INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Features, Analysis Techniques, and Detection Methods of Cryptojacking Malware: A Survey

Laith M Kadhum [a,b], Ahmad Firdaus [a], Syifak Izhar Hisham [a,*], Waheed Mushtaq [a], Mohd Faizal Ab Razak [a]

[a] *Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Pahang, Malaysia*
[b] *University of Kufa, Najaf, Iraq*
*Corresponding author: *syifakizhar@umpsa.edu.my*

Abstract— **Various types of malwares are capable of bringing harm to users. The list of types are root exploits, botnets, trojans, spyware, worms, viruses, ransomware, and cryptojacking. Cryptojacking is a significant proportion of cyberattacks in which exploiters mine cryptocurrencies using the victim's devices, for instance, smartphones, tablets, servers, or computers. It is also defined as the illegal utilization of victim resources (CPU, RAM, and GPU) to mine cryptocurrencies without detection. The purpose of cryptojacking, along with numerous other forms of cybercrime, is monetary gain. Furthermore, it also intended to stay concealed from the victim's viewpoint. Following this crime, to the author's knowledge, a paper focusing solely on a review of cryptojacking research is still unavailable. This paper presents cryptojacking detection information to address this deficiency, including methods, detection, analysis techniques, and features. As cryptojacking malware is a type that executes its activities using the network, most of the analysis and features fall into dynamic activities. However, static analysis is also included in the security researcher's option. The codes that are involved are opcode and JavaScript. This demonstrates that these two languages are vital programming languages to focus on to detect cryptojacking. Moreover, the researchers also begin to adopt deep learning in their experiments to detect cryptojacking malware. This paper also examines potential future developments in the detection of cryptojacking.**

Keywords— **Cryptojacking; cryptocurrencies; distribution; detection.**

Manuscript received 27 Sep. 2023; revised 5 Dec. 2023; accepted 12 Feb. 2024. Date of publication 31 May 2024.

I. Introduction

Cryptocurrencies are digital asset currencies that are supported by cryptographic protocols. They facilitate safer online transactions without any need for intermediaries [1]. Crypto refers to the diverse encryption algorithms and cryptographic techniques that protect these entries and manage third-party intermediaries. To own cryptocurrencies, we need to mine or purchase them from cryptocurrency exchanges. For mining-based cryptocurrencies (namely bitcoin, Ethereum, or Litecoin), the type of algorithm is called proof of work (PoW). PoW denotes entities that participate in the "voting" process and demonstrate they have solved a moderately tricky puzzle. To verify a transaction and add it to the distributed ledger, participants should compute a Proof of Work (PoW). The participants must sacrifice their hardware (computer or asic machine) and high electricity consumption to solve the puzzle's complexity. Once the puzzle is solved, the participants will receive cryptocurrency as a reward, and then the transaction of cryptocurrency to the receiver from the sender is finally completed. After this, a type of malware called cryptojacking exists.

Malware includes botnets, root exploits, ransomware, Trojans, viruses, ransomware, cryptojacking, cryptocurrency miners, and adware [2], [3]. Cryptojacking is distinct from other forms of malware. Their primary objective is not to inflict harm to the attacked system. However, they complete assigned tasks and illegally earn cryptocurrency using the victim's resources. In 2013, browser-based crypto mining was introduced. It was an alternative revenue model to advertisements. Cybercriminals employed this tactic extensively during the 2017 cryptocurrency boom [4].

From another point of view, there is an existing relationship between cryptocurrency and Ultra-Wide Band (UWB). UWB is a new short-range wireless technology that has the potential to revolutionize how we interact with devices. UWB could enable many new applications, including exact location tracking, high-speed data transfer, and crypto-currency mining. One potential use case for UWB is cryptojacking, whereby UWB-enabled devices could be used to mine for cryptocurrency without the knowledge or approval of the user. It would allow criminals to surreptitiously generate revenue by using other people's devices to do the mining for them. UWB is still in its early days, so it remains to be seen how this new technology will be used. However, if UWB does take off, cryptojacking could become a significant problem. UWB can potentially revolutionize many aspects of our lives, including how we interact with devices.

Coinhive was first introduced in September 2017. It mines the Monero cryptocurrency. Coinhive proposed to use crypto mining method to earn revenue from the websites instead of ads. Website owners can embed JavaScript code to mine Monero. This way, the company can mine cryptocurrency while users are simply browsing the website. The creators of the website still capable of gaining profit and supporting their businesses without users being bothered [2] Whenever consumers visit any website integrated with the embedded mining coding script, the hardware resources (RAM or CPU) of the victim's website are utilized to initiate the crypto-mining process on their behalf. However, website owners are incentivized to employ system resources to do complicated computational jobs, not users. Additionally, hackers profited from the code's placement on the hacked websites.

Inexperienced users would click and accidentally install scripts without verifying the changes, leaving their system vulnerable to attack [6], [7]. This situation calls for more study in crypto jacking detection. The methodology for this study is described in this part, along with how the research articles were found and categorized. Additionally, this section highlights the studies' interest in crypto jacking detection. The assault seen in Fig. 1 illustrates the initial step in the attack from the attacker to utilize the victim's or website owner's hardware illegally. The attacker then mines Bitcoin illegitimately using the victim's hardware (CPU, GPU, or ASIC machine) in order to earn cryptocurrency as a reward.

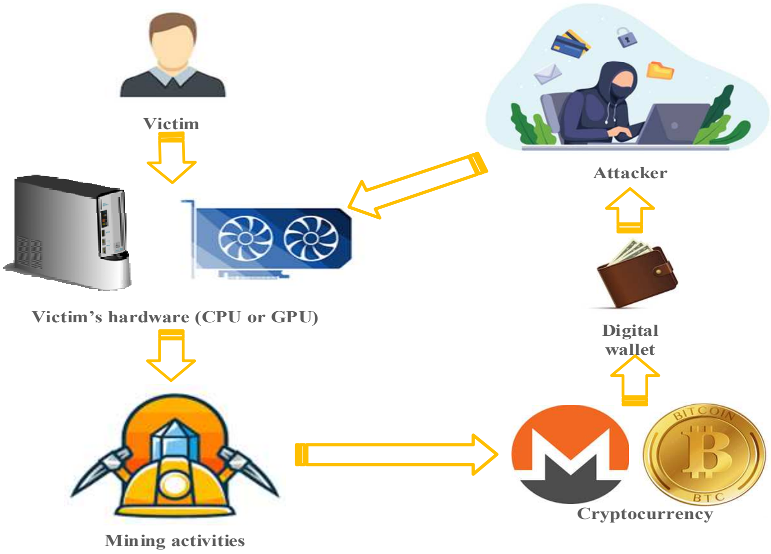

Fig. 1 Cryptojacking attack

Despite the tremendous significance in cryptojacking detection, no study has yet analyzed previous cryptojacking research. Therefore, it is necessary to examine all aspects of cryptojacking, including research issues, characteristics, and detections. The contributions made by this study are as follows:

1) Provides a taxonomy of cryptojacking research interests.
2) Explain the detection approaches in cryptojacking research.
3) Provides analysis techniques and its features in cryptojacking detection.

II. Material and Method

This study compiles and analyses previous research on browser-based cryptojacking detection from 2012 to 2021. Fig. 2 illustrates the data collection process flowchart. Using "Cryptojacking" and "Crypto mining" as our primary search terms, we searched three databases for relevant research papers: ScienceDirect, IEEE Xplore Google Scholar, and Research Gate. We selected these databases because they contain numerous research papers on various topics, thereby enhancing the visibility of research papers. This research disclosed approximately 2,000 records from a variety of article types, including journals, books, and conferences. Nonetheless, almost all of these document's hail from different scientific fields that have only a tenuous connection to fileless cryptojacking.

We employ the keywords "fileless cryptojacking", "Fileless based crypto mining", and "browser-based cryptojacking" to retrieve more relevant data in this area, after removing duplicates and unrelated topics. Seventy highly relevant and high-quality experimental papers on browser-based cryptojacking detection serve as the study's foundation. After locating all published documents from the inputs, we intricately examined and analyzed all these, taking notes on the problem statement and detection technique. The subsequent section displays and discusses our analysis's outcomes.

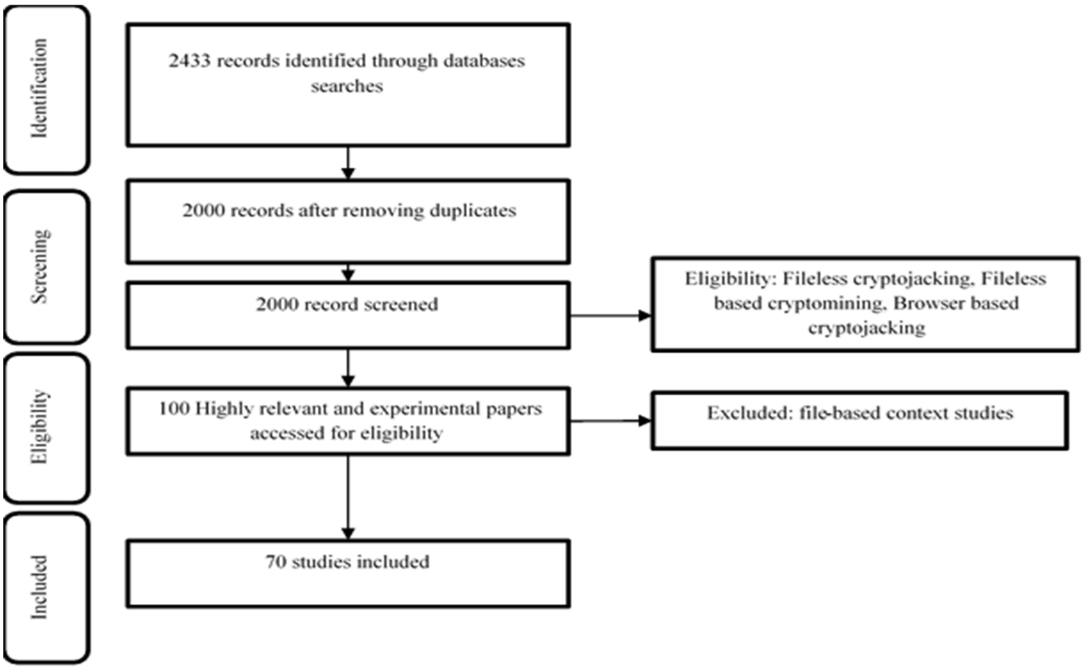
Identification
2433 records identified through databases searches
Screening
2000 records after removing duplicates
2000 record screened
Eligibility: Fileless cryptojacking, Fileless based cryptomining, Browser based cryptojacking
Eligibility
100 Highly relevant and experimental papers accessed for eligibility
Excluded: file-based context studies
Included
70 studies included

Fig. 2 The flow chart of the procedure in collecting data

A. Data Analysis

Upon an assessment of the accumulated documents, this research was able to categorize the research issue and known vulnerabilities detection methods into distinct groups. Categorization requires a comprehensive understanding of the browser-based cryptojacking detection paradigm in terms of trending research issues and detection strategies [9, 10]. This study splits the data assessment into diverse research problems.

B. Research Problems

This paper identifies recent study issues in browser-based cryptojacking to demonstrate the current interest in this topic. Based on the research problems we can extract from the papers, we can categorize them according to the primary research problems they are attempting to solve. This categorization facilitates a greater comprehension of browser-based cryptojacking in light of current research issues. Methods, detection, features, datasets, and code make up the five categories of research problems. The distribution of papers is depicted in Fig. 3 based on various research problem categories. Presented in Fig. 3, the utmost notable diversity is a set that the that most of documents, 36 percent, address issues related to the detection of fileless malware within a system. The second-largest distribution, at 21 percent, consists of papers that focus on available methods and techniques for detecting cryptojacking. The third largest distribution, at 16 percent, consists of papers with either publicly accessible or publicly modified large datasets. The fourth-largest distribution (12 percent) focuses on papers that discuss features, such as CPU, GPU, and memory, to detect and analyze whether a system is under a fileless malware attack or not. The remaining distributions, code (8 percent) and miscellaneous (7 percent), focus on code analysis issues. For a better understanding of the growth of the literature review over time, Figure 4 illustrates the total sum of the various research problem categories between the years of 2011 and 2020.

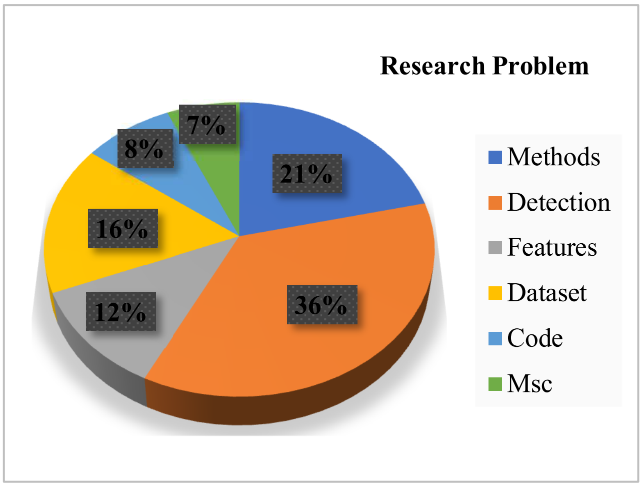

Fig. 3 Distribution of papers according to the research problem categories

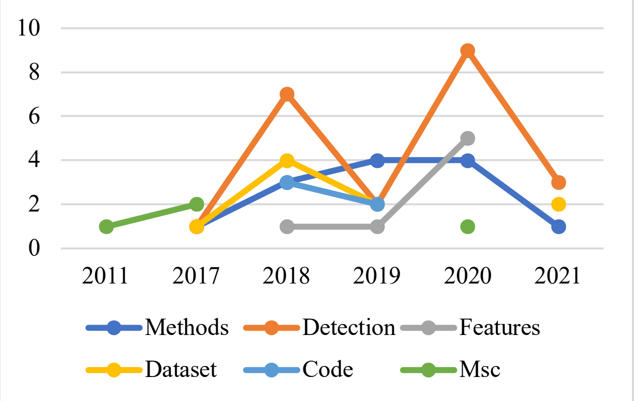

Fig. 4 The sum value of distinct study topic categories between 2011 and 2021

The increasing interest in methods, detection, and characteristics as research topics is depicted in Figure 4. In the features part, only seven papers addressing the issue. This is one of the largest research gaps regarding fileless malware. In contrast, the number of papers devoted to datasets, codes, and miscellaneous research problems has increased marginally over the years, from seven to ten in total. Consequently, these analyses demonstrate that most of the current work in fileless malware detection focuses on methodological and detection-category-related research problems. However, since features are rarely discussed, this could be an intriguing topic to investigate.

C. *Taxonomy of Research Interest in Cryptojacking*

This section reviews the previous articles of cryptojacking investigation. This work suggests a taxonomy of cryptojacking comprised of two categories: a) detection approach; and b) analysis techniques and their characteristics. This taxonomy provides a detailed insight into the current cryptojacking research, as shown in Fig. 5.

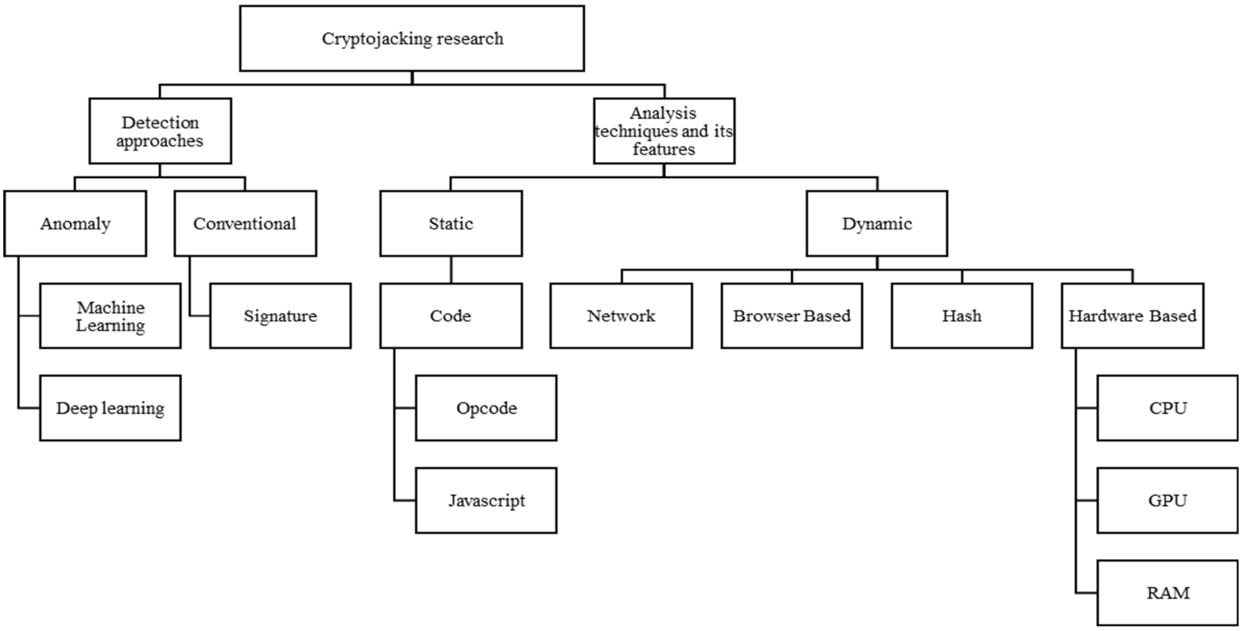

Fig. 5 Taxonomy of cryptojacking interests

TABLE I
DISTRIBUTION OF DETECTION APPROACHES IN RESEARCH PAPERS

Detection Approach	**Classifier**	**Information**
Anomaly (machine learning)	Decision Tree	A Decision Tree is an algorithm in which the training data is constantly divided into subsets based on a predetermined parameter. In this type of learning, the input and output are explicitly described.
	Naïve Bayes	A naive Bayes classifier refers to an algorithm that categorizes things using Bayes' theorem. Strong or naive independence between the attributes of data elements is an assumption made by naive Bayes classifiers.
	Support Vector Machines (SVM)	SVM analyses data for classification and regression analysis.
	Random Forest	Random Forest is a tree-predictor-based technique for supervised ensemble machine learning.
	K Nearest neighbor (KNN)	It considers the k value, or the number of neighbors, which is typically chosen by the user and can significantly impact the algorithm's accuracy. Additionally, KNN requires a distance metric to measure the similarity between data points.
Anomaly (deep learning)	Convolutional neural network (CNN)	A neural network model that enables the extraction of more accurate representations of image content.
	LSTM (Long Short-term memory)	Part of the recurrent neural network (RNN) architecture is designed to handle traditional RNNs' vanishing gradient problem. LSTM networks use special memory cells that can store information over long periods and selectively forget or remember that information when needed. This allows them to capture long-term dependencies in sequential data, such as language, speech, and time-series.
Conventional Approach (signature)	Execution Emulation	The code emulation method of malware detection examines a file's behavior by simulating its execution in a virtual environment.
	Heuristics	Heuristics is a technique designed to solve a problem more quickly when traditional methods are too slow or to find an approximation of a solution when traditional methods cannot find an exact solution.
	YARA	YARA is an open-source solution that lets malware analysts and researchers identify and classify malware.

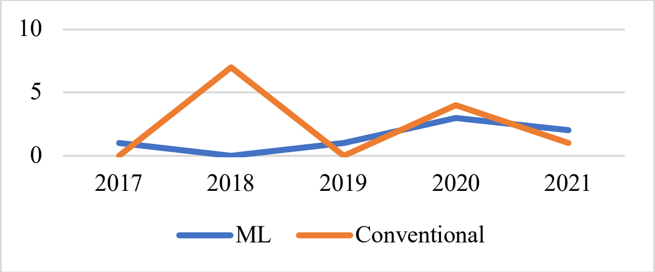

Fig. 6 The accumulated number of various categories of detection strategies from 2017 to 2021

D. Detection Approach

This section explores the current interest in detection approaches in cryptojacking research. The methods are classified into two types: anomaly methods and conventional methods. Then, the category of anomalies is further divided into machine learning and deep learning.

E. Anomaly

Anomaly is a term that refers to something or situation that is abnormal than usual or known as different than others [3], [4]. Hence, an anomaly in cryptojacking detection is an experiment to detect any unusual activities, which can detect new or unknown cryptojacking malware activities [5], [6]. There are two methods for detecting anomalies: machine learning and deep learning [3], [7]Machine learning alludes to a program learning from the training dataset and using techniques and algorithms to perform an action without being explicitly programmed. The phrase "deep learning approach" refers to a sophisticated set of algorithms that utilize information from the human brain.

F. Conventional

Another detection category is the conventional approach, also known as the signature approach [8]. It analyzes and evaluates the signatures of known attacks in the dataset by observing events and looking for intrusions. Unlike anomalies, which merely match known signatures from the current signature database, this method is unable to find undiscovered malware. Table 1 tabulates the descriptions of unconventional and conventional approaches, followed by examples with descriptions. Fig. 6 shows the interest in machine learning (ML) and traditional approaches over the years. It shows that, compared to traditional ML, the utilization of ML has been increasing throughout the years.

III. Result and Discussion

A. Analysis Techniques and Its Features

This section lays down the features extracted from the 70 research papers. Table 2 tabulates the list of features of static and dynamic Static analysis is an experiment that reverse engineers the application and analyses the source code without executing the application [9], [10]. Meanwhile, dynamic analysis is an experiment that perform and execute the application in controlled environment (real hardware, virtual hardware or sandbox) [11]. Table 2 lists that two codes have been used by security practitioners in the detection of cryptojacking: opcode and JavaScript. In dynamic analysis, on the other hand, the dynamic features that have been studied in the past are network, browser, hash, and hardware (CPU, GPU, and RAM) based.

TABLE II
CRYPTOJACKING FEATURES AND THEIR DETAILS

Features	Features Detail
Network and hash [12]	src and dst Ips, src and dst port numbers, protocol, packet size, hash-rate
CPU and RAM [13]	CPU usage, RAM, Average Quadratic Deviation, Operative Memory, CPU Power
Network [14]	Traffic volume and flow times
CPU [15]	CPU usage at user and system kernel level, CPU idle, CPU servicing hardware and software interrupts, and CPU’s hypervisor
CPU [16]	CPU usage
CPU and hash [17]	CPU usage, Usage of WebAssembly and WebWorkers, Hash and URL
CPU and RAM [18]	CPU and memory usage, CPU usage even after the website is closed and less inbound traffic
CPU [19], [20]	CPU utilization
Hash [21]	Hash libraries, cumulative time of websites spent on hashing
Opcode and GPU [22]	Rarely used opcodes
Legitimate User Applications [23], [24], [20]	Usage of legitimate applications to complete the operation
JavaScript code injections [25], [26], [27], [12], [19]	JavaScript injection is a method by which we can input and utilise our own JavaScript code in a page, whether by submitting the form in the address bar or by locating an XSS vulnerability on a website.

B. Challenges in Detecting Cryptojacking

As malware became even more widely recognized and anti-virus programs became more sensitive enough to detect malware through patterns, the criminal element felt the need to make it more difficult to identify these programs as they attempted to infiltrate target systems. Hence, in order to detect cryptojacking, security practitioners need to face challenges and obstacles specific to cryptojacking malware. The challenge is that the criminal adopts a virtual private network (VPN) to hide their server information in order to go undetected [14], [28].

IV. Conclusion

This section discusses future work in cryptojacking detection, which is multi-feature detection. It is important to detect multiple types of attacks and to detect cryptojacking more effectively. This enables the researcher to modify discovered attacks efficaciously and swiftly as they discover known attacks against the feature they pertain to. Consequently, security investigators are no longer needed to train distinct strategies for different software projects.

In conclusion, cryptojacking malware detection occupies a significant position within the system security industry. Since internet usage and vulnerabilities are expanding at a rapid rate, this topic requires extensive investigation. Detection of cryptojacking provides large organizations with a sense of security, as they will be able to prevent numerous system attacks. This research paper outlines all available detection techniques and characteristics. This can aid future researchers in developing new techniques with a unique combination of characteristics. Between 2017 to 2021, the review methodology evaluates a large number of cryptojacking malware detection research papers. This study describes a taxonomy of cryptojacking interests by classifying existing works into two primary groups.: detection approaches (anomaly and conventional), their analysis (static and dynamic), and the types of features associated with each.

Acknowledgment

This work is funded by the collaboration matching grant between Universiti Malaysia Pahang (UMP) (RDU222410) and Kaneka Malaysia (UIC220830).

References

[1] E. Juma Adwan and B. Ali Alsaeed, "Cloud Computing adoption in the financial banking sector-A systematic litreture review (2011-2021)," *International Journal of Advanced Science Computing and Engineering (IJASCE)*, vol. 4, no. 1, pp. 48–55, 2022.

[2] European Union Agency for Cybersecurity, "Cryptojacking - Cryptomining in the browser." Accessed: Apr. 22, 2021. [Online]. Available: https://www.enisa.europa.eu/publications/info-notes/cryptojacking-cryptomining-in-the-browser

[3] D. Nincarean Eh Phon, A. Firdaus, M. F. Ab Razak, S. Kasim, A. H. Basori, and T. Sutikno, "Augmented reality: effect on conceptual change of scientific," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1537–1544, Dec. 2019, doi:10.11591/eei.v8i4.1625.

[4] A. Firdaus *et al.*, "Selecting root exploit features using flying animal-inspired decision," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 7, no. 4, pp. 628–638, 2019, doi:10.11591/ijeei.v7i4.1146.

[5] M. Abbas and H. Ghous, "Early Detection of Breast Cancer Tumors using Linear Discriminant Analysis Feature Selection with Different Machine Learning Classification Methods," *Computer Science & Engineering: An International Journal*, vol. 12, no. 1, pp. 171–186, 2022, doi: 10.5121/cseij.2022.12117.

[6] M. Sajjad, M. Pasha, and U. Pasha, "Parametric Evaluation of E-Health Systems," *International Journal of Information Systems and Computer Technologies (IJISCT)*, vol. 1, no. January, pp. 31–37, 2022.

[7] A. Firdaus *et al.*, "Adaboost-multilayer perceptron to predict the student's performance in software engineering," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1556–1562, 2019, doi:10.11591/eei.v8i4.1432.

[8] M. Sulistiyono, L. A. Wirasakti, and Y. Pristyanto, "The Effect of Adaptive Synthetic and Information Gain on C4. 5 and Naive Bayes in Imbalance Class Dataset," *International Journal of Advanced Science Computing and Engineering (IJASCE)*, vol. 4, no. 1, pp. 1–11, 2022.

[9] A. Karim, V. Chang, and A. Firdaus, "Android botnets: A proof-of-concept using hybrid analysis approach," *Journal of Organizational and End User Computing*, vol. 32, no. 3, pp. 50–67, 2020, doi:10.4018/JOEUC.2020070105.

[10] C. A. Che Yahaya, A. Firdaus, S. Mohamad, F. Ernawan, and M. F. A. Razak, "Automated Feature Selection using Boruta Algorithm to Detect Mobile Malware," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 5, pp. 9029–9036, 2020, doi: 10.30534/ijatcse/2020/307952020.

[11] R. Jusoh, A. Firdaus, S. Anwar, M. Z. Osman, M. F. Darmawan, and M. F. Ab Razak, "Malware detection using static analysis in Android: a review of FeCO (features, classification, and obfuscation)," *PeerJ Comput Sci*, vol. 7, no. e522, pp. 1–54, 2021, doi: 10.7717/peerj-cs.522.

[12] Y. Feng, D. Sisodia, and J. Li, "POSTER: Content-Agnostic Identification of Cryptojacking in Network Traffic," in *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (CCS)*, New York, NY, USA: ACM, Oct. 2020, pp. 907–909. doi: 10.1145/3320269.3405440.

[13] D. Tanana, "Behavior-Based Detection of Cryptojacking Malware," *Proceedings - 2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBEREIT 2020*, pp. 543–545, 2020, doi: 10.1109/USBEREIT48449.2020.9117732.

[14] J. S.-V. and P. B.-R. J. Z. i. Muñoz, "Detecting cryptocurrency miners with NetFlow/IPFIX network measurements," in *IEEE International Symposium on Measurements & Networking (M&N)*, 2019, pp. 1–6. doi: 10.1109/IWMN.2019.8804995.

[15] F. Gomes and M. Correia, "Cryptojacking Detection with CPU Usage Metrics," *2020 IEEE 19th International Symposium on Network Computing and Applications, NCA 2020*, 2020, doi:10.1109/NCA51143.2020.9306696.

[16] S. Eskandari, A. Leoutsarakos, T. Mursch, and J. Clark, "A First Look at Browser-Based Cryptojacking," *Proceedings - 3rd IEEE European Symposium on Security and Privacy Workshops, EURO S and PW 2018*, pp. 58–66, 2018, doi: 10.1109/EuroSPW.2018.00014.

[17] M. Musch, C. Wressnegger, M. Johns, and K. Rieck, "Web-based Cryptojacking in the Wild," Aug. 2018. [Online]. Available: http://arxiv.org/abs/1808.09474

[18] A. Abdul Aziz, S. Ngah, Y. Ti Dun, and T. Fui Bee, "Coinhive's Monero Drive-by Crypto-jacking," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jun. 2020. doi: 10.1088/1757-899X/769/1/012065.

[19] V. S. K. A. Nukala, "Website Cryptojacking Detection Using Machine Learning : IEEE CNS 20 Poster," *2020 IEEE Conference on Communications and Network Security, CNS 2020*, pp. 1–2, 2020, doi:10.1109/CNS48642.2020.9162342.

[20] I. Petrov, L. Invernizzi, and E. Bursztein, "CoinPolice: Detecting hidden cryptojacking attacks with neural networks," *ArXiv*, 2020.

[21] G. Hong *et al.*, "How you get shot in the back: A systematical study about cryptojacking in the real world," *Proceedings of the ACM Conference on Computer and Communications Security*, pp. 1701–1713, 2018, doi: 10.1145/3243734.3243840.

[22] H. Darabian *et al.*, "Detecting Cryptomining Malware: a Deep Learning Approach for Static and Dynamic Analysis," *J Grid Comput*, vol. 18, no. 2, pp. 293–303, Jun. 2020, doi: 10.1007/s10723-020-09510-6.

[23] B. N. Sanjay, D. C. Rakshith, R. B. Akash, and V. V. Hegde, "An Approach to Detect Fileless Malware and Defend its Evasive mechanisms," *Proceedings 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions, CSITSS 2018*, pp. 234–239, 2018, doi: 10.1109/CSITSS.2018.8768769.

[24] Vala Khushali, "A Review on Fileless Malware Analysis Techniques," *International Journal of Engineering Research and*, vol. V9, no. 05, pp. 46–49, 2020, doi: 10.17577/ijertv9is050068.

[25] D. Draghicescu, A. Caranica, A. Vulpe, and O. Fratu, "Crypto-Mining Application Fingerprinting Method," in *International Conference on Communications (COMM)*, IEEE, 2018, pp. 543–546. doi:10.1109/iccomm.2018.8484745.

[26] M. Saad, A. Khormali, and A. Mohaisen, "End-to-End Analysis of In-Browser Cryptojacking," 2018, [Online]. Available: http://arxiv.org/abs/1809.02152

[27] G. Hong *et al.*, "How you get shot in the back: A systematical study about cryptojacking in the real world," in *Proceedings of the ACM Conference on Computer and Communications Security*, Association for Computing Machinery, Oct. 2018, pp. 1701–1713. doi:10.1145/3243734.3243840.

[28] M. Caprolu, S. Raponi, G. Oligeri, and R. di Pietro, "Cryptomining makes noise: Detecting cryptojacking via Machine Learning," *Comput Commun*, vol. 171, pp. 126–139, Apr. 2021, doi:10.1016/j.comcom.2021.02.016.