A Review on Malware Detection and Analyzing Techniques

[1]Rajender Singh,[2]Rajendra Kachhwaha,[3]Arjun Choudhary
[1,3]Sardar Patel Police University of Police Security and Criminal Justice, Jodhpur,
[2]Dept. of CSE, MBM Engg. College, Jai Narain Vyas University, Jodhpur, India.

Received: Feb. 11, 2018 Accepted: March 04, 2018

ABSTRACT *Malware is not defined in single word. It is collection of malicious code or instructions which spread through the connected system or internet. It's using for gain illegally economic benefits and to damage other computer or network system. Malware detection is an important role in the cyber security. At present some anti malware software are used to detect malware, these are signature-based methods who cannot provide accurate result of malware attacks. Many metamorphic and polymorphic techniques are used to conceal the behavior of malicious program. These are the serious challenges to global security threat. Presently various malware detection techniques are available such as Heuristic based, Signature based and behavior based techniques. Most of the anti virus vendor uses signature based detection techniques, who already have known and well documented data base of signature value. Obfuscation and polymorphism technique impede the primary stage detection.*

Keywords: *malicious code, obfuscation, polymorphism, clustering.*

1. Introduction

The malware is mix up of two word Malicious and software. Malware is a malicious code that propagates over the connected systems in network [1].This scenario is increasing day by day with advanced computing technology and communication network. Malware can be considered as the entity in which new features can be easily added to enhance its dark side effects in the form of various attacks. It is software who added any code, temper, or destroy from a software system and steal a confidential information. It's motivationto harm or subvert the applications of the system.To protect from these malware in the Internet, computer system vendor of anti-malware software heavily rely on the automatic analysis and anti-virus tool[2].Malware developers are use obfuscation technique to conceal their signature code. Therefore the traditional anti- virus cannot capable to detect it because mostly dependent on signature based detection.

Malware is major threats in respect to global security and big challenges on the internet. The millions of website and computer currently infected with malware. The malware characterized according to their function as replication, propagation, obfuscation technique and corruption the system[5].

2. Malware techniques

Obfuscation techniques are mostly usedby attacker in now days. In this technique the malware code is conceal from anti-virus , firewall and IDS/IPS. These techniques change the in the program and add malicious code while not change in original code of program. These code harder to analyzed and perceive. These program run on system the code remain the same but instruction swap.

Polymorphism - If a program to seem completely different every times it replicated, however keeping the initial code intact, these are the polymorphic malware. A polymorphic malware consists of encrypted malicious code at the side of the decipherment module. Polymorphic code may be a methodology currently usually enforced in malware that uses a polymorphic generator to change the code whereas keeping the initial formula intact[4]. A typical implementation of a polymorphic code is to code malware and embrace the encryption/decryption inside the code. Polymorphic malwares have specially designed mutation engines.

Metamorphism - The metamorphic malware is capable of adjusting itself to a totally new in each instance that doesn't have something common to its original . This behavior makes it the foremost difficult malware to analysis. It capable to mutate while spreading across the network[7].

3. Malwareclassification

Malware classified according to their nature, workingand propagation.

Network based – Spyware –spyware could be a reasonably malware that's put in in secret on a user pc for the aim of aggregation info regarding users while not their data.

Adware- conjointly known as advertising-supported package whose practicality is to displays or downloads the advertisements to a laptop once the installation of malicious package or application.

Botnet- A botnet is remotely controlled autonomous code. it's sometimes a zombie

program that is controlled for any network infrastructure.

Trojan horse –First time it seems a genuine or useful software but in real it is a malware that corrupt the system and steals the data.

Snifer - Sniffers are a unit laptop programs that may intercept and record traffic over a network. Snifer capture every packet and convert in their original form[5].

Ordinary based Malware - virus – It is a harmful program which replicate itself and attached with application program. It's not requirea internet connection to propagate.

Worm - it is a package code that has the flexibility of self-replicating on victim pc. Worms area unit independent; they don't want for a number program to begin lifecycle.

Logic bomb - It could be a package program that remains inactivewhen a desire situation is met. The foremost common substance for a slag code could be a date and time. The slag code checks and updated information for should be activated[5].

4. Malware detectionTechnique

Malware detection technique are use for identify the malware and repair or remove it. The best way to analysis and monitor to malware in virtual environment e.g. sandbox. The malware detection techniques are classifiedin some categories,ordinary-based detection and signature-based detection. Ordinary based detection techniques are work on the system before examine normal condition on the system and what changes occur after a program execute.

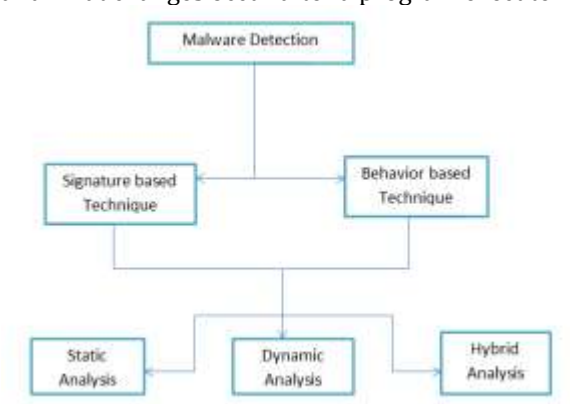

Fig.1 Malware detection technique

A. **Signature based malware detection -** Most of anti-malware based on signature based malware detection technique. These signatures are created by examining the disassembled code of malware binary. Numerous disassemblers and debuggers are offered that facilitate in disassembling the moveable executables. Disassembled code is analyzed and features are extracted. These features are employed in constructing the signature of specific malware family. The commercial antivirus looks the signature whois sequence of byte in within the malware code.

B. **Behavior based detection technique -** These technique analysis the behavior of known and unknown malware. These technique include various parameter such that source address of malware, type of attachment and read, write operation. These techniques are also classified in three categories[6].

Some Challenges and Difficulties in analyzing malware:-[2]

- Large volume
- Obfuscation
- False Positives
- Detection speed
- API calls

5. Survey on Related Work

"A survey on malware detection using data mining techniques" Yanfang ye et al.2017 survey on malware detection purpose a intelligent malware detection method. It divides in two steps feature extraction and classification/clustering. The performance depend on feature extraction and clustering. These are critically stage of further malware analysis. These paper provide comprehensive investigation on feature extraction and classification/ clustering[4].

"Malware Detection Using Machine Learning" [8]DragosGavrilutused a lot off perceptron algorithms.These paper propose a versatile framework which employ various machine algorithm and differentiate malware file and clean file basically aim to minimize the false positives. These paper approaches is one side cascade perceptron and second is generalized perceptron. The idea behind in this approach scaling up process to enable work a large no. datasets of malware infected and clean file. For using various algorithms, he obtain the accuracy of 69.90%-96.18%.

"A Static Malware Detection System Using Data Mining Methods" [Baldangombo et al. 2013] First of all they extract the feature based on API Function, PE headers and DLLs. These methods based on J 48 Decision Trees, Naive Bayes, and SVM. In these paper two major techniques are use such as Signature Based Detection and Heuristic Based Detection. These techniques are applicable well in respect to known malware.

"Malware Detection Module using Machine Learning Algorithms to Assist in Centralized Security in Enterprise Networks," These paper define the malware is executable or system library files these are the form of viruses, worms, Trojans, and these are design to breach information and compromising with system. (Singhal and Raul, International Journal of Network Security & Its Applications (IJNSA), Vol.4, 2012)

"Breach detection system testing methodology" In this paper advance attacker how can bypass the security layers and create unknown malware. Researcher use combine approach where the one side setup a virtual server and another side is real .In these paper testing on in the wild threats and zero day threats.[Z Balazs, S Miladinov, C Pickard, IEEE 2014]

"Antimalware Software: Do we Measure Resilience? "[9]this paper describes resilience of a antimalware.This paper describe the various examined concept of resilience. It applicable to cyber network. The development of interchanging set of metrics that adequately measure resilience. In this paper examined current tests of antimalware tool these tests follow resilience metrics guidelines. [RichardFord, Marco Carvalho, Liam Mayron, Matt Bishop,IEEE, 2013]

"Malware behavioral analysis analysissystem ; TWMAN"[10]These are define an analysis process real operating system. Malware investigator are use virtual environment but some malware compromises with virtual machine. They can not provide a perfect orreliable environment. These problems are facing at present so new tool developed Taiwan Malware Analysis Net (TWMAN). These tool use for malware behavior analysis. There are two sandbox are use in which on is VM based and another is real operating system based. It perform on 4840 type malware.

"Classification of Malware Based on String and Function Feature Selection" *These paper describe a new method automated detecting and classification. In this approach it's use a pattern recognition algorithm. In this process combine static features with printable string information. It's result give a classification result. It works on 1400 unpacked malware and give a 98% classification accuracy[7].*

"SubVirt: Implementing malware with virtual machines" In these paper focus on developing virtual machine based rootkit. These are new type of malware who run on virtual machine. It's working flow subvert Windows or Linux as target system use of VMBR. In this paper define how to avoid it and what strategy use to defend a system from these serious threat. It also provide a lose point of a system and possible attacks from these malicious software[6].

"A Survey on Techniques in Detection and Analyzing Malware Executables" These paper focus on recent trend of malware, their classification on based their working .A general study obtain most of the malware comes from the using internet including downloading and surfing. Malwares are differentiates according to their payload, enabling vulnerability & propagation mechanism. Focus on different variants of malware who already exist in cyber space.

6. Conclusion

Malware is big and challenges in cyber security field it faces big threat for system and network security. The major role of malware to steal personal and private information, corrupting or disabling our security system.In this survey paper malware detection techniques have been described. The issues with traditional signature based detection are also highlighted. This paper explains about static, dynamic and hybrid analysis. These provide us about various malware techniques and code obfuscation technology.

7. References

[1.]https://content.iospress.com/articles/journal-of-computer-security/jcs410 (Automatic analysis of malware behavior using machine learning), downloaded on 25 Oct. 2017

[2.]http://ieeexplore.ieee.org/document/7346730/(Behavior analysis of malware using machine learning) accessed on 25 Oct. 2017

[3.] http://ieeexplore.ieee.org/document/7841031/ (Towards an effective and efficient malware detection system) last accessed 11 sept.2017

[4.]https://dl.acm.org/citation.cfm?doid=3101309.3073559 (A Survey on Malware Detection Using Data Mining Techniques) last accessed 11 sept. 2017

[5.]AComparison of Malware Detection Techniques Based on Hidden Markov Model

[6.]Vinod P. V.Laxmi,M.S.Gaur: Survey on Malware Detection Methods, 3rd Hackers‟

[7.]KirtiMathur,A Survey on Techniques in Detection and Analyzing Malware Executables,IJARCS,2013

[8.]"Malware Detection Using Machine Learning" DragosGavrilut et al. 2013

[9.] Antimalware Software: Do we Measure Resilience. [Richard Ford, Marco Carvalho, Liam Mayron, Matt Bishop,IEEE, 2013]

[10.]"Malware behavioral analysis analysis system; TWMAN"[Hsien-De Huang et al. Twain govt. journal]