IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141126

AI-POWERED MALWARE DETECTION SYSTEM

Shubham N. Bawa[1], Prof. Pravin I. Patil [2], Prof. Manoj V. Nikum*[3]

Student, MCA Department, SJRIT DONDAICHA, KBC NMU JALGAON, Maharashtra, India[1]

Assistant Professor, MCA Department, SJRIT DONDAICHA, KBC NMU JALGAON, Maharashtra, India.[2]

Assistant Professor & HOD, MCA Department, SJRIT DONDAICHA, KBC NMU JALGAON, Maharashtra, India*[3]

Abstract: The rapid growth of cyber threats has left traditional signature-based malware detection methods less effective against new and complex attacks. To address this issue, the AI-Powered Malware Detection System identifies malicious software using machine learning, focusing on behavior instead of static signatures. Developed in Python, the system detects and classifies malware and non-malicious software using algorithms like Random Forest and XG Boost. By training on datasets such as the Microsoft Malware dataset and CIC-MalMem, the model identifies complex patterns in system behavior, including file operations, network activity, and process interactions associated with malware. The features extracted are then processed to create a high-performance model that detects malware with low false positives. This system is also resilient to future variant developments, making it more effective than traditional methods. With applications in cybersecurity defense systems, enterprise IT infrastructure, and cloud security, this paper enhances proactive malware detection and improves system resilience against cyberattacks.

Keywords: Malware, cloud security, cyberattacks, Ai, machine learning, cybersecurity.

I. INTRODUCTION

Cyber-security is a collection of technological procedures and practices that guard data, hardware, software, and networks against damage, intrusion, and unauthorized access. Malware is a set of malicious programming codes or scripts and intrusive software that is built to destroy targeted computer systems and programs or mobile and web applications using different forms including computer viruses, worms, ransomware, rootkits, Trojans, dialers, adware, spyware and key loggers. The first computer-based virus was discovered on Apple II machines called "Elk Cloner" in 1982, developed by a 15-year-old high school student Rich Skrenta. Basit Farooq Alvi and Amjad Farooq Alvi who were two brothers and wanted to prove that PC is not immune, wrote a PC-based stealth virus called "Brain" in 1986 [2].

These viruses were able to replicate with the use of floppy disks by inserting the infected. Cyber-attacks often target sensitive information and lead to financial losses which severely defames the impacted organizations. When Cyber-attacks target critical infrastructure like healthcare systems, they may even turn into life-threatening consequences. Artificial Intelligence helps in advancements regarding malware detection and prevention while providing opportunities to develop robust, efficient, and scalable malware recognition modules. However, questions around data privacy, algorithmic accountability, and the potential for misuse have been existing around the corner, and it goes without saying that the adoption of these advanced technologies is not without growing areas of concern. That is crystal clear; the traditional techniques are not enough to protect us from this evolving landscape of cyber threats, and therefore, turning towards more intelligent systems looks almost inevitable. It is not just a theoretical fact but has real-world implications which can severely affect organizations, individuals, and even nations. AI's tailored user education approach also holds great potential for creating a security-aware workforce by customizing training materials to meet unique learning requirements and knowledge gaps. This paper is going to cover the comprehensive report of technical advancements and their practical applications, and also critically engages the limitations, challenges, and ethical considerations involved in adopting these new technologies for cybersecurity.

II. RESEARCH METHODOLOGY

In this study, we surveyed to investigate the current role of applications of AI and ML in malware detection for achieving cyber-security. A collection of various research papers, books, and conference proceedings was analyzed during this survey. We mainly focused on literature published between 2018 and 2024 to ensure coverage of only recent developments in the respective field. The selected literature was then thoroughly reviewed to extract information pertinent to some of our research questions which are as follows [7]:

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.141126

- What are the major challenges for malware detection using AI?
- What are the developing technologies used by malware authors?
- How is sophisticated malware impacting static and dynamic analysis?
- What are the limitations of existing malware repositories?
- What features are optimal for training an AI model?
- Which Al models are most successful for malware detection and what are their advantages and limitations?

An extensive literature search was conducted in the ACM Digital Library, Google Scholar, IEEE Xplore, and Scopus database. ACM Digital Library provided access to Association for Computing Machinery (ACM) journals, proceedings, and conferences. Similarly, IEEE Xplore was chosen because it provides access to Institute of Electrical and Electronics Engineers (IEEE) conference papers and journals. Both ACM and IEEE are predominant in cyber-security. Scopus provided the most extensive database for the related subject. We also filtered publication topics including Systems and Data Security for Springer Link and Publication Topics for IEEE Explore and Computer Science and Security for Science Direct. Search keywords were mainly malware detection, artificial intelligence, and machine learning joined with offensive, network security, information security, and cyber-attack. An approx. of 235 studies were found during the initial search. Once the search processes were completed, we had to go through a screening process for finding relevant papers based on the paper title at first followed by reading and understanding the abstract and conclusion from screened papers. The results were limited by cutting off snowballing for articles older than 2018 to focus on the latest research in a rapidly evolving field. The articles included were needed to contain descriptions of malware based on machine learning or Al functionality and thereafter the literature was analyzed with respect to the application areas of AI and ML in cyber-security. Thorough analysis identified the most relevant AI/ML techniques and their applications were studied in areas such as intrusion detection, malware detection, security automation, network security, threat intelligence, security management, cyber-attack prediction, vulnerability management, and associated awareness education.

III. LITERATURE REVIEW

Aggravation in the digital economy and infrastructure due to technological advancements results in a visible increase in cyber-attacks that could carry grave implications. The majority of cyber-attacks are made up of Malware attacks, accounting for 43% of the total, with 5.6 billion attacks [3]. Some of the Malware that has been discovered in the past are The Nimda, Morris Worm, Sasser, ILOVEYOU, Welchia, Melissa, Code Red, Slammer, Commwarrior-A, Crypto Locker, and Stuxnet. These computer viruses can propagate by download, malicious intent, installation of commercial software, or even by clicking a predefined link and thus affect any government, data center, commercial, laboratory, enterprise, or organizational software application. Identifying, preventing, detecting, addressing, and recording cyber-attacks to avoid future security breaches, advisory organizations like the National Institute of Standards and Technologies (NIST) are regularly advising to use of more proactive and adaptive approaches by moving towards real-time assessments, data-driven analysis, and continuous monitoring. According to the "2021 Sonic Wall Cyber Threat Report," there was an increase in 62% of global ransomware attacks in 2020, with over 304 million ransomware attacks. Thus there is a need for effective cyber security measures, driving the development of innovative techniques such as Artificial Intelligence (AI) and Machine Learning (ML) algorithms. Distribution of malware from January 2022 to June 2022 was observed as 6% IoT malware, 19% crypto-jacking, 28% malicious intrusions, 53% malicious office and PDFs, 105% ransomware, and 167% encrypted threats.

The mapping of the state of Al-powered malware has also been performed to recognize AI-enhanced attacks carried out by malware. Malware types that conceal themselves from detection using AI techniques could also be identified, to get a better understanding of the maturity of those attacks, and to develop the algorithms and methods involved in those attacks.

A. Signature-Based Detection

This method relies on known digital indicators as a database of known malware signatures to identify suspicious behavior or threats. Whenever a software piece matches a signature in the database, the system alerts it as malicious. This technique is only effective for finding known malware and is not efficient with polymorphic malware. Therefore a list of indicators of compromise (IOCs), is supposed to be maintained in a database, which can eventually be used to identify a breach.

B. Static File Analysis

This technique refers to the examination of the code of a file, without running it, to identify any kind of malicious content. This evaluation of finding malicious objects could be done on file names, strings such as IP addresses, hashes, and file header data [4]. Proficient security teams are now using additional techniques to alert about advanced malware that might go unidentified during normal static analysis.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.141126

C. Dynamic Malware Analysis/ Sandboxing

This analysis is a closed system that enables security professionals to analyze and execute suspected malicious code in a safe environment called a sandbox [4]. Intel PT (Processor Trace) is applied in generation sandboxing to access the full execution flow of the potentially malicious artifact and analyze it using a complete "trace" alongside examining changes to virtual memory during execution. The process helps in studying malware without the risk of infection into their system or letting them escape into the enterprise network.

D. Heuristic Analysis

It does malware detection of new or modified malware by analyzing the behavior of a program or code that might not have a known signature. As soon as the software is guessed to exhibit characteristics similar to malware, it alerts it as a potentially malicious object.

E. Machine Learning and Artificial Intelligence

This technology can detect previously unknown threats and adapt to new malware variants while analyzing huge amounts of data and identifying patterns to classify them as favorable or malicious.

F. Check summing/Cyclic Redundancy Check (CRC)

This technique offers calculation on a collection of data, such as a file, to confirm its integrity and can be effective for identifying corruption in data. CRC is one of the most common checksums used that involves analysis of both the position and value of a group of data.

G. Honeypots

This system is designed to mimic a software application that entices targets to malware. As soon as they get infected by the malware, security professionals can study it and design defenses to address these specific vulnerabilities or threats accordingly for their real system.A malware honeypot is similar to an application programming interface (API) that draws out malware attacks in a controlled and non-threatening environment.

H. Intrusion Detection and Prevention Systems (IDS/IPS)

Suspicious activities on network traffic are monitored and thus alert the administrators of potential security breaches, while detected threats are blocked by IPS. Both can be host-based (HIDS/HIPS) or network-based (NIDS/NIPS) [5].

IV. RESEARCH ISSUES

The increasing adoption of AI and ML in cyber security has many associated challenges and barriers to its implementation [7]. Some of the most frequently reported challenges are said to be the lack of understanding of the technology, reported to be 36.9%, shortage of skilled personnel, reported to be 34%, and High costs as a significant barrier to adoption, with 29.1% of total surveyed organizations. One of the most common issues is that machine learning might generate false positives or false negatives, which ultimately reduces its reliability and efficiency leaving out concerns for the researchers working in the respective field.

V. CONCLUSION AND FUTURE WORK

Artificial Intelligence (AI) and Machine Learning (ML) have already begun to reshape the cyber-security landscape and have the potential to revolutionize the field in multiple dimensions. Windows-based systems are the major targets of malware by cyber attackers. Techniques that are most effective for detecting malware attacks have to be Machine learning and deep learning. By understanding different types of malware accompanied by the engagement of a multi-layered approach to security, organizations can remarkably limit the risk of falling victim to a malware attack. However, Al could be less effective if it relies only on historical data, which might reduce its impacts and further prevent it from adapting to innovative attack methods. Artificial intelligence still is dependent on human knowledge or interference since it finds it difficult to envision contextual differences and can be mistaken in interpreting user behaviors and intentions. Ultimately, despite Al's tremendous potential, overcoming its present limitations requires careful balancing. While focusing on the future of AI and ML in cyber-security, several fascinating research paths are emerging [8][9][10][11][12][13]. The other advanced technologies like block chain and quantum computing when integrated with AI and ML will provide the potential to promise an avenue for further exploration. Future work must also delve into the ethical considerations, which would aim to formulate strategies for responsible usage and transparent decision-making processes while not just focusing entirely on the technical aspects. Incident response, proactive threat hunting, and disaster recovery are areas that are currently ripe for AI and ML applications, which are opening new frontiers in cyber-security measures. Al's processing power has made it feasible to identify potential threats in advance, and its tailored advice promotes a Cyberwar

DOI: 10.17148/IJARCCE.2025.141126

culture. Thus it is very evident that progress in this field is not going to happen without encountering its challenges. Biases, adversarial flaws, and false positives can undermine effectiveness and confidence. The right mix between AI's

REFERENCES

[1]. Daniel Gibert, Carles Mateu, Jordi Planes., The rise of machine learning for detection and classification of malware: Research developments, trends and challenges, Journal of Network and Computer Applications Volume 153, 1 March 2020, 102526 https://doi.org/10.1016/j.jnca.2019.102526
[2]. Kaspersky: A Brief History of Computer Viruses & What the Future Holds
[3]. Gary Smith, April 10, 2024 +95 Cyber Security Breach Statistics 2024, station
[4]. Kurt Baker, Malware Analysis, April 17, 2023: crowdstrike
[5]. Perception Point Malware Detection: 7 Methods and Security Solutions that Use Them
[6]. Mohamed, Engineering Cogent (2023), 10: 2272358https://doi.org/10.1080/23311916.2023.2272358 https://doi.org/10.1080/23311916.2023.2272358
[7]. Matthew G. Gaber, Mohiuddin Ahmed, and Helge Janicke. 2024. Malware Detection with Artificial Intelligence: A Systematic Literature Review. ACM Comput. Surv. 56, 6, Article 148 (January 2024), 33 pages. https://doi.org/10.1145/3638552
[8]. Abdullahi, M., Baashar, Y., Alhussian, H., Alwadain, A., Aziz, N., Capretz, L. F., & Abdulkadir, S. J. (2022). Detecting cybersecurity attacks in internet of things using artificial intelligence methods: A systematic literature review. Electronics, 11(2), 198. https://doi.org/10.3390/electronics 11020198
[9]. 9. Gupta, S., Sabitha, A. S., & Punhani, R. (2019). Cyber Security Threat Intelligence using Data Mining Techniques and Artificial Intelligence. In International Journal of Recent Technology and Engineering pp. 6133-6140). Issue 3. (IJRTE) (Vol. 8. https://doi.org/10.35940/ijrte.c5675.098319
[10]. R.Sri Devi, M. Mohan Kumar, Cyber Security Affairs in Empowering Technologies. (2019). In International Journal of Innovative Technology and Exploring Engineering (Vol. 8, Issue 10S, pp. 1-7). https://doi.org/10.35940/ijitee.j1001.08810s19
[11]. Saudi, M. M.. Sukardi, S., Abd Aziz. N. A. A., Ahmad, A., & Husainiamer, M. Afif. (2019). Malware Classification for Cyber Physical System (CPS) based on Phylogenetics. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 1, pp. 3666-3670). https://doi.org/10.35940/ijeat.a2711.109119
[12]. Joshma K J, & Sankar P, V. (2024). Phishing Website Detection. In Indian Journal of Data Mining (Vol. 4, Issue 1, pp. 38-41). https://doi.org/10.54105/ijdm.a1642.04010524
[13]. Rathore, R., & Shrivastava, Dr. N. (2023). Network Anomaly Detection System using Deep Learning with Feature Selection Through PSO. In International Journal of Emerging Science and Engineering (Vol. 11. Issue 5. pp. 1-6). https://doi.org/10.35940/ijese.f2531.0411523