

Anomaly Detection of DDoS Attack in Bitcoin Transactions using CNN-LSTM-Hybrid Model

Santhosh Narayanan R¹, S. Devi², Renuga S³, Santosh R⁴ and Aditiya S⁵

¹⁻⁵B. Tech Information Technology, Coimbatore Institute of Technology, Tamil Nadu

Email: santhoshnarayanan36@gmail.com, devi.s@cit.edu.in, 71762207041@cit.edu.in, 71762207042@cit.edu.in, 71762207002@cit.edu.in

Abstract— Bitcoin, the initial use of blockchain technology, revolutionized online finance by providing a decentralized, open, and secure method of processing transactions. Blockchain's innovative architecture has influenced many industries by providing trustless data management systems and eliminating central authorities. While blockchain networks like Bitcoin have become ubiquitous and mainstream, they have also become more attractive targets for more sophisticated cyber attacks as well. Generalized attacks include 51% attacks, Sybil attacks, Distributed Denial of Service (DDoS) attacks, selfish mining, eclipse attacks, and time jacking. These attacks can seriously erode the stability of a network, data integrity, and user trust. DDoS attacks, in particular, aim at overloading the Bitcoin network with a staggering amount of malicious requests, thereby slowing down operations and genuine transactions. To enable identification and countermeasures for such threats effectively, deep learning techniques are under exploration. Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and combinations of CNN- LSTM models are promising solutions. CNNs are efficient at extracting spatial patterns, while LSTMs are efficient at handling sequential data. The CNN-LSTM model combines the two when they are applied in tandem to yield high accuracy real-time DDoS detection. Hybrid mode enhances responses and increases the ability to detect threats in complex blockchain systems like Bitcoin.

Keywords— Deep Learning, Bitcoin Transactions, DDoS Attack Detection, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Hybrid CNN-LSTM Model.

I. INTRODUCTION

A. Blockchain

Blockchain is a secure and decentralized technology for data storage that functions as a public ledger, documenting transactions in an ongoing chain of blocks. Blockchain, first proposed in 2008 and implemented in 2009 as the foundational technology for Bitcoin, allows for peer-to-peer transactions without the need for intermediaries. For security, it employs asymmetric cryptography and utilizes distributed consensus algorithms to ensure ledger consistency. The key characteristics of blockchain technology are decentralization, immutability, and anonymity. These traits render it suitable for use in a variety of areas, including digital payments, smart contracts, public services, IoT, and security services. [1]

B. Bitcoin

Developed in 2009, Bitcoin was the first cryptocurrency. It uses blockchain technology and was intended to be independent of all governments and central banks. As of September 2021, there were about 18.8 million Bitcoin tokens in use, out of a maximum of 21 million.[2]

C. DDOS Attack

Distributed Denial of Service attack is when someone knowingly attempts to halt the normal functioning of a server, service, or network by sending excess internet traffic to it. Lots of requests are sent in large numbers through multiple infected computers at the same time (the computers being a botnet most of the time) to a particular server so that this happens. The server cannot process legitimate requests since the heavy utilization of resources results in a denial of service to legitimate users. The attack employs compromised computers remotely controlled by an attacker (a Bot-master) to direct the traffic, which is difficult to detect as an attack source.[5][6]

D. DDoS Attacks in Bitcoin Network

A Bitcoin DDoS attack on the network uses botnets to send traffic requests to nodes, which implies:

- Interference or delay in propagating and authenticating transactions.
- Interference with the Bitcoin consensus, and mining procedure.
- Increased transaction fees, decreased scalability, and decreased network stability.

E. DDoS Attacks in Bitcoin Network

A Bitcoin DDoS attack on the network uses botnets to send traffic requests to nodes, which implies:

- Interference or delay in propagating and authenticating transactions.
- Interference with the Bitcoin consensus, and mining procedure.
- Increased transaction fees, decreased scalability, and decreased network stability.

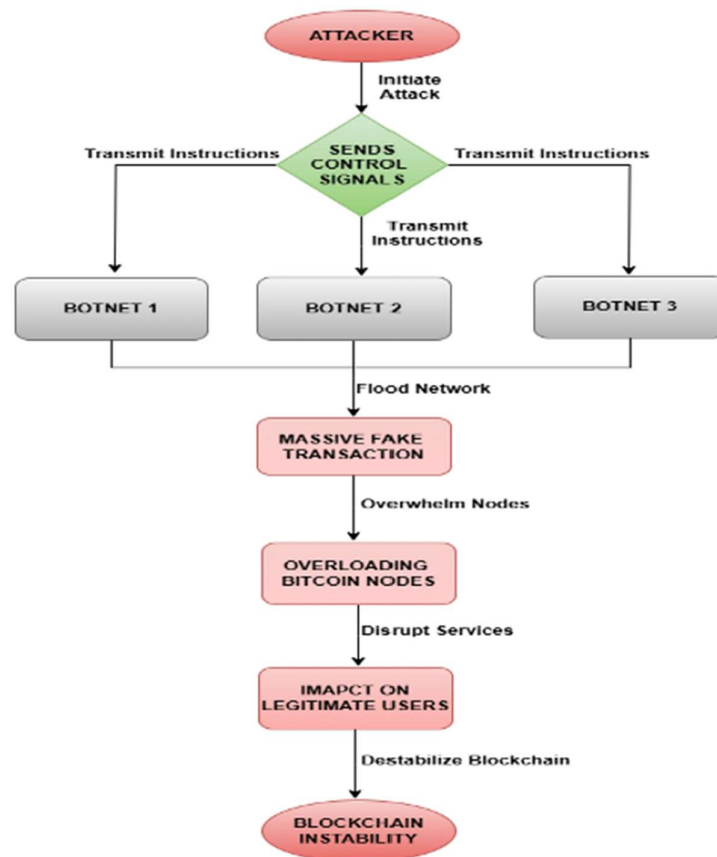


Fig 1: DDOS Attack in Bitcoin Network

Figure 1 illustrates, Distributed Denial-of-Service (DDoS) attack on the Bitcoin protocol is an intentional attempt to clog nodes and miners with an avalanche of fake transactions or network requests, thus causing monstrous traffic, unnecessary fees, and systemic instability. The attack is implemented by the perpetrator by transmitting control messages to an enormous army of hijacked computers (botnets or zombie nodes), which does so unaware. The directive forces the botnets to start making and transmitting large volumes of spam or unnecessary transactions, clogging up the Bitcoin network. Average users cannot handle transactions efficiently owing to mem pool congestion, CPU and bandwidth use, and slowdown in transaction propagation. As the prolonged attack overwhelms Bitcoin nodes with the intense load of transaction requests, it results in lagged block confirmation time, increased transaction fees owing to congestion, and overall network lag. Long-term network congestion could lead to blockchain instability, which further may trigger attacks like forks, orphan blocks, and even double-spending attacks. A Bitcoin DDoS attack ultimately decays its basic operations, affecting nodes, users, and miners, undermining the security and integrity of the network. Efficient countermeasures like peer-to-peer filtering, mem pool optimization, node rate-limiting, and transaction prioritization are essential in managing such attacks and maintaining the immunity of the Bitcoin blockchain.

III. RELATED WORKS

Various models and algorithms, each using different datasets and approaches, have been used to identify and counter attacks in blockchain networks. The Multi Neural Memory Network (MNMN) Classifier [9] effectively limits the scope of Sybil nodes with a success rate of 97% by identifying and classifying them based on simulated blockchain traffic datasets. Likewise, Universal Unique Identifier (UUID) [10] technique tracks frequency of block transmission to identify and classify Sybil nodes with a success rate of 91% on a test blockchain network. TransWalk [11], using One-Class SVM for classification and building directed multi-transaction graph based on transaction-amount and timestamp skewed sampling, has a success rate of 82.78% according to figures from Etherscan. Based on a Bitcoin service dataset augmented with CTGAN, a Hybrid Deep Learning model of RNN and LSTM has been utilized to detect DDoS attacks by doing binary classification with accuracy of 95.84%.

The Edge-Featured Graph Attention Network (EGAT) [12], using Etherscan data, builds star-shaped subgraphs depending on transaction history with attributes like Value, Gas, and Timestamp and classifies them with accuracy of 98.1%. With the datasets like CIC-DDoS2019, NSL-KDD, and CIC-IDS2017, a Contractive Autoencoder [13] was used for the classification of DDoS attacks to separate normal and malicious traffic using reconstruction loss metrics.

The accuracy of the model was 97.58%, 96.08%, and 92.45% with varying datasets. Although its precision is not being specifically called upon, a Random Forest (ML) model [14] that was trained on simulated blockchain data (70 days from SimBlock) and tested on 30 days of simulated and real attack data can identify anomalies based on a variety of features like orphan block rate, block times, revenue-per-hour, and forks. Static verification of smart contract EVM bytecode is achieved through symbolic execution with the Z3 SMT solver [15], and pre-execution vulnerability detection is achieved through control flow analysis with 100% True Positive Rate and 0% False Positive Rate. With 92.6% accuracy in the training set and 95.3% accuracy in the test set using ϕ -entropy divergence, an ISMOTe [16] used CNN/Bi-LSTM model with multihead attention used on actual blockchain traffic data on Ethereum 2.0 and Hyperledger efficiently identifies and segregates malicious nodes to prevent routing table poisoning by identifying in real-time. Finally, with low overhead and high probability of detecting attacks, a Gossip-based Protocol [17] using blockchain data and empirical internet traces detects bad opinions quickly and possesses an extremely robust resistance to Eclipse attacks.

IV. PROPOSED MODEL

The figure 2 shows the workflow of the CNN- LSTM hybrid model in detecting DDoS attacks on Bitcoin transactions. The workflow starts with the gathering of datasets, which are preceded by preprocessing raw data and putting them into correct forms. This is followed by Feature Selection where the most relevant features are determined, thus increasing model efficiency. The Proposed Model (CNN+ LSTM+ Hybrid) is then built on selected features. Model optimization is performed by Hyperparameter Tuning for improved performance. The data is divided for testing and learning in 80% and 20% ratio respectively for test assurance and learning. Model comparison is performed after the test based on performance metrics. Finally, the output is compared with the output of any other Hybrid, CNN, and LSTM models for test efficiency. A high-efficiency and robust intrusion detection system is assured through such a procedure.

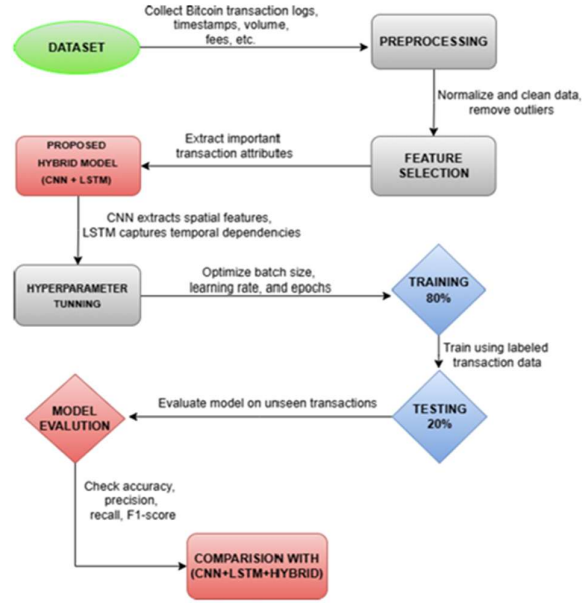


Fig 2: Proposed Model

A. Dataset

Real-world DDoS traffic scrutinized by CICFlowMeter-V3 falls within the realm of the CIC-DDoS2019 dataset. Flows within networks have labels consisting of a timestamp, IP address, a port, a protocol, and attack categories. Reflection- and exploitation-type TCP and UDP attacks are presented. [6][8][14] The dataset further provides derived features like Bitcoin mean input/output amounts, and transaction features like hash values, in- and out- going edges, and Bitcoin flow metrics. The dataset can even be used to study patterns in transactions and discover anomalies in blockchain networks since the dataset contains labels that identify unusual transactions.

B. Feature Correlation

TABLE I: FEATURE CORRELATION OF DATASETS

DDoS Attack Dataset (CICDDoS2019)	Bitcoin Transaction Dataset
Duration of Flow	Total BTC (Duration of Flow Mean)
Total Forward Packets	Indegree (Total Forward Packets)
Total Backward Packets	Outdegree (Total Backward Packets)
Total Length of Forward Packets	In BTC (Total Length of Forward Packets)
Total Length of Backward Packets	Out BTC (Total Length of Backward Packets)
Flow Bytes per second	In BTC (Flow Bytes per second Mean)
Flow Packets per second	Out BTC (Flow Packets per second Mean)

The CICDDoS2019 dataset represents key characteristics of network traffic in opposition to DDoS attacks, and Bitcoin transaction dataset represents data about blockchain transactions. Due to the similarities in the datasets

structure, they may be compared as seen in Table I. "Total Forward Packets" and "Total Backward Packets" of network traffic are being contrasted with "Indegree" and "Outdegree" of Bitcoin transactions showing incoming and outgoing links. Similarly, "Total Length of Forward Packets" and "Total Length of Backward Packets" become "In BTC" and "Out BTC," i.e., transaction direction in Bitcoin. Network traffic "Duration of Flow" becomes "Total BTC," or the total of Bitcoin. Furthermore, "Flow Bytes per second" and "Flow Packets per second" are likewise equal to "Mean In BTC" and "Mean Out BTC," which denote mean flow rates of Bitcoin. This correspondence maps characteristics of DDoS attacks to the space of Bitcoin transactions, in which identifying outliers is simpler and conducting security analysis is possible in a more straightforward manner, as exemplified in Table I.

C. Feature Selection

To enhance the performance of an Intrusion Detection System (IDS), there is a need to identify those features that better discriminate between genuine and malicious traffic. Feature selection is thus an important process. The bar chart in the above page shows feature selection scores as obtained by the Select KBest method, where features are ordered based on statistical significance and calculated using the chi-square (χ^2) test. The equation used in computation of the chi-square score is:

$$\chi^2 = \sum (O - E)^2 / E \quad (1)$$

where O is the observed count of a feature, and E is the expected count. The larger the values of chi-square, the higher the relevance in classification.

V. MODELS

A. CNN (Convolutional Neural Network)

CNNs are mainly employed for the extraction of spatial features from sequential datasets. It utilizes 1D Convolutional layers (Conv1D) to identify local patterns within the input data. MaxPooling1D minimizes dimensionality and focuses on crucial features. Batch Normalization contributes to stabilizing and speeding up the training process. The Flatten Layer transforms the 2D feature maps into a 1D feature vector. [7] The Softmax function helps classify transactions:

$$P(y_i) = e^{(z_i)} / \sum_j e^{(z_j)} \quad (2)$$

where $P(y_i)$ is the Probability of class i , z_i is the CNN layer output, and n is the Total number of classes. Transactions exceeding a set probability threshold are flagged as anomalies

B. LSTM (Long Short-Term Memory)

LSTM networks, a more sophisticated type of RNNs, prove to be efficient in identifying anomalies in Bitcoin transactions impacted by DDoS attacks. LSTMs, in contrast to conventional RNNs, address the vanishing gradient issue through memory cells and gating mechanisms, which allows them to learn long-term dependencies. DDoS attacks inundate the Bitcoin network with an overload of transactions, causing operational disruptions. [7][8] LSTM models identify anomalies through an examination of transaction behavior, frequency, and volume. The Mean Squared Error (MSE) assists in recognizing deviations:

$$MSE = \sum (x_t - \hat{x}_t)^2 / n \quad (3)$$

Where x_t is the Actual transaction value, $\hat{x}_t$ is the predicted transaction value, and n is the Number of transactions. Additionally, LSTM autoencoders use the reconstruction error:

$$RE_t = \|X_t - \hat{X}_t\|^2 \quad (4)$$

Where X_t is the Original sequence, $\hat{X}_t$ is the reconstructed sequence. If $RE_t > \theta$, the transaction is flagged as an anomaly.

C. Hybrid Architecture (CNN-LSTM)

This model is made up of two simultaneous branches: Its local feature extraction process is performed by the CNN branch. The LSTM branch is utilized only to learn long-term dependencies. The outputs of the two branches are concatenated and then passed to the dense layers. Fully connected layers enhance the features already obtained. A sigmoid activation function at the output layer guarantees binary classification. [7]

VI. EXPERIMENTAL RESULTS AND ANALYSIS

A. Performance Metrics

To accurately estimate machine learning and deep learning models reasonably, the right performance measures are to be employed. We essentially utilized precision (P), accuracy (A), recall (R), and F1-score (F1) as performance metrics throughout this paper. Precision is how much a positive prediction is actually correct. Precision is calculated using equation (19), and it will range from 0 to 1.

$$\text{Precision} = TP / (TP + FP) \quad (1)$$

$$\text{Recall} = TP / (TP + FN) \quad (2)$$

$$\text{F1-Score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (3)$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

VII. RESULT

To detect DDoS attacks in Bitcoin transactions, it is essential to analyze key characteristics of network traffic. The Total Length of Forward Packets (19,000) and the Total Length of Backward Packets (14,000) assist in identifying unusual transaction volumes. Flow Bytes (8,000) and Flow Packets (6,000) record excessive data transfers, which are a common indicator of attacks. Unusual traffic patterns are detected by Total Forward Packets (5,000) and Total Backward Packets (2,000). Flow Duration (4,000) assists in identifying sessions that are unusually long or short, which is presented in Table II.

TABLE II: FEATURE CORRELATION OF DATASETS

Features Name	Scores
Flow Duration	4000
Total Forward Packets	5000
Total Backward Packets	2000
Total Length of Forward Packets	19000
Total Length of Backward Packets	14000
Flow Bytes	8000
Flow Packets	6000

The Table-II is the performance of three deep models (LSTM, CNN, and Hybrid CNN-LSTM) to detect DDoS attacks from Bitcoin transactions. For usual transactions (Class 0) and attack transactions (Class 1), each of the three models is considered with accuracy, precision, recall, and F1-score.

A. LSTM Model

Table III suggests that the model stands at 96.24% accuracy level, which proves to be the testimony of fine performance in the identification of temporal relationships in transactional data. Given precision of usual transactions at 0.9440 and precision of attack transactions at 0.9755, it seems that the model is even more confident towards detecting attacks. Normal transaction recall of 0.9651 suggests that although most of the normal transactions are being caught by the model, there can be possibilities of false positives as well. (Refer to Fig. 3)

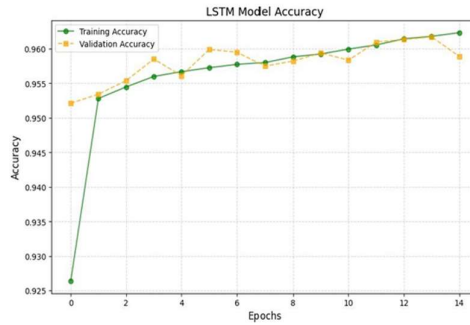


Fig-3 LSTM Model Accuracy

B. CNN Model

With 98.84% accuracy of Table III, CNN surpasses LSTM. CNN performs better classification because it can extract structural and spatial information from transaction data. False alarms are significantly reduced by the CNN, and accuracy is 0.9832 (Class 0) and 0.9919 (Class 1). Since the recall for both classes is significantly higher than 0.98, it is easy to comprehend that the model is catching most of the transactions either as normal or as attacks. The F1-score values (0.9858 and 0.9901) also talk about the capability of the model in identifying the patterns underlying the DDoS anomalies. (Refer to Fig. 4)

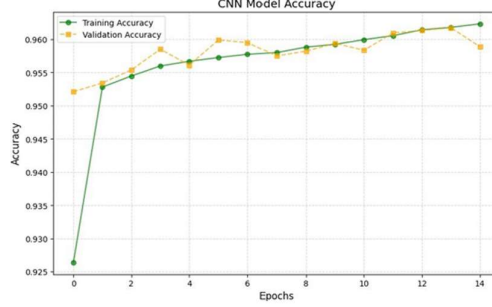


Fig-4 LSTM Model Accuracy

C. Hybrid Model

With a precision of more than 0.988 for each of the two classes, it is ensured that practically no misclassifications are there. With accuracy of 99.23% (Table III), the Hybrid CNN-LSTM model outperforms any one of the separate models. The Hybrid CNN-LSTM model leverages the ability of CNN to learn spatial relationships and LSTM's ability to learn sequential patterns. The recall values for Class 0 and Class 1 as 0.9930 and 0.9917 show an extremely low false negative rate, indicating that almost all the attack transactions are detected correctly. With F1- scores of 0.9905 and 0.9934, the model shows an improved precision-recall trade-off. (Refer to Fig. 5)

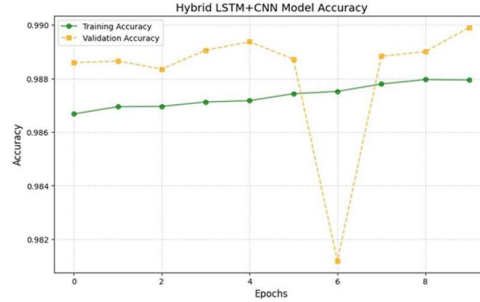


Fig-5 Hybrid Model Accuracy

TABLE III: RESULT

Model	Accuracy	Precision (Class 0)	Recall (Class 0)	F1-Score (Class 0)	Precision (Class 1)	Recall (Class 1)	F1-Score (Class 1)
LSTM	96.24%	0.944033	0.965123	0.954461	0.975510	0.960443	0.967918
CNN	98.84%	0.983228	0.988451	0.985833	0.991987	0.988343	0.990162
Hybrid (CNN + LSTM)	99.23%	0.988153	0.993029	0.990585	0.995164	0.991769	0.993464

VIII. CONCLUSION AND FUTURE DIRECTION

Introduced a deep learning approach for anomaly detection to identify Bitcoin transaction DDoS attacks with LSTM, CNN, and a hybrid (LSTM+CNN) model. The hybrid model performed better than the individual models

with maximum accuracy (99.23%), precision (0.9917), recall (0.9924), and F1-score (0.9920), proving that it has the ability to identify fraud and cyber-attacks. By synergizing LSTM's ability to learn sequence patterns and CNN's power in spatial feature learning, we also dramatically enhanced anomaly detection in blockchain networks. Following work can be concentrated on real-time deployment and scalability, advanced feature engineering with reinforcement learning, and integration of blockchain security features using smart contracts. Additionally, the integration of explainable AI (XAI) for better interpretability, network traffic analysis along with multi-modal data fusion, and expansion of attack detection models like GANs will further improve the efficiency of systems.

REFERENCES

- [1] Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017, June). An overview of blockchain technology: Architecture, consensus, and future trends. In 2017 IEEE international congress on big data (BigData congress) (pp. 557-564). Ieee.
- [2] Davis, J. (2011). The crypto-currency. The New Yorker, 87.
- [3] Akilandeswari, R., & Malathi, S. (2022). Design and implementation of controlling with preventing DDOS attacks using bitcoin by Ethereum block chain technology. *Journal of Transportation Security*, 15(3), 281-297.
- [4] Ali, T. E., Chong, Y. W., & Manickam, S. (2023). Machine learning techniques to detect a DDoS attack in SDN: A systematic review. *Applied Sciences*, 13(5), 3183.
- [5] Tuan, T. A., Long, H. V., Son, L. H., Kumar, R., Priyadarshini, I., & Son, N. T. K. (2020). Performance evaluation of Botnet DDoS attack detection using machine learning. *Evolutionary Intelligence*, 13(2), 283-294.
- [6] Aktar, S., & Nur, A. Y. (2023). Towards DDoS attack detection using deep learning approach. *Computers & Security*, 129, 103251.
- [7] Akhtar, M. S., & Feng, T. (2022). Detection of malware by deep learning as CNN-LSTM machine learning techniques in real time. *Symmetry*, 14(11), 2308.
- [8] Kumar, D., Pateriya, R. K., Gupta, R. K., Dehalwar, V., & Sharma, A. (2023). DDoS detection using deep learning. *Procedia Computer Science*, 218, 2420-2429.
- [9] Kirupanithi, D. N., Antonidoss, A., & Subathra, G. (2023). Automated Classification of The Sybil Attack in BlockChain Network Using Multi Neural Memory Network Classifier. *Appl. Math*, 17(2), 323-336.
- [10] Raghav, N., & Bhola, A. K. Detecting Sybil Attack in Blockchain and Preventing through Universal Unique Identifier in Health Care Sector for privacy preservation.
- [11] Xiong, A., Tong, Y., Jiang, C., Guo, S., Shao, S., Huang, J., ... & Qi, B. (2023). Ethereum phishing detection based on graph neural networks. *IET Blockchain*.
- [12] Zhou, X., Yang, W., & Tian, X. (2023). Detecting phishing accounts on Ethereum based on transaction records and EGAT. *Electronics*, 12(4), 993.
- [13] Aktar, S., & Nur, A. Y. (2023). Towards DDoS attack detection using deep learning approach. *Computers & Security*, 129, 103251.
- [14] Peterson, M., Andel, T., & Benton, R. (2022, March). Towards detection of selfish mining using machine learning. In *International Conference on Cyber Warfare and Security* (Vol. 17, No. 1, pp. 237-243).
- [15] Chinen, Y., Yanai, N., Cruz, J. P., & Okamura, S. (2021). Ra: A static analysis tool for analyzing reentrancy attacks in ethereum smart contracts. *Journal of Information Processing*, 29, 537-547.
- [16] Dai, Q., Zhang, B., & Dong, S. (2022). Eclipse attack detection for blockchain network layer based on deep feature extraction. *Wireless Communications and Mobile Computing*, 2022(1), 1451813.
- [17] Alangot, B., Reijsbergen, D., Venugopalan, S., Szalachowski, P., & Yeo, K. S. (2021). Decentralized and lightweight approach to detect eclipse attacks on proof of work blockchains. *IEEE Transactions on Network and Service Management*, 18(2), 1659-1672.