Review

Not peer-reviewed version

Ransomware Detection using Machine Learning: Survey

Amjad Raizza * and Abdulmohsen Algarni

Posted Date: 23 May 2023

doi: 10.20944/preprints202305.1635.v1

Keywords: Machine Learning; Ransomware Techniques; Cybersecurity; Ransomware Detection; Ransomware Attacks

Article

Ransomware Detection using Machine Learning: Survey

Amjad ALRaizza [1,†,‡,*], **Abdulmohsen Algarni** [2,‡]

[1] Information Systems Department, King Khalid University, Abha 62529, Saudi Arabia
[*] Correspondence: amjadraizza@gmail.com

Abstract: significant security threats to Ransomware attacks provide serious security hazards to personal and corporate data and information. The owners of computer-based resources suffer serious verification and privacy violations, monetary losses, and reputational damage due to a successful ransomware assault. As a result, it is reported critically, accurately, and swiftly identifying ransomware. Numerous methods have been proposed for ransomware, each with pros and cons. The main objective of this study is to discuss current trends and potential future debates on automated ransomware detection. The document includes an overview of ransomware, a timeline of assaults, and details on their background. It also provides a comprehensive study of existing methods for identifying, avoiding, minimizing, and recovering from ransomware. An analysis of studies between 2017 and 2022 is another advantage of the study. This provides readers with up-to-date knowledge of the most recent developments in ransomware detection. It also highlights advancements in methods for combating ransomware attacks. In conclusion, this study highlights unanswered concerns and potential research challenges in ransomware detection.

Keywords: machine learning; ransomware techniques; cybersecurity; ransomware detection; ransomware attacks

1. Introduction

The rapid proliferation of ransomware attacks has emerged as one of the most significant cybersecurity threats facing organizations today. In recent years, ransomware has become an increasingly popular tool for cybercriminals to extort money from victims by encrypting their data and demanding payment for a decryption key. The impact of ransomware attacks has been felt across all industries, from healthcare and finance to government and education. Given the high stakes involved, it is crucial to understand the nature of ransomware attacks, how they spread, and the potential consequences of falling victim to one [1]. The importance of research in this area cannot be overstated. With the threat of ransomware attacks continuing to grow, there is a pressing need for scholars and practitioners to delve deeper into the problem and identify effective strategies for prevention and mitigation. This paper aims to contribute to this effort by providing a comprehensive overview of the ransomware threat landscape, analyzing the factors that contribute to the spread of ransomware, and exploring potential avenues for future research. By shedding light on this critical issue, we hope to help individuals and organizations better protect themselves against ransomware attacks and mitigate the potential damage caused by these malicious programs [1].

The paper is novel in the following areas: Section 2 introduces the ransomware concept and how it works. It also discusses the different types of ransomware attacks, such as encrypting ransomware, Locker ransomware, and Scareware. Section 3 provides an in-depth analysis of the evolution of ransomware over a period of twelve years. Section 4 provides an overview of the existing ransomware detection techniques, including signature-based detection, behavior-based detection, and machine learning-based detection. Also, discuss the different evaluation metrics used for measuring the performance of machine learning models for ransomware detection. Section 5 focuses on the use of machine learning techniques for ransomware detection. It discusses the different machine learning algorithms used for this purpose, such as decision trees, random forests, support vector machines, and neural networks. It also has different features used for ransomware detection using machine learning. It will also cover the techniques used for feature selection. And provide studies of machine learning-based ransomware detection systems developed by researchers. It discusses the methodology used, the performance achieved, and the limitations of each system. Section 6 discusses the challenges in collecting and preprocessing data for ransomware detection using machine learning. Section 7 discusses the challenges in developing effective machine learning-based ransomware detection systems. It also highlights future directions in this field, such as developing more robust and accurate models, incorporating real-time detection capabilities, and addressing the issue of adversarial attacks. Section 8 concludes what has been achieved in this novel. This novel offers a valuable resource for researchers and practitioners interested in developing effective ransomware detection systems using machine learning techniques.

2. Background

Ransomware encrypts information or computer systems and prevents unauthorized users from accessing them. Ransomware attacks use tactics, techniques, and procedures that could lock the computer or encrypt data and are challenging for a computer professional to undo. It might also steal private information from victims' PCs and network systems. Individual PCs, commercial systems (and the data and software they contain), and industrial control systems are all potential targets for ransomware. Additionally, it emphasizes the variety of sensors that Internet of Things (IoT) users employ [1]. A ransomware attack employs private key encryption to prevent authorized users from accessing a system or data unless they pay a ransom (cash), typically in Bitcoin [2]. Ransomware operations may include data exfiltration techniques. Hackers steal private information from vulnerable networks and threaten to release it if the owner does not pay a ransom. The infection is disseminated through malicious advertising, email attachments, and connections to rogue websites. The attacker also sends a file (or files) with instructions for paying the ransom. Once the attacker has verified that the ransom has been paid, the victim can access the decryption key [3].

Files with encryption or ransomware infections frequently include extensions. Locky, Cryptolocker, Vault, Micro, Encrypted, TTTT, XYZ, ZZZ, Petya, etc. Each file's extension indicates the sort of ransomware that affected it. Examples of ransomware include WannaCry, WannaCry.F, Fusob, TorrentLocker, CryptoWall, CryptoTear, and Reveton [4]. Figure 1 illustrates the classification of ransomware into three categories: scareware, locker ransomware, and crypto-ransomware [2,4].

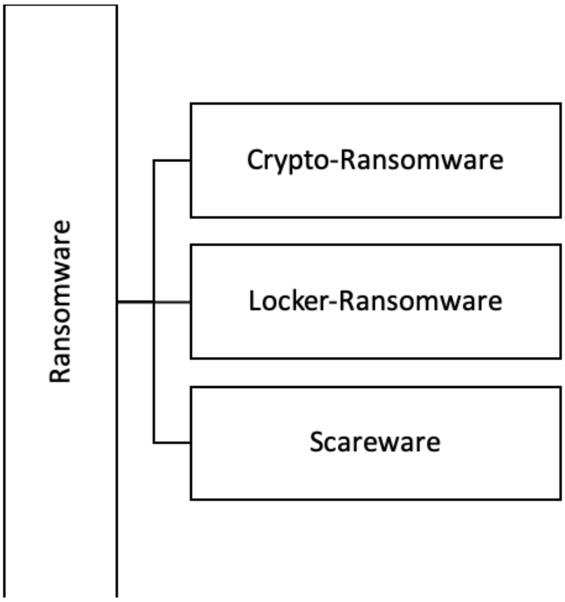
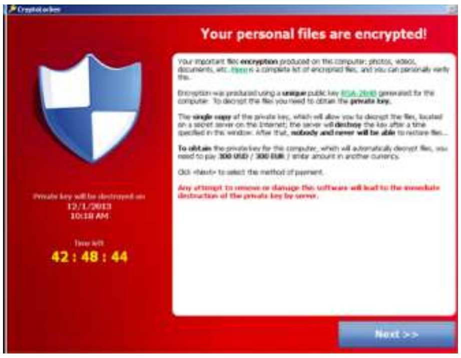


Figure 1. Types of ransomware.

Crypto is the most prevalent ransomware that targets computer systems and networks. Ransomware encrypts files and data using symmetric and asymmetric encryption algorithms. Even if the malicious software is removed from an infected computer or a compromised storage device is introduced into another system, crypto-ransomware renders encrypted data unusable. Because the malware frequently does not corrupt imported essential data, the compromised device can still be used to pay the ransom [4]. Figure 2 provides a visual representation of Crypto-Ransomware, a form of malicious software that is becoming increasingly prevalent in cyberattacks. [4].



CryptoLocker

Your personal files are encrypted!

Your important files encryption produced on this computer: photos, videos, documents, etc. Here is a complete list of encrypted files, and you can personally verify this.

Encryption was produced using a unique public key RSA-2048 generated for this computer. To decrypt the files you need to obtain the private key.

The single copy of the private key, which will allow you to decrypt the files, located on a secret server on the Internet; the server will destroy the key after a time specified in this window. After that, nobody and never will be able to restore files...

To obtain the private key for this computer, which will automatically decrypt files, you need to pay 300 USD / 300 EUR / similar amount in another currency.

Click <Next> to select the method of payment.

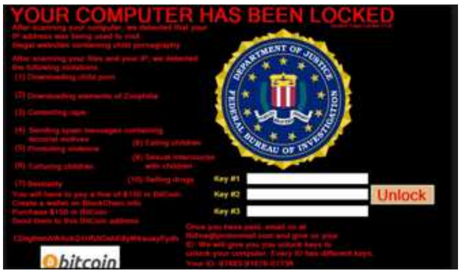
Any attempt to remove or damage this software will lead to the immediate destruction of the private key by server.

Private key will be destroyed on
12/1/2013
10:18 AM

Time left
42 : 48 : 44

Figure 2. Crypto-Ransomware.

However, by locking a computer or other device and demanding money, locker ransomware prevents its owner from using it. The workstation is affected by the Locker ransomware, but saved data is not rendered inaccessible. Once the malicious program has been eliminated, the data has not been altered. The data is often recoverable by connecting the infected storage device, such as a hard drive, to another machine. Individuals wanting to extort money from assault victims will not be drawn to locker ransomware. Figure 3 provides a visual representation of the Locker-Ransomware, a form of malicious software that is becoming increasingly prevalent in cyberattacks. [4].



YOUR COMPUTER HAS BEEN LOCKED

After scanning your computer, we detected that your IP address was being used to visit illegal websites containing child pornography

After scanning your files and your IP, we detected the following violations

(1) Downloading child porn

(2) Downloading elements of Zoophilia

(3) Committing rape

(4) Sending spam messages containing terrorist motives

(5) Promoting violence

(6) Torturing children

(7) Bestiality

(8) Eating children

(9) Sexual intercourse with children

(10) Selling drugs

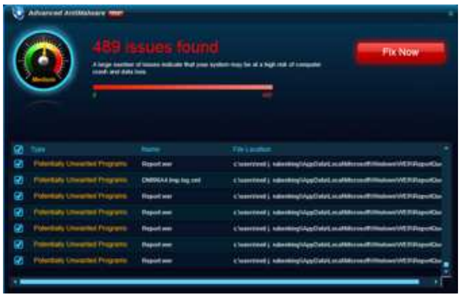
You will have to pay a fine of $150 in BitCoin.
Create a wallet on BlockChain.info
Purchase $150 in BitCoin
Send them to this BitCoin address

[illegible]

Once you have paid, email us at [illegible] and give us your ID. We will give you you unlock keys to unlock your computer. Every ID has different keys.
Your ID: [illegible]

Figure 3. Locker-Ransomware.

Scareware preys on its victims by informing them that their machines have been hijacked and promising to eradicate the ransomware using a false antivirus program backed by the attacker. Numerous innocent consumers buy and install fake antivirus software due to the scareware alert's frequent appearance [5]. Human-operated malware and ransomware without data are further different from ransomware. Cybercriminals also employ human-operated ransomware to break into networks or cloud infrastructure, carry out privilege escalation, and launch attacks on sensitive data. Instead of simply one system, the attack actively targets an entire organization. Attackers typically access a whole IT system, move laterally, and exploit flaws by improper security configurations. Ultimately unauthorized access to privileged users' credentials leads to ransomware assaults on IT systems that enable crucial corporate activities [3,4]. Figure 4 provides a visual representation of the Scareware, a form of malicious software that is becoming increasingly prevalent in cyberattacks. [4].



Advanced AntiMalware

489 issues found

A large number of issues indicate that your system may be at a high risk of computer crash and data loss.

Fix Now

☑ Type	Name	File Location
☑ Potentially Unwanted Programs	Report.wer	[illegible]
☑ Potentially Unwanted Programs	[illegible]	[illegible]
☑ Potentially Unwanted Programs	Report.wer	[illegible]
☑ Potentially Unwanted Programs	Report.wer	[illegible]
☑ Potentially Unwanted Programs	Report.wer	[illegible]
☑ Potentially Unwanted Programs	Report.wer	[illegible]
☑ Potentially Unwanted Programs	Report.wer	[illegible]

Figure 4. Scareware ransomware.

However, ransomware without files uses a native and reliable system to launch attacks. It is difficult to identify the attack because no code needs to be placed on the victim's machine for it to work. As a result, anti-ransomware technologies do not find any suspicious files to trace during an attack. Depending on the attacker's intentions, file-based and human-operated ransomware can encrypt, lock, or leak data from files [2].

Ransomware poses a danger to businesses' technology and files. Until the ransom is paid, infected files or compromised devices are locked out of reach, typically with Bitcoin. The decryption key is frequently withheld even after a victim pays the ransom the hackers want. They periodically try to use the attacker's key to decrypt the data, which damages the system's stored files. Technology advancements like ransomware development kits, ransomware-as-a-service, and bitcoins are to blame for the ongoing rise in ransomware attacks on desktop PCs, networks, and mobile devices [2]. Attacks using ransomware cost businesses and individuals hundreds of millions yearly [3]. New malware kinds are created thanks to the enormous cash benefits that hackers gain from ransomware assaults. Since 2013, numerous ransomware variants have appeared. Therefore new, effective, and reliable techniques are needed to detect, prevent, and mitigate ransomware attacks. Different ransomware strains cannot be created using conventional antivirus software or other intrusion detection systems. Significant loss People and companies endure significant losses for this [1,6].

3. Evolution of Ransomware

Ransomware attacks have been around since the late 1980s; Joseph Popp showcased the first instance of ransomware. This attack utilized symmetric-key encryption to take control of victims' hard drives and request a ransom. The flaw in this system was that the same key was used for encryption and decryption, making it vulnerable. As a result, it would be possible to research the AIDS ransomware (also known as PC Cyborg) to find the decryption key and create a solution for the malware's encryption. Ransomware attacks have continued evolving and becoming more sophisticated in recent years, making them a significant threat to individuals and organizations [7]. A brief timeline of various potent ransomware attacks is shown in Table 1. The table, an excerpt from a timeline of the most significant ransomware attacks from 2012 to 2023, contains essential information on the evolution of ransomware based on the year the ransomware first appeared, its name, and its primary description [2,3,7].

Table 1. Brief chronology of major ransomware attacks from 2012 to 2022.

References	Year	Name of the ransomware	Description
[4]	1989	AIDS Trojan	The first known ransomware attack, the AIDS Trojan, was distributed on floppy disks and demanded a payment of $ 189 to unlock infected files.
[5]	2012	Reveton	ransomware that posed as law enforcement and demanded payment for supposed illegal activities.
[7]	2013	CryptoLocker	one of the first widespread ransomware attacks that used encryption to lock victims' files.
[8]	2014	CryptoWall	A variant of CryptoLocker that caused millions of dollars in damages.
[3]	2015	TeslaCrypt	A ransomware strain that targeted gamers and encrypted game-related files.
[9]	2016	Locky	Ransomware that was spread through malicious email attachments.
[3]	2017	WannaCry	A ransomware attack affecting over 200,000 systems across 150 different countries.
[10]	2018	SamSam	A ransomware attack that targeted hospitals, municipalities, and other organizations.
[3]	2019	Ryuk	A ransomware attack that caused significant damage to several companies and organizations.
[11]	2020	Maze	A ransomware attack that encrypted victims' files and threatened to leak sensitive data if the ransom was not paid.

References	Year	Name of the ransomware	Description
[3]	2021	REvil/Sodinokibi	A ransomware attack that targeted Kaseya, a software company, and affected over 1,500 businesses worldwide.
[12]	2022	Royal Ransomware	A ransomware attack that encrypted victims and demands a ransom payment in order to decrypt them, targets businesses, governments, and healthcare organizations, and the victims are mostly from the United States.
[12]	2023	LockBit Ransomware	A ransomware attack that encrypts the files and demands payment in exchange for the decryption key. often in conjunction with phishing emails or other social engineering techniques.

Ransomware has become a popular tool for cybercriminals to extort money from individuals and organizations. As technology advances, preventing such attacks is more challenging. It is essential to remain vigilant and take appropriate measures to protect against these threats, such as keeping software up-to-date and regularly backing up important data [5]. There are six levels, which can be summarized as follows, and they are adapted from [13] and shown in figure 5.

- Distribution Campaign: The attacker silently induces the victim to download the infection-starting dropper code. The attacker uses methods including email phishing, social engineering, and others.
- Malicious code injection: During this phase, the target's computer is infected with ransomware, and malicious code is downloaded.
- Malicious Payload Staging: Ransomware sets up persistence by inserting the system.
- Scan Checks for encryption on the target computer and any network-accessible resources.
- Encryption: The process of encrypting all of the selected documents begins.
- Payday: They cannot access the victim's data, and a notification seeking payment is visible on the screen of the targeted device.

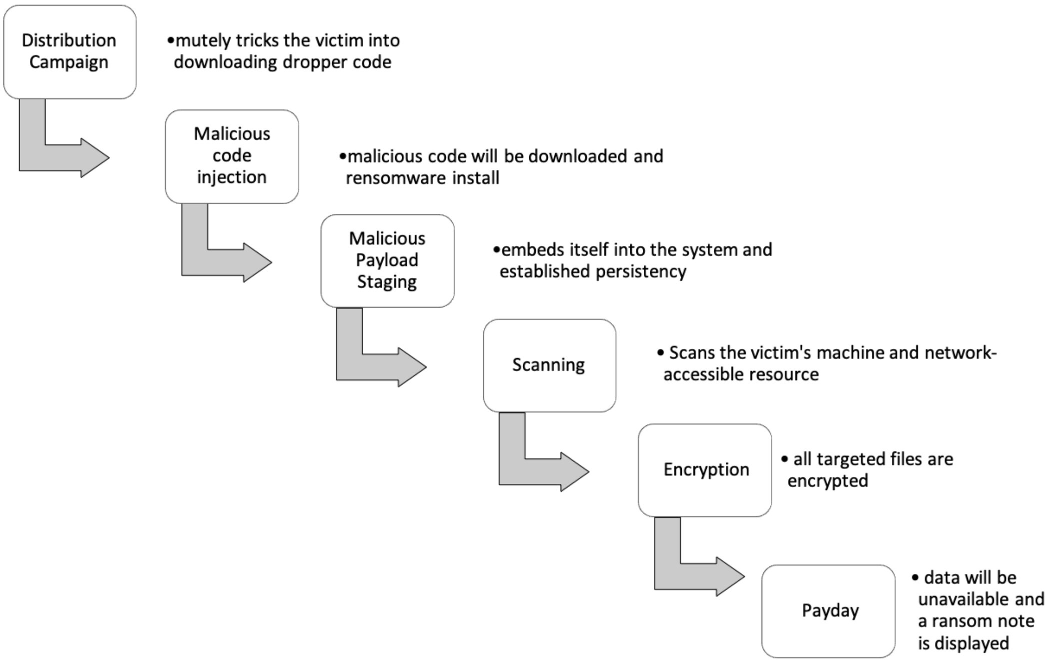

Distribution Campaign
•mutely tricks the victim into downloading dropper code
Malicious code injection
•malicious code will be downloaded and rensomware install
Malicious Payload Staging
•embeds itself into the system and established persistency
Scanning
• Scans the victim's machine and network-accessible resource
Encryption
• all targeted files are encrypted
Payday
• data will be unavailable and a ransom note is displayed

Figure 5. Six levels of a ransomware attack.

4. Ransomware Detection Techniques.

The two main types of ransomware detection techniques are automated and manual. Employing technologies to identify and report ransomware attacks is a prerequisite for automated methods. These tools are typically software programs that have the potential to be able to stop attacks. Techniques for manual detection focus on routinely scanning data and devices for indicators of attacks. Checking to see if a malware attack has not modified data or stopped authorized users from accessing their devices or files includes looking at any changes to file extensions, the accessibility of devices and files by authorized users, and any changes to file extensions. Figure 6 shows the Ransomware detection taxonomy.

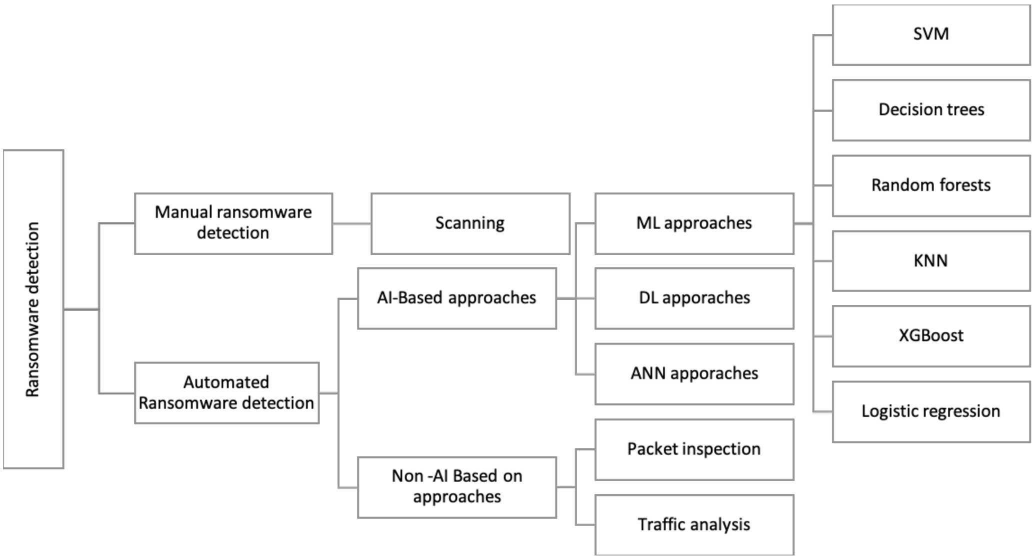


Figure 6. Ransomware detection taxonomy.

Artificial intelligence-based machine learning approaches such as behavioral techniques and static and dynamic analysis, deep learning, and artificial neural networks are used to detect ransomware attacks automatically [14]. Deep learning algorithms are addressing the shortcomings of supervised ransomware detection technologies. Increasing the accuracy and dependability of results from ransomware detection activity is the goal. Deep learning techniques are used for unstructured data to generate features automatically [6,15]. Artificial neural network approaches are well suited for identifying various ransomware data (text or picture) kinds and variants because of their general uses. Neural networks are the best choice for adjusting to new ransomware data and spotting zero-day attacks because they can learn continuously [5,6,16]. Ransomware can be found using non-AI approaches like packet inspection and traffic analysis. One strategy that seeks to observe online honeypots or fake computers and net network intention is to establish a honeypot folder and keep an eye out for any alterations that would indicate the presence of ransomware. Early ransomware detection is essential to lessen its effects and stop additional harm [2]. This section will discuss different ransomware detection techniques proposed in the literature and their strengths, weaknesses, and limitations.

4.1. Signature-Based Detection

Signature-based detection is a traditional approach that relies on identifying known ransomware signatures or patterns in the code or behavior of the malware. This approach is based on creating a database of known ransomware signatures or marks and scanning the system or network for matching signatures or patterns. If a match is found, the ransomware is flagged as malicious and appropriate actions are taken [17,18]. One benefit of signature-based detection is the simplicity and effectiveness in detecting known ransomware variants. However, this approach is limited by its inability to detect new or unknown ransomware variants that do not match existing signatures or patterns. Moreover, attackers can easily evade signature-based detection by modifying the code or behavior of the ransomware to avoid detection [14].

4.2. Heuristic-Based Detection

Heuristic-based detection is a more advanced approach that identifies ransomware behavior patterns or anomalies indicative of malicious activity. This approach is based on creating rules or heuristics that describe typical ransomware behavior and then monitoring the system or network for any deviations or anomalies from these rules. If such variations or abnormalities are detected, the ransomware is flagged as suspicious or malicious, and appropriate actions are taken [17,18]. One of the advantages of heuristic-based detection is its ability to detect new or unknown ransomware variants that do not match any existing signatures or patterns. Moreover, this approach is less prone to false positives than signature-based detection, as it relies on detecting actual behavior patterns rather than static code signatures. However, heuristic-based detection is limited by its reliance on predefined rules or heuristics, which may only capture some possible ransomware behavior patterns or anomalies. Moreover, attackers can easily evade heuristic-based detection by modifying the behavior of the ransomware to avoid detection [14].

4.3. Machine Learning-Based Detection

Machine learning-based detection is a more advanced approach that relies on training a machine learning model to detect ransomware based on its behavior patterns or features. This approach is based on collecting a large dataset of benign and malicious samples, extracting relevant features from them, and then training a machine learning model to classify new samples as peaceful or hostile based on their characteristics [17,18]. Machine learning-based detection has several benefits, including its ability to detect new or unknown ransomware variants that do not match existing signatures or patterns and to adapt to changing ransomware behavior patterns over time. Moreover, this approach is less prone to false positives than signature-based and heuristic-based detection, as it relies on detecting actual behavior patterns rather than static code signatures or predefined rules. However, machine learning-based detection is limited by its reliance on a large and representative dataset of training

doi:10.20944/preprints202305.1635.v1

samples and by its susceptibility to adversarial attacks that can manipulate the features or behavior of the ransomware to evade detection [14].

4.4. Network-Based Detection

A network-based detection is an approach that relies on monitoring the network traffic for suspicious or malicious activity that may be indicative of a ransomware attack. This approach is based on analyzing the network traffic for anomalies or patterns characteristic of ransomware, such as large volumes of outbound traffic, unusual network connections, or network traffic encryption [17,18]. One of the advantages of network-based detection is its ability to detect ransomware activity even if the malware has not yet infected the system or if the ransomware is using non-standard encryption methods. Moreover, this approach is less prone to false positives than other detection approaches, as it relies on detecting actual network traffic patterns rather than static code signatures or predefined rules. However, network-based detection is limited by its reliance on network traffic analysis tools that may not be available or may not capture all ransomware activity. Moreover, attackers can easily evade network-based detection by encrypting their network traffic or using stealthy communication channels [14].

4.5. Hybrid Detection

Hybrid detection is an approach that combines different ransomware detection techniques to improve the overall detection accuracy and speed. This approach combines the strengths of other detection techniques, such as signature-based, heuristic-based, machine learning-based, and network-based detection, to create a more robust and effective detection system [17,18]. One of the advantages of hybrid detection is its ability to overcome the limitations of individual detection approaches and to improve the overall detection accuracy and speed. Moreover, this approach is less prone to false positives and negatives than unique detection approaches, as it combines different sources of information and analysis. However, hybrid detection is limited by its complexity and resource requirements, as it requires integrating and coordinating other detection systems and tools [14].

Evaluating the performance of machine learning models for ransomware detection is crucial to determine their effectiveness in detecting and preventing its spread. In this section, we will discuss different evaluation metrics used for measuring the performance of machine learning models for ransomware detection, including accuracy, precision, recall, F1-score, and ROC curve.

- Accuracy: Accuracy is the most straightforward evaluation metric, representing 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 greatly outweigh the positive models [19,20].
- Precision: Out of all samples predicted to be positive (recognized as ransomware by the algorithm), precision is the percentage of true positives (samples of malware successfully identified). The ratio of true positives to the total of true and false positives is known as precision. A model with a high precision score will have a low false positive rate, making it less likely to label innocent files as ransomware mistakenly [20].
- Recall: Recall counts the number of positive samples in the collection that are true positives. The ratio of true positives to true and false negatives is computed. A high recall score suggests that the model has a low incidence of false negatives, which makes it less likely to fail to detect actual ransomware samples [20,21].
- ROC curve: The performance of a binary classifier as the discrimination threshold is changed is graphically represented by the receiver operating characteristic (ROC) curve. At various threshold values, it plots the actual positive rate (TPR) versus the false positive rate (FPR). The model's overall performance is assessed using the area under the ROC curve (AUC), with higher AUC values indicating better performance [22].

doi:10.20944/preprints202305.1635.v1

5. Machine Learning for Ransomware Detection

A particular kind of artificial intelligence known as machine learning enables computer systems to improve their performance on a given job without being explicitly taught. The malicious ransomware malware encrypts a victim's files and demands payment for the decryption key. Due to their rising prevalence and severity, machine-learning techniques are increasingly needed to identify and stop ransomware attacks. Table 2 lists the machine learning methods that are employed. Decision trees, random forests, support vector machines, and neural networks are just a few machine-learning machines that can detect ransomware. Each method has advantages and disadvantages, and the best approach depends on the situation and the data [1,6].

Table 2. Machine learning algorithms.

References	Algorithms	Characteristics
[23,24]	Decision trees	Decision trees can be trained on features such as file modifications, network traffic, and system calls to distinguish between ransomware and benign software behavior. The decision tree that results can then be used to determine whether new data contains ransomware.
[23,24]	Random forests	In order to guarantee that each tree in the forest has the same distribution and is dependent on the values of a randomly selected random vector, this strategy uses an ensembled method that combines tree predictors.Performance may be enhanced in comparison to standalone decision trees.Using a network of decision trees, the random forest approach is used to select and forecast the input data type.
[25,26]	Support vector machines	Support vector machines can be trained on features such as system calls, network traffic, and file behavior to distinguish between ransomware and benign software behavior. After that, it is possible to determine whether new data constitutes ransomware using the resultant Support vector machines. Support vector machines are handy when the data is high-dimensional and non-linearly separable, often in ransomware detection.
[27,28]	Neural networks	Like a biological brain, neural networks can find patterns in vast amounts of data. After getting the raw input, multi-layer neural network algorithms performed internal operations to extract and choose features. They have a mechanism for feature extraction and selection as a result. An input layer, an output layer holding the categorized variables, and a hidden layer comprise a primary neural network. The layers create an interconnected network of neurons.

Decision trees are a simple and intuitive machine learning algorithm that can be used for classification tasks, including ransomware detection. Decision trees work by recursively partitioning the data into subsets based on the values of the features and creating a tree-like structure representing the decision-making process. Both categorical and continuous components can be handled by decision trees, which are simple to interpret but susceptible to overfitting and sensitive to minute changes in the data [23]. Decision trees can be used to detect ransomware attacks by analyzing patterns in computer system events and identifying indicators of a potential attack. A decision tree is a type of algorithm that creates a tree-like model of decisions and their possible consequences. In the context of ransomware detection, a decision tree can be trained on a dataset of known ransomware attacks and non-malicious events to identify common patterns and attributes that are associated with ransomware attacks. For example, a decision tree might analyze events such as file access, network traffic, and system processes to determine whether they are indicative of a ransomware attack. By considering multiple attributes and their relationships, a decision tree can determine the likelihood of a particular event sequence being a ransomware attack. Once a decision tree has been trained on a dataset, it can be used to detect ransomware attacks in real-time by analyzing incoming system events and comparing them to the patterns identified in the training data. If a sequence of events matches the pattern of a ransomware attack, the system can trigger an alert or take other defensive actions to prevent the attack from spreading. Overall, decision trees can be a useful tool for detecting ransomware attacks by analyzing patterns in system events and identifying indicators of a potential attack. However, they are just one of many tools and techniques that can be used to defend against ransomware and should be used in conjunction with other cybersecurity best practices. [14,21,23] Decision trees can be extended to create random forests, enhancing performance and less overfitting. Using random selections of the features and data, random forests build numerous decision trees and then aggregate their predictions. Random forests are less prone to overfitting than decision trees and can handle high-dimensional data, but they can be computationally expensive and challenging to interpret [24]. Support vector machines are reliable machine learning techniques that can be utilized for ransomware detection and classification and regression applications. Support vector machines operate by identifying the hyperplane that divides the data into distinct classes according to the values of the features as thoroughly as possible. Support vector machines can effectively handle high-dimensional data. They can accept both linear and nonlinear borders, but the choice of the kernel function and its parameters may impact them [25]. Neural networks are sophisticated and adaptable machine learning algorithms that may be applied to various tasks, including ransomware detection. Neural networks comprise many layers of interconnected nodes, or neurons that may learn to identify patterns in input and anticipate outcomes. Complex and nonlinear interactions between the constituents and the goal variable can be handled using neural networks. Nevertheless, they can be computationally expensive and require many training data [27]. The choice of a machine learning algorithm for ransomware detection depends on the specific problem and data available. Decision trees, random forests, Support vector machines, and neural networks are all effective options, and researchers have successfully used each of these algorithms for ransomware detection in different contexts [5,14].

5.1. 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 different features used for ransomware detection using machine learning and the techniques used for feature selection, such as principal component analysis and correlation analysis [29,30].

5.1.1. Features Used for Ransomware Detection

- File access patterns: Ransomware typically accesses and encrypts files in a particular pattern, such as alphabetical order, extension type, or creation date. This behavior can be detected using file access patterns as features [31].
- System calls: Ransomware commonly uses system calls to perform its malicious activities, such as reading and writing files, creating processes, and network communication. System call traces can be extracted and used as features for detection [23].
- Network traffic: A command-and-control (C&C) server is frequently used by ransomware to deliver and receive orders. The analysis of network traffic can provide valuable features for detecting ransomware [27].
- Behavioral analysis: Behavioral analysis involves monitoring the behavior of running processes and identifying anomalies that indicate malicious activity. Features such as process creation, termination, and file access can be used [1].

- Static analysis: Static analysis involves looking at the executable file's source code to spot malicious activity. Features such as code size, entropy, and string patterns can be used for this purpose [17].

5.1.2. Feature Selection Techniques

- Principal component analysis: This technique is used to reduce the dimensionality of a dataset by identifying the most critical features that explain the majority of the variance in the data. Principles component analysis can help identify redundant or irrelevant features and select the most informative ones for ransomware detection [28].
- Correlation analysis: Correlation analysis is a technique used to identify the correlation between features in a dataset. Highly correlated features may be redundant and can be removed to simplify the model and improve performance [11].

5.2. Machine Learning Detection Studies

Preventing ransomware is challenging for several reasons. The way ransomware functions is the same as benign software, which acts covertly. Ransomware detection in zero-day assaults is, therefore, crucial at this time. The primary objectives are to avoid ransomware-caused system damage, identify zero-day (previously unidentified) malware, and minimize detection. Ransomware can be found using a variety of tools and methodologies. Methods based on static analysis decompose source code without running it. They have a lot of false positives and cannot find ransomware that's been disguised. Attackers frequently create new variations and modify their codes using various packaging techniques. To solve these issues, researchers use dynamic behavior analysis methods that monitor interactions between the executed code and a virtual environment. However, these detection methods are cumbersome and memory-intensive. Machine learning is ideal for analyzing any process or application's behavior. The following are some machine learning-based detection systems that follow highly traditional methodologies: Table 3 summarizes previous studies on Machine learning techniques (behavioral techniques) for ransomware detection from 2017 to 2022

Table 3. Studies on Machine learning techniques (behavioral techniques) for ransomware detection from 2017 to 2022.

References	Year	Author	Resolved the Issue	Utilized Technique	Result	Limitation
[32]	2017	Zahra & Sha	Detecting a ransomware attack using Cryptowall	Blocklisting of command and control (C and C) servers	The web proxy server, which acts as the TCP/IP traffic gateway, extracts the TCP/IP header.	The model's efficacy and precision in identifying ransomware and its attack techniques against various operating system environments were not demonstrated through implementation.

Table 3. *Cont.*

References	Year	Author	Resolved the Issue	Utilized Technique	Result	Limitation
[33]	2018	Shaukat & Ribeiro	detection of ransomware	(RansomWall) A layered and hybrid mechanism	effective at identifying zero-day attacks	N/A
[34]	2019	Makinde et al.	To determine whether an actual network system is vulnerable to a ransomware assault	Learning Machines	Correlation greater than 0.8	It imitated the behavior of a small group of users.
[35]	2019	Ahmad et al.	Differentiating Locky ransomware users	Utilizing parallel classifiers, a behavioral approach to ransomware detection	Highly reliable detection with a low proportion of false positives	N/A
[36]	2022	Singh et al.	Discovery of new ransomware families and classification of newly discovered ransomware assaults	Checks process memory access privileges to enable rapid and accurate malware detection	Between 81.38% and 96.28% accuracy.	N/A

An application's normal behavior is assessed from a user and resource perspective. A baseline for normal behavior is established based on what is thought to be the typical or routine operation of a computer system or network. Indicators of usual activity include logins, file access, user and file behaviors, resource utilization, and other significant indicators [1]. The length of the learning process is determined by the amount of data needed to build a baseline to represent typical system behavior. The tool investigates behavioral outliers from the baseline's depiction of the typical behavioral pattern. A ransomware detection and prevention model was created for unstructured datasets derived from Ecuadorian Control and Regulatory Institution (EcuCERT) logs [37]. The methodology uses musing to spot peculiar behavioral patterns connected to Windows malware. Feature selection was applied to the Log data to extract the most beneficial and discriminating information that indicates a ransomware attack. The extracted data represents that autonomous learning algorithms in Ransomware are swiftly and precisely identified using the input feature set and algorithms that mimic abnormal behavioral patterns. Code obfuscation tools and new polymorphic variants have been developed to signature additions in identifying ransomware attacks, which are constantly evolving [33]. Since generic malware attack vectors cannot effectively capture the particular behavioral traits of cryptographic ransomware, they are insufficient or inaccurate for ransomware detection. The suggested approach, RansomWall, is a hybrid system that uses static and dynamic analytics to present a novel set of properties that mimic ransomware activity. The technique allows for early ransomware detection while utilizing a strong trap layer to detect zero-day attacks. RansomWall with Gradient Tree Boosting Algorithm demonstrated a detection rate of 98.25% and an incredibly low (almost nil) false positive rate when tested against 574 samples of 12 cryptographic ransomware running on the Microsoft Windows operating system. It also has a detection rate of less than 10% for 30 zero-day attack samples compared to 60 VirusTotal security engines. One version of behavioral detection methodologies uses a machine learning baseline model for simulating and forecasting the specific network user behavior pattern at the micro level to identify potential scenarios that could indicate a vulnerability or a true ransomware assault [34]. The goal was to find a simple network system's vulnerability to a ransomware attack. Comparing the outcomes from the simulated network and the log data from the server in the existing network system reveals a realistic model with a correlation above 0.8. This method's drawback was that it only adequately captured the activity of a small percentage of users. Future studies should focus on mimicking user behavior over a large user base using big data analytics tools. A more recent method of behavioral ransomware detection used two parallel classifiers [35]. to distinguish between the several Locky ransomware variants. The technique focused on early detection based on behavioral analysis of ransomware network traffic to prevent ransomware from connecting to command-and-control servers and carrying out damaging payloads. The study employed a dedicated network to collect information and extract important details from network traffic. Using data at the packet and datagram levels, two different (parallel) classifiers analyze the extracted properties of the Locky ransomware family. The results of the studies show that the technology has a high level of success in detecting ransomware activities on the network. Furthermore, it permits an extreme lexicon with a low percentage of false positives. Using command and control (C&C), server blocklist ransomware attacks as the means of communication, communication, and behavioral analysis of the ransomware in an IoT environment [32]. provided a strategy for identifying Cryptowall ransomware attacks that is domain-specific. The operation obtains the TCP/IP header from the web proxy server, which serves as the TCP/IP traffic gateway. Furthermore, it retrieves source and destination IPs and compares them to IPs of forbidden Command-and-Control servers. Ransomware is identified if the source or destination IPs match an attack targeting Internet of Things devices. However, the model was not used to demonstrate how well it could spot ransomware and its attack vectors against different operating system environments. Using a very recent technique of behavioral-based detection that uses access privileges in process memory, ransomware may now be quickly and accurately detected [21,36]. It is possible to categorize new ransomware attacks and find malware families that have not yet been recognized by looking at a file or application's access privileges and the area of memory it intends to access. Examining the behavior and ascertaining the purposes of lawful files and applications before executing them is beneficial. The experimental results employing these several approaches show good detection accuracy, ranging from 81.38% to 96.28%. Table 4 summarizes previous studies on Machine learning techniques (static and dynamic analysis) for ransomware detection from 2017 to 2022

Table 4. Studies on Machine learning techniques (static and dynamic analysis) for ransomware detection from 2017 to 2022.

References	Year	Author	Problem Addressed	Method Used	Result
[25]	2017	Rahman and Hasan	Enhanced ransomware detection method	Using support vector machines as an analysis tool	Better ransomware detection is achieved with an integrated approach than static or dynamic analysis used separately.
[21]	2018	Dehghantanha et al.	Windows ransomware detection that is quick and accurate	Netconverse (classifier using j48 decision tree)	97.1% actual positive detection rate
[38]	2019	Jasmin	Separating ransomware traffic and regular traffic	Algorithms used in logistic regression include random forest and support vector machine.	The best detection rate is 99.9% for the random forest, with 0% false positives.
[39]	2019	Ameer	Detection of ransomware.	Analyses that are static and dynamic.	100% detection and classification precision
[24]	2020	Khammas	Detection of ransomware.	Random forest method.	97.74% of samples are detected.
[29]	2020	Hwang et al.	An improved method of detecting ransomware.	Random forest and Markov models	97.3% overall accuracy, 4.8% for false positives, and 1.5% for false negatives.
[40]	2022	Talabani and Abdulhadi	Tools for detecting ransomware that involves data mining and machine learning approaches have poor accuracy.	Decision Table and PARTially Decided Decision Tree.	Recall (96%), accuracy (96.01%), F-measure (95.6%), and precision (95.9%)

Several improved machine-learning approaches have been applied for accurate and efficient ransomware detection. These methods are meant to address the drawbacks of the current ML-based ransomware detection tools. One of these advancements is the challenge detection systems (such as sandbox analysis and pipelines) face in isolating a sample and handling the wait time for isolated ransomware samples to be evaluated [41]. The approach predicts ransomware using a dataset containing 30,000 attributes as independent variables. Five qualities that were obtained through feature selection were used in the support vector machine technique. The approach provides a respectable 88.2% accuracy rate in ransomware detection. To reduce the number of false positives, this hybrid technique combines the "guilt by association" hypothesis with content-, metadata-, and behavior-based analysis. Giving the user control over recovery is necessary, and file versioning in cloud storage is used to halt the process. The only duty of the end user is to keep track of the recovery. Users are given classification information so they may make educated decisions and prevent false positives. The method results in more accurate detection and reliable recovery. An innovative method for detecting network-level ransomware uses machine learning, certificate information, and network connection information [42]. The technique can be used with system-level monitoring to detect ransomware outbreaks early. The method uses connection-, encryption-, and certificate-based network traffic characteristics to extract and model ransomware features. It is a feature model that uses support vector machines, logistic regression, and random forest to distinguish ransomware traffic. According to experimental findings on various datasets, the random forest has the best detection rate of 99.9% and the lowest rate of false positives. Another more effective detection method is a decision tree model based on big data technology that uses Argus for packet preprocessing, combining, and malware file identification [43]. The flow replaced the packet data, resulting in a 1000-fold (1000:1) reduction in data size. Feature selection and concatenation were used to extract and aggregate the attributes of the actual network traffic. In order to improve classification accuracy, the technique made use of six feature selection techniques. Machine learning has recently been creatively applied to monitor Android device power usage as a ransomware detection technique [21]. The suggested method measures how much energy particular Android processes use to distinguish ransomware from valuable programs. Data on the ransomware's unique local energy fingerprint is gathered and analyzed to accomplish this. According to experimental findings, the approach offers high detection and precision rates of 95,6%, and 89%, respectively. Additionally, it outperforms the K-Nearest Neighbor, Neural Network, Support Vector Machine, and Random Forest regarding the accuracy, recall rate, precision rate, and F-measure. Another superior option is the cutting-edge, portable RanDroid approach for automatically detecting polymorphic ransomware [44]. The method compares the structural similarity of pieces obtained from an application with a collection of threat information from well-known ransomware variants to detect new ransomware variants on Android devices. Image Similarity Measurement (ISM) and String Similarity Measurement (SSM) are the two similarity measures used. Using language analysis, the app's behavioral attributes and picture textural strings are mined for additional information. The strategy reduced ransomware threats without changing the Android OS or its underlying security module while addressing the constraints of static analysis. The methodology can detect ransomware using evasive tactics like complex codes or dynamic payloads, according to an analysis of the method based on 950 malware samples. According to a related study, a strategy combining static and dynamic analysis can help identify and separate Android ransomware from other malware [39]. We looked at network-based features, text, and permissions using static analysis. Furthermore, dynamic analysis was performed on the system call, CPU, and memory logs. The strategy's effectiveness in reducing evasive ransomware assaults is demonstrated by experiments using traits from malicious and benign samples. Additionally, it is 100 percent accurate at classifying and identifying unknown ransomware.

6. Data Collection and Preprocessing

Machine learning techniques have been increasingly used for ransomware detection, but collecting and preprocessing data presents several challenges [45,46].

One of the main challenges in collecting data for ransomware detection is the need for publicly available datasets that include real-world ransomware samples. This is due to the sensitive nature of the data and the fact that many victims are reluctant to report ransomware attacks. As a result, researchers often rely on synthetic datasets or datasets generated from sandbox environments, which may not accurately reflect the complexity and variability of real-world ransomware attacks [3].

Another challenge is the diversity of ransomware families and variants, which require a large and diverse dataset to ensure adequate coverage. Ransomware behavior can also vary depending on the victim's system and network environment, making generalizing detection models across different contexts challenging [2,45].

Preprocessing data for ransomware detection also presents several challenges. Ransomware often employs obfuscation techniques to evade detection, such as encrypting the payload or using anti-analysis mechanisms. This can make extracting relevant data features difficult and identifying patterns distinguishing ransomware from benign software. In addition, ransomware may use legitimate system functions that are difficult to distinguish from malicious behavior, requiring advanced feature engineering and modeling techniques [45].

Despite these challenges, several datasets have been used to train and evaluate ransomware detection models.

Collecting and preprocessing data for ransomware detection using s machine learning presents several challenges, including the lack of real-world datasets, the diversity of ransomware families and variants, and the obfuscation techniques used by ransomware. However, several datasets have been developed to address these challenges, providing valuable resources for training and evaluating ransomware detection models [45].

7. Challenge and Future Directions

Developing effective machine learning-based ransomware detection systems is challenging due to several factors. This section will discuss the challenges of developing such systems and highlight the future directions in this field. Challenges in developing effective machine learning-based ransomware detection systems:

- Data quality and quantity: A vast amount of high-quality data is needed to train machine learning models effectively. However, obtaining high-quality data for ransomware detection is challenging due to the limited availability of labeled ransomware samples [45,46].
- Rapidly evolving ransomware: Ransomware is a constantly changing threat, with new variants and attack techniques being developed regularly. This makes it challenging to build machine learning models that can detect all ransomware accurately and quickly [47].
- Adversarial attacks involve modifying the input data to bypass the machine learning model's detection capabilities. Malicious attacks can be used to evade ransomware detection systems, making the systems less effective [47].
- Real-time detection requirements: Ransomware can spread rapidly and cause significant damage within a short time. Therefore, ransomware detection systems must be able to detect ransomware in real-time to prevent further spread and damage [48].

7.1. Future Work in This Field

- Developing more robust and accurate models: Researchers must build more substantial and precise machine-learning models that detect a wide range of ransomware variants and attack techniques. This can be achieved through advanced techniques such as deep learning and ensemble learning [4,45,49].
- Incorporating real-time detection capabilities: Ransomware detection systems must incorporate real-time detection capabilities to quickly identify and prevent ransomware attacks. This can be achieved through the use of real-time monitoring and analysis techniques [46].

- Addressing the issue of adversarial attacks: Researchers need to develop machine learning models that are robust to malicious attacks. This can be achieved through techniques such as negative training and defensive distillation [45,47].
- Collaboration and sharing of data: Collaboration and sharing of data among researchers and organizations can help develop more effective ransomware detection systems. This can help build more comprehensive datasets for training and testing machine learning models [47].
- Developing effective machine learning-based ransomware detection systems is challenging for several reasons. However, with advanced techniques and collaboration among researchers and organizations, it is possible to develop more robust and accurate ransomware detection systems [45].

8. Conclusion

Ransomware attacks have caused significant harm to computer systems and the data they manage, resulting in unauthorized access, disclosure, and destruction of important and sensitive information. These attacks have led to substantial financial losses and reputational damage for both individuals and businesses. In response, various methods have been suggested to detect ransomware accurately, quickly, and dependably. This novel provides readers with a historical background and timeline of ransomware attacks, as well as a discussion of the issue's context. The recent literature review offers an up-to-date understanding of automated ransomware detection approaches. This knowledge will help readers stay current on the latest advances in automated ransomware detection, prevention, mitigation, and recovery. Additionally, this study discusses future research directions, highlighting open issues and potential research problems for those interested in studying ransomware detection, prevention, mitigation, and recovery.

References

1. Celdrán, A.H.; Sánchez, P.M.S.; Castillo, M.A.; Bovet, G.; Pérez, G.M.; Stiller, B. Intelligent and behavioral-based detection of malware in IoT spectrum sensors. *International Journal of Information Security* **2022**, pp. 1–21.
2. Chesti, I.A.; Humayun, M.; Sama, N.U.; Jhanjhi, N. Evolution, mitigation, and prevention of ransomware. In Proceedings of the 2020 2nd International Conference on Computer and Information Sciences (ICCIS). IEEE, 2020, pp. 1–6.
3. Philip, K.; Sakir, S.; Domhnall, C. Evolution of ransomware. *IET Netw* **2018**, *7*, 321–327.
4. Jegede, A.; Fadele, A.; Onoja, M.; Aimufua, G.; Mazadu, I.J. Trends and Future Directions in Automated Ransomware Detection. *Journal of Computing and Social Informatics* **2022**, *1*, 17–41.
5. Brewer, R. Ransomware attacks: detection, prevention and cure. *Network Security* **2016**, *2016*, 5–9.
6. Bello, I.; Chiroma, H.; Abdullahi, U.A.; Gital, A.Y.; Jauro, F.; Khan, A.; Okesola, J.O.; Abdulhamid, S.M. 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* **2021**, *12*, 8699–8717.
7. Scaife, N.; Carter, H.; Traynor, P.; Butler, K.R. Cryptolock (and drop it): stopping ransomware attacks on user data. In Proceedings of the 2016 IEEE 36th international conference on distributed computing systems (ICDCS). IEEE, 2016, pp. 303–312.
8. Sgandurra, D.; Muñoz-González, L.; Mohsen, R.; Lupu, E.C. Automated dynamic analysis of ransomware: Benefits, limitations and use for detection. *arXiv preprint arXiv:1609.03020* **2016**.
9. Prakash, K.P.; Nafis, T.; Biswas, S.S. Preventive Measures and Incident Response for Locky Ransomware. *International Journal of Advanced Research in Computer Science* **2017**, *8*.
10. Paquet-Clouston, M.; Haslhofer, B.; Dupont, B. Ransomware payments in the bitcoin ecosystem. *Journal of Cybersecurity* **2019**, *5*, tyz003.
11. Kok, S.; Abdullah, A.; Jhanjhi, N.; Supramaniam, M. Ransomware, threat and detection techniques: A review. *Int. J. Comput. Sci. Netw. Secur* **2019**, *19*, 136.

12. Thakran, E.; Kumari, A. Impact of "Ransomware" on critical infrastructure due to pandemic. *Available at SSRN 4361110* **2023**.
13. Ahmed, Y.A.; Huda, S.; Al-rimy, B.A.S.; Alharbi, N.; Saeed, F.; Ghaleb, F.A.; Ali, I.M. A weighted minimum redundancy maximum relevance technique for ransomware early detection in industrial IoT. *Sustainability* **2022**, *14*, 1231.
14. Akhtar, M.S.; Feng, T. Malware Analysis and Detection Using Machine Learning Algorithms. *Symmetry* **2022**, *14*, 2304.
15. Sharmeen, S.; Ahmed, Y.A.; Huda, S.; Koçer, B.Ş.; Hassan, M.M. Avoiding future digital extortion through robust protection against ransomware threats using deep learning based adaptive approaches. *IEEE Access* **2020**, *8*, 24522–24534.
16. Swami, S.; Swami, M.; Nidhi, N. Ransomware Detection System and Analysis Using Latest Tool. *International Journal of Advanced Research in Science, Communication and Technology* **2021**, *7*, 2581–9429.
17. Yamany, B.; Elsayed, M.S.; Jurcut, A.D.; Abdelbaki, N.; Azer, M.A. A New Scheme for Ransomware Classification and Clustering Using Static Features. *Electronics* **2022**, *11*, 3307.
18. Yamany, B.; Azer, M.A.; Abdelbaki, N. Ransomware Clustering and Classification using Similarity Matrix. In Proceedings of the 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC). IEEE, 2022, pp. 41–46.
19. Kok, S.; Azween, A.; Jhanjhi, N. Evaluation metric for crypto-ransomware detection using machine learning. *Journal of Information Security and Applications* **2020**, *55*, 102646.
20. Masum, M.; Faruk, M.J.H.; Shahriar, H.; Qian, K.; Lo, D.; Adnan, M.I. Ransomware classification and detection with machine learning algorithms. In Proceedings of the 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, 2022, pp. 0316–0322.
21. Azmoodeh, A.; Dehghantanha, A.; Conti, M.; Choo, K.K.R. Detecting crypto-ransomware in IoT networks based on energy consumption footprint. *Journal of Ambient Intelligence and Humanized Computing* **2018**, *9*, 1141–1152.
22. Edis, D.; Hayman, T.; Vatsa, A. Understanding Complex Malware. In Proceedings of the 2021 IEEE Integrated STEM Education Conference (ISEC). IEEE, 2021, pp. 1–2.
23. Ullah, F.; Javaid, Q.; Salam, A.; Ahmad, M.; Sarwar, N.; Shah, D.; Abrar, M. Modified decision tree technique for ransomware detection at runtime through API Calls. *Scientific Programming* **2020**, *2020*.
24. Khammas, B.M. Ransomware detection using random forest technique. *ICT Express* **2020**, *6*, 325–331.
25. Ghouti, L.; Imam, M. Malware classification using compact image features and multiclass support vector machines. *IET Information Security* **2020**, *14*, 419–429.
26. Arunkumar, M.; Kumar, K.A. GOSVM: Gannet optimization based support vector machine for malicious attack detection in cloud environment. *International Journal of Information Technology* **2023**, pp. 1–8.
27. Madani, H.; Ouerdi, N.; Boumesaoud, A.; Azizi, A. Classification of ransomware using different types of neural networks. *Scientific Reports* **2022**, *12*, 1–11.
28. Arivudainambi, D.; KA, V.K.; Visu, P.; et al. Malware traffic classification using principal component analysis and artificial neural network for extreme surveillance. *Computer Communications* **2019**, *147*, 50–57.
29. Hwang, J.; Kim, J.; Lee, S.; Kim, K. Two-stage ransomware detection using dynamic analysis and machine learning techniques. *Wireless Personal Communications* **2020**, *112*, 2597–2609.
30. Dargahi, T.; Dehghantanha, A.; Bahrami, P.N.; Conti, M.; Bianchi, G.; Benedetto, L. A Cyber-Kill-Chain based taxonomy of crypto-ransomware features. *Journal of Computer Virology and Hacking Techniques* **2019**, *15*, 277–305.
31. Sheen, S.; Asmitha, K.; Venkatesan, S. R-Sentry: Deception based ransomware detection using file access patterns. *Computers and Electrical Engineering* **2022**, *103*, 108346.
32. Zahra, A.; Shah, M.A. 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). IEEE, 2017, pp. 1–6.
33. Shaukat, S.K.; Ribeiro, V.J. 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). IEEE, 2018, pp. 356–363.

34. Makinde, O.; Sangodoyin, A.; Mohammed, B.; Neagu, D.; Adamu, U. Distributed network behaviour prediction using machine learning and agent-based micro simulation. In Proceedings of the 2019 7th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE, 2019, pp. 182–188.
35. Almashhadani, A.O.; Kaiiali, M.; Sezer, S.; O'Kane, P. A multi-classifier network-based crypto ransomware detection system: A case study of locky ransomware. *IEEE access* **2019**, *7*, 47053–47067.
36. Singh, A.; Ikuesan, R.A.; Venter, H. Ransomware detection using process memory. *arXiv preprint arXiv:2203.16871* **2022**.
37. Silva, J.A.H.; Hernández-Alvarez, M. Large scale ransomware detection by cognitive security. In Proceedings of the 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM). IEEE, 2017, pp. 1–4.
38. Modi, J. Detecting ransomware in encrypted network traffic using machine learning. PhD thesis, 2019.
39. Ameer, M. Android ransomware detection using machine learning techniques to mitigate adversarial evasion attacks. *Capital University of Science and Technology, Islamabad, Pakistan* **2019**.
40. Talabani, H.S.; Abdulhadi, H.M.T. Bitcoin ransomware detection employing rule-based algorithms. *Science Journal of University of Zakho* **2022**, *10*, 5–10.
41. Adamu, U.; Awan, I. Ransomware prediction using supervised learning algorithms. In Proceedings of the 2019 7th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE, 2019, pp. 57–63.
42. Modi, J. Detecting ransomware in encrypted network traffic using machine learning. PhD thesis, 2019.
43. Wan, Y.L.; Chang, J.C.; Chen, R.J.; Wang, S.J. 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). IEEE, 2018, pp. 85–88.
44. Alzahrani, A.; Alshehri, A.; Alshahrani, H.; Alharthi, R.; Fu, H.; Liu, A.; Zhu, Y. Randroid: Structural similarity approach for detecting ransomware applications in android platform. In Proceedings of the 2018 IEEE International Conference on Electro/Information Technology (EIT). IEEE, 2018, pp. 0892–0897.
45. Beaman, C.; Barkworth, A.; Akande, T.D.; Hakak, S.; Khan, M.K. Ransomware: Recent advances, analysis, challenges and future research directions. *Computers & security* **2021**, *111*, 102490.
46. McIntosh, T.; Kayes, A.; Chen, Y.P.P.; Ng, A.; Watters, P. Ransomware mitigation in the modern era: A comprehensive review, research challenges, and future directions. *ACM Computing Surveys (CSUR)* **2021**, *54*, 1–36.
47. Aboaoja, F.A.; Zainal, A.; Ghaleb, F.A.; Al-rimy, B.A.S.; Eisa, T.A.E.; Elnour, A.A.H. Malware detection issues, challenges, and future directions: A survey. *Applied Sciences* **2022**, *12*, 8482.
48. Gorment, N.Z.; Selamat, A.; Cheng, L.K.; Krejcar, O. Machine Learning Algorithm for Malware Detection: Taxonomy, Current Challenges and Future Directions. *IEEE Access* **2023**.
49. Kapoor, A.; Gupta, A.; Gupta, R.; Tanwar, S.; Sharma, G.; Davidson, I.E. Ransomware detection, avoidance, and mitigation scheme: a review and future directions. *Sustainability* **2021**, *14*, 8.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.