Detection of Malware and Malware-based Attacks using AI Approaches

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

by

Parthajit Borah

Enrollment No. CSP17103

Registration No. TZ166804 of 2016

Department of Computer Science and Engineering

School of Engineering, Tezpur University

Tezpur, Assam, India - 784028

October, 2024

However, the interpretability of GNNs remains an ongoing concern. Understanding the representations learned by GNNs is essential for building trust in their decision-making processes. As GNNs generate complex, high-dimensional embeddings to represent malware graphs, deciphering these representations to provide clear, human-understandable insights is challenging. Addressing this issue is vital for enhancing the transparency and acceptance of GNN-based detection systems in practical applications.

Chapter 10

Conclusion and Future Direction

This dissertation makes six significant contributions to the field of malware detection and malware-based attacks. In this chapter, we summarize these contributions and suggest directions for future research.

10.1 Conclusion

Following conclusions are drawn from the contributions in this dissertation.

- The first contribution, detailed in Chapter 3, involves the creation of two malware feature datasets: TUMALWD and TUANROMD. These datasets, designed for different platforms, support the validation of malware detection methods. The dataset generation framework includes three phases: data collection and storage, data analysis, and feature engineering. An automated honeynet system and client-server architecture facilitate efficient data collection and storage. Comprehensive malware analysis is conducted using Analysis Client and Analysis Server modules, which store analysis reports in a database. The final phase, feature engineering, involves preprocessing to remove empty analysis reports and extract meaningful features. For TUANROMD, permission-based and API-based features are extracted, while TUMALWD focuses on host-based and network-based features. These datasets provide robust resources for validating malware detection methods. In addition to the Windows and Android malware datasets, two more feature datasets have been introduced to this research. These include an image-based malware feature dataset and a function call graph (FCG) dataset,

both derived from collected raw malware binaries.

- In Chapter 4, we detail a feature selection approach called FSR, designed to handle imprecise data. FSR uses the concepts of indiscernibility relation, set approximation, and attribute dependency to determine a relevant feature subset by minimizing redundancies between features and maximizing the relevance between features and class. The effectiveness of FSR is assessed using a classifier on several real-world datasets, showing better classification accuracy and execution time performance compared to other methods. FSR employs rough set theory to select the most effective subset of features from a given feature set, enhancing the overall efficiency and performance of the classification process.

- In Chapter 5, we have presented a cost-effective method for ransomware detection, ERAND, which operates on an optimal feature space to yield the best possible accuracy for the ransomware class as a whole and for each variant. High-dimensional data can pose challenges such as the curse of dimensionality and overfitting, which can degrade performance and increase memory and computational costs. ERAND addresses these issues by removing irrelevant features through an ensemble feature selection approach. This method combines the results of multiple feature selection algorithms using a consensus function, thereby eliminating biases and generating an optimal subset of features. The final set of features is used to evaluate classifier performance, and the multi-objective evolutionary method NSGA-II is applied to compute the best set of weighting factors for the classifiers based on their performance across 11 ransomware variants.

  In Chapter 6, the ever-evolving landscape of malicious software is addressed by developing effective detection methods focused on identifying important features. An ensemble approach is proposed called FRAMC to identify key features that significantly contribute to malware detection. FRAMC exploits the Markov chain algorithm while aggregating individual feature rankings given by participating base ranker algorithms. The effectiveness of FRAMC is assessed using different types of classifiers on various real-world malware datasets. The outcomes of our analysis demonstrate that FRAMC excels in performance compared to other methods. It finds an optimal number of features for both Windows and Android malware. While FRAMC shows promise, further investigation into different aggregating techniques might produce even better outcomes. Exploring various ensemble tactics or fusion approaches could improve the approach's flexibility and robustness.

- In Chapter 7, a parallel KNN algorithm called TUKNN is proposed, designed to handle voluminous data with high accuracy. KNN is a widely used classification algorithm and is well-suited for parallel implementation due to its numerous independent operations. When dealing with large training and testing datasets, execution speed can become a bottleneck, making parallel implementation advantageous. We implemented TUKNN on the CUDA framework, exploring various proximity measures in parallel to identify the most accurate one. Measures include Euclidean distance, Manhattan distance, Kulczynski distance, cosine similarity, Chebyshev distance, Soergel distance, Sorensen, and Tanimoto. Notably, for the binary classification of the Ransomware Dataset, GPU computation using Chebyshev distance was 40.86 times faster than CPU computation. For high accuracy in binary classification, Kulczynski or Soergel Distance is recommended, while for multi-class classification, Kulczynski Distance is preferred.

- In Chapter 8, various deep learning models are evaluated to counter and mitigate malware threats. Deep learning models, particularly convolutional neural networks, can continuously adapt and learn from new data, which is crucial for combating evolving malware. These models can detect previously unseen malware by learning patterns and anomalies from large datasets and can automatically extract intricate features from raw data without manual feature engineering. In our work, we evaluated eight different convolutional neural network architectures for their effectiveness in malware classification. The comprehensive evaluation using three benchmark datasets demonstrates the superiority of our proposed framework over traditional machine learning algorithms.

- In Chapter 9, propose a malware defense solution based on Graph Neural Networks (GNNs). This novel strategy leverages graph-based representations and neural networks to address malware threats. Malware attacks are represented as graphs, with nodes representing different functions and edges capturing their interactions. Function call graphs, which illustrate the high-level actions and execution flow of malware, are effective in identifying complex attacks and code obfuscation techniques. Features extracted from these graphs, such as specific function sequences and call frequencies, are used to train models for automated malware detection. Our study assesses the performance of GNN models using a function call graph dataset of malware, demonstrating their potential in enhancing malware defense strategies.

10.2 Future Directions

In this section, some of the possible directions for future research in this domain are outlined.

- Explainable AI (XAI) is an important aspect that needs due attention in future research. Incorporating explainable AI techniques into malware detection models would enhance the interpretability of these models. This would make it easier for researchers to understand how decisions are made, identify potential false positives or false negatives, and gain trust in the automated detection systems. By providing clear explanations for the model's decisions, XAI can also aid in refining and improving the detection algorithms.

- The issue of adversarial robustness in malware detection becomes increasingly critical as malware authors continuously evolve their evasion techniques. In this scenario, focusing on enhancing detection models against adversarial attacks is essential to maintain their effectiveness. Incorporating adversarial training, developing robust feature extraction methods, and enhancing the adversarial robustness of graph neural networks are key areas of focus.

- Developing models capable of real-time malware detection and response is crucial. Optimization of algorithms and leveraging edge computing are essential to ensure swift detection and mitigation of malware threats, reducing potential damage. Fine-tuning detection algorithms for faster and more accurate performance and integrating edge computing to process data closer to the source can minimize latency. Additionally, exploring hardware integration-based solutions can enhance detection capabilities by utilizing specialized hardware accelerators for real-time analysis. Incorporation of these advancements into current solutions and studying their performance are kept under future works.

- Exploring self-supervised and unsupervised approaches for malware detection is an important aspect that needs due attention. These methods can utilize large amounts of unlabeled data, which can enhance detection capabilities without heavily relying on labeled datasets. Self-supervised learning can help models generate useful representations from raw data, improving their ability to identify new malware patterns. Unsupervised learning techniques, such as clustering and anomaly detection, can reveal hidden structures and

relationships in the data, aiding in the detection of unknown malware. Incorporation of these advanced techniques into current solutions and studying their performance are kept under future works.