Acta Technica *62*, No. 3B/2017, 17–26

Study on P2P network virus inspection based on wavelet neutral network

Zhao Nannan[1], Chen Jinjian[1]

Abstract. In order to improve the inspecting accuracy of P2P network virus, the application of wavelet neutral network on it is studied in depth. Firstly, the basic theory of wavelet neutral network is studied. Secondly, the training algorithm of wavelet neutral network is analyzed. Thirdly, the P2P network virus inspection model is constructed. Then P2P network virus inspecting algorithm procedure based on wavelet neutral network is designed. Finally, simulation is carried out, and results show that the wavelet neutral network can improve the inspection accuracy of P2P network virus effectively.

Key words. P2P network virus, inspection algorithm, wavelet neutral network.

1. Introduction

With development of network virus, the computer resources have suffered huge losses and destruction. P2P environment offers convenient resource sharing and convenient communication, the network virus gets the good intrusion opportunity at same time, therefore the anti virus task of P2P network is even more difficult. P2P (Peer-to-peer computing) technology is a main and wide application in file sharing network and instant message communication network. Because of logic adjacent nodes in P2P network geographical position has far adjoining distance, however the number of nodes enters P2P network is every big, therefore the unwanted code spread through P2P system has big range, and wide coverage, then the losses caused is very big. Taking virus as example, anti virus ability of every node in P2P network is different, if a node is infected by virus, and the virus is spread to nearby close node through internal sharing and communicating mechanism. In short time the congestion even paralysis of network can be caused, the sharing information will loss, and the confidential information is stolen, even the whole network can be controlled by network virus. Detecting the P2P network virus is very important, it is necessary to choose an effective algorithm, and then the inspecting efficiency can be improved. The inspection of P2P network virus can be disturbed by many factors because of

[1]Guangdong Ocean University Cunjin College, Zhanjiang, 524094, China; e-mail: znn_923@163.com

every reasons, and while the wavelet transform is a new signal processing method, which has good time-frequency local characteristic, therefore it can be applied in the inspection of P2P network virus [1].

2. Basic theory of wavelet neutral network

Wavelet neutral network belongs to a kind of feed-forward type neutral network, which has connection structure. The basic idea of wavelet neutral network is using wavelet function as basic function, and the expansion and translation coefficients of wavelet corresponds to weight coefficients from input layer to hidden layer and threshold of hidden layer. Basic structure of wavelet neutral network is shown in Fig. 1 [2].

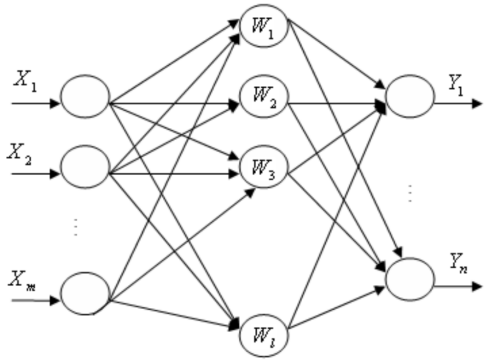

Fig. 1. Theory diagram of wavelet neutral network

Mathematical formulation of wavelet neutral network is expressed as

$$Y_i = f_i(X) = \sum_{j=1}^{l} \omega_{ij} \Phi_j \left(\sum_{k=1}^{m} b_{jk} x_k - r_j \right), \tag{1}$$

where X denotes the inputting vector of wavelet neutral network, which can be expressed as $X = [X_1, X_2, X_3, \cdots, X_m]$, Y denotes the outputting vector of wavelet neutral network, that can be defined as $Y = [Y_1, Y_2, Y_3, \cdots, Y_n]$, ω_{ij} denotes the weight value from node i to node j of wavelet neutral network, Φ_j denotes the wavelet function corresponding to node j of hidden layer of wavelet neutral network, b_{jk} stands for the expansion coefficient of wavelet basic function of wavelet neutral network and r_j iss the translation coefficient of wavelet basic function of wavelet neutral network.

Symbols l, m and n denote the number of nodes in hidden layer, number of nodes in inputting layer, and number of nodes in outputting layer, respectively, of Mexico hat wavelet neutral network. The wavelet function is used as the reward function of

wavelet neutral element of node, the corresponding expression being listed as [3]

$$\Phi(x) = (1 - x^2)e^{-x^2/2}\,. \tag{2}$$

2.1. Training algorithm of wavelet neutral network

In order to improve the reliability of inspecting P2P network virus and reduce the response time, the online training of parameters of wavelet neutral network can be carried out effectively. The gradient descent method and recurrence least square method are applied in the online training of wavelet neutral network, and the objective function can be expressed as

$$H(\theta) = \frac{1}{2}(Y(\theta) - X(\theta))^2 = \frac{1}{2}(E(\theta))^2\,, \tag{3}$$

where $E(\theta)$ denotes the outputting error.

2.1.1. Training algorithm of wavelet neutral network. Because the output of wavelet neutral network and weight value can be denoted as linear relation, the solution can be based on least square method. The recurrence least square method is used to train the weight coefficients of wavelet neutral network, then the inspecting correctness of P2P network virus can be ensured. In addition, falling into the local optimal solution can be avoided, the recurrence least square method can be expressed as [4]

$$\omega(\theta) = \omega(\theta - 1) + (1 + X^{\mathrm{T}}(\theta)P(\theta - 1))^{-1}P(\theta - 1)X(\theta)e(\theta)\,, \tag{4}$$

where

$$P(\theta - 1) = (X(\theta - 1)X^{\mathrm{T}}(\theta - 1))^{-1}$$

and

$$X(\theta) = [\Phi_1(\theta), \Phi_2(\theta), \Phi_3(\theta), \cdots, \Phi_n(\theta)]\,,$$

$X(\theta)$ denoting the outputting variable of hidden layer.

2.1.2. Training algorithm of translation coefficient and expansion coefficient. During the training procession of translation factor b_{jk} and expansion factor r_j of wavelet neutral network, the momentum term can be introduced, and the weight value can change smoothly, the convergence time of wavelet neutral network can be reduced, and the dynamical characteristics of wavelet neutral network can be improved. The gradient descent method can be used to train b_{jk} and r_j, then falling into local minimum value can be avoided and the corresponding training algorithm can be described as follows:

$$\Delta b_{jk}(\theta) = -\lambda\frac{\partial H(\theta)}{\partial m_{jk}} + \beta\Delta b_{jk}(\theta - 1)\,, \tag{5}$$

$$\Delta r_j(\theta) = -\lambda\frac{\partial H(\theta)}{\partial r_j} + \beta\Delta r_j(\theta - 1)\,, \tag{6}$$

where λ denotes the learning rate of wavelet neutral network and β denotes the momentum factor.

3. P2P network virus inspection model

3.1. Packet capture of network card

Packet capture module mainly assigns packet capture software Winpcap, all data packet fro inspecting captured through local network card based on interface provided by it. This inspection system only capture possible data packet of P2P flow. The P2P network virus inspection mainly considers the flow of P2P, and the inspector diagram of data is shown in Fig. 2. Figure 2 shows the whole frame of P2P network virus inspector system [5].

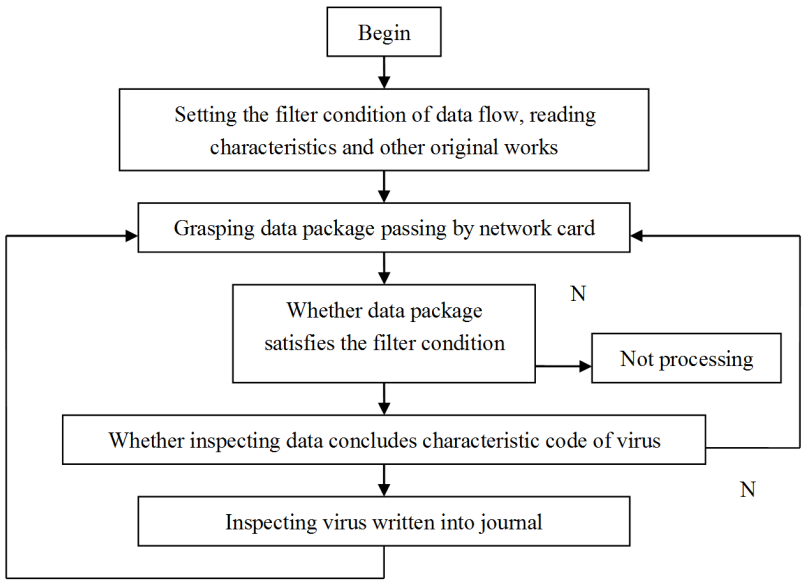

Fig. 2. Flow chart of P2P network virus

3.2. Preprocessing of data packet

Data from network card is the unprocessed data packet information, and can not inspect the virus directly, and therefore the preprocessing of data packet captured is necessary. This is a complex and important part, it is critical to get the data of application layer. The preprocessing progress concludes head of data, IP datagram fragmentation, protocol packet analysis of transmission layer, and data analysis of application layer, the data flow of P2P is separated based on protocol field, and then the unnecessary data processing can be reduced, and the corresponding algorithm

procedure is shown in Fig. 3.

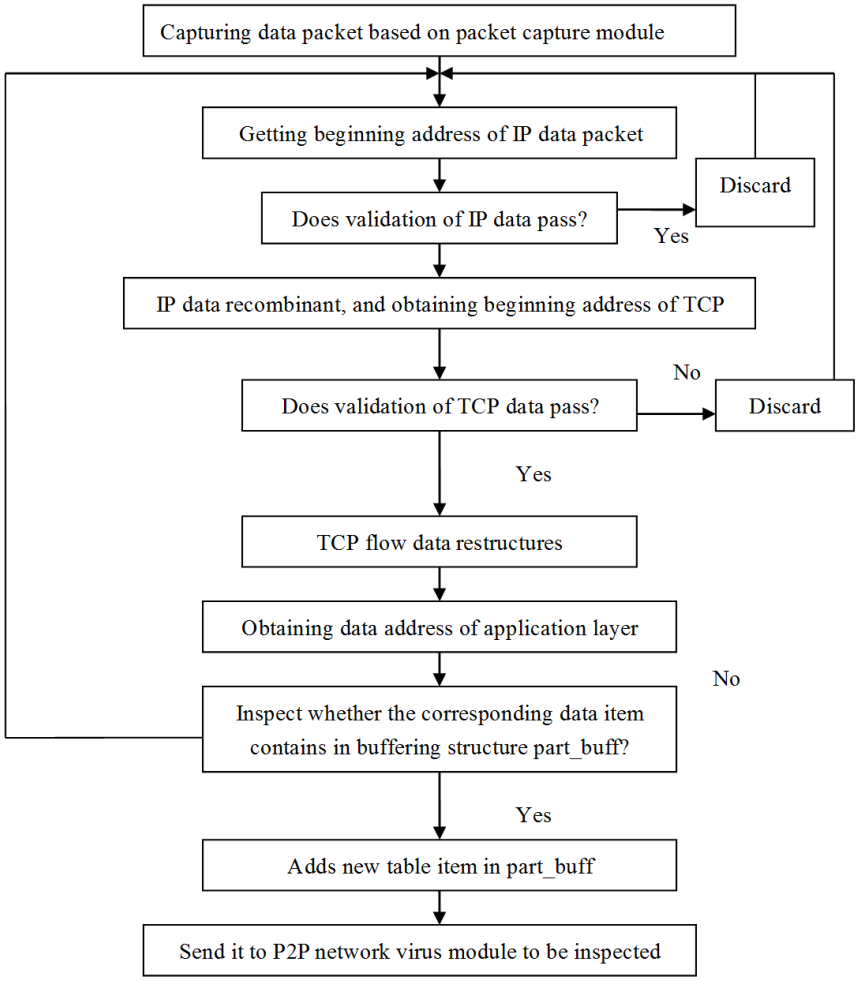

Fig. 3. Data packet preprocessing module

3.3. Inspection of P2P network virus

Data information of application layer can be obtained based on procession mentioned above, because block mechanism and subcontracting mechanism of P2P software file, current data information only inspect coincidence of part information and a part of certain characteristic code in virus database [6]. Then system can consider that it belongs to a part of virus characteristic coded divided, then this part of data can be buffer stored. These data combines with later data information can enter into inspecting part to be inspected, the flow chart of P2P network is shown in Fig. 4.

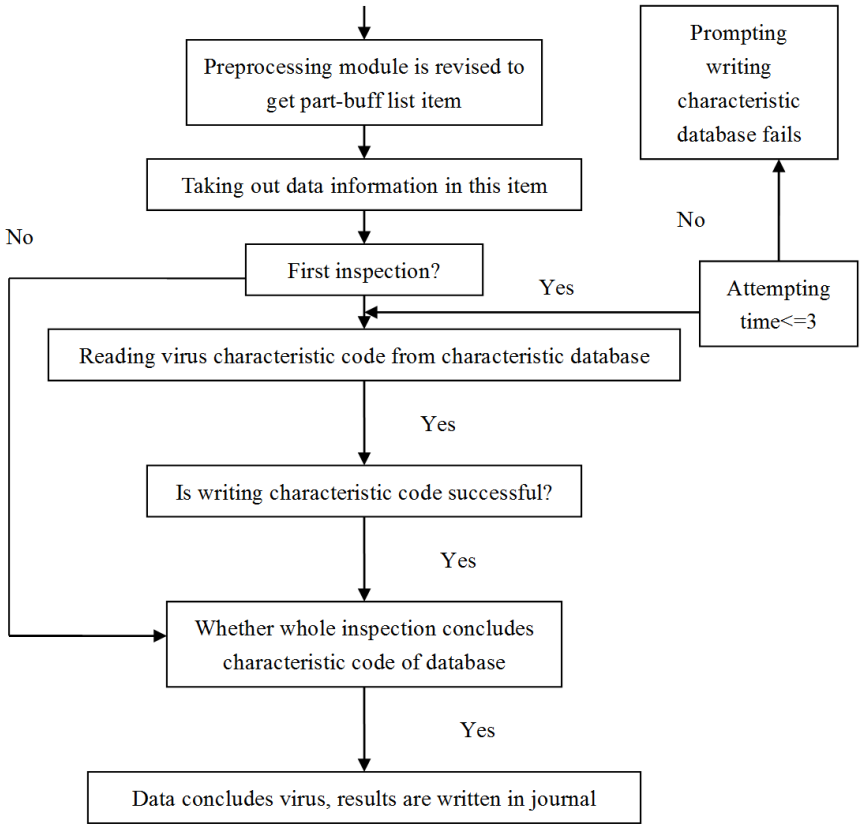

Fig. 4. Processing module of P2P network inspecting module

4. P2P network virus inspecting algorithm procedure based on wavelet neutral network

The algorithm procedure of P2P network virus inspecting is listed as follows [7]:

Step 1: The P2P network virus data is written into inspecting system.

Step 2: The normalization preprocessing is carried out for P2P network virus data.

Step 3: Choosing 150 P2P network virus data from network virus data processed, and obtaining training and testing samples according to all kinds of experimental plans.

Step 4: The wavelet neutral network is trained based on training results from Step 3, and results are stored in the wavelet neutral network.

Step 5: The results obtained from step 3 are input into wavelet neutral network to carry out test, based on output of outputting layer nodes to predict the P2P network virus.

5. Case study

The experiment can analyze the spreading procession of simulation virus in P2P network, the experiment is carried out in local network. The experiments include four general PC machine and a router and several meters network line. Four PC machines are connected by router to construct a local network, and PC machine can communicate with each other. The operating system is Windows XP, capture packet and issue packet software use Wincap software, Wincap is desined based on Win32 platform, which can capture open source library provided by network data packet.

The P2P network virus inspection experiment is carried out through setting different block length, and in this experiment total number of characteristic code is set as 10 sections mentioned above.

The training is carried out for wavelet neutral network firstly, and then the nonlinear mapping relationship between output and input of wavelet neutral network can be established. Ten neutral elements are set in the hidden layer of wavelet neutral network.

The gradient descent method and recurrence least square method are trained for wavelet neutral network, and the training algorithm is compiled by MATLAB software. The parameter setting of wavelet neutral network is listed as follows: $\lambda = 0.02$, $\eta = 0.30$, permitted error of algorithm is taken as 0.001. The wavelet neutral network is trained based on 150 P2P network viruses, and the iteration curve is shown in Fig. 5. As seen from Fig. 5, wavelet neutral network algorithm converges to the iteration error predefined, and the results show that the wavelet neutral network has quick convergence speed.

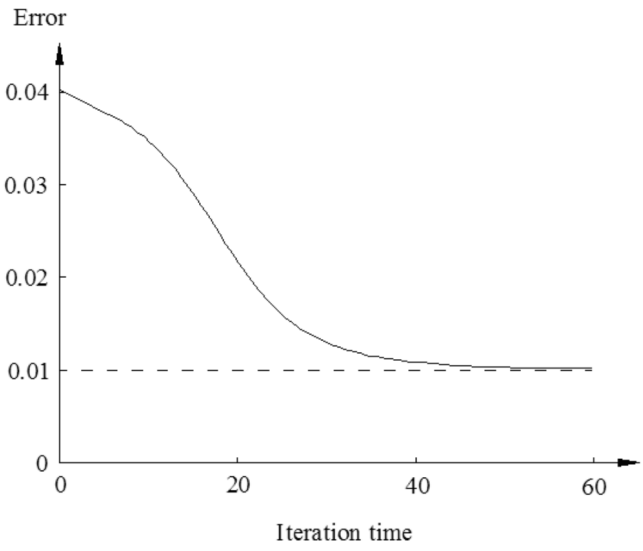

Fig. 5. Iteration simulation curve of wavelet neutral network training

In order to verify the effectiveness of wavelet neutral network on the P2P network virus inspection, 10 P2P network virus data is used as simulation data, the corresponding verifying test is carried out, and results are shown in Table 1.

Table 1. The distribution of the dynamic comfort

Block size of file/Byte	Total number of file block	Number of inspecting characteristic code	Whether is virus or not
8	44056	0	Yes
8	22528	1	Yes
16	11264	2	No
16	5632	4	Yes
32	2816	7	Yes
32	1408	9	No
64	704	10	No
64	352	21	No
128	176	32	Yes
128	88	46	Yes

As seen from Table 1, the P2P network virus inspection results agree with real testing results, and the correctness reaches 99 %. Therefore, the wavelet neutral network can improve the inspecting correctness of P2P network virus, and can reach good inspection precision.

6. Conclusion

P2P network is facing the threat of virus, and the virus can spread through applying generating virus in sharing file and normal upload activity of customer. Therefore it is necessary to establish a effective algorithm to inspect the P2P network virus, the wavelet neutral network can inspect the P2P network virus effectively based on simulation results, which can improve the inspecting correctness of P2P network virus.

References

[1] L. Z. Zhu: *Based on the BP neural network and ant colony algorithm virus detection.* Computer Programming Skills & Maintenance (2014), No. 18, 95–97.

[2] H. D. Ping, B. Di: *A modified wavelet neutral network model for measuring goodwill.* International Journal of Hybrid Information Technology *7* (2014), No. 4, 201–212.

[3] X. T. Deng, R. Yuan, Z. Xiao, T. Li, K. Wang, L. Li: *Fault location in loop dis-*

tribution network using SVM technology. International Journal of Electrical Power & Energy Systems *65* (2015), 254–261.

[4] A. Tjahyanto, D. P. Wulandari, Y. K. Suprapto, M. H. Purnomo: *Gamelan instrument sound recognition using spectral and facial features of the first harmonic frequency.* Acoustical Science and Technology *36* (2015), No. 1, 12–23.

[5] F. M. Rasheed: *Modelling virus propagation in P2P networks.* International Journal of Computer Science Issues *9* (2012), No. 2, 580–587.

[6] W. Tarng, C. K. Chou, K. L. Ou: *A P2P botnet virus detection system based on data-mining algorithms.* International Journal of Computer Science & Information Technology *4* (2012), No. 5, 51–65.

[7] X. G. Qiu, B. H. Wang, S. F. Gong: *Study of P2P network model based on distributed agent memory mechanism.* Journal of Computer Applications *30* (2010), No. 6, 1483–1485.

Received July 30, 2017