Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 02, NCRTCS - 2026

A Study of the Relationship Between Malware Detection Methods and Cyber Attack Prevention

Understanding the Role of Malware Detection in Preventing Cyber Threats

Asst Prof Ashwini Yogesh Dhanave
Dept of Computer Science
ASM's CSIT Pimpri, Pune - 411007

***Abstract* - This study explores the critical causal relationship between malware detection methodologies and the efficacy of cyber attack prevention. By analyzing the transition from reactive signature-based detection to proactive AI-driven behavioral analysis, the research identifies how detection speed and accuracy serve as the primary variables in successful threat neutralization.**

Keywords - Malware detection, Cyber-attack prevention, Cyber threats Detection methods.

INTRODUCTION

In today's world, cyber threats are growing rapidly, and malware is becoming more advanced and difficult to detect. Because of this, traditional security methods are no longer enough. This research focuses on combining dynamic malware analysis, Software Defined Networks (SDNs), and machine learning to build stronger network security.

SDNs provide a centralized and programmable network structure, which makes it easier to monitor and control network traffic. This gives security team's better visibility into what is happening inside the network. By integrating dynamic malware analysis into SDN-based networks, security systems can study malware behavior in a flexible and controlled way, helping to identify new and evolving threats early.

Machine learning further improves this approach by enabling automated malware detection and response. It helps security systems recognize suspicious behavior and react quickly without human intervention. Malware is analyzed in isolated environments within the SDN so that its behavior can be closely observed without risking the safety of the entire network.

Using intelligent traffic control, SDN controllers route suspicious network traffic to specific security checkpoints for detailed inspection. Machine learning algorithms then analyze malware behavior during execution by monitoring system calls, file changes, and network communication patterns. This allows the system to understand how the malware operates and how it spreads.

The combination of SDNs and machine learning makes it possible to detect previously unknown malware by identifying unusual behavior patterns. The study also highlights the importance of using real-time threat intelligence, which helps SDN controllers respond quickly to new cyber threats. Based on the analysis, SDN controllers can automatically update network rules, isolate infected devices, or block harmful traffic.

Another important aspect of this framework is continuous learning. Machine learning models are regularly updated with new data so they become more accurate and adaptable over time. This ensures that the security system stays effective even as malware techniques change.

In conclusion, this research presents an integrated security framework that uses the strengths of dynamic malware analysis, SDNs, and machine learning. By combining SDN's flexible network control with machine learning's analytical power, the proposed approach aims to provide real-time detection and prevention of malware in modern network environments.

Dynamic malware analysis refers to studying malicious software in a safe, controlled environment—such as a sandbox—to understand how it behaves and what damage it can cause. Software Defined Networks (SDNs) allow centralized and programmable network management, offering improved control, monitoring, and security visibility.

RELATED WORK-

In today's digital world, malware has become a serious and growing threat. There are many different types of malware found on the internet, and research shows that malware has evolved rapidly over the past ten years. This rapid growth has caused major financial losses to individuals and organizations.

Malware is a harmful type of software that can severely damage a user's computer system. It can affect the system in several ways, such as slowing down performance, stealing data, or causing system crashes. Because of this, identifying malware at an early stage is very important.

Many studies propose the use of Artificial Intelligence (AI) techniques to detect whether a file downloaded from the internet is malicious or safe. These methods help identify different types of malware before they infect the user's system. The proposed approach is capable of detecting various malware types, including Adware, Trojans, Backdoors, Obfuscated malware, Multidrop, Robot, Spam malware, and Ransomware.

Malicious software is widespread due to the large number of computer users who are constantly exposed to threats from sources such as the internet, local networks, and portable storage devices. Depending on its severity, malware can cause minor issues or serious damage, such as data theft, system malfunction, or complete system failure.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 02, NCRTCS - 2026

- Malware often appears as executable files or system

library files and includes threats such as viruses,worms,and Trojans. All of these aim to break system security and violate user privacy. Today, malware is one of the most common and dangerous cyber attacks.

Signature-Based Detection This method relies on a database of known malware signatures to identify suspicious behaviour or threats, utilising known digital indicators. 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 maintained in a database, which can eventually be used to identify a breach. B. Static File Analysis This technique involves examining the code of a file without running it to identify any 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 standard static analysis. 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 artefact and analyse it using a complete "trace" alongside examining changes to virtual memory during execution. The process helps in studying malware without risking infection on their system or allowing it to escape into the enterprise network. D. Heuristic Analysis It detects malware by analysing the behaviour of new or modified malware, which may not have a known signature. As soon as the software exhibits characteristics similar to those of malware, it is flagged as a potentially malicious object. E. Machine Learning and Artificial Intelligence This technology can detect previously unknown threats and adapt to new malware variants by analysing vast amounts of data and identifying patterns to classify them as either favourable or malicious. F. Checksumming/Cyclic Redundancy Check (CRC) This technique involves calculating a checksum of a collection of data, such as a file, to verify its integrity and can be effective in identifying data corruption. CRC is one of the most common checksums used, which involves analysing 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 defences 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, alerting administrators to potential security breaches, while the IPS blocks detected threats. Both can be host-based (HIDS/HIPS) or network-based (NIDS/NIPS) [5]. I. Endpoint Protection Platforms (EPP) EPP offers centralised management and protection of endpoint devices, including desktops, laptops, and mobile devices, by providing anti-malware, antivirus, and other security features such as application control, device control, and data loss prevention (DLP). To enhance malware detection, Artificial Intelligence (AI) and Machine Learning (ML) techniques have been employed, with deep learning being a prominent approach within AI. Advanced deep learning algorithms, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), are trained on vast datasets, ultimately excelling in identifying and categorising malware with the highest precision. Furthermore, for enhancing training capabilities, generative models such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) are incorporated to create high-quality synthetic data that mimics real-world patterns. Remarkable progress has been achieved by AI and ML in strengthening measures for network security by offering advanced capabilities to detect and counteract threats such as unauthorized network infiltrations, Distributed Denial of Service (DDoS) attacks, and Advanced Persistent Threat (APT) cyberattacks with increased efficiency and accuracy [6]. Fig.2.2. A Comprehensive Malware Detection Strategy Involves Multiple Layers of Protection.

III. RESEARCH METHODOLOGY In this study, we surveyed to investigate the current role of AI and ML applications in malware detection for achieving cybersecurity. A collection of various research papers, books, and conference proceedings was analysed 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]: ▪ What are the significant challenges for malware detection using AI? ▪ What do malware authors use to develop these technologies? ▪ 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 AI 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 databases. ACM Digital Library provides access to Association for Computing Machinery (ACM) journals, proceedings, and conferences.

IV. RESEARCH ISSUES The increasing adoption of AI and ML in cybersecurity presents numerous challenges and barriers associated with their implementation [7]. Some of the most frequently reported challenges are the lack of understanding of the technology (36.9%), shortage of skilled personnel (34%), and High costs as a significant barrier to adoption (29.1%) among the total surveyed organisations. Malware Detection Using Artificial Intelligence: Techniques, Research Issues and Future Directions 4 Published By: Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) One of the most common issues is that machine learning may generate false positives or false negatives, which ultimately reduces its reliability and efficiency, leaving concerns for researchers working in the respective field.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 02, NCRTCS - 2026

Significant actions that are needed to address challenges faced by this technology are as follows:

V.CONCLUSION AND FUTURE WORK Artificial Intelligence (AI) and Machine Learning (ML) have already begun to reshape the cybersecurity landscape and have the potential to revolutionise the field in multiple dimensions. Windows-based systems are a significant target of malware by cyber attackers. Techniques that are most effective for detecting malware attacks include machine learning and deep learning. By understanding different types of malware and employing a multi-layered approach to security, organisations can significantly reduce the risk of falling victim to a malware attack. However, AI could be less effective if it relies solely on historical data, which may reduce its impact and hinder its ability to adapt to innovative attack methods. Artificial intelligence is still dependent on human knowledge or intervention, as it struggles to recognise contextual differences and can misinterpret user behaviours and intentions. Ultimately, despite AI's tremendous potential, overcoming its present limitations requires careful balancing. While focusing on the future of AI and ML in cybersecurity, several fascinating research paths are emerging .Other advanced technologies, such as blockchain and quantum computing, when integrated with AI and ML, will provide a potential avenue for further exploration. Future work must also explore the ethical considerations, aiming to formulate strategies for responsible usage and transparent decision-making processes, while not focusing solely on 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 cybersecurity measures. AI's processing power has made it feasible to identify potential threats in advance, and its tailored advice promotes a culture of cyberwarfare. Thus, it is evident that progress in this field will not occur without encountering its challenges. Biases, adversarial flaws, and false positives can undermine effectiveness and confidence. The right balance between AI's advantages and human abilities must be found to optimise its benefits and minimise its disadvantages. The development of artificial intelligence (AI) in cybersecurity has been examined in various roles, categories of solutions, specific use cases, and types of AI approaches.

3. PROPOSED METHODOLOGY

- The proposed methodology introduces a high-level network security framework that uses **Artificial Intelligence (AI)** to improve threat detection and response. The main goal is to overcome the limitations of traditional security methods and strengthen overall protection against cyber threats.
- This framework combines multiple AI models, including **supervised and unsupervised learning techniques**, to create a hybrid detection approach. Using more than one model helps improve accuracy and reduces the weaknesses of individual methods.
- To ensure reliability, the system is designed to be strong against **adversarial attacks**, where attackers try to trick AI models. The AI models are regularly updated and trained using adversarial training techniques so they can better resist manipulation and evolving threats.

  The methodology also includes **automated response mechanisms** to quickly handle detected threats. These automated actions may include isolating infected systems, blocking malicious traffic, or updating security rules. Such rapid responses help reduce reaction time and limit the overall impact of security incidents.

Data processing Data processing in malware detection involves the systematic analysis and manipulation of raw data to extract meaningful insights for identifying and combating malicious software threats. This process encompasses various stages, including data cleaning to remove noise and inconsistencies, feature extraction to capture relevant characteristics of malware samples, and feature engineering to enhance the discriminatory power of extracted features. Additionally, data labeling assigns class labels to indicate whether a sample is malicious or benign, enabling supervised learning algorithms to train and evaluate effectively. Data processing also includes techniques such as data augmentation to increase dataset diversity and splitting the data into training, validation, and test sets to assess model performance accurately. By carefully processing and preparing the data, cyber-security professionals can develop robust detection models capable of accurately identifying and mitigating malware threats..

1) Data Collection

In this research, data collection is carried out using multiple techniques to ensure accurate and reliable malware detection. Malware samples are collected from trusted malware repositories and threat intelligence sources, which provide verified examples of known malware variants. These samples are analyzed to extract key features such as file hashes, file attributes, API calls, and byte patterns, which help in identifying and classifying malware behavior.

Along with malware samples, benign (safe) software samples are also collected to create a balanced dataset. This balanced dataset is essential for effectively training and evaluating the malware detection models used in this study.

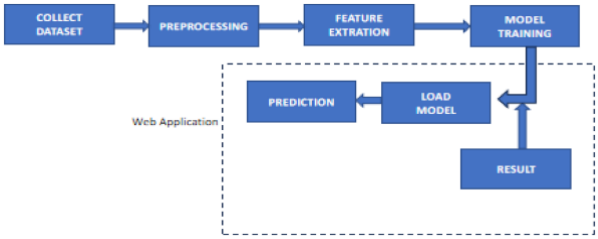
To gain deeper insights into malware behavior, dynamic analysis techniques such as sandboxing and emulation are employed. Malware samples are executed in a controlled and isolated environment, allowing observation of their runtime behavior without affecting real systems. During this process, activities such as file modifications, system library changes, process injections, and network communications are monitored and recorded.

The combination of static and dynamic data collection techniques provides a comprehensive understanding of malware characteristics and behavior. This collected data is then used for analysis, training machine learning models, and evaluating the effectiveness of malware detection in preventing cyber attacks.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 02, NCRTCS - 2026

Block diagram



Block diagram of the proposed system is shown Figure

3.4 Feature Extraction

Feature extraction in malware detection is the process of identifying and selecting important characteristics from malware samples that help in understanding and recognizing malicious software. These features describe how the malware behaves, how it is structured, and what properties it has. By using these features, machine learning and AI models can effectively differentiate between harmless (benign) files and harmful (malicious) files.

The feature extraction process depends on the type of data being analyzed. It may include **static features**, such as file attributes examined without executing the program; **dynamic features**, which observe the behavior of the malware while it is running; and **network-based features**, which analyze communication patterns in network traffic.

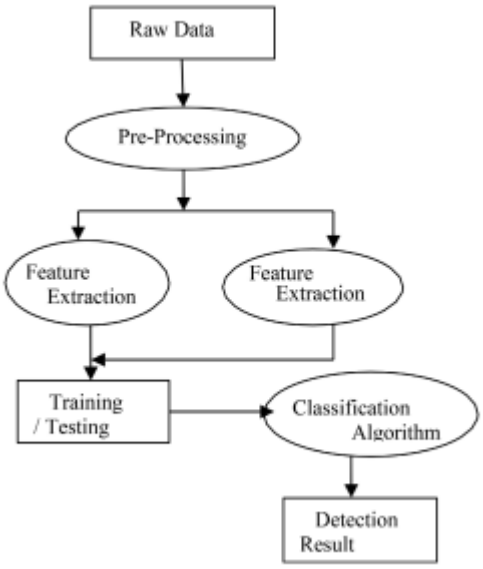
Commonly extracted features from malware samples include file hashes, file size, file type, specific API calls used by the program, embedded strings, byte sequences, code structure, and metadata. These features provide meaningful information that helps improve the accuracy of malware detection systems.

3.5 Model Training and Model Loading in Malware Detection

In malware detection, **model training** refers to the process of teaching a machine learning algorithm to identify malicious and benign software using labeled datasets. During training, the algorithm analyzes the extracted features from malware samples and learns patterns that distinguish malicious software from normal programs.

This process involves repeatedly providing training data to the model and adjusting its internal parameters to reduce errors in prediction. The goal is to minimize the difference between the model's predictions and the actual labels. Common machine learning techniques used for malware detection include decision trees, random forests, support vector machines (SVM), and neural networks.

Once the model is trained and performs well on the training data, it is evaluated using separate validation and test datasets. This step ensures that the model can accurately detect previously unseen malware samples. After successful evaluation, the trained model can be loaded and used in real-time malware detection systems to identify potential threats.



3.7 Prediction

Prediction in malware detection using the Gradient Boosting Algorithm (GBA) refers to using the trained model to determine whether a given file or software sample is malicious or benign. Once the GBA model has been trained on labeled data and optimized to reduce prediction errors, it can be used to classify new and previously unseen samples.

During the prediction phase, the features extracted from an input file are provided to the trained GBA model. The model analyzes these features and generates a probability or confidence score indicating the likelihood that the sample is malicious. A predefined decision threshold is then applied to this score to classify the sample as either malicious or benign.

Gradient Boosting is well suited for malware detection because it can handle complex datasets and identify subtle patterns in data. This capability is especially important in cybersecurity, where malware often closely resembles legitimate software. By using the predictive strength of GBA, cybersecurity professionals can improve real-time malware detection and response, thereby enhancing the overall security of computer systems and networks.

Table 1. Gradient boosting.

ALGORITHM	ACCURACY	PRECISION	RECALL
GRADIENT BOOSTING	99	99	98.5
SVC	91	90	92
LSTM	94	94	95

From **Table 1**, it is observed that the **Gradient Boosting Algorithm (GBA)** outperforms the other two algorithms used in the comparative analysis. In terms of evaluation metrics such as accuracy, precision, and recall, the proposed Gradient Boosting model achieves **99% accuracy**, **99% precision**, and **98.5% recall**, demonstrating its superior performance in malware detection.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 02, NCRTCS - 2026

1) Results and Discussion

The overall results of this study indicate that the Gradient Boosting Algorithm is a highly effective approach for malware detection. Its strong predictive capability and consistent performance enable more accurate identification of malicious software when compared to other models. These results suggest that GBA can significantly enhance the ability of cybersecurity professionals to detect, analyze, and mitigate malware threats in an increasingly complex threat environment.

Furthermore, the findings encourage continued research in this domain to further improve GBA-based malware detection systems. Future work may focus on advanced feature engineering techniques, optimization methods, and ensemble learning strategies to enhance detection accuracy and computational efficiency. The performance comparison of the algorithms is illustrated in **Figure** .

CONCLUSION

All things considered, the blend of man-made intelligence into network security tends to be a momentous method for managing to address the consistently creating scene of computerized risks. The advantages introduced by computer-based intelligence advancements are critical, Table 1. Gradient boosting. 3. Performance analysis. outfitting relationship with updated capacities with regards to risk acknowledgment, progressing checking, and adaptable response frameworks. As we investigate the complexities of the state-of-the-art electronic environment, obviously a reliance solely on standard organization security measures is insufficient. Artificial intelligence adds to an adjustment of standpoint, engaging a more proactive and dynamic insurance technique.

REFERENCES

[1] Gomez-Rodriguez, J.R.; Sandoval-Arechiga, R.;
[2] Ibarra-Delgado, S.; Rodriguez-Abdala, V.I.; Vazquez-Avila, J.L. and Parra-Michel, R. (2021).
[3] A survey of software-defined networks-on-chip: Motivations, challenges and opportunities. Micro Machines, 12, 183. [Google Scholar] [CrossRef] [PubMed]. [7]
[4] Ruaro, M.; Caimi, L.L. and Moraes, F.G. (2020). A systemic and secure SDN framework for NoCbasedmany-cores. IEEE Access, 8, 105997–106008 [8]
[5] Ruaro, M.; Caimi, L.L. and Moraes, F.G. (2020). SDN-based secure application admission and execution for many-cores. IEEE Access, 8, 177296–177306.names; do not use “et al.”.
[6] V. Mythily, P. Sukumar, V. Akshaya, S. Dinesh, Harivardhini & M. Nandhini-Malware detection and prevention using machine learning
[7] 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
[8] Gary Smith, April 10, 2024: +95 Cyber Security Breach Statistics 2024, station
[9] Perception Point: Malware Detection: 7 Methods and Security Solutions that Use Them
[10] D. Du, Y. Sun, Y. Ma , F. Xiao,” A Novel Approach to Detect Malware Variants Based on Classified Behaviors”, IEEE, 81770 – 81782 ,2019.
[11] Uikey, R., Gyanchandani, M., “Survey on Classification Techniques Applied to Intrusion Detection System and its Comparative Analysis”.
[12] International Conference on Communication and Electronics Systems,2019.
[13] ”A survey on machine learning-based malware detection in executable files”. Journal of Systems Architecture, vol 112, 2020.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)