Journal of Theoretical and Applied Information Technology
31st January 2013. Vol. 47 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

TROJAN HORSE HIDING TECHNOLOGY RESEARCH BASED ON K - SUMMARY CLUSTERING ALGORITHM

DAN LIU, SHIXIA MA, ZU HUA GUO, XIU LAN WANG

Dept.of Computer Science, Henan Mechanic and Electric Engineering College, XinXiang 453002, Henan, China

ABSTRACT

The rapid development of computer networks has brought great convenience to people's learning, work and life. However, because the network has the characteristics of sharing and openness, and the computer systems often use the open operating system, computer networks are more vulnerably subject to attack from hackers. The network security issues have been particularly prominent. Using of Trojan technology to attack and destroy the network system has become the greatest threat to the security of information networks. On the basis of analyzing the life characteristics, working principle and hiding technique of Trojans, the paper gives the Trojan removal method. Meanwhile, the paper proposes a K-summary clustering algorithm. The algorithm can accurately locate the Trojan position and significantly saving the positioning running time, which has certain practical significance to improve the security of the network.

Keywords: *Trojan Principle, Hiding Techniques, K-summary Clustering Algorithm, Hang Horse Way, Trojan Positioning*

1. INTRODUCTION

With the rapid development and wide application of computer and network technology, computer and communication networks have become indispensable basic part. They have brought great convenience to the learning, work, life and entertainment of people and play an important role in national security, economic development, national education, and modern management. However, because of the inherent weaknesses of computer system and information network system, the openness and sharing makes them vulnerable to outside attack. And the security and confidentiality of the information also are subjected to the serious threat [1]. Network security problem has become the greatest concern of the world's governments, enterprises and vast network users. Among them, using of Trojan invasion, control and destroy the networks information systems is one of the typical performance to produce the network and information security problems.

With the increase of Internet users, the influence of Trojan is also growing. The in-depth study to Trojans can make up for the lack of Trojan knowledge and raise awareness of Trojan and the vigilance of network security, which can greatly improve network security. The studies of international information security start early, while network construction in China has developed rapidly, but also made great achievements. However, due to lacking of a good program to solve the issue of network security and insufficient investment of the network security construction, network security issues are still quite serious. According to the statistics of National Computer Virus urgent treatment center virus sample library, new virus samples add to 2.99 million in 2009, which are 3.2 times than the 2008 new virus [2]. And web pages Trojans appear composite trend. Information and network security capabilities are still in the development stage in China, and many applications systems are still in the undefended state. So the ability of our network system against attack is weak, which forces us to further improve network security and defense capability and increase to the Trojans research and prevention efforts on behalf of the Trojans virus. In addition, network security [3] issues are the biggest factor restricting the rapid development of Internet commerce degree.

The research of this paper enables people to be no longer strangers in front of the Trojans. Through analyzing of Trojan hidden technology, it is easy to see the Trojan prototype to improve the prevention and killing capability of people on the Trojans. Through understanding of the Trojan works and Trojan habits Trojans, people will have a clearer understanding of the Trojans which allows people to easily find the Trojan hiding place, and raise the safety awareness of the majority of Internet users. This method of work greatly reduces the possibility of stealing by the Trojans, which have to practical significance for the improvement of the security of the entire network.

2. TROJAN OVERVIEW

2.1 Trojan working principle

Trojan [4] [5] is a section code which is used to remotely control another computer through writing a program (Trojans program). The Trojans have two executable programs: a program is server that is control terminal; another program is client, which is the control end. The computer being planted is part of the "server", the so-called "hacker" enters and runs a "server" computer using of the "controller". The client end is the part controlling the target host which is installed on the computer of hackers. It mainly connects with the server installing the Trojan which is used to control a remote computer.

After the Trojans executed on the target computer, the Trojans will open a common port to listen. When a client requests a connection, the response will produce from the server, then a connection will establish as follows. The clients control the target computer through a series of instructions. Because Hacker uses of a variety of means to hide the Trojan, the Trojans are not found. Once the server running Trojan connects with the control terminal, the control terminal will enjoy most of privileges of the server such as increasing password to the computer, browsing, moving, copying, deleting files, modifying the registry, changing the configuration of your computer and so on. Trojan is a tool for remote control with the following characteristics including remote portability, spontaneity and controllability [6]. The majority will not directly produce harm on the computer, but to control.

2.2 Hiding Method of Trojan

Trojan is a virus program based on remote control which has strong crypticity and harmfulness. It can control or monitor you in the secrecy of state. In order to avoid being found, most network Trojan use of the hiding processing techniques to hide themselves. The early network Trojans are mostly hidden in the task bar. Now, with the development of hiding technology, Trojan mostly uses of the core embedding method to achieve hiding themselves. The current Trojan hiding technology includes two types which are communication hiding and host hiding. Communication hiding consists of channel hiding and transfer content hiding. Host hiding are many ways such as hiding files, hiding core connection module, hiding processes, hiding network connection and so on.

2.2.1 Communication hiding

The traditional Trojan communication hiding methods merely encrypt the transfer contents, but they can not hide the channel. While hidden network channel technology can either hide channel or hide the content. If it choose the 80-port for communicating allowed by the majority of the security policy, it can easily penetrate the firewall to escape the detection of the intrusion detection system and has a strong concealment. The traditional channel covert technology is carried out between the processes [7], but the covert channel may also be used on the network. Because there is a lot of redundancy in the TCP / IP protocol suite, this information redundancy can be taken advantage to create a covert channel. Trojan precisely uses of these network covert channel to open the gates of network security [8].

2.2.2 Host hiding

Host hiding refers to concealment means that Trojans programme written by hackers avoid users finding them. The appropriate testing programmes can be replaced by the Trojans. The Trojan can be embedded in the legal programmes [9]. And host hiding can be achieved through using of the defects.

2.2.3 Latest hiding technology

The latest hiding technology [10] [11] loads Trojan by modifying the dynamic-link library (DLL) or modifying the virtual device driver (VXD). The Trojan replaces the pre-modification DLL for the modified DLL and filters the other function calls. Under normal circumstances, the DLL is only used for monitoring. Only upon receiving the request, it will activate itself attack. This Trojan does not require adding new content and does not need to open a new port, and also does not increase the new process. So the conventional method can't monitor it.

3. NETWORK TROJAN DETECTION ALGORITHM

K-summary algorithm improves k-prototypes algorithm and prompts the k-means algorithm. It can handle mixed-attribute data sets, and improves the clustering accuracy,

The space complexity of K-summary algorithm and the k-prototypes algorithm are almost the same. The difference between the two is different in cluster representation and distance calculation method. But k-prototypes algorithm only needs a comparison with the frequency biggest attribute value and time consuming is slightly less then k-summary algorithm. For data sets with classification properties, the clustering quality of k-summary algorithm is slightly better.

Clustering summary information CSI can be adopted for being represented for cluster. Each cluster set two CSI: OldCSI and NewCSI. The former is mainly used to save the last iteration results, and used to calculate the distance in the current iteration. NewCSI is used to save the dynamic change of CSI in time of adding object.

The steps of k-summary are as follows:

The first step: the initialization that is to choose k objects and create CSI of k clusters;

The second step: dividing the object to the closest cluster;

The third step: recalculating CSI of each cluster;

The fourth step: repeating the second step and the third step until completing convergence of the selecting measure function.

The formal description of K-summary algorithm is as follows:

```
int _tmain(int argc, _TCHAR* argv[])
{
 Open(Data_base);
 Init(k);
 While(1)
{
  p=MoveFirst();/
  While(!eof())
 {
   Ci=FindNearestCluster(p); \
   AddTupleToCluster(p,Ci);
   p=MoveNext();
 }
  InitCluster();
}
 return 0;
}
```

(2) The time and space complexity of k - summary algorithm depends on the size of the data set N, attribute number m, clustering (CSI) number k and the size of each CSI, iterations time μ. Assume that every classification properties have a different values and the operation number is the same between different object on the numerical attribute part. But for classification properties, there need to compare the object classification properties with different attributes in CSI, so the operation frequency is different of different objects.

For detection the scalability of algorithm, two subsets T1 and T2 are randomly chosen in the T, T1 contains 38838 records and T2 contains 19542 records. Table 6.2 gives the average execution time and average clustering accuracy of k-summary algorithm in T, T1, and T2 under repeatedly operating 10 times.

Table 1: Time Of K - Summary Algorithm In Different Scale Data Set

Clustering Son set	T	T1	T2
average execution time （s）	2586	253	107
average clustering accuracy （%）	98.67	99.62	99

The results show that the algorithm execution time is the approximate sexual function of the size of data set. The algorithm has very good scalability. The algorithm is applied to network Trojan detection, which can save a lot of testing time and greatly improve the efficiency of the detection Trojans.

4. TROJAN REMOVAL

4.1 Check The Network Connection

Some network managers have abundant experience, and they often check network connection and find Trojan. Most Trojans are intelligent, and they can forwardly listen in port and can connect with the specified IP and port. We can click Start -> Run -> cmd, then type “netstat-an” and “Enter”. There will show all IP which establish a connection with the local computer and the listening ports [44]. They contain four parts: local address (address of the local computer), proto (connection method), foreign address (addresses connecting with local computer), state (port status). Through adopting this approach, the network connection situation of this computer can completely be monitored, and then Trojans can be searched in the open ports, such as 54320(Back Orifice 2000), 7626 (Glacier Trojan) and so on.

4.2 Backup The Most Common Process And Investigate The Suspicious Process

Firstly, we backup a process list, and then find suspicious process through the comparison. The specific operation is as follows: After the computer is booted, firstly we should backup process and then carry out any other operation. The aim is to backup the process before the other applications will be loaded. Then, we can click start - > run - > CMD - > tasklist/SVC and type enter. There will show the current image name, PID, service list. Then input \"tasklist /?\" to display other parameters of the order, which can find out the suspicious process through comparing.

4.3 Check Registry

(1) Ordinary users are unknown to the registry, but Trojan horse often patronize there. In order to achieve startup with the system and other functions, Trojan horse will modify the registry. 1)Click start - > run - > regedit, type “enter” and open the registry editor, then check all the key value beginning of the "run" in HKEY_CURRENT_USER \ Software \ Microsoft \ Windows \ CurrentVersion; 2)Check whether a suspicious content exists; 3)Check HKEY_LOCAL_MACHINE \ Software \ Microsoft \ Windows \ CurrentVersion \ Explorer, respectively open the contents of the son key “Run”, “User Shell Folders”, “RunOnce”, “Shell Folders”, “RunOnce” and “RunServices”. Once the procedures being not familiar are found, we should be aware of them.

(2) Some Trojans achieve self-starting by adding service items. The checking method is as follows. 1) open the registry, and check the suspected primary key in HKEY_LOCAL_MACHINE \ SOFTWARE \ Microsoft \ Windows \ CurrentVersion \ Runservices; 2) Find the suspicious key in HKEY_LOCAL_MACHINE \ System \ CurrentControlSet \ Services Viewkey; 3)Delete or disable the services items added by the Trojans;4) Click Start -> Run -> Services.msc and open the service settings window, which displays the name, description, status, startup type and logon type information of all service items in the system; 5)Find the service started by the Trojans inside the name; 6) click right key -> Properties -> General -> Startup type -> Disable, and click OK, when the Trojans item is closed.

4.4 View System.Ini And Win.Ini File In The System

The Win.ini and System.ini in System INI file are frequently visited by the Trojans. The checking method is as follows. 1) Type "% systemroot%" in “Run” command and enter into the "Windows" folder after type “enter”; 2) find the win.ini file in the file. And in this file, see if there exist the "load = xxx.exe, run = xxx. The exe " statement (xxx.exe is program name of Trojan), because this is likely to be the main program of Trojan. Similarly, in the system.ini file to find whether there is "Shell = aaa.exe". The default should be "Shell = Explorer.exe", if found, this file could be infected with the Trojans.3) In addition, the key value “HKEY_LOCAL_MACHINE \ Software \ classes \ exefile \ shell \ open \ command” of registry may also be loaded the Trojans. For example, the key is modified into the "X: \ windows \ system \ a.exe"% 1 "%".

4.5 Check The System Account

The attacker can control your computer through setting the account. The method for testing the account is as follows. 1) Click start - > run - > CMD - > net user and type enter to check users; 2) Input the user name of "net user” and check the user's authority. Usually, only “Administrator” is attributed to the administrators group besides, others are not administrators group. When you find a system built-in user belongs to the administrators group, it can be sure that your computer has been invaded. You can use the command "net user name/del" just to delete it. If it detects the presence of Trojan horse, the following way can be used to clear the Trojan [5]:

(1) Open the task manager and close the Trojan process.

(2) Check the son key in registry such as Run, User Shell Folders, Shell Folders and RunServices and so on. Firstly backup registry, writes down the starting item address, and then deletes the suspicious items.

(4) Normally this kind of documents exists in these folders such as WINNT, SYSTEM, and SYSTEM32. They do not usually exist alone, and probably copy over by a mother file. Check the suspected files including “*. Bat, *. Exe or *. Com file” in hard disk. If there exist, delete them.

(5) Check registry HKEY_LOCAL_MACHINE and several key items (such as Local Page) in HKEY_CURRENT_USER \ software \ Microsoft \ Internet Explorer \ Main. If they are modified, and then change it back.

(6) Check whether the default open program of used file types including Hkey_Classes_Root \ txtfile \ DefaultIcon and Hkey_Classes_Root \ txtfile \ shell \ command is modified in the registry. If they are changed, change them back. A lot of

virus is loaded when the users open the text file through modifying the*.TXT file's default open program.

5. CONCLUSION

This paper simply introduces the work principle and the hidden way of Trojan, and uses of k - summary algorithm to rapidly search positioning and analyze Trojan. This algorithm can accurately position the position of Trojan and timely repair of security vulnerabilities, so that the Trojans have no chance of survival, which improves the killing efficiency of Trojan and saves running time. This is very important practical value and significance to improve the entire network safety.

REFERENCES:

[1] Lorine A. Hughes, Gregory J. DeLone, "Viruses, Worms, and Trojan Horses: Serious Crimes, Nuisance, or Both?", *Social Science Computer Review*, Vol. 25, No.1, 2007, pp. 78-98.

[2] Jin Xu, Zhao Xue-jun, Ma Shi-xia, "Analysis and Prevention of Network Trojan", *Computer Knowledge and Technology*, Vol. 6, No. 23, 2010, pp. 6450-6452.

[3] Xu Ming, Yang Tong, Zheng Lian-qing, Zhang Chuan-rong, "Concealing technology of Trojan horses and prevention", *Computer Engineering and Design*, Vol. 32, No.2, 2011, pp.489-496

[4] Lansheng Han , Xiao Qu , Yuan Li , Tao Yin, "The Probability of Trojan Attacks on Multi-level Security Strategy based Network", *Journal of Networks*, Vol. 7, No. 2,2012, pp.300-304

[5] Ghossoon. M. W. Al-Saadoon , Hilal M. Y. Al-Bayatti, "A Comparison of Trojan Virus Behavior in Linux and Windows Operating Systems", *World of Computer Science and Information Technology Journal* , Vol. 1, No. 3, 2011,pp.56-62

[6] Yongchang Ren , Tao Xing , Guanghui Cao , Xu , E. , Xiaoji Chen "Research and practice on the cooperative concealing technology of Trojan horses", *Networking and Digital Society (ICNDS)* ,Vol.1, 2010, pp.216-219

[7] Cohen F， "A Cryptographic checksum for integrity protection", *Computers and Security*, Vol. 6, No. 6, 1987, pp.505-510

[8] Cohen F, "On the implications of computer virus and methods of defense", *Computers and Security*, Vol. 7, No. 2,1988，pp.167-184

[9] Cohen F， "Models of practical denfenses against computer vimscs", *Computers and Security*, 1989, Vol. 8, No. 2,pp.149-160

[10] Thimbleby H, Anderson S， Cairns P, "A framework for modeling Trojans and computer virus infection", *The Computer Journal*, Vol.44, No. 7,1998, pp. 444-458.

[11] Liu Lan, Gao Yue-xiang, " Study and Analysis of Concealing Technology for Trojan Horses", *Communications Technology*, Vol.43, No.4, 2010, pp.78-80.