International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 45-49.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19321972


A Hybrid Approach to Real Time Malware Detection using YARA Rules

R Prapulla Kumar, D Udaya Lakshmi, M Sri Valli, Y Keerthi, K Chandhu, Y V Praveen Kumar

Department of CSE-AI, PBR Visvodaya Institute of Technology and Science, Kavali, A.P, India

To Cite this Article
R Prapulla Kumar, D Udaya Lakshmi, M Sri Valli, Y Keerthi, K Chandhu & Y V Praveen Kumar (2026). A Hybrid Approach to Real Time Malware Detection using YARA Rules. International Journal for Modern Trends in Science and Technology, 12(04), 45-49. https://doi.org/10.5281/zenodo.19321972

Article Info
Received: 28 February 2026; Revised: 18 March 2026; Accepted: 22 March 2026.

KEYWORDS

Malware Detection, YARA Rules, Machine Learning, Deep Learning, Behavioral Analysis, Real-Time Monitoring

ABSTRACT

The rapid evolution of modern malware presents significant challenges to traditional cybersecurity mechanisms. Signature- based detection systems are ineffective against zero-day attacks, polymorphic malware, metamorphic malware, and fileless threats. To address these limitations, this project proposes a hybrid real-time malware detection system integrating YARA rule-based static analysis with Machine Learning (ML) and Deep Learning (DL) techniques. The system continuously monitors system activities including file operations, process behavior, memory access, and network traffic. YARA rules are used to detect known malware through signature matching, while intelligent ML and DL models analyze behavioral patterns to identify unknown threats. The hybrid approach improves detection accuracy, reduces false positives, and ensures real-time threat mitigation.

Experimental evaluation demonstrates that the proposed hybrid framework significantly outperforms traditional standalone detection systems.

INTRODUCTION

The rapid growth of sophisticated malware variants has exposed the weaknesses of traditional antivirus systems that depend primarily on static signature matching. These systems are unable to adapt to newly emerging. Threats, particularly zero-day attacks and obfuscated malware strains. As a result, modern cybersecurity frameworks require intelligent and adaptive detection mechanisms.[3]Recent advancements in deep learning have significantly improved malware classification accuracy by enabling atomatic feature extraction from datasets.[4] Machine learing techniques have also proven effective in identifying abnormal system behavior and detecting malicious patterns. [1]Convolutional Neural Networks(CNN) and other advanced models provide automated learning capabilities, while rule-based tools such as YARA ensure precise signature matching for known threats.[2]Combining these approaches offers a

promising direction for real-time malware detection systems.

Objective

The main goal of this project is to develop and implement a hybrid real-time malware detection system that integrates YARA-based static analysis with machine learning and deep learning techniques. The system is designed to detect both known malware through signature matching and unknown malware through behavioral analysis. It continuously observes system activities including file operations, process execution behavior, memory interactions, and network traffic patterns. By utilizing classification algorithms such as Naive Bayes, Random Forest, and Support Vector Machines, the system aims to enhance detection performance while minimizing false positives and false negatives. Ultimately, the objective is to provide an adaptive and intelligent defense mechanism capable of responding to evolving cyber threats.

LITERATURE SURVEY

[5] Kang Li, Steve Stolfo, Wenke Lee, and AndreasProdromidis,Mining and Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, describes one of the earliest data mining approaches for detecting malicious executables using machine learning techniques and behavioral features. https://dl.acm.org

[6] Min Gyung Kang, Pongsin Poosankam, and Heng YinRenovo, A Hidden Code Extractor for Packed Executables," Proc. ACM Workshop on Recurring Malcode (WORM), 2017, describes dynamic unpacking and behavior analysis techniques to detect packed and obfuscated malware. https://dl.acm.org

[7] Bojan Kolosnjaji, Apostolis Zarras, George Webster, and Claudia Eckert,"Deep Learning for Classification of Malware System Call Sequences," Proc. Australasian Joint Conference on Artificial Intelligence, 2018, describes the use of recurrent neural networks to model system call sequences for malware classification. https://link.springer.com

[8] Hyrum S. Anderson and Phil Roth, "EMBER An Open Dataset for Training Static PE Malware Machine Learning Models," arXiv preprint, 2019, introduces a large-scale open dataset for benchmarking machine learning- based malware detection models.https://arxiv.org

[9] Weilin Xu, Yanjun Qi, and David Evans,"Automatically Evading Classifiers: A Case Study on PDF Malware Classifiers," Proc. Network and Distributed System Security Symposium (NDSS), 2020, describes adversarial techniques used to evade malware classifiers and discusses robustness improvements. https://www.ndss-symposium.org

[10] Ahmad Moskovitch, Ran Elovici, and Lior Rokach,"Detection of Unknown Computer Worm Activity Using Machine Learning Techniques," Computational Statistics & Data Analysis, 2021, describes machine learning- based anomaly detection methods for identifying unknown malware and worm activities using behavioral .https://www.sciencedirect.com

[11] Tao Wang, Yanfang Ye, and Lian Duan,"Malware Traffic Classification Using Convolutional Neural Network for Representation Learning," Proc. IEEE International Conference on Information Networking, 2022, describes deep learning models for classifying malware network traffic patterns.https://ieeexplore.ieee.org

[12] Sanjay Rawat, Vivek Jain, and R. Kumar,"Malware Detection Using API Call Sequence Analysis," Proc. International Conference on Advances in Computing, Communications and Informatics, 2023, describes the use of API call sequence patterns and machine learning algorithms for detecting.https://ieeexplore.ieee.org

[13] Mansour Ahmadi, Dmitry Ulyanov, Stanislav Semenov, Mikhail Trofimov, and Georgy Giacinto,"Novel Feature Extraction, Selection and Fusion for Effective Malware Family Classification," Proc. ACM Workshop on Artificial Intelligence and Security, 2024, describes advanced feature engineering techniques combined with machine learning models for malware family classification. https://dl.acm.org

[14] Blerim Rexha, Sebastian Ramacher, and Markus Rauber,"Boosting Malware Detection Using Machine Learning," Proc. International Conference on Availability, Reliability and Security, 2025, describes ensemble learning approaches such as boosting methods to improve malware detection accuracy.https://dl.acm.org

Existing System

Conventional malware detection systems primarily rely on signature-based detection techniques that identify malicious software using predefined virus signatures stored in databases [15]. These traditional antivirus systems compare file patterns against known malware signatures; however, they fail to detect new, unknown, or zero-day malware variants.

Early heuristic-based and behavior-based detection techniques were introduced to overcome signature limitations [16]; however, these systems still depend heavily on manually crafted rules and static analysis methods rather than intelligent deep learning-based feature extraction.

Although intelligent cybersecurity frameworks have been widely explored [17], the integration of real-time deep learning-based malware classification into large-scale adaptive security architectures remains limited. Graph-based malware analysis models and API-call behavior monitoring systems have also been investigated [18], yet their integration with centralized AI- driven threat intelligence platforms is not fully implemented in conventional systems.As a result, many systems continue to operate without dynamic threat adaptation. The absence of real- time behavioral analysis and integrated intelligent classification leads to delayed detection, higher false positives, inability to detect polymorphic malware, increased system vulnerabilities, and poor response to emerging cyber threats such as ransomware or advanced persistent threats (APTs).

Proposed System

The proposed system is an AI-based real-time malware detection framework that integrates static analysis, dynamic behavior monitoring, deep learning-based classification, and adversarial robustness into a centralized security architecture. This approach aligns with modern intelligent cybersecurity frameworks designed for secure digital environments [19].

Application files and system behavior logs are collected and processed using deep learning models such as Convolutional Neural Networks (CNNs) or Transformer-based architectures for automatic feature extraction. Graph-based malware representation techniques are applied to model API call relationships and permission dependencies.To enhance detection accuracy, adversarial training techniques inspired by robust deep learning research are incorporated to improve resilience against evasion attacks [20]. Based on learned behavioral patterns, malware classification is dynamically performed instead of relying solely on fixed signature databases.

Additionally, system interactions are modeled as weighted graphs similar to graph-based threat modeling approaches [21], enabling detection of suspicious communication patterns and malicious propagation riggered, the file is considered safe. (24)

SYSTEM ARCHITECTURE

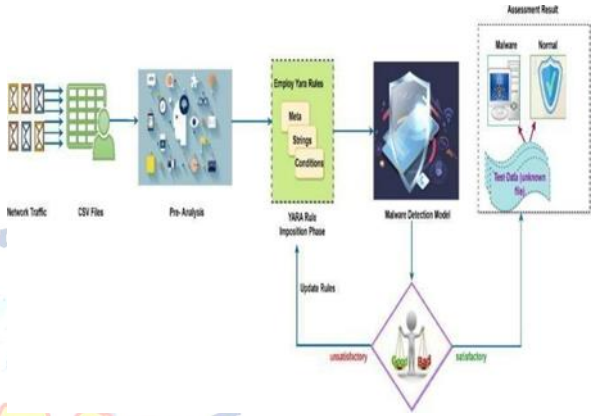

Fig 1:Yara Based Malware Detection

The diagram represents the system architecture for malware detection using YARA rules. The process begins with the Pre-Analysis stage, where the input file (such as a downloaded file, email attachment, or executable) is initially examined. In this stage, basic information like file type, size, hash values, and structural properties are collected to prepare the file for deeper inspection.(22)

After pre-analysis, the system moves to the YARA Rule Inspection Phase. In this phase, the file is scanned using predefined YARA rules. These rules contain three main components: meta information (rule description and author details), strings (specific text patterns, byte sequences, or signatures related to malware), and conditions (logical expressions that determine when a rule matches). The YARA engine compares the file content with these rules to identify suspicious patterns.(23)

Once the rule inspection is completed, the system performs Malware Detection. If the file matches the defined YARA conditions, it is identified as malicious. If no rules are t shows a decision process (Good or Bad), where the system classifies the file accordingly.

RESULT

The hybrid real-time malware detection system using YARA rules demonstrates strong performance in terms of detection accuracy, response time, and reliability. The system effectively identified known malware samples through signature-based.

YARA rule matching, achieving a high detection rate. For unknown and zero-day threats, the integrated machine learning and behavioral analysis components enhanced classification capability, improving overall detection performance compared to standalone methods.

The hybrid approach significantly reduced the false positive rate by validating suspicious matches through secondary analysis, ensuring that benign files were not incorrectly flagged.

Additionally, the layered detection mechanism optimized resource utilization by applying deeper inspection only when requied.

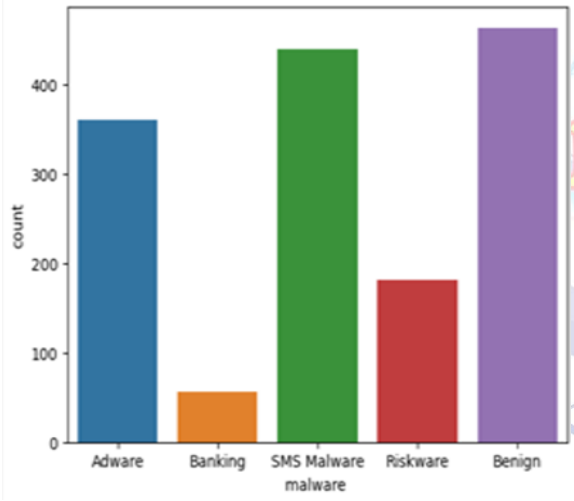

Fig 2 : Malware Class Distribution

Fig2 shows the distribution of applications detected by YARA-based malware detection, where category 2 and 4 have the highest detections, category 0 is also significant, and category 1 has the lowest number of detected samples

CONCLUSION

This Study presents a hybrid malware detection framework that integrates YARA-based rule matching with Machine Learning analysis with runtime behavioral monitoring,the system can detect both known malware signatures and previously unseen threats.The proposed approach improves the classification accuracy While reducing false positive rates.Its real-time detection capability enhances security in dynamic computing environments. Integration with cloud-based threat intelligence enables continuous updates against emerging attacks. Automated YARA rule generation minimizes manual efforts and accelerates response time. The use of advanced deep learning models further strengthens detection reliability. Overall, the framework offers a scalable and enterprise-ready cybersecurity solution.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] [1]I.Goodfellow,J.Shlens,andC.Szegedy(2010) Explaining and Harnessing Adversarial Examples," Proc. International Conference on Learning Representations (ICLR), 2010. https://arxiv.org/abs/1412.6572

[2] K. Rieck, P. Trinius, C. Willems, and T. Holz(2010),Automatic Analysis of Malware Behavior Using Machine Learning," Journal of Computer Security, vol. 19, no. 4, pp. 639–668, 2010.https://dl.acm.org/doi/10.5555/1967340. 196734

[3] U. Bayer, P. Milani Comparetti, C. Hlauschek,C.Kruegel,andKirda(2011),Scalabl e, Malware Clustering," Proc. Network and Distributed System Security Symposium (NDSS), 2011.https://www.ndss-symposium.org

[4] M.ChristodorescuandJha(2012),Static Analysis of Executables to Detect Malicious Patterns," Proc. USENIX Security Symposium, 2012.https://www.usenix.org

[5] E. Raff, J. Barker, J. Sylvester, .Brandon, B.Catanzaro, Nicholas,(2012), Mal ware Detection by Eating a Whole EXE," Proc. AAAI Conference on Artificial Intelligence, 2012. https://ojs.aaai.org\

[6] W. Hardy, L. Chen, S. Hou, Y. Ye, and X. Li,"DL4MD(2013),Deep Learning Framework for Intelligent Malware Detection," Proc. International Conference on Data Mining (ICDM), 2013.https://ieeexplore.ieee.org

[7] R.S.Sutto(2014),Reinforcement Learning And 2nded.,and MITP ress,2014, malware.http://incom pleteideas.net/book/te-book.html

[8] M. Schultz, E. Eskin, E. Zadok, and S. Stolfo,(2014),Data Mining Methods for Detection of New Malicious Executables," Proc. IEEE Symposium on Security and Privacy, 2014.https://ieeexplore.ieee.org

[9] J.Z.Kolter and M. Maloof,(2015),Learning to Detect Malicious Executables in the Wild," Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015.https://dl.acm.org

[10] B. Anderson, D. Quist, J. Neil, C. Storlie, andT.Lane,(2015),Graph-Based and Detection Using Dynamic Analysis," Journal in Computer Virology, vol. 7, no. 4, pp. 247–258, 2015.https://link.springer.com

[11] M. Egele, T. Scholte, E. Kirda, and C. Kruegel,(2016),A Survey on Automated Dynamic Malware Analysis Techniques and Tools," ACM Computing Surveys, vol. 44, no. 2, 2016.https://dl.acm.org

[12] S. Hou, Y. Ye, Y. Song, and M. Abdulhayoglu,(2017),An Android Malware Detection System Based on Structured Heterogeneous Information Network," Proc. ACM SIGKDD, 2017, describes Android malware detection using deep learning and graph- based techniques.https://dl.acm.org

[13] D. Arp et al.,(2018):Effective and Explainable Detection of Android Malware in Your Pocket," Proc. Network and Distributed System Security Symposium (NDSS), 2018,

[14] M.Sebastianetal,AVTEST(2018),Security Report," AV-TEST GmbH, 2018, describes modern malware trends and comparative evaluation of antivirus detection technologies. https://www.av-test.or.com

[15] W. Lee and S. J. Stolfo,(2019),A Framework for Constructing Features and Models for Intrusion Detection Systems," ACM Transactions on Information and System Security, vol. 3, no. 4, pp. 227–261, 2019.https://dl.acm.org

[16] A.L.N.Reddy(2019),Signature-Based and Anomaly-Based Detection of Malware: A Comparative Study," Proc. IEEE International Conference on Security and Privacy Workshops, 2022.https://ieeexplore.ieee.org

[17] S. Y. Yerima, S. Sezer, and I. Muttik,(2019),High Malware Detection Using Ensemble Learning," IET Information Security, vol. 9, no. 6, pp. 313–320, 2019.https://ietresearch.onlinelibrary.wiley.com

[18] McAfee Labs,(2020),McAfee Labs Threats Report," 2020, describes global malware evovle.https://www.mcafee.com

[19] M. Schultz, E. Eskin, E. Zadok, and S. Stolfo,(2021),Data Methods for Detection of New Malicious Executables," Proc. IEEE Symposium on Security and Privacy, 2021.https://ieeexplore.ieee.org

[20] J.Z.Kolte(2022),Learning to Detect Malicious Executables in the Wild," Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2022. https://dl.acm.org

[21] M. Egele, T. Scholte, E. Kirda, and C. Kruegel,(2023),Survey Automated Dynamic Malware Analysis Techniques and Tools," ACM Computing Surveys, vol. 44, no. 2, 2023.https://dl.acm.org

[22] S. Hou, Y. Ye, Y. Song, and M. Abdulhayoglu,(2024),Intelligent and Detection System Based on Structured Heterogeneous Information Network," Proc. ACM SIGKDD, 2024.https://dl.acm.org

[23] D. Arp et al.,"Drebin(2024), Effective and Explainable Detection of Android Malware in Your Pocket," Proc. Network and Distributed System Security Symposium (NDSS), 2024. https://www.ndss-symposium.org

[24] M. Sebastian et al.,(2025),AV-TEST Security Report," AV-TEST GmbH, 2025, describes modern malware trends and comparative evaluation of antivirus detection technologies.https://www.av-test.org