DeCrypt: Revealing Anonymized Transactions

# Niranjan Ambi

*KIT’s College of Engineering (Autonomous)*

Kolhapur, Maharashtra, India [niranjan.kitcoek@gmail.com](mailto:niranjan.kitcoek@gmail.com)

# Srushti Shete

Tirth Kaktar

*KIT’s College of Engineering (Autonomous)*

Kolhapur, Maharashtra, India [tirthkatkar17@gmail.com](mailto:tirthkatkar17@gmail.com)

# Sakshi Patil

*KIT’s College of Engineering (Autonomous)*

Kolhapur, Maharashtra, India [sakshirpatil.24@gmail.com](mailto:sakshirpatil.24@gmail.com)

# Uma Gurav

*KIT’s College of Engineering (Autonomous)*

Kolhapur, Maharashtra, India [srushtidshete28@gmail.com](mailto:srushtidshete28@gmail.com)

*KIT’s College of Engineering (Autonomous)*

Kolhapur, Maharashtra, India [gurav.uma@kitcoek.in](mailto:gurav.uma@kitcoek.in)

***Abstract*—While cryptocurrencies like Bitcoin were initially designed to offer the user maximum anonymity, transparency of blockchain transactions and immutability of said transactions have made deanonymization-an active area of research-a process of identifying individuals behind pseudonymous addresses. This paper looks at methodologies to deanonymize cryptocurrency transactions, focusing on transaction graph analysis, address clustering, IP tracking, and inference techniques based on machine learning. This is in line with the research conducted on the effectiveness of privacy- augmenting technologies, like the mixing services offered by coins in countering this process. Deanonymization is useful for compliance in terms of AML and regulations as well as conducting crime investigations but results in severe privacy issues. This paper, therefore, aims to critically review the inherent trade-offs between privacy and security, shedding light on the legal and ethical implications, as well as the technological implications, of deanonymizing blockchain transactions in the broader context of financial innovation and regulation. The study will outline some areas that weigh up privacy with security, and the legal and ethical implications of deanonymizing blockchain transactions.**

***Index Terms*—YOLOv8, Image Segmentation, OCR, Number Plate, ResNet-50.**

1. Introduction

Cryptocurrency has undoubtedly revolutionized the world of finance by introducing a decentralized and safe method for transferring value anywhere in the world. However, due to its anonymous nature and lack of central oversight, cryptocur- rency has become the prime target for cybercrime and illegal activities. Crimes related to money laundering, ransomware, and other illicit transactions in dark web marketplaces existed because of the existence of cryptocurrency. All these crimes called for mechanisms that will discover as well as prevent fraud in the cryptographic networks.

Though these techniques are very effective in the classi- cal financial systems, they pose enormous challenges when applied to cryptocurrencies. For example, most conventional techniques don’t perform well with pseudonymous transac- tions, large and unstructured data volumes, and explosive change in the tactics of criminals. Moreover, most techniques

used by the present methods rely on centralized monitor- ing systems, which is starkly contrasting with decentralized blockchain networks.

Two significant challenges are identified in this issue. First, most cryptocurrency exchanges operate within the law, with illegal activities representing only a small fraction of transac- tions. This results in a problem of sample imbalance, where legitimate transactions vastly outnumber illegal ones. Second, there exists a significant amount of untagged data—many illegal transactions remain undetected, creating a dataset filled with unclassified information. These challenges underscore the complexity of identifying the boundary between legal and illegal transactions

To tackle these issues, a self-supervised model based on mutual information is proposed. This approach addresses the untagged data while improving the detection of illegal activity. By combining mutual information and self-supervised learning, the methodology outlined aims to more effectively distinguish between legitimate and unlawful cryptocurrency transactions

By Section-II, which highlights the Related Works. Section- III represents the Proposed Method, and Section-IV gives an overview of the Results and Discussion. This paper concludes with Section-V, offering the final insights in the Conclusion.

1. Related Work

Deanonymization of cryptocurrency has been an active area of research due to the pseudonymous nature of blockchain transactions. The field focuses on uncovering the identities of users involved in transactions through various analytical techniques. Several methods have been developed and refined over time to achieve this, each leveraging the transparent and traceable characteristics of public blockchain networks like Bitcoin. Key techniques and related works are discussed below:

1. Transaction Graph Analysis: Researchers like Meiklejohn et al. (2013) have used transaction graphs to link addresses, revealing clusters associated with known entities.
2. Address Clustering: Multi-input clustering, as proposed by Androulaki et al. (2013), groups addresses by identifying shared inputs in transactions, effectively linking them to indi- vidual users.
3. Temporal and Spatial Analysis: Spagnuolo et al. (2014) used temporal patterns to identify recurring behaviors, signif- icantly reducing anonymity.
4. Network Layer Deanonymization: Biryukov et al. (2014) demonstrated how tracing IP addresses within the Bitcoin network can associate transactions with specific users.
5. Heuristic Techniques: Heuristics like change address iden- tification, employed by the Fistful of Bitcoins project (2013), have been effective in deanonymizing users through behavioral patterns.
6. Machine Learning: Jourdan et al. (2018) applied machine learning models to classify transactions, significantly improv- ing accuracy in identifying users.
7. Blockchain Analytics Tools: Tools like Chainalysis use a combination of graph analysis, heuristics, and machine learn- ing to track and identify users, with successful applications in law enforcement.
8. Coin Mixing Analysis: Mo¨ser et al. (2017) showed that mixing services, despite offering enhanced privacy, can often be deanonymized through transaction flow analysis.
9. Cross-Service Correlation: Koshy et al. (2014) highlighted how correlating data across exchanges and wallet providers can identify users by tracing fund flows across platforms. I.

TABLE I: Literature Review

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Author** | **Title** | **Methodology** |
| 2023 | Kewei Zhao, Guixin Dong, Dong Bian | Detection of Illegal Trans- actions of Cryptocurrency Based on Mutual Informa- tion [2] | The document addresses the challenge of detecting il- legal cryptocurrency transactions using a novel method that combines Graph Neural Networks (GNN) and mu- tual information. The key steps are 1. Graph Convo- lutional Network (GCN): Cryptocurrency transactions are modeled as a graph, and GCN is used to classify transactions (legal or illegal) by processing relationships between transactions. 2.Self-Supervised Learning: To handle the large amount of unlabeled data, the method generates pseudo-labels through self-supervised learn- ing, allowing the model to learn from both labeled and unlabeled data. 3.Mutual Information-Based Loss Function: A new loss function, incorporating mutual information and cross-entropy, is introduced to address the issue of data imbalance, ensuring more accurate detection of illegal transactions, which are a minority in the dataset |

|  |  |  |  |
| --- | --- | --- | --- |
| 2023 | Aditya Kuppa, Jack Nicholls, Nhien-An Le-Khac | The next phase of identify- ing illicit activity in Bitcoin [3] | The methodology in this paper involves reviewing and analyzing existing techniques for detecting illicit activity in the Bitcoin network, focusing on deanonymization heuristics like multi-input clustering, change address detection, and peeling. It also explores how machine learning (ML) and deep learning (DL) can complement these methods to improve detection. The authors assess the effectiveness of these techniques alongside privacy- enhancing tools and include real-world case studies to demonstrate their application. |
| 2020 | Joshua Ellul, Jonathan Galea, Max Ganado, Stephen Mccarthy, Gordon J. Pace | Regulating Blockchain, DLT and Smart Contracts [4] | The authors review global regulatory approaches to cryptocurrency, blockchain, and DLT, focusing on fi- nancial regulations, AML directives, and cybersecurity measures. Case studies like the Mt. Gox hack and Silk Road incident highlight regulatory challenges. To address these, the authors propose a novel framework from the Malta Digital Innovation Authority (MDIA), featuring independent technology audits, system certifi- cations, and roles such as Technical Administrators and Forensic Nodes to tackle blockchain-specific issue |
| 2018 | Dr Mahdi H. Miraz, Maaruf Ali | Applications of Blockchain Technology beyond Cryp- tocurrency [5] | In this paper, a literature review is conducted to explore blockchain technology applications beyond cryptocur- rency. The authors outline fundamental blockchain con- cepts, including Proof-of-Work and public key systems for privacy. The methodology involves reviewing recent research and case studies to assess blockchain’s use in non-monetary areas such as cloud storage, healthcare, decentralized voting, and the Internet of Things (IoT). |
| 2023 | Jiajun Zhou, Chenkai Hu, Jianlei Chi, Jiajing Wu and Meng Shen | Behavior-aware Account De-anonymization on Ethereum Interaction Graph [6] | The research paper proposes a methodology for deanonymizing Ethereum accounts using Ethident, an end-to-end graph neural network (GNN) framework de- signed for account de-anonymization based on behavior patterns. |
| 2023 | Jack Nicholls, Aditya Kuppa and Nhien-An Le-Khac | The next phase of identify- ing illicit activity in Bitcoin [7] | The research paper proposes the methodology for de- tecting illicit activities in Bitcoin through the analysis of transaction data. The key approach relies on heuristics that attempt to deanonymize users by clustering inputs and outputs, such as the multi-input heuristic and change address detection. These heuristics aim to identify users by grouping addresses under the assumption that inputs in a transaction belong to a single entity. However, these methods struggle when encountering advanced techniques like CoinJoin or Peeling chains, which ob- scure transaction flows by mixing inputs from different users or iteratively sending small amounts to various addresses. |

1. Proposed Method

Our framework consists of three main components: transac- tion pattern analysis, entity resolution, and identity mapping. 1.Transaction Pattern Analysis The Transaction Pattern Analysis stage is crucial for identifying anomalous behavior in

cryptocurrency transactions, which may indicate illicit activi- ties such as money laundering, fraud, or other financial crimes. In this stage, an Isolation Forest-based anomaly detection system is employed to flag transactions or addresses that deviate from typical behavior patterns observed in legitimate

transactions. Key Features Analyzed:Transaction Volume and Frequency

1. Volume: Refers to the amount of cryptocurrency being transferred in each transaction. Abnormalities in transaction volumes, such as unusually large or small transfers compared to typical user behavior, can raise red flags. Example: Large, infrequent transfers may indicate an attempt to launder money, while small, repetitive transfers could point to a ”peeling chain” used to obfuscate the origin of funds.
2. Frequency: The number of transactions made by an ad- dress over a certain period. High-frequency transactions over a short time window may suggest suspicious behavior, such as rapid fund transfers across multiple addresses to evade detection. Example: A high transaction frequency combined with a low transaction volume might signal ”smurfing,” where criminals divide a large sum of money into smaller, less noticeable transfers.

Deposit/Withdrawal Ratios This feature examines the ratio of deposits (incoming funds) to withdrawals (outgoing funds) for each address or entity. Abnormal deposit-to-withdrawal patterns can indicate money laundering or mixing services. Example: A significantly higher withdrawal ratio compared to deposits, especially when funds are frequently transferred to new addresses, could suggest that an address is acting as an intermediary for laundering funds. Monitoring the flow of funds from one address to another, particularly when large withdrawals are split into smaller amounts, can provide in- sights into potential illicit fund movement. Network Centrality Metrics

Centrality measures are used to assess the importance or influence of an address within the broader transaction network. In the context of anomaly detection, high centrality scores might point to addresses that play a central role in a suspicious network of transactions. Betweenness Centrality: Measures how often an address serves as a bridge in the shortest paths between other addresses. High betweenness can indicate that an address is being used to move funds between many other addresses, possibly acting as a mixing service or a laundering hub. Degree Centrality: Represents the number of direct connections an address has. A high degree may indicate that an address is interacting with a disproportionate number of other addresses, which could be a sign of malicious activity. Closeness Centrality: This metric evaluates how close an address is to all other addresses in the network. Anomalies here may show that certain addresses are key nodes in illegal operations or centralized illicit financial structures. Temporal Patterns

Time-based patterns in transactions can offer insights into suspicious behaviors. For example, irregular transaction tim- ings, such as large transfers made at unusual hours or repeated transactions at fixed intervals, can be indicative of automated laundering schemes or bot-controlled addresses. Example: A pattern of transferring funds at the same time each day may suggest automated money laundering activities. Seasonal variations: Changes in transaction behavior over time, such as a spike in activity following known events (e.g., ransomware

attacks), can also indicate potential involvement in illicit activ- ities. Isolation Forest-Based Anomaly Detection An Isolation Forest is particularly well-suited for detecting anomalies in high-dimensional data like cryptocurrency transactions. Unlike other clustering algorithms that aim to group similar points, Isolation Forest explicitly identifies anomalies by isolating them in the data structure. The key idea is that anomalies are few and different, making them easier to isolate than normal observations.

How It Works: The algorithm randomly selects features (e.g., transaction volume, frequency, etc.) and splits the data based on these features. It then constructs a tree structure, where anomalies are isolated closer to the root of the tree, as they require fewer splits to separate them from the rest of the data. Efficiency: Isolation Forests can efficiently handle large- scale datasets and high-dimensional data, making them ideal for detecting suspicious patterns in massive cryptocurrency networks. Benefits: This method is unsupervised, meaning it does not require labeled training data (which is often difficult to obtain for illicit transactions). It also works well in environments where anomalies are rare and have distinct patterns from normal transactions. Application of Features in Detection The key features outlined above (transaction volume, frequency, deposit/withdrawal ratios, network central- ity, and temporal patterns) are fed into the Isolation Forest model, which then assesses each transaction or address for anomalies. Transactions that deviate significantly from the ”normal” patterns learned by the model are flagged for further investigation.

Scoring: Each transaction is given an anomaly score by the Isolation Forest. Higher scores indicate a greater likelihood that the transaction is abnormal and may involve illicit activity.

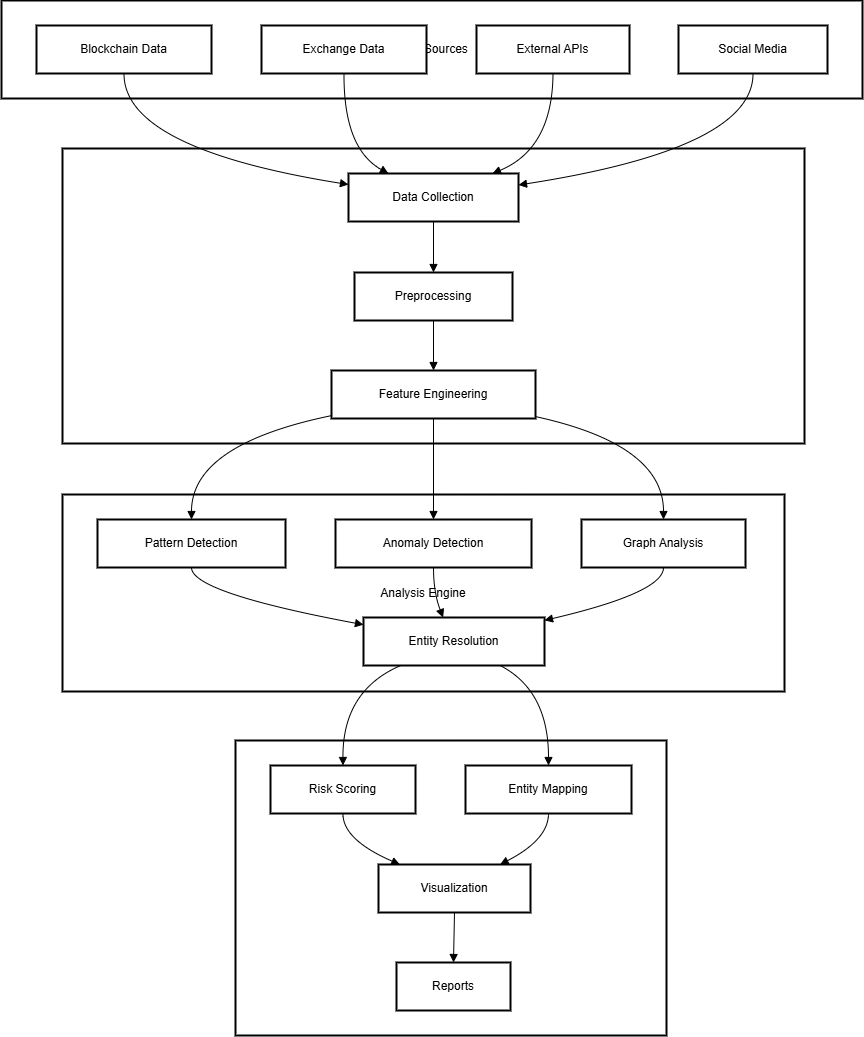


Fig. 1: Block diagram of the proposed work

during nighttime scenarios, contributing to the system’s ro- bustness in 24/7 surveillance environments.

*A. Identity Mapping Techniques*

3.4 Identification Mapping Techniques The identification mapping stage of cryptocurrency deanonymization is done through a number of techniques that combine on-chain transac- tion information with off-chain information to link blockchain addresses to real identities. This involves using collected data from a variety of exchanges, web intelligence sources, and network analysis, which would form a multi-dimensional identity mapping with blockchain transactions.

1. Exchange Data Correlation Withdrawal Pattern Analysis This approach monitors outgoing and incoming flow from known cryptocurrency exchanges through transaction tracking. Analysis of frequent withdrawal patterns, such as constant small or one-time large withdrawals, helps infer user inter- action possibilities, such as users having gone through the KYC process as part of regulatory requirements. Patterns with regular transaction interaction with the exchange accounts are flagged since they would help reveal other user activity streams and maybe provide identity data. KYC Data Matching: Most regulated exchanges have adopted the policy of Know Your Customer, collecting the identity data in order to com- ply with regulatory standards. Researchers who partner with such exchanges can use KYC-related addresses that associate blockchain addresses with identified personal identities. For instance, using hashed data from the exchanges, it can be matched with the addresses of the transactions and, conse- quently, identify direct or indirect links between blockchain activity and the verified persons.
2. Web Intelligence Social media analysis will be possible through public social media profiles containing information related to cryptocurrency wallets available through voluntary actions by users when they share addresses for donations, payments, or reputation. Researchers can find addresses dis- closed by users through mining social media data and link these addresses to their online identities. Forum and Post Correlation: one finds BitcoinTalk or other communities that focus on cryptocurrencies, where wallet addresses are used for payments, discussion of transactions or advertisements of ser- vices. Thus, correlation of such posts would add another layer of identity mapping through user-generated content to user accounts on forums. Public Declarations: There are individuals and groups that publicly declare their cryptocurrency addresses at their websites, blogs, or pages within communities, for soliciting donations or publishing transparent transactions. A researcher could crawl or mine the data and associate declared addresses to certain entities while growing the graph of addressable identities .
3. Network Analysis Co-spending patterns: Co-spending analysis entails examining addresses that are also spent to- gether in transactions. It reflects the common ownership or control whereby several addresses appear as inputs in the same transaction; this is normally controlled by a single entity. It is effectively useful in addressing clustering and

helps group addresses under one user by spending patterns. Temporal correlation: This methodology analyzes the timing of transactions that take place between addresses. It looks to find temporal patterns that might indicate identity, such as consistent and regular time differences between interaction or equivalent delays of identical transactions. User behavior is usually indicated by periodicity or consistency in delays of transaction typically means automatic behavior by one user. Subprofiles concerning user activity assist in profile creation, usually indicating whether the transactions belong to a larger financial activity pattern-for instance, payroll or investment withdrawals. Graph-Based Clustering: Graph-based clustering uses graph theory to form clusters between the addresses based on strong transaction ties. Researchers can visualize the relationship between addresses as a graph and then identify groups of addresses that are likely controlled by the same user. Algorithms, such as Louvain or modularity-based methods, reveal a community within the transaction network that is indicative of coordinated transaction behavior and might be associated with a single identity or entity.

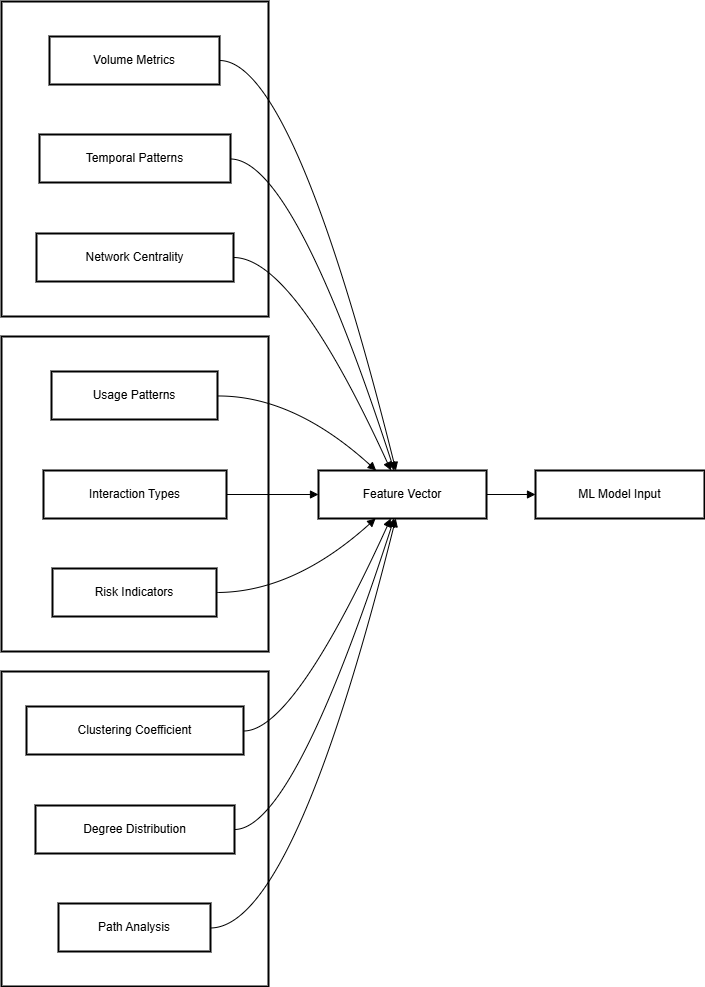


Fig. 2: Block diagram of the proposed work

REFERENCES

1. Miraz, M. H., & Ali, M. (2018). Applications of Blockchain Technology Beyond Cryptocurrency. \*Annals of Emerging Technologies in Computing (AETIC)\*, 2(1).

2. Blossey, G., Eisenhardt, J., & Hahn, G. (2019). Blockchain Technology in Supply Chain Management: An Application Perspective. In \*Proceedings of the 52nd Hawaii International Conference on System Sciences\*.

3. McGhin, T., Choo, K.-K. R., Liu, C. Z., & He, D. (2019). Blockchain in Healthcare Applications: Research Challenges and Opportunities. \*Journal of Network and Computer Applications\*, 135, 62–75.

4. Linoy, S., Stakhanova, N., & Ray, S. (2021). De-Anonymizing Ethereum Blockchain Smart Contracts Through Code Attribution. \*International Journal of Network Management\*, 31(1), e2130.

5. Yuan, Q., Huang, B., Zhang, J., Wu, J., Zhang, H., & Zhang, X. (2020). Detecting Phishing Scams on Ethereum Based on Transaction Records. In \*2020 IEEE International Symposium on Circuits and Systems (ISCAS)\* (pp. 1–5). IEEE.

6. Huang, Y., Wang, H., Wu, L., Tyson, G., Luo, X., Zhang, R., Liu, X., Huang, G., & Jiang, X. (2020). Understanding (Mis)Behavior on the EOSIO Blockchain. \*Proceedings of the ACM on Measurement and Analysis of Computing Systems\*, 4(2), 1–28.

7. Li, Y., Cai, Y., Tian, H., Xue, G., & Zheng, Z. (2020). Identifying Illicit Addresses in Bitcoin Network. In \*International Conference on Blockchain and Trustworthy Systems\* (pp. 99–111). Springer.

8. Linoy, S., Stakhanova, N., & Ray, S. (2021). De-Anonymizing Ethereum Blockchain Smart Contracts Through Code Attribution. \*International Journal of Network Management\*, 31(1), e2130.

9. Wu, J., Yuan, Q., Lin, D., You, W., Chen, W., Chen, C., & Zheng, Z. (2020). Who Are the Phishers? Phishing Scam Detection on Ethereum via Network Embedding. \*IEEE Transactions on Systems, Man, and Cybernetics: Systems\*.

10. Bartoletti, M., Pes, B., & Serusi, S. (2018). Data Mining for Detecting Bitcoin Ponzi Schemes. In \*2018 Crypto Valley Conference on Blockchain Technology (CVCBT)\* (pp. 75–84). IEEE.

11. Chen, L., Peng, J., Liu, Y., Li, J., Xie, F., & Zheng, Z. (2020). Phishing Scams Detection in Ethereum Transaction Network. \*ACM Transactions on Internet Technology (TOIT)\*, 21(1), 1–16.

12. Yuan, Q., Huang, B., Zhang, J., Wu, J., Zhang, H., & Zhang, X. (2020). Detecting Phishing Scams on Ethereum Based on Transaction Records. In \*2020 IEEE International Symposium on Circuits and Systems (ISCAS)\* (pp. 1–5). IEEE.

13. Huang, Y., Wang, H., Wu, L., Tyson, G., Luo, X., Zhang, R., Liu, X., Huang, G., & Jiang, X. (2020). Understanding (Mis)Behavior on the EOSIO Blockchain. \*Proceedings of the ACM on Measurement and Analysis of Computing Systems\*, 4(2), 1–28.

14. Fu, B., Yu, X., & Feng, T. (2022). CT-GCN: A Phishing Identification Model for Blockchain Cryptocurrency Transactions. \*International Journal of Information Security\*, 21, 1223–1232.

15. Huang, T., Lin, D., & Wu, J. (2022). Ethereum Account Classification Based on Graph Convolutional Network. \*IEEE Transactions on Circuits and Systems II: Express Briefs\*, 69, 2528–2532.

16. Cui, W., & Gao, C. (2023). WTEye: On-Chain Wash Trade Detection and Quantification for ERC20 Cryptocurrencies. \*Blockchain Research & Applications\*, 4, 100108.

17. Ghosh, A., Gupta, S., Dua, A., & Kumar, N. (2020). Security of Cryptocurrencies in Blockchain Technology: State-of-the-Art, Challenges, and Future Prospects. \*Journal of Network and Computer Applications\*, 163, 102635.

18. Farrugia, S., Ellul, J., & Azzopardi, G. (2020). Detection of Illicit Accounts Over the Ethereum Blockchain. \*Expert Systems with Applications\*, 150, 113318.

19. Kumar, N., Singh, A., Handa, A., & Shukla, S.K. (2020). Detecting Malicious Accounts on the Ethereum Blockchain with Supervised Learning. In \*Proceedings of the Cyber Security Cryptography and Machine Learning: Fourth International Symposium (CSCML 2020)\* (pp. 94–109). Springer, Berlin/Heidelberg.

20. Gu, Z., Lin, D., & Wu, J. (2022). On-Chain Analysis-Based Detection of Abnormal Transaction Amount on Cryptocurrency Exchanges. \*Physica A: Statistical Mechanics and its Applications\*, 604, 127799.

21. Ammer, M.A., & Aldhyani, T.H. (2022). Deep Learning Algorithm to Predict Cryptocurrency Fluctuation Prices: Increasing Investment Awareness. \*Electronics\*, 11, 2349.

22. Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., & Tang, J. (2021). Self-Supervised Learning: Generative or Contrastive. \*IEEE Transactions on Knowledge and Data Engineering\*, 35, 857–876.

23. Cao, K., Wei, C., Gaidon, A., Arechiga, N., & Ma, T. (2019). Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss. In \*Advances in Neural Information Processing Systems\* (Vol. 32, pp. 1567–1578).

24. Gai, K., Wu, Y., Zhu, L., Zhang, Z., & Qiu, M. (2019). Differential Privacy-Based Blockchain for Industrial Internet-of-Things. \*IEEE Transactions on Industrial Informatics\*, 16, 4156–4165.

25. Reynolds, S. (2022). Crypto.com’s Stolen Ether Being Mixed Through Tornado Cash. Retrieved June 27, 2022, from https://www.coindesk.com/business/2022/01/18/cryptocoms-stolen-etherbeing-laundered-via-tornado-cash/

26. Sharma, R. (2022). Decentralized Finance (DeFi) Definition — Investopia. Retrieved June 30, 2022, from https://www.investopedia.com/decentralized-finance-defi-5113835

27. The Block (2022). Value Locked - Ethereum and Binance Smart Chain - The Block. Retrieved June 27, 2022, from https://www.theblock.co/data/decentralized-finance/total-value-locked-tvl

28. The Law Society (2022). Anti-Money Laundering — The Law Society. Retrieved June 27, 2022, from https://www.lawsociety.org.uk/topics/anti-money-laundering

29. Hudson Intelligence (2022). Peel Chain — Cryptocurrency Investigation - Hudson Intelligence. Retrieved June 27, 2022, from https://www.fraudinvestigation.net/cryptocurrency/tracing/peel-chain