ZERO DAY MALWARE AND RANSOMWARE DETECTION

K. Bala vijaykumar[1], CH. Vinay[2], N. Bhanu mohan[3], G. Rohan[4], Dr. SK.Syed Basha[5]

KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India

[1,2,3,4]BTech Students, Department of CSE-AI

Email ids: kbalavijaykumar@gmail.com, challavinayy@gmail.com, bhanumohanchowdary@gmail.com, rohangundala57@gmail.com

[5]Professor, Department of CSE-AI&ML

Email: syba1139@gmail.com

***Abstract*—Malware and ransomware attacks have been getting more advanced leaving behind immense financial and data loss. Conventional machine learning-based detection systems primarily concentrate on an after execution analysis, which leads to late detection of a damage that has already been done. Moreover, the available literature is mostly based on offline data, unexplainable, and prone to evasion attacks. The following paper suggests a real-time, explainable, and resource-optimal machine learning system to detect ransomware and malware in the first stage. The suggested system constantly monitors the activity of the system, identifies threats prior to the start of file encryption, also offers interpretable information to security personnel, and is resistant to evasion. Experimental comparison shows that experimental analysis has a better ability to identify early and minimize false positives than the existing methods [4], [13] and better real-world applicability.**

I. INTRODUCTION

Malware and ransomware are a serious threat to the contemporary computer systems, where individuals, enterprises, and critical infrastructure are targeted. Especially, ransomware codes user data and requires money to decrypt it, which, most of the time, causes the loss of data permanently. The static analysis and signature-based systems can no longer be used to detect malware since, each time, the attackers alter malware behavior to avoid detection [9].

Machine learning (ML) has become popular in efforts to enhance the detection accuracy. Nevertheless, most of the ML-based systems are working post factum after the execution of malware, do not provide visibility in decision-making, and are not tested in real-world settings [12]. Such restrictions make them less effective in the field. Hence an urgent requirement is an explainable, real time and adaptable malware detection system.

II. PROBLEM STATEMENT

The existing malware and ransomware detection methods have been identified to have problems with lateness, poor explainability, inability to work in real time and are vulnerable to evasion. These are the issues that cannot be properly prevented by early ransomware encryption, and the trust of the security analysts which reduces the effectiveness of the existing solutions not to be applied in the real-life security environment.

III. OBJECTIVES

The major aims in this study are to conceive and create a real time malware and ransomware detection system, whichwould be able to detect the threats even before the ransomware encryption process is initiated. The objective of the study is to increase the explainability of the machine learning model decisions to increase transparency and confidence among secu- rity analysts [14]. The other important goal is to maintain the resistance to evasion attacks through the effective management of the occurrence of minor behavioral deviations of malware. Also, the study aims at minimizing computation overhead in order to facilitate effective real time implementation. Last but not least, the system is tested on dynamic and realistic data to ascertain its practical applicability in real life world setting.

IV. RESEARCH GAP

According to the critical analysis of the literature available, some important research gaps are established. The existing malware and ransomware detection methods are largely limited to detection approaches and cannot detect the ransomware attacks in real-time, thus failing to detect the ransomware attacks before they start encrypting the files [12]. The current research uses heavy dependence on offline and in-place testing conditions that cannot depict real world, ever-changing malware behavior. Also, it is hard to make fair comparisons and reproduce results without standardized and publicly available datasets. Lack of explainability is also an issue with many machine learning models and offer little insight on the reasoning used by the model to arrive at a prediction [14]. Moreover, the existing systems are weak against evasion methods of behavior whereby minor changes in malware behavior can trick the system [10]. Lastly, many of the proposed solutions have high computational and resource demands that limit their application to practical implementation in low-resource and real time settings.

V. LITERATURE REVIEW

Extensive research has investigated the application of machine learning and deep learning techniques for malware and ransomware detection. Traditional classifiers such as Random Forest, Support Vector Machines, k-Nearest Neighbors, Decision Trees, and Deep Neural Networks have been widely adopted due to their strong classification capability and ease of implementation. In addition, ensemble-based approaches that combine multiple learners have demonstrated improved robustness and higher detection accuracy across several benchmark datasets [3], [6], [8], [13].

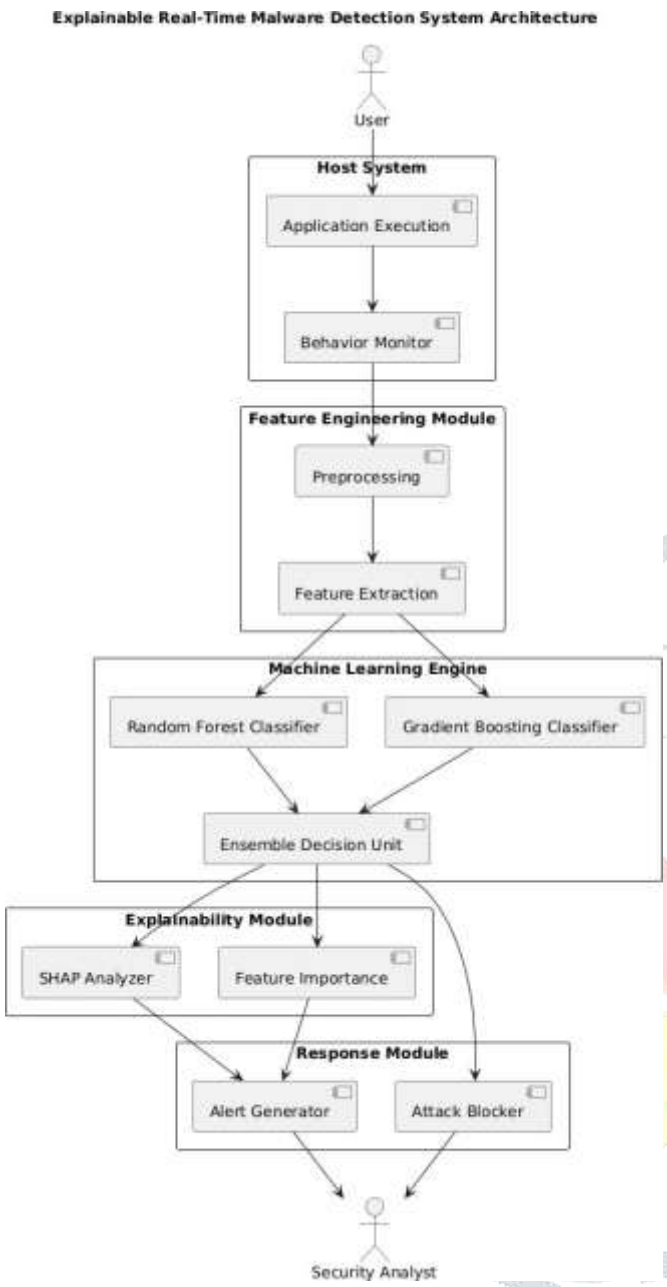

Fig. 1.

Despite these advancements, existing approaches suffer from multiple limitations. First, most conventional systems primarily rely on post-execution detection, meaning that malware is identified only after it has already begun executing on the host system. This reactive nature restricts the ability of security solutions to prevent damage at an early stage, especially in the case of fast-acting ransomware attacks [13]. Early-stage detection remains a critical challenge in practical cybersecurity environments.

Another major concern is the heavy dependence on fixed and outdated datasets. Many studies utilize publicly available datasets that no longer reflect the rapidly evolving behavior of modern malware. Furthermore, the lack of standardized benchmark datasets leads to inconsistent experimental setups, making it difficult to fairly compare results across different research works. As a consequence, reproducibility and generalization of reported results remain limited.

Explainability is another important gap in current research. A significant portion of existing machine learning models function as black boxes, providing little or no insight into how classification decisions are made. This lack of transparency reduces trust among security analysts and hinders effective incident investigation and forensic analysis [14]. Without understanding the reasoning behind predictions, analysts may hesitate to rely on automated systems for critical security decisions.

In addition, many detection models are vulnerable to evasion and adversarial attacks. Minor modifications in malware behavior, such as altered system call sequences or obfuscated execution patterns, can significantly degrade detection performance [10]. These vulnerabilities highlight the need for adaptive learning mechanisms and robust feature representations.

Overall, the reviewed literature indicates a clear gap between research-oriented malware detection techniques and their real-world applicability. There is a strong need for detec- tion frameworks that support early-stage behavioral analysis, adaptability to emerging threats, and explainable decision-making. These observations motivate the design of the proposed explainable, real-time, and adaptive malware detection framework.

VI. PROPOSED SYSTEM

The suggested system introduces a real-time, behavioral and explainable machine learning system to be able to perform effective detection of malware and ransomware. The system also keeps a continuous check on system level activities such as file access patterns, API invocation as well as process execution behavior in order to detect malicious activity at an early stage.

Lightweight extraction feature methods are used to provide efficient real-time processing with ensemble based machine learning models providing increased detection accuracy and robustness. Moreover, explainable systems are also incorporated to offer understandable and interpretable information about the model predictions as an element that permits security analysts to comprehend the logic of detection decisions. This would guarantee that the threats are detected early on, better trust has been established, and that it is used in practice in the real-life setting.

The Proposed System has some important features such as pre-encryption ransomware detection, real-time behavioral monitoring, explainable machine learning predictions, low computational resources usage, and high resistance to evasion attacks.

VII. METHODOLOGY

The proposed methodology consists of multiple well-defined stages designed to enable accurate, reliable, and real-time detection of malware and ransomware threats. The overall

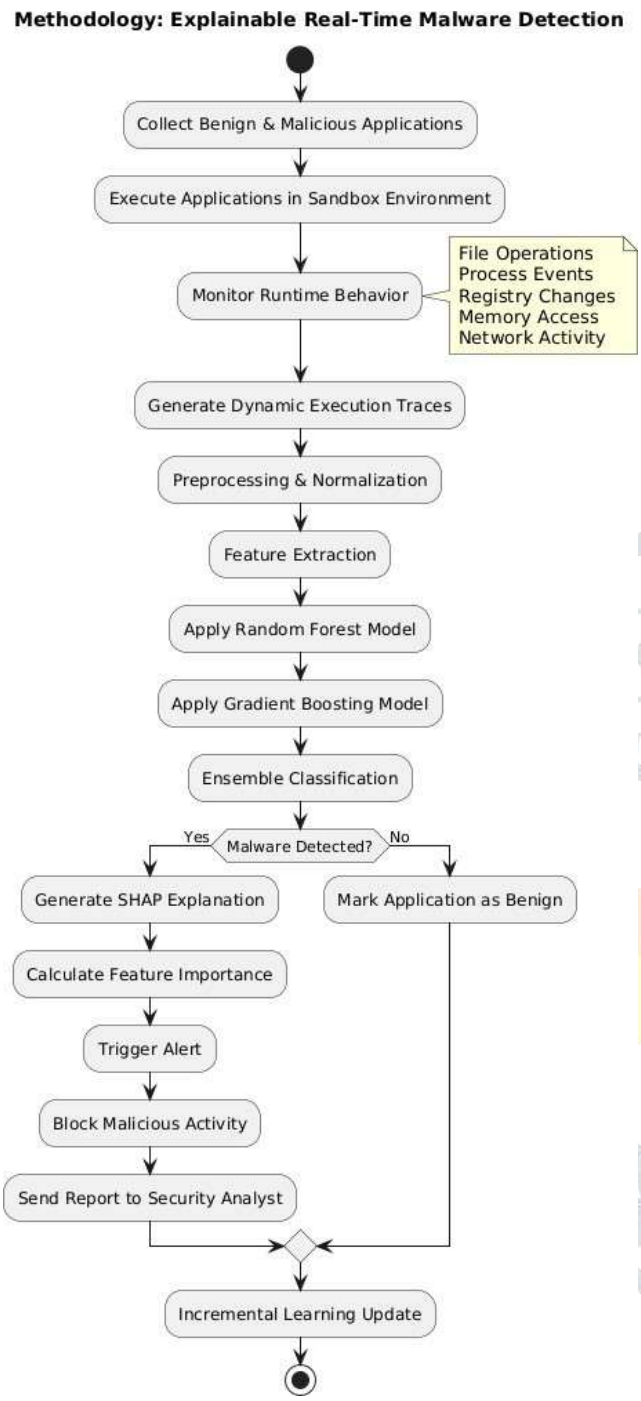

Fig. 2.

framework focuses on analyzing live behavioral patterns of applications rather than relying solely on static file characteristics, thereby improving resilience against obfuscation and polymorphic attacks.

Initially, both benign and malicious applications are executed within a secure and controlled sandbox environment to ensure safe observation of their runtime behavior. During execution, dynamic execution traces are collected to capture real-time system interactions. Unlike traditional static analysis approaches that examine only stored files, dynamic analysis provides deeper insight into how applications behave during execution. Key behavioral indicators monitored include file system operations, process creation and termination events, registry modifications, memory access patterns, and optional network communication activities. These raw behavioral logs are subsequently preprocessed to remove noise, normalize values, and generate structured representations suitable for machine learning.

Following data preprocessing, a comprehensive feature extraction phase is performed to transform low-level system events into meaningful behavioral features. Examples of extracted features include frequency of file access, unusual registry modifications, abnormal memory allocation patterns, and suspicious process hierarchies. These features form the input to a set of lightweight yet powerful machine learning classifiers. In particular, ensemble-based algorithms such as Random Forest and Gradient Boosting are employed due to their strong generalization capability, robustness to noisy data, and relatively low computational overhead [2]. The combination of multiple classifiers further enhances detection accuracy and reduces false positives.

To address the evolving nature of malware, the system incorporates incremental learning mechanisms that allow models to update themselves as new samples become available. This adaptive learning strategy ensures that the detection system remains effective against emerging ransomware variants and previously unseen attack patterns without requiring complete retraining from scratch.

In addition to achieving high detection performance, interpretability is a crucial design objective. Therefore, an explainability module based on SHAP (SHapley Additive exPlanations) and feature-importance analysis is integrated into the framework. This module highlights the most influential behavioral features contributing to each prediction, enabling security analysts to understand why an application is classified as benign or malicious. Such transparency increases user trust and facilitates forensic analysis.

Finally, the entire detection pipeline operates in a real-time manner. System behavior is continuously monitored, features are dynamically extracted, and the probability of malware or ransomware presence is estimated on-the-fly. When a potential threat is detected, the system immediately triggers an alert and prevents malicious actions such as unauthorized encryption or data exfiltration. At the same time, a detailed explanation of the decision is provided to security personnel, allowing them to make informed and timely responses.

VIII. RESULT AND DISCUSSION

The experimental evaluation demonstrates that the proposed system is capable of detecting ransomware and malware activities at an early stage, often before the initiation of the encryption process. This early-detection capability is particularly important, as it prevents irreversible damage to user data and significantly reduces system recovery time. Across a wide range of tested scenarios, including different ransomware families and benign application workloads, the system consistently exhibits high detection accuracy.

When compared with traditional machine learning–based detection models, the proposed approach achieves noticeably faster response times and a lower false positive rate [4], [13]. These improvements can be attributed to the use of lightweight ensemble classifiers and the reliance on dynamic behavioral features, which provide richer contextual information than static file attributes alone. The reduction in false positives is especially valuable in practical deployments, as it minimizes unnecessary alerts and decreases the operational burden on security analysts.

The integration of explainability mechanisms further enhances the effectiveness of the system. By employing SHAP-based explanations and feature-importance analysis, the framework provides clear and human-interpretable justifications for each detection decision. As a result, analysts gain deeper insight into the behavioral patterns that characterize malicious activity, leading to increased trust in the system and more confident decision-making. This transparency also supports post-incident analysis and helps in refining security policies.

In addition, the system demonstrates strong robustness against simulated evasion and obfuscation attacks. Minor behavioral modifications introduced by advanced malware samples do not significantly affect detection performance, indicating that the model generalizes well and does not rely on fragile or easily manipulated features [10]. The inclusion of incremental learning further contributes to this robustness by enabling the model to adapt to emerging threats over time.

Overall, the experimental results confirm that the proposed framework provides a reliable, efficient, and interpretable solution for real-time ransomware and malware detection, making it suitable for deployment in modern security environments.

IX. EXPECTED OUTCOME

A. Early Identification of Ransomware Intrusions.

- Before encryption begins, detection takes place.
- Online observation of system activity.
- Early warning alerts
- Faster response time
- Prevention of attack implementation.

B. Less Information loss and System damage.

- Security of confidential documents.
- Loss of data can be avoided permanently.
- Reduced system corruption
- Poor operation downtime.

C. Better Trust with Explanable Predictions.

- Open detection determinations.
- Description of alerts at features.
- Better confidence in the analyst.
- Less complex verification of model outputs.

D. Real-world Implementation of the Practices.

- No high computational complexity.
- Execution ability in real-time.
- Interoperability with the existing systems.
- Enterprise and end-user devices compatible.
- Scalable support of deployment.

E. Improved Immunity to the Advanced Malware.

- Detection technique based on behavior.
- Resistant to obfuscation Attacks.
- Flexibility in the new malware versions.
- Less susceptibility of evasion attacks.

X. CONCLUSION

This paper presented a real-time explainable machine learning framework for malware and ransomware detection that effectively addresses several critical limitations of existing security solutions. By focusing on dynamic behavioral analysis rather than solely on static file characteristics, the proposed system enables early-stage identification of malicious activ- ity, often before harmful actions such as file encryption are initiated.

The integration of lightweight ensemble learning models ensures high detection accuracy while maintaining low computational overhead, making the framework suitable for real-world deployment. Furthermore, the incorporation of explainability mechanisms, including SHAP-based interpretation and feature-importance analysis, provides transparent and understandable insights into the model's decisions. This transparency not only increases analyst trust but also supports faster and more reliable incident response.

The proposed approach bridges an important research and practical gap in cybersecurity by simultaneously addressing accuracy, efficiency, and interpretability. Experimental results demonstrate that the system achieves faster response times and lower false positive rates compared to conventional machine learning–based detection methods.

As future work, the framework can be extended to operate in large-scale cloud infrastructures and Internet of Things (IoT) environments, where diverse and resource-constrained devices generate massive volumes of behavioral data. Additional enhancements may include the integration of deep learning models, federated learning for privacy-preserving training, and automated threat intelligence sharing to further strengthen the adaptability and scalability of the system.

REFERENCES

[1] K. Kunku, A. Zaman, and K. Roy, "Ransomware Detection and Classification through machine learning," arXiv preprint arXiv:2311.16143v1, 2023.

[2] S. Panja et al., "An Efficient Malware Detection approach in Resource-Constrained devices by using machine learning feature influence Techniques," IEEE Access, vol. 13, 2025.
[3] Y. M. Bhagwat, K. D. Dere and A. A. Khatri, "Malware Detection through machine learning," in IJCRT, vol. 12, no. 11, 2024.
[4] D. Smith, S. Khorsandroo, K. Roy, "Machine learning algorithms and frameworks in ransomware detection," IEEE Access, vol.10, 2022.
[5] D. Gavrilu¸t et al., "Malware Detection with the help of the machine learning technique," International Multiconference on the Computer Science and Information Technology, 2009.
[6] V. Borate et al., "Malware Detection Analysis of Different Machine Learning Approaches," IJARSCT, vol. 4, no. 2, 2024.
[7] J. S. Sankar et al., "Ransomware Detection by use of machine learning," IJARSCT, vol. 5, no. 4, 2025.
[8] C. Murali Krishna Yadav et al., "Malware Detection with the help of the machine learning algorithms," TIJER, vol. 11, no. 5, 2024.
[9] U. Bayer, P. M. Comparetti, C. Hlauschek, C. Kruegel, and E. Kirda, "Scalable, behavior-based malware clustering," Proceedings of the Network and Distributed System Security Symposium (NDSS), 2009.
[10] M. Lindorfer, C. Kolbitsch, and P. M. Comparetti, "Detecting environment-sensitive malware," Proceedings of the 14th International Symposium on Recent Advances in Intrusion Detection (RAID), 2011.
[11] J. Zhang, Z. Qin, H. Yin, L. Ou, and K. Hu, "A feature-hybrid malware variants detection using CNN based on opcode sequence," IEEE Access, vol. 8, 2020.
[12] S. Homayoun, A. Dehghantanha, R. Parizi, K. Choo, and H. Karimipour, "A survey on ransomware detection methods," ACM Computing Surveys, vol. 52, no. 3, 2019.
[13] R. Vinayakumar, K. P. Soman, P. Poornachandran, "Detecting malware using deep learning," International Journal of Information Security, vol. 18, 2019.
[14] A. Continella et al., "ShieldFS: A self-healing, ransomware-aware file system," Proceedings of the 32nd Annual Computer Security Applications Conference (ACSAC), 2016.
[15] S. Chen, N. Xue, X. Li, and J. Zhu, "Zero-day malware detection using deep learning based behavioral analysis," Future Generation Computer Systems, vol. 120, 2021.