EMPLOYABILIY OF BIG DATA AND HEURISTIC BASED TOOLS IN DETECTING AND MITIGATING MALWARE

Gautam Anand

Student, Sanskriti School, Chanakyapuri

ABSTRACT

Malware is not defined in a single word. It is a collection of malicious code or instructions which spread through the connected system or the Internet. It's used illegally for gaining economic benefits and damaging other computers or network systems. Malware detection is an essential role in cybersecurity. At present, some antimalware softwares are used to detect Malware; these are signature-based methods which cannot provide an accurate result of malware attacks. Many Metamorphic and polymorphic techniques are used to conceal the Behaviour of malicious program. These are the severe challenges to a global security threat. Presently various malware detection techniques are available such as Heuristic-based, Signature-based and Behaviour-based techniques. Most of the antivirus vendors use signature-based detection techniques, which already have known and well-documented database of the signature value. Obfuscation and polymorphism techniques impede the first stage detection.

1. INTRODUCTION

The Malware is mix up of two words Malicious and Software. Malware is a destructive code that spreads over the associated frameworks in the system [1]. This situation is expanding step by step with cutting edge. Malware can be considered as the element in which new highlights can be effectively added to improve its clouded reactions to different assaults. It is software which adds any code, tampers or destroys software system and steals confidential information. Its motive is to harm or subvert the applications of the system. To protect from this Malware on the Internet, computer system vendor of antimalware software heavily relies on the automatic analysis and antivirus tools [2]. Malware developers are using obfuscation technique to conceal their signature code. Therefore, traditional antivirus cannot be capable of detecting it as it is mostly dependent on Signature-based detection. Malware is a major threat with respect to global security and significant challenge on the Internet. The millions of websites and computers are currently infected with Malware. The Malwares are characterized according to their function as replication, propagation, obfuscation technique and corruption the system [5].

2. MALWARE TECHNIQUES

Obfuscation techniques are mostly used by an attacker nowadays. In this technique, the malware code is concealed from antivirus, firewall and IDS/IPS. These techniques change the program and add malicious code while not original code shift of program. This code is harder to analyze and perceive. These program run on a system where the law remains the same but instructionswap. Polymorphism - If a program seems completely different every time it replicated, however keeping the initial code intact, these are the polymorphic Malwares. A polymorphic malware consists of encrypted malicious code at the side of the decipherment module. Polymorphic code maybe a methodology currently usually enforced in Malware that uses a polymorphic generator to change the system, whereas keeping the original formula intact [4]. A typical implementation of a polymorphic code is to code Malware and embrace the encryption/decryption inside the system. Polymorphic malwares have specially designed mutation engines. Metamorphism - The metamorphic Malware is capable of adjusting itself to a totally new instance that doesn't have something familiar to its original. This Behaviour makes it the foremost severe Malware to analyse. It was capable of mutating while spreading across the network [7].

3. MALWARE CLASSIFICATION

Malwares are classified according to their nature, and propagation. Network-based – Spyware –spyware could be malware that's put in secretly on user's pc for the aim of aggregation of information regarding users while not their data. Adware- conjointly known as an advertising-supported package whose practicality is to display or download the advertisements to a laptop once the installation of malicious package or application.

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

Sniffer - Sniffers are unit laptop programs that may intercept and record traffic over a network. Sniffer captures every packet and converts to their original form [5].

Ordinary based Malware - virus – It is a harmful program which replicates itself and attached with the application program. It does not require an 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 inactive when a desired situation is met. The first common substance for a slag code could be a date and time. The slag code checks and updated information should be activated [5].

4. MALWARE DETECTION TECHNIQUE

The malware detection technique is used to identify the Malware and repair or remove it. The best way to analyse and monitor to Malware is in a virtual environment, e.g. sandbox. The malware detection techniques are classified in some categories, ordinary-based detection and signature-based detection. Ordinary based detection techniques work on the system before examining reasonable condition on the system and what changes occur after program execution.

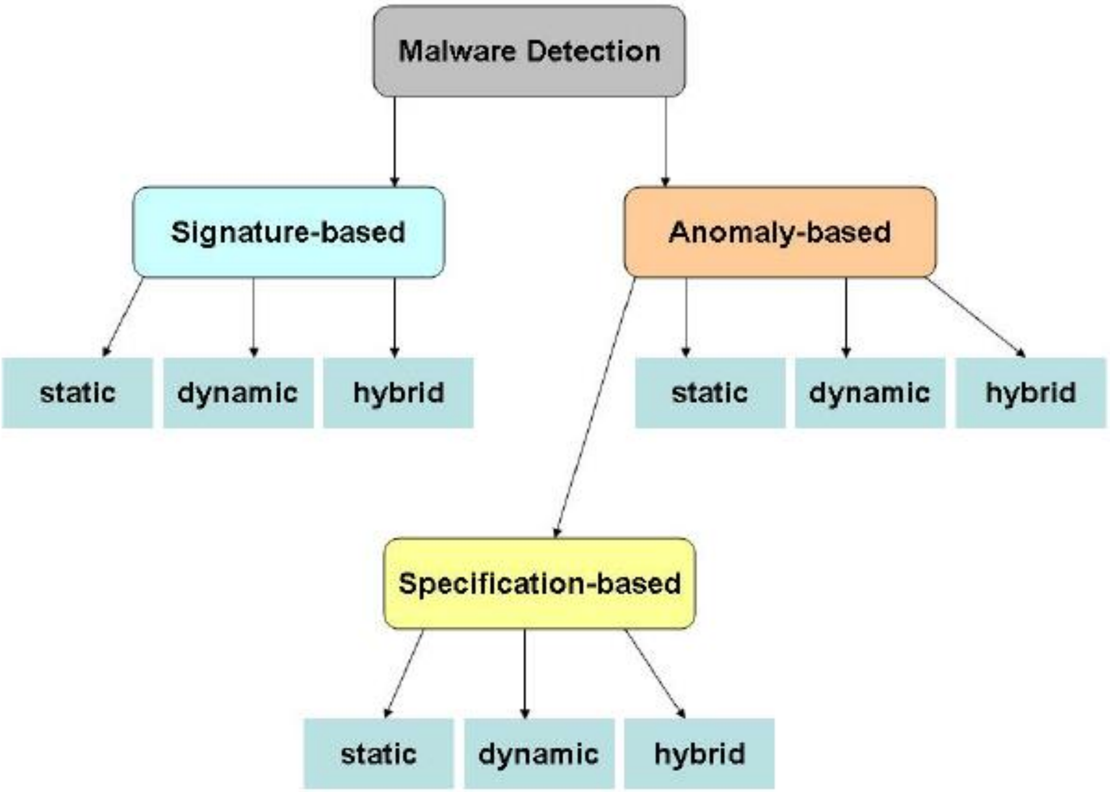

Fig 1: Malware detection technique

A. Detection on the basis of Signature–

Most of antimalware are based on Signature-based malware detection technique. These marks are made by looking at the dismantled code of Malware paired. Various disassemblers and debuggers are offered that encourage disassembling the moveable executables. Dismantled law is examined, and highlights are extricated. These highlights are utilized in developing the Signature of a specific malware family. The commercial antivirus looks the Signature which is the sequence of a byte in within the malware code.

B. Detection on the basis of Behaviour-

This technique analyses the Behaviour of known and unknown Malwares .These techniques include various parameters such that source address of Malware, type of attachment and read, write operation. These techniques are also classified into 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 is an intelligent malware detection method. It divides into two steps feature extraction and classification/clustering. The performance depends on feature extraction and clustering. This critical stage of further malware analysis. These papers provide a comprehensive investigation on feature extraction and classification/ clustering [4].“Malware Detection Using Machine Learning”[8]DragosGavrilut used a lot off perceptron algorithms. These papers r propose a versatile framework which employs various machine algorithms and differentiate malware file and clean file basically with aim to minimize the false positives. These paper approaches are one side cascade perceptron and second is generalized perceptron. The idea behind this approach is scaling-up process to enable work on a large no. Datasets of Malware infected and clean file. For using various algorithms, he obtains 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 APIFunction, PE headers and DLLs. These methods based on J 48 Decision Trees, Naive Bayes, and SVM. In this paper, two primary techniques are used, such as Signature-Based Detection and HeuristicBased Detection. These techniques are applicable well with respect to known Malware.

“Malware Detection Module using MachineLearning Algorithms to Assist in Centralized Security in Enterprise Networks,” These papers define the Malware is executable or system library files these are the form of viruses, worms,Trojans, and these are designed to breach information and compromising with the 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 this paper testing on the wild threats and zero-day threats. [Z Balazs, S Miladinov, C Pickard,IEEE 2014]

“Antimalware Software: Do we MeasureResilience? ”[9]this paper describes the Resilience of an antimalware. This paper describes the various examined concepts of Resilience. It was applicable to cyber network. The development of interchanging set of metrics that adequately measure Resilience. In this paper examined current tests of the antimalware tool, these tests follow resilience metrics guidelines. [RichardFord, Marco Carvalho,Liam Mayron, Matt Bishop, IEEE, 2013]“Malware behavioural analysis system;TWMAN"[10]These define an analysis process real operating system. Malware investigator is using a virtual environment, but some malware compromises with the virtual machine. They cannot provide a perfect or reliable environment. These problems are being faced at present, so a new tool developed Taiwan Malware Analysis Net(TWMAN). These tools are used for malware behaviour analysis. There are two sandboxes used in which one is VM based, and another is a real operating system found. It performs on 4840 type malware. “Classification of Malware Based on String andFunction Feature Selection" This paper describes a new method automated detecting and classification. In this approach, it uses a pattern reognition algorithm. In this process, combining static features with printable string information. Its result gives a classification result. It works on1400 unpacked malware and provides a 98%classification accuracy [7].“SubVirt: Implementing malware with virtual machines In this paper focus is on developing a virtual machine-based root kit. These are a new type of Malware which run on a virtual machine. It’s working flow subvert Windows or Linux as target system use of VMBR. In this paper it’s is defined how to avoid it and what strategy is used to defend a system from this severe threat. It also provides a loose point of a system and possible attacks from this malicious software [6].“A Survey on Techniques in Detection and Analysing Malware Executables” These papers focus on the recent trend of Malware, their classification based on their working. A general study obtains most of the Malware comes from using Internet, including downloading and surfing. are differentiated according to their payload, enabling vulnerability & propagation mechanism. Focus is on different variants of Malware which already exist in the cyberspace.

6. CONCLUSION

Malware is significant a challenge in cyber security field; a big threat for system and network security. The primary role of Malware is 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 with various malware techniques and code obfuscation technology.