Developing a Behavioral Analysis Framework for the Detection and Mitigation of Ransomware Attacks

1st Kaniz Fatima Daya
Dept. of Computer Science
King Fahd University of Petroleum and Minerals.
Dhahran, Kingdom of Saudi Arabia
g202319170@kfupm.edu.sa

***Abstract*—Files can be held hostage by malicious software called ransomware, which demands a ransom for their release. This growing threat has affected over three million people recently. This study compares deep learning and machine learning algorithms for detecting and mitigating ransomware attacks using the proposed method. Random Forest, KNN, and CNN algorithms are integrated with data pretreatment and feature engineering techniques to enhance accuracy which is 98%. Among these, CNN is expected to achieve the highest accuracy in ransomware detection. Additionally, the security system continuously monitors for suspicious activity. This research can aid researchers, vendors, endpoint security providers, and anti-malware developers in utilizing deep learning algorithms to create more robust defense solutions against ransomware, capable of detecting and stopping attacks before they occur.**

I. Introduction

Malicious software known as ransomware is made to encrypt or lock down files on a victim's computer, making the data unreadable until a ransom is paid. Over time, this type of cyberattack has changed dramatically, taking use of sophisticated encryption methods and finding weaknesses in both personal and corporate information systems. The sophistication of ransomware attacks has increased, and they frequently use evasive and polymorphic tactics to get around conventional protection measures. Such attacks have the potential to have catastrophic financial and operational effects, including large monetary losses, data breaches, and operational interruptions. The dynamic ransomware threat landscape demands the use of cutting-edge technologies like machine learning, deep learning, and fuzzy hashing to strengthen cybersecurity defenses. These tactics are essential for effective detection and mitigation.

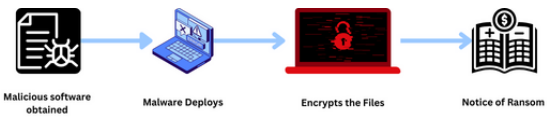


Fig. 1. General workflow of Ransomware Attack.

The fundamental procedure of a ransomware assault is broken down into four sequential parts in Figure 1. First, the malicious software is unintentionally downloaded by the victim; this frequently happens via phishing emails. The malware then installs itself on the victim's computer and begins to operate covertly. After it is installed, the ransomware uses sophisticated algorithms to encrypt the victim's data, rendering them unreadable. Ultimately, a ransom notice is sent to the victim, requesting money (usually in cryptocurrency) in return for the decryption key that unlocks the encrypted files.

Because it may encrypt or lockout data, ransomware poses a serious danger to organizational information systems. This makes these systems utterly unusable. For instance, businesses and individuals lost around $4 billion as a result of the WannaCry virus [4]. Concern is mounting over the increase in ransomware assaults on hospital infrastructures, particularly Integrated Clinical Environments (ICE). Critical healthcare procedures may be interfered with by these attacks, jeopardizing patient care and data security. This problem has been made worse by the COVID-19 outbreak; since November 2020, ransomware assaults on healthcare facilities have increased by 45% [2]. It targets governmental services like police stations and hospitals as well as private individuals, businesses, and organizations. Over three million customers were impacted by ransomware between 2019 and 2020, and in 2019 ransom payments topped $25 billion [1]. In the first quarter of 2020, ransom payments increased by 33% when compared to the same period in 2019. With 15.6% of all ransomware assaults in 2020 being caused by the Sodinokibi (REvil) virus, it has emerged as one of the most dangerous threats. Among the notable events is the attack on Harvest Food Distributors, which led to a $7.5 million ransom demand [3].

Since it encrypts files on victims' systems and demands ransom payments to unlock them, ransomware has grown to be a lucrative and widespread type of cyberattack. There is an urgent need for strong defense measures against threats like the WannaCry ransomware, as high-profile assaults like this one have caused enormous financial harm [5]. To tackle this, techniques such as machine learning (ML) and deep learning (DL) have been put forth; however, for them to be effectively trained and tested, large amounts of different ransomware data

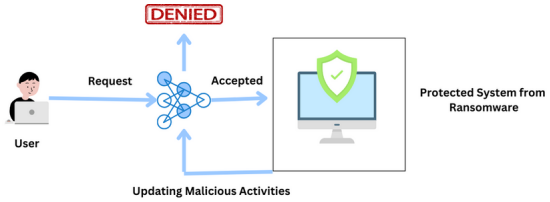


Fig. 2. General workflow of Ransomware Attack.

are needed [2].

The pre-attack can be detected by our Proposed system, which can also keep track of its defense within the system. To identify the attack, we contrast three machine learning and deep learning models. We make use of the convolution neural network, K-nearest neighbors (KNN), and random forest classifier. When comparing the three, CNN does better. and carefully extract the data from the secured system to determine whether any malicious activity has occurred.

II. Literature Review

In order to combat ransomware, the paper presents a novel strategy that makes use of the NTFS file system's Alternative Data Streams (ADS). With ADS, data streams can be concealed inside files, reducing ransomware's ability to access them. Although malicious apps frequently employ this technique, it can also be utilized to shield valid user data from encryption. The Protected ADS User File (PAUF), a novel file structure created by the proposed FREEDOM (Fast Recovery and Detection Mechanism) system, uses ADS to protect user data. This structure, which comprises of an ADS file and a linker file, makes sure that ransomware only targets the ADS file and leaves the actual data safe. A kernel-level driver monitor built into the system finds and stops illegal changes to PAUF files [5]. To detect new or unknown ransomware variants by comparing them to known samples, the study presents a method for ransomware detection utilizing fuzzy hashing, which maintains similarity and provides fuzzy hash values. Three fuzzy hashing techniques—SSDEEP, SDHASH, and mvHASH-B—were assessed. These techniques were tested on a corpus of WannaCry ransomware strains to see how well they could identify commonalities [4].

The study suggests a density-based machine learning technique and intelligent K-Nearest Neighbors (KNN) for identifying ransomware pre-attacks on endpoint systems. This method improves anti-malware and anti-ransomware programs by more accurately anticipating ransomware than existing machine learning methods [1]. The FedDICE framework combines FL and Software-Defined Networking (SDN). Through the use of dynamic network programmability offered by SDN, malicious devices can be isolated and removed from the network. FedDICE preserves data privacy and performs on par with centralized learning techniques [2].

The study investigates how to restrict network traffic and slow the spread of ransomware using SDN, more especially with the Ryu controller and Open Virtual Switch (OVS). Programmatic network management, made possible by SDN, enables real-time configuration adjustments to isolate and stop ransomware from spreading. By preventing TCP connections from infected devices, the SDN-based solution used in this study successfully lowered ransomware traffic records by up to 73.97% and decreased the propagation of the Sodinokibi virus by 17.13% [3]. SSDEEP generates hash values using a rolling hash based on the Adler32 function and variable-sized blocks, which makes it efficient for files with comparable content. SDHASH is useful for finding uncommon characteristics in comparable data since it finds statistically-improbable features and stores them using Bloom filters. In order to ensure that slight modifications to the input do not impact the similarity detection, mvHASH-B generates hash values using majority voting, run-length encoding, and Bloom filters [4].

The FREEDOM system was put to the test against a variety of ransomware samples, such as GandCrab and WannaCry. The outcomes showed that the solution successfully guarded user data kept in ADS, stopped linker file encryption, and stopped ransomware processes that tried to alter files that were protected. With very little performance overhead, the system achieved high accuracy and efficiency, making it a workable tool for ransomware defense [5]. The algorithm's better accuracy in predicting ransomware pre-attacks was assessed using a collection of ransomware samples. The findings demonstrated that, in comparison to conventional techniques, the suggested KNN and density-based algorithm produced greater accuracy [1]. FedDICE was tested utilizing clinical environment network traffic data, concentrating on four ransomware families: Petya, WannaCry, BadRabbit, and PowerGhost. The findings showed that FedDICE could successfully identify ransomware propagation in both IID (independent and identically distributed) and non-IID data sets, obtaining excellent accuracy and low false-negative rates [2].

Methodology

There are numerous methods and algorithms for detecting ransomware that are based on Windows. These detection techniques are doing admirably. But, there is a need to enhance malware detection methods and algorithms due to the ongoing rise in zero-day attacks and the attacks on endpoint systems by the latest wave of ransomware families [1]. In this study, we applied the Random Forest Classifier, K-nearest neighbors (KNN), Convolution Neural Network method, and Convolution Neural Network with exceptional accuracy results for ransomware detection.

Figure represents a comprehensive machine learning workflow, starting with Data Collection, where data is gathered from various sources. This data is then imported into the system as a CSV file Import Dataset.csv File. Next, the dataset is checked for any missing values Check Missing Values to ensure data completeness and quality. Following this, Feature Engineering is performed to create or modify features, enhancing the dataset with meaningful information, which is then refined through Feature Extraction to retain only the most relevant features. The core of the process is the

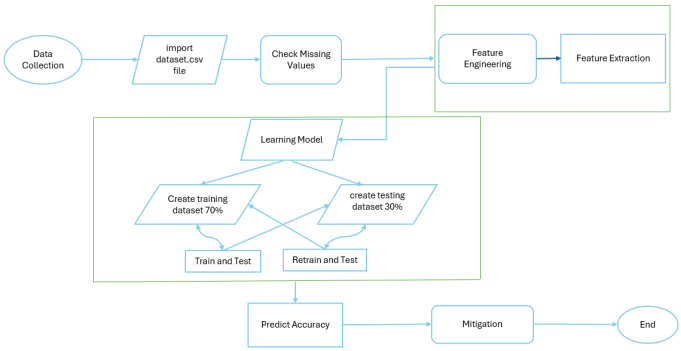


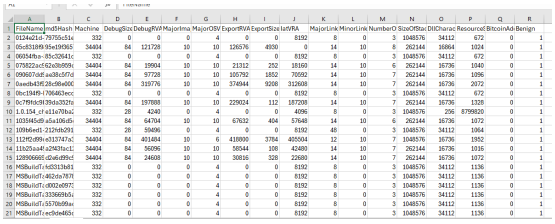
Fig. 3. Ransomware attack detection algorithm for protection.

Learning Model step, where the dataset is split into training and testing sets, with 70% used for training and 30% for testing. The model undergoes an initial Train and Test phase, where it is trained on the training dataset and evaluated on the testing dataset. The density-based method and KNN are used in this paper [1]. This study makes use of feedforward neural networks, support vector machines, and logistic regression [2].

If the performance is unsatisfactory, the model is iteratively Retrained and Tested to improve its accuracy. The Predict Accuracy step calculates the model's prediction accuracy, and any issues identified are addressed through Mitigation. Finally, the process concludes with the End step, marking the completion of the workflow. This systematic approach ensures a robust and effective machine learning model development.

A. Dataset Selection

we use open sourch data set from kaggel.



B. Evaluation Metrics Selection

True positives (TP): Happen when a positive data point is correctly predicted by the model. (geeksforgeeks Organization, 2023).

True negatives (TN): Happen when a negative data point is correctly predicted by the model. (geeksforgeeks Organization, 2023).

False positives (FP): Occur when a positive data point is wrongly predicted by the model. (geeksforgeeks Organization, 2023).

False negatives (FN): Happen when a negative data item is mispredicted by the model. (geeksforgeeks Organization, 2023).

Sensitivity (Recall): Quantifies the percentage of positive cases that the model successfully detected.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{1}$$

where FN represents the quantity of False Negatives (cases in which the model predicts the negative class given a positive case wrongly). (geeksforgeeks Organization, 2023).

Precision: Calculates the percentage of optimistic forecasts that came true.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

where FP is the quantity of False Positives, or cases in which the model predicts the positive class for a negative case in error. (geeksforgeeks Organization, 2023).

F1 Score: A single statistic that balances the trade-off between precision and recall is the harmonic mean of the two. (geeksforgeeks Organization, 2023).

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

Accuracy: Determines what percentage of all forecasts positive and negative were accurate.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

the number of True Negatives (TN) is the number of cases where the model accurately predicts the negative class. (geeksforgeeks Organization, 2023).

In actuality, the confusion matrix—a table that compares the actual and predicted classes to summarize the performance of a classification model—is where the values of TP, FP, TN, and FN are found.

22910	1779
2167	16883

TABLE I
CONFUSION MATRIX FOR RANDOM FOREST CLASSIFIER.

10508	170
272	7796

TABLE II
CONFUSION MATRIX FOR K-NEAREST NEIGHBORS (KNN).

10489	189
157	7911

TABLE III
CONFUSION MATRIX FOR CONVOLUTION NEURAL NETWORK.

To evaluate the efficacy of classification models, we employ classification reports, a machine learning performance evaluation tool.

	Precision	Recall	F1-score	Support
Class 0	0.91	0.93	0.92	24689
Class 1	0.90	0.89	0.90	19050
macro avg	0.91	0.91	0.91	43739
weighted avg	0.91	0.91	0.91	43739

TABLE IV
CLASSIFICATION REPORT FOR RANDOM FOREST CLASSIFIER.

	Precision	Recall	F1-score	Support
Class 0	0.97	0.98	0.98	10678
Class 1	0.98	0.97	0.97	8068
macro avg	0.98	0.98	0.98	18746
weighted avg	0.98	0.98	0.98	18746

TABLE V
CLASSIFICATION REPORT FOR K-NEAREST NEIGHBORS (KNN).

RESULTS ANALYSIS AND DISCUSSION

The Convolutional Neural Network (CNN) outperforms the Random Forest Classifier and K-nearest Neighbors (KNN) in all performance metrics, demonstrating its superior effectiveness in ransomware detection.

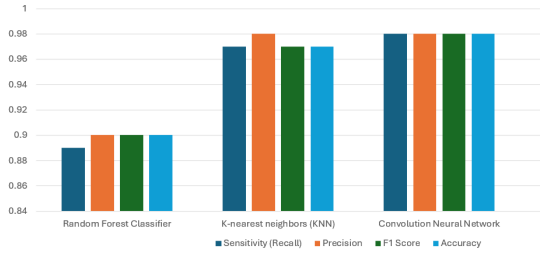


Fig. 4. Evaluation Metrics Results.

Model	Sensitivity (Recall)	Precision	F1 Score	Accuracy
RF	0.89	0.90	0.90	0.90
KNN	0.97	0.98	0.97	0.97
CNN	0.98	0.98	0.98	0.98

TABLE VI
COMPARISON OF EVALUATION METRICS.

CONCLUSION

This study used machine learning and deep learning to suggest an effective way to identify and mitigate ransomware utilizing the FREEDOM system [5]. The three Machine Deep Learning algorithms Random Forest Classifier, KNN, and CNN were used in this detection approach to assess the success rate of each method's similarity identification on the gathered ransomware assault. For this specific ransomware attack, all three detection techniques yielded extremely high similarity detection findings (over 89%). The algorithm's accuracy is 90%, 97%, and 98%, in that order.In comparison to the other two Random Forest Classifiers and KNN, the CNN yielded comparatively superior outcomes. This suggested CNN is a quick and effective way to identify the attack.

	Precision	Recall	F1-score	Support
Class 0	0.99	0.98	0.98	10678
Class 1	0.98	0.98	0.98	8068
macro avg	0.98	0.98	0.98	18746
weighted avg	0.98	0.98	0.98	18746

TABLE VII
CLASSIFICATION REPORT FOR CONVOLUTION NEURAL NETWORK.

ACKNOWLEDGMENT

I want to start by thanking Allah Ta'ala for giving me the ability to finish the Project. After that, I would want to express my profound gratitude to our honored instructor Waleed Al-Gobi, an assistant professor in the King Fahd University of Petroleum and Minerals' Department of Information and Computer Science. His motivation has encouraged me to finish. Lastly, I want to express my gratitude to my family and parents for their unwavering love and wonderful support throughout my studies.

REFERENCES

[1] Jian Du, Sajid Hussain Raza, Mudassar Ahmad, Iqbal Alam, Saadat Hanif Dar, and Muhammad Asif Habib, "Digital Forensics as Advanced Ransomware Pre-Attack Detection Algorithm for Endpoint Data Protection," in). International Journal of Safety and Security Engineering,2021 .

[2] Chandra Thapa, Kallol Krishna Karmakar, Alberto Huertas Celdran, Seyit Camtepe, Vijay Varadharajan, "FedDICE: A Ransomware Spread Detection in a Distributed Integrated Clinical Environment Using Federated Learning and SDN Based Mitigation," in Quality, Reliability, Security and Robustness in Heterogeneous Systems Conference paper, 2021.

[3] Rusydi Umar, Imam Riadi, Ridho Surya Kusuma, "Mitigating Sodinokibi Ransomware Attack on Cloud Network Using Software-DefinedNetworking (SDN)," in International Journal of Safety and Security Engineering,2021.

[4] Nitin Naik, Paul Jenkins and Nick Savage, "A Ransomware Detection Method Using Fuzzy Hashing for Mitigating the Risk of Occlusion of Information Systems," IEEE International Symposium on Systems Engineering (ISSE),2019.

[5] Joshua Morris, Dan Lin, and Marcellus Smith, "Fight Virus Like a Virus: A New Defense Method Against File-Encrypting Ransomware," Journal of ResearchGate,2021.

[6] geeksforgeeks Organization. (2023, 5 5). metrics-for-machine-learning-model. Retrieved from https://www.geeksforgeeks.org/metrics-for-machine-learning-model/

APPENDIX

The deep learning and machine learning code: https://colab.research.google.com/drive/1cc5QfY$_e$1$QvkR55P89Xi4nUY$ $sharing$

The Dataset: https://www.kaggle.com/datasets/amdj3dax/ransomware-detection-data-set .