International Journal of Hybrid Information Technology
Vol.8, No.7 (2015), pp.435-442
http://dx.doi.org/10.14257/ijhit.2015.8.7.39

Study on Active Defense Technology and Detection Algorithm of Virus Based on Computer Network

Gao Rui

Tianjin Urban Construction Managerment Vocational Technology College
tjgaorui1980@163.com

Abstract

Computer network virus has been manufactured in bulk, which can affect the normal work and study of people, and therefore the defense and detection of computer network virus should be concerned, advanced defense ability of computer network virus under the environment of computer internet can be improved. Firstly, the basic characteristics of computer network virus are analyzed. Secondly, main computer network virus propagation models are studied. Thirdly, active defense technology and detection model of computer network virus are established based on supported vector machine, and corresponding algorithm procedure is constructed. Finally, the simulation experiments are carried out, and results show that the detection model can obtain better detection effect of the computer network virus.

Keywords**:** *Computer network virus; internet; supported vector machine*

1. Introduction

Now the computer network virus is getting stronger and stronger, virus technology are becoming more and more advanced. Moreover, computer network virus manufacturing at home has become group and industrialization, then the computer network virus has been manufactured in bulk. Computer network virus has affected the life, study and work seriously. Therefore the defense and confrontation of computer network virus should be concerned. In recent years computer internet technology has developed quickly and been applied widely, and it has penetrated into all aspects of social production and life. In the internet age the work and life of people is convenient and efficient, people's distances will not remote and work has become free and relaxed. However computer internet suffers all kinds of attack and threat. The computer network virus is a kind of threat means with biggest threat, worst effect and strongest damage capability. In recent years the computer viral infection and attack can cause information leakage and damage of user information. Massive network paralysis is nothing new. In order to resist the attack of computer network virus and maintain safety operation of computer internet some countermeasures and methods are put forward. However, there are large amounts of computer virus, and computer can quickly spread, and it can make variation and evolution constantly, therefore the defense of computer network is more difficult. In order to improve the ability to resist the virus, the relating characteristics and corresponding defense styles of computer network virus should be studied in depth, then advanced defense ability of computer network virus under the environment of computer internet can be improved [1-3].

2. Basic Characteristics of Computer Network Virus

The computer network virus is a group of computer instructions that can damage computer software and hardware function, and affect normal use of computer, it has

ISSN: 1738-9968 IJHIT

capacity to replicate itself. The computer network virus has the following characteristics [4]:

(1) Latent

The latent is an important characteristic for computer virus. Before outbreak of computer virus, computer network virus generally hides in the critical parts of computer system. Because the computer system generally cannot appear exception during the latent period, these computer viruses can get away the scan and check of system safe software, and it can lead to attention of clients. However the latent period is over, computer network virus can reproduce, spread and infect other parts of computer system quickly.

Explosive growth generally makes the infected computer failure in a short time, and the hardware and software was destroyed. The longer the latent period of computer network virus is, the deeper the virus is, and the more difficult the defense is. Once the computer network virus broke out, the computer system can cause serious harm.

(2) Destructive

The main harm of computer network virus mainly is embodied in destructive effect. After the computer network virus infects the computer system, the malicious code is copied in large quantities and executed, and system resources are occupied and the normal operation of programmer is affected in serious case. The important files in computer can be deleted, then the data can lose and computer collapse. In addition, computer network virus can steal and let out password, and make the property and privacy security of user threatened.

(3) Infectious

The computer network virus has the infectious characteristic. The computer infected by virus or the built-in malignant programmer can infect other normal computer through all kinds of methods and channels. For example, the computer network virus can enter the normal computer from the host computer through all kinds of mobile media and internet, and is copied in the computer system infected, and infects other computer system at any time. Finally the computer system connecting with it can be infected, and lose the normal working ability.

(4) Hidden

Computer network virus want to avoid the killing and guarding of anti-virus software and computer system firewall, and masqueraded as application code, and is transplanted into normal programmer. Because the program fragments are little, and can be shielded by normal programmer, therefore it is difficult to find out them. With the development of computer network virus protection technology, people can scan and find out the virus hidden in the computer through feature code and malicious code. However computer network virus evolves constantly, the encryption is carried out through feature code and malicious code, and then the feature matching of computer network virus is harder, the probability that does not be found out by anti-virus software can be improved.

3. Main Computer Network Virus Propagation Models

(1) SIS model

SIS model can make nodes of network divide into easy infection status and infection status, which are denoted by S and I, and the nodes of S status and I status can change each other. Once the nodes of S status are infected as the nodes of I status, and the nodes of I status can become the nodes through killing the virus. And the corresponding model is shown in figure 1.

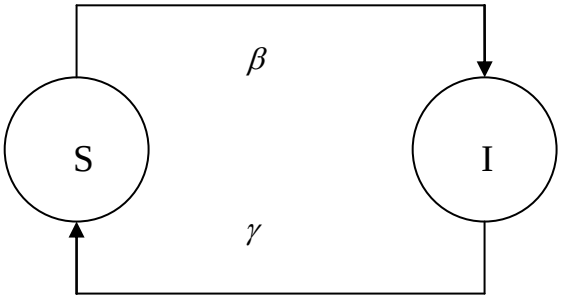

Figure 1. Diagram of SIS Model

Where β denotes the probability that is infected by virus, γ denotes the probability that is checked by the virus.

And the corresponding mathematical model is expressed as follows [5]:

$$\begin{cases} \dfrac{dS}{dt} = -\beta SI + \gamma I \\ \dfrac{dI}{dt} = \beta SI - \gamma I \\ S(0) = N - I_0 \\ I(0) = I_0 \end{cases} \quad (1)$$

where N denotes the total number of nodes, βSI denotes the increment that the nodes of S status are changed into nodes of I status, γI denotes the increment that the nodes of I status are changed into nodes of S status, I_0 denotes the number of computers infected under the original status.

(2) SIR model

SIR model is amended based on SIS model and makes up the disadvantages of SIS model. This kind of model can divide the nodes of network into easy infection status, infection status and immune status, which can denotes by S, I and R. SIR model is shown in figure 2.

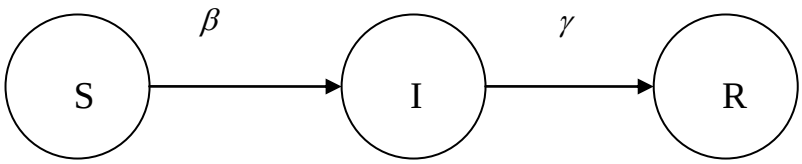

Figure 2. Diagram of SIR Model

The mathematical model of SIR model is expressed as follows [6]:

$$\begin{cases} \dfrac{dS}{dt} = -\beta SI \\ \dfrac{dR}{dt} = \gamma I \\ S(0) = N - I_0 \\ I(0) = I_0, R(0) = 0 \end{cases} \tag{2}$$

SIR model can make up the disadvantages of SIS model, and the immune status can be considered. However SIR exists the following disadvantages.

(3) SEIR model

SEIR model is improved based on SIR model, the latent state is introduced based on S, I and R status, which is denoted by E, the nodes in this status has been infected, and there is no infected nodes. In practical application, the viral infection delays the nodes with virus get symptoms after it is infected for a long time, and the virus can spread. In addition, SEIR model synthesizes the characteristics of SIS model and SIR model, S status and I status can change each other, the SEIR model is shown in figure 3.

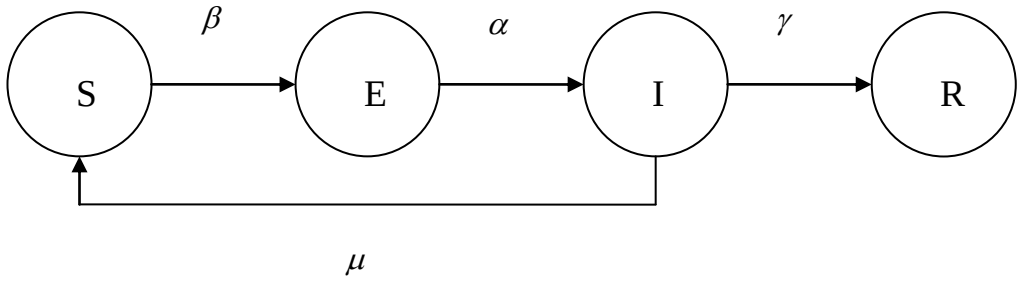

Figure 3. Diagram of SIR Model

Where α denotes the probability that latent status changes to infected status, μ denotes the probability that infected status returns to easy infection status. The corresponding mathematical model is expressed as follows:

$$\begin{cases} \dfrac{dS}{dt} = -\beta SI + \mu I \\ \dfrac{dE}{dt} = \beta SI - \alpha E \\ \dfrac{dI}{dt} = \alpha E - (\mu + \gamma) I \\ \dfrac{dR}{dt} = \gamma I \\ S(0) = N - I_0, E(0) = 0, I(0) = I_0, R(0) = 0 \end{cases} \tag{3}$$

4. Active Defense Technology and Detection Algorithm of Virus

According to feature attributes of computer network virus the eigenvalue matrix is obtained, which is expressed as follows [7]:

$$X_{mn} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (4)$$

Where n denotes the number of procedure samples, m denotes the number of virus features.

Before detection of computer network virus the effect of dimension difference should be eliminated, the normalizing processing should be carried out for eigenvalue, which is expressed as follows:

$$R_{mn} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix} \quad (5)$$

Where r_{ij} $(1 \le i \le n, 1 \le j \le m)$.

The supported vector machine is put forward based on structural risk minimization theory, which is fit for classification of small sample and nonlinear problem. It can be applied in the active defense and detection of computer virus. Classification model of supported vector machine is listed as follows:

The program characteristic quantity is defined by x_{ij}, the program category is defined by y_i, the supported vector machine is to find out optimal classification plane, and all samples can be divided well, and the classification distance is maximum. The optimal separate plane is expressed as follows [8]:

$$\min \ \phi(\omega) = \frac{1}{2}\|\omega\|^2 \quad (6)$$

The constraint condition is listed as follows:

$$f(x) \ge 1 \ i = 1, 2, \cdots, n \quad (7)$$

Where $f(x_i)$ denotes classification plane.

$$f(x) = \omega \cdot x_i + b \ i = 1, 2, \cdots, n \quad (8)$$

Then the decision plane of the optimal classification plane is expressed as follows:

$$g(x) = \mathrm{sgn}(\omega \cdot x_i + b) \quad (9)$$

The Lagrange multiplier is introduced, and the defense and detection of computer network virus can be changed to the following mathematical model:

$$Q(a) = \sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i,j=1}^{n} a_i a_j y_i y_j (x_i \cdot x_j) \quad (10)$$

Where a_i is Lagrange multiplier.

The constraint conditions are listed as follows:

$$\begin{cases} \sum_{i=1}^{n} y_i a_i = 0 \\ a_i \geq 0 \end{cases} \tag{11}$$

The supported vector machine of defense and detection of computer network virus is expressed as follows:

$$f(x) = \text{sgn}(\sum_{i=1}^{n} a_i^* y_i k(x_i \cdot x) + b^*) \tag{12}$$

Where $k(x_i \cdot x)$ denotes the core function of supported vector machine, the Gaussian kernel function is applied in this model.

The defense and detection of computer network virus concludes training and detection stages, which is listed as follows [9]:

Step 1: collect needed program data of computer virus.

Step 2: extract the program features of computer virus, and carry out quantization process, and generate the feature matrix.

Step 3: carry out normalizing processing for characteristic data, and delete the difference between dimension and magnitude for characteristic value.

Step 4: divide data into training collection and testing collection, the training collection is used in constructing the defense and detection model of computer virus, the testing collection can be used in performance detection of model.

Step 5: train the training collection based on supported vector machine, and search optimal parameters of supported vector machine based on network searching algorithm, and construct the defense and detection model of computer virus.

Step 6: detect the testing collection of computer network virus based on optimal detection model, and verify the effectiveness of model.

Step 7: construct the detection model of computer network virus for defense of computer network safety.

5. Experiments and Results Analysis

In order to detect the performance of active defense and detection of computer virus, 12000 viruses of 10 classes are collected from internet. Based on these samples, a normal small type virus database is established in form of database, every item of virus database corresponds to a virus sample, which concludes order, name, class and other properties. All samples can be divided into training and testing samples. The real distribution of computer **network** virus data is shown in table 1.

Table 1. Real Computer Network Virus Data Distribution

Type of sample	Normal program	Computer network virus program
Training sample	9000	1000
Testing sample	1200	800

The active defense and detection of computer network virus algorithm programmer is compiled by MATLAB software, and RBF neural network and Bayesian method are used to carry out detection of computer virus. The correct rate and under-reporting rate are used as evaluation standard, every detection algorithm operates 12 times, and the optimal results are used as final detection results of computer virus. The correct rate and under-reporting rate are calculated by the following expression:

$$\text{correct rate} = \frac{n}{N} \times 100\% \tag{13}$$

$$\text{under} - \text{reporting rate} = \frac{N-n}{N} \times 100\% \tag{14}$$

where n denotes the computer network virus detected, N denotes the total number of computer virus.

The penalty coefficient C and core function width σ should be optimized for supported vector machine, and the training samples are input into the supported vector machine for studying, and the grid search algorithm is used for searching optimal parameters, then the optimal parameters can be obtained as follows: $C = 80$, $\sigma = 1.05$, then the optimal detection model of computer network virus can be established.

The detection results of testing collection based on different methods are shown in table 2.

Table 2. Real Computer Network Virus Data Distribution

Defense and detection method	Correct rate/%	Under-reporting rate/%
Supported vector machine	85.2	14.8
RBF neural network	89.6	10.4
Bayesian method	98.5	1.5

As seen from table 2, the correct rate of computer network virus detection based on supported vector machine can achieve 98.5%, which is higher than that based on RBF neural network and Bayesian method. And the studying speed based on supported vector machine is higher than that of other defense and detection method, which can suit for the real time requirement of computer network virus detection. Therefore the active defense and detection model of computer network virus is an effective method for detecting the attacking activity of computer virus.

7. Conclusions

Active defense and detection of computer network virus is a main technical problem for network. The corresponding active defense and detection model of computer network virus is established, the host computer can achieve zero touch with computer virus. The defense and detection algorithm of computer network virus is constructed. The supported vector machine is used to construct the corresponding model, simulation results show that the correct rate and under--reporting rate of computer network virus decreases, which can be applied in the active defense and detection of computer network.

References

[1] B. Y. Zhang, J. P. Yin, S. L. Wang. Neurocomputing, 5, 137 (**2014**).
[2] W. H. Chang, S. Y. Yang, C. L. Lin. Nanomedicine: Nanotechnology, Biology and Medicine, 8, 9 (**2013**).
[3] X. Han, Q. L. Tan. Applied Mathematics and Computation, 6, 217 (**2010**).
[4] N. A. Seresht and R. Azmi. Engineering Applications of Artificial Intelligence, 10, 35 (**2014**).
[5] C. Laorden, X. Ugarte-Pedrero and I. Santos. Information Sciences, 1, 277 (**2014**).
[6] J. Y. Kwon, J. S. Hong, M. J. Kim. Journal of Virological Methods, 15, 206 (**2014**).

[7] Y. H. Shao, W. J. Chen, J. J. Zhang. Pattern Recognition, 9, 47 (**2014**).
[8] G. E. Güraksın, H. Haklı, H. Uğuz. Applied Soft Computing, 11, 24 (2014).
[9] S. H. Choi, B. E. Min, E. G. Song. Neurocomputing, 22, 140 (2014).

Author

Gao Rui is a lecturer in Tianjin Urban Construction Managerment Vocational Technology College. His research direction is computer application and development.

