International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(REVIEW ARTICLE)

Check for updates

Advancements in automated malware analysis: evaluating the efficacy of open-source tools in detecting and mitigating emerging malware threats to US businesses

John Oluwafemi Ogun *

Department of Information Systems & Business Analytics, Hankamer School of Business, Baylor University, Waco, TX

International Journal of Science and Research Archive, 2024, 12(02), 1958–1964

Publication history: Received on 04 July 2024; revised on 13 August 2024; accepted on 15 August 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.12.2.1488

Abstract

Malware, short for malicious software, represents a significant and evolving threat to computer systems, targeting individuals, corporations, and governments globally. This paper explores the multifaceted nature of malware, which includes viruses, worms, Trojans, and more, and delves into how they compromise systems by disrupting services, stealing sensitive data, and denying access. Modern malware is increasingly sophisticated, evading traditional detection methods and posing challenges to cybersecurity professionals. This review outlines key methodologies in malware analysis, including MARE (Malware Analysis Reverse Engineering) and SAMA (Systematic Approach to Malware Analysis), which offer systematic frameworks for understanding and mitigating malware threats. Additionally, the paper highlights the challenges of malware analysis, such as the complexity of advanced malware variants and the limitations of current detection techniques. By examining the types of malwares, from ransomware to keyloggers, and discussing the signs of an attack, the paper underscores the importance of ongoing research and the development of more robust analytical tools. The insights provided aim to enhance the preparedness of IT professionals in combating emerging threats, emphasizing the necessity of a comprehensive understanding of malware behavior for effective defense strategies.

Keywords: Malware; Cybersecurity; Ransomware; Dynamic Analysis; MARE (Malware Analysis Reverse Engineering); SAMA (Systematic Approach to Malware Analysis)

1. Introduction

"The term "malware" is short for malicious software, and it is any type of software designed to do unwanted and malicious actions on a computer system. Examples of malware include viruses, worms, Trojans, logical bombs, rootkits, and spyware. This malware exists in a variety of forms, from custom designed to attack a specific system to generic self-replication probes that attack every available target" (Wangen, 2015).

There are more applications on the internet that posse's malware. Malicious software poses a serious security threat to computer systems. In the last decade, malicious software has become a dangerous tool employed by threatening actors to mount cyberattacks on private companies, government agencies, and individuals. Ransomware, for example, is increasingly used to attack major US companies, organizations, and individuals (Maglaras et al., 2021). Malware usage is on the increase, and we can group malware into distinct categories based on their behaviors. Malware comes in various forms, it can be an executable binary, a script, or codes with a malicious aim. Malware intends to gain access to a system, disrupt system services, steal sensitive information, destroy resources, and deny services. The idea that malware is only transmitted into a system through counterfeit or cracked software disguised as genuine software is much farther from the truth. Genuine software or programs can have malware embedded in them. They serve as a wrapper for the malware and once run on the system, the malware is also run along with the software (Ijaz et al., 2019). In the process of downloading genuine software from a website, malicious software may be downloaded alongside.

* Corresponding author: John O Ogun

Malware apart from being executable codes can also function as a malicious content downloader e.g., PDF and PHP links gaining control of a system and acting as a catalyst to facilitate further execution of malicious software on the system.

The systems security analyst has put in place several measures to counter malware including numerous anti-malware defense measures, however, cybercriminals and malicious hackers are on the increase, especially with money, fame, and prestige (respect) as the end goal. This makes malicious hackers and cyber criminals up their game from the traditionally popular forms of malware to a more sophisticated one, capable of penetrating the anti-malware defense mechanism put in place to protect the system and network. Cybercriminals are well informed of the increasing attack surfaces, the ever-increasing innovative technology being introduced in this era of internet of things (IoT), and the time gap before security issues surrounding such technology can be perfected, and even at that, zero-day vulnerabilities still exist due to the unpredictable human nature. All they must do is play around with the security architecture of the new technology in search of vulnerabilities to exploit. In the US and other countries of the world, malicious hackers are famous for breaking into system networks of organizations, companies, corporations, and even government agencies, ransomware attacks are common in US, examples include the Colonial Pipeline ransomware attack of 2021, Baltimore ransomware attack in 2019 and WannaCry ransomware attack. using discrete infections, but no one is spared, not even the average web user. The integration of the IoTs into our technology has made it even more difficult for cybersecurity experts to prevent most of the malware that targets various application domains (Aryal et al., 2021). Devoting time and resources to understanding malware analysis will help to first detect the nature of the present malware and secondly, provide valuable insight into new and emerging ones. Gone are the days when malware analysis was manually performed by system security experts in a strenuous and time-consuming manner. Today, open-source malware analysis tools are used to carry out this operation automatically (Liu et al., 2022). By malware analysis, we mean the process put in place to understand the behavior and intent of a strange file, link, program, or software, and the feedback from the analysis helps detect and prevent what could be a potential threat. In cases where a cyber-attack has occurred or is in process, malware analysis is crucial to understand the nature of the attack and how best to recover from it. The pre-emptive deployment of malware analysis especially on emerging malware programs can provide security analysts more insight into ways to best protect and forestall them.

This document intends to provide information on the current state of malware analysis on information assets to IT professionals in the US and around the world. By highlighting pertinent subjects on malware analysis, this review offers various results and additions to the literature.

2. Literature review

Ijaz et al. 2019, analyzed malware based on static as well as dynamic features. “In static analysis, the executable file is analyzed on structure bases without executing it in a controlled environment. The executable file has many static attributes like different sections and memory compactness. Portable Executable PEFILE is a python library that extracts static features from executable files. In dynamic analysis, malware behavior is analyzed in a dynamic controlled environment. When the malware executes, it changes the registry key maliciously and takes the privileged mode of the operating system. In dynamic analysis, the software has full access to all the resources to execute in a controlled environment. In this environment, the software can change the registry keys of the computer and run in debugger mode. At the end of malware execution, the dynamic environment reverts to its previous snapshot, which is created at the start of the environment setup. Cuckoo Sandbox is a controlled environment that consists of three parts, one is the host, the other is a guest virtual machine, and the third one is an agent.” Cuckoo, which has a logging function was used for dynamic malware analysis.

Kolbitch et al., (2009) extracted dynamic system calls from the malware, and a controlled environment analysis was performed. Based on the system call, the exe file is categorized as malicious or benign. Dynamic analysis of the malware's activities is frequently obscured.

The executable code's structure was provided by David et el., (2017) who also identified six significant static properties. These six characteristics were used to categorize files as malicious or benign. “Compilation Time, File Info, Section Alignment, Size of Image, File Alignment, and Size of Header” are the six crucial features.

Chumachenko (2017) used machine learning to analyze and categorize malware using API calls and return codes. Nine separate malware families were utilized for analysis, accurately classifying each one. The Cuckoo Sandbox was used to extract millions of features from malware, and a few hours after the features were loaded in the system, the result was produced.

Ravula, et al., (2011) built a thorough framework to categorize and identify malware using machine learning and data mining approaches, protecting crucial data from dangerous attacks. By examining both anomaly-based and signature-based features, they suggested a comprehensive and efficient method to detect and classify malware.

In Ijaz et al., (2019) "Cuckoo sandbox was used for dynamic analysis of malware and extracts their behavior at run time during execution. The aim was to isolate the actual system from the testing environment and extract desired information from malware execution. The extracted features include Summary information, Files, API call during execution, registry keys, IP addresses, and DNS queries."

3. Malware analysis methodologies

3.1. MARE (Malware Analysis Reverse Engineering)

MARE resides between the Detection and Code Analysis and Reverse Engineering phases of Malware Defense (M. D), MARE introduces the logical steps taken in each process to help analysts produce an output that is repeatable, objective, and applicable, with the intent of understanding the analysis. The detection phase is the initial stage experienced following the infection. Malware scanners, such as Virus Total, are used to examine the questioned malware during this stage. This is done to determine whether the malware poses an established threat or an unknown one. The signature created by the malware's cryptographic hash is used as the basis for the verification (Kiachidis and Baltatzis, 2021).

The detection phase is followed by the isolation and extraction phases. This stage aims to extract, isolate, and safeguard the malware. This is due to the requirement for transferring malware securely to an environment for behavioral analysis. Additionally, during this step, analysts attempt to understand the nature of the malware and determine whether it is a rootkit. The extraction process for rootkits is different. The phase of behavioral analysis follows. Analysts attempt to notice and record the modifications that the malware's functionality and execution make to the system during this phase. This is conducted because the malware's alterations can be used to deduce its destructive intent. File manipulation, registry hacking, library modification, connections being made, etc. are some of the modifications that analysts need to consider. Before and following the malware's execution, analysts capture screenshots. The variations between these snapshots show the adjustments done because of virus functionality. For this step, automated dynamic analysis is also helpful. It aids in expediting the entire process, and depending on the outcomes, analysts may go on to personally inspect the virus (e.g., when the malware seeks user interaction). Depending on the results of the Code Analysis and Reverse Engineering phase after this one, it could be necessary to repeat this step (Kiachidis and Baltatzis, 2021).

The final process is code analysis and reverse engineering. The use of assembly language is crucial for this stage. Disassembling and debugging is key to it. Understanding the way the malware reacts in relation to the code is the core objective. Typically, it begins with the recognition of strings and goes on from there. When performed by an experienced analyst, this phase reveals a lot.

3.2. SAMA (Systematic Approach to Malware Analysis)

Based on the dynamic nature of malware, the MARE strategy falls short of the required standard to address the challenges that emanate due to the complexity of the techniques initiated. As such, the SAMA strategy complements the MARE strategy to curb these emerging challenges. The SAMA strategy retained the four (4) different phases discussed by Kiachidis and Baltatzis, (2021) and renamed them:

- Initial Actions.
- Classification.
- Static and Dynamic Code Analysis.
- Behavioral Analysis.

The goal is to provide a framework that can analyze contemporary, sophisticated malware. This is required due to the systematic structure of the iterative process which must be a reliable and inflexible technique capable of helping analysts learn the necessary information from a particular malware (Higuera, 2020).

The analysts should be conversant with the operations of the system used to conduct the analysis method during the Initial Actions phase of SAMA. Systems come in both virtual and physical forms and in either scenario, snapshots must be taken to give the analyst the option of going back to a clean condition. Additionally, the required hash values must be

generated and stored to preserve the integrity of the snapshots. The analyst has not yet started analyzing the relevant sample itself (Or-Meir et al., 2019).

The step of classification follows. During this phase, fundamental static analysis procedures create a sufficient technique, which is necessary for the subsequent phases. In this phase, the sample is not analyzed, and the primary goal is to confirm whether they must continue the analysis process or whether there have been any concrete results to date or evidence of the sample's goodness and otherwise.

The next phase is code analysis, where advanced static and dynamic analysis techniques are used to accomplish the objective of examining the sample's code. The analysts will find this to be the most challenging and complicated step to complete. The conclusion of this phase is necessary since its findings offer a valuable understanding of how the sample functions and reveal obscure or hidden aspects that would otherwise go undetected for the highly sophisticated and complicated modern malware (Kiachidis and Baltatzis, 2021).

The phase of behavioral analysis marks the end of SAMA. As the name suggests, approaches for dynamic and memory analysis are used in this phase. Behavioral analysis requires a secure lab environment to run the analysis, and every change it makes to the system (registry changes, connections made, etc.) is logged and examined.

4. Challenges in malware analysis

Malware analysis is a tedious task. Obfuscation and evasion techniques, backed with diverse behavior patterns, new, and advanced malware variants, and time requirements for malware analysis even make it more tedious.

Singh and Singh, (2018) two approaches to malware analysis are signature-based (without executing the file) and behavior-based (running the file mostly in a controlled environment). Whitepapers from security experts have shown that signature-based detection techniques have become obsolete and cannot detect new malware variants, behavior-based analysis in which malware files are executed for capturing behavioral artifacts has proven resourceful. Though a possibility that complex obfuscated malware can cheat the execution environments like sandboxes, and debuggers by not executing actual behavior. Notwithstanding, behavior-based system detection is far better than signature-based malware detection systems. Behavior-based systems detection is, however, slow and time-consuming, which is a major concern.

Najmi et al., (2012) in their paper “challenges in high accuracy of malware detection” discussed how new malware is derived from previous malware, and with large similarities becomes the new variant. In addition, they stated the insertion of garbage to confuse analysts with fake API calls, encryption of the vital details within the malware body, and the use of “packing” make malware analysis difficult and time-consuming. “Packing is a method to compress a Windows executable without having the user manually decompress them. The purpose is legit since it is used by benign Windows executables as well, but it has already been exploited by malicious hackers for compressing the size plus encryptions.”

Ficco, (2020) made use of the hybrid approach for malware analysis which uses generic and specialized detectors. “In particular, the work presented different methods to optimally combine both generic and specialized detectors during the analysis process, which was used to increase the predictability of the detection strategy, as well as improve the detection rate in presence of unknown malware families and its provided better detection performance in the absence of a constant re-training of detector needed to cope with the evolution of malware”. The considered features are extracted by using the Cuckoo sandbox. “An alpha-count mechanism that explores how the length of the observation time window can affect the detection accuracy and speed of different combinations of detectors when the malware analysis was developed. An extended experimental campaign was conducted on both an open-source sandbox and an Android smartphone with different malware datasets.”

5. Types of malwares

- Wiper Malware: just as the name suggests, wiper malware is designed to remotely erase every piece of data in the hard disk of the intended victim system. This malware attack seeks to destroy data with the goal of destroying evidence, sabotage, or financial gain. Wiper malware has been used in the past (e.g., to attack Saudi Aramco and Qatar's RasGas oil companies in 2012 and the ‘Olympic Destroyer’ used to attack the Winter Olympics hosted by South Korea in 2018). Recently, this malware has been increasingly becoming popular. This can be credited to ongoing Ukraine and Russian wars. In this year alone, no fewer than six (6) forms of wiper malware attacks have been deployed on Ukraine institutions and organizations some of them are (Whisper-

Kill, Whisper-Gate, Hermetic-Wiper, Isaac-Wiper, Caddy-Wiper, Couple-Zero, and Acid-Rain). The approach commonly used to deploy this malware is to enumerate the filesystem and overwrite the selected files with data (Revay, 2022).

- Ransomware: This malicious software is gaining more popularity, especially in the cryptocurrency market where malicious hackers use encryption to lock out crypto exchange companies from their own network. This paralyzed the activities of the company until an agreed sum is paid to the hacker and in most cases releases the decryption key. In other cases, the cybercriminal gains access to the company's crypto wallets and transfers all or a significant sum of crypto coins (e.g., Bitcoin or Ethereum) into another private wallet. Since all transactions on the blockchain are visible and traceable, the cybercriminal forces the company to a ransom payoff before releasing the coins back to the company's wallets. Ransomware occurs in all industries whether private or public.
- IP Spoofing: In this type of malware, the intruder gains unauthorized access to the system by camouflaging the IP address. The intruder sends a message using an untrusted source and makes it look as if the message is from a trusted source or host (Gupta, 2013).
- Adware: Every computer user, especially online shoppers, has contended with this type of malware at one point or the other. Adware is a form of malware that tracks a computer user's online activity through algorithm and pattern matching to determine the type of ads to serve them. That is, it utilizes pattern matching to capture what the user views online and tailors related ads to the user based on that. Adware cannot be termed as malicious in nature, however, the constant sending of unsolicited ads and tracking of user records, locations, and even friends can be viewed as an invasion of user's privacy. Another concern with adware is the safety of all the information collected over time, some of which is sensitive and could be sold for monetary gain (Kim and Solomon, 2023).
- Browser Hijacker: though this term is used for several malicious programs, browser hijackers mostly describe software used to manipulate internet explorer settings. The attacker uses external code to manipulate this setting making the user unsuspecting of these changes. This malware is sometimes categorized as a form of social engineering (Gupta, 2013).
- Trojan: This malware operates in a unique way called appeasement. Appeasement is the process of attraction through enticement. Trojan malware works in this same way. It disguises itself as an attractive useful software or program with enticing features that are irresistible for users. As unsuspecting users download the software the malware is activated and immediately begins the process of taking over the user's computer for malicious intent.
- Worms: These are self-replicating, self-producing malicious software programs that replicate from one host machine to another. Though like a virus, and perhaps used interchangeably sometimes, a worm is in fact different from a virus. While most viruses self-replicate using the computer resource (host), a worm can self-replicate without the need for a host. A worm is a powerful piece of hardware in the hand of an attacker as it gradually corrupts the network and gets primed to launch several attacks such as DDoS, stealing confidential information, and ransomware. This attribute makes worms one of the most dangerous types of malicious software (Kim and Solomon, 2023).
- Keylogger: This program functions as the custodian of keystrokes which are then passed to a third party. As a keystroke custodian, once installed on a system, the keylogger monitors the user's activities. All keystrokes on the keyboard are recorded. In the hands of an attacker, this can be used to steal valuable information such as passwords, financial data, and much other confidential information.

6. Signs of malware attack

A malware attack can be mild or catastrophic depending on the plan and expertise of the attacker and the infrastructure being attacked. Overall, most malware attacks exhibit a trace of these signs before becoming a full-blown disaster. These include.

- Gradual decrease in the speed or efficiency of a system for days or weeks. This could be because the malware is spreading through the networks which hinders performance.
- Another sign of a malware attack is continuous freezing and, in some cases, crashing the system. There is the possibility for the user to mistake it for a bug (software or hardware problem), however, with malware analysis, this could be easily determined and fixed.
- A common way to detect malware attacks is the appearance of a new program, file, or icon on the computer which the user did not install. This can be a sign of a virus attack (a macro virus that can create a software program on the computer). This can also be a sign of a worm attack (malware capable of replicating itself in the system network).

- Other signs of malware can be the system not shutting down or powering up, an unsolicited program running, or a program in use shutting down by itself. The system heating up unnecessarily and a lot more.

To prevent system breakdowns that might affect critical business functions, a defensive anti-malware security system should be installed on all systems and networks. Also, there should be a routine malware analysis check.

7. Merits of malware analysis

Malware analysis tools have been the frontline instrument systems security analysts use to detect, dissect, and better understand the intent and purpose of any form of malware attack.

- In cases of malware attacks, malware analysis tools help in troubleshooting the attack and classify it according to the level of severity. The information provided will be what the security analyst will act on to repel the attack.
- Malware analysis armed the security analysts' necessary indicators to respond swiftly to an attack, these indicators could be, the nature of the attack, malware type(s), behavior, pattern, signature, and even the targeted system architecture.
- Malware analysis helps to improve the preparedness of security analysts to combat an attack. Its ability to identify the signatures of emerging malware makes it an indispensable tool for companies in the bid to fortify their system security architecture.

8. Conclusion

This document on malware analysis provides evolving insights into the different types of malwares, the analytical tools used to identify them, and the challenges in malware analysis. Understanding malware analysis, the various types of malwares, and the unique features attributed to each malware are important to building a secured system architecture capable of defending and protecting users' systems and networks from malware invasion. However, it is worth mentioning that this paper is not in itself exhaustive, rather it is meant to provide valuable background information on malware analysis which can lead to further research on the topic.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Aryal K., Gupta M. & Abdelsalam M. (2021). A survey on Adversarial Attacks for Malware Analysis. Computer and Information Sciences.

[2] Chumachenko, K. (2017). Machine Learning Methods for Malware Detection and Classification.

[3] David, B., Filiol, E., & Gallienne, K. (2017). Structural analysis of binary executable headers for malware detection optimization. Journal of Computer Virology and Hacking Techniques, 13(2), 87-93.

[4] Ficco, M. (2020). "Comparing API Call Sequence Algorithms for Malware Detection". 10.1007/978-3-030-44038-1_77.

[5] Gupta S. (2013) "Types of Malwares and its Analysis". International Journal of Scientific & Engineering Research Volume 4, Issue 1, http://www.ijser.org

[6] Higuera J. B., Aramburu C. A., Higuera J. R. B., Urban M. A. S. & Montalvo J. A. S. (2020). "Systematic approach to Malware analysis (SAMA)," Appl. Sci., doi: 10.3390/app10041360

[7] Ijaz M., Durad M. H. & Ismail M. (2019). Static and Dynamic Malware Analysis Using Machine Learning. FFProceedings of 2019 16th International Bhurban Conference on Applied Sciences & Technology (IBCAST).

[8] Ioannis G. Kiachidis and Dr. Dimitrios A. Baltatzis (2021). "Comparative Review of Malware Analysis Methodologies". International Journal of Network Security & Its Application (IJNSA), Vol. 13, No. 6.

[9] Kim D. and Solomon M., (2023) "Fundamentals of Information Systems Security". Jones & Bartlett Learning. 4th Edition.

[10] Kolbitsch, C., Comparetti, P. M., Kruegel, C., Kirda, E., Zhou, X. Y., & Wang, X. (2009). Effective and Efficient Malware Detection at the End Host. In USENIX security symposium (Vol. 4, No. 1, pp. 351-366).

[11] Liu Y., Li J, Liu B., Gao X. & Liu X. (2022). Malware detection method based on image analysis and generative adversarial networks. Concurrency and Computation: Practice and Experience, vol. 34, Issu. 22.

[12] Maglaras L., Kantzavelou I, and Ferrag M. (2021). "Cyber Security of Critical Infrastructures". (Available at: https://www.mdpi.com/journal/applsci/special_issues/Cyber Security Critical Infrastructures).

[13] Najmi, A., Mohd, M., and Anazida, Z., (2012) "Challenges in high accuracy of malware detection". EP – 125, DO - 10.1109/ICSGRC.2012.6287147.

[14] Or-meir O., Nissim N., Elovici, Y. & Rokach, L. (2019). Dynamic Malware Analysis in the Modern Era—A State of the Art Survey. ACM Computing Surveys, Vol. 52, No. 5.

[15] Rathnayaka C. & Jamdagni A. (2017). "An efficient approach for advanced malware analysis using memory forensic technique," Proceedings of the 16Th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, The 11Th Ieee International Conference on Big Data Science and Engineering, And The 14Th Ieee International Conference on Embedded Software and Systems, DOI: 10.1109/Trustcom/ BigDataSE/ ICESS.2017.365

[16] Ravula R., Chan C. and J. Liszka K. (2011). Dynamic Analysis of Malware Using Decision Trees. International Conference on Knowledge Discovery and Information, pg. 74-83. DOI: 10.5220/0003660200740083

[17] Revay, G. (2022) An Overview of the Increasing Wiper Malware Threat. Fortinet.

[18] Serpanos D., Michalopoulos P., Xenos G., & Ieronymakis V. (2021). Sisyfos: A Modular and Extendable Open Malware Analysis Platform. MDLI.

[19] Singh, J. and Singh, J. (2018). "Challenges of Malware Analysis: Obfuscation Techniques". International Journal of Information Security Science. Vol.7, No.3.

[20] Wangen G. (2015). "The Role of Malware in Reported Cyber Espionage: A Review of the Impact and Mechanism". Norwegian Information Security Laboratory, Center for Cyber and Information Security. Gjøvik University College, Teknologivn. www.mdpi.com/journal/information