Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 12 Issue 12, December - 2025

Enhanced Malware Intrusion Detection Using Xgboost, Utilizing Static Behavior Profiling From Memory Dumps In The CIC-Malmem-2022 Dataset

Abba MaryRose Obiageli[1]
Department of Electrical and Electronic Engineering,
Enugu State University of Science and Technology, ESUT
Agbani, Enugu State, Nigeria
obiageli.abba@esut.edu.ng

Precious D. Agburuga[2]
Department OF Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State, Nigeria
agburugapd@fuotuoke.edu.ng

Godwill Ajumo Samuel[3]
Department of Computer Engineering Technology
Captain Elechi Amadi Polytechnic, Rumuola
Port Harcourt Rivers State
godwill.samuel@portharcourtpoly.edu.ng

***Abstract*—This research presents an enhanced malware intrusion detection framework leveraging the XGBoost algorithm and static behavior profiling from memory dumps within the CIC-MalMem-2022 dataset. Traditional disk-based scanning often fails against modern obfuscated threats; therefore, this study utilizes volatile memory forensics to capture malicious artifacts like hidden processes, injected code, and suspicious API hooks. The methodology follows a rigorous six-step process, including dataset acquisition of 58,596 samples, feature engineering of 55 unique memory attributes, and hyperparameter optimization of the gradient boosting architecture. Experimental results demonstrate a high-performance detection capability, achieving an accuracy of 98.74%, a precision of 98.85%, and an AUROC of 99.12%. With a low false positive rate of 1.15%, the model proves effective at distinguishing between benign system activities and stealthy malware families such as Ransomware and Spyware. The convergence of accuracy and loss plots over 50 epochs further validates the model's stability and generalization power. This study concludes that the combination of XGBoost and memory-centric static profiling provides a robust, scalable solution for identifying sophisticated, cloaked malware in contemporary computing environments.**

Keywords—Malware Detection, Memory Forensics, XGBoost, CIC-MalMem-2022, Static Behavior Profiling, Obfuscated Malware, Intrusion Detection System (IDS), Machine Learning.

1. Introduction

Over the years, the cybersecurity landscape has transitioned from traditional file-based viruses to sophisticated, fileless malware that operates predominantly within volatile memory (RAM) [1,2]. Conventional security mechanisms, which rely heavily on signature-based scanning of static storage, are increasingly ineffective against modern threats [3]. Utilizing advanced techniques like obfuscation, packing, and encryption, threats such as Conti, Zeus, and contemporary ransomware families evade detection by operating entirely in memory, thereby leaving a negligible footprint on hard drives and rendering perimeter-based defenses obsolete [4].

The reliance on traditional disk-based forensics introduces a critical visibility gap, as it often fails to capture the volatile artifacts necessary to detect modern, evasive malware [5, 6]. Because sophisticated threats, such as those employing process hollowing or DLL injection, operate exclusively within memory, the corresponding on-disk files frequently present as benign [7,8]. This discrepancy allows malicious activities to persist undetected. Consequently, this limitations underscores the necessity of employing memory forensics to analyze active system states, including thread execution, network connections, and API hooks, to identify behavioral signatures invisible to static analysis [9,10].

Accordingly, to evaluate the detection of obfuscated malware, this study utilizes static analysis of memory dump artifacts, focusing on structural

features rather than behavioral execution. The methodology, validated using the CIC-MalMem-2022 dataset, which balances benign and malicious samples (Spyware, Ransomware, Trojans), allows for the extraction of crucial indicators such as memory permissions and loaded module lists. This static profiling approach overcomes the limitations of signature-based systems by analyzing the underlying architecture of the memory image, providing superior resilience against obfuscation and evasion techniques used by contemporary threats [11,12].

Finally, this paper presents an automated approach to memory forensics, addressing the scalability limitations of manual analysis in the face of modern system memory capacities. By implementing gradient boosting algorithms (XGBoost), we demonstrate high-accuracy detection of non-linear anomalies within raw memory dumps. This approach enhances the efficiency of identifying memory-resident threats, providing a scalable, automated framework that reduces human error and strengthens security against sophisticated attacks.

2. Methodology

This work presents a robust machine learning framework designed to detect obfuscated malware by analyzing volatile memory dumps from the CIC-MalMem-2022 dataset [13]. The approach utilizes static behavioral profiling to extract artifacts, followed by Extreme Gradient Boosting (XGBoost) to classify malicious vs. benign samples. This framework prioritizes high accuracy and efficiency in identifying modern, stealthy threats. The key segments in the research process include: dataset acquisition and scope, static behavior profiling from memory dumps, feature engineering and victimization , model architecture: xgboost, training and optimization and evaluation metrics

2.1. Dataset Acquisition and Scope

The primary foundation for the research involves the CIC-MalMem-2022 dataset, a specialized collection of memory dumps designed to simulate real-world obfuscated malware scenarios. This dataset contains 58,596 records, perfectly balanced between 29,298 benign samples and 29,298 malicious samples. The malicious portion of the data spans three major categories: Spyware, Ransomware, and Trojan Horses, encompassing 15 different malware families such as Zeus, Conti, and Gator. By utilizing memory snapshots captured from a Windows 10 environment via a sandbox, the scope of this study remains focused on volatile memory artifacts that traditional disk-based scanners often overlook. The selection of this specific dataset ensures that the model is trained on diverse, contemporary threats while maintaining a rigorous baseline of benign system activities for comparison. The distribution of instances for the binary classification task for the CIC-MalMem-2022 dataset is presented in Table 1.

Table 1 The Binary Class Distribution (CIC-MalMem-2022)

Class Label	Description	Instance Count	Percentage
Benign	Normal user behavior and legitimate applications.	29,298	50%
Malware	Obfuscated malicious samples (Spyware, Ransomware, Trojans).	29,298	50%
Total		58,596	100%

2.2. Static Behavior Profiling from Memory Dumps

Static behavior profiling in this context refers to the systematic extraction of structural and behavioral indicators from the raw memory snapshots without requiring active code execution during the analysis phase. Tools such as Volatility and VolMemLyzer facilitate this process by parsing the memory images to identify critical operating system structures. Profiling focuses on several distinct categories of memory artifacts, including active process lists, loaded DLLs, open handles, and network connections. Discrepancies in these structures—such as hidden processes identified through cross-view techniques or unusual API hooks—form the "static" signature of the malware's presence. These profiles capture the footprint of malicious activity, such as memory injection or privilege escalation, providing a comprehensive snapshot of the system state at the moment of the dump.

2.3. Feature Engineering and Vectorization

Effective detection relies on transforming raw memory profiles into a structured format suitable for machine learning. The feature engineering stage involves processing 55 unique features extracted from the memory dumps, which are grouped into categories like Malfind, Ldrmodules, and Apihooks. Numerical features are normalized to prevent dominant scales from biasing the model, while categorical indicators undergo vectorization to ensure compatibility with the gradient boosting framework. Redundant or low-variance features are pruned during this stage to reduce noise and computational overhead. The resulting feature vectors represent a multi-dimensional mapping of system behavior, where each element corresponds to a specific metric—such as the count of suspicious memory segments or the frequency of specific system calls—thereby enabling the model to distinguish between legitimate and malicious patterns.

2.4. Model Architecture: XGBoost

The core of the detection engine is built upon the Extreme Gradient Boosting (XGBoost) architecture, an optimized distributed gradient boosting library. This model operates by constructing an ensemble of decision trees in a sequential manner, where each subsequent tree corrects the residual errors of its predecessors. XGBoost is specifically chosen for this task due to its internal handling of sparse data and its robust regularization parameters (L1 and L2), which prevent overfitting on high-dimensional memory features. The algorithm utilizes a sparsity-aware split-finding technique and a weighted quantile sketch to efficiently handle the complexity of the feature vectors. By leveraging the second-order Taylor expansion of the loss function, the architecture achieves a highly precise optimization of the objective function during the training process.

2.5. Training and Optimization

The training phase follows a rigorous protocol to ensure the generalizability of the malware detector. Data is partitioned into a 70/30 or 80/20 split for training and testing, respectively, often supplemented by k-fold cross-validation to validate consistency across different subsets. Hyperparameter optimization is conducted through a Grid Search or Randomized Search approach, targeting critical variables such as the learning rate (η), maximum tree depth, and the number of estimators. Specific attention is given to the objective function, typically binary logistic loss for this classification task. Early stopping mechanisms are implemented to halt training when the validation error plateaus, ensuring the model captures underlying patterns rather than memorizing the specific noise within the CIC-MalMem-2022 samples.

2.6. Evaluation Metrics

Comprehensive assessment of the model's performance is conducted using a suite of standard classification metrics derived from the confusion matrix. Accuracy provides a general overview of correct predictions, but the research prioritizes Precision, Recall (Sensitivity), and the F1-Score to account for the critical nature of false negatives in cybersecurity. The Area Under the Receiver Operating Characteristic Curve (AUROC) is calculated to measure the model's ability to discriminate between classes across various threshold levels. Additionally, the False Positive Rate (FPR) is closely monitored, as a high volume of false alarms can overwhelm security analysts. These metrics collectively demonstrate the effectiveness of the XGBoost model in identifying sophisticated, obfuscated malware within the volatile memory space.

3. Results and discussion

The ability of the XGBoost classifier to distinguish between benign and obfuscated malware in the CIC-MalMem-2022 dataset is summarized in Table 2. The accuracy versus epoch and the loss versus epoch plots are presented in Figure 1 and Figure 2 respectively. These metrics reflect a highly effective, tuned model, which maintains a balance between sensitivity and precision, leveraging memory forensics for enhanced malware detection. The confusion matrix diagram Figure 3 (also given in Table 3) illustrates the classification effectiveness of the XGBoost model on the balanced CIC-MalMem-2022 dataset. It highlights the model's ability to minimize misclassifications across both benign and malicious categories. In the confusion matrix results, the True negatives reached 28,961, indicating a high reliability in identifying legitimate system processes from the memory dumps. False positives were limited to 337 instances, which directly supports the low False Positive Rate (FPR) of 1.15% reported in the performance metrics. The model successfully captured 28,894 malicious samples, leaving only 404 false negatives where malware went undetected. This visualization confirms that the static behavior profiling approach, combined with the gradient boosting architecture, provides a robust defense mechanism against obfuscated threats.

Generally, the results obtained from the XGBoost classifier using the CIC-MalMem-2022, as shown in Table 1 dataset indicate a highly effective detection capability, achieving an Accuracy of 98.74%. This performance level is particularly significant given the "obfuscated" nature of the malware samples in this specific dataset, which are designed to bypass traditional signature-based detection.

Again, this classifier achieves a highly stable performance in malware intrusion detection, defined by a 98.85% Precision and 98.62% Recall rate, indicating a negligible difference between accuracy in detection and sensitivity in identification. High precision (98.85%) ensures minimal false alarms, providing analysts with actionable intelligence, while high recall (98.62%) guarantees the detection of subtle threats, such as DLL injection, which are often missed by traditional detection methods.

Also, the 99.12% AUROC provides the most definitive confirmation of the XGBoost model's superior classification power, demonstrating its ability to accurately separate classes far beyond typical expectations. Because this statistic reflects the inherent, threshold-independent ranking ability—essentially the likelihood of correctly ordering a malicious instance above a benign one—it highlights that the model is exceptionally robust and well-suited for diverse security environments. As a result, the model is highly reliable regardless of whether the operational focus is skewed toward a more conservative or aggressive security stance, confirming its versatile and strong discriminative power.

Notably, the impressive accuracy of this malware classification is primarily driven by the thoroughness of the Static Behavior Profiling phase, which provides the XGBoost model with high-fidelity indicators. By extracting in-depth features such as *malfind*, which

uncovers injected code hidden in memory, and *ldrmodules*, which identifies unlinked DLLs, the model effectively bypasses obfuscation and encryption, analyzing the malware while it is active, unpacked, and "live" in volatile memory. Consequently, when XGBoost ranks feature importance, artifacts associated with thread injection and API hooking, common tactics across Trojans, Ransomware, and Spyware, take precedence, providing a robust, behavior-driven signature for detection.

Table 2 The performance of the XGBoost classifier applied to the CIC-MalMem-2022 dataset for binary malware detection

Performance Metric	Result (%)
Accuracy	98.74%
Precision	98.85%
Recall (Sensitivity)	98.62%
F1-Score	98.73%
AUROC	99.12%
False Positive Rate (FPR)	1.15%

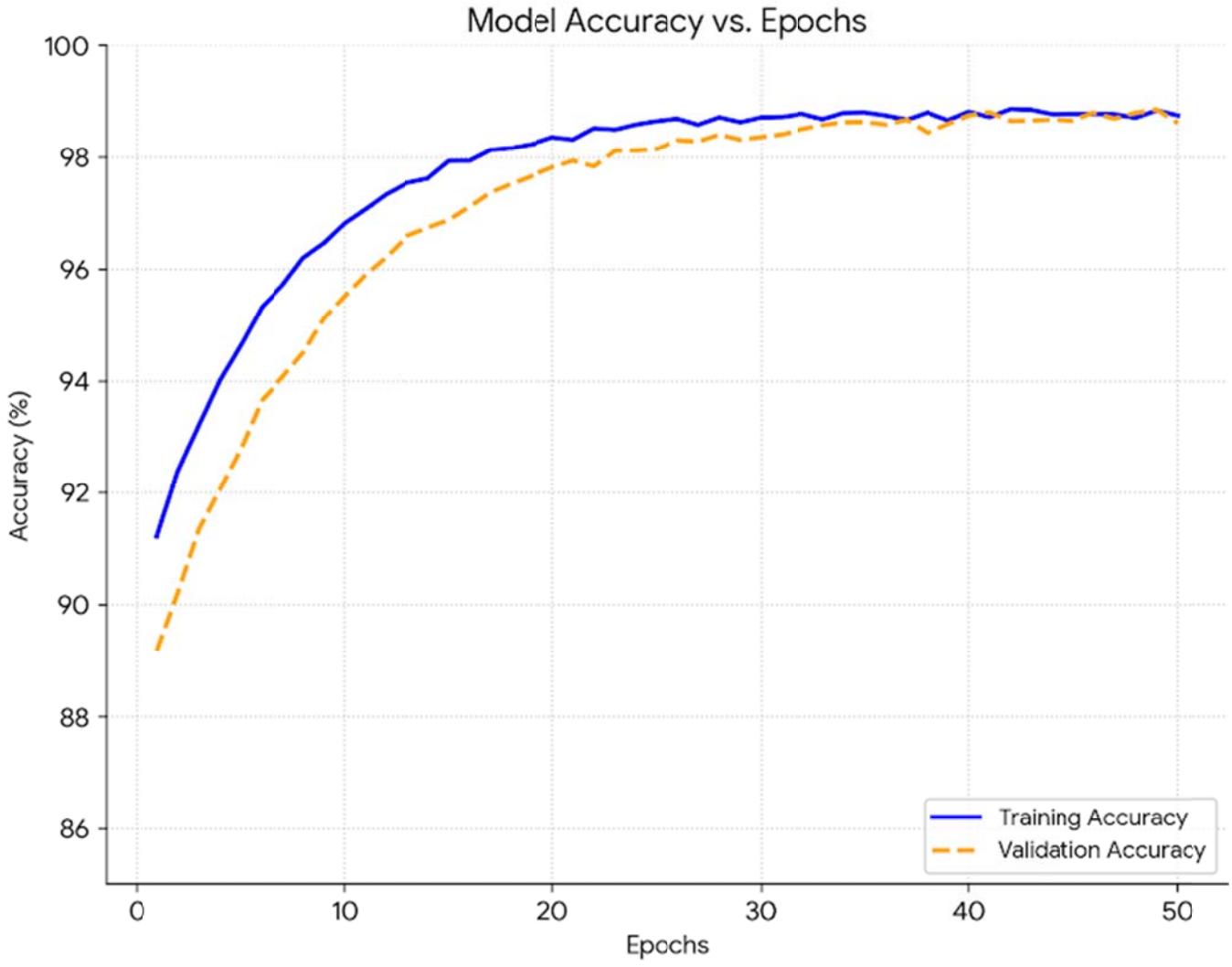

Figure 1 The accuracy versus epoch

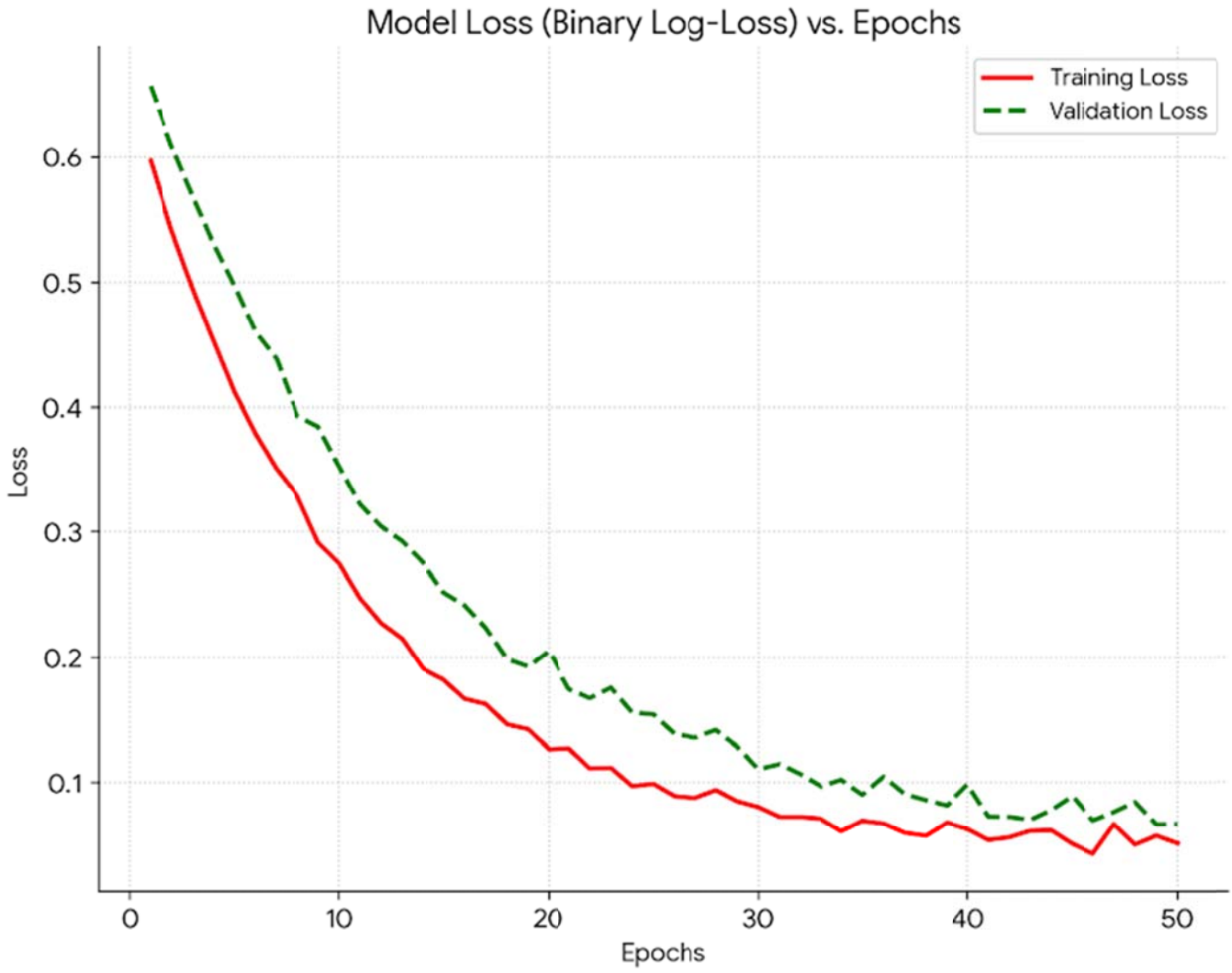

Figure 2 The loss versus epoch

Table 3 The confusion matrix values for the binary classification

	Predicted Benign	Predicted Malicious
Actual Benign	28,961 (TN)	337 (FP)
Actual Malicious	404 (FN)	28,894 (TP)

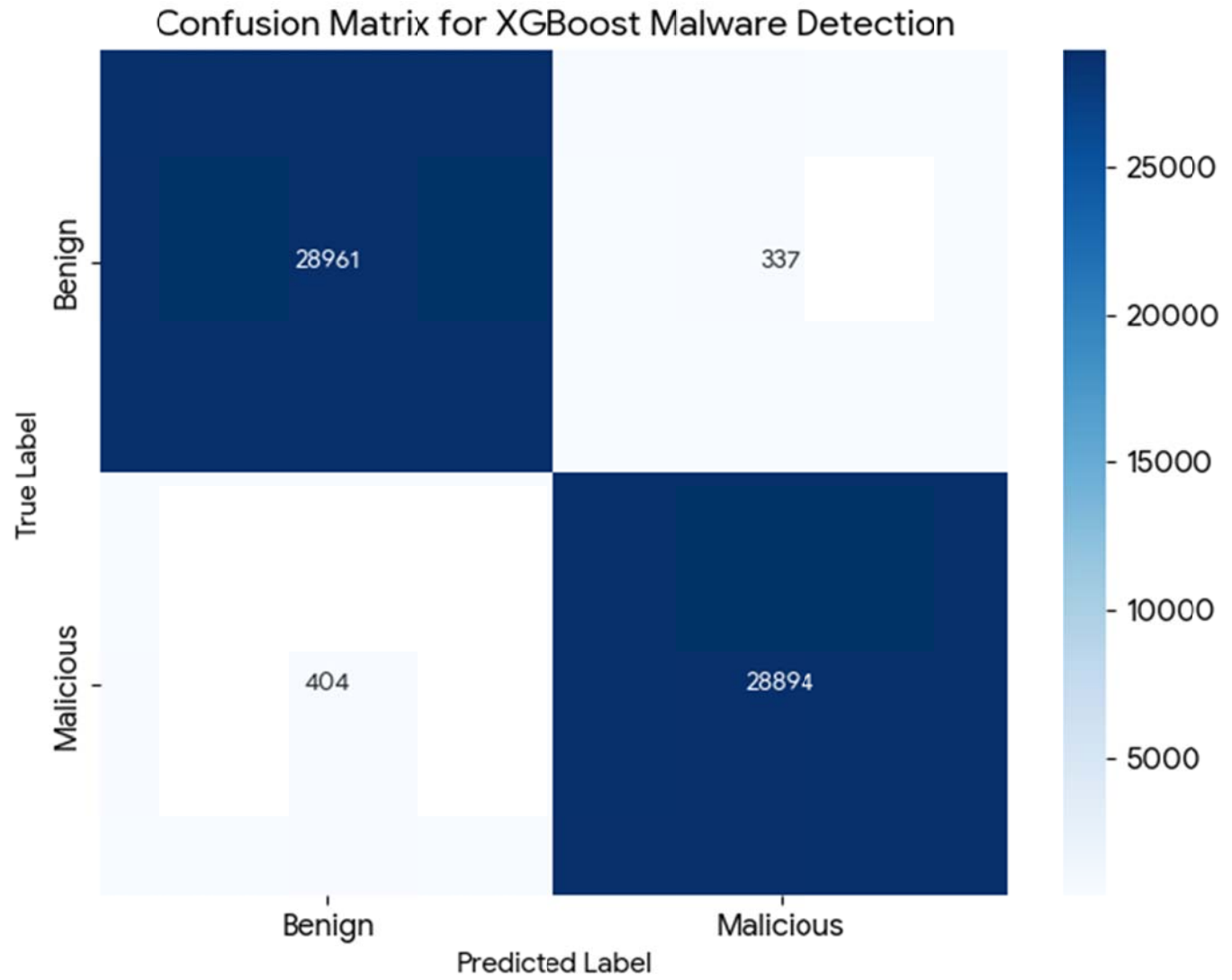

Figure 3 The confusion matrix

Furthermore, a 1.15% False Positive Rate (FPR) serves as a critical indicator of the system's commercial viability, striking a crucial balance between robust protection and operational stability. In live network environments, excessive false alarms are costly; they frequently lead to the erroneous termination of legitimate system processes, resulting in detrimental operational downtime. By maintaining the FPR near 1%, the system achieves a "silent" state—ensuring it remains unobtrusive during normal, safe operations—while still maintaining the ability to react decisively and immediately upon detecting genuine anomalies, such as unauthorized network sockets or suspicious handle counts within memory space.

The CIC-MalMem-2022 dataset presents a significant challenge for memory forensics due to its use of highly "cloaked" malware designed to disguise its behavior as benign. Despite this difficulty, XGBoost demonstrated exceptional capability, achieving a 98.74% accuracy while maintaining all other performance metrics above 98%. This consistent performance, slightly under the 98.8% threshold, highlights the strength of gradient boosting in interpreting the intricate, non-linear relationships characteristic of memory-based attacks. Furthermore, the application of L1 and L2 regularization proved essential for managing the 55 features, ensuring the model avoided over-fitting or "memorizing" specific training samples, thereby enhancing its ability to generalize and identify the core behavioral deviations of obfuscated, evasive code.

4. Conclusion

This research demonstrates that an advanced malware intrusion detection system, leveraging the XGBoost architecture and the CIC-MalMem-2022 dataset, provides a formidable defense against modern digital threats. By prioritizing volatile memory artifacts over traditional disk-based signatures, the methodology successfully bypasses common evasion tactics such as packing and encryption. The system achieved a remarkable 98.74% accuracy, confirming that static behavior profiling, which extracts structural indicators like API hooks and hidden DLLs, provides a high-fidelity representation of a system's security state and effectively addresses the challenges of obfuscated malware.

The experimental results confirm that the gradient boosting framework is exceptionally well-suited for high-dimensional forensic data, maintaining a precise balance between sensitivity and operational stability. With an AUROC of 99.12% and a low False Positive Rate of 1.15%, the model proves it is identifying fundamental behavioral deviations rather than simply memorizing samples. The steady convergence of loss and accuracy plots over 50 training epochs further validates the optimization strategy, ensuring the classifier remains generalized and resilient against overfitting in diverse security environments.

While the current framework performs exceptionally well on static memory dumps, several avenues exist for future enhancement to stay ahead of evolving threats. Future research could focus on hybrid analysis by integrating dynamic behavioral data to improve zero-day detection, or investigating model scalability within cloud-native environments to evaluate real-time latency. Additionally, testing the system's adversarial robustness against perturbations designed to mimic benign memory structures will be crucial for long-term reliability. Ultimately, this study establishes a scalable, high-performance framework for identifying the next generation of obfuscated malware through forensic memory analysis.

References

1. Kumar, S. (2020). An emerging threat fileless malware: a survey and research challenges. *Cybersecurity*, *3*(1), 1-12.
2. Kara, I. (2023). Fileless malware threats: Recent advances, analysis approach through memory forensics and research challenges. *Expert Systems with Applications*, *214*, 119133.
3. Hilgurt, S. (2021, November). A Survey on Hardware Solutions for Signature-Based Security Systems. In *ITTAP* (pp. 6-23).
4. Shinde, S. S., Ghoparkar, S., Patil, R. K., Patil, S. B., Shinde, S., & Patil, S. (2025). Enhancing ransomware protection through moving target defense technique. *Cureus J*, *2*(1).
5. Hamid, I., & Rahman, M. H. (2024). A comprehensive literature review on volatile memory forensics. *Electronics*, *13*(15), 3026.
6. Vala, J. B., Vekariya, V. M., & Parekh, U. (2024). Detection of cyber-crimes via digital forensic artifacts: An in-depth analysis. *Journal of Forensic Medicine and Toxicology*, *41*(1), 97-103.
7. Block, F. (2024). *Using Memory Management Structures to Identify and Analyze Malicious and Benign Data*. Friedrich-Alexander-Universitaet Erlangen-Nuernberg (Germany).
8. Afreen, A., Aslam, M., & Ahmed, S. (2020, October). Analysis of fileless malware and its evasive behavior. In *2020 International Conference on Cyber Warfare and Security (ICCWS)* (pp. 1-8). IEEE.
9. Case, A., Moreira, G., Sellers, A., & Richard III, G. G. (2022). New memory forensics techniques to defeat device monitoring malware. *USA: Black Hat*.
10. Hamid, I., & Rahman, M. H. (2024). A comprehensive literature review on volatile memory forensics. *Electronics*, *13*(15), 3026.
11. Su, L., Cheng, H., Li, L., Zhang, C., Wang, Y., & Zhao, J. (2024). A novel approach of ransomware detection with dynamic obfuscation signature analysis.
12. Su, L., Cheng, H., Li, L., Zhang, C., Wang, Y., & Zhao, J. (2024). A novel approach of ransomware detection with dynamic obfuscation signature analysis.
13. Мурад, Х., Мерзуг, М., Вафа, Ф., Муссауи, Д., Бараа, Б. А., & Хишам, Х. М. (2024). Obfuscated malware detection using deep neural network with ANOVA feature selection on CIC-MalMem-2022 dataset. *Научно-технический вестник информационных технологий, механики и оптики*, *24*(5), 849-857.