iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.65915

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

Blockchain-Enabled Security Framework against Ransomware Attacks Using Machine Learning

Prince Yadav[1], Harsh Patel[2], Meraj Shaikh[3], Arpit Dahat[4], Dr. Vinod Wadne[5]
Dept. of IT, JSPM BSIOTR, Pune, India

Abstract: ***In this presentable paper, we first describe and analyze the background of a certain discipline by reviewing the recent scientific literature on this subject and summarizing it about issues of the new research. Ransomware through reasoning can also be considered the last significant stage for cyber extortion blocking access to the resources of the target organization until submission by the latter to certain coercion or payment. There is a separate class of malware known as ransomware. When a computer or some other device suffers a ransomware malware attack, such device is either locked/held hostage or the data within the device is encrypted. Ransom demands (usually a small sum) are placed on the victims for providing the data order for the data translation in the form of a decryption key. Due to these attacks, surveillance means and protective programs are recommended given the prevention of ransomware epidemic outbreaks. The most vulnerable targets are probably those classed as organizations, such as financial institutes and healthcare sectors. Blockchain technology prevents tempering which makes it much more effective than the traditional centralized approach of data storage. Such aspects of blockchain technology can enhance the security perimeter for the detection and prevention of ransomware attacks even more. The objectives of the research are to demonstrate the extent of the problem and to show how problems are identifiable within datasets, which is through the application of machine learning. In this paper, we propose a new security framework that applies machine learning to prevent ransomware attacks and is based on the principles of blockchain technology.***

Keywords: Blockchain technology, Ransomware attacks, Cybersecurity, Machine learning, Data integrity, Network traffic analysis.

I. INTRODUCTION

Ransomware attacks are one of the most dangerous and rapidly growing threats that one can encounter within the field of information technology security. During a ransomware attack, the assailants employ certain software that inflicts restrictions on access to important information belonging to the victim while at the same time making payment demands for access to the information. Malware attacks of this nature have become a troubling issue in the world today with several areas like health care, banking, and government under siege due to the sensitive information present in those sectors. Also, these attacks do not only interfere with the day-to-day activities of many businesses but they incur costs, damage the intended images, and cause strategic operational problems for the concerned businesses for a while even after the attack. With the aggressive escalation trends of organized cybercrime and particularly the tension witnessed in cyber raids, it becomes imperative for organizations to upgrade their security systems to combat such threats proactively. Classic security vulnerabilities are oriented more towards the objectives of a ransomware attack. The objectives in this case are detection and restoring the data which has been encrypted. Though this is useful in some cases, such methodologies do not help to mitigate the threat posed by ransomware before its occurrence. In addition, recovering the data or paying out the ransom puts the victim at discretion which is far risky and unfair to the victims. The importance of having an all-encapsulating, 360-degree security policy in place has never been as stark as it is now. This underlines the need for a preventative and detective solution to the challenges of ransomware infection. The features of blockchain such as its decentralized, transparent, and unalterable nature make it possible to solve the challenges of ransomware attacks. Since data on the blockchain can be stored in a manner that does not allow alteration of its contents, these systems can be utilized to develop an effective backup strategy that will protect all vital information during an ongoing ransomware assault. Also, since applications of the blockchain are distributed, there is less risk of attack which makes it impossible for the hackers to infiltrate the system easily.

We develop a Blockchain-Enhanced Security Framework against Ransomware that integrates the advantages of blockchain technology and machine learning for containment and recovery after a ransomware attack. This framework is designed to extend integrated security beyond one of solely reactively responding to threats. To be specific, we also incorporate a modicum of blockchain technology so as to do secure backup storage to eliminate the possibility of paying any ransom in order to retrieve important information. Early stage containment of ransomware is also carried out with the help of machine learning which detects malicious files, with special attention to portable executable (PE) files, the common weapon for any ransomware tactician.

II. LITERATURE SURVEY

Title : Ransomware: Ez a Survey and niego tendencje
Sana Aurangzeb, Muhammad Aleem, Muhammad Azhar Iqbal, Muhammad Arshad Islam
This survey is on the increasing menace posed by ransomware, a form of malware that either encrypts or deletes files on a computer and demands for the payment of a ransom, in most cases in bitcoin. The paper synthesizes 40 studies completed over four years about Windows-based ransomware and includes structural features, attack methods, methods of payment addressing. The authors claim future studies should concentrate on the formulation of techniques and strategies that would deal with the menace of ransomware at it's earliest stages and prevent its escalation.

Title: Crypto-ransomware Detection Using Machine Learning
Models in File-sharing Network Scenario with Encrypted
Traffic. Authors: Eduardo Berrueta et al.
In this work named as A Machine Learning-Based Tool for Ransomware Detection in File-Sharing Networks, the authors propose a machine learning detection tool for ransomware in file sharing networks. It examines the traffic of clients and file server to uncover the 'ransomware behavior' of file replacement. The model that has been evaluated with the majority of over seventy ransomware types and the user's actual traffic can successfully identify both known and novel types of ransomware threats, including the detection in encrypted traffic.

Title: A Survey on Detection Techniques for Cryptographic Ransomware. Authors: Eduardo Berrueta et al.
This paper provides a review of different methods used for detecting crypto-ransomware, which targets user files by encrypting them and demanding for a ransom. The research discusses the classification of detection algorithms with respect to the inputs and decision-making processes used in their operation, and the history of development of algorithms for detecting ransomware. The authors also reiterate the importance of routine changes to the detection strategies, considering the persistent advancement of ransomware attacks.

Title: Blockchain for Internet of Things: A Survey
Authors: Hong-Ning Dai, Zibin Zheng, Yan Zhang
The authors present a survey paper on how blockchain technology can be integrated with IoT. They term it 'blockchain of things (BCoT)' and analyze the capabilities of this technology in relation to impact of IoT in terms of decentralization, security, and privacy concerns. Then they describe BCoT architecture and examine its prospects in the 5G and and the components of industrial internet of things (IIoT) applications and also present future perspectives on this convergence.

III. METHODOLOGY

A. Overview of the Proposed Framework

This research proposes a Blockchain-Enabled Security Framework for ransomware detection and prevention in smart healthcare systems. The framework blockchain-enabled security framework against ransomware attacks , leverages the combined power of blockchain and machine learning to mitigate ransomware attacks proactively. The methodology involves three main components: continuous monitoring, malware detection, and file recovery.

1) Continuous Monitoring and Prevention blockchain-enabled security framework against ransomware attacks continuously monitors the system for suspicious files and activity. When potential ransomware behaviours are detected, the system automatically removes suspicious files, minimizing infection risk and stopping the attack before it spreads.
2) Detection of Ransomware in Portable Executable Files Ransomware typically leverages Portable Executable (PE) files to gain access and initiate encryption. Our system focuses on detecting ransomware in these files using a combination of machine learning techniques, including feature extraction, selection, correlation analysis, and classification. By scanning PE files through a React-based interface, we can classify each file as either legitimate or infected based on its features.
3) Backup and Recovery via Blockchain
4) User data is backed up on a decentralized storage network integrated with blockchain, ensuring data integrity and availability. This backup system prevents data loss by allowing users to recover their files without paying ransom in case of an attack.
5) Secure Ransom Payments In worst-case scenarios where data recovery is impossible, the framework provides a secure blockchain-based mechanism for ransom payments, ensuring safe and traceable transactions.

B. Analysis of Existing System

Existing systems primarily focus on post-attack detection and recovery, often relying on signature-based methods that struggle against advanced, obfuscated ransomware variants. They lack proactive prevention measures, instead reacting only after ransomware has infiltrated. Centralized storage in these systems creates a single point of failure, leaving data vulnerable during attacks. Additionally, traditional recovery can be slow and costly, and without secure transaction protocols, ransom payments are at risk. This reactive approach limits effectiveness in high-security environments like healthcare, where rapid response and data integrity are critical.

C. System Architecture

The blockchain-enabled security framework against ransomware attacks system architecture is designed to address both the antecedence and aftermath of ransomware attacks. It consists of several interconnected modules:

1) User Module: Users log into the system to back up data and check files for infections. Using a React-based interface, they can scan executable files to determine if they are infected. In cases where data cannot be recovered, users may opt to make a secure ransom payment.

2) Attacker Module (Demo): This module demonstrates the behaviour of ransomware, including Locker and Crypto ransomware types. It provides a real-time demonstration of ransomware's impact and recovery options.

3) Ransomware Detection Module: This module utilizes machine learning to detect ransomware in PE files. Through a series of steps—data preprocessing, feature selection, correlation analysis, and classification—the system detects if files are legitimate or infected.

4) Blockchain-Integrated Data Backup and Recovery Module:This component ensures that user data is securely stored on a decentralized network. Blockchain records content identifiers (CIDs) generated by IPFS for each backup file, ensuring that files remain tamper-proof and retrievable in case of ransomware incidents.

5) Payment and Transaction Module:For scenarios requiring ransom payment, this module enables secure blockchain transactions to the attacker, making payments transparent and traceable.

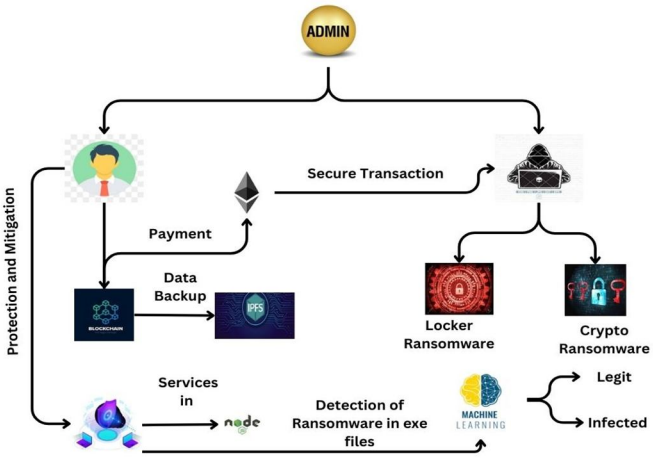


Figure 1 System Architecture

D. Technical Architecture

The blockchain-enabled security framework against ransomware attacks approach has a decentralized architecture toward preventing, detecting, and recovering from ransomware attacks in smart healthcare. The React-based frontend of the built application allows for scanning Portable Executable (PE) files and makes data backups easier to manage. The core detection engine makes use of machine learning for feature extraction and selection, with variance thresholding and correlation analysis included to ensure data reduction with an accuracy of ransomware detection. A wrapper over multiple algorithms of a Lazy classifier evaluates different models in search of the one that best classifies a file as legitimate or infected: KNN, SVM, or even Random Forest.

Finally, it implements decentralized storage by using a Content Identifier for each individual file, leading to effective content-addressable access of the content in the nodes across the network, thus reducing the impact on single points of storage. Thus, blockchain technology is then employed to secure the CIDs in an immutable ledger that would ensure integrity and traceability of the backed-up files, thereby being tamper-proof. Smart contracts further automate backup and recovery-related actions; therefore, they add reliability to the system.

In scenarios in which recovery cannot be supported otherwise, a blockchain-based module for secure payment allows traceable ransom transactions; therefore, more security on the critical ransom payments is provided. A continuous monitoring mechanism catches suspicious activity real time. Infected files are automatically removed to prevent further spread. This hybrid architecture of machine learning and blockchain-backed decentralized storage offers complete ransomware protection for smart healthcare environments.

IV. ALGORITHM DEVELOPMENT

The blockchain-enabled security framework against ransomware attacks algorithm integrates machine learning and blockchain technologies for proactive detection, classification, and data recovery against ransomware attacks within the framework. Below follows a detailed step-by-step outline of the algorithm development process.

A. Data Collection and Preprocessing

Dataset Selection: Collect a suitable dataset, with both normal and ransomware-infected Portable Executable (PE) files. This dataset will be used to train and test the developed machine learning models.

Data Cleaning:

- Outlier Removal: Applied Z-score and IQR values, for identification of removing outliers, thus enhancing data quality.
- Handling Missing: Values: The null or missing values in the dataset are addressed either by replacing them with mean/mode imputation or by deleting rows affected for completeness and consistency.

Feature Extraction

- Extracting important features from each PE file. Some of them are as follows:
- File Size and Entropy: Indications of potential obfuscation techniques
- Header Information: The attributes in the PE file's header such as entry point, timestamp, etc. reflecting structural patterns utilized commonly by malware.
- Imported/Exported Functions: List of imported/exported functions, which is commonly modified by ransomware.
- Section Characteristics: Metrics associated with sections in the PE file, for example, .text, .data and .rsrc sections, to trace behavioural signatures of ransomware.

B. Feature Selection and Optimization

1. Variance Threshold: The variance threshold helps in the removal of low-variance features that have very little contribution to information during the decision-making process of the model. This means that only those features will remain that highly vary for good differentiations between the legitimate files and the infected files.
2. Correlation Analysis: After variance thresholding, correlation analysis is performed to find the features that are more correlated with a correlation coefficient > 0.8. Highly correlated features cause multicollinearity. Due to multicollinearity, the accuracy of the model might be reduced. To remove multicollinearity, one feature from each pair of correlated attributes is deleted, which increases model efficiency.

C. Classification using Lazy Classifier Algorithm

1) Lazy Classifier Evaluation: Lazy classifier algorithm is applied on several machine learning classifiers to the refined feature set. A lazy classifier is a wrapper that evaluates different algorithms without a specific model having been defined in advance.

2) Model Candidates

These are the classifiers evaluated on their accuracy, precision, recall, and F1-score values:

a) KNN-K-Nearest Neighbours: It classifies the sample based on its nearest neighbours, which helps to further identify the similarity between known and new samples. It depends on finding the closest neighbours for the classification of samples.

Euclidean Distance for K-Nearest Neighbors (KNN) Classification

- KNN classification calculates the Euclidean distance between a test point and all points in the training data.
- A*i* and B*i* represent the coordinates of points A and B in each dimension *i*.
- (A*i*−B*i*)^2calculates the squared difference in each dimension.

$$\text{Euclidean Distance}(\mathbf{A},\mathbf{B}) = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2}$$

- Summation over all dimensions gives the total squared difference.
- Square root provides the final Euclidean distance.

b) *Random Forest:* It is an ensemble learning method that uses multiple decision trees, which can handle complex feature interactions.

Entropy Calculation for Random Forest and Decision Tree Classification

- Entropy measures the impurity in a dataset and is used for splitting nodes in decision trees.
- p*i*: The probability of an instance belonging to class *i* in the dataset.

$$H = -\sum_{i=1}^{c} p_i \log_2(p_i)$$

- Summation: This iterates over all classes *c* in the dataset, calculating the contribution of each class to the entropy.
- Logarithmic Term: The log base 2 is typically used to express entropy in bits.

c) Naive Bayes: It's a probabilistic classifier that assumes features are independent of one another and involves simple classification tasks

d) Support Vector Machine (SVM): It develops a hyperplane in such a way that it can maximize separation between classes, and this data is clearly separable.

3) Optimal Model Selection: Using the best model from the evaluation results regarding Lazy classifier's accuracy and F1-score as the primary classifier, this model performs the task in detecting ransomware. Then this classifier is trained on the entire dataset and applied for real-time use in classifying files.

D. Integration with Blockchain for Decentralized Backup and Security

1) Data backup on IPFS and Blockchain: Classification The framework takes a backup of the classified data on IPFS-a decentralized network for storage. Each file that is taken as a backup generates within the IPFS a Content Identifier called CID that will refer to the file when it retrieves it.

2) CID can be stored in Blockchain: The CID is saved on a blockchain in a safe manner. This results in an immutable, tamper-proof record of where each file on IPFS is located. It also makes sure that in case of a ransomware attack, this link ensures that the data integrity might be recovered quickly.

3) Secure Payment Protocol (if needed): Extreme scenarios, where data recovery is impossible, are guarded by a blockchain-based protocol for ransom payments. The last resort ensures tracing and security in case of ransom payments being necessary.

V. IMPLEMENTATION

The implementation of the Blockchain-Enabled Security Framework against Ransomware Attacks Using Machine Learning project can be broken down into detailed steps to develop, integrate, and deploy the system's components, combining blockchain technology and machine learning for proactive ransomware protection.

A. Set Up Development Environment

1) Install Node.js and React for frontend and backend development.

2) Install Visual Studio Code for code development and management.

3) Install Ganache and MetaMask for local blockchain development, enabling testing and simulation of blockchain transactions.

B. Blockchain Framework Development

1) Set up a blockchain network using Ganache to have an in-house record of all backup transactions and also user data management.

2) Smart Contract Development: Write smart contracts in Solidity for secure data backup, restoring functionality, and optional ransom payment if needed.

3) Blockchain Integration: MetaMask will integrate for secure authentication of the user and blockchain transactions. Ganache will connect with MetaMask for testing the transactions locally.

4) Data Backup System Blockchain-based Data Backup Securely Blockchain System Review the backup immutability and recoverability in a ransomware attack by utilizing data backup system.

C. Machine Learning Model Development

The first step in creating a machine learning model for ransomware detection stands in data preprocessing. For instance, a honeypot dataset undergoes its data-cleaning process, which means getting rid of noise or outliers and null values that would otherwise hinder the quality of the data. Thereafter, normalization of data through standardization is done towards training the data. Relevant feature extraction comes next; in this case however, we are interested in the specific details of the ransomware's behavior within the portable executable (PE) files and why these details are important in classifying these files as clean or infected.

A set of machine learning algorithms is used to classify files during the training and optimization phase of model application. The dataset is subdivided in a training and testing one in order to train the model and to adjust its configurations to achieve the maximum accuracy possible. The model performance is assessed using several indicators such as accuracy, F1-score among others which enables choosing the best performing model. Last but not least, the trained and verified model is put to active use in the backend system for live scanning and classification of files which helps strengthen the defense of the system against ransomware attacks and other threats.

D. Frontend and User Interface Development

The design of the user interface is intended to leverage the React framework to enhance user experience with respect to features such as file scanning, viewing alerts, taking backups, and control for recovering lost files. This interface will be coupled with a blockchain network to handle secure backup transaction and also with a machine learning model in order to allow scanning and detection of ransomware in real-time. Access to the system will be provided through secure means using MetaMask so that the blockchain can only be accessed by authorized users and the features of the system can be utilized.

E. File Monitoring and Detection

An always-on file scanning service will be established in order to prevent or minimize the effects of ransomware attacks. This service works in conjunction with the machine learning model responsible for evaluating the files and designating them as infected or clean in real-time. In such an event, there is a high possibility that the infected file will be cleaned, restored, executed or otherwise used, and hence such file will be quarantined or deleted by the application and the user will be notified through their appropriate interface in advance of damage.

F. Backup and Recovery Mechanism

Regular backups will be programmed for the purpose of storing data onto the blockchain, so that there will be deposit of data which will be safe from any alterations. In case of an assault, a protection feature will assist in getting the information back in these recorded blockchain backups. The smart contracts will manage the automatic operation services for backup planning and also will activate restoration in case of need, hence making the process of backup and recovery effective for management of data in a secure manner.

G. Testing and Optimization

Functional testing of all the essential features including detection, backup, and recovery will be performed comprehensively. Security testing will evaluate the system by placing it under attack and determining whether unauthorized access and ransomware threats are successful. Furthermore, performance testing will look into how efficiently the machine learning model and the transaction speed of the blockchain work in the system to ensure that there is fast, effective real-time processing throughout the system.

VI. FUTURE ENHANCEMENTS

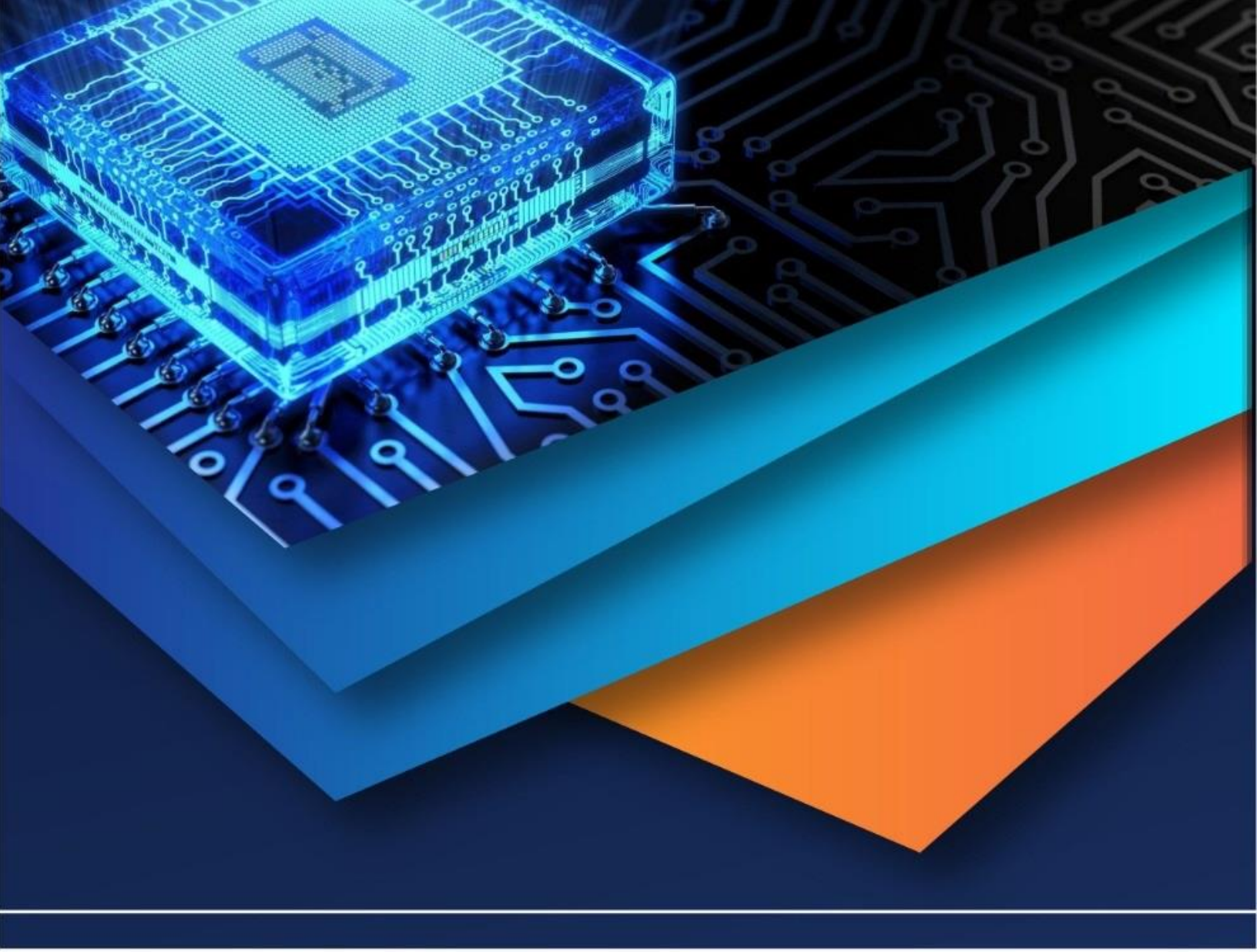
1) *Crypto Ransomware Decryption:* Plans to enhance data recovery by decrypting files affected by Crypto ransomware without paying a ransom.
2) *Exploration of Advanced ML Models:* Additional ML models may be tested to improve the detection and classification accuracy further.
3) *Improved Recovery Capabilities:* Future iterations could focus on strengthening recovery procedures to address evolving ransomware techniques.

VII. CONCLUSION

The Blockchain-Based Security Framework to Ransomware Using Machine Learning, a system capable of responding to ransomware attacks through the project presented here integrates the concepts of machine learning and blockchain in addressing the challenge of ransomware in a preventative measure and a remedial measure to an extent. Powered by a React user interface, the system leverages sophisticated feature extraction, variance thresholding, and correlation to scan Portable Executable (PE) files to provide appropriate data for classification. To maximize the performance capabilities of the framework in separating good and bad files, a Lazy classifier algorithm is employed to determine the best machine learning model suitable for the task. Moreover, IPFS is used to provide data and files in a distributed manner, while distributed ledger system of blockchain is used to track Content Identifiers (CIDs) enabling assurance of safety, accessibility and backup of data eliminating the challenges that come with the hazards of centralized storage. The system also implements an automatic response mechanism along with a continuous monitoring system that is able to contain suspicious acts and curb the spread of an infection. The blockchain-enabled security framework against ransomware attacks , in contrast to other existing systems, is a developed system that is decentralised and aimed at responding to attacks before they take place, with specific regard to the exceptional preparedness cut-out especially in health care. Enhancements in the fierceness of the warfare would most likely incorporate some aspects of deep learning and more changeable parameters to improve efficiency and responsiveness to new forms of ransomware**.** Taken all together, blockchain-enabled security framework against ransomware attacks represents a solid structure wherein machine learning and blockchain technologies are interconnected to provide a viable and dependable solution to ransomware.

REFERENCES

[1] S. S. Chakkaravarthy, D. Sangeetha, M. V. Cruz, V. Vaidehi, and B. Raman, “Design of intrusion detection honeypot using social leop ard algorithm to detect IoT ransomware attacks,” IEEE Access, vol. 8, pp. 169944–169956, 2020.
[2] M. Wazid, A. K. Das, V. Odelu, N. Kumar, and W. Susilo, “Secure remote user authenticated key establishment protocol for smart home environment,” IEEE Trans. Dependable Secure Comput., vol. 17, no. 2, pp. 391–406, Mar./Apr. 2020.
[3] S. Tian, W. Yang, J. M. L. Grange, P. Wang, W. Huang, and Z. Ye, “Smart healthcare: Making medical care more intelligent,” Global Health J., vol. 3, no. 3, pp. 62–65, 2019.
[4] E. Berrueta, D. Morato, E. Magana, and M. Izal, “A survey on detec tion techniques for cryptographic ransomware,” IEEE Access, vol. 7, pp. 144925–144944, 2019.
[5] D. Farhat and M. S. Awan, “A brief survey on ransomware with the perspective of Internet security threat reports,” in Proc. 9th Int. Symp. Digit. Forensics Security (ISDFS), Elazig, Turkey, 2021, pp. 1–6.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY