Ransomware Detection Using Deep Q-Network and L2PGD Attack Analysis on a Custom Dataset

Niringiye Godfrey[1], HoonJae Lee[2*], ByungGook Lee[3]
1. Mr, Dept. computer engineering, Graduate School, Dongseo university
2. Prof., Dept. information security, Dongseo university
3. Prof., Dept. computer engineering, Dongseo university

Abstract

In the current fast changing cyberspace, ransomware has continued to be a formidable threat. Through this research, using deep reinforcement learning and adversarial attack models, we undertook performance analysis evaluation of a locally constructed ransomware dataset. The dataset contained key dynamic features that were extracted from raw ransomware samples processed in Cuckoo sandbox environment. Our approach combined supervised learning for initial detection and Deep Q-Network (DQN) algorithm for adaptive behavioral analysis. An L2 Projected Gradient Descent (L2PGD) adversarial attack was then carried out to evaluate the robustness of both security and stability of the ransomware detection model. The results that were obtained demonstrated that Deep Reinforcement Learning (DRL) can effectively classify samples as benign and ransomware. Moreover, the successful adversarial attack underscores the need for improved robustness measures in artificial intelligence models.

Keywords: Artificial Intelligence (AI)|Custom Dataset | Supervised Learning | Deep Q-Network |L2 Projected Gradient Descent |Adversarial Attack

I. INTRODUCTION

In today's cybersecurity landscape, ransomware continues to pose a significant threat. According to the federal bureau of investigations (FBI) reports 14 out of 16 critical infrastructure sectors have been targeted by ransomware [1]. Moreover, in South Korea, there have been enhanced cybersecurity efforts such as cyber threat intelligence sharing platform and increase in ransomware attacks is one of the major reasons behind its establishment [2]. The Internet crime complaint center (IC3) of the FBI collects data to respond to cybercrimes of which ransomware is among [3]. Even though these efforts and resources have been invested, ransomware attacks skyrocketed by staggering 74% worldwide in 2023 alone [4]. To curb this rapidly evolving threat, advanced AI and machine learning solutions both in industry and research fields have been proven to be promising solutions. However, in research community, existing AI based Ransomware detection studies are faced with significant limitations. Many of these studies utilize generic ransomware datasets. These datasets often are less effective in real world scenarios since they do not capture unique characteristics of ransomware attacks [11]. Furthermore, there is limited research in addressing adversarial attacks that greatly exploit the performance of the AI models of ransomware detection systems [12]. Our research contributes to addressing these challenges by exploring the construction of custom ransomware datasets, deep

* This work was supported by Dongseo University, "Dongseo Cluster Project (type 1)" Research Fund of 2024 (DSU-20240008)

Manuscript : 2024. 11. 20
Revised : 2024. 12. 05
Confirmation of Publication: 2025.12.26
Corresponding Author : HoonJae LEE, e-mail : hjlee@dongseo.ac.kr

reinforcement learning and adversarial attack analysis. Our dataset consists of custom ransomware features, and thus aims to strengthen robustness and reliability of AI based ransomware detection models. Moreover, through incorporation of techniques that detect and mitigate adversarial attacks, we address adversarial threats and achieve improved overall security and resilience of AI powered ransomware detection systems.

II. Background and Motivation

Several studies in the literature carried out research on AI-based ransomware detection [5-8]. Gazzan, M. et al. [5] explored novel ransomware detection exploiting uncertainty and calibration quality measures using deep learning. Through utilization of Bayesian methods as well as dropout techniques, they enhanced ransomware detection model accuracy at the same time preventing overfitting. This is very crucial for model performance on unseen ransomware data. However, the study relied on generic datasets that may not fully capture the unique characteristics of ransomware attacks, reducing its effectiveness in real-world scenarios. Singh et al. [6] employed the RANSOMNET+ model for ransomware classification on cloud-encrypted data, achieving high accuracy and low loss, demonstrating the model's effectiveness in detecting ransomware. Nevertheless, the research did not address potential vulnerabilities to adversarial attacks, which could undermine the model's performance.

Taha et al. [7] proposed a framework for detecting ransomware during the early stages through API call feature extraction. They successfully detected ransomware early in the attack lifecycle, improving response times. However, due to the framework relying on specific API calls may limit its applicability across different ransomware variants. Furthermore, it did not consider adversarial attack scenarios. Moujoud, L et al [8] in their state-of-the-art survey on ransomware detection using machine learning techniques surveyed various machine learning and Deep Q-Network based ransomware detection techniques. They provided a comprehensive overview of the methods that are employed in ransomware detection. Their research, however, can be criticized for lack of practical implementations. In our research we address this and bridge the gap between theory and practice.

To address the challenges in the above studies, we undertake a comprehensive study involving the design and development of a real-world ransomware dataset. We then evaluate the dataset using a Deep Reinforcement Learning model to assess its performance in detecting ransomware attacks. Furthermore, we implement techniques that detect and mitigate adversarial attacks on detection model attacks and contribute to achieving resilience and security of the AI powered ransomware detection systems.

III. PROPOSED METHOD

Our proposed approach in Figure 1 includes 5 major stages of environmental construction, raw sample collection, preprocessing, labeling, and evaluation using DRL and adversarial attack analysis.

1. Construction of Cuckoo Sandbox Environment

We set up Cuckoo Sandbox for the dynamic analysis of the ransomware samples in a nested virtualization manner on a dedicated machine as shown in Figure 2. Through this environment detailed behavioral features were extracted.

2. Collecting and preprocessing raw ransomware samples

1209 raw samples of ransomware were collected from malwares.com. A python script configured with Abuse.ch API was used for this purpose. We then performed data preprocessing through 5 stages of raw sample processing, feature extraction, data cleaning, transformation, and augmentation. Z−score and SMOTE were used for normalization and addressing class imbalance respectively.

3. Labeling of data

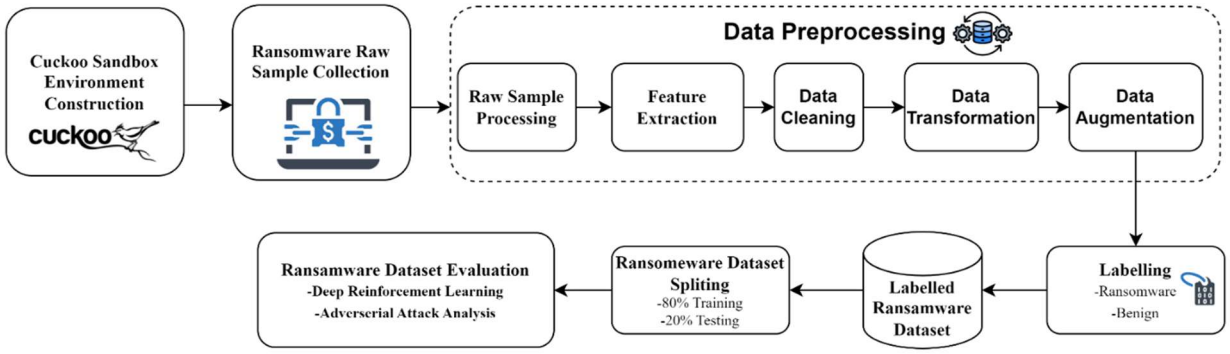


Figure 1. Our Proposed framework for Deep Reinforcement Learning and Adversarial Attack Analysis for Ransomware Detection Using a Custom Dataset

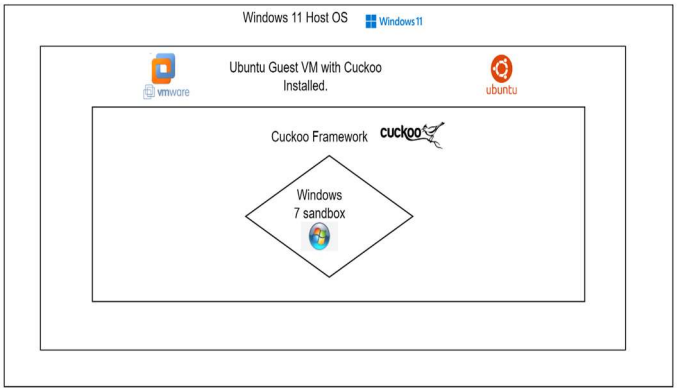


Figure 2. Our experimental setup

We labelled samples as ransomware (1) if their behavior in Cuckoo sandbox exceeded 0, otherwise benign. This resulted in a labeled dataset that was stored in CSV format. We split this dataset into 80% for training and 20% for testing. After that we performed analysis using DRL models and adversarial attack analysis.

4. Deep Reinforcement Learning and Adversarial Attack Analysis for Ransomware Detection

Due to its capacity to continuously learn and improve, reinforcement learning is well suited for detecting rapidly evolving ransomware attacks. In this research we explore DRL integrating supervised neural network and DQN for initial classification and for ongoing learning from environment interactions respectively. To enhance learning, the DQN receives rewards based

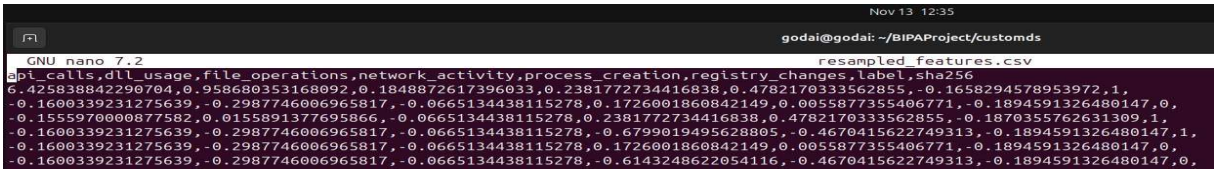


Figure 3. Part of resulting dataset after Data Preprocessing stage

on classification of samples as ransomware or benign. Features from the dataset represent the state in the reinforcement learning environment and this is where the training begins with an observation space. For decision making, the agent retrieves relevant features, and this is the starting State (S=1). Observations are classified, learned policy π is applied and rewards are received based on (R=+1) for correct classifications and (R=−1) otherwise with additional incentives for exploration. Overtime their reward system improves classification accuracy. The agent then increments space index (S=S+1), checks if S < = n to continue and if not, the process is concluded. Based on the awards received, the agent updates its policy at the end of one learning episode. To ensure effective learning for classification between benign and ransomware actions, the Q−learning update rule adjusts Q−values based on observed and estimated future rewards. Key Parameters of Learning Rate (α), Number of Iterations (Timesteps), Epsilon (ε) were selected for successful training of our DQN agent. To ensure a balance between effective perturbation and maintaining the perturbation within the L2 norm constraints our learning rate was set as 0.0001. To ensure convergence of the perturbation and thorough exploration of perturbations, Number of Iterations was set as 8772 timesteps. To control the magnitude of the adversarial changes and achieve maximum allowable perturbation, Epsilon was set to 0.1.

To assess our DQN model, L2PGD attack was carried out. To maximize prediction errors, small perturbations were introduced. Our attack algorithm is as follows.

Start

1 Initialize the perturbation δ to zero.

2 For t = 1 to T:

3 Compute the gradient of the loss with respect to the perturbation:

$g_t = \nabla_\delta L(\theta, x+\delta_t, y)$

4 Update the perturbation using gradient ascent: $\delta_{t+1} = \delta_t + \alpha \cdot g_t / (||g_t||_2)$ [9]

5 Project the perturbation onto the L2 ball if needed: $\delta = \epsilon \cdot \delta / ||\delta||_2$ if $||\delta||_2 > \epsilon$

6 The final adversarial example is $x_{adv} = x + \delta_T$ [10]

7 End

IV. Results and Analysis

1. Dataset Features in the dataset

To identify key variables that can influence model performance, feature

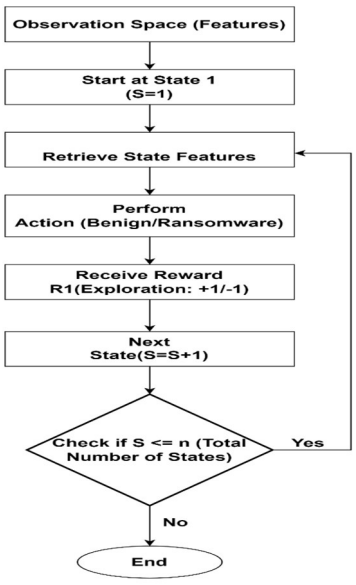


Figure 4. Training process of DQN agent

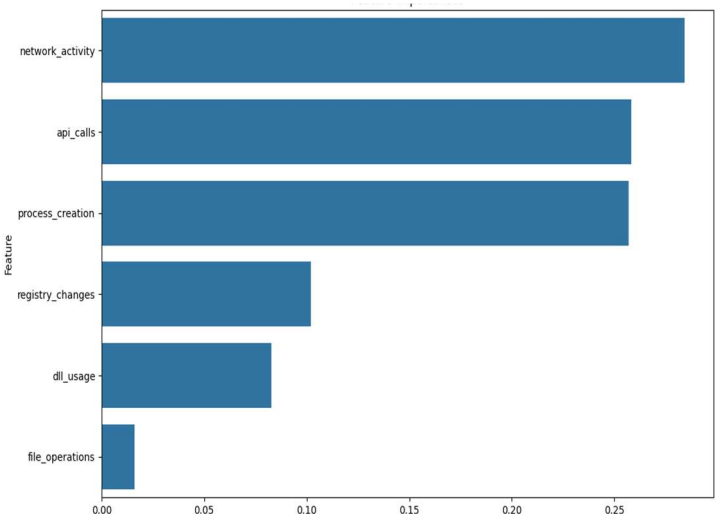


Figure 5. A plot of feature importance

importance is plotted in Figure 5. This is crucial in pinpointing anomalies as well as relationships. Network activity, API Calls and Process creation were the top 3 most important features.

2. Training of Neural Network

We trained our neural network on 100 epochs. Approximately 81.80% and 81.97% training accuracy and validation accuracy were achieved respectively as indicated in

Figure 6. Neural Network Model Performance Evaluation Results

Figure 6. Loss decreased to 0.3845 during the training too. The values of key metrics of precision (0.76), recall (0.94) and F1−score (0.84) indicated the model is well tuned.

3. Reinforcement Learning Component

With parameters configured in Section III, the DQN agent was trained in a custom Gym environment. The training loss of 0.255 and 2167 updates suggested active learning as indicated in Figure 8. To assess learning progress, rewards overtime was plotted. To track the rewards obtained by the agent over time we plot Reward vs. Timesteps Plot for our DRL Model in Figure 7. This plot helps in understanding the learning progress of the reinforcement learning agent. It indicates how well the agent is optimizing the reward function and provides insights into the efficiency of the training process.

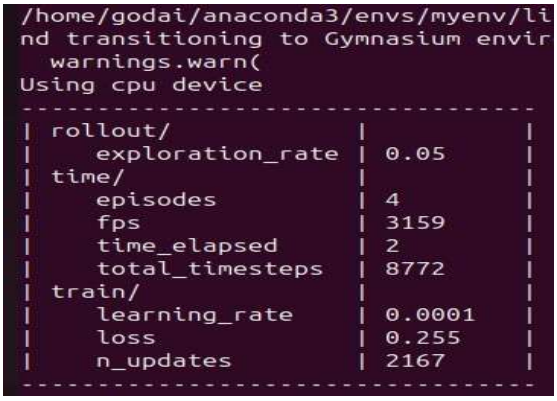


Figure 8. DQN agent training results

4. Adversarial Attack Training component

To evaluate model robustness, original and perturbed samples were analyzed.

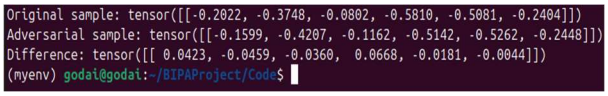


Figure 9. Adversarial Attack Analysis

Differences between the two were noticed, indicating specific changes made to mislead the model as indicated in Figure 9. This emphasizes the importance of assessing model vulnerabilities. Furthermore, to demonstrate the resilience of the model, visualizations of original and adversarial samples were plotted in Figure 10.

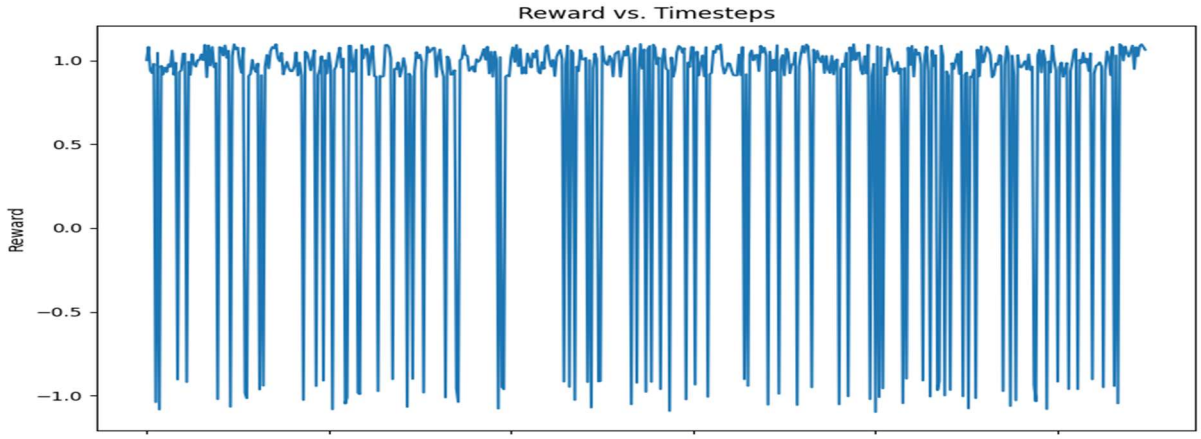


Figure 7. DQN agent training plot

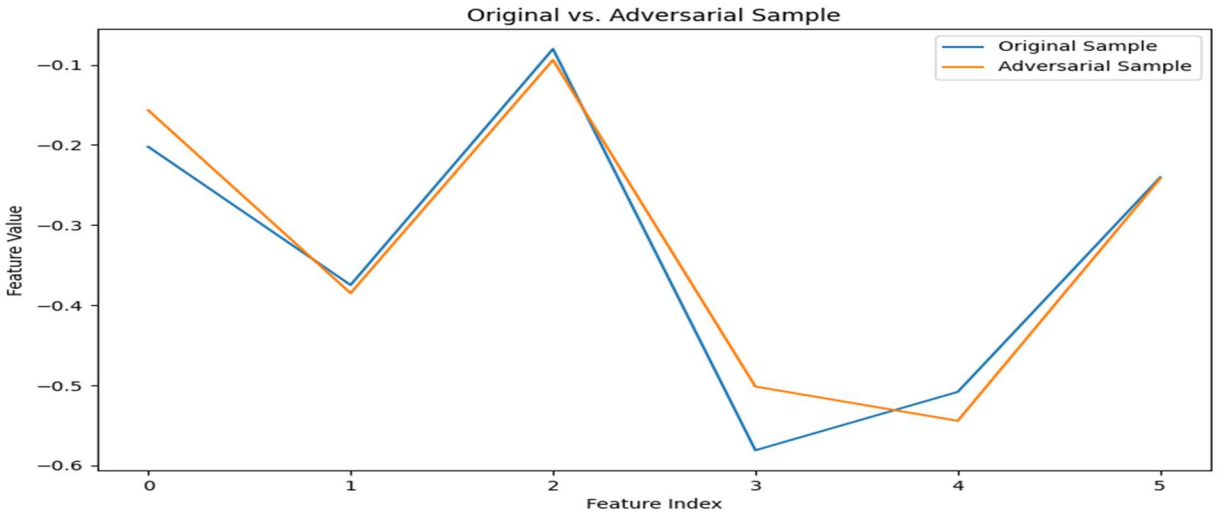


Figure 10. Original and Adversarial samples

V. Conclusion and Future works

Through this research, we demonstrated a successful construction of custom ransomware dataset and evaluated it using both deep reinforcement learning and adversarial attack. Our approach combined supervised learning for initial detection and reinforcement learning facilitated by DQN model for adaptive behavior analysis. An L2PGDP adversarial attack was then carried out to evaluate the robustness of ransomware detection model towards its security and stability. Our findings proved that deep reinforcement learning models can effectively classify benign and ransomware, however the success of our adversarial attack exposed the susceptibility of the model to adversarial attacks. Thus, our future work will focus on enhancing the robustness of deep reinforcement models against adversarial attacks.

REFERENCES

[1] Federal Bureau of Investigation, "Cyber Crime: Ransomware 2023," 2023, https://www.fbi.gov/news/stories/cyber-crime-ransomware-2023.(accessed Nov., 11, 2024).

[2] Korea Internet & Security Agency. "Cyber Threat Analysis & Sharing (C-TAS), "https://www.krcert.or.kr/en/subPage.do?menuNo=205098. (accessed Nov., 11, 2024).

[3]Federal Bureau of Investigation. "Internet Crime Complaint Center (IC3) Annual Report," 2023. https://www.ic3.gov/AnnualReport/Reports (accessed Nov., 11, 2024).

[4] Morgan, Steve. 2023. "Ransomware Statistics and Trends," Cybersecurity Ventures. https://cybersecurityventures.com/cybersecurity-almanac-2023/.(accessed Nov., 11, 2024).

[5] Gazzan, M., & Sheldon, F. T. (2024). "Novel Ransomware Detection Exploiting Uncertainty and Calibration Quality Measures Using Deep Learning," *Information*, vol. 15, no. 5, p. 262, 2024.

[6]Singh, A., Mushtaq, Z., Abosaq, H. A., Faraj, S. N. M., & Irfan, M. (2023). "Enhancing Ransomware Attack Detection Using Transfer Learning and Deep

Learning Ensemble Models on Cloud-Encrypted Data," *Electronics*, vol. 12, no. 18, 3899, Sep. 2023.

[7] Taha, N. N., and N. A. Abdullah. 2024. "Dynamic Analysis-Based Early-stage Ransomware Detection Using Deep Learning." *International Journal of Intelligent Engineering and Systems,* Vol. 17, No. 6, 2024.

[8] Moujoud, L., M. Ayache, and A. Belmekki. 2023. "A State-of-the-Art Survey on Ransomware Detection using Machine Learning and Deep Learning Techniques," *In Studies in Computational Intelligence,* pp. 183-200, Aug. 2023.

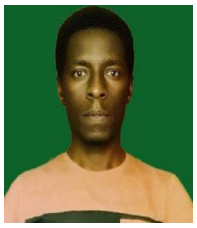
[9] Madry, Aleksander, et al. "Towards Deep Learning Models Resistant to Adversarial Attacks," *Proceedings of the 35th International Conference on Machine Learning (ICML),* 2018.

[10] Kurakin, Alexey, Ian Goodfellow, and Samy Bengio. "Adversarial Examples in the Physical World," *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2017.

[11] Connolly, L. Y., Wall, D. S., Lang, M., & Oddson, B. (2020). "Empirical Study of Ransomware Attacks on Organizations: An Assessment of Severity and Salient Factors Affecting Vulnerability," *Journal of Cybersecurity*, vol. 6, no. 1, tyaa023, 2020.

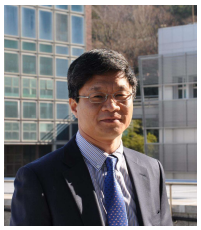
[12] Papernot, N., McDaniel, P., Wu, X., Jha, S., & Swami, A. (2016). Distillation as a Defense to Adversarial Perturbations Against Deep Neural Networks. IEEE Symposium on Security and Privacy. [Online]. Available: https://arxiv.org/abs/1511.04508

Authors



Niringiye Godfrey
He received his M.sc. degree in Data Communications and Software Engineering from Makerere University, Uganda in 2021.

HoonJae LEE
Professor, Department of Information Security, Dongseo university



ByungGook LEE
Professor, Department of Computer Engineering, Dongseo university