



Please cite the Published Version

AliAhmad, Anas, Eleyan, Derar, Eleyan, Amna , Bejaoui, Tarek, Zolkipli, Mohamad Fadli and Al-Khalidi, Mohammed  (2023) Malware detection issues, future trends and challenges: a survey. In: 2023 International Symposium on Networks, Computers and Communications (ISNCC), 23 October 2023 - 26 October 2023, Doha, Qatar.

DOI: <https://doi.org/10.1109/isncc58260.2023.10323624>

Publisher: IEEE

Version: Accepted Version

Downloaded from: <https://e-space.mmu.ac.uk/633522/>

Usage rights:  In Copyright

Additional Information: © 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Malware Detection Issues, Future Trends and Challenges: A Survey

Anas AliAhmad

Department of Investigations and
Cybercrime, College of Graduate
Studies, Palestine Technical University
a.n.aliahmad2@students.ptuk.edu.ps

Tarek Bejaoui

Computer Engineering Department,
University of Carthage, Tunisia
tarek.bejaoui@ieeee.org

Derar Eleyan

Applied Computing Department,
Palestine Technical University,
Kadoorie, Palestine
d.eleyan@ptuk.edu.ps

Mohamad Fadli Zolkipli

School of Computing,UUM CAS,
University Utara
Malaysiam.fadli.zolkipli@uum.edu.my

Amna Eleyan

Department of Computing and
Mathematics, Manchester Metropolitan
University, Manchester, United
Kingdom

a.eleyan@mmu.ac.uk

Mohammed Al-Khalidi

Department of Computing and
Mathematics, Manchester Metropolitan
University, Manchester, United
Kingdom

m.al-khalidi@mmu.ac.uk

Abstract— This paper focuses on the challenges and issues of detecting malware in to-day's world where cyberattacks continue to grow in number and complexity. The paper reviews current trends and technologies in malware detection and the limitations of existing detection methods such as signature-based detection and heuristic analysis. The emergence of new types of malware, such as file-less malware, is also discussed, along with the need for real-time detection and response. The research methodology used in this paper is presented, which includes a literature review of recent papers on the topic, keyword searches, and analysis and representation methods used in each study. In this paper, the authors aim to address the key issues and challenges in detecting malware today, the current trends and technologies in malware detection, and the limitations of existing methods. They also explore emerging threats and trends in malware attacks and highlight future directions for research and development in the field. To achieve this, the authors use a research methodology that involves a literature review of recent papers related to the topic. They focus on detecting and analyzing methods, as well as representation and ex-traction methods used in each study. Finally, they classify the literature re-view, and through reading and criticism, highlight future trends and problems in the field of malware detection.

Keywords— malware analysis; malware feature; future solutions to malware; Malware detection methods

Introduction

The provision and growth of commercial Internet services and services that provide services that provide services provided by financial services, financial banking and financial services[1]. The cost of the damage is expected to be \$6 trillion, according to the official annual Internet Crime Report of the Cyber Security Enterprises [2]. One of the data collection methods that researchers use is dynamic analysis and dynamic analysis, and then it is collected in one set of features, all in order to increase the accuracy of malware detection. And then a combination of the advantages and disadvantages of each of the two analyses was taken through the hybrid analysis approach [3]. The analysis of this malware is similar to the game of catch, because the person who may have written this malware has written it in a way that some experienced people can thwart it. From some The main components of the static analysis are re-verse engineering and

malware [4]. In cyberspace, there have been numerous instances of security incidents and cyber attacks. One such incident occurred in May 2017 when the WannaCry ransomware exploited system vulnerabilities to infect and damage the computers of hundreds of thousands of users in several countries. On May 9, 2021, the United States declared a state of emergency when a cyber-attack on the country's largest fuel pipeline operator caused it to go offline.[5] There are two ways to detect programs which are signature based programs and behavior detectors, in programs based on some classes are hash and byte signatures and signature inference, especially malicious software in this case, they run and are very fast. The most common hash functions are MD5 and SHA1[6].The increasing frequency and sophistication of malware attacks: Malware attacks continue to grow in both number and complexity, making it increasingly challenging to detect and prevent them. The limitations of current malware detection techniques: Existing malware detection methods such as signature-based detection and heuristic analysis have limitations in terms of their ability to detect new and unknown malware. The emergence of new types of malware: The rise of new technologies and platforms has given rise to new types of malware, such as fileless malware, which can evade traditional detection techniques. The need for real-time detection and response: As malware attacks become more advanced and automated, there is a growing need for real-time detection and response to prevent damage and mitigate the impact of attacks.

The research questions:

- 1.What are the key issues and challenges in detecting malware today?
- 2.What are the current trends and technologies in malware detection?
- 3.What are the limitations of current malware detection methods?
- 4.What are the emerging threats and trends in malware attacks?
- 5.What are the future directions for malware detection research and development?

I. RESEARCH METHODOLOGY

The methodology was followed as in Figure 1, to present this paper, this was done through a set of methods, In this report, we focus on recent Recently written papers are reviewed for the benefit of a limit or set of limits for current revisions, and then reviewed. that we need new papers to review the literature, and secondly with the first step as well. Specific keywords were used to collect empirical papers related to the same topic. Thirdly, according to the detection and analysis methods along with the representation and extraction methods that were used in each study separately. Finally, a literature review was classified. Operations were carried out in the last stage, which is reading and criticism, to obtain The final result and highlighting future trends and problems in the field of malware detection, reading and classification, and Figure 1 illustrates the methodology used [7].

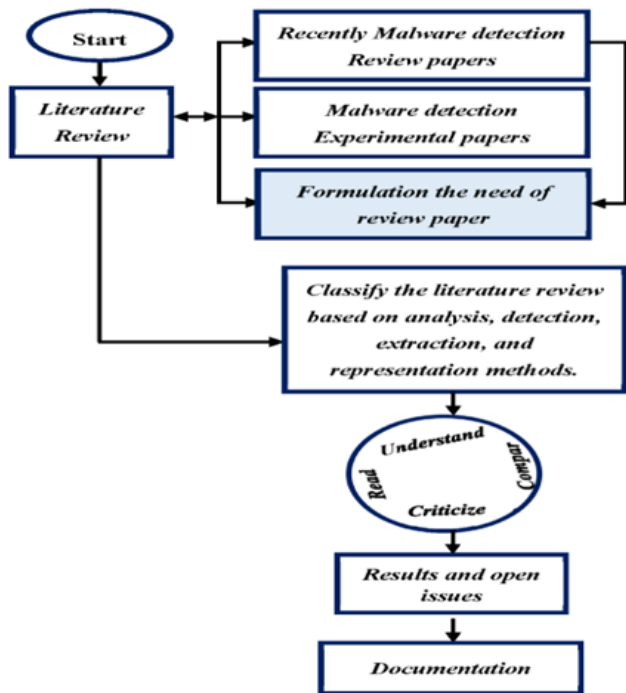


Fig. 1. Research Methodology

The number of records that have been excluded has been identified and included exceptions. Figure 2 illustrates this scheme.

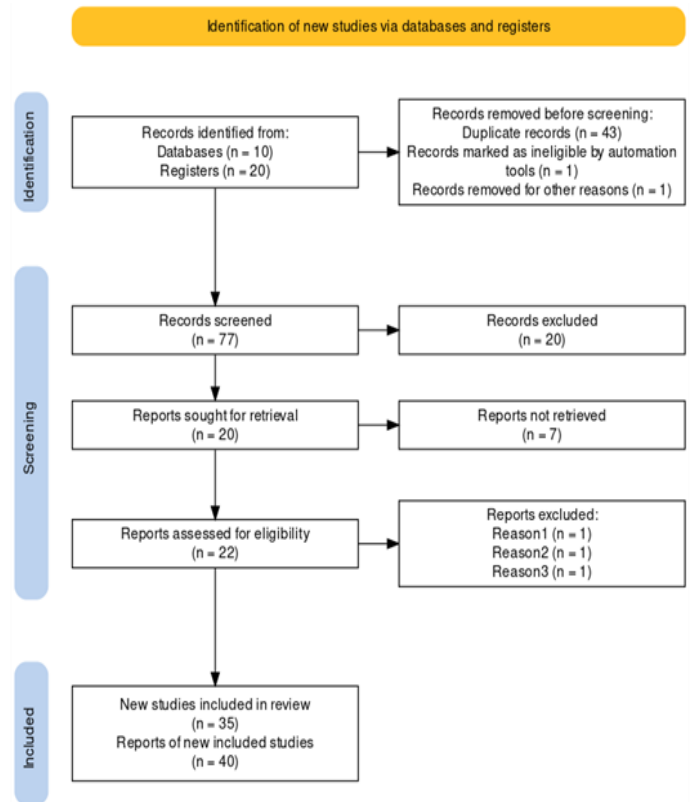


Fig. 2. Prisma framework flow diagram.

TABLE 1. Databases & keyword

Publisher	Articles	Keywords
IEEE	6	malware
MDPI	2	analysis;
Sites	1	malware
sciencedirect	10	feature; future
springer	7	solutions to
Applied Security Research(taylor)	1	malware;
		Malware
		detection
		methods.
arxiv	2	
Cybersecurity	1	
mecs	1	
Wiley Online Library	1	
mendel soft computer journal	1	
uti	1	
advanced science	1	

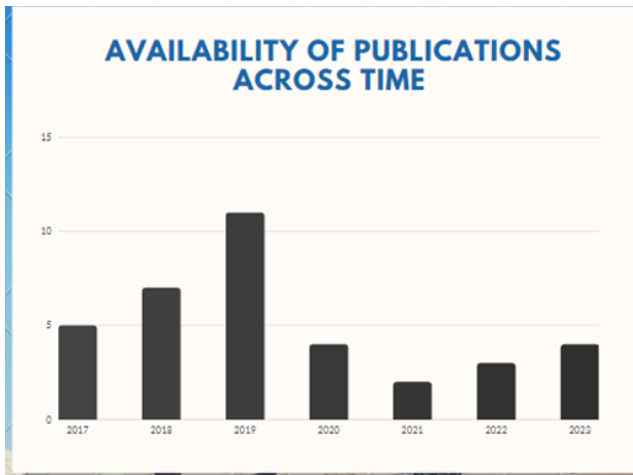


Fig. 3. Availability Of Publications Across Time.

II. RELATED WORK

In the malware community, many review papers have been prepared and com-piled to find out definitions of some malware and how malware evolves to be more complex[7] Some authors[8][9] presented some methods such as dynamic, static, and hybrid as analysis methods, and they divided the detection methods into a number of divisions, for example specification-“based”, signature-“based”, ma-chine-“learning”-“based”, and multimedia-“based”. Malware can be defined as any file that can cause harm to a computer, and it means malicious software. One of the first ways to detect this software is static analysis that analyzes files without executing. Another method is dynamic analysis that analyzes and monitors file behavior during execution. There are many malicious programs such as Trojan horses, viruses, backdoors, adware that are intrusive, ransomware[10] . We will investigate some methods in analyzing malware, study methods for detecting and classifying these programs, compare different criteria such as supported file formats, software used and environment analysis, as well as knowing the features that have been extracted, classifying malicious programs, and representing features, knowing the algorithm used The results appear in the Table 1 [10].

TABLE 2. Comparison of various malware detection/classification techniques

Study	"Detection/Classification"	"Analysis type"	"Analysis Environment"	"File Format Supported"	"Software Used"
[11]	Classification	"Dynamic"	"VM"	"DDL,DOC,E XE,PDF, PPT,HTML,X SL and URL"	"Customized Cuckoo Sandbox InetSim, Mongo DB, WEKA"
[6]	Classification	Dynamic	"VM"	"DLL,EXE,DOC,PDF PPT,HTML,X SL And URL"	AutoMal, MaLabel
[12]	Detection	"Static and Dynamic"	"VM"	"Portable Executable"	"Cuckoo Sandbox, WEKA"
[13]	Classification	FFRI data st	-	-	-
[14]	Detection	Dynamic	"VM"	Portable executable	PIN tool

One of the reasons for the difficulty of accessing the reasons for the difficulty of access and dealing with its growth and development in complexity and size with programs is its continuous growth in complexity and size [15]. Difficulty analyzing malware, and therefore difficult to accurately discover

the solution to it [6]. A classification of cyber damages has been presented through a comprehensive evaluation of the analysis and literature with a set of cyber incidents in order to communicate a deeper picture Direct and indirect problems and damages that people and companies may face. The next steps are to expand the field of re-search by making models and designing them to be asset-oriented., which will encourage organizations Focus on the fundamentals of its essence [16]. Having redundant or unreasonable information in datasets is a software problem [17][18]. The main Malware analysis aims to capture additional files and features in order to use them later in security improvements as much as possible, thus detecting any malicious software on the device that is not required[15] . The following problem is considered to be more prevalent with viruses, spyware and malware (33%), as it constitutes, impersonation of organizations' pages in e-mail messages or via the Internet constitutes (27%), and ransomware constitutes (17%) [16]. To solve this problem, a set of procedures must be taken as follows. First, do not open any suspicious or unofficial message, check the re-quired permissions before downloading or opening any file, and do not click on unverified links[19].

III. CLASSIFICATION OF METHODS FOR DETECTING AND ANALYZING MALWARE

In this section we will describe the classification of malware detection and analysis methods. Malware detection is presented in depth classification And every known discovery approach is offered subtypes as well as The relationship between each subtype of data types being presented and disclosed. Figure 4 shows malware analysis and detection classification [7].

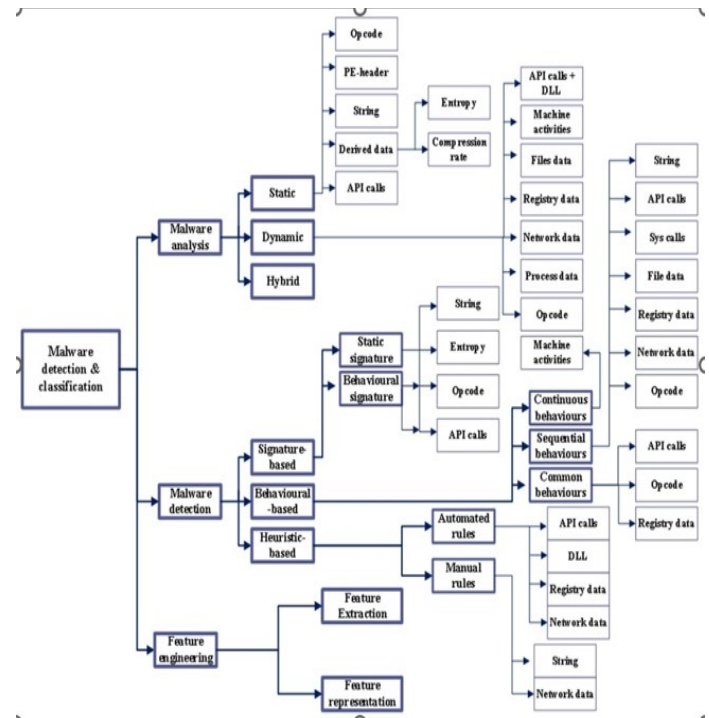


Fig. 4. Malware Detection and Classification Analysis.

A. Data types and malware analysis methods

Static Analysis

Static analysis has been used extensively by not revealing the underlying code, but the executable files that do not run. We can collect static data set by static analysis, including “PE header data”. [20][21][22][23] Some The data that was accessed such as entropy based on compression ratio and series [24][25][26][27]. In addition, some tools are used for static analysis such as “IDA Pro” and advanced “Python” modules compile and call API and static opcode [28][29][30][31][32]. Although static analysis is able to follow all possible execution methods and paths, it is affected by coding and packaging techniques.

Dynamic Analysis

A large number of researchers carried out dynamic analysis to collect different types of data in order to distinguish malicious programs and benign files, by running the files that accept execution in virtual machines (VMs), isolated environments, or simulation programs, and all in order This is to monitor the file during its working period. After that, the required dynamic data is collected [33]. Also, several types of data were collected by the dynamic analysis approach. We can represent some malicious activities dynamically using the behavior of the executable file and by saving memory images during runtime. Finally, The behavior or nature of files that are executable can be known by compiling API calls that have been called[34],[35],[36],[37],[38].

Hybrid Analysis

In In some previous papers and studies, some of the extracted data were collected through both dynamic and static analysis in order to reduce the defects of the approach and to reach a higher detection rate. Various tools such as “Cuckoo sandbox, IDA pro disassembler, and OlleyDbg”. In order to collect static and dynamic data, and then create a set of mixed features based on a set of types of static and dynamic data, and then create a number of mixed features based on a set Information, data and examples opcode, string, “API” call, and others [39],[40],[41],[42],[43].

B. An environment that is suitable for the deployment and execution of malware in IoT devices

IoT devices are vulnerable to various forms of attacks, such as "physical, network-layer, and application-layer attacks", which take advantage of weaknesses in the system. Malware can be executed on these devices due to several reasons, including outdated firmware, weak authentication, connectivity issues, and the limitations of resource-constrained devices [44].

Challenges of adversarial attacks for PE malware

This section focuses on the challenges posed by adversarial attacks for PE mal-ware. We begin by introducing the general concept and taxonomy of adversarial attacks, which have been studied extensively in the field of image classification tasks. We then highlight the distinct challenges that arise when applying adversarial attacks to PE malware, which differ from those encountered in other fields such as images, audios, and texts[45].

Sandboxing techniques

The process of detecting malware involves making informed decisions that ultimately lead to the identification of a malicious program. Malware researchers commonly employ a sandbox environment to execute code obtained from unknown attachments or suspicious URLs, in order to observe

the behavior of the malware code. By executing the code in a controlled environment, researchers can better understand the tactics, techniques, and procedures used by malware to evade detection and compromise systems. This enables them to develop more effective strategies for detecting and mitigating malware threats.[46].

IV. DISCUSSION

Despite the popularity of the n-gram method for feature extraction among authors, it results in a high number of features, which creates challenges for N-gram-based techniques. Frequency-dependent feature extraction methods so that it works to reduce The number of positives and features extracted by taking the majority only, and in most this happens, and the result is often overcoming the problem of high-dimensional space, and despite all that, some jamming methods are able to adjust certain features and frequencies in Each of them, the frequency-based meth-od can be overridden by a variable. It is difficult or impossible to make a rare graph of each section of the malware, but some studies have used the A graph-based feature extraction method is utilized to create generic graphs based on shared characteristics.. Most of the models that are based on graphs have a great degree of complexity with the passage of time. Examinations have been made for the problem of matching certainty, and a proposal for a solution has been developed Time matching, and a proposal for a solution was developed in the representation of the stage when representing some of the graph studies that were created as vectors [7]. Ac-cording to the literature, researchers have utilized various techniques such as opcode analysis, thread behavior analysis, and feature derivation to create static signatures.. Operation codes and “API calls” have been used to generate dynamic signatures. In general, malware can be detected by predictions derived earlier, and the use of dynamic and static signatures may not be sufficient to combat malware. The authors tried to obtain common behaviors among the malware group, for example opcodes and API calls. One of the most promising solutions to overcome these vulnerabilities is the behavior-based approach, but if we rely on these behaviors, it will lead the proposed models to misclassify the malware that Similar functions work as innocence or mimicry. Subsequently, the samples that suffer from high lying rates. Otherwise, malicious software is able to know The nature of its implementation through evasion techniques, after which its behavior changes to become similar to some benign behaviors, or it terminates its implementation, which leads to its representation through unrepresentative behaviors. It is difficult to extract sufficient features, a difficult process, and it has a very significant impact on detection models. Mal-ware and its classification. Malware represents Certain behaviors that depend on names, duplicate extracted properties, or sequences can lead to malware detection and differentiation of models most susceptible to obfuscation techniques that were used to update sequences, frequencies, and names of previous characteristics. A number of Some researchers have provided training for their models by recruitment them [7]. Key issues and challenges in detecting malware today include the increasing complexity and sophistication of malware, the use of encryption and obfuscation techniques by malware authors to evade detection, the rapid evolution and proliferation of new malware variants, and the difficulty of detecting zero-day attacks. Current trends and technologies in malware detection include the use of machine learning and artificial intelligence

algorithms to analyze behavior patterns, the use of sandboxing techniques to isolate and analyze malware samples in a safe environment, the use of cloud-based malware detection services, and the integration of threat intelligence feeds and data analytics tools. Limitations of current malware detection methods include the potential for false positives and false negatives, the limited effectiveness of signature-based detection methods against new and un-known malware variants, the difficulty of detecting polymorphic and metamorphic malware, and the performance impact of some detection techniques on system re-sources. Emerging threats and trends in malware attacks include the use of fileless malware that resides in memory and leaves no trace on disk, the increasing use of ransomware as a tool for extortion and data theft, the targeting of Internet of Things (IoT) devices, and the use of social engineering techniques to trick users into installing malware. Future directions for malware detection research and development include the development of more sophisticated machine learning algorithms that can detect and respond to emerging threats in real-time, the use of blockchain technology to improve malware detection and analysis, the development of more effective methods for detecting fileless malware, and the integration of artificial intelligence and machine learning techniques with traditional signature-based detection methods. Additionally, the use of decentralized and distributed systems to improve the scalability and efficiency of malware detection and response is an area of active research. The difference between the work and the previous work

The previous paper talks in general about the security threats that computer systems and users may be exposed to through malicious programs (malware), and indicates the importance of conducting an analysis of these programs to understand their behavior and try to protect systems from them. It also introduces some malware analysis methods, such as static, dynamic, and hybrid analysis, and indicates some types of malware detection techniques, such as signature pattern detection, behavioral detection, and subjective detection. As for this paper, it presents a new type of classification that is necessary to understand the relationship between the methods of analyzing malicious programs and the types of data used in these analyses. The paragraph focuses on providing a detailed taxonomy of malware analysis methods in greater detail, relates each category of analysis methods to specific types of data that can be used in these analyzes, and also introduces a new classification of data representation methods used in malware analysis. This classification is intended to provide a more detailed view of the research community in the field of malware analysis and detection.

V. CONCLUSIONS

In this survey, we made a comprehensive review on the trends and development of malware detection and analysis methods. This survey focused on points of view that are often ignored or considered and partially studied through previous surveys. Such as discovering the benefits of each type of data according to the analysis method that was used, which Provides an in-depth classification of malware detection methods where detection methods are given priority over predictive, behavioral, or heuristic detection methods [7]. Also, demonstration programs show about machine learning techniques There are five contributions to our work.

1. We suggested that the reviewed works be organized according to three approaches.

- “The objective of the analysis”
- “The type of features extracted of samples”
- “Machine learning algorithms used to process these features”

2. We accessed a list of literature that talks about the analysis of malware for PE by means of an automatic selection process that was developed according to the proposed classifications, in addition to providing a detailed and comparative analysis of the work that was studied and surveyed We will focus on the most important current problems in machine learning for malware analysis: what processes need to be considered for some of the features and datasets used and the anti-analytics techniques used by the malware.

3. We've identified some thematic trends around interesting features and targets like malware screening and attribution.

4. We introduced Meaning and modern definition of economics of malware analysis in relation to exploits and investigated performance metrics [15].

In this Review, we presented a dynamic description in order to analyze the malicious programs, so that it will be the nature of the behavior that was worked on during the operation of the malicious programs was shown. Because of the packing techniques and code obfuscation, the performance of the static analysis is lower when differentiating it from dynamic analysis, because the malware is executed in a custom environment, most studies conclude that making an API call is one of the key features that will describe “malicious behavior”. There are a set of ways to represent the API call, including frequency and binary representation. In future work, we will infer some of the most important positives or features of the behavior and tools that were accessed and created during the work and implementation of malicious programs and propose a highly efficient system for classifying programs harmful and detected [10]

REFERENCES

- [1] L. Caviglione et al., “Tight Arms Race: Overview of Current Malware Threats and Trends in Their Detection,” *IEEE Access*, vol. 9, pp. 5371–5396, 2021, doi: 10.1109/ACCESS.2020.3048319.
- [2] B. M. Iqbal, Juneed; Beigh, “Cybercrime in India: Trends and Challenges,” *Int. J. Innov. Adv. Comput. Sci.*, vol. 6, no. 12, pp. 187–196, 2017.
- [3] R. Sihwail, K. Omar, and K. A. Z. Ariffin, “A survey on malware analysis techniques: Static, dynamic, hybrid and memory analysis,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 4–2, pp. 1662–1671, 2018, doi: 10.18517/ijaseit.8.4-2.6827.
- [4] M. Egele, T. Scholte, E. Kirda, and C. Kruegel, “A survey on automated dynamic malware-analysis techniques and tools,” *ACM Comput. Surv.*, vol. 44, no. 2, 2012, doi: 10.1145/2089125.2089126.
- [5] A. Yang et al., “Application of meta-learning in cyberspace security: a survey,” *Digit. Commun. Networks*, vol. 9, no. 1, pp. 67–78, 2023, doi: 10.1016/j.dcan.2022.03.007.
- [6] A. Mohaisen, O. Alrawi, and M. Mohaisen, “AMAL: High-fidelity, behavior-based automated malware analysis and classification,” *Comput. Secur.*, vol. 52, pp. 251–266, 2015, doi: 10.1016/j.cose.2015.04.001.
- [7] F. A. Aboaja, A. Zainal, F. A. Ghaleb, B. A. S. Al-rimy, T. A. E. Eisa, and A. A. H. Elnour, “Malware Detection Issues, Challenges, and Future Directions: A Survey,” *Appl. Sci.*, vol. 12, no. 17, 2022, doi: 10.3390/app12178482.

- [8] R. Tahir, "A Study on Malware and Malware Detection Techniques," *Int. J. Educ. Manag. Eng.*, vol. 8, no. 2, pp. 20–30, 2018, doi: 10.5815/ijeme.2018.02.03.
- [9] D. Gibert, C. Mateu, and J. Planes, "The rise of machine learning for detection and classification of malware: Research developments, trends and challenges," *J. Netw. Comput. Appl.*, vol. 153, no. December 2019, p. 102526, 2020, doi: 10.1016/j.jnca.2019.102526.
- [10] S. C. Satapathy, V. Bhateja, S. K. Udgata, and P. K. Pattnaik, "Preface," *Adv. Intell. Syst. Comput.*, vol. 515, pp. v–vii, 2017, doi: 10.1007/978-981-10-3153-3.
- [11] R. S. Pircoveanu, S. S. Hansen, T. M. T. Larsen, M. Stevanovic, and J. M. Pedersen, "Analysis of malware behaviour classification from ML," 2015.
- [12] B. Zones, "Subject Index," pp. 703–711, 2004, doi: 10.1016/s0169-3158(04)80011-2.
- [13] N. Kawaguchi and K. Omote, "Malware function classification using apis in initial behavior," *Proc. - 2015 10th Asia Jt. Conf. Inf. Secur. AsiaJCIS 2015*, pp. 138–144, 2015, doi: 10.1109/AsiaJCIS.2015.15.
- [14] M. Ozsoy, C. Donovan, I. Gorelik, N. Abu-Ghazaleh, and D. Ponomarev, "Malware-aware processors: A framework for efficient online malware detection," 2015 IEEE 21st Int. Symp. High Perform. Comput. Archit. HPCA 2015, pp. 651–661, 2015, doi: 10.1109/HPCA.2015.7056070.
- [15] D. Ucci, L. Aniello, and R. Baldoni, "Survey of machine learning techniques for malware analysis," *Comput. Secur.*, vol. 81, pp. 123–147, 2019, doi: 10.1016/j.cose.2018.11.001.
- [16] I. Agrafiotis, J. R. C. Nurse, M. Goldsmith, S. Creese, and D. Upton, "A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate," *J. Cybersecurity*, vol. 4, no. 1, pp. 1–15, 2018, doi: 10.1093/cybsec/tyy006.
- [17] Y. A. Ahmed, B. Koçer, S. Huda, B. A. Saleh Al-rimy, and M. M. Hassan, "A system call refinement-based enhanced Minimum Redundancy Maximum Relevance method for ransomware early detection," *J. Netw. Comput. Appl.*, vol. 167, p. 102753, 2020, doi: 10.1016/j.jnca.2020.102753.
- [18] F. Parrales-Bravo, J. Torres-Urresto, D. Avila-Maldonado, and J. Barzola-Monteses, "Relevant and Non-Redundant Feature Subset Selection Applied to the Detection of Malware in a Network," *ETCM 2021 - 5th Ecuador Tech. Chapters Meet.*, Oct. 2021, doi: 10.1109/ETCM53643.2021.9590777.
- [19] "الحماية من البرمجيات الخبيثة والبرمجيات الخبيثة: كل ما تحتاج إلى معرفته," <https://me.kaspersky.com/resource-center/threats/malware-protection> (accessed Apr. 29, 2023).
- [20] S. Naz and D. K. Singh, "Review of Machine Learning Methods for Windows Malware Detection," 2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019, pp. 1–6, 2019, doi: 10.1109/ICCCNT45670.2019.8944796.
- [21] Kamesh and N. Sakthi Priya, "A survey of cyber crimes Yanping," *Secur. Commun. Networks*, vol. 5, no. June, pp. 422–437, 2012, doi: 10.1002/sec.
- [22] E. Amer and I. Zelinka, "An ensemble-based malware detection model using minimum feature set," *Mendel*, vol. 25, no. 2, pp. 1–10, 2019, doi: 10.13164/mendel.2019.2.001.
- [23] T. Denzer, A. Shalaginov, and G. O. Dyrkolbotn, "Intelligent Windows Malware Type Detection based on Multiple Sources of Dynamic Characteristics," *Nis. J.*, vol. Vol. 12.20, no. November, 2019.
- [24] P. Seshagiri, A. Vazhayil, and P. Sriram, "AMA: Static Code Analysis of Web Page for the Detection of Malicious Scripts," *Procedia Comput. Sci.*, vol. 93, no. September, pp. 768–773, 2016, doi: 10.1016/j.procs.2016.07.291.
- [25] D. Wael, S. G. Sayed, and N. AbdelBaki, "Enhanced approach to detect malicious VB script files based on data mining techniques," *Procedia Comput. Sci.*, vol. 141, pp. 552–558, 2018, doi: 10.1016/j.procs.2018.10.127.
- [26] Y. T. Ling, N. F. M. Sani, M. T. Abdullah, and N. A. W. A. Hamid, "Nonnegative matrix factorization and metamorphic malware detection," *J. Comput. Virol. Hacking Tech.*, vol. 15, no. 3, pp. 195–208, 2019, doi: 10.1007/s11416-019-00331-0.
- [27] A. Kumar, K. S. Kuppusamy, and G. Aghila, "A learning model to detect maliciousness of portable executable using integrated feature set," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 31, no. 2, pp. 252–265, 2019, doi: 10.1016/j.jksuci.2017.01.003.
- [28] I. D. Odol, V. E. X. Df, P. Ud, and L. D. F. Lu, "+ Hxulwlvf Phwdprusklf Pdozduh Ghwhfwlrq Edvhg Rq Vwdwlvwlvf Ri Dvvhpeo \ Lqvwxwfwlrqv Xvlqj Fodvviilfdwlrq".
- [29] H. Hashemi, A. Azmoodeh, A. Hamzeh, and S. Hashemi, "Graph embedding as a new approach for unknown malware detection," *J. Comput. Virol. Hacking Tech.*, vol. 13, no. 3, pp. 153–166, 2017, doi: 10.1007/s11416-016-0278-y.
- [30] S. Euh, H. Lee, D. Kim, and D. Hwang, "Comparative analysis of low-dimensional features and tree-based ensembles for malware detection systems," *IEEE Access*, vol. 8, pp. 76796–76808, 2020, doi: 10.1109/ACCESS.2020.2986014.
- [31] A. Khalilian, A. Nourazar, M. Vahidi-Asl, and H. Haghighi, "G3MD: Mining frequent opcode sub-graphs for metamorphic malware detection of existing families," *Expert Syst. Appl.*, vol. 112, pp. 15–33, 2018, doi: 10.1016/j.eswa.2018.06.012.
- [32] R. Lu, "Malware Detection with LSTM using Opcode Language," 2019, [Online]. Available: <http://arxiv.org/abs/1906.04593>
- [33] S. P. Choudhary and M. D. Vidyarthi, "A Simple Method for Detection of Metamorphic Malware using Dynamic Analysis and Text Mining," *Procedia Comput. Sci.*, vol. 54, pp. 265–270, 2015, doi: 10.1016/j.procs.2015.06.031.
- [34] H. S. Galal, Y. B. Mahdy, and M. A. Atia, "Behavior-based features model for malware detection," *J. Comput. Virol. Hacking Tech.*, vol. 12, no. 2, pp. 59–67, 2016, doi: 10.1007/s11416-015-0244-0.
- [35] R. Mosli, R. Li, B. Yuan, and Y. Pan, "Automated malware detection using artifacts in forensic memory images," 2016 IEEE Symp. Technol. Homel. Secur. HST 2016, 2016, doi: 10.1109/THS.2016.7568881.
- [36] J. Hwang, J. Kim, S. Lee, and K. Kim, "Two-Stage Ransomware Detection Using Dynamic Analysis and Machine Learning Techniques," *Wirel. Pers. Commun.*, vol. 112, no. 4, pp. 2597–2609, 2020, doi: 10.1007/s11277-020-07166-9.
- [37] M. A. Jerlin and K. Marimuthu, "A New Malware Detection System Using Machine Learning Techniques for API Call Sequences," *J. Appl. Secur. Res.*, vol. 13, no. 1, pp. 45–62, 2018, doi: 10.1080/19361610.2018.1387734.
- [38] H. Kim, J. Kim, Y. Kim, I. Kim, K. J. Kim, and H. Kim, "Improvement of malware detection and classification using API call sequence alignment and visualization," *Cluster Comput.*, vol. 22, pp. 921–929, 2019, doi: 10.1007/s10586-017-1110-2.
- [39] J. B. Fraley and M. Figueroa, "Polymorphic malware detection using topological feature extraction with data mining," *Conf. Proc. - IEEE SOUTHEASTCON*, vol. 2016-July, 2016, doi: 10.1109/SECON.2016.7506685.
- [40] B. Ndiabanje, K. H. Kim, Y. J. Kang, H. H. Kim, T. Y. Kim, and H. J. Lee, "Cross-method-based analysis and classification of malicious behavior by API calls extraction," *Appl. Sci.*, vol. 9, no. 2, 2019, doi: 10.3390/app9020239.
- [41] W. Zhong and F. Gu, "A multi-level deep learning system for malware detection," *Expert Syst. Appl.*, vol. 133, pp. 151–162, 2019, doi: 10.1016/j.eswa.2019.04.064.
- [42] X. Huang, L. Ma, W. Yang, and Y. Zhong, "A Method for Windows Malware Detection Based on Deep Learning," *J. Signal Process. Syst.*, vol. 93, no. 2–3, pp. 265–273, 2021, doi: 10.1007/s11265-020-01588-1.
- [43] A. Damodaran, F. Di Troia, C. A. Visaggio, T. H. Austin, and M. Stamp, "A comparison of static, dynamic, and hybrid analysis for malware detection," *J. Comput. Virol. Hacking Tech.*, vol. 13, no. 1, pp. 1–12, 2017, doi: 10.1007/s11416-015-0261-z.
- [44] M. Venkatasubramanian, A. H. Lashkari, and S. Hakak, "IoT Malware Analysis Using Federated Learning: A Comprehensive Survey," *IEEE Access*, vol. 11, no. January, pp. 5004–5018, 2023, doi: 10.1109/ACCESS.2023.3235389.
- [45] X. Ling et al., "Adversarial attacks against Windows PE malware detection: A survey of the state-of-the-art," *Comput. Secur.*, vol. 128, 2023, doi: 10.1016/j.cose.2023.103134.
- [46] M. Gopinath and S. C. Sethuraman, "A comprehensive survey on deep learning based malware detection techniques," *Comput. Sci. Rev.*, vol. 47, p. 100529, 2023, doi: 10.1016/j.cosrev.2022.100529.