GJMRR

Global Journal of Multidisciplinary Research and Reviews (GJMRR)

Open-Access, Peer-Reviewed, Double Blind, Multidisciplinary Online Journal

Volume 1, Issue 2, February 2026

MACHINE LEARNING APPROACHES TO DETECT SYBIL ATTACKS AND NETWORK MANIPULATION IN BLOCKCHAIN

Priyesha Dnyaneshwar Gharat[1] and Dr. Lalit Kumar Khatri[2]

[1]Research Scholar, Department of Information Technology

[2]Professor, Department of Information Technology

Sunrise University Alwar (Raj.) India

Abstract

Blockchain networks are susceptible to Sybil attacks and other network manipulations, which threaten consensus and integrity. Traditional security mechanisms alone are insufficient for detecting such attacks. This review examines machine learning (ML) techniques for identifying malicious nodes and anomalous behaviors in blockchain networks. Both supervised and unsupervised approaches, including hybrid and graph-based methods, are discussed. A comparative summary highlights datasets, methodologies, and detection performance, providing insights into research gaps and future directions.

Keywords: Blockchain security, Sybil attack detection, Network manipulation.

I. INTRODUCTION

Blockchain technology is the foundation of decentralized applications, cryptocurrencies, and distributed finance. Its security relies on majority honest participation and the immutability of transaction ledgers. However, blockchain networks are vulnerable to attacks such as:

1. **Sybil attacks**: where a single adversary creates multiple fake identities to influence consensus or manipulate votes.
2. **Eclipse attacks**: isolating nodes to control their perception of the network.
3. **Transaction flooding or spamming**: overwhelming network resources.

Machine learning provides powerful tools for detecting these attacks by analyzing node behaviors, network topology, and transactional patterns. Unlike traditional cryptographic defenses, ML can adaptively identify anomalies and detect previously unseen attack strategies.

II. EXISTING STUDIES

Several studies have explored ML-based approaches to Sybil detection in blockchain:

1. **Graph-based detection**: Chen et al. (2020) used network metrics such as node degree and connectivity to detect Sybil nodes in Bitcoin.
2. **Neural networks and autoencoders**: Zhang et al. (2022) applied autoencoders to capture abnormal transaction patterns in Ethereum.
3. **Graph Neural Networks (GNNs)**: Li et al. (2021) modeled Ethereum transaction networks to detect clusters of malicious nodes.
4. **Support Vector Machines (SVM)**: Kumar et al. (2019) demonstrated SVM efficiency in identifying anomalous nodes in peer-to-peer networks.
5. **Semi-supervised learning**: Singh et al. (2021) leveraged partial labels to detect Sybil attacks without requiring fully labeled datasets.

These studies indicate that both network topology features and transaction behavior analysis are critical for effective detection.

III. BLOCKCHAIN DATA ANALYSIS AND MACHINE LEARNING APPROACHES

Blockchain networks generate massive volumes of data every second, including transactional records, node interactions, consensus updates, and network topology changes. Analyzing this data to detect security threats such as Sybil attacks, double-spending, and other forms of network manipulation requires specialized Blockchain Data Analysis (BDA) techniques combined with Machine Learning (ML) approaches. BDA involves extracting meaningful features from blockchain networks, such as the number of transactions per node, the frequency and timing of interactions, connectivity patterns, node centrality, transaction amounts, and propagation delays. These features provide insight into normal and anomalous behaviors across the network. Machine learning techniques can then use these features to identify suspicious activity that may indicate malicious nodes or coordinated attacks.

One of the primary advantages of ML is its ability to adaptively learn patterns from historical and real-time data, enabling it to detect previously unseen attacks. Supervised learning algorithms, such as Random Forests, Support Vector Machines (SVMs), and deep neural networks, are commonly employed when labeled datasets are available. In supervised learning, the model is trained on a dataset containing examples of both legitimate nodes and Sybil nodes or other malicious actors, allowing it to learn the distinguishing characteristics of malicious behavior.

Random Forests are particularly effective because they can handle high-dimensional feature spaces and capture complex interactions among features, while SVMs are efficient for detecting outliers in transactional and network patterns. Deep learning models, including multi-layer neural networks, can automatically extract nonlinear relationships and subtle behavioral patterns that are difficult to capture through traditional statistical methods.

In situations where labeled data is scarce or unavailable, unsupervised learning techniques become essential. Unsupervised algorithms, such as k-Means clustering, DBSCAN, and autoencoders, rely on the assumption that normal node behavior forms dense clusters in feature space, while malicious nodes appear as outliers. For example, autoencoders can reconstruct input features of legitimate nodes with low error, but when presented with anomalous or malicious nodes, the reconstruction error increases, allowing detection. Clustering methods group nodes with similar connectivity patterns or transaction behaviors and flag nodes that do not conform to typical patterns as potential Sybil or malicious nodes. Graph-based unsupervised methods leverage the network structure itself, analyzing the graph of node interactions to identify nodes that form abnormal clusters, maintain unusually high or low connectivity, or attempt to manipulate consensus mechanisms.

A particularly promising approach combines graph-based analysis with advanced machine learning, such as Graph Neural Networks (GNNs). Blockchain networks can naturally be represented as graphs, where nodes correspond to users or miners and edges represent transactions or communication links. GNNs can capture both local and global structural information in the graph, allowing them to identify patterns of collusion, abnormal connectivity, or rapid creation of new identities indicative of Sybil attacks. By analyzing both the node features and their relational information, GNNs provide a powerful framework for detecting coordinated attacks that might not be visible when considering nodes individually. Semi-supervised learning can also be applied in graph contexts, where a small subset of labeled nodes propagates information across the network to classify unlabeled nodes, reducing the dependency on fully labeled datasets. Feature engineering plays a critical role in the effectiveness of ML-based BDA. Features derived from transaction behavior include the number of incoming and outgoing transactions, average transaction amounts, transaction frequency, and timing intervals.

Features derived from network topology include node degree, clustering coefficients, betweenness centrality, closeness centrality, and the number of distinct peer connections. Temporal features, such as sudden spikes in transaction volume or bursts of new node creation, can also indicate attacks. The combination of these features allows ML models to differentiate between legitimate network growth and malicious attempts to manipulate the blockchain. In addition to detecting attacks, BDA can provide insights into overall network health, efficiency, and resilience by analyzing patterns of connectivity, consensus participation, and transaction propagation.

Despite their potential, ML approaches face challenges in the blockchain context. The decentralized and pseudonymous nature of blockchain networks makes it difficult to obtain ground truth for training supervised models. Attackers can also adapt their strategies over time, requiring models that are adaptive and capable of learning in real-time. Moreover, the scale of blockchain networks, particularly in cryptocurrencies like Bitcoin or Ethereum, demands computationally efficient algorithms that can handle millions of transactions and thousands of nodes without significant latency. False positives are another concern; misclassifying legitimate nodes as malicious can reduce trust in the network and negatively impact its functionality. Therefore, researchers have explored hybrid approaches that combine ML with rule-based heuristics, consensus validation, or cryptographic verification to enhance accuracy and reduce errors.

Several real-world applications demonstrate the effectiveness of BDA combined with ML. In Bitcoin networks, researchers have successfully used graph-based metrics and supervised learning to identify clusters of malicious nodes attempting double-spending attacks. Ethereum networks have been analyzed with GNNs and autoencoders to detect unusual transaction patterns, including bot-based token manipulations and coordinated Sybil attacks on decentralized applications. Simulation environments and testnets have also been used to generate synthetic datasets with injected attacks, allowing models to be trained and evaluated in controlled settings. These studies show that ML-based BDA can achieve detection accuracies exceeding 90% in many cases, particularly when combining multiple feature types and leveraging relational information. Looking forward, integrating reinforcement learning into blockchain security frameworks offers promising avenues. Reinforcement learning agents could monitor network behavior continuously and learn optimal detection and mitigation strategies dynamically.

Additionally, lightweight ML models optimized for deployment on resource-constrained blockchain nodes could enhance real-time detection without compromising network performance. Standardizing benchmark datasets for Sybil and network manipulation attacks would facilitate better comparisons of ML approaches and accelerate research progress. Combining ML with cryptographic and consensus-layer defenses could also create multi-layered security frameworks capable of addressing both known and novel attack vectors.

Blockchain Data Analysis (BDA) combined with machine learning approaches represents a powerful strategy for detecting Sybil attacks and other network manipulations. By extracting meaningful features from transactions and network topologies, and applying supervised, unsupervised, or hybrid ML models, it is possible to identify malicious behavior effectively. Graph-based methods, particularly GNNs, enhance the ability to detect coordinated attacks, while semi-supervised approaches reduce reliance on labeled data. Despite challenges such as data scarcity, adaptive adversaries, and computational constraints, ongoing research and development continue to advance the field, offering scalable, accurate, and adaptive solutions for blockchain network security. The integration of BDA and ML not only improves attack detection but also provides insights into network health and resilience, ultimately contributing to the robustness of decentralized systems.

IV. SUPERVISED LEARNING

Supervised algorithms require labeled datasets to classify nodes or transactions:

1. **Random Forest (RF)**: Handles high-dimensional features and captures complex patterns.
2. **Support Vector Machines (SVM)**: Efficient in separating benign and malicious behaviors.
3. **Deep Neural Networks (DNN)**: Learn complex nonlinear relationships in transactional and network data.

V. UNSUPERVISED LEARNING

Unsupervised approaches detect anomalies without labeled data:

1. **k-Means Clustering**: Groups nodes based on similarity metrics; outliers indicate potential Sybil nodes.
2. **Autoencoders**: Capture deviations in transactional patterns using reconstruction errors.

3. **Graph-based anomaly detection**: Exploits network connectivity and transaction flows to identify suspicious nodes.

VI. HYBRID AND GRAPH-BASED APPROACHES

1. **Graph Neural Networks (GNNs)**: Model node interactions and detect Sybil clusters.
2. **Semi-supervised learning**: Uses limited labeled data combined with network features for robust detection.

VII. BENEFITS, LIMITATIONS, AND ANALYSIS

1. **Adaptive detection**: ML can detect new attack patterns without explicit rule-based programming.
2. **Scalable analysis**: Graph-based and clustering methods handle large networks.
3. **Integration potential**: ML techniques can complement cryptographic and consensus-based defenses.
4. **Data labeling**: Supervised models require labeled datasets, often scarce in blockchain networks.
5. **Dynamic adversaries**: Attackers continuously change strategies, requiring adaptive ML models.
6. **False positives**: Misclassification of legitimate nodes can reduce network trust.

VIII. ANALYSIS

1. Graph-based and neural network approaches generally achieve higher detection accuracy.
2. Semi-supervised methods reduce dependency on labeled data while maintaining performance.
3. Feature selection, including node degree, transaction frequency, and connectivity, strongly influences detection effectiveness.

Comparative Summary Table

Study	ML Approach	Dataset	Features	Accuracy	Key Contribution
Chen et al., 2020	Random Forest	Bitcoin	Node degree, transaction count	92%	Graph-based Sybil detection

Li et al., 2021	GNN	Ethereum	Transaction graph, temporal patterns	95%	Detected Sybil clusters using relational info
Kumar et al., 2019	SVM	Simulated P2P	Node connectivity, message propagation	88%	Low-cost classification of malicious nodes
Zhang et al., 2022	Autoencoder	Ethereum	Transaction amount, frequency	90%	Anomaly detection via reconstruction error
Singh et al., 2021	Semi-supervised	Blockchain testnet	Node interactions, consensus behavior	93%	Reduced labeled data requirements

IX. FUTURE RESEARCH DIRECTIONS

1. Reinforcement learning for adaptive and real-time detection.
2. Integration of ML with cryptographic methods for stronger security.
3. Benchmark datasets for Sybil and network manipulation attacks.
4. Lightweight ML models for resource-constrained nodes.

X. CONCLUSION

Machine learning techniques offer promising solutions for detecting Sybil attacks and network manipulations in blockchain networks. Supervised, unsupervised, and hybrid approaches can model node behavior and transaction patterns effectively. Future research should focus on real-time adaptability, scalability, and integration with existing blockchain security mechanisms to enhance decentralized network resilience.

REFERENCES

1. Chen, X., Li, Y., & Zhang, W. (2020). Detecting Sybil attacks in Bitcoin using graph-based features. *IEEE Access*, 8, 155213–155223.

2. Chen, Y., Guo, S., & Li, P. (2021). Machine learning-based detection of network manipulation in blockchain-based decentralized applications. *IEEE Access, 9*, 122456–122468.
3. Kumar, R., Singh, A., & Gupta, S. (2019). SVM-based approach for detecting Sybil nodes in peer-to-peer networks. *Journal of Network and Computer Applications*, 137, 35–45.
4. Li, H., Wang, Y., & Xu, J. (2022). Detecting malicious nodes in blockchain networks using ensemble learning. *Future Internet*, *14*(3), 78.
5. Li, P., Zhao, H., & Wang, J. (2021). Graph neural networks for Sybil detection in Ethereum networks. *Computers & Security*, 104, 102205.
6. Patel, R., & Bhattacharya, S. (2021). Graph-based semi-supervised learning for Sybil attack detection in permissionless blockchains. *Applied Soft Computing*, *107*, 107387.
7. Singh, P., Rathi, A., & Sharma, V. (2021). Semi-supervised learning for Sybil attack detection in blockchain networks. *Journal of Information Security and Applications*, 59, 102833.
8. Wang, Y., Sun, H., & Li, F. (2020). Machine learning for Sybil attack detection in decentralized networks: A survey. *IEEE Communications Surveys & Tutorials*, *22*(4), 2605–2628.
9. Yu, J., Zhang, Q., & Chen, X. (2019). Anomaly detection in blockchain systems using deep learning. *IEEE Transactions on Network and Service Management*, *16*(4), 1485–1497.
10. Zhang, T., Liu, K., & Chen, L. (2022). Autoencoder-based anomaly detection in blockchain transactions. *Future Generation Computer Systems*, 133, 217–228.