Decentralized Malware Attacks Detection using Blockchain

Sheela S[1*], *Shalini* S[2], *Harsha* D[3], *Chandrashekar* V T[3] and *Ayush* Goyal[3]

[1]Assistant Professor, Global Academy of Technology, Department of Computer Science and Engineering, Karnataka, Bangalore-560098, India
[2]Associate Professor, Global Academy of Technology, Department of Computer Science and Engineering, Karnataka, Bangalore-560098, India
[3]Global Academy of Technology, Department of Computer Science and Engineering, Karnataka, Bangalore-560098, India

Abstract. This research introduces an approach to detect malware attacks using blockchain technology that integrates signature-based and behavioral-based methods. The proposed system uses a decentralized blockchain network to share and store malware signatures and behavioral patterns. This enables faster and more efficient detection of new malware files. The signature-based method involves storing the signatures in the blockchain and the sharing of the signature of malware files among the user nodes of the p2p blockchain network, while the behavioral-based approach analyzes the behavior and actions of files in a separate virtualized environment to identify suspicious patterns. This system addresses the limitations of conventional signature-based methods, which can be evaded by polymorphic malware, and behavioral-based methods, which may generate false positives. The results of the evaluation indicate that the proposed system achieves high detection rates while maintaining low false positives. Overall, the proposed system offers an effective and efficient approach to malware detection by utilizing the strengths of both signature-based and behavioral-based methods and utilizing the security and transparency benefits of blockchain technology.

1 Introduction

Malware detection is a crucial aspect of cybersecurity as it enables the identification and prevention of harmful software from causing damage to computer systems and networks. With the increasing sophistication of malware, traditional detection methods have become less effective, leading to the need for innovative solutions.

*Corresponding author: sheela.s@gat.ac.in

One such solution is the use the decentralized blockchain technology in conjunction with signature-based and behavioral analysis methods for malware detection. Blockchain's distributed and immutable nature provides a secure and transparent platform for storing and sharing information related to malware, such as signature patterns and behavioral data. Signature-based detection involves the use of pre-defined patterns, or signatures, to identify malware. Behavioral analysis, on the other hand, focuses on identifying unusual or malicious behavior of software to detect previously unknown malware. By combining these two methods and utilizing the blockchain's capabilities, it is possible to create a powerful and reliable detection system for malware.

This paper explores the potential of using blockchain for malware detection, specifically focusing on signature-based and behavioral analysis methods. We will discuss the benefits and limitations of this approach and provide a detailed analysis of the technology and how it can be implemented in practice.

2 Blockchain technology

Blockchain technology has emerged as a revolutionary innovation that has the potential to transform various industries. It is a decentralized and also distributed ledger that secures and transparently documents transactions. Unlike traditional centralized systems, blockchain allows for secure peer-to-peer transactions without the need for intermediaries.

The blockchain concept was first introduced in the year 2008 as a major key component of the cryptocurrency, Bitcoin. However, the application of blockchain technology extends beyond digitalized currencies. Blockchain can be used to track and manage digital assets, establish smart contracts, and enhance the security of data storage and transfer.

At its core, a blockchain is a collection of a chain of blocks, where each block in the blockchain network contains a set of transactions. When a single block is appended to the chain in the blockchain network, it is impossible to alter the information contained within it. This is because every block in the blockchain network is linked to the previous block forming the chain of blocks, forming an immutable chain of records. This decentralized and tamper-proof nature of blockchain makes it ideal for applications requiring high security and transparency.

3 Literature survey

Malware detection has been a challenging task due to the ever-evolving nature of malicious software. Blockchain technology has emerged as a promising solution to enhance the security of malware detection systems. Gu et al. (2018) proposed a consortium blockchain-based malware detection system for mobile devices. The system uses a consensus algorithm to detect malware and share the results with other devices in the network. The authors tested their system on a dataset of 5,000 Android apps and achieved an accuracy of 96.3% [1]. Kumar et al. (2019) developed a multimodal malware detection technique for Android IoT devices. The system uses various features, such as permissions, API calls, and network traffic, to detect malware. The authors evaluated their system on a dataset of 2,000 Android apps and achieved an accuracy of 97.8% [2]. Alotaibi (2021) proposed a blockchain-based malware detection system for the Internet of Medical Things (IoMT). The system uses a deep learning model and the Ruzicka index to detect malware in medical devices. The authors tested their system on a dataset of 1,000 malware samples and achieved an accuracy of 99.3% [3]. Punithavathi et al. (2022) proposed a crypto hash-based malware detection system for the IoMT framework. The system uses a hashing algorithm to detect and classify malware in medical devices. The authors tested their system on a dataset of 4,000 malware samples and achieved an accuracy of 98.5% [4]. Raje et al. (2017) proposed a decentralized firewall for malware detection. The system uses a blockchain to distribute the firewall rules and detect malicious traffic in a peer-to-peer network. The authors tested their system on a simulated network and showed it can detect malware effectively [5]. Rana et al. (2019) evaluated the performance of machine-learning models for Android malware detection on the Ethereum blockchain. The authors showed that their system can achieve high accuracy while maintaining low computational overhead.[6] Anita and Vijayalakshmi (2019) surveyed various security attacks on blockchain and discussed the potential impact of these attacks on different blockchain applications, including malware detection.[7] Fuji et al. (2020) proposed a blockchain-based malware detection method that uses shared signatures of suspected malware files. The authors tested their system on a dataset of 10,000 malware samples and achieved an accuracy of 97.3%.[8] Homayoun et al. (2019) proposed a blockchain-based framework for detecting malicious mobile applications in app stores. The system uses a consensus algorithm to verify the authenticity of apps and detect malware. The authors evaluated their system on a dataset of 10,000 apps and achieved an accuracy of 98.7%.[9] Moubarak et al. (2018) proposed a K-ary malware that uses blockchain to evade detection by traditional antivirus systems. The authors showed that their malware can bypass several popular antivirus programs and argued that blockchain-based malware can pose a significant cybersecurity threat.[10] Saad et al. (2019) explored the use of machine learning in malware detection and discussed the challenges and limitations of current machine-learning approaches. The authors argued that a combination of different techniques, such as static and dynamic analysis, can improve the accuracy of malware detection.[11]

4 Proposed system

Malware is a persistent and evolving threat to the security of computer systems. Traditional antivirus software is no longer enough to protect against sophisticated malware attacks. As a result, there has been a growing interest in using emerging technologies such as blockchain to increase the effectiveness and accuracy of detection of the malware.

The method proposed here involves using of blockchain technology to store malware signatures and Hybrid analysis in a virtualized environment that can be used to detect malware in files and URLs in a computer system. There are two types of users in the proposed system one is a non-testing node user and the other is a testing node user. The non-testing node user can simply check a file for malicious content and update the signature in the blockchain without being part of the testing node and can also check the maliciousness of the file using Hybrid analysis. The testing node user is responsible for running the Hybrid Analysis process on the system and detecting potential malware. The testing node user can also add new malware signatures to the blockchain when new threats are identified.

Since the user is registered as a testing node, he can participate in the voting for a percentage of malware for the malware signatures in the blockchain.
The proposed system is divided into five steps:

- Signature Generation
- Signature Storage
- Hybrid Analysis
- Signature Verification
- Notification and Quarantine

Below is the detailed flow diagram and detailed explanation of the above-mentioned steps :

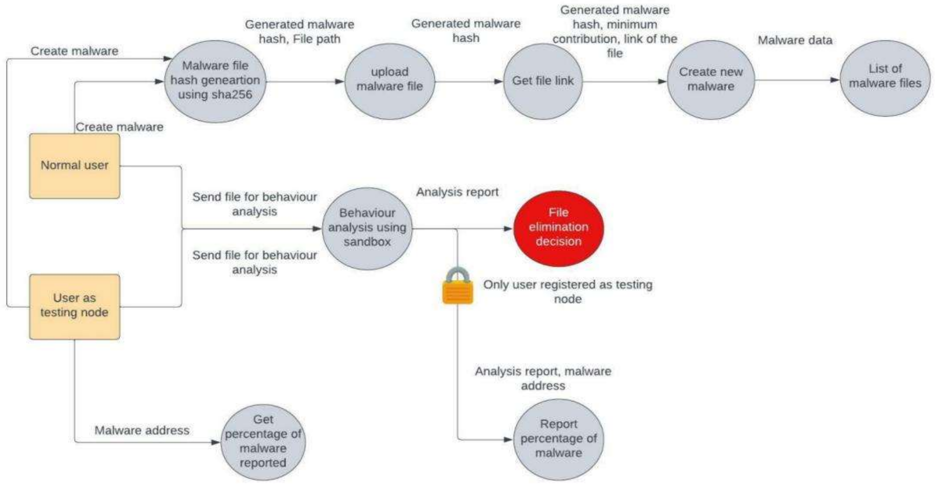

Create malware
Malware file hash geneartion using sha256
Generated malware hash, File path
upload malware file
Generated malware hash
Get file link
Generated malware hash, minimum contribution, link of the file
Create new malware
Malware data
List of malware files
Create malware
Normal user
Send file for behaviour analysis
Behaviour analysis using sandbox
Analysis report
File elimination decision
Send file for behaviour analysis
Only user registered as testing node
User as testing node
Analysis report, malware address
Malware address
Get percentage of malware reported
Report percentage of malware

Fig. 1. Flow diagram of the proposed system

4.1 Signature generation

The first step in the proposed method is to generate a unique signature for each malware by analyzing its code and behavior. The signature can include information such as the file name, hash value, and behavior pattern. The goal is to create a unique identifier that can be used to distinguish the malware from other files on the system.

The signature generation process can be done using various techniques such as static and dynamic analysis. Static analysis involves the examination of the code and the file structure, while dynamic analysis deals with monitoring the file behavior during execution. By combining both techniques, it is possible to generate a signature that accurately represents the behavior of the malware. In the proposed system we have used the SHA256 algorithm to generate the signature (hash) of the files.

4.2 Signature storage

Once the signature is generated, it can be stored on the blockchain. Blockchain technology provides a secure and decentralized way of storing information. By storing the malware signatures on the blockchain, it ensures that the signature cannot be tampered with or modified. It also allows for easy retrieval and sharing of the signature among different nodes.

Here every block in the blockchain contains the following information :->

- Malware file hash.
- The number of testing nodes voted the file as malicious.
- The number of testing nodes voted the file as benign.
- Address of the testing node who voted the file as malicious.
- Address of the testing node who voted the file as benign.

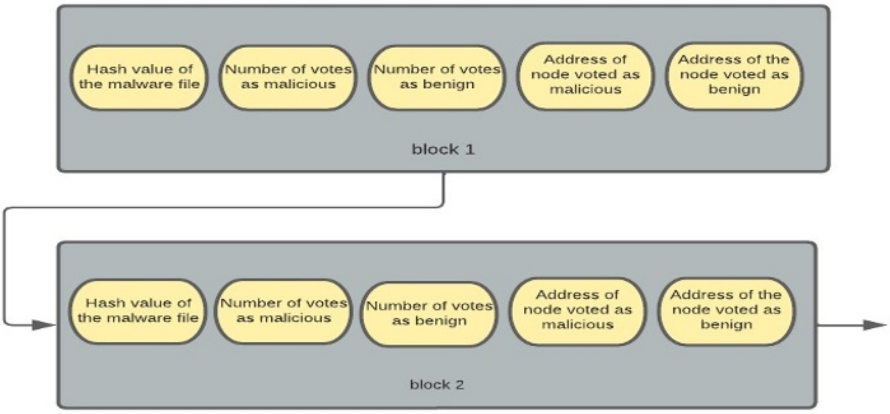

Hash value of the malware file
Number of votes as malicious
Number of votes as benign
Address of node voted as malicious
Address of the node voted as benign
block 1
Hash value of the malware file
Number of votes as malicious
Number of votes as benign
Address of node voted as malicious
Address of the node voted as benign
block 2

Fig. 2. Block containing the information about the malware files.

Apart from that user can also get the link to download the files by using the hash or signatures of the files.

4.3 Hybrid analysis

The next step is to perform a Hybrid Analysis of the system to detect malware. Hybrid Analysis combines both static analysis and also dynamic analysis techniques to identify potential malware threats. Static analysis includes an examination of the code and file structure, while dynamic analysis includes scanning the file behavior during execution. By combining both techniques, Hybrid Analysis can detect malware that may be missed by traditional antivirus software.

Hybrid Analysis can also use machine learning algorithms to analyze the behavior of the system and identify potential threats. Large data sets can be analyzed by machine learning to find trends that may be hard for people to notice. This may aid in enhancing the quickness and precision of malware identification.

After completing the Hybrid Analysis, a detailed analysis report can be downloaded to decide on voting the file as malicious or benign.

4.4 Signature verification

The step-in signature verification if Hybrid Analysis identifies a potential malware, the signature of the file is verified on the blockchain to ensure that it matches with the stored signature. This helps to prevent false positives and false negatives. If the signature matches, it confirms that the file is malware and triggers the appropriate response.

The verification process can be done using various techniques such as cryptographic hashes and digital signatures. These techniques can ensure the integrity and authenticity of the signature, making it more difficult for attackers to manipulate the signature.

4.5 Notification and quarantine

Notification: If the Hybrid Analysis process identifies potential malware on the system, the system can generate a notification to alert the user or administrator of the threat. The notification can include information such as the type of malware, the file name, and the location of the infected file. The notification can be sent through various means such as email, text message, or pop-up window.

The notification process is crucial as it helps the user to take immediate action to mitigate the threat. This can include removing the infected file or quarantining the infected file to prevent further damage.

Quarantine: Quarantine is a crucial step in the malware detection process. When the system identifies potential malware, it can quarantine the infected file to prevent it from causing further harm to the system. Quarantine involves isolating the infected file in a secure location where it cannot interact with other files on the system. The quarantine can help to prevent the malware from spreading to other parts of the system and causing more damage. It can also allow the user to safely remove the infected file without the risk of it causing further harm. During the quarantine process, the system can also perform additional analysis on the infected file to determine the type and severity of the malware. This can help the user to take appropriate action to remove the malware from the system and prevent it from causing further damage.

Below is the detailed system design of the proposed system for detecting malware attacks in the blockchain network.

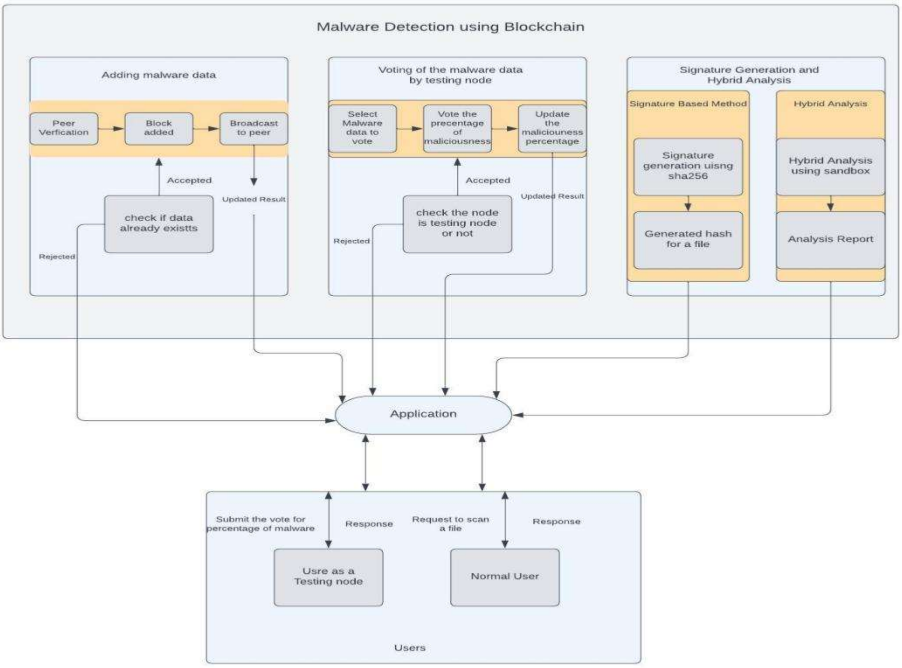

Malware Detection using Blockchain
Adding malware data
Peer Verfication
Block added
Broadcast to peer
Accepted
Updated Result
check if data already existts
Rejected
Voting of the malware data by testing node
Select Malware data to vote
Vote the precentage of maliciousness
Update the maliciouness percentage
Accepted
Updated Result
check the node is testing node or not
Rejected
Signature Generation and Hybrid Analysis
Signature Based Method
Signature generation uisng sha256
Generated hash for a file
Hybrid Analysis
Hybrid Analysis using sandbox
Analysis Report
Application
Submit the vote for percentage of malware
Response
Request to scan a file
Response
Usre as a Testing node
Normal User
Users

Fig. 3. System design of the proposed system

Users will be prompted with an interface. There are two users, a testing node user and a non-testing node user who just wants to know if the file or the URL is malicious. Users who have a meta mask account connected to the hosted website can see the listed testing nodes and files that are available in the system. If the user wants to test the file or URL he can just test and leave. If the user wants to be part of the network, then click on add a testing node, user will be verified and authenticated through meta mask and added a testing node. Tester nodes can vote for the existing files. When the network comes to a consensus, a file is assigned the percentage of maliciousness according to consent. The user will also have an option to just check the file and can choose to download it. The file is decided as malicious based on a consensus of all nodes in the testing network.

The proposed system has several advantages over traditional antivirus software, including:
Improved accuracy: By using blockchain technology, the system can store a large number of malware signatures that can be used to detect new threats. This will decrease the number of false hits and increase the precision of malware identification.
Faster detection: The Hybrid Analysis process allows for faster detection of malware by running the analysis in a virtualized environment.
Decentralized: The system is decentralized, which means that it is not controlled by a single entity. This reduces the risk of a single point of failure and makes the system more resilient to attacks.
Community-driven: The system allows for community participation, which means that users can add new malware signatures and participate in the voting process. This makes the system more transparent and accountable.
Improved privacy: By using blockchain technology, the system can ensure that the malware signatures are stored securely and that user data is kept private.

Overall, the proposed system can significantly improve the precision and effectiveness of detecting malware, while also providing a more decentralized and community-driven approach to cybersecurity.

5 Results and discussion

The proposed approach to malware detection using blockchain technology achieves high detection rates while maintaining low false positives by integrating signature-based and behavioral-based methods. The system utilizes a decentralized and distributed blockchain network to share and store malware signatures and behavioral patterns, enabling faster and more efficient detection of new malware files.

To evaluate the above-proposed system we have proposed a CNN model to detect the malware and benign ratio. The CNN model explores the raw bytes of an EXE file to create an image of EXE.

The input to the above CNN-based model is the list of files we have uploaded to the blockchain network. Here the files in the blockchain network act as test data to the above model and as a result it will output the malicious and benign ratio percentage which can be compared to the malicious percentage of the file present in the blockchain network. The below images show the result of the above proposed CNN experiment.

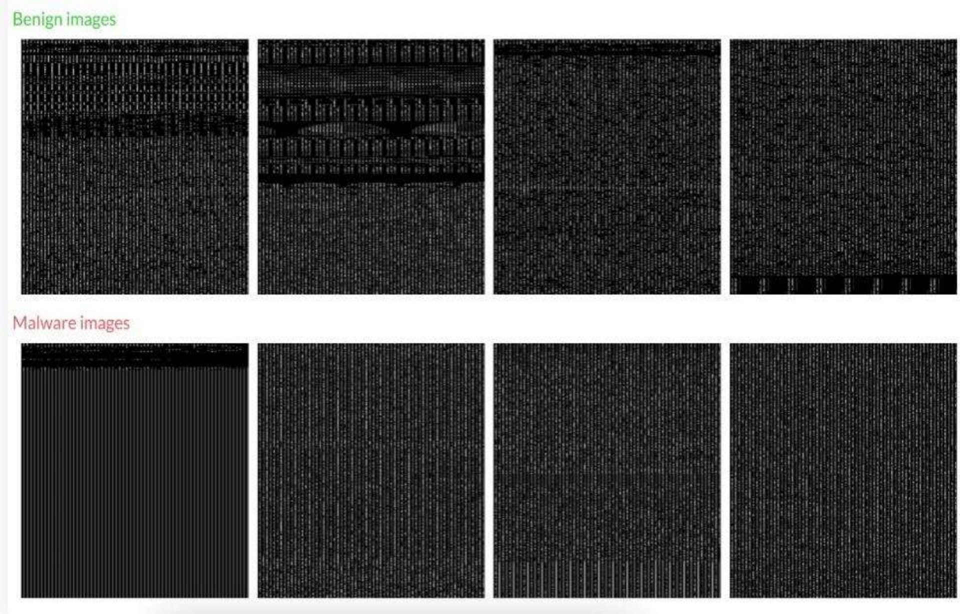


Fig. 4. Benign and malware images of test data

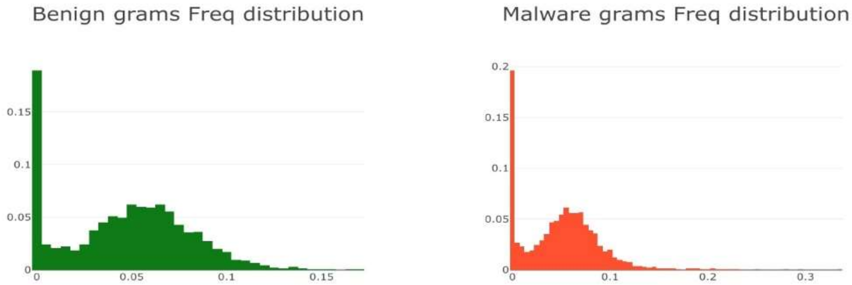
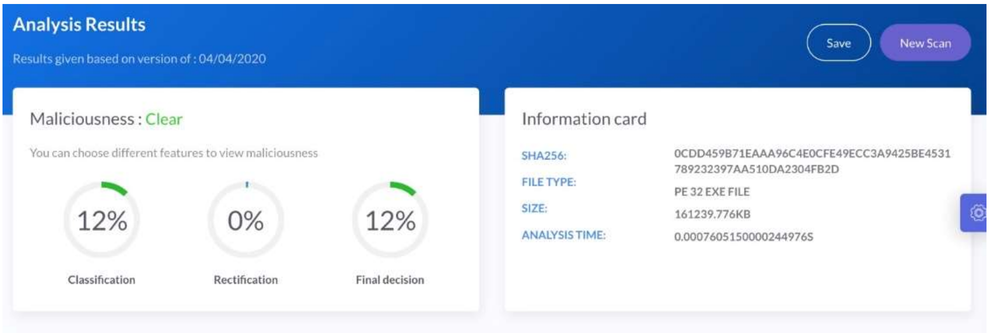


Fig. 5. Benign and malware frequency distribution



Analysis Results

Results given based on version of : 04/04/2020

Maliciousness : Clear

You can choose different features to view maliciousness

12% Classification

0% Rectification

12% Final decision

Information card

SHA256: 0CDD459B71EAAA96C4E0CFE49ECC3A9425BE4531 789232397AA510DA2304FB2D

FILE TYPE: PE 32 EXE FILE

SIZE: 161239.776KB

ANALYSIS TIME: 0.0007605150000244976S

Fig. 6. Maliciousness Result of the test data along with the test data information card

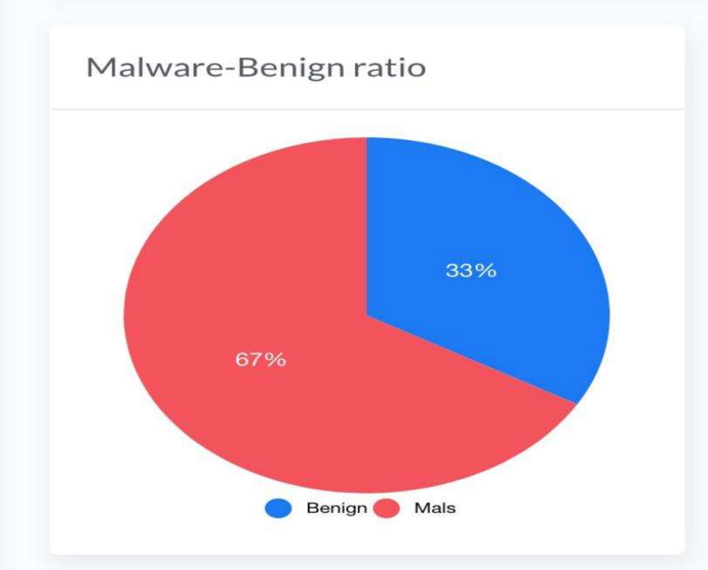


Fig. 7. Final result showing the Malware and Benign ratio of the test data

Discussion: The proposed approach combines both methods to provide a more robust and effective solution to malware detection. The system stores the signatures of suspected malware files in the blockchain and shares them among the users of the peer-to-peer blockchain network, enabling faster and more efficient detection of new malware files. The behavioral-based method analyzes the behavior of the file in a separate virtualized environment, reducing the risk of false positives.

The use of blockchain technology enhances the security and transparency of the system, making it more difficult for attackers to manipulate or compromise the detection process. The decentralized and distributed nature of the blockchain network ensures that the system is resilient to attacks and provides a tamper-proof record of all transactions. However, the proposed approach may have some limitations. For example, the system may face scalability issues due to a large amount of data that needs to be stored in the blockchain. Additionally, the system may require a significant amount of computing resources to analyze the behavior of files in a separate virtualized environment, which may impact the performance of the system.

6 Conclusion

In conclusion, the proposed system that incorporates blockchain technology and Hybrid Analysis is a promising solution to improve the accuracy and effectiveness of malware detection. By using blockchain, the system can store a vast number of malware signatures and involve community participation, making it more decentralized and transparent. Furthermore, the Hybrid Analysis process enables faster detection of malware by running the analysis in a virtualized environment. The proposed system has several advantages over traditional antivirus software, such as improved accuracy, faster detection, a community-driven approach, and improved privacy.

The proposed system can be further enhanced by incorporating scanning, analysis, and identifying patterns in malware behavior using artificial intelligence (AI) and machine learning (ML) algorithms. AI and ML algorithms can assist in identifying new threats and detecting sophisticated malware that traditional antivirus software cannot detect. Additionally, integrating a real-time monitoring system can further improve the detection process by continuously monitoring the system for potential malware threats. Finally, adding a feature to automatically update malware signatures in the blockchain can further enhance the system's accuracy and efficiency.

References

1. J. Gu, B. Sun, X. Du, J. Wang, Y. Zhuang, Z. Wang, Consortium blockchain-based malware detection in mobile devices, IEEE Access. 6 12118–12128 (2018).
2. R. Kumar, X. Zhang, W. Wang, R.U. Khan, J. Kumar, A. Sharif, A multimodal malware detection technique for Android IOT devices using various features, IEEE Access. 7 64411–64430 (2019).
3. A.S. Alotaibi, Biserial Miyaguchi–preneel blockchain-based Ruzicka-indexed deep perceptive learning for malware detection in IOMT, Sensors. 21 7119 (2021).
4. R. Punithavathi, K. Venkatachalam, M. Masud, M. A. AlZain, M. Abouhawwash, Crypto hash based malware detection in IOMT framework, Intelligent Automation & Soft Computing. 34 559–574 (2022).
5. S. Raje, S. Vaderia, N. Wilson, R. Panigrahi, Decentralised firewall for malware detection, 2017 International Conference on Advances in Computing, Communication and Control (ICAC3). (2017).
6. M.S. Rana, C. Gudla, A.H. Sung, Evaluating machine learning models on the Ethereum Blockchain for Android Malware detection, Advances in Intelligent Systems and Computing. 446–461 (2019).
7. N. Anita., M. Vijayalakshmi., Blockchain security attack: A brief survey, 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). (2019).
8. R. Fuji, S. Usuzaki, K. Aburada, H. Yamaba, T. Katayama, M. Park, et al., Blockchain-based malware detection method using shared signatures of suspected malware files, Advances in Networked-Based Information Systems. 305–316 (2019).
9. S. Homayoun, A. Dehghantanha, R.M. Parizi, K.-K.R. Choo, A blockchain-based framework for detecting malicious mobile applications in App Stores, 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). (2019).
10. J. Moubarak, M. Chamoun, E. Filiol, Developing a K-ary malware using blockchain, NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium. (2018).
11. S. Saad, W. Briguglio, H. Elmiligi, The curious case of machine learning in malware detection, Proceedings of the 5th International Conference on Information Systems Security and Privacy. (2019).