Blockchain and Cryptocurrency, Vol. 3, Issue 1, March 2025, pp. 65-74

IFSA Publishing

Blockchain and Cryptocurrency

Published by IFSA Publishing, S. L., 2025
https://bc-ifsa-journal.com/

Don't Hold Your Breath: An Empirical Dive into Bitcoin Money Laundering Detection

[*] Otilia Maria MUNTEAN and Ciprian PUNGILĂ
West University of Timisoara, Faculty of Mathematics and Informatics,
B-dul Vasile Pârvan 4, Timișoara, 300223, Romania
[*] E-mail: otilia.muntean97@e-uvt.ro, ciprian.pungila@e-uvt.ro

Received: 30 June 2024 *Accepted:* 10 February 2025
Revised: 9 January 2025 *Published:* 24 March 2025

Abstract: Money laundering represents a substantial threat to the integrity of global financial systems. With the advent and rapid advancement of blockchain technology, new opportunities for illicit activities have emerged. The rapid pace of blockchain development has often outstripped the evolution of corresponding regulatory frameworks, creating an environment conductive to financial crimes. Existing anti-money laundering (AML) methods frequently struggle to address the intricate and dynamic characteristics of blockchain transactions, underscoring the need for ongoing reassessment of both technological and legislative strategies. This research presents a comprehensive taxonomy based on an analysis of 150 academic papers published between 2019 and the moment of writing, alongside a comparative evaluation of three existing tools designed to process blockchain data with the purpose of detecting money laundering activities.

Keywords: Blockchain, Bitcoin, Data extraction, Database, Money laundering.

1. Introduction

In the last few years, the prelevance of money laundering has risen substantially, driven by the rapid expansion of financial systems and the emergence of cryptocurrencies. This form of crime poses a serious threat to the stability of economic and social systems while undermining the security of public and private financial institutions.

The ongoing battle against cybercriminals engaging in money laundering requires rigorous analysis of assets and the origins of funds to identify in order to disrupt illegitimate financial flows. The adoption of cryptocurrencies such as Bitcoin has further complicated these efforts, as criminals increasingly exploit advanced technologies and innovate techniques to conceal their illicit activities.

This research aims to provide a comprehensive taxonomy of existing literature focused on money laundering detection within Bitcoin blockchain. After gaining a deeper insight of the literature on this topic, we present a comparison between three existent tools which parse blockchain data and extract it in different ways in order to further use the data for money laundering detection techniques.

2. State of the Art

Although concerns about money laundering through cryptocurrencies have been growing, the amount of literature addressing this issue remains surprisingly limited. The interdisciplinary nature of investigating money laundering within blockchain systems poses a significant barrier to in-depth research and understanding. Conducting effective studies in this field requires expertise spanning computer science, finance, law and criminology. This complexity often

discourages researchers, resulting in fragmented efforts that fail to address the dimensions of the problem comprehensively.

Additionally, the dynamic and ever-evolving nature of blockchain technology presents further challenges. Combating money laundering in this context can be seen as an ongoing adversarial interaction between researchers, developers, investigators and cybercriminals. While regulatory frameworks struggle to keep pace with rapid technological advancements, illicit actors continue to adapt and innovate, developing sophisticated methods to exploit vulnerabilities in the system. This constant evolution hinders the production of timely and relevant academic contributions.

To address this gap, we developed a comprehensive search strategy to systematically review existing academic studies on money laundering detection within the Bitcoin blockchain. This approach ensures the identification and analysis of relevant literature, providing a clearer understanding of current state of research in this context. Furthermore, we recognize the importance of parsing Bitcoin blockchain data to gain a deeper understanding of transactions in relation to users. This step is essential in the process of detecting money laundering, yet we have identified a noticeable gap in available resources and tools to facilitate this process. We identified several distinct approaches in the development of tools for blockchain data processing; however, we faced numerous challenges in making them fully operational. Additionally, we will present a comparison of three blockchain data processing tools, focusing on their data processing speed, which is a critical factor given the immense size of the Bitcoin blockchain dataset.

2.1. Selection Criteria

The aim of the research: We focused on identifying research projects that describe various money laundering detection methods, highlighting their advantages and providing comparative analyses of specific tools and alternatives. Additionally, we included studies that, while not addressing technical approaches, offer detailed legislative or economic perspectives by examining differences in enforced regulations that may inadvertently enable financial cybercrimes. The selection process was conducted with great care to ensure that the chosen papers closely align with our subject of interest.

Source selection: To ensure the credibility and reliability of our references, we prioritized selecting trustworthy sources. Our focus was on peer-reviewed academic papers, which undergo a rigorous review process, published in reputable journals or by well-established institutions such as Elsevier, IEEE, ACM, and Springer. Additionally, we incorporated industry reports from leading companies in this field, leveraging their expertise on current trends and practices. This approach aimed to establish a robust foundation of evidence-based information to underpin our research.

Dataset selection: The existing literature covers a diverse range of datasets, some of which are open-source, while others are acquired by the authors from reliable sources. It is important to note that privacy regulations, such as GDPR in Europe, require data preprocessing, which in some cases may impact data quality. Nonetheless, the most frequently encountered datasets in the reviewed studies are the Bitcoin blockchain real dataset and the Elliptic dataset [1].

Search terms: We conducted our search using the Google Scholar search engine, using keywords such as “bitcoin”, “blockchain”, “money laundering” and “detection”. The research process began in early April 2024 and was continued until late December 2024. We have limited our search to publications between 2019 and 2024, up to the time of writing. The purpose of this study was to develop a taxonomy that is both recent and comprehensive, ensuring its relevance to the current state of the field.

Selection and Disqualification Criteria: We have selected peer-reviewed publications due to their generally rigorous review process. It is important to note that papers written in languages other than English were excluded from the outset for accessibility reasons. Additionally, non-academic resources, such as news articles, blog posts, opinion pieces or non-peer-reviewed publications were also excluded from our consideration.

2.2. The Selection Process

The Fig. 1 is designed to illustrate the selection process we followed to curate a collection of peer-reviewed and relevant projects. We have selected the top 150 articles representing the latest research findings on our topic. Out of those, we have initially excluded 22 papers for not being peer-reviewed, as we place significant value on peer-reviewed papers due to their rigorous evaluation process, which ensures credibility, accuracy and reliability of the research project. This is particularly important in our study to build a strong foundation of trustworthy data and to draw conclusions based on thoroughly validated findings.

After applying the peer-review exclusion criteria, 128 articles remained. From these, we excluded an additional 6 papers due to duplication. The next step in the selection process involved a thorough analysis of each paper, beginning with the title, abstract, and content. Through this process, we excluded another 57 articles that were not relevant to our topic of interest. This left us with a final set of 66 papers, which constitute the core selection for our study.

Out of the 66 papers in the final selection, 47 papers were published in journals and 19 are presented as different conference proceedings. Table 1 shows a better understanding of the publication trend over the last few years, up to the moment of writing.

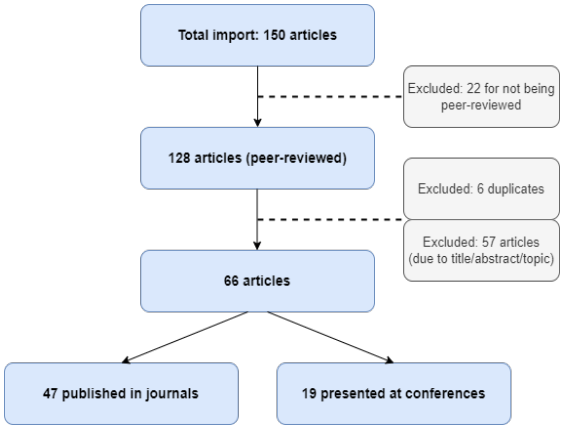

Fig. 1. Selection process.

Table 1. Publication trend.

Year	Journal	Conference	Total
2019	4	4	8
2020	3	4	7
2021	14	5	19
2022	10	1	11
2023	12	3	15
2024	4	2	6

The total number of publications rose sharply from 2019 (8 publications) to 2021, where it peaked at 19. This indicates heightened interest in the topic during this period, likely driven by increased global attention to cryptocurrency-related financial crimes and advancements in blockchain analysis techniques. Journals consistently outperformed conferences in terms of publication numbers, particularly in 2021 (14 journal articles vs. 5 conference papers) and 2023 (12 journal articles vs. 3 conference papers). This trend might reflect a preference for in-depth, peer-reviewed analyses in this domain, which journals typically provide. Another possible reason could be the limited number of academic conferences focused on blockchain, as business-oriented blockchain conferences significantly outnumber academic ones.

While the decline in 2024 is notable, emerging threats and innovations in blockchain technology may reignite interest in this area. Future research could focus on integrating AI-driven tools for detection, exploring cross-chain laundering methods, or addressing regulatory gaps.

3. Literature Review

In our previous paper, which presented a taxonomy based on the analysis of 120 articles [62], we categorized the papers into four broad groups according to the approaches employed: supervised learning, unsupervised learning, hybrid learning, and data mining or data analysis methods.

In this research paper, we have expanded the selection of papers to include 150 studies, as previously mentioned. To gain a deeper understanding of the current state of the literature on our topic of interest, we have categorized them into more specific groups. Fig. 2 highlights the most commonly utilized techniques for detecting money laundering in the Bitcoin blockchain divided into two major categories: technical approaches and non-technical approaches.

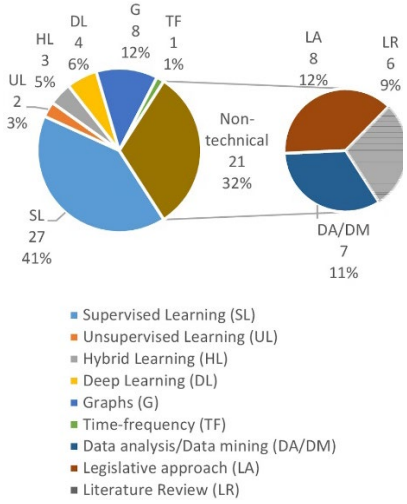

Fig. 2. Domain classification.

The majority of techniques fall into the technical category, with Supervised Learning (SL) being the most frequently employed approach, accounting for 41 % of the studies. This highlights the dominance of machine learning models trained on labeled datasets for detecting suspicious transactions. Graphs (G) follow at 12 %, indicating the importance of network analysis in understanding transactional relationships and identifying anomalies. Other technical methods include Deep Learning (DL) at 6 %, Hybrid Learning (HL) at 5 %, and Unsupervised Learning (UL) at 3 %, which together demonstrate the diversity of machine learning techniques applied in this domain. Time-Frequency (TF) methods, however, are rarely used, making up just 1 % of the studies.

Non-technical approaches make up 32 % of the taxonomy and are categorized into Legislative Approaches (LA) (12 %), Data Analysis/Data Mining (DA/DM) (11 %), and Literature Reviews (LR) (9 %). These methods focus on regulatory frameworks, exploratory analyses, and critical reviews of existing research, underscoring the importance of a broader contextual understanding of money laundering detection beyond purely algorithmic methods.

Table 2 presents a taxonomy on analyzed papers alongside with their reference number for further investigation. This structured representation provides insight into the diversity of methodologies employed in this research area. In both previous illustration and table, we can observe that **supervised learning** approaches are represented by the largest number of references, demonstrating the significant reliance on labeled data to train predictive models for detecting suspicious transactions. It underscores the dominance

of machine learning techniques in tackling this challenge. **Unsupervised learning** methods were highlighted in fewer studies: this approach focuses on identifying patterns or anomalies in unlabeled data, showcasing its potential for detecting new money laundering schemes. **Hybrid learning** combines multiple learning paradigms reflecting the efforts to enhance detection accuracy and model robustness. **Deep learning** approaches highlight the use of advanced neural networks for capturing complex patterns in transactional data, indicating growing interest in leveraging Artificial Intelligence (AI) for blockchain analytics and forensics. The research projects using graphs analyze the structure and relationships within transaction networks, emphasizing the importance of network-based insights in identifying suspicious activity. We have observed a growing interest in graph-based approaches due to their ability to model data as graph structures. This layout simplifies the process of tracking transactions to identify patterns: wallets are represented as nodes, and transactional flows are depicted as outgoing arcs connecting the source wallet to the destination wallet. However, despite its intuitive and visually informative nature, this representation poses significant challenges in practice, particularly given that the number of Bitcoin wallets has already surpassed 460 million [68]. Time-frequency Analysis, with only one reference, is a less explored technique, suggesting a potential gap in research for methods that consider temporal or spectral data.

Table. 2. Classification with references to analyzed papers.

	Approach	References
Technical	Supervised learning	[2-28]
	Unsupervised learning	[29, 30]
	Hybrid learning	[31-33]
	Deep learning	[34, 37]
	Graphs	[38, 45]
	Time-frequency	[46]
Non-technical	Data analysis / Data mining	[47-53]
	Legislative	[54-61]
	Literature review	[62-67]

From the non-technical category, the **data analysis and data mining** studies focus on extracting insights from data without employing advanced machine learning or algorithmic models, offering a foundational understanding of trends and behaviors. These papers explain the concept of knowledge extraction from data originating from diverse and different sources: from blockchain data to social media platforms. The aim of analyzing such diverse data is to create patterns for suspicious activities within the blockchain network, for identifying suspicious individuals or to understand different trends within the network. The papers under the **legislative** umbrella tackle the importance of regulatory frameworks when it comes to money laundering within the cryptocurrency domain, highlighting the interplay between technology and policy. There is also a growing emphasis on highlighting legislative inconsistencies between countries worldwide, which create loopholes that can be exploited for money laundering while remaining on the edge of legality. The last category we have to mention here is represented by the papers **reviewing the literature**. Those papers do not come with any technical developments, but contribute to the research in this area, providing comprehensive overview of the field, identifying gaps, discussing challenges and indicating future directions.

3.1. Datasets Used

An essential aspect of this analysis is the dataset utilized in the previously chosen academic papers. We identified seven distinct situations, as illustrated in the Fig. 3.

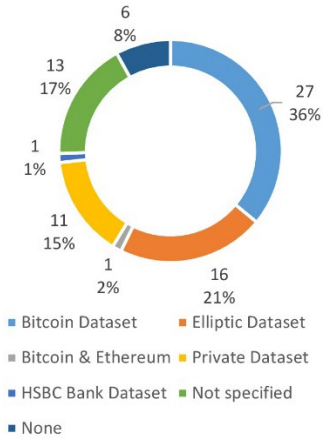

Fig. 3. Datasets used in the selected projects.

The most commonly used dataset, found in 36 % papers is the Bitcoin dataset. This dataset is represented by the real blockchain dataset containing transactions, addresses and information such as block size, block version, block header hashes, Merkle root, timestamps, nonces and other metadata. Using data from the Bitcoin blockchain results in findings that closely reflect real-world scenarios, thereby enhancing the relevance and practical applicability of these studies. Additionally, one paper indicated that its methodology could be applied to both Bitcoin and Ethereum blockchains, so we classified it as using mixed datasets. Another widely used dataset is the Elliptic database, featured in 16 papers, representing 21 % of our sample. The Elliptic dataset links Bitcoin transactions to real-world entities, categorizing them into illicit (scams, malware, terrorist organizations, Ponzi schemes, etc.) and legitimate categories (such as exchanges, wallet providers, and miners). Thirteen

papers (17 % of the total) did not specify a dataset, as they focused on general concepts applicable to various datasets. One study used a dataset from HSBC Bank, which is privately held, requiring pre-processing to address confidentiality concerns. Eleven papers relied on private databases provided by different industrial entities. There were 6 papers which did not claim any dataset used, which is caused by the fact that they were not discussing any technical aspect.

4. Experiments

4.1. Understanding the Importance of Parsing Data within Blockchain Networks

During our intensive study on the 150 papers, we have reached the conclusion that being able to process blockchain data and parse it effectively is a crucial step and advancement in the fight against money laundering. At the moment of writing, the Bitcoin blockchain reached almost 630 GB of data [69]. Downloading the entire Bitcoin blockchain took around 10 days to be completed. This process depends on peer-to-peer downloading from nodes, whose availability is uncertain, which can impact both the download speed and duration. Moreover, the files containing the blockchain data are system dependent, meaning that their size depends on what system they are stored on. As an example of the time required for these operations, a tool we developed to export transactions from the blockchain took approximately 9 hours to extract all transactions from four specified batches of **.blk** files (50, 100, 500, and 1,000 files). Additionally, we aimed to determine the number of blocks in each batch. For this task, the tool we created required roughly 10 hours to process all the files and return the block counts.

Parsing blockchain data is essential for money laundering investigations because it enables law enforcement and financial analysts to uncover and understand complex illicit activities within the blockchain's decentralized structure. Blockchain technology offers a high degree of transparency and immutability, recording every transaction on a public ledger. However, this transparency does not automatically translate into accessibility or usability without proper parsing tools and techniques. Parsing blockchain data allows investigators to extract, process, and analyze transaction information, such as sender and receiver addresses, transaction amounts, timestamps, and associated metadata, to identify suspicious activity and trace the movement of illicit funds. Money laundering often involves sophisticated schemes to obscure the origin, flow, and destination of illicit funds. Criminals may use tactics such as "mixers" or "tumblers" to combine and obscure transactions, chain-hopping to move funds across different blockchains, or using thousands of microtransactions to create a trail that is difficult to follow. Parsing blockchain data enables investigators to detect these patterns, identify anomalies, and link transactions to real-world entities or wallets associated with illegal activities. By reconstructing the flow of funds and identifying links between transactions, investigators can expose the networks involved in money laundering.

Moreover, parsing blockchain data is vital for collaboration with private-sector entities such as cryptocurrency exchanges, financial institutions, and blockchain intelligence firms. Many of these organizations rely on parsed data to comply with anti-money laundering (AML) regulations, monitor transactions, and report suspicious activities. Accurate and efficient parsing ensures that investigators have reliable insights to build strong legal cases, identify perpetrators, and recover stolen or laundered assets. As the use of cryptocurrencies in illicit finance continues to evolve, the ability to parse blockchain data remains a critical tool in the fight against money laundering and other financial crimes.

Several authors have highlighted a lack of functional tools available for effectively parsing blockchain data. We identified only a limited selection of such tools, most of which were either outdated and no longer usable or incomplete, requiring significant modifications to deliver satisfactory results. The goal of this experiment is to highlight the challenging and time-consuming process of handling blockchain data. Parsing the data to extract meaningful insights – whether as graphs, readable text, or a relational database – is the foundational step in investigating financial cybercrimes within the Bitcoin blockchain. This step is crucial for any research in the field, and we aim to emphasize the need for advancements in this area.

4.2. Tools

The first tool we are going to use in our experiment is **Bitcoin Core** [70], serving as a full-node client that downloads and verifies the entire blockchain. It ensures all transactions and blocks comply with the Bitcoin protocol's rules, enhancing security and decentralization. By running Bitcoin Core, users can independently validate their transactions without relying on third parties, maintaining full control and improving privacy. In addition to being a full node, Bitcoin Core includes a built-in wallet for securely storing, sending, and receiving Bitcoin. It also provides tools for developers to interact with the blockchain using APIs and supports testnets for experimentation and research. While resource-intensive, Bitcoin Core is vital for users and businesses seeking the highest levels of security, privacy, and contribution to the decentralization of the Bitcoin network. First of all, we have used Bitcoin Core to download all the blockchain data, process which took around 10 days, as also stated before. Our study focused on the folder containing all Bitcoin blocks, starting from the genesis block and extending

to the most recent one. These blocks are stored in files, each holding multiple blocks, with the number of blocks per file varying. Each Bitcoin block contains essential data that supports the network's operation, including the block header. The header provides a summary of the block's content through fields such as the version, the previous block's hash, the Merkle root, the timestamp, and the nonce. Additionally, each block includes a list of transactions, the block height, size, reward, difficulty target, and sometimes an extra nonce.

The Fig. 4 illustrates the components of a blockchain block, provided for a better understanding of its components.

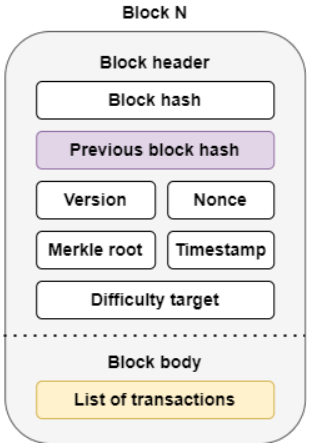

Fig. 4. The components of a Bitcoin-alike block.

Another tool we utilized is **BitcoinDatabaseGenerator** [71], a high-performance data transfer tool designed to copy blockchain data from Bitcoin Core files into a SQL Server database. This tool demonstrates its efficiency by employing multithreading techniques to accelerate the data transfer process into the database. Like any other tool we encountered, this one also has its drawbacks. One notable limitation is that its last update dates back to 2017. To address this, we worked on updating the tool, incorporating modern technological advancements, and ensuring it is fully functional. One characteristic of this tool, which can be a significant drawback in some cases, is that it always begins parsing from the very first block. It does not allow you to start parsing data from a specific block or select a particular interval, making it inefficient in terms of both time and storage size.

Another tool we utilized is **Blockchain-parser** [72], a straightforward script designed to parse **blk<XYZ>.dat** files from the Bitcoin blockchain database. This tool is highly useful as it enables a text-based view of the raw blockchain database files. In contrast to the previously mentioned tool, Blockchain-parser allows you to begin parsing from any **blk<XYZ>.dat** file of your choice, without the need to start from the beginning. Both tools are valuable depending on how you intend to use the data – whether you require it in a database or as plain text.

The final tool we will use for testing is **BTCGraphConstruction** [73], a Python-based tool designed to build a transaction graph by extracting data from the Bitcoin blockchain. While the tool lacks detailed documentation regarding necessary additional modules and usage instructions, its underlying logic appears potentially efficient for graph construction purposes. Notably, this is the only tool we found specifically focused on constructing transaction graphs.

4.3. Speedup Test

We begin our experiment with downloading the blockchain data using Bitcoin core. The block files containing the blockchain blocks have the size of 666 GB, containing approximately 4,700 **blk<XYZ>.dat** files.

The next step in our research is to process the blockchain data using 3 different tools: one that copies the data into a relational database, one that converts the data into readable format and another one that creates a transaction graph based from the blockchain blocks.

Our experimental setup consists of an Intel Core i9 13900H CPU, 48 GB of RAM, and an external Samsung SSD for storing blockchain blocks.

Due to some issues with the functionality of the tool employing the graph-based approach, it could not be fully operational for our intended tests One significant issue was the lack of proper documentation, which required us to invest considerable time in making the tool operational, including installing all necessary dependencies and additional modules and thoroughly investigating its functionality. Despite resolving these issues and successfully running the tool, we discovered that the amount of data it parsed was insufficient for inclusion in our research. As a result, we decided to exclude this tool from the study.

We proceeded with our research using the other two tools. The Table 3 details the exact time taken by each tool to process the first 50, 100, 500, and 1000.blk files. A significant difference in processing time between the two tools is evident. From a time-efficiency perspective, storing blockchain data in relational databases may prove to be a more effective solution.

Table. 3. Time in seconds/number of **blk<XYZ>.dat** files processed.

Number of blk files processed	Blockchain-parser	Bitcoin Database Generator
50	856 seconds	25.74 seconds
100	2,931 seconds	26.28 seconds
500	10,214 seconds	22.34 seconds
1000	20,288 seconds	21.56 seconds

Table 4 presents the dimensions of the four batches of blk files, along with the total number of blocks and transactions. We wanted to mention this, because, as previously stated, the **blk<XYZ>.dat** files are system dependent, therefore there might be discrepancies when tested on different setups.

Table 4. Total number of blocks, transactions and size of the **blk<XYZ>.dat** files processed.

Number of blkXXXX.dat files	Size in GB	Total number of blocks	Total number of transactions
0-49 (50)	7.08 GB	228,146 blocks	~15 million transactions
0-99 (100)	14.3 GB	274,296 blocks	~29 million transactions
0-499 (500)	71.0 GB	408,698 blocks	~124 million transactions
0-999 (1000)	141 GB	485,511 blocks	~254 million transactions

After considering the size of the files and the time required to process each batch, we created the following illustration to represent the bandwidth utilized during this operation. Bandwidth refers to the maximum amount of data that can be transmitted over a network in a given amount of time. It is a critical factor in determining the speed and efficiency of data transfer, especially for large-scale operations like processing blockchain files.

Storing blockchain data in relational databases proves to be a more efficient and effective approach, as demonstrated by the significant time savings in processing large batches of blockchain files. As shown in the table, the Bitcoin Database Generator consistently outperforms the Blockchain-parser tool, particularly as the file count increases. This efficiency stems from the structured nature of relational databases, which are optimized for querying and managing large datasets, making them a natural fit for handling blockchain information.

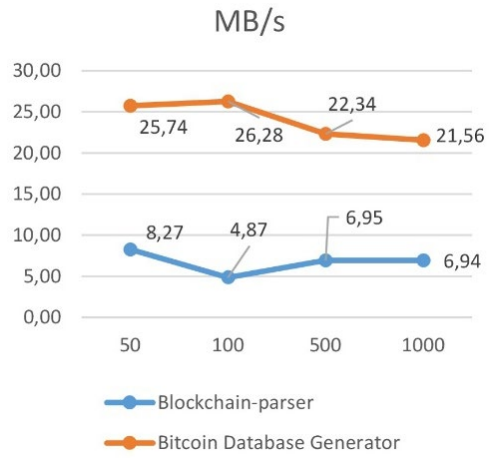

Fig. 4. Bandwidth while parsing blockchain data.

For criminal investigations, especially those focused on detecting and combating money laundering, relational databases offer additional advantages. Their ability to organize and query data systematically allows for more effective pattern recognition, link analysis, and data extraction. Investigators can quickly identify suspicious transactions, trace the flow of funds, and correlate findings with other financial data. This structured and scalable approach not only accelerates investigations but also ensures that crucial insights are not overlooked due to inefficiencies in data processing. Therefore, leveraging relational databases is a critical step toward enhancing the analysis and usability of blockchain data in the context of financial crime investigations.

5. Conclusions

Money laundering remains a significant challenge for global financial systems, with the rise of cryptocurrencies like Bitcoin introducing new complexities. Continuous research into money laundering detection tools, particularly for the Bitcoin blockchain, is therefore crucial for safeguarding the integrity of financial systems. Advancing these tools ensures that investigators can stay ahead of evolving criminal tactics, enabling the identification of suspicious transactions, criminal networks, and laundering patterns in a rapidly growing digital economy.

Efficiently processing blockchain data is a fundamental requirement for combating money laundering and other cybercrimes. The Bitcoin blockchain, as well as other cryptocurrency systems, generates immense amounts of data daily. Without the ability to effectively parse, analyze, and visualize this data, law enforcement and financial institutions risk missing critical evidence hidden within the blockchain. Continuous improvements in processing tools allow for faster and more accurate identification of anomalies, such as unusual transaction volumes, circular transactions, or accounts engaged in obfuscation techniques like mixing services. These insights enable the timely interception of illicit activities, ensuring that criminals face greater barriers to exploiting the blockchain.

Moreover, developing tools to detect and analyze blockchain transactions can contribute to the creation of comprehensive transaction patterns that can prove to be of great help when fighting against money laundering. By systematically identifying red flags, such as irregular transaction timings or links to known illicit entities, researchers can establish a framework of predictive patterns. These patterns, once shared with financial institutions and regulatory bodies, can not only help disrupt existing money laundering schemes but also discourage criminals from using blockchain platforms in the first place. Such advancements are critical not only for Bitcoin but also for the broader cryptocurrency ecosystem, promoting transparency

and trust among users while reducing its appeal as a medium for financial crime.

References

[1]. Elliptic Dataset, https://www.kaggle.com/datasets/ellipticco/elliptic-data-set

[2]. Joana Lorenz, et al., Machine learning methods to detect money laundering in the Bitcoin blockchain in the presence of label scarcity, in *Proceedings of the First ACM International Conference on AI in Finance (ICAIF'20)*, 2020, pp. 1-8.

[3]. A. A. Badawi, et al., Detection of money laundering in bitcoin transactions, in *Proceedings of the 4th Smart Cities Symposium (SCS'21), Online Conference*, Bahrain, 2021, pp. 458-464.

[4]. I. Alarab, et al., Comparative analysis using supervised learning methods for anti-money laundering in bitcoin, in *Proceedings of the 5th International Conference on Machine Learning Technologies (ICMLT'20)*, 2020, pp. 11-17.

[5]. J. Alotibi, et al., Money laundering detection using machine learning and deep learning, International *Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 13, Issue 10, 2022.

[6]. I. Alarab, et al., Competence of graph convolutional networks for anti-money laundering in bitcoin blockchain, in *Proceedings of the 5th International Conference on Machine Learning Technologies (ICMLT'20)*, 2020, pp. 23-27.

[7]. M, Weber, et al., Anti-money laundering in bitcoin experimenting with graph convolutional networks for financial forensics, *arXiv Preprint*, 2019, arXiv:1908.02591v1.

[8]. K. Kolesnikova, et al., Analysis of bitcoin transactions to detect illegal transactions using convolutional neural networks, in *Proceedings of the IEEE International Conference on Smart Information Systems and Technologies (SIST'19)*, Nur-Sultan, Kazakhstan, 2021, pp. 1-6.

[9]. S. Ouyang, et al., Bitcoin money laundering detection via subgraph contrastive learning, *Entropy*, Vol. 26, Issue 3, 2024, 211.

[10]. D. Vassalo, et al., Application of Gradient Boosting Algorithms for Anti-Money Cryptocurrencies, *Springer Computer Science*, Vol. 2, 2021, 143.

[11]. L. Yu, et al., Who are the Money Launderers? Money laundering detection on blockchain via mutual learning based graph neural network, in *Proceedings of the International Joint Conference on Neural Networks (IJCNN'23)*, Gold Coast, Australia, 2023, pp. 1-8.

[12]. I. Alarab, et al., Graph-Based LSTM for Anti-Money Laundering: Experimenting Temporal Graph Convolutional Network with Bitcoin Data, *Neural Processing Letters,* Volume 55, 2022, pp. 689-707.

[13]. W. W. Lo, et al., Inspection-L: self-supervised GNN node embeddings for money laundering detection in bitcoin, *Appl. Intell.*, Vol. 53, 2023, pp. 19406-19417.

[14]. C. Lee, et al., Toward detecting Illegal Transactions on Bitcoin Using Machine-Learning Methods, *Springer: Communications in Computer and Information Science,* Vol 1156, 2019, pp. 520-533.

[15]. B. Chen, et al., Bitcoin theft detection based on supervised machine learning algorithms, *Security and Communications Networks Journal*, 2021.

[16]. E. V. Feldman, et al., Bitcoin Abnormal Transaction Detection Based on Machine Learning, *Chelyabinsk Phys, Math. Journal,* Vol. 6, 2021, pp. 119-132.

[17]. P. Nerurkar, Illegal Activity Detection on Bitcoin Transaction Using Deep Learning, *Springer: Soft Computing*, Vol. 23, 2023, pp. 5503-5520.

[18]. C. Oliveira, et al., GuiltyWalker: Distance to illicit nodes in the Bitcoin network, *arXiv preprint*, 2021, arXiv:2102.05373.

[19]. M. Bhowmik, et al., Comparative Study for Machine Learning Algorithms for Fraud Detection in Blockchain, in *Proceedings of the 5th International Conference on Computing Methodologies and Communication (ICCMC'21)*, 2021, pp. 539-541.

[20]. C. Guo, et al., LB-GLAT: Long-term bi-graph layer attention convolutional network for anti-money laundering in transactional blockchain, *Mathematics*, Vol. 11, Issue 18, 2023, 3927.

[21]. N. Nayyer, et al., A new framework for fraud detection in bitcoin transactions: through ensemble stacking model in smart cities, *IEEE Access*, Vol. 11, 2023, pp. 90916-90938.

[22]. P. Nerurkar, et al., Supervised learning model for identifying illegal activities in Bitcoin, *Springer: Applied Intelligence,* Vol. 51, 2021, pp. 3824-3843.

[23]. N. B. Bynagari, et al., Anti-money laundering recognition though the gradient boosting classifier, *Academy of Accounting and Financial Studies Journal*, Vol. 25, Issue 5, 2021, pp. 1-11.

[24]. Mohammad Javad Shayegan et al., A Collective Anomaly Detection Crypto Wallet Frauds on Bitcoin Network, *Symmetry 2022*, Volume 14, Issue 2, 2022, 328.

[25]. Hao Hua Sun Yin et al., Regulating Cryptocurrencies: A Supervised Machine Learning Approach to De-anonymizing the Bitcoin Blockchain, *Journal of Management Information Systems*, Volume 36, Issue 1, 2019, pp. 37-73.

[26]. Eric Pettersson Ruiz et al., Combating money laundering with machine learning-applicability of supervised-learning algorithms at cryptocurrency e, *Journal of Money Laundering Control*, Vol. 25, Issue 4, 2021, pp. 766-778.

[27]. Pranav Nerurkar et al., Detecting Illicit Entities in Bitcoin using Supervised Learning of Ensemble Decision Trees, in *Proceedings of the International Conference on Information Communication and Management (ICICM),* 2020, pp. 25-30.

[28]. Udit Agarwal et al., Blockchain and crypto forensics: Investigating crypto frauds, *International Journal of Network Management*, Vol. 34, Issue 2, 2023, e2255.

[29]. Guangyi Yang et al., Anti-money laundering supervision by intelligent algorithm, *Computers & Security*, Vol. 132, 2023, 103344.

[30]. Hilmar Pall Stefansson et al., Detecting potential money laundering addresses in the Bitcoin blockchain using unsupervised machine learning, in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 1562-1571.

[31]. Yining Hu et al., Characterizing and Detecting Money Laundering Activities on the Bitcoin Network, *arXiv: Computer Science – Social and Information Networks,* 2019.

[32]. Palita Humranan et al., A Study on GCN using Focal Loss on Class-Imbalanced Bitcoin Transaction for Anti-Money Laundering Detection, in *Proceedings of the International Electrical Engineering Congress (iEECON),* 2023, pp. 101 – 104.

[33]. Danilo Labanca et al., Amaretto: An active Learning Framework for Money Laundering Detection, *IEEE Access,* Vol. 10, 2022, pp. 41720 – 41739.

[34]. Qasim Umer et al., Ensemble Deep Learning Based Prediction of Fraudulent Cryptocurrency Transactions, *IEEE Access*, Vol 11, 2023, pp. 95213-95224.

[35]. Jialin Song et al., HBTBD: A Heterogeneous Bitcoin Transaction Behavior Dataset for Anti-Money Laundering, *Journal of Applied Sciences*, Vol. 13, Issue 15, 2023, 8766.

[36]. Khalid Alkhatib et al., Anti-Laundering Approach for Bitcoin Transactions, in *Proceedings of the International Conference on Information and Communication Systems (ICICS)*, 2023, pp. 1-6.

[37]. Zhuoming Gu et al., On-chain analysis-based detection of abnormal transaction amount on cryptocurrency exchanges, *Physica A: Statistical Mechanics and its Applications*, Vol. 604, 2022, 127799.

[38]. Jeyakumar Samantha Tharani et al., Graph Based Visualisation Techniques for Analysis of Blockchain Transactions, in *Proceedings of the IEEE 46th Conference on Local Computer Networks (LCN),* 2021, pp. 427 – 430.

[39]. Qianyu Wang et al., GraphALM: Active Learning for Detecting Money Laundering Transactions on Blockchain Networks, *IEEE Network*, 10.1109, 2024.

[40]. Anuraj Mohan et al., Improving anti-money laundering in bitcoin using evolving graph convolutions and deep neural decision forest, *Journal of Data Technologies and Applications*, Volume 57, Issue 3, 2023, pp. 313-329.

[41]. Clynton Tomacheski et al., The nilcatenation problem and its application for detecting money laundering activities in cryptocurrency networks, *International Transactions in Operational Research*, Vol. 31, 2023, pp. 3559-4407.

[42]. Simone Marasi and Stefano Ferretti, Anti-Money Laundering in Cryptocurrencies through Graph Neural Networks: A Comparative Study, in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC),* 2024, pp. 272 – 277.

[43]. Bogdan Dumitrescu et al, Anomaly Detection in Graphs of Bank Transactions for Anti Money Laundering Applications, *IEEE Access*, Vol. 10, 2022, pp. 47699-47714.

[44]. Lijun Xiao et al., CTDM: cryptocurrency abnormal transaction detection method with spatio-temporal and global representation, *Journal of Soft Computing in Decision Making and in Modeling in Economics*, Vol. 27, 2023, pp. 11647-11660.

[45]. Yuhang Zhang et al., Transaction Community Identification in Bitcoin, in *Proceedings of the International Symposium on Computational Intelligence and Design (ISCID),* 2020, pp. 140 – 144.

[46]. Utku Gorkem Ketenci et al., A Time-Frequency Based Suspicious Activity Detection for Anti-Money Laundering, *IEEE Access*, Vol. 9, 2021, pp. 59957-59967.

[47]. Ammar Oad, Blockchain-enabled Transaction Scanning Method for Money Laundering Detection, *Journal of Electronics*, Vol. 10, Issue 15, 2021, 1766.

[48]. Sidharth Samanta et al., A framework to Build User Profile on Cryptocurrency Data for Detection of Money Laundering Activities, in *Proceedings of the International Conference on Information Technology*, 10.1109, 2019, pp. 425 – 429.

[49]. Lingxiao Yang et al., An abnormal Transaction Detection Mechanism on Bitcoin, in *Proceedings of the International Conference on Networking and Network Applications*, 2019, pp. 542 – 549.

[50]. Jesse Crawford et al., Knowing your Bitcoin Customer: Money Laundering in the Bitcoin Economy, *in Proceedings of the 13th International Conference on Systematic Approaches to Digital Forensic Engineering (SADFE'20),* 2020, pp. 38 – 45.

[51]. Youssef Elmougy et al., Demystifying Fraudulent Transactions and Illicit Nodes in the Bitcoin Network for Financial Forensics, in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining,* 2023, pp. 3979-3990.

[52]. Huy Tran Tien et al., Blockchain-Data Mining Fusion for Financial Anomaly Detection: A Brief Review, *Procedia Computer Science,* Vol. 235, 2023, pp. 478-483.

[53]. Gaspare Jucan Sicignano, Money Laundering using Cryptocurrency: the case of Bitcoin! *Athens Journal of Law*, Vol. 7, Issue 2, 2021, pp. 253-264.

[54]. Christian Leuprecht et al., Virtual money laundering: policy implications of the proliferation in the illicit use of cryptocurrency, *Journal of Financial Crime,* Vol. 30, Issue 4, 2022, pp. 1036-1054.

[55]. Christoph Wronka, Cyber-laundering: the change of money laundering in the digital age, *Journal of Money Laundering Control*, Vol. 25, Issue 2, 2022, pp. 330-344.

[56]. Tiara Putri et al., Inadequate Cryptocurrency and Money Laundering Regulations in Indonesia (Comparative Law of US and Germany), *Yustisia Jurnal Hukum*, Vol. 12, Issue 2, 2023, 129.

[57]. Cuneyt G. Akcora et al., How to Not Get Caught When You Launder Money on Blockchain?, *Cryptography and Security*, *arXiv:2010.15082*, 2020.

[58]. Thomas A. Frick, Virtual and cryptocurrencies – regulatory and anti-money laundering approaches in the European Union and in Switzerland, *ERA Forum*, Vol. 20, 2019, pp. 99-112.

[59]. Iulia Oana Florea and Maria Nitu, Money Laundering Through Cryptocurrencies, *The Romanian Economic Journal*, Vol. 22, Issue 76, 2020, pp. 66-71.

[60]. Serhii Hrytsai, Cryptocurrency in the declarations of government officials: A toolkit for money laundering (trends and experience of counteraction by the example of Ukraine), *Access to Justice in Eastern Europe Journal (AJEE),* Vol. 3, 2023, pp. 1-27.

[61]. Christopher P. Buttigieg et al., Anti-money laundering regulation of crypto assets in Europe's smallest member state, *Law and Financial Markets Review*, Vol. 13, 2019, pp. 211-227.

[62]. Otilia Muntean and Ciprian Pungila, Finding Needles in a Haystack: a Taxonomy on Money Laundering on the Bitcoin Blockchain, in *Proceedings of the 3rd Blockchain and Cryptocurrency Conference (B2C'24),* 2024, pp. 58-64.

[63]. Jiajing Wu et al., Analysis of cryptocurrency transactions from a network perspective: An overview, *Elsevier: Journal of Network and Computer Applications*, Vol. 190, 2021, 103139.

[64]. Chang-Yi Lin, A Systematic Review of Detecting Illicit Bitcoin Transactions, *Procedia Computer Science*, Vol. 207, 2022, pp. 3217-3225.

[65]. Japinye A. et al., Integrating Machine Learning in Anti Money Laundering through Crypto: A Comprehensive Performance Review, *European Journal of Accounting Auditing and Finance Research*, Vol. 12, Issue 4, 2024, pp. 54-80.

[66]. Bekach Youssef et al., State of the Art Literature on Anti-money Laundering Using Machine Learning and

Deep Learning Techniques, in Proceedings of the *3rd International Conference on Artificial Intelligence and Computer Vision*, 2023, pp. 77-90.
[67]. Dattatray Vishnu Kute et al., Deep Learning and Explainable Artificial Intelligence Techniques Applied for Detecting Money Laundering–A Critical Review, *IEEE Access,* Vol. 9, 2021, pp. 82300-82317.
[68]. Chainalysis, https://www.chainalysis.com/blog/bitcoin-addresses/, last accessed on 21 December 2024
[69]. Blockchain.com, https://www.blockchain.com/explorer/charts/blocks-size, last accessed on 30 December 2024
[70]. Bitcoin Core, https://bitcoin.org/en/bitcoin-core/, last accessed on 30 December 2024
[71]. BitcoinDatabaseGenerator, https://github.com/ladimolnar/BitcoinDatabaseGenerator, last accessed on 30 December 2024
[72]. Blockchain-parser, https://github.com/ragestack/blockchain-parser, last accessed on 30 December 2024
[73]. BTCGraphConstruction, https://github.com/hugoschnoering2/BTCGraphConstruction, last accessed on 30 December 2024