Malware Analysis Detection

RAHUL[1], RAVINDRA SONI[2]
[1,2]*Dept. of Computer Science and Engineering, Poornima College of Engineering, Jaipur*

Abstract -- The term malware stands for malicious software. It is a program installed on a system without the knowledge of owner of the system. It is basically installed by the third party with the intention to steal some private data from the system.One of the major and serious threats on the Internet today is malicious software, often referred to as a malware. The malwares being designed by attackers are polymorphic and metamorphic which have the ability to change their code as they propagate. Moreover, the diversity and volume of their variants severely undermine the effectiveness of traditional defenses which typically use signature based techniques and are unable to detect the previously unknown malicious executable. The variants of malware families share typical behavioral patterns reflecting their origin and purpose. The behavioral patterns obtained either statically or dynamically can be exploited to detect and classify unknown malwares into their known families using machine learning techniques. This survey paper provides an overview of techniques for analyzing and classifying the malwares.

Indexed Terms -- Malware, Malware Analysis, Static Malware Analysis, Dynamic Malware Analysis, exploited etc.

I. INTRODUCTION

Now a day, web becomes an important a part of the way of life of the many folks. On web several services square measure accessible and are increasing day by day.

A lot of and a lot of folk's square measure creating use of those services. On-line banking or advertising square measure the samples of the business services of the web.

Even as within the physical world, there square measure folks on the web with malevolent intents by taking advantage of legitimate users whenever cash is concerned. Malware like code of malicious intent helps these folks accomplishing their goals. McAfee catalogs over one hundred,000 new malware samples on a daily basis suggests that regarding sixty nine new threats each minute or regarding one threat per second. With the rise in promptly accessible and complicated tools, the new generation cyber threats/attacks have become a lot of targeted, persistent and unknown. Attackers exploit vulnerabilities in internet services, browsers and operative systems, or use social engineering techniques to create users run the malicious code so as to unfold malwares.

Malware authors use obfuscation techniques like dead code The term “malware” here is being employed because the generic name for the category of code that's malicious, together with viruses, Trojans, worms, and spyware. Malware authors use generators, incorporate libraries, and borrow code from others—there exists a strong network for exchange, and a few malware authors take time to scan and perceive previous approaches by (Arief & Besnard, 2003.) (Fred Cohen's) original definition of a computer program as of 1983 was: "a program that may 'infect' alternative programs by modifying them to incorporate a presumably evolved copy of itself." He updated this definition a year later in 1984 in his paper entitled: "Computer Viruses – Theories and Experiments”.

According to BBC News on-line, 2004 malware could be a general term for a chunk of code inserted into associate degree data system to cause hurt to it.

II. MALWARE DETECTION TECHNIQUE

a) Static analysis detection technique:

Software while not death penalty it's known as static analysis. Static analysis techniques will be applied on totally different representations of a program. Static analysis tools can even be used on the binary illustration of a program. Once assembling the ASCII text file of a program into a binary practicable, some data gets lost. This loss of knowledge any complicates the task of analyzing the code.

The process of inspecting a given binary while not death penalty it's largely conducted manually. as an example, if the ASCII text file is accessible many fascinating data, like information structures and used functions will be extracted. This data gets lost once the ASCII text file has been compiled into a binary practicable and so impedes any analysis. There square measure totally different techniques used for static malware analysis.

b) Signature based detection technique:

This technique is additionally referred to as pattern matching or string or mask or process technique. A signature could be a little bit of sequence injected within the computer program by malware writers that unambiguously identifies a specific malware. To observe a malware within the code, the malware detector seek for a antecedently given signature within the code.

c) File based heuristic analysis:

Also referred to as file analysis. During this technique, the file is analyzed deeply just like the contents, purpose, destination, operating of file. If the file contains commands to delete or hurt alternative file, than it's thought-about as malicious.

d) Weight based heuristic analysis:

It is the lot of ancient technique. Every application is weighted in keeping with the danger it's going to possess. If the weighted price exceeds the predefined threshold price, then the appliance contains malicious code.

e) Advantage of Static Analysis:

Static analysis is quick and safe; additionally it gathers the structure of code of program below scrutiny. If static analysis will calculate the malicious behavior within the application then this data will then be used for future security mechanism.

f) Dynamic-Analysis detection technique:

A given malware sample will be dead among a controlled setting and observation its actions so as to investigate the malicious behavior that is named dynamic malware analysis. Since Dynamic Malware Analysis is performed throughout runtime and malware unpacks itself, dynamic malware analysis evades the restrictions of static.

g) Advantage of Dynamic Analysis:

One will simply observe the unknown malware by merely analyzing the behavior of the appliance.

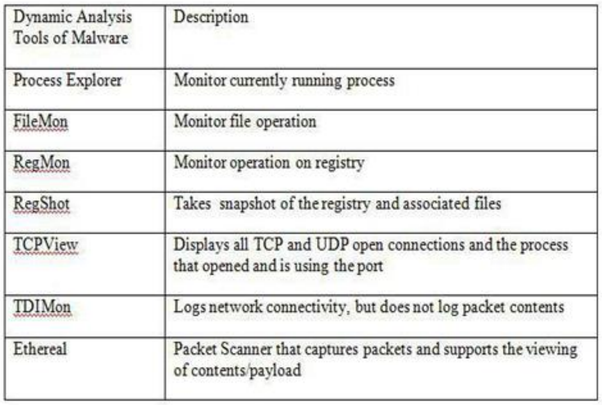
h) Hybrid analysis detection technique:

This technique is that the combination of each static analysis and dynamic analysis [6]. The procedure it follows it that it 1st checks for any malware signature.

i) Trends in Malware:

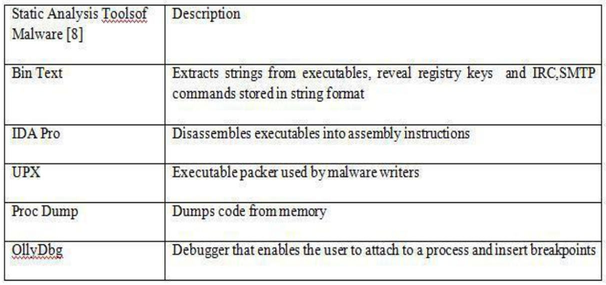
Malware is growing increasingly sophisticated. Malware authors seek to make their tools undetectable. Virtually every known offensive technique has been incorporated into malware to make it more difficult to defend against. Malware authors often seek to deliver several components in a single malware payload. Such additional components can include kernel level drivers designed to hide the presence of the malware, and malware client and server components to provide proxy services through an infected computer. One technique for embedding these additional components within Windows malware is to make use of the resource sections within Windows binaries. Malware may choose to create its own installation directory deep within the install program’s hierarchy in an attempt to hide from curious users. Various techniques also exist to prevent installed antivirus programs from detecting a newly infected computer. A crude yet effective method is to modify a system’s host’s file to add entries for hosts known to be associated with antivirus.

III. SOME STATIC ANALYSIS AND DYNAMIC ANALYSIS TOOLS



Dynamic Analysis Tools of Malware	Description
Process Explorer	Monitor currently running process
FileMon	Monitor file operation
RegMon	Monitor operation on registry
RegShot	Takes snapshot of the registry and associated files
TCPView	Displays all TCP and UDP open connections and the process that opened and is using the port
TDIMon	Logs network connectivity, but does not log packet contents
Ethereal	Packet Scanner that captures packets and supports the viewing of contents/payload

Fig. 1: - Dynamic Analysis Tools



Static Analysis Toolsof Malware [8]	Description
Bin Text	Extracts strings from executables, reveal registry keys and IRC,SMTP commands stored in string format
IDA Pro	Disassembles executables into assembly instructions
UPX	Executable packer used by malware writers
Proc Dump	Dumps code from memory
OllyDbg	Debugger that enables the user to attach to a process and insert breakpoints

Fig. 2: - Static Analysis Tools

IV. LIMITATIONS OF MALWARE

a) Limitation of Static Malware Analysis:

Generally, the source code of malware samples is not readily available. That reduces the applicable static analysis techniques for malware analysis to those that retrieve the information from the binary representation of the malware. Consider, for example, that most malware attacks hosts executing instructions in the IA32 instruction set. The disassembly of such programs might result in ambiguous results if the binary employs self-modifying code techniques.

b) Limitation of Dynamic Malware Analysis:

Static analysis tools are complex. To function properly they need to have a semantic understanding of the code, its dependencies, configuration files, and many other moving pieces that may not be written in same programming language. They must do this while effectively juggling speed with accuracy and reducing the number of false positives to be usable. Their effectiveness is greatly challenged by dynamically-typed languages like JavaScript and Python where simply inspecting an object at compile time may not be able to reveal its class/type.

V. METHODOLOGY OF MALWARE ANALYSIS DETECTION

Based on suggestions given by various authors for writing a literature review paper, the following steps were adopted to search the sources for the review. The initial stage for the review covered doctoral thesis from various international universities because they have been reviewed at higher exams. Online library has been used as a source for all doctoral and master theses. Keywords such as “malware”, “malware analysis” have been used to identify relevant thesis for the study. For the second stage leading journals and international conference papers were selected as these have gone through scientific peer reviews in order to be accepted at journals or conference proceedings.

VI. CONCLUSION

Rhetorical associate degree analysis of evidences and residual traces of crimes is an ancient, tried and productive field within the realm of investigation. the newest inclusion in crime is with the appearance of computers, communication and networking. The trend of growth in malware attack is increasing a lot of rapidly. Networks became a lot of vulnerable and square measure below constant malware attacks. From a lone system (PC) to a complete organization network, nobody is inevitable from the present sabotage. Monetary siphoning of bank a/c, stalking, calumny, duping square measure some samples of malware attack (cybercrime). because of this reality, rhetorical digital analysis of such crimes has gained vast importance in investigation recently.

REFERENCES

[1] Arun Lakhotia, Aditya Kapoor, Eric Uday , "Are Metamorphic Viruses Really Invincible ? Part 2", Virus Bulletin, January 2005.

[2] Robin Sharp, An Introduction to Malware, Spring 2012. Retrieved on April, 10, 2013 http://orbit.dtu.dk/fedora/objects/orbit:82364/datastreams/file_4918204/content

[3] A. H. Sung, J. Xu, P. Chavez and S.Mukkamala: Static Analyzer of Vicious Executables (SAVE), Proceedings of the 20th Annual Computer Security Applications Conference (ACSAC"04), IEEE.J.Rabek, R.Khazan, S.Lewandowski and R.Cunningham.

Detection of injected, dynamically generated, and obfuscated malicious code.

In Proceedings of the 2003 ACM Workshop on Rapid Malcode, pages 76–82, 2003.

[4] G. McGraw and G. Morrisett. Attacking malicious code: A report to the infosec research council. IEEE Software, 17(5):33–44, 2000.