WORM'S BEHAVIOR ANALYSIS SYSTEM

TEH WEI HAO

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

© Universiti Teknikal Malaysia Melaka

BORANG PENGESAHAN STATUS TESIS

JUDUL: WORM'S BEHAVIOR ANALYSIS SYSTEM

SESI PENGAJIAN: SESI 2010/2011

Saya TEH WEI HAO mengaku membenarkan tesis (PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan penukaran antara institusi pengajian tinggi.
4. Sila tandakan(/)

_____	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
_____	TERHAD	(Mengandungi maklumat TERHAD yang ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
___/___	TIDAK TERHAD	

(TANDATANGAN PENULIS)

(TANDATANGAN PENYELIA)

Alamat tetap: No.16-1,JLN Orkid,86100 Ayer Hitam,BatuPahat, Johor, Malaysia

Tarikh: 04.07.2011

EN OTHMAN BIN MOHD

Nama Penyelia

Tarikh: 04.07.2011

© Universiti Teknikal Malaysia Melaka

WORM'S BEHAVIOR ANALYSIS SYSTEM

TEH WEI HAO

This report is submitted in partial fulfillment of the requirements for the

Bachelor of Computer Science (Computer Networking)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2011

© Universiti Teknikal Malaysia Melaka

DECLARATION

I hereby declare that this project entitled

WORM'S BEHAVIOR ANALYSIS SYSTEM

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT:_____ DATE: 04.07.2011

(TEH WEI HAO)

SUPERVISOR:_____ DATE: 04-07-2011

(EN OTHMAN BIN MOHD)

© Universiti Teknikal Malaysia Melaka

DEDICATION

To my beloved parents…

© Universiti Teknikal Malaysia Melaka

ACKNOWLEDGEMENTS

I would like to show my gratitude and appreciationto my supervisors, Mr Othman Bin Mohd and Mr Zulkiflee Bin Muslim for all ideas and advices in guiding me throughout the project.

I would also like to thank my family members especially my parents. They have been giving me moral supports and all sorts of material supports throughout my years studying in this University.

Last but not least, I would like to say thank you to all my friends and course mates for their kindness in sharing knowledge and resources.

Thanks a lot.

© Universiti Teknikal Malaysia Melaka

ABSTRACT

A computer worm is a self-replicating malware computer program, which uses a computer network to send copies of itself to other nodes (computers on the network) and it may do so without any user intervention. This is due to security shortcomings on the target computer. The purpose of this project is to develop an analysis system that can use to analyze the computer worm's behavior based on the output generated by packet sniffer software known as "tcpdump". Text file generated by tcpdump is not directly related to any database and statistical analysis. As a result, analysis work becomes inconvenient and less reliable. In this project, a system will be developed to import the output of tcpdump into database and use the data in database to conduct statistical analysis job. The methodology used in this project is waterfall model so that the system can be developed smoothly and fulfill all the scopes and objectives of this project.

© Universiti Teknikal Malaysia Melaka

ABSTRAK

Worm komputer adalah satu jenis malware komputer program yang boleh replikasi diri dengan menggunakan rangkaian komputer untuk menghantar replikasinya ke node-node yang lain dan boleh menjalani aktiviti tersebut dengan tidak diketahui oleh pengguna. Ini adalah kerana kekurangan keselamatan dalam target komputer. Tujuan projek ini ialah membangunkan satu analisis sistem yang boleh digunakan untuk menganlisis perilaku worm komputer berdasarkan output yang dihasilkan oleh tcpdump. Teks fail yang dihasilkan oleh tcpdump tidak mempunyai sebarang perhubungan dengan database dan statistik analisis. Sehubungan itu, kerja analisis menjadi susah dan tidak handal. Dalam projek ini, satu system and dibangunkan untuk import tcpdump output ke dalam database dan menggunakan data tersebut untuk menjalankan kerja statistik analisis. Metodologi yang digunakan dalam projek ini adalah waterfall model. Metodologi ini digunakan supaya sistem dapat dibangunkan dengan lancar dan memenuhi semua skop dan objektif dalam projek ini.

© Universiti Teknikal Malaysia Melaka

Table of Contents

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF TABLES

© Universiti Teknikal Malaysia Melaka

LIST OF FIGURES

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF ABBREVIATIONS

ALPHABET	WORDS	EXPLANATION
A	ARP	Address Resolution Protocol
D	DOS	Denial of Services
H	HTTP	Hypertext Transfer Protocol
I	IPv6	Internet Protocol Version 6
	IPv4	Internet Protocol Version 4
	IETF	Internet Engineering Task Force
	ICMP	Internet Control Message Protocol
O	OOA	Object-oriented analysis
	OOD	Object-oriented design
R	RFC	Request for Comments
T	TCP/UDP	Transmission Control Protocol/ User Datagram Protocol
	TCP/IP	Transmission Control Protocol/ Internet Protocol
U	URL	Uniform Resource Locator
	UML	Unified Modeling Language

© Universiti Teknikal Malaysia Melaka

LIST OF ATTACHMENTS

© Universiti Teknikal Malaysia Melaka

CHAPTER I

INTRODUCTION

1.1 Project background

IPv6 is a version of the Internet Protocol that is designed to succeed IPv4, the first publicly used Internet Protocol in operation since 1981. IPv6 is an Internet Layer protocol for packet-switched internetworking. The main driving force for the redesign of Internet Protocol was the foreseeable IPv4 address exhaustion (Hossam Afifi, 1999). IPv6 was developed by the IETF, and is described in Internet standard document RFC 2460, published in December 1998. Since the replacement of IPv4 to IPv6 is inevitable, network security on IPv6 network should also be enhanced to protect the network. One of the protection methods is the implementation of IDS.

Tcpdump is a common packet analyzer that runs under the command line. It allows the user to intercept and display TCP/IP and other packets being transmitted or received over a network to which the computer is attached. It was originally written in 1987 by Van Jacobson, Craig Leres and Steven McCanne who were, at the time, working in the Lawrence Berkeley Laboratory Network Research Group. Distributed under the BSD license, tcpdump is free software. Tcpdump works on most Unix-like operating systems: Linux, Solaris, BSD, Mac OS X, HP-UX and AIX. In those systems, tcpdump uses the libpcap library to capture packets. There is also a port of tcpdump for Windows called WinDump and uses WinPcap, which is a port of libpcap to Windows.

© Universiti Teknikal Malaysia Melaka

Tcpdump provides the ability to analyze network behavior, performance and applications that generate or receive network traffic. It can also be used for analyzing the network infrastructure itself by determining whether all necessary routing is occurring properly, allowing the user to further isolate the source of a problem. It is also possible to use tcpdump for the specific purpose of intercepting and displaying the communications of another user or computer. A user with the necessary privileges on a system acting as a router or gateway through which unencrypted traffic such as Telnet or HTTP passes can use tcpdump to view login IDs, passwords, the URLs and content of websites being viewed, or any other unencrypted information. The user may optionally apply a BPF-based filter to limit the number of packets seen by tcpdump. This renders the output more usable on networks with a high volume of traffic.

Malware, short for malicious software, (sometimes referred to as pestware[1]) is a software designed to secretly access a computer system without the owner's informed consent. The expression is a general term used by computer professionals to mean a variety of forms of hostile, intrusive, or annoying software or program code[2]. Software is considered to be malware based on the perceived intent of the creator rather than any particular features. Malware includes computer viruses, worms, Trojan horses, spyware, dishonets adware, scareware, crimeware, most rootkits, and other malicious and unwanted software or program. In law, malware is sometimes known as a computer contaminant, for instance in the legal codes of several U.S. states, including California and West Virginia.[3][4]

In this project, types of the malware will be using to release into the network is computer worms. A computer worm is a self-replicating malware computer program, which uses a computer network to send copies of itself to other nodes (computers on the network) and it may do so without any user intervention. This is due to security shortcomings on the target computer. Unlike a computer virus, it does not need to attach itself to an existing program. Worms almost always cause at least some harm to the network, even if only by consuming bandwidth, whereas viruses almost always corrupt or modify files on a targeted computer.

© Universiti Teknikal Malaysia Melaka

The actual term "worm"' was first used in John Brunner's 1975 novel, *The Shockwave Rider*. In that novel, Nichlas Haflinger designs and sets off a data-gathering worm in an act of revenge against the powerful men who run a national electronic information web that induces mass conformity. "You have the biggest-ever worm loose in the net and it automatically sabotages any attempt to monitor it... There's never been a worm with that tough a head or that long a tail!"[5]. In November 2, 1988, Robert Tappan Morris, a Cornell University computer science graduate student, unleashed what became known as the Morris worm, disrupting perhaps 10% of the computers then on the Internet and prompting the formation of the CERT Coordination Center CERT[6] and Phage mailing list. Morris himself became the first person tried and convicted under the 1986 Computer Fraud and Abuse Act[7].

In this project, tcpdump will be installed as packet sniffer to monitor the IPv6 network, which is infected by worms. Text file generated will be used for further analysis. A system will be developed to read the text file and save into database. Graph will be generated by the system to analyze the traffic pattern of the network based on worms attack.

1.2 Problem statement

Currently, there been little effort to study the traffic pattern of real environment infected by worms in IPv6 network. Main reason of this phenomenon is due to lack of demand from network users. Despite years of forecasts of "doom and gloom" when discussing the famous exhaustion of IPv4 addresses there has been little uptake of IPv6 with the exception of countries such as Japan which has been very active in promoting it[8, 9]. This has meant that very little non-research traffic has actually travelled over IPv6 and hence the impetus for new malware attack making use of IPv6 features has been absent.

Since most internet attack focus in IPv4 network, there is a need to figure out how worms affected network traffic in IPv6 network. In this project, computer

© Universiti Teknikal Malaysia Melaka

worms will be released into an IPv6 network. An open source software called Tcpdump will be used to sniff the network and monitor how the worms affect the network. Output generated by tcpdump will be used for further analysis.

Normally, tcpdump will produces one line of output for each packet that it sees. For each connection, tcpdump will always displays (except in very special cases) a timestamp, a source IP address, a destination IP address, and some additional information about the packet (such as protocol and port information). Further, tcpdump has a default standard output based on the protocol [10]. On the tcpdump output in Listing A, we can see the UDP, TCP, ICMP, and ARP protocol information.

```
#tcpdump
tcpdump: listening on hme0
11:17:36.965387 192.168.10.1.1028 > 192.168.10.5.700: udp 82
11:17:42.645580 pine.tree.com.34342 > birch.tree.com.telnet:
=>S 1819099388:1819099388(0)
win 8760 <mss 1460> (DF)
11:17:50.011072 pine.tree.com > oak.tree.com: icmp: echo request (DF)
11:17:50.011091 oak.tree.com > pine.tree.com: icmp: echo reply (DF)
11:17:55.870599 arp who-has 192.168.10.1 tell 192.168.10.55
```

Listing A show the example output generated by tcpdump.

It is quite inconvenient if the tcpdump users read the output line by line and often cause confusing among tcpdump users. So, there is a need to develop a system to read the output from tcpdump, not only to make the analysis work easier and also more systematically. The research problem and research question are shown in the table 1.1 and table 1.2.

Table 1.1 shows the research problems in this project.

No	Research Problem
RP 1	Traffic pattern in IPv6 environment is one of the major problem

© Universiti Teknikal Malaysia Melaka

	when it infected by worms.
RP 2	There is no system to transferring data from tcpdump to Microsoft SQL Server database for data analysis.

Table 1.2 shows the research problems and research questions in this project.

RP	No	Research Question
1	1	What is the worms' behavior under IPv6 environment?
2	2	How to analyze and evaluate data generated by tcpdump in a systematic way?

1.3 Objective

Objectives of PSM project that is going to be conducted as follows:

- To analyses the traffic pattern of real environment based on worm attack: Graph will be generated using the system developed for analysis.
- To develop a data analysis system based on worm attack: A system will be developed using Microsoft Visual Basic to read the text file generated by the tcpdump and save the data into mysql.
- To evaluate the worms' behavior: traffic pattern of the infected network will be analyzed based on graph generated in the system to evaluate worms' behavior.
- Table 1.3 shows the research problems, research question, and research objectives in this project.

Table 1.3 Research problems, research questions, and research objectives in this project

RP	RQ	RO	Objective
1	1	1	To analyses the of real traffic environment based on worm attack.
2	2	2	To develop a data analysis system based on worms attack.
2	2	2	To evaluate the behavior of worms under IPv6 environment.

© Universiti Teknikal Malaysia Melaka

1.4 Scope

Scope of PSM project that is going to be conducted as follows:

- Data capturing will be conducted under IPv6 environment at Security Lab FTMK: Data capturing process will be conducted in IPv6 network owned by Encik Zulkiflee Bin Muslim inside Security Lab.
- Three HP workstations are required in this project: 1 workstation with Linux Fedora 14 as operating system will be installed tcpdump to monitor the network traffic while the others 2 HP workstations will be installed with Microsoft Windows as operating system and used to release worms into network.
- Tcpdump will be used in this project to sniff the network.
- Data capturing duration will be conducted for 3 days.

1.5 Project significance

To give a clear explain how worms affect in IPv6 environment and process tcpdump output in a systematic manner.

1.6 Expected Output

A data analysis system will be developed for users to read the tcpdump output systematically and make the analyzing job easier.

© Universiti Teknikal Malaysia Melaka

1.7 Conclusion

In conclusion, this project will develop a data analysis system that based on worms attack in IPv6 environment. The purpose of this project is to hope that it is able to help tcpdump users to read its output in a systematic manner and learn how worms affect our network. In activities of next chapter, I will do more research about worms and tcpdump. At the same time, literature review will also be done. Project methodology will also be clearly explained in chapter II.

© Universiti Teknikal Malaysia Melaka