. Proceeding of ANNIE-2002, New York, ASME Press, 2002. 579-584.
[2] Yang Ke, Wang Li-Ping, Fang Ding-Yi. Program behavior anomaly detection based on neural network. Dalian Ligong Daxue Xuebao/Journal of Dalian University of Technology, v 45, n SUPPL., October, 2005, p S136-S141.
[3] Azadi Avenue, Tehran, Iran. RT-UNNID: A practical solution to real-time network-based intrusion detection using unsupervised neural networks. Computers & Security, Volume 25, Issue 6, September 2006, Pages 459-468.
[4] Guisong Liu, Zhang Yi and Shangming Yang. A hierarchical intrusion detection model based on the PCA neural networks Neurocomputing, Volume 70, Issues 7-9, March 2007, Pages 1561-1568.
[5] WeiShengJun, HuChangZhen, JiangFei. intrusion detection method [J] based on BP neural network improved algorithm. Computer engineering and application, 2005, (7) : 154-158.
[6] A. K. Ghosh, A. Schwartzbard. A study in using neural networks for anomaly and misuse detection [A]. In Proceedings of 8th USENIX Security Symposium [C], San Washington: USENIX Association, 1999, 23-36.
[7] Adrian G. Bors. Introduction of the Radial Basis Function(RBF) Networks. University of York UK. : Rbf.pdf.
[8] LuTao, ChenDeZhao. Radial basis network research progress and review [J]. Computer engineering and application, 2005, (4) : - 62.
[9] Elman, J.L. Finding structure in time. Cognitive Science, 1990, 14(2): 179-211.
[10] Sun Microsystems. Sun SHIELD Basic Security Module Guide.[BE/OL].
[11] S. A. Hofmeyr, S. Forrest, A. Somayaji. Intrusion detection usingsequences of system calls[J], Journal of Computer Security, 1998,(3) 151-180.
[12] Cunningham R K, Lippmann R P, Fried D. J, et al. Evaluating Intrusion Detection Systems without Attacking Your Friends: The 1998 DARPA Intrusion Detection Evaluation. Proceedings of Third Conference and Workshop on Intrusion Detection and Response. San Diego: CA, 1999.10-21.
[13] Lippmann R, Haines J W, Fried D J, et al. The 1999 DARPA Off-Line Intrusion Detection Evaluation. Computer Networks, 2000,30(2). 14-26.