VISUALIZATION OF MALWARE BEHAVIOR USING MATRIX

MUHAMMAD HAFIZUL HELMI BIN MOHD ZURIN

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS TESIS

JUDUL: Visualization of Malware Behavior Using Matrix

SESI PENGAJIAN: 2016 / 2017

Saya MUHAMMAD HAFIZUL HELMI BIN MOHD ZURIN
(HURUF BESAR)

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 pertukaran 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 telah ditentukan oleh organisasi/ badan di mana penyelidikan dijalankan)

___/___ TIDAK TERHAD

(TANDATANGAN PENULIS)

Alamat tetap: 110, Kampung Jawa,

84500, Panchor, Muar, Johor.

Tarikh: 25/8/2017

(TANDATANGAN PENYELIA)

DR. SITI RAHAYU SELAMAT

Tarikh: 25/8/2017

CATATAN: * Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)

** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

VISUALIZATION OF MALWARE BEHAVIOR USING MATRIX

MUHAMMAD HAFIZUL HELMI BIN MOHD ZURIN

This report is submitted in partial fulfilment of the requirement for the

Bachelor of Computer Science (Computer Networking)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2017

DECLARATION

I hereby declare that this project report entitled

VISUALIZATION OF MALWARE BEHAVIOR USING MATRIX

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

STUDENT : ____________________________
(MUHAMMAD HAFIZUL HELMI BIN MOHD ZURIN)

Date: 25/8/2017

SUPERVISOR : ____________________________
(DR SITI RAHAYU SELAMAT)

Date: 25/8/2017

DEDICATION

To my beloved parents thank you very much and a alot
for always supporting me
and being there when I am feeling down

To my loyal friend thank you
for sharing your knowledge, motivate me and helping me
in completing this project

To my supervisor thank you
for encouraging, motivating and believing
in me

ACKNOWLEDGEMENTS

I would like to show my gratefulness to Allah SWT, who with His willing give me the opportunity to complete this Final Year Project which titled Visualisation of Malware Behavior Using Matrix. Next, I would like to express how thankful I am to Dr Siti Rahayu Selamat as my supervisor who had guided a lot of task during this semester in completing this Final Year Project. Deepest thanks to my mother give motivation and appreciation to my father, family and my supportive friends and others for their assistance, encouragement, constructive suggestion and full support for the report completion, from the beginning till the end. Last but not least, my thanks to the members of Faculty of Information Communication and Technology UTeM, for commitment and cooperation during my Final Year project.

ABSTRACT

Malware is a type of malicious program that replicate from host machine and propagate through network. It can take form of executable code, scripts, active content and other software. The development of new malware is increases every year. We need to analyze the malware behavior in order to detect their attack pattern. However, malware behavior is hard to understand by non-technical viewers. This research will perform analysis for malware behavior and construct matrix for malware behavior to provide better understanding. The method used in this research consists of five approaches. First, the network environment will be set up in this research. After that, the malware attack is activated. The network traffic data will be collected. Then, all network traffic data will be analyzed. Finally, matrix will be constructed in order to visualize the malware behavior. The expectation by the end of this project is to represent the malware behavior by visualize it using matrix. Hence, this will facilitate an administrator to identify the behavior of malware during the threat analysis. Besides that, it can provide better view for others to understand malware behavior in visual form.

ABSTRAK

Malware adalah sejenis program yang boleh memberi kesan buruk kepada komputer mangsa dan ia boleh disebarkan melalui rangkaian. Ia juga boleh disebarkan dalam bentuk kod, skrip, kandungan aktif dan perisian lain. Perkembangan malware baru meningkat setiap tahun. Kita perlu mengenalpasti tingkah laku malware untuk mengesan cara ia menyerang. Walau bagaimanapun, tingkah laku malware sukar difahami. Kajian ini akan menjalankan analisis untuk tingkah laku malware dan membina jadual matriks untuk memberikan pemahaman yang lebih baik. Kaedah yang digunakan dalam kajian ini terdiri daripada lima pendekatan. Pertama, menyediakan persekitaran rangkaian. Selepas itu, serangan malware akan diaktifkan. Data trafik rangkaian akan dikumpulkan. Kemudian, semua data trafik rangkaian akan dianalisis. Akhir sekali, jadual matriks akan dibina untuk menggambarkan tingkah laku malware. Harapan pada akhir projek ini adalah memberikan pemahaman tentang tingkah laku malware dengan menggunakan jadual matriks. Oleh itu, ia memudahkan dalam mengenal pasti tingkah laku malware semasa proses menganalisis. Selain itu, ia dapat memberikan pandangan yang lebih baik untuk orang lain memahami tingkah laku malware dalam bentuk visual.

TABLE OF CONTENTS

CHAPTER VI TESTING

CHAPTER VII PROJECT CONLUSION

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Background Study

Malware is short for malicious software. It is referring to any software that is inserted without any authorize into a computer system to comprome the confidentiality, integrity, or availability of the victim's data, applications, or operating systems. Malware is malicious code as any code added, changed, or removed from a software system in order to intentionally cause harm or subvert the intended function of the system (McGraw & Morrisett, 2000).

The number of new type of malware released has increased day by day. Malware is not only executed in windows operating system. It also can be executed in smartphone, tablet, and other operating system such as macOS and Linux. Since Windows is used widely, the statistics shows the highest amount of malware attack was occurred in Windows operating system. Malware can be classified based on their behavior. There are two approaches towards analyzing a malware sample which is dynamic analysis and static analysis. Dynamic analysis is a technique for studying the behavior of a malware sample while the sample is being executed. However, static analysis is a technique that enables the study of a sample without the need for sample execution (Band & Antenna, 2014). Based on this problem, we need to expose to users on malware behavior. However, malware behavior is hard to understand by non-technical viewers. Visualization on malware behavior is needed to give more understanding on how they attack and affect the system.

Nowadays, many existing method of visualizing malware behavior have been done previously. Malware behavior visualization could possibly open up a new paradigm for malware research. There are currently 4 methods of malware visualization. These are Malware Treemap, Malware Threadgraph, Malware Image, and visualization of Executables for Reversing and Analysis (VERA) (Band & Antenna, 2014). In this research, a new technique to visualize malware behavior using matrix is presented.

1.2 Problem Statement

Malware behavior should be documented in the visual form that can be used in presentation process. Besides that, it can provide better understanding for others to translate malware behavior in visual form.

Table 1.1: Problem Statement

No	Project Problem
PP1	Malware behavior is hard to understand by non-technical viewers

1.3 Project Questions

Based on the problem statements, two project questions (PQ) are constructed as shown in Table 1.1 below.

Table 1.2: Project Question

PP	PQ	Project Question (PQ)
PP1	PQ1	How could we identify the malware behavior?
	PQ2	What is the effective visualization technique?

1.4 Project Objective

In order to solve the problem identified as in Section 1.1, two project objectives (PO) are derived as shown in Table 1.2.

Table 1.3: Project Objective

PP	PQ	PO	Project Objective (PO)
PP1	PQ1	PO1	To analyze malware behavior
	PQ2	PO2	To construct matrix for malware behavior visualization

1.5 Project Scope

The scope for this project are:

1. The data used in this project is limited to the types of malware that is discovered and tested.
2. The result acchieved are based on the data in a controlled environment experiment and testing.

1.6 Expected Output

The expectation by the end of this project is to represent the malware behavior by Visualize it using matrix. Hence, this will facilitate an administrator to identify the behavior of malware during the threat analysis.

1.7 Report Organization

Chapter 1: Introduction

This chapter explained about the definition, background, problem statement, objective, scope and expected output related to the malware.

Chapter 2: Literature review

This chapter explained about malware, malware behavior analysis, and the visualization techniques of malware behavior. It will help to more understanding about malware behavior and the methods to identify the behavior for various types of malware.

Chapter 3: Methodology

This chapter provide a decision of the method or what analysis techniques to be used for experimental part. With the certain analysis technique, it helps to know about the malware behavior. It also will involve about the method to visualize it.

Chapter 4: Design and implementation

The design of visualize malware behavior in matrix form is describe in details on how it works carried out. The sample of result and output will be providing.

Chapter 5: Testing and analysis

On the testing and analysis part, it explains about the method use and procedure on how to test and analyze the experiment. After the visualizing technique was identified, we compare the result with the other techniques.

Chapter 6: Conclusion

This chapter combining the entire chapter in a final documentation and state the contribution that able to provide for future works.

1.8 Summary

The increasing of malware variants in each day seems to be serious problem for all computer users. We should pay enough attention on this situation. Malware detection is one of the actions that can be taken. By knowing their behavior, we can easily know the type of malware based on their behavior. To get better understanding, presentation of malware behavior should be done visually.

Visualization in the form of matrix will be presented in this research. Related work about visualization technique of malware behavior will be explained in the next chapter of this research

.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

This chapter will discuss about the literature review regarding all the sub topics in the framework as shown in Figure 2.1.

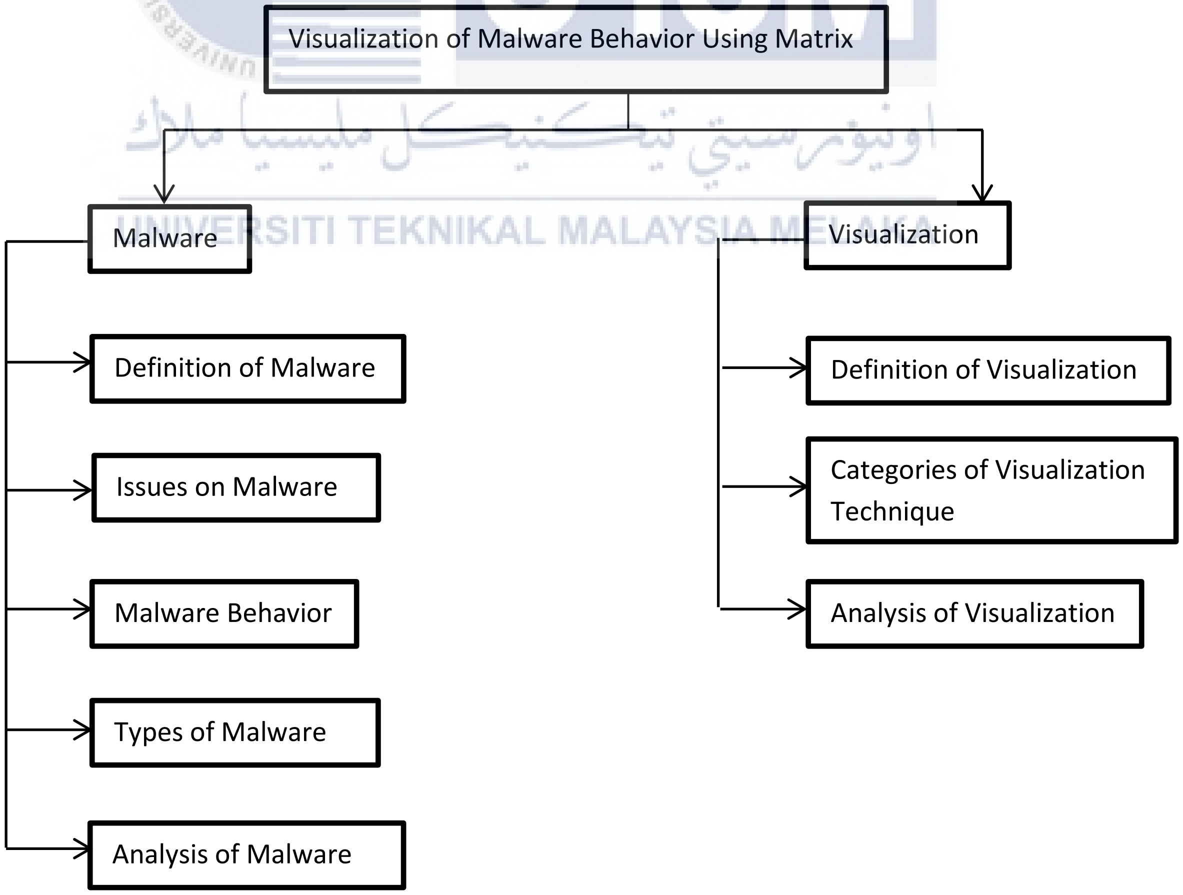
Visualization of Malware Behavior Using Matrix
Malware
Definition of Malware
Issues on Malware
Malware Behavior
Types of Malware
Analysis of Malware
Visualization
Definition of Visualization
Categories of Visualization Technique
Analysis of Visualization

Figure 2.1: Framework of Literature Review

Figure 2.1 shows the topics that will be elaborated and analyzed in this chapter. Two main topics are defined namely malware and visualization.

2.2 Malware

In this section, the definition, type, and issues related malware behavior are elaborated and analyzed.

2.2.1 Definition of Malware

There are millions of new malware was developed each year. Many researchers defined malware with different words. There are several definitions of malware defined by different authors was shown in Table 2.1.

Table 2.1: Definition of Malware

Author	Definition
Rutkowska, 2006	A piece of code which changes the behavior of either the operating system kernel or some security sensitive applications, without a user consent that it is then impossible to detect those changes using a documented features of the operating system or the application
Kramer & Bradfield, 2010	A software that harmfully attacks other software where to harmfully attack can be observed to mean to cause the actual behavior to differ from the intended behavior
Moser, 2007	Software that deliberately fulfills the harmful intent of an attacker is commonly referred to as malicious software or malware
Science, 2010	Malware is short for malicious software that represents the category of programs designed to infiltrate a computer system without the owner's consent.
Grégio & Santos, 2011	A set of malicious applications or codes, such as worms, viruses, trojans and bots to attack system in order to disrupt them, steal sensitive, financial information or even to use them as a disguise in other attacks, with directed target or not
Makandar & Patrot, 2015	A computer virus this is also a name given to a group of malicious data to all types of malicious data like virus, worm, Trojan and so on
Sikorski & Honig, 2012	Malicious software, or malware, can be defined as any software that does something that causes harm to a user, computer, or network
Symantec Corp, 2012	A software designed to attack and disable, damage or disrupt computers, computer systems, or networks.

Table 2.1 shows several different definition of malware by different authors. They have different opinion about what actually malware is. Based on the definition, this project defines malware as software that contain malicious code that can causes bad effect to computer user, computer system or computer network. Malware have been developed in many different types and each type have different characteristic. The next section will discuss about several types of malware.

2.2.2 Issues on Malware

First viruses started to be created in the early 1970s, when ARPANET, the forerunner of the Internet, was the main and wider interconnection network available. They had the form of experimental self-replicating programs, initially ideated as jokes between colleagues in laboratories. The first virus to be executed outside the single computer or lab where it was created was written in 1981, and injected in a game on a floppy disk as a practical joke. Before computer networks became widespread, most viruses spread on removable media, particularly floppy disks (Tiziano Santoro, 2010).

The effect of malicious data affect the various computer networks, infrastructures, services, file sharing, online social networking, and Bluetooth wireless networks (Makandar & Patrot, 2015).. Malware has infected every corner of the Internet, and is now can affect the social networks and mobile devices too. In 2010 alone, 286 million different types of malware were responsible for more than 3 billion total attacks on computer users, staggering numbers that are just one simple measure of malware’s impact (Symantec Corp, 2012). This become worst as the rapid increased on the new malware development as shown in Figure 2.2.