Ransomware Classification and Detection With Machine Learning Algorithms

Mohammad Masum∗, Md Jobair Hossain Faruk†, Hossain Shahriar‡
Kai Qian§, Dan Lo§, Muhaiminul Islam Adnan¶

∗School of Data Science, Kennesaw State University, USA
†Department of Software Engineering and Game Development, Kennesaw State University, USA
‡Department of Information Technology, Kennesaw State University, USA
§Department Computer Science; Kennesaw State University, USA
¶Institute of Natural Sciences, United International University, Bangladesh.
Email: {*mmasum∗, mhossa21†*}@students.kennesaw.edu | {*hshahria‡, kqian§, dlo2§*}@kennesaw.edu
{*adnan08mi¶*}@gmail.com

2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC) | 978-1-6654-8303-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CCWC54503.2022.9720869

***Abstract*— Malicious attacks, malware, and ransomware families pose critical security issues to cybersecurity, and it may cause catastrophic damages to computer systems, data centers, web, and mobile applications across various industries and businesses. Traditional anti-ransomware systems struggle to fight against newly created sophisticated attacks. Therefore, state-of-the-art techniques like traditional and neural network-based architectures can be immensely utilized in the development of innovative ransomware solutions. In this paper, we present a feature selection-based framework with adopting different machine learning algorithms including neural network-based architectures to classify the security level for ransomware detection and prevention. We applied multiple machine learning algorithms: Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), Logistic Regression (LR) as well as Neural Network (NN)-based classifiers on a selected number of features for ransomware classification. We performed all the experiments on one ransomware dataset to evaluate our proposed framework. The experimental results demonstrate that RF classifiers outperform other methods in terms of accuracy, F-beta, and precision scores.**

Keywords*— *Ransomware Classification, Feature Selection, Machine Learning, Neural Network, Cybersecurity

I. Introduction

Malicious applications or attacks, malware and ransomware families for instance, consistently endures to pose critical security issues to cybersecurity and it may cause catastrophic damages to computer systems, data centers, web, and mobile applications across various industries and businesses[1]–[3]. Most ransomware is designed to block and prevent targeted victims from accessing computer data by applying an indestructible encrypting methodology that can be decrypted by the attacker itself solely. Removing the ransomware leads the victim to irreversible losses, as a result, victims are forced to pay according to the attacker's demands [4]. Failure or denial to comply with the attacker's demand will lead to losing data permanently. With the help of modern technology, attackers are transforming conventional ransomware into emerging ransomware families which is more difficult in reversing a ransomware infection [5].

Ransomware is a sophisticated and variants threat affecting users worldwide that limits users from accessing their system or data, either by locking the system's screen or by encrypting and the users' files unless a ransom is paid [2]. Two primary forms of ransomware based on attack approaches include locker ransomware that denies access to the computer or device and crypto ransomware that prevents access to files or data [6]. After these attacks, it is incredibly difficult to revert without paying the extortion. Traditional ransomware detection techniques including event-based, statistical-based, and data-centric-based techniques are not adequate to combat. Therefore, implementing the highest level of optimal protection and security by adopting futuristic technology against such advanced malicious attacks should be imperative for the research community.

Novel technology, machine learning for instance in ransomware detection is a new research topic and can be immensely utilized in the development of innovative ransomware solutions [7]. Employing the application of Machine Learning (ML) methodologies enables automatic detection of malware including ransomware through their dynamic behaviors and enhances security [8]. Algorithms such as Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), Logistic Regression (LR), and Neural Network (NN)-based architectures have potential efficacy for ransomware classification and detection [9]. In this study, we conduct a comprehensive assessment and investigates the machine learning techniques for the classification of ransomware. The primary contributions of the paper as follows:

978-1-6654-8303-2/22/$31.00 ©2022 IEEE