LAUTECH Journal of Computing and Informatics (LAUJCI) – ISSN : 2714-4194
Volume 4. No.1, March 2024 – www.laujci.lautech.edu.ng

A COMPREHENSIVE REVIEW ON MACHINE LEARNING TECHNIQUES FOR THE IDENTIFICATION OF RANSOMWARE ATTACKS IN COMPUTER NETWORKS

AYINLA O. M
Postgraduate Researcher,
Department of Computer Science
Al-Hikmah University, Ilorin, Nigeria
ayinlamutiu2019@gmail.com

OYELAKIN A. M
Lecturer, Department of
Computer Science, Crescent
University, Abeokuta, Nigeria
moruff.oyelakin@cuab.edu.ng

OLOMU J. O.
Postgraduate Researcher,
Department of Computer Science
Al-Hikmah University, Ilorin, Nigeria
luability4u@gmail.com

ABSTRACT
Ransomware attacks have been identified as one of the serious threats in the cyber space. The malware poses serious security challenges to corporate networks and internet users worldwide. In response, several machine learning techniques have gained popularity for the classification of ransomware in the internet space when compared with signature-based approaches. This paper presented a comprehensive review of various studies that focus on the use of machine learning techniques for the identification of ransomware attacks in computer networks. The study collected relevant literature from various research databases by using some specific keywords and search strings that are deeply related to the topic. A good number of literatures that were obtained, were sorted and studied. The literatures were organised in different sections, arranged chronologically from the most recent to relatively older works. The publication years for the reviewed papers ranges from 2017 to 2023. The review began by exploring some relevant concepts and then shifted ground to machine learning algorithms that have been proposed for ransomware attacks identification. Thereafter, the performances of the different learning techniques used for the identification of ransomware attacks in computer networks were reported. The study argued that the review can serve as insights for future researches in this cyber security area.

Key words: Ransomware, Malware Attacks, Learning Algorithms, Malware analysis, Internet Resources.

1. INTRODUCTION

Online communities are faced with a wide range of security threats and attacks. Ransomware attacks are good examples of such attacks. Ransomware is a form of malicious software that has evolved significantly since its initial appearance in the late 1980s. The malware was originally used for personal blackmail; it has now escalated to corporate extortion, posing serious threats to business networks and the overall internet ecosystem (Khammas, 2022). These attacks have become more frequent and sophisticated in recent years, causing financial losses, operational disruptions, and reputational damage to businesses (Adamu & Awan, 2019). Ransomware operates by encrypting vital data and demanding a ransom for its release, rendering traditional signature-based detection methods ineffective. As a result, there is a need to explore more advanced strategies to combat this rapidly evolving threat.

Horduna, Lăzărescu and Simion (2023) argued that Machine learning (ML) techniques are very promising for building models that can handle ransomware identification on corporate networks and the internet. ML approach can be used to analyse large datasets for the detection of patterns, anomalies, and behavioral indicators associated with ransomware operations (Horduna et al., 2023). ML systems can be trained on diverse datasets, consisting of known ransomware samples and benign data, enabling them to differentiate between normal and malicious network activities and promptly identify ransomware threats. Considering the evolving ransomware threat landscape as well as the dynamic nature of

corporate networks, and the internet, this research seeks to identify the different ML-based methods that have been used to classify ransomware attacks.

This study aims to review various ML techniques that have been proposed for ransomware identification in networks. Through a comprehensive literature review, we aim to assess the strengths, limitations, and potential areas for improvement of existing ML-based ransomware identification approaches and then propose improve ML-based ransomware identification models in the future. The findings of this research can inform the design of more effective machine learning-based protection mechanisms against ransomware attacks and enhance the overall cybersecurity posture of enterprises.

How Ransomware Attacks Work

In a ransomware attack, private key encryption is used so as to prevent authorized users from accessing a system or data until a ransom (usually paid in Bitcoin) is provided (Chesti, Humayun, Sama, &Jhanjhi, 2020). It has to be pointed out that infected files typically have specific extensions, such as Locky, Cryptolocker, Vault, Micro, Encrypted, TTTT, XYZ, ZZZ, Petya, and others, indicating the type of ransomware that has affected them. Some examples of ransomware include WannaCry, WannaCry.F, Fusob, TorrentLocker, CryptoWall, CryptoTear, and Reveton (Jegede, Fadele, Onoja, Aimufua, &Mazadu, 2022). Ransomware can be classified into three, namely: scareware, locker ransomware, and crypto-ransomware (Chesti et al., 2020; Jegede et al., 2022; Raizza et al., 2023).

Crypto ransomware is the most widespread type of ransomware that targets computer systems and networks. It utilizes both symmetric and asymmetric encryption algorithms to encrypt files and data. Even if the ransomware is removed from an infected computer or if a compromised storage device is connected to another system, the encrypted data remains inaccessible. However, in some cases, the malicious software may spare important files to enable the victim to pay the ransom using the compromised device (Jegede et al., 2022). Figure 1 visually depicts Crypto-Ransomware, malicious software that has been increasingly prevalent in cyberattacks (Jegede et al., 2022; Kovács, 2022).

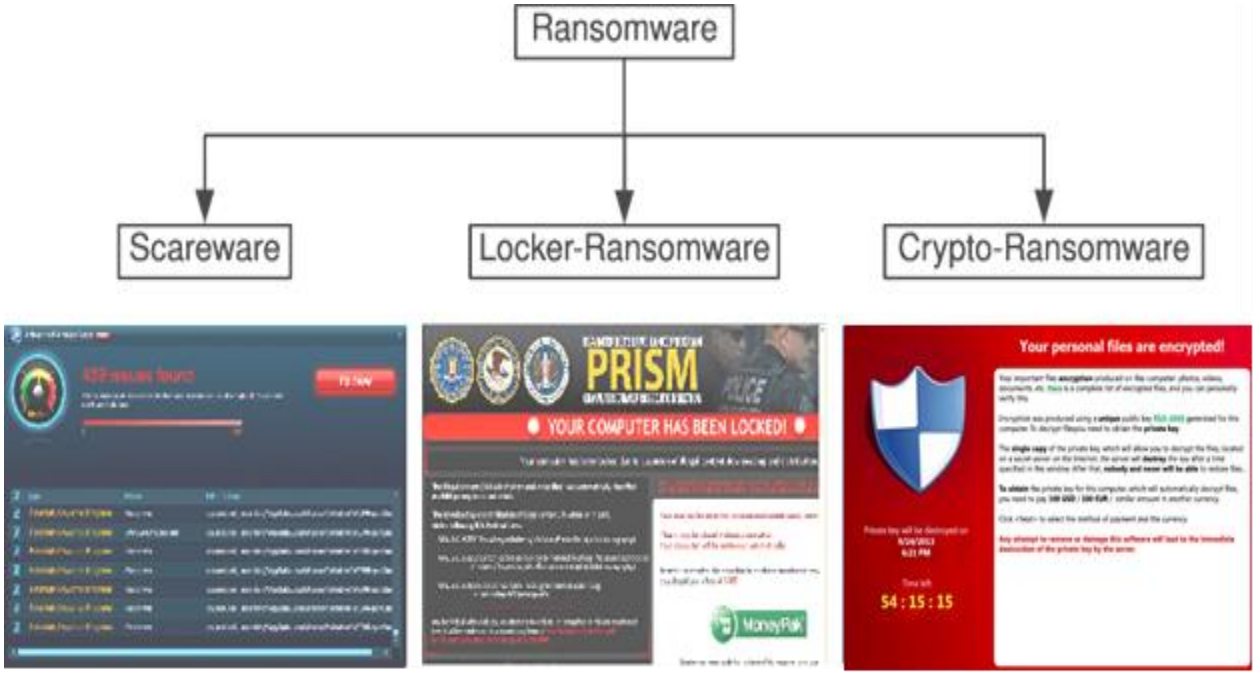

Figure 1: Categories of Ransomware (Kovács, 2022)

Locker ransomware operates differently from crypto ransomware, as it locks the victim's computer or device and demands a ransom to regain access. While the workstation is affected by Locker ransomware, the saved data remains accessible and unaltered. Once the malicious program is removed, the data can be recovered by connecting the infected storage device, such as a hard drive, to another machine. This type of ransomware is less attractive to extortionists seeking to demand money from their victims.

Scareware preys on its victims by falsely alerting them that their machines have been hijacked by ransomware. It tricks users into believing that their issues can be resolved by purchasing and installing a fake antivirus program, which is actually backed by the attacker. This deceptive tactic often leads innocent consumers to buy and install the fraudulent antivirus software due to the persistent appearance of scareware alerts (Brewer, 2016). On the other hand, human-operated malware and ransomware without data differ from typical ransomware. Cybercriminals employ human-operated ransomware to infiltrate networks or cloud infrastructure, perform privilege escalation, and target sensitive data. Instead of focusing on a single system, this type of attack actively targets an entire organization. Attackers gain access to the entire IT system, move laterally, and exploit security vulnerabilities caused by improper configurations. Ultimately, unauthorized access to privileged users' credentials paves the way for ransomware attacks on critical corporate systems that support essential business activities (Philip et al., 2018; Jegede et al., 2022).

Scareware is a non-file-based ransomware operates using the system's native and dependable functionalities to carry out attacks. This type of ransomware does not require the installation of any code on the victim's machine, making it challenging to detect during an attack. Consequently, anti-ransomware technologies may not identify any suspicious files to track while the attack is underway. Depending on the attacker's objectives, both file-based and human-operated ransomware can employ techniques such as encryption, locking, or data leakage from files (Chesti et al., 2020).

2. REVIEWED STUDIES

Machine learning methods have attracted considerable interest in the detection of ransomware in networks. This section provides is structured under some sub-sections and it involves a review of relevant and recent studies that have reported the use of ML-based techniques for ransomware identification.

Ransomware Detection Techniques

There are two primary categories of ransomware detection techniques: automated and manual. Automated methods require the use of technologies to recognize and report ransomware attacks, often in the form of software programs capable of preventing such attacks. On the other hand, manual detection techniques involve regular scanning of data and devices to identify potential indicators of attacks. This includes checking for any alterations to file extensions, verifying the accessibility of devices and files by authorized users, and monitoring changes in file extensions to ensure that data has not been tampered with or access to files has not been hindered by malware attacks. Amin (2017) proposed different techniques that can be used to address Ransomware attacks in networks. The author pointed out that the approaches are very promising in ransomware attack detection.

Cawthra, Ekstrom, Lusty, Sexton and John (2020) argued that destructive malware such as Ransomware attacks, malicious insider activity are part of the reasons why modern organizations have to step up their game in cyber security. Akhtar and Feng (2022) have opined that Artificial intelligence-based machine learning methods, including behavioral

techniques, static and dynamic analysis, deep learning, and artificial neural networks, are employed to automatically detect ransomware attacks. Deep learning algorithms are specifically addressing the limitations of supervised ransomware detection technologies, aiming to enhance the accuracy and reliability of detection results. They excel in handling unstructured data and automatically generating features (Bello et al., 2021; Sharmeen et al., 2020). Furthermore, artificial neural network (ANN) approaches are very suitable for identifying various types and variants of ransomware, whether in text or image format, due to their versatility. Their continuous learning ability makes them ideal for adapting to new ransomware data and detecting zero-day attacks (Brewer, 2016; Bello et al., 2021; Swami et al., 2021).

Ransomware detection can also be accomplished using non-AI techniques, such as packet inspection and traffic analysis. One effective strategy involves the use of online honeypots or simulated computers to observe and monitor network activity. By setting up a honeypot folder and monitoring any changes that may indicate the presence of ransomware, early detection can be achieved, minimizing its impact and preventing further damage (Chesti et al., 2020). To further report different detection techniques for ransomware, the following sub-sections report various ransomware detection approaches as pointed out in the literature.

Signature-based Detection Method of Ransomware Attacks

Signature-based detection is a conventional method that relies on identifying known ransomware signatures or patterns present in the code or behavior of the malware. It involves creating a database of recognized ransomware signatures and scanning the system or network for any matches. If a match is found, the ransomware is flagged as malicious, and appropriate actions are taken (Yamany, Elsayed, Jurcut, Abdelbaki, &Azer, 2022a; Yamany, Azer, & Abdelbaki, 2022b). One advantage of signature-based detection is its simplicity and effectiveness in detecting known ransomware variants. However, it has limitations, as it cannot detect new or unknown ransomware variants that do not match existing signatures or patterns. Furthermore, attackers can easily evade this approach by modifying the ransomware's code or behavior to avoid detection (Akhtar et al., 2022).

Heuristic-based Detection Method of Ransomware Attacks

Heuristic-based detection is a more advanced approach that identifies ransomware behavior patterns or anomalies indicative of malicious activity. It involves creating rules or heuristics that describe typical ransomware behavior and then monitoring the system or network for any deviations from these rules. If variations or abnormalities are detected, the ransomware is flagged as suspicious or malicious, and appropriate actions are taken (Yamany et al., 2022a; Yamany et al., 2022b). One advantage of heuristic-based detection is its ability to detect new or unknown ransomware variants that do not match existing signatures or patterns. It is also less prone to false positives compared to signature-based detection, as it focuses on detecting actual behavior patterns rather than static code signatures. However, heuristic-based detection has limitations, as it relies on predefined rules that may only capture some ransomware behavior patterns or anomalies. Additionally, attackers can evade detection by modifying the ransomware's behavior (Akhtar et al., 2022).

Machine Learning-based Detection Approach

Machine learning-based detection is a more sophisticated approach that involves training a machine learning model to detect ransomware based on behavior patterns or features. It requires a large dataset of benign and malicious samples to extract relevant features and train the model to classify new samples as either benign or malicious based on their characteristics (Yamany et al., 2022a; Yamany et al., 2022b). Machine learning-based detection has several

advantages, including its ability to detect new or unknown ransomware variants and adapt to changing behavior patterns over time. It is also less prone to false positives compared to signature-based and heuristic-based detection since it focuses on actual behavior patterns. However, this approach requires a representative dataset for training and is susceptible to adversarial attacks that can manipulate ransomware features to evade detection (Akhtar et al., 2022). Connolly, Wall.,Lang and Oddson (2020) carried out a study that involved an emperical analysis of Ransomware Attacks On Organizations: The authors did an assessment of the level of severity as well as some of the reasons responsible for vulnerability.

Network-based Detection Technique

Network-based detection relies on monitoring network traffic for suspicious or malicious activity indicative of a ransomware attack. It analyzes network traffic for anomalies or patterns characteristic of ransomware, such as large volumes of outbound traffic or unusual network connections (Yamany et al., 2022a; Yamany et al., 2022a). One advantage of network-based detection is its ability to detect ransomware activity even before the malware infects the system or if non-standard encryption methods are used. It is also less prone to false positives since it focuses on actual network traffic patterns. However, network-based detection requires effective network traffic analysis tools and may not capture all ransomware activity. Attackers can also evade detection by encrypting their network traffic or using stealthy communication channels (Akhtar et al., 2022).

Hybrid Detection Method

Hybrid detection is an approach that combines various ransomware detection techniques to enhance overall detection accuracy and speed. It integrates the strengths of signature-based, heuristic-based, machine learning-based, and network-based detection to create a more robust and effective detection system (Yamany et al., 2022a; Yamany et al., 2022b). One advantage of hybrid detection is its ability to overcome the limitations of individual detection approaches and improve overall accuracy and speed. It is also less prone to false positives and negatives compared to single detection methods since it combines different sources of information and analysis. However, hybrid detection is complex and resource-intensive, as it requires integrating and coordinating multiple detection systems and tools (Akhtar et al., 2022).

Metrics for Evaluating ML-based Ransomware Detection Models

Evaluating the performance of machine learning models for ransomware detection is vital. In this section, some of the basic metrics that researchers do use to evaluate the performances of ML-based models are highlighted. Some of them include: accuracy, precision, recall, F1-score, and ROC curve.

Accuracy: Accuracy is a straightforward evaluation metric that represents the percentage of correct predictions made by the model. It is calculated as the ratio of accurate predictions to the total number of predictions. However, accuracy can be misleading when dealing with imbalanced datasets, where negative samples significantly outnumber positive ones (Kok, Azween, & Jhanjhi, 2020; Masum, Faruk, Shahriar, Qian, Lo, & Adnan, 2022).

Precision: Precision refers to the percentage of true positives (correctly identified ransomware samples) out of all the samples predicted to be positive (identified as ransomware by the algorithm). The ratio of true positives to the total of true and false

positives is calculated as precision. A model with a high precision score will have a low false positive rate, reducing the likelihood of misclassifying innocent files as ransomware (Masum et al., 2022).

Recall: Recall counts the number of positive samples in the dataset that are correctly identified as true positives. The ratio of true positives to true and false negatives is computed as recall. A high recall score indicates that the model has a low incidence of false negatives, meaning it is less likely to miss actual ransomware samples (Masum et al., 2022; Azmoodeh, Dehghantanha, Conti, & Choo, 2018).

F1-score: The F1-score is a metric that combines both precision and recall to provide a balanced evaluation of the model's performance. It is the harmonic mean of precision and recall and helps assess the overall accuracy of the model in identifying both positive and negative samples (Kok et al., 2020).

ROC curve: The receiver operating characteristic (ROC) curve visually represents the performance of a binary classifier as the discrimination threshold is varied. It plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold values. The model's overall performance is evaluated using the area under the ROC curve (AUC), where higher AUC values indicate better performance (Edis, Hayman, & Vatsa, 2021).

Table 1. Studies on Machine learning techniques for ransomware detection from 2017 to 2022

Year of Publication	Author	Resolved the Issue	Utilized Technique	Result	Limitation
2022	Singh et al.	Identification of previously unrecognized ransomware families and categorization of recently detected ransomware attacks.	Utilizes process memory access privileges to achieve fast and precise detection of malware.	The accuracy ranges from 81.38% to 96.28%.	N/A
2022	Ahmed et al.	Distinguishing users affected by the Locky ransomware.	Using multiple classifiers in parallel for ransomware detection based on behavior.	Achieving high accuracy in detection with a minimal number of false positives.	NA
2019	Makinde et al.	To assess the susceptibility of a real network	Machine learning	Correlation exceeding	It replicated the

		system to a ransomware attack.	algorithms	0.8.	behavior of a limited number of users.
2018	Shaukat & Ribeiro	Ransomware identification.	(RansomWall) An integrated and hybrid approach.	Highly proficient in detecting previously unknown attacks.	NA
2017	Zahra & Shah	Detecting ransomware attacks with Cryptowall.	Blocking command and control (C&C) servers through a blocklist.	The TCP/IP header is extracted from the web proxy server, which functions as the gateway for TCP/IP traffic.	The effectiveness and accuracy of the model in detecting ransomware and its attack methods on different operating system environments were not shown through practical implementation.
2022	Singh et al.	Identification of previously unrecognized ransomware families and categorization of recently detected ransomware attacks.	Utilizes process memory access privileges to achieve fast and precise detection of malware.	The accuracy ranges from 81.38% to 96.28%.	N/A

The normal behavior of an application is evaluated from both the user's perspective and the resource perspective. A standard baseline for normal behavior is established by considering the typical or expected operation of a computer system or network. Indicators of regular

activity encompass logins, file access, user and file behaviors, resource usage, and other relevant indicators (Celdran et al., 2022).

The process involves extracting the TCP/IP header from the web proxy server, which acts as the gateway for TCP/IP traffic. It then checks the source and destination IPs against a list of forbidden Command-and-Control servers to identify ransomware. However, the model's effectiveness in detecting ransomware across various operating system environments was not demonstrated. A newer technique based on behavioral analysis and process memory access privileges now allows for rapid and accurate ransomware detection (Azmoodeh et al., 2018; Singh et al., 2022). By analyzing the access privileges and memory access patterns of files and applications, it is feasible to classify new ransomware attacks and identify previously unknown malware families. Examining the behavior and intentions of legitimate files and applications before execution proves to be advantageous. Experimental results using these various approaches demonstrate promising detection accuracy, ranging from 81.38% to 96.28%.

The duration of the learning process depends on the data required to establish a baseline representing the usual system behavior. The tool examines behavioral deviations from this baseline to identify outliers. A ransomware detection and prevention model was developed for unstructured datasets sourced from the logs of the Ecuadorian Control and Regulatory Institution (EcuCERT) (Silva & Hernandez-Alvarez, 2017).

The approach utilizes observation to detect unusual behavioral patterns associated with Windows malware. Feature selection was employed on the Log data to extract the most useful and distinctive information indicative of a ransomware attack. The extracted data enables autonomous learning algorithms to quickly and accurately identify ransomware using input features and algorithms that mimic abnormal behavioral patterns. As ransomware attacks continuously evolve with code obfuscation tools and new polymorphic variants, signature-based approaches struggle to keep up with their identification (Shaukat et al., 2018).

Traditional malware detection methods that are designed for generic attack vectors are not effective in accurately identifying cryptographic ransomware due to its unique behavioral characteristics. To address this limitation, the proposed RansomWall approach combines static and dynamic analytics to create a novel set of properties that mimic ransomware behavior. The technique enables early detection of ransomware while also providing a robust trap layer to detect previously unknown zero-day attacks. When tested against a significant number of ransomware samples, RansomWall with the Gradient Tree Boosting Algorithm achieved a high detection rate of 98.25% and an extremely low false positive rate. Additionally, the approach demonstrated impressive performance in detecting zero-day attacks compared to popular security engines. Another version of behavioral detection methodologies uses a machine learning baseline model to simulate and predict specific network user behavior patterns, enabling the identification of potential vulnerabilities or true ransomware attacks (Makinde et al., 2019).

The objective was to identify vulnerabilities in a basic network system concerning ransomware attacks. By comparing results from a simulated network with log data from the actual network system, a realistic model with a correlation above 0.8 was obtained. However, the method had limitations in accurately capturing the behavior of only a small percentage of users. To improve detection, future research should concentrate on emulating user behavior across a larger user base using big data analytics tools. Additionally, a more recent behavioral

ransomware detection approach employed two parallel classifiers (Almashhadani, Kaiiali, Sezer& O'Kane, 2019).

The research aimed to differentiate various Locky ransomware variants by focusing on early detection through behavioral analysis of ransomware network traffic. To prevent ransomware from connecting to harmful command-and-control servers and executing malicious actions, the study used a dedicated network to gather and extract crucial information from the network traffic. By analyzing the extracted properties of the Locky ransomware family using two parallel classifiers at the packet and datagram levels, the technology demonstrated high success in detecting ransomware activities on the network, while maintaining a low percentage of false positives. Additionally, the study explored ransomware attacks in an IoT environment, particularly focusing on communication and behavioral analysis using command and control (CC) server blocklists (Zahra & Shah, 2017).

Table 2. Summary of previous studies on Machine learning techniques for ransomware detection from 2017 to 2022

Year	Author	Problem Addressed	Method Used	Result
2022	Talabani & Abdulhadi	Data mining and machine learning methods for detecting ransomware show low accuracy levels.	Decision Table and Partially Decided Decision Tree.	Recall (96%), accuracy (96.01%), F-measure (95.6%), and precision (95.9%)
2020	Khammas	Ransomware identification.	The method involving the use of random forest.	97.74% of the samples are identified.
2020	Hwang et al.	A more advanced approach to ransomware detection.	the utilization of random forest in combination with Markov models.	The overall accuracy is 97.3%, with a false positive rate of 4.8% and a false negative rate of 1.5%.
2020	Ghouti et al.	Improved approach for identifying ransomware.	Employing support vector machines for analysis.	An integrated approach yields improved ransomware detection compared to using static or dynamic analysis independently.
2019	Modi	Distinguishing	Random forest and	The random

		between ransomware traffic and regular network traffic.	support vector machine are among the algorithms utilized in logistic regression.	forest achieves a detection rate of 99.9% with no false positives, which is the highest among the tested techniques.
2019	Ameer	Ransomware identification.	Examinations involving both static and dynamic aspects	Perfect detection and classification accuracy.
2018	Azmoodeh et al.	Fast and precise detection of Windows ransomware.	Netconverse, a classifier employing the J48 decision tree.	A 97.1% rate of correctly detecting positive instances.

Various enhanced machine-learning techniques have been employed to achieve precise and effective ransomware detection, aiming to overcome the limitations of existing ML-based detection tools. Among these advancements is the difficulty that detection systems, like sandbox analysis and pipelines, encounter in isolating and evaluating ransomware samples, leading to delays in the evaluation process. Adamu and Awan (2019) used a dataset with 30,000 attributes as independent variables to predict ransomware. By employing feature selection, five key qualities are identified and applied in the support vector machine technique, resulting in an 88.2% accuracy rate for ransomware detection. To reduce false positives, this hybrid approach combines the "guilt by association" hypothesis with content, metadata, and behavior-based analysis. The method also involves giving users control over recovery, employing file versioning in cloud storage to halt the process, and providing users with classification information to make informed decisions and avoid false positives. Additionally, an innovative approach for network-level ransomware detection utilizes machine learning, certificate information, and network connection details. (Modi, 2019).

The method is applicable for early detection of ransomware outbreaks through system-level monitoring. It utilizes network traffic characteristics related to connection, encryption, and certificates to model and extract ransomware features. The feature model employs support vector machines, logistic regression, and random forest to differentiate ransomware traffic, with the random forest showing the highest detection rate of 99.9% and the lowest false positive rate based on experimental results. Additionally, a more efficient detection method is a decision tree model that utilizes big data technology, Argus, for packet preprocessing, combining, and identifying malware files. (Wan, Chang, Chen & Wang, 2018).

The flow replaced packet data, resulting in a significant reduction in data size by a factor of 1000. Feature selection and concatenation were employed to gather and combine attributes from real network traffic, and six feature selection techniques were utilized to enhance classification accuracy. In a recent application, machine learning was ingeniously used to monitor the power usage of Android devices as a ransomware detection method (Azmoodeh et al., 2018). The proposed technique assesses the energy consumption of specific Android processes to differentiate ransomware from legitimate programs. By analyzing the

ransomware's unique energy pattern, the method achieves high detection (95.6%) and precision (89%) rates. Moreover, it surpasses other methods like K-Nearest Neighbor, Neural Network, Support Vector Machine, and Random Forest in terms of accuracy, recall rate, precision rate, and F-measure. Another excellent option is the advanced and portable RanDroid approach, designed to automatically detect polymorphic ransomware. (Alzahrani, Alshehri, Alshahrani, &Alharthi, 2018).

The technique compares pieces obtained from an application with a database of known ransomware variants to detect new ransomware on Android devices. It utilizes Image Similarity Measurement (ISM) and String Similarity Measurement (SSM) for comparison and employs language analysis to extract behavioral attributes and image textural strings for additional insights. This approach effectively detects ransomware using complex codes or dynamic payloads without modifying the Android OS or its core security module, and it successfully separates ransomware from other types of malware through a combination of static and dynamic analysis. (Ameer, 2019).

A significant research focus is on using supervised learning algorithms to detect ransomware. Li, Rios, and Trajković (2020) introduced RansomNet, a deep learning method that effectively classified ransomware network traffic with high accuracy. Their approach utilized a convolutional neural network (CNN) trained on a diverse dataset of ransomware samples, showcasing the efficacy of supervised learning in distinguishing between benign and ransomware traffic.

In the field of unsupervised learning, Hwang, Kim, Lee and Kim (2020) focused on clustering ransomware behavioral profiles to identify similar ransomware samples. They employed various unsupervised clustering algorithms, such as K-means and DBSCAN, to group ransomware samples based on behavioral characteristics. The results demonstrated the potential of unsupervised learning techniques in detecting and categorizing ransomware variants.

Semi-supervised learning approaches have also been explored for ransomware identification. Li et al. (2020) proposed a framework that combined supervised and unsupervised learning to enhance ransomware detection accuracy. They used labeled data to train a classifier and then utilized an unsupervised clustering algorithm to refine the classification results. The findings highlighted the effectiveness of incorporating semi-supervised learning in ransomware identification tasks.

To evaluate the performance of ML techniques for ransomware identification, it is crucial to consider multiple metrics, including accuracy, precision, recall, and F1-score. Li et al. (2020) emphasized the importance of using these metrics to obtain a comprehensive evaluation, allowing researchers to compare the effectiveness of different ML techniques and assess their suitability for real-world deployment.

3. METHODOLOGY

This study collected relevant literature on ransomware attacks from various research databases by using specific keywords related to the topic. A total of 80 pieces of literature were obtained, and 46 of them were deemed particularly relevant to the subject. The literature was organised and arranged chronologically from the most recent to relatively older works. The study period for this research spans from 2017 to 2023.Then, a review of pertinent works was conducted.

Research Sources/Repositories Used
The research repositories or sites used for sourcing for the relevant studies include: Science Direct, Research gate, IEEE Explore, ACM Conferences, Google search engine, and Google scholar.

Keywords used for the searches
The searches were conducted using a wide range of search strings. The search phrase used was: "Machine learning techniques" AND "Ransomware detection" AND "Corporate networks", "Artificial intelligence algorithms" OR "Data mining methods" OR "Predictive modeling approaches", "Ransomware identification" AND "Cybersecurity threat detection" AND "Internet security". This search phrase was considered so as to obtain a good number of relevant studies in research repositories for the review being carried out. ML models can learn to identify and classify ransomware instances with a higher degree of accuracy and speed (Horduna et al., 2023). The researchers focused on studies or literature published in the English language and contained in journals (both printed and electronic), white papers, conference proceedings, and books.

The research in the field of machine learning techniques for the identification of ransomware attacks in networks has revealed several critical gaps in existing works. Firstly, many prior studies have predominantly focused on traditional signature-based methods for ransomware detection, overlooking the potential of machine learning approaches. This gap arises due to the rapid evolution of ransomware variants, which often evade signature-based systems, necessitating more adaptive and data-driven solutions. Secondly, there is a shortage of comprehensive comparative analyses among various machine learning algorithms, hindering the identification of the most effective techniques for ransomware detection. Addressing these gaps is essential to develop robust, adaptive, and efficient solutions for detecting ransomware attacks in modern network environments.

Machine Learning for Ransomware Identification
Machine learning is a form of artificial intelligence that allows computer systems to improve their performance on a task without explicit instruction. As ransomware attacks become more prevalent and harmful, the use of machine-learning techniques is on the rise to detect and prevent such attacks. Table 3 provides a summary of the machine learning algorithms utilized, including decision trees, random forests, support vector machines, and neural networks. Each approach has its strengths and weaknesses, and the most suitable method depends on the specific scenario and data. (Celdran et al., 2022; Bello et al., 2021).

Table 3. Machine learning algorithms

Author	**Learning Algorithm Used**	**Characteristics**
(Ullah, Javaid, Salam, Ahmad, Sarwar, Shah & Abrar, 2020; Khammas, 2020) .	Decision trees	Decision trees can undergo training using features like file modifications, network traffic, and system calls to differentiate between ransomware and benign software behavior. The resulting decision tree can subsequently be applied to identify whether new data exhibits

		characteristics of ransomware.
(Ullah et al., 2020; Khammas, 2020) .	Random forests	To ensure that each tree in the forest has the same distribution and relies on the values of a randomly selected random vector, this approach employs an ensemble method that combines tree predictors. This can lead to improved performance compared to individual decision trees. Utilizing a network of decision trees, the random forest method is employed to select and predict the type of input data.
Ghouti, & Imam, 2020; Arunkumar& Kumar, 2023.	Support vector machines	Support vector machines can be trained on characteristics like system calls, network traffic, and file behavior to differentiate between ransomware and benign software behavior. Then, the resulting support vector machines can be utilized to identify whether new data represents ransomware. Support vector machines are particularly useful in ransomware detection when dealing with high-dimensional and non-linearly separable data.
Madani, Ouerdi, Boumesaoud, & Azizi, 2022; Arivudainambi, KA, Visu, et al., 2019.	Neural networks	Similar to the human brain, neural networks have the ability to detect patterns within extensive datasets. Upon receiving raw input, multi-layer neural network algorithms conduct internal operations to extract and select features, enabling a mechanism for feature extraction and selection. The primary neural network consists of an input layer, an output layer containing categorized variables, and a hidden layer, all collaborating to form an interconnected network of neurons.

Detailed Explanation of Learning Algorithms

Ullah et al. (2020) pointed out that decision trees are a simple and easy-to-understand machine learning algorithm used for ransomware detection. They recursively divide data into subsets based on feature values, creating a tree-like structure representing the decision-making process. Decision trees can analyze patterns in computer system events to identify potential ransomware attacks. However, they are prone to overfitting and sensitive to small data changes. (Ullah et al., 2020).

Decision trees serve as effective tools for ransomware detection by examining patterns in computer system events and identifying potential attack indicators. This algorithm creates a tree-like model that makes decisions based on various attributes and their relationships. In the

context of ransomware detection, a decision tree can be trained on a dataset containing both known ransomware attacks and non-malicious events to recognize common patterns associated with ransomware. For instance, the decision tree may analyze file access, network traffic, and system processes to determine if they indicate a ransomware attack. By considering multiple attributes, the decision tree assesses the likelihood of a sequence of events being a ransomware attack. Once trained, the decision tree can be deployed in real-time to detect ransomware attacks by analyzing incoming system events and comparing them to the identified patterns in the training data. If the events match the ransomware attack pattern, the system can trigger an alert or take defensive actions to prevent the spread of the attack. While decision trees are valuable tools for ransomware detection, it is essential to use them in conjunction with other cybersecurity best practices for comprehensive defense against ransomware (Akhtar et al., 2022; Azmoodeh et al., 2018; Ullah et al., 2020).

Random forests are an extension of decision trees that improve performance and reduce overfitting. By creating multiple decision trees with random feature and data selections and combining their predictions, random forests achieve better generalization. While random forests are less prone to overfitting compared to individual decision trees and can handle high-dimensional data, they come with computational complexity and interpretability challenges (Khammas, 2020).

Support vector machines are reliable for ransomware detection, as they find a hyperplane to separate data into distinct classes based on features. They can handle both linear and nonlinear boundaries, but the choice of the kernel function and parameters can affect their performance. (Ghouti et al., 2020).

Neural networks are sophisticated algorithms that excel in pattern recognition tasks, including ransomware detection. They comprise interconnected nodes that learn from input data and predict outcomes, handling complex and nonlinear relationships. However, they can be computationally expensive and require a large amount of training data. (Madani et al., 2022).

The choice of a machine learning algorithm for ransomware detection depends on the specific problem and available data. Decision trees, random forests, support vector machines, and neural networks are all effective options and have been successfully used for ransomware detection in various scenarios. (Brewer, 2016; Akhtar et al., 2022).

Feature Extraction and Selection

Machine learning techniques have been increasingly used to detect ransomware due to their ability to learn behavior patterns and detect anomalies. In this section, we will discuss the various features used for ransomware detection using machine learning, as well as the techniques used for feature selection, such as principal component analysis and correlation analysis (Hwang, Kim, Lee, & Kim, 2020; Dargahi, Dehghantanha, Bahrami, Conti, Bianchi, & Benedetto, 2019).

Attributes Employed for Ransomware Classification

File Access Patterns: Ransomware exhibits specific patterns when accessing and encrypting files, such as following alphabetical order, organizing by extension type, or based on creation date. *Alphabetical Order:* Measuring the degree to which files are accessed or encrypted in alphabetical order. *Extension-Based*: Analyzing the patterns of file access or encryption based on file extensions. *Creation Date:* Examining the order in which files are accessed or encrypted relative to their creation date. These discernible patterns can be utilized as features for detection purposes (Sheen, Asmitha & Venkatesan, 2022).

System Calls: Ransomware commonly employs system calls to carry out its malicious activities, including reading and writing files, creating processes, and communicating over networks Some features extracted from system calls may include: *Frequency of Calls:* Counting how often certain system calls related to file operations, process creation, or network communication occur. *Sequences of Calls:* Analyzing the sequences and patterns of system calls made during ransomware execution. *Resource Usage:* Examining the resource consumption associated with specific system calls. The traces of system calls can be extracted and used as features for detecting ransomware (Ullah et al., 2020).

Network Traffic: Ransomware frequently communicates with a command-and-control (CC) server to receive and deliver instructions. Analyzing network traffic can provide valuable features for identifying ransomware. Analyzing network traffic for ransomware detection can involve the extraction of various features, such as: *Communication Patterns:* Identifying unusual communication patterns between the infected host and a command-and-control server, such as frequency, timing, or size of data transfers. *Anomalies:* Detecting deviations from typical network traffic behavior, which may indicate ransomware activity. *IP and Domain Analysis:* Examining IP addresses and domains associated with network traffic to identify known malicious entities. (Madani et al., 2022).

Behavioral Analysis: The approach of behavioral analysis involves monitoring the actions of running processes and identifying anomalies indicative of malicious activity. Features like process creation, termination, and file access can be utilized in this context; Behavioral analysis involves monitoring running processes to identify malicious anomalies. Features in this context can include: *Process Creation:* Counting the number of processes created within a specific timeframe. *File Access Patterns:* Analyzing which files processes access and whether these patterns deviate from normal behavior. *API Calls:* Examining API calls made by processes for abnormal activities. (Celdran et al., 2022).

Static Analysis: In static analysis, the source code of the executable file is examined to detect any malicious behavior. Features such as code size, entropy, and string patterns can be employed for this purpose; *Code Size:* Measuring the size of the code within an executable file. *Entropy:* Calculating entropy to assess the randomness of code or data sections. *String Patterns:* Identifying specific strings or patterns within the code that are indicative of ransomware behavior. (Yamany et al., 2022).

These specific features and feature engineering techniques represent the diverse approaches used in ransomware detection, illustrating how researchers leverage various aspects of file behavior, system activity, network communication, and code analysis to enhance the accuracy of their detection models

4. CHALLENGES IN ML STUDIES FOR RANSOM CLASSIFICATION

Creating effective machine learning-based ransomware detection systems poses several challenges. This section will delve into these difficulties and discuss potential future directions in the field. Here are the challenges involved in developing effective machine learning-based ransomware detection systems:

Data Quality and Quantity: Training machine learning models effectively demands a vast amount of high-quality data. However, obtaining such data for ransomware detection is challenging due to the limited availability of labeled ransomware samples. The scarcity of labeled ransomware samples can hinder the training of effective machine learning models. Using insufficient or low-quality data may result in false positives or negatives. Collaboration

with cybersecurity organizations and sharing threat intelligence can help improve data quality and quantity (Beaman et al., 2021; McIntosh et al., 2021).

Rapidly Evolving Ransomware: Ransomware is an ever-changing threat, with new variants and attack techniques constantly emerging. Ransomware is highly adaptive, making it difficult to develop models that can keep up with new variants. Delayed detection can lead to significant damage. This dynamic nature makes it tough to build machine learning models that can accurately and swiftly detect all types of ransomware (Aboaoja, Zainal, Ghaleb, Al-rimy, Eisa&Elnour, 2022).

Adversarial Attacks: Adversarial attacks involve modifying input data to bypass the machine learning model's detection capabilities. Adversarial attacks can bypass machine learning-based detection systems, reducing their effectiveness and reliability. Malicious actors can utilize such attacks to evade ransomware detection systems, compromising their effectiveness (Aboaoja et al., 2022).

Real-Time Detection Requirements: Ransomware can spread rapidly, causing significant damage within a short period. Ransomware can propagate quickly, necessitating real-time detection to prevent further damage. Balancing detection accuracy with speed is crucial to minimize false positives while reacting promptly to threats. As a result, ransomware detection systems must be capable of real-time detection to prevent further spread and mitigate the damage caused (Gorment, Selamat, Cheng, & Krejcar, 2023).

5. LIMITATIONS OF ML TECHNIQUES IN ATTACK CLASSIFICATION

The research on machine learning techniques for the identification of ransomware attacks in networks exhibits several limitations. Firstly, the availability of high-quality labeled data remains a challenge, as ransomware attacks are relatively rare, leading to potential issues with the generalization of models. Secondly, the imbalanced nature of ransomware datasets, where benign instances significantly outnumber malicious ones, can lead to skewed model performance and practical deployment challenges. Thirdly, the rapidly evolving nature of ransomware, with new variants and evasion techniques emerging regularly, raises concerns about the adaptability of machine learning models to keep up with evolving threats. Fourthly, the computational intensity of some machine learning techniques, such as deep learning models, may hinder their practicality in resource-constrained environments. Lastly, ethical and privacy concerns associated with network monitoring and the potential for false positives require careful consideration in the deployment of machine learning-based ransomware detection systems.

6. DISCUSSION OF FINDINGS

Based on the various studies reviewed, this study agreed that application of machine learning techniques for ransomware detection has become increasingly common. However, this practice faces several challenges in data collection and preprocessing (Beaman, Barkworth, Akande, Hakak, & Khan, 2021; McIntosh, Kayes, Chen, Ng & Watters, 2021). One major challenge lies in obtaining publicly available datasets that contain real-world ransomware samples, as victims are often hesitant to report such attacks due to their sensitive nature. Consequently, researchers often resort to synthetic datasets or those generated from sandbox environments, which might not fully capture the complexity and variability of actual ransomware attacks (Philip et al., 2018). Additionally, the diversity of ransomware families and variants requires a large and diverse dataset to ensure comprehensive coverage. Moreover, ransomware behavior can differ based on the victim's system and network

environment, making it difficult to generalize detection models across various contexts (Chesti et al., 2020; Beaman et al., 2021).

Preprocessing of datasets for ransomware detection presents its own set of challenges. Ransomware often employs obfuscation techniques to evade detection, such as encrypting the payload or utilizing anti-analysis mechanisms. This can complicate the extraction of relevant data features and identifying patterns that distinguish ransomware from benign software. Furthermore, ransomware may employ legitimate system functions that are hard to differentiate from malicious behavior, necessitating advanced feature engineering and modeling techniques (Beaman et al., 2021).

7. CONCLUSION AND FUTURE WORK

This review highlights the significant role that machine learning techniques play in the identification of ransomware on corporate networks and the internet. Ransomware attacks have become a major concern in the cybersecurity landscape, causing considerable damage to individuals and organizations alike. As a result, the development of effective ransomware detection methods has become a critical area of research.

Throughout this review, various traditional machine learning algorithms, including decision trees, random forests, support vector machines, neural networks and deep learning approaches, have been explored for their potential in ransomware detection. These techniques have demonstrated promising results in analyzing system events, network traffic, and file behavior to identify ransomware patterns effectively.

However, the review also emphasised the challenges associated with ransomware detection. The availability of high-quality and diverse datasets remains a significant obstacle, as real-world ransomware samples are often challenging to obtain due to the sensitive nature of the data. Additionally, the rapidly evolving nature of ransomware and the potential for adversarial attacks require continuous updates and improvements to machine learning models.

To enhance ransomware detection on corporate networks and the internet, it is crucial for researchers, cybersecurity professionals, and organizations to collaborate and share knowledge and resources. By working together, the development of more robust and accurate machine learning-based ransomware detection systems can be achieved.

Looking ahead, future research in this area should focus on addressing the data challenges, exploring novel feature selection techniques, and further refining machine learning models to keep pace with the ever-evolving ransomware landscape. Additionally, efforts to implement real-time detection capabilities and enhance the resilience of the detection systems against adversarial attacks will be instrumental in bolstering cybersecurity defenses.

In conclusion, machine learning techniques offer great promise in the fight against ransomware on corporate networks and the internet. By addressing the identified challenges and pursuing collaborative efforts, the cybersecurity community can make significant strides in detecting and mitigating ransomware threats, ultimately safeguarding critical data and systems from potential harm or damage

8. REFERENCES

[1] Aboaoja, F. A., Zainal, A., Ghaleb, F. A., Al-rimy, B. A. S., Eisa, T. A. E., &Elnour, A. A. H. (2022). Malware detection issues, challenges, and future directions: A survey. *Applied Sciences*, 12, 8482.

[2] Adamu, U., & Awan, I. (2019). Ransomware prediction using supervised learning algorithms. *In Proceedings of the 2019 7th International Conference on Future Internet of Things and Cloud (FiCloud)* (pp. 57-63). IEEE.

[3] Ahmed, Y. A., Huda, S., Al-rimy, B. A. S., Alharbi, N., Saeed, F., Ghaleb, F. A., & Ali, I. M. (2022). A weighted minimum redundancy maximum relevance technique for ransomware early detection in industrial IoT. *Sustainability*, 14, 1231.

[4] Akhtar, M. S., & Feng, T. (2022). Malware Analysis and Detection Using Machine Learning Algorithms. *Symmetry*, 14, 2304.

[5] Almashhadani, A. O., Kaiiali, M., Sezer, S., & O'Kane, P. (2019). A multi-classifier network-based crypto ransomware detection system: A case study of Locky ransomware. *IEEE Access*, 7, 47053-47067.

[6] Alzahrani, A., Alshehri, A., Alshahrani, H., & Alharthi, R., et al. (2018). Randroid: Structural similarity approach for detecting ransomware applications in the Android platform. *In Proceedings of the 2018 IEEE International Conference on Electro/Information Technology (EIT)* (pp. 0892-0897). *IEEE*.

[7] Ameer, M. (2019). Android ransomware detection using machine learning techniques to mitigate adversarial evasion attacks. *Capital University of Science and Technology, Islamabad, Pakistan*.

[8] Amin Kharraz (2017). Techniques and Solutions for Addressing Ransomware Attacks, *A PhD thesis submitted for the award of Doctor of Philosophy in the field of Information Assurance, College of Computer and Information Science Northeastern University* retrieved from https://repository.library.northeastern.edu/files/neu:cj82rg77w/fulltext.pdf

[9] Arivudainambi, D., KA, V. K., Visu, P., et al. (2019). Malware traffic classification using principal component analysis and artificial neural network for extreme surveillance. *Computer Communications*, 147, 50-57.

[10] Arunkumar, M., & Kumar, K. A. (2023). GOSVM: Gannet optimization based support vector machine for malicious attack detection in cloud environment. *International Journal of Information Technology*, 1-8.

[11] Azmoodeh, A., Dehghantanha, A., Conti, M., & Choo, K. K. R. (2018). Detecting crypto-ransomware in IoT networks based on energy consumption footprint. *Journal of Ambient Intelligence and Humanized Computing*, 9, 1141-1152.

[12] Beaman, C., Barkworth, A., Akande, T. D., Hakak, S., & Khan, M. K. (2021). Ransomware: Recent advances, analysis, challenges and future research directions. *Computers & Security*, 111, 102490.

[13] Bello, I., Chiroma, H., Abdullahi, U. A., Gital, A. Y., Jauro, F., Khan, A., ... & Abdulhamid, S. M. (2021). Detecting ransomware attacks using intelligent algorithms: Recent development and next direction from deep learning and big data perspectives. *Journal of Ambient Intelligence and Humanized Computing*, 12, 8699-8717.

[14] Brewer, R. (2016). Ransomware attacks: detection, prevention and cure. *Network Security*, 2016, 5-9.

[15] Cawthra Jennifer, Ekstrom Michael, Lusty Lauren, Sexton Julian & John Sweetnam (2020). Data Integrity: Detecting and Responding to Ransomware And Other Destructive Events, *Nist Special Publication* 1800-26, Https://Nvlpubs.Nist.Gov/Nistpubs/Specialpublications/Nist.Sp.1800-26.Pdf

[16] Celdran, A. H., Sanchez, P. M. S., Castillo, M. A., Bovet, G., Perez, G. M., & Stiller, B. (2022). Intelligent and behavioral-based detection of malware in IoT spectrum sensors. *International Journal of Information Security*, 1-21.

[17] Chesti, I. A., Humayun, M., Sama, N. U., &Jhanjhi, N. (2020). Evolution, mitigation, and prevention of ransomware. *In Proceedings of the 2020 2nd International Conference on Computer and Information Sciences (ICCIS)* (pp. 1-6). *IEEE*.

[18] Connollylena Yuryna, Wall David S.,Lang Michael &Oddson Bruce (2020).An Empirical Study of Ransomware Attacks On Organizations: An Assessment of Severity and Salient Factors Affecting Vulnerability, *Journal of Cybersecurity*, 1–18, Doi: 10.1093/Cybsec/Tyaa023

[19] Dargahi, T., Dehghantanha, A., Bahrami, P. N., Conti, M., Bianchi, G., & Benedetto, L. (2019). A Cyber-Kill-Chain based taxonomy of crypto-ransomware features. *Journal of Computer Virology and Hacking Techniques*, 15, 277-305.

[20] Edis, D., Hayman, T., &Vatsa, A. (2021). Understanding Complex Malware. *In Proceedings of the 2021 IEEE Integrated STEM Education Conference (ISEC)* (pp. 1-2). IEEE.

[21] Ghouti, L., & Imam, M. (2020). Malware classification using compact image features and multiclass support vector machines. *IET Information Security*, 14, 419-429.

[22] Gorment, N. Z., Selamat, A., Cheng, L. K., &Krejcar, O. (2023). Machine Learning Algorithm for Malware Detection: Taxonomy, Current Challenges and Future Directions. *IEEE Access*.

[23] Horduna, M., Lăzărescu, S. M., &Simion, E. (2023). A note on machine learning applied in ransomware detection. *Cryptology ePrint Archive*.

[24] Hwang, J., Kim, J., Lee, S., & Kim, K. (2020). Two-stage ransomware detection using dynamic analysis and machine learning techniques. *Wireless Personal Communications*, 112, 2597-2609.

[25] Jegede, A., Fadele, A., Onoja, M., Aimufua, G., &Mazadu, I. J. (2022). Trends and Future Directions in Automated Ransomware Detection. *Journal of Computing and Social Informatics*, 1, 17-41.

[26] Khammas, B. M. (2020). Ransomware detection using random forest technique. *ICT Express*, 6, 325-331.

[27] Khammas, B. M. (2022). Comparative analysis of various machine learning algorithms for ransomware detection. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 20(1), 43-51.

[28] Kok, S., Azween, A., &Jhanjhi, N. (2020). Evaluation metric for crypto-ransomware detection using machine learning. *Journal of Information Security and Applications*, 55, 102646.

[29] Kovács, A. M. 2022. Ransomware: a comprehensive study of the exponentially increasing cybersecurity threat. *Insights into Regional Development*, 4(2), 96-104.

[30] Li Z, Rios ALG &Trajkovic L (2020) Detecting internet worms, ransomware, and blackouts using recurrent neural networks. *2020 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, Toronto, ON, Canada*, pp 2165–2172

[31] Madani, H., Ouerdi, N., Boumesaoud, A., &Azizi, A. (2022). Classification of ransomware using different types of neural networks. *Scientific Reports*, 12, 1-11.

[32] Makinde, O., Sangodoyin, A., Mohammed, B., Neagu, D., &Adamu, U. (2019). Distributed network behavior prediction using machine learning and agent-based microsimulation. *In Proceedings of the 2019 7th International Conference on Future Internet of Things and Cloud (FiCloud)* (pp. 182-188). *IEEE*.

[33] Masum, M., Faruk, M. J. H., Shahriar, H., Qian, K., Lo, D., & Adnan, M. I. (2022). Ransomware classification and detection with machine learning algorithms. *In Proceedings of the 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0316-0322). *IEEE*.

[34] McIntosh, T., Kayes, A., Chen, Y. P. P., Ng, A., & Watters, P. (2021). Ransomware mitigation in the modern era: A comprehensive review, research challenges, and future directions. *ACM Computing Surveys (CSUR)*, 54, 1-36.

[35] Modi, J. (2019). Detecting ransomware in encrypted network traffic using machine learning. *PhD thesis*.

[36] Philip, K., Sakir, S., & Domhnall, C. (2018). Evolution of ransomware. *IET Netw*, 7, 321-327.

[37] Raizza, A., & Algarni, A. (2023). Ransomware Detection using Machine Learning: Survey.

[38] Sharmeen, S., Ahmed, Y. A., Huda, S., Koger, B. S., & Hassan, M. M. (2020). Avoiding future digital extortion through robust protection against ransomware threats using deep learning based adaptive approaches. *IEEE Access*, 8, 24522-24534.

[39] Shaukat, S. K., & Ribeiro, V. J. (2018). RansomWall: A layered defense system against cryptographic ransomware attacks using machine learning. *In Proceedings of the 2018 10th international conference on communication systems & networks (COMSNETS)* (pp. 356-363). IEEE.

[40] Sheen, S., Asmitha, K., &Venkatesan, S. (2022). R-Sentry: Deception-based ransomware detection using file access patterns. *Computers and Electrical Engineering*, 103, 108346.

[41] Silva, J. A. H., & Hernandez-Alvarez, M. (2017). Large scale ransomware detection by cognitive security. *In Proceedings of the 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)* (pp. 1-4). *IEEE*.

[42] Singh, A., Ikuesan, R. A., & Venter, H. (2022). Ransomware detection using process memory. *arXiv preprint arXiv*:2203.16871.

[43] Swami, S., Swami, M., &Nidhi, N. (2021). Ransomware Detection System and Analysis Using Latest Tool. *International Journal of Advanced Research in Science, Communication, and Technology*, 7, 2581-9429.

[44] Talabani, H. S., &Abdulhadi, H. M. T. (2022). Bitcoin ransomware detection employing rule-based algorithms. *Science Journal of the University of Zakho*, 10, 5-10.

[45] Ullah, F., Javaid, Q., Salam, A., Ahmad, M., Sarwar, N., Shah, D., &Abrar, M. (2020). Modified decision tree technique for ransomware detection at runtime through API Calls. *Scientific Programming*, 2020.

[46] Wan, Y. L., Chang, J. C., Chen, R. J., & Wang, S. J. (2018). Feature-selection-based ransomware detection with machine learning of data analysis. *In Proceedings of the 2018 3rd international conference on computer and communication systems (ICCCS)* (pp. 85-88). *IEEE*.

[47] Yamany, B., Azer, M. A., & Abdelbaki, N. (2022a). Ransomware Clustering and Classification using Similarity Matrix. In Proceedings of the 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (pp. 41-46). IEEE.

[48] Yamany, B., Elsayed, M. S., Jurcut, A. D., Abdelbaki, N., & Azer, M. A. (2022b). A New Scheme for Ransomware Classification and Clustering Using Static Features. *Electronics*, 11, 3307.

[49] Zahra, A., & Shah, M. A. (2017). IoT-based ransomware growth rate evaluation and detection using command and control blacklisting. *In Proceedings of the 2017 23rd international conference on automation and computing (icac)* (pp. 1-6). *IEEE*.

Author's Brief Profile

Ayinla O. M. (MNCS) is currently a master student in the Department of Computer Science at Al-Hikmah University, Ilorin, Nigeria. He received the Bachelor's degree in Mathematics and Computer Science from

National Open University of Nigeria (NOUN). Currently a Postgraduate Researcher, Department of Computer Science Al-Hikmah University, Nigeria. His research areas include Machine Learning, Cybersecurity, Data and Computer Networks and Computational Sciences. He is happily married with kids. He can be reached through e-mail ayinlamutiu2019@gmail.com.

Oyelakin A. M. (Ph.D, MNCS) received his National Diploma (ND) and First Degree in Computer Science from Federal Polytechnic, Offa and University of Ilorin respectively. He finished with Distinction and Second Class Upper respectively. He had his Master in Computer Science from University of Lagos, Akoka, Lagos in 2014. He equally obtained PhD in Computer Science from University of Ilorin, Ilorin, Nigeria in 2021. Prior to joining academic on full-time basis, he has worked in different capacities as IT personnel in Information Technology, Oil Servicing and Aviation companies. He has published more than thirty-five papers in refereed journals and conference proceedings. His areas of research interest include Data and Computer Networks, Security/Privacy, Mobile Computing, Machine Learning and Medical Image Segmentation. He is a member of Nigeria Computer Science and Internet Society. He is happily married with kids. He can be reached through his official email: moruff.oyelakin@cuab.edu.ng.

Olomu J. O. is currently a master student in the Department of Computer Science at Al-Hikmah University, Ilorin, Nigeria. He holds a Postgraduate Diploma in Computer Science from University of Ilorin, Ilorin, Nigeria; and he's currently a Postgraduate Researcher, Department of Computer Science, Al-Hikmah University, Ilorin, Nigeria, Nigeria. His research areas include Machine Learning, Cybersecurity and Computational Sciences. He is happily married with kids. He can be reached by phone through email luability4u@gmail.com