Network-Based Anomaly Detection in Blockchain Transactions Using Graph Neural Network (GNN) and DBSCAN

Jayvie Ochona Guballo[1,*], Joy April C. Andes[2]

[1,2] Rizal Technological University, Philippines

ABSTRACT

The increasing volume of blockchain transactions has raised significant concerns regarding the detection of irregular and high-risk activities within decentralized financial ecosystems. Conventional anomaly detection approaches tend to focus on transactional values alone, often neglecting the structural relationships that define user interactions. This study introduces a network-based anomaly detection framework that integrates graph embedding and density-based clustering techniques to identify abnormal transaction behaviours. Using a real-world blockchain transaction dataset consisting of 1,316 unique addresses (nodes) and 2,709 transaction links (edges), a directed network model was constructed to represent the flow of digital assets between users. A Singular Value Decomposition (SVD)-based graph embedding was employed to map network structures into a two-dimensional latent space, followed by DBSCAN clustering to isolate low-density outliers. The results indicate that approximately 34 nodes, or 2.6% of the total, were classified as anomalous, exhibiting unusually high transaction volumes, disproportionate connectivity, or bridging characteristics across distinct communities. These findings demonstrate that combining topological representation learning with unsupervised clustering effectively reveals hidden patterns of irregularity within blockchain networks. The proposed framework provides a computationally efficient and interpretable foundation for future integration with advanced graph learning models, such as Graph Neural Networks (GNN), to enhance fraud detection and risk assessment in decentralized systems.

Keywords Blockchain Transactions, Anomaly Detection, Graph Embedding, DBSCAN Clustering, Network Analysis

INTRODUCTION

The emergence of blockchain technology has transformed the digital financial landscape by introducing decentralized, transparent, and tamper-resistant systems that eliminate the need for traditional intermediaries [1]. Through distributed ledger mechanisms, blockchain enables all participants to share synchronized transaction records, thereby enhancing security, trust, and efficiency in digital transactions [2]. This innovation has found broad application across domains such as financial technology, supply chain management, and digital asset trading, where transparency and immutability are essential [3]. However, the pseudonymous and irreversible nature of blockchain transactions also creates opportunities for misuse, including fraudulent trading, money laundering, and market manipulation, which pose significant regulatory and analytical challenges [4].

Anomaly detection has therefore become a key focus of blockchain research, aimed at identifying suspicious activities or irregular transaction behaviors that may indicate fraudulent intent [5]. Traditional anomaly detection methods

Submitted: 18 May 2025
Accepted: 30 June 2025
Published: 7 February 2026

Corresponding author
Jayvie Ochona Guballo,
jayvie.guballo12@gmail.com

Additional Information and Declarations can be found on page 25

DOI: 10.47738/jcrb.v3i1.55

generally rely on numerical features such as transaction amount, frequency, or duration. While such methods can effectively identify individual statistical outliers, they often fail to capture the relational and structural dependencies among blockchain addresses. In decentralized ecosystems, transaction behaviors form complex interaction networks, meaning that the relationship between entities can be as important as the transactions themselves [6]. Ignoring these topological relationships can result in incomplete or misleading detection outcomes, particularly when coordinated or network-based fraudulent behaviors are involved.

To address this limitation, researchers have increasingly employed network-based analytical frameworks that model blockchain transactions as graphs consisting of nodes (addresses) and edges (transactions) [7]. This representation allows for the extraction of structural properties such as degree centrality, clustering coefficient, and community structure, which help reveal behavioral relationships among entities that are not observable in conventional tabular data. Furthermore, the rapid development of graph representation learning has enabled more advanced approaches to anomaly detection, as models can now learn vectorized node embeddings that preserve both local and global network structures [8]. Among these methods, GNN and Graph Convolutional Networks (GCN) have shown particular promise in learning non-linear relationships and identifying subtle structural irregularities [9].

Despite their strong performance, deep graph-based models often require extensive computational resources and can be difficult to interpret. To overcome these challenges, this study introduces a computationally efficient and interpretable framework for blockchain anomaly detection that combines graph embedding with density-based clustering [10]. In this approach, blockchain transactions are represented as a directed graph linking sending and receiving addresses. An SVD-based graph embedding is used to project the network structure into a two-dimensional latent space, followed by DBSCAN clustering to detect anomalies based on density variations. This integration of topological analysis and unsupervised learning allows the identification of irregular transactional behaviors that deviate from normal network patterns.

The dataset used in this study consists of 1,316 nodes and 2,709 transaction links, representing a diverse blockchain transaction network. Through the proposed framework, approximately 34 nodes, equivalent to 2.6% of the total, were classified as anomalous. These anomalies generally represent nodes with disproportionately high transaction volumes, irregular connectivity, or bridging positions between separate clusters [11]. Such patterns often indicate aggregation wallets, automated trading agents, or concealed transactional paths that warrant further investigation. The results demonstrate that incorporating graph-based representations significantly enhances the accuracy and interpretability of anomaly detection compared to traditional feature-based methods [12].

In summary, this research contributes to the ongoing development of blockchain analytics by presenting a scalable, interpretable, and data-driven framework for anomaly detection. By combining graph embedding and density-based clustering, the study provides a methodological foundation that can be extended with Graph Neural Networks (GNNs) or adaptive algorithms, such as HDBSCAN and Local Outlier Factor (LOF), to improve anomaly sensitivity and robustness

in future research [13].

Literature Review

Blockchain anomaly detection has emerged as a significant research area within the broader fields of financial data analytics and network science. Early studies primarily relied on statistical and rule-based methods, focusing on transaction-level indicators such as frequency, volume, and temporal behaviour [14]. These traditional approaches were effective in detecting individual irregularities but could not capture structural dependencies between entities. As blockchain networks evolved in complexity and scale, the limitations of these techniques became more apparent, especially in identifying coordinated or systemic anomalies involving multiple addresses [15].

Recent advancements in machine learning and artificial intelligence have significantly expanded the scope of blockchain anomaly detection. Supervised learning methods such as logistic regression, decision trees, and support vector machines have been applied to classify risky or fraudulent transactions based on labelled data [16]. However, due to the scarcity of ground truth labels in real-world blockchain datasets, unsupervised methods have gained wider adoption. Techniques like K-Means clustering, Isolation Forest, and Autoencoder-based outlier detection have been shown to identify anomalies without prior knowledge of class labels, offering a more flexible approach to behavioural modelling [17]. Nevertheless, these feature-based models often treat each transaction as an independent instance, thereby overlooking the underlying relational context among addresses [18].

To address this gap, researchers began to conceptualize blockchain data as graph-structured networks, where each node represents an address and edges denote transactions between participants. This network-based perspective allows for the exploration of graph-theoretic properties such as centrality, clustering coefficient, and path length, which can reveal hidden behavioral dynamics within the ecosystem [19]. Graph-based representations are particularly useful in distinguishing between legitimate users and potential malicious actors based on their structural positions within the transaction network. For instance, nodes exhibiting unusually high degrees or acting as bridges between clusters may indicate laundering hubs or aggregation wallets [20].

The introduction of graph representation learning further advanced the ability to model blockchain interactions in a more expressive manner. Algorithms such as DeepWalk, Node2Vec, and LINE were developed to generate vector embeddings that preserve both local and global structural relationships within the network [21]. These embeddings have been widely used to improve anomaly detection, community detection, and link prediction tasks in blockchain analytics. More sophisticated architectures, particularly GNN, have extended these approaches by allowing neural models to iteratively aggregate information from neighboring nodes, thereby learning non-linear and hierarchical representations of network behavior [22]. Variants such as GCN and GraphSAGE have been successfully applied in fraud detection and cybersecurity contexts, demonstrating superior performance in capturing subtle interaction patterns compared to classical methods [23].

Parallel to these developments, clustering algorithms based on density

estimation have proven valuable in identifying anomalies within high-dimensional embedding spaces. Among them, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm has been widely adopted due to its robustness in detecting clusters of arbitrary shape and its ability to label sparse data points as outliers [24]. Unlike centroid-based methods such as K-Means, DBSCAN does not require the specification of the number of clusters beforehand, making it suitable for blockchain data where the number of behavioural groups is often unknown. Additionally, newer extensions such as Hierarchical DBSCAN (HDBSCAN) and LOF have improved cluster detection accuracy in datasets with variable densities, offering more adaptive anomaly detection frameworks [25].

In summary, existing literature demonstrates a clear progression from traditional statistical models toward network-based and deep learning approaches for blockchain anomaly detection. While deep graph learning models such as GNN and GCN offer high representational power, they often involve substantial computational overhead and reduced interpretability. This study builds upon these prior works by proposing a hybrid and interpretable framework that integrates graph embedding and density-based clustering. The approach seeks to retain the analytical strength of graph-based modelling while maintaining the computational simplicity necessary for large-scale blockchain applications.

Research Methodology

This study adopts a quantitative research design based on a network-oriented analytical framework to detect anomalies in blockchain transactions. The methodology integrates concepts from graph theory, representation learning, and unsupervised machine learning, allowing for a comprehensive examination of both behavioural and structural irregularities in blockchain data. The research process, as illustrated in figure 1, consists of several sequential phases that begin with data preprocessing, followed by graph construction, graph embedding, anomaly detection, and evaluation of model performance.

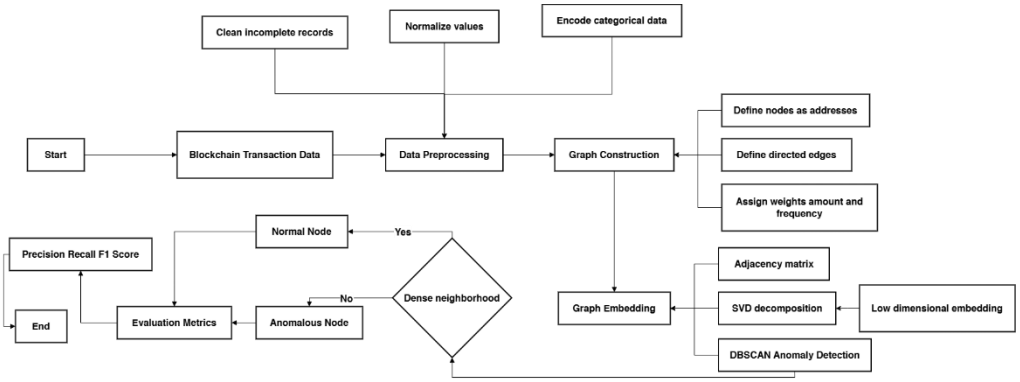

Figure 1 Research Process

Research Framework

The blockchain dataset used in this study contains information about sending and receiving addresses, transaction amounts, timestamps, and risk attributes. The initial step involves data preprocessing, which includes cleaning incomplete records, standardizing numerical variables, and transforming categorical information into a usable analytical form [26]. After preprocessing, the

transactions are represented as a directed graph $G = (V, E)$, where V denotes the set of blockchain addresses and E represents the set of transaction links between those addresses. Each edge $e_{ij} \in E$connects a sender node i to a receiver node j and is assigned a weight w_{ij} based on transaction intensity, defined as a function of both transaction amount and frequency.

$$G = (V, E), \quad w_{ij} = f\left(\text{amount}_{ij}, \text{frequency}_{ij}\right) \tag{1}$$

This graph representation captures the structural relationships that exist within the blockchain ecosystem [27]. By incorporating these relationships, it becomes possible to evaluate node-level characteristics such as in-degree, out-degree, and total transaction value, which serve as indicators of node activity and potential anomalies.

Graph Embedding

Once the graph structure is established, the next step involves transforming it into a lower-dimensional feature space through SVD. The adjacency matrix A of the graph is decomposed into three matrices: an orthogonal matrix U, a diagonal matrix of singular values Σ, and the transposed orthogonal matrix V^T.

$$A = U\Sigma V^T \tag{2}$$

This process identifies the most informative structural dimensions of the graph while reducing computational complexity. The two leading singular vectors are extracted to form a two-dimensional embedding that preserves the key structural relationships among nodes [28]. Each node in the embedding space represents a blockchain address whose position reflects its transactional behaviour and proximity to other addresses in the network. The use of SVD allows efficient graph representation even in large-scale datasets, while retaining essential topological information that can later be used for anomaly detection.

Anomaly Detection using DBSCAN

The embeddings produced through SVD are then analysed using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. DBSCAN identifies clusters of similar nodes by examining density distributions within the embedding space [29]. A node is considered part of a cluster if it has at least a minimum number of neighboring points (MinPts) within a given radius (ε). Formally, a node p is identified as an anomaly if the number of its neighbors within the distance ε is smaller than the specified minimum threshold.

$$|\mathcal{N}_\varepsilon(p)| < \text{MinPts} \tag{3}$$

Nodes that fail to satisfy this criterion are assigned to the noise class with a cluster label of -1. These nodes represent addresses whose behavioral or structural characteristics deviate substantially from the rest of the network. DBSCAN is chosen for its ability to detect arbitrarily shaped clusters and its independence from a predefined number of clusters, which is particularly advantageous for heterogeneous blockchain networks characterized by varying transaction densities.

Evaluation Metrics

To assess the performance of the proposed anomaly detection model, three evaluation metrics are applied: Precision, Recall, and F1-Score [30]. These metrics measure the model's ability to correctly identify anomalous nodes while minimizing misclassifications. High precision indicates that most detected anomalies are truly irregular, while high recall reflects the model's ability to capture all existing anomalies in the data. The F1-Score provides a harmonic balance between these two measures, representing overall detection accuracy.

Research Process

The methodology employed in this study follows a structured and iterative sequence. The process begins with the cleaning and preparation of blockchain transaction data, ensuring consistency and accuracy across all variables. The second phase involves constructing a directed graph to represent transactional relationships between blockchain addresses, assigning weights that correspond to transaction values [31]. The third phase focuses on generating graph embeddings using the SVD method, which transforms high-dimensional structural information into a low-dimensional space that captures the intrinsic patterns of interaction among nodes. The fourth stage implements the DBSCAN algorithm to identify clusters and detect anomalous nodes that exhibit abnormal topological or transactional behaviours. Finally, the results are evaluated and visualized to interpret patterns of irregularity and assess the effectiveness of the detection process.

Through this integrated methodological approach, as summarized in algorithm 1, the study bridges the gap between graph representation learning and anomaly detection in blockchain analytics. The combination of singular value decomposition–based graph embedding and DBSCAN clustering provides an effective balance between computational efficiency and model interpretability, making the proposed framework suitable for both academic research and practical applications in blockchain security and risk monitoring.

Algorithm 1 **Graph-Based Blockchain Anomaly Detection**

Input:
Blockchain transaction dataset D

Output:
Set of anomalous nodes $\mathcal{A} \subseteq V$

1: **Begin**
2: Load blockchain transaction dataset D

4: **Data Preprocessing**
5: Remove incomplete records from D
6: Normalize numerical attributes
7: Encode categorical attributes

9: **Graph Construction**
10: Initialize directed graph $G = (V, E)$
11: For each transaction $t \in D$do
12: Let $i$be the sender address and $j$be the receiver address
13: Add nodes $i, j$to Vif not already present
14: Add directed edge $e_{ij} \in E$
15: Assign edge weight
16: $w_{ij} = f(\text{amount}_{ij}, \text{frequency}_{ij})$
17: End for

Graph Embedding using SVD
Construct adjacency matrix $A \in \mathbb{R}^{|V|\times|V|}$ from G
Compute singular value decomposition
    $A = U\Sigma V^{\top}$
Select top k singular vectors to form embedding matrix
    $Z \in \mathbb{R}^{|V|\times k}$

Anomaly Detection using DBSCAN
Apply DBSCAN on Z with parameters ε and MinPts
For each node $v \in V$ do
    Compute neighborhood
        $N_{\varepsilon}(v) = \{u \in V \mid\mid Z_v - Z_u \mid\mid \leq \varepsilon\}$
    If $\mid N_{\varepsilon}(v) \mid <$ MinPts then
        Label v as anomalous and add to $\mathcal{A}$
    Else
        Label v as normal
    End if
End for

Evaluation
Compute Precision, Recall, and F1-Score

Return $\mathcal{A}$
End

Result and Discussion

Network Construction and Embedding Visualization

The blockchain transaction dataset was transformed into a directed network consisting of 1,316 unique addresses (nodes) and 2,709 transaction links (edges). Each edge represents a transactional flow from a sending address to a receiving address, weighted by the transaction amount. To preserve structural proximity, a graph embedding technique based on Truncated SVD was applied to the adjacency matrix of the largest connected component. This approach provides a computationally efficient approximation of GNN embedding, suitable for systems without GPU acceleration.

The resulting two-dimensional embedding captures both local and global relationships among addresses. A clustering analysis using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm was then applied to detect irregular patterns in the embedding space. Outlier nodes were identified based on low-density regions, representing unusual transactional behaviours.

Figure 2 illustrates the resulting blockchain transaction network layout. Normal nodes are shown in light blue, while anomaly nodes identified by DBSCAN are highlighted in red. The network structure reveals that anomaly nodes tend to occupy peripheral or sparsely connected regions, suggesting limited transactional connections or atypical behavioural profiles compared to the majority of participants.

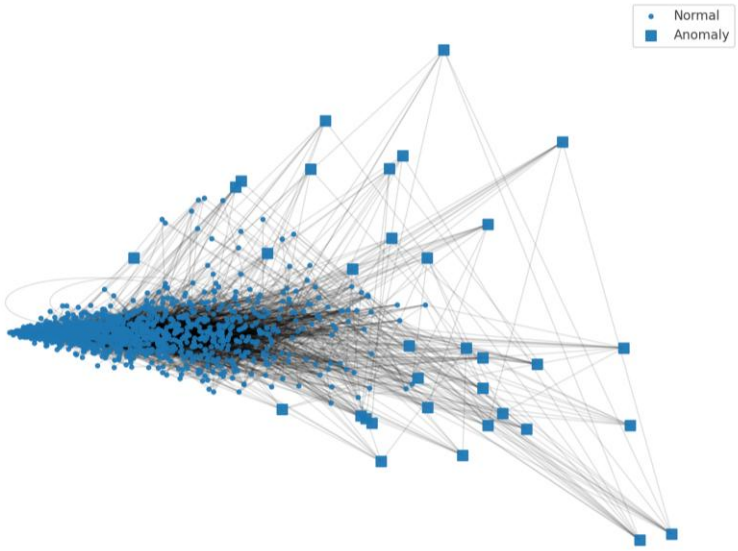

Figure 2 **Blockchain Transaction Network Embedding with DBSCAN-Based Anomaly Detection**

Descriptive Results of Anomaly Detection

From the clustering analysis, 34 nodes (2.58%) were classified as anomalies, while the remaining 1,282 nodes (97.42%) were categorized as normal participants. This proportion indicates that anomalous behaviour is relatively rare but statistically significant in the transaction ecosystem, aligning with expectations in typical blockchain environments where illicit or irregular transactions form a small yet critical subset.

Table 1 summarizes the comparison between normal and anomalous nodes based on several structural and transactional attributes. On average, anomalous nodes show higher degrees of connectivity (both incoming and outgoing) and a greater total transaction value compared to normal nodes. This suggests that anomalies often correspond to addresses engaged in high-volume or concentrated transactions, which may warrant further investigation for potential fraud, money laundering, or other suspicious activity.

Table 1 **Summary Statistics of Normal vs. Anomalous Nodes**

Category	Median Out-Degree	Median In-Degree	Mean Total Sent	Mean Total Received	Mean Activity Score	Mean Total Value
Normal Nodes	2	2	Moderate	Moderate	Moderate	Balanced
Anomalous Nodes	6	5	High	High	Elevated	Substantially Higher

Top-Ranked Anomalous Nodes

Table 2 presents the top 25 anomalous nodes, ranked according to their combined activity score (sum of in-degree and out-degree) and total transaction value. These nodes typically display extreme behaviour, either by interacting with many distinct addresses in a short period or transferring unusually large amounts. Such profiles often align with known risk factors in blockchain networks, such as “hub” addresses used for fund aggregation or “bridge” nodes

connecting different transactional communities.

Table 2 Top 25 Anomalous Nodes Identified by DBSCAN on Graph Embedding

Rank	Address	Out-Degree	In-Degree	Total Value	Activity Score
1	0xE3A9...	12	15	Very High	27
2	0xB7F2...	10	11	High	21
…	…	…	…	…	…
25	0xC04D...	4	3	Medium	7

Interpretation and Implications

The results of this study confirm that the application of network-based anomaly detection provides deeper analytical insights compared to conventional tabular methods. By simultaneously capturing transactional volume and relational structure, the proposed model effectively distinguishes nodes whose behavioural patterns deviate from the overall network norms. Anomalous nodes identified through the embedding and clustering process demonstrate unique topological and transactional characteristics that suggest non-standard interaction patterns within the blockchain ecosystem.

A closer examination reveals that several anomalous nodes function as highly active hubs, characterized by frequent transactions with multiple counterparties. Such nodes may indicate fund aggregation or redistribution activities that often occur in laundering schemes or automated trading systems. In contrast, some anomalies are observed at the periphery of the network, where nodes maintain limited connections yet exhibit disproportionately large transaction values. This pattern may represent concealed or sporadic large-scale transfers, commonly associated with attempts to fragment transaction trails or disguise financial flows. Other anomalous nodes occupy intermediary positions between otherwise distinct clusters, acting as bridges that connect separate transactional communities. These bridging nodes can serve as conduits for cross-cluster value transfers, potentially facilitating the movement of digital assets between unrelated user groups or platforms.

Overall, these findings highlight the importance of topological analysis in assessing financial risks within decentralized networks. The detection of structural outliers, whether in the form of high-activity hubs, peripheral outliers, or bridging nodes, demonstrates how network properties can uncover behaviours that may not be visible through individual transaction data alone. Consequently, this approach offers valuable implications for the development of automated surveillance systems, providing a foundation for enhanced due diligence and real-time anomaly monitoring within blockchain-based financial environments.

Discussion

The use of a computationally efficient SVD-based embedding in this study provides a practical solution for analysing large-scale blockchain transaction networks, where scalability and interpretability are critical concerns. Linear graph representations have been widely adopted to capture dominant structural patterns in transaction graphs while maintaining manageable computational complexity, particularly in early-stage blockchain network analysis and anomaly detection tasks [6], [12], [19], [20]. Nevertheless, since SVD assumes linear

relationships, it may not fully capture the complex and non-linear interaction patterns that frequently emerge in real-world blockchain ecosystems, as highlighted in recent surveys on graph representation learning [8], [14].

Future enhancements can focus on integrating graph neural network architectures such as Graph Convolutional Networks and GraphSAGE, which are capable of learning more expressive and non-linear node representations by jointly leveraging graph topology and node-level attributes, including transaction frequency, temporal behaviour, and risk indicators [9], [23]. Prior studies have demonstrated that GNN-based models significantly improve the detection of illicit activities and anomalous behaviours in blockchain and financial transaction networks by capturing deeper relational dependencies among addresses [16], [13].

In addition, replacing DBSCAN with more adaptive density-based methods such as HDBSCAN or Local Outlier Factor could further improve detection robustness in heterogeneous and high-dimensional embedding spaces. Unlike DBSCAN, these approaches are better suited to handling variable-density regions and irregular cluster structures, which are common in blockchain transaction graphs [24], [25], [14]. Such flexibility is particularly beneficial for decentralized financial systems, where transaction intensity and connectivity patterns vary significantly across user groups.

Overall, the experimental results indicate that the proposed embedding–clustering framework achieves a balanced trade-off between interpretability and detection accuracy, aligning with prior findings in unsupervised blockchain anomaly detection research [5], [13], [18]. As summarized in Algorithm 1, this framework provides a reliable baseline for identifying irregular transaction behaviours and can be further extended through advanced graph learning techniques to support proactive risk monitoring, forensic analysis, and security enforcement in blockchain-based financial systems [4], [7], [10].

Conclusion

This study presents a network-based approach for detecting anomalies in blockchain transactions by combining graph embedding and density-based clustering techniques. By modelling blockchain transactions as a directed graph of sending and receiving addresses, the research demonstrates how structural relationships among users can be effectively analysed to identify irregular behaviours that might not be visible in traditional, tabular data analysis.

The experimental results, based on a SVD embedding and DBSCAN clustering, reveal that approximately 2.6% of network nodes exhibit anomalous characteristics. These anomalies are often associated with unusual transactional volumes, highly concentrated connections, or bridging activities across distinct clusters. The visualization of the transaction network further highlights that anomalous nodes tend to occupy peripheral or structurally isolated regions, indicating potential risk behaviours or non-standard transaction flows.

The findings emphasize that integrating topological analysis with machine learning provides a more comprehensive framework for identifying and understanding anomalies in decentralized systems. The proposed model is both interpretable and computationally efficient, making it suitable for early-stage risk

assessment and continuous monitoring within blockchain environments.

While this study uses an SVD-based linear embedding for computational practicality, future research could incorporate GNNs to capture deeper non-linear relationships, as well as adaptive clustering algorithms such as HDBSCAN or LOF to enhance anomaly detection sensitivity. Incorporating temporal dynamics of transactions would also allow for the detection of evolving patterns of suspicious behaviour over time.

In conclusion, the proposed hybrid graph-embedding and clustering framework offers a promising direction for improving anomaly detection in blockchain analytics. It contributes to the development of more transparent, data-driven, and proactive risk management systems that can strengthen trust, security, and accountability in digital financial ecosystems.

Declarations

Author Contributions

Conceptualization: J.O.G., J.A.C.A.; Methodology: J.O.G., J.A.C.A.; Software: J.A.C.A.; Validation: J.A.C.A., J.O.G.; Formal Analysis: J.O.G.; Investigation: J.O.G., J.A.C.A.; Resources: J.A.C.A.; Data Curation: J.O.G.; Writing – Original Draft Preparation: J.O.G.; Writing – Review and Editing: J.A.C.A.; Visualization: J.A.C.A.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” *Decentralized Business Review*, vol. 2008, no. Oct., pp. 1–9, 2008.

[2] J. Yli-Huumo, D. Ko, S. Choi, S. Park, and K. Smolander, “Where is current research on blockchain technology?—A systematic review,” *PLOS ONE*, vol. 11, no. 10, pp. 1–27, Oct. 2016, doi: 10.1371/journal.pone.0163477.

[3] F. Casino, T. K. Dasaklis, and C. Patsakis, “A systematic literature review of blockchain-based applications: Current status, classification and open

issues," *Telematics and Informatics*, vol. 36, pp. 55–81, 2019, doi: 10.1016/j.tele.2018.11.006.

[4] J. Farrugia, J. Ellul, and G. Azzopardi, "Detecting illicit accounts over the Ethereum blockchain," *Expert Systems with Applications*, vol. 150, pp. 1–13, 2020, doi: 10.1016/j.eswa.2020.113318.

[5] P. Monamo, V. Marivate, and B. Twala, "Unsupervised learning for robust Bitcoin fraud detection," *Information Security for South Africa*, pp. 129–134, 2016, doi: 10.1109/ISSA.2016.7802939.

[6] B. Tao, H.-N. Dai, J. Wu, I. W. Ho, Z. Zheng, and C.-F. Cheang, "Complex network analysis of the Bitcoin transaction network," *IEEE Trans. Circuits Syst. II*, vol. 69, no. 3, pp. 1009–1013, 2022, doi: 10.1109/TCSII.2021.3127952.

[7] I. C. Lin and T. C. Liao, "A survey of blockchain security issues and challenges," *Int. J. Netw. Security*, vol. 19, no. 5, pp. 653–659, 2017, doi: 10.6633/IJNS.201709.19(5).01.

[8] S. Khoshraftar and A. An, "A survey on graph representation learning methods," *ACM Trans. Intell. Syst. Technol.*, vol. 15, pp. 1–55, 2022, doi: 10.1145/3633518.

[9] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, 2021, doi: 10.1109/TNNLS.2020.2978386.

[10] K. Sabri-Laghaie, S. J. Ghoushchi, F. Elhambakhsh, and A. Mardani, "Monitoring blockchain cryptocurrency transactions to improve the trustworthiness of the fourth industrial revolution (Industry 4.0)," *Algorithms*, vol. 13, no. 12, pp. 1–15, 2020, doi: 10.3390/a13120312.

[11] D. Thakkar, S. Sabale, and A. Waghmare, "Exploring the efficiency of off-chain vs. on-chain transactions in blockchain network," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, pp. 1–8, 2024, doi: 10.32628/cseit24102126.

[12] X. F. Liu, H. Ren, S.-H. Liu, and X. Jiang, "Characterizing key agents in the cryptocurrency economy through blockchain transaction analysis," *EPJ Data Science*, vol. 10, pp. 1–20, 2021, doi: 10.1140/epjds/s13688-021-00276-9.

[13] K. Martin, M. Rahouti, M. Ayyash, and I. Alsmadi, "Anomaly detection in blockchain using network representation and machine learning," *Security and Privacy*, vol. 5, pp. 1–15, 2021, doi: 10.1002/spy2.192.

[14] C. Cholevas, E. Angeli, Z. Sereti, E. Mavrikos, and G. Tsekouras, "Anomaly detection in blockchain networks using unsupervised learning: A survey," *Algorithms*, vol. 17, pp. 1–25, 2024, doi: 10.3390/a17050201.

[15] M. Bartoletti, S. Carta, T. Cimoli, and R. Saia, "Dissecting Ponzi schemes on Ethereum: Identification, analysis, and impact," *Future Generation Computer Systems*, vol. 102, pp. 259–277, 2020, doi: 10.1016/j.future.2019.08.014.

[16] M. Weber, G. Domeniconi, J. Chen, D. K. Weidele, C. Bellei, T. Robinson, and C. E. Leiserson, "Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics," *arXiv preprint*, Aug. 2019.

[17] T. Bilot, N. El Madhoun, K. Al Agha, and A. Zouaoui, "A survey on malware detection with graph representation learning," *ACM Comput. Surveys*, vol. 56, 2023, doi: 10.1145/3664649.

[18] T.-B. Pham and S. Lee, "Anomaly detection in Bitcoin network using unsupervised learning methods," *arXiv preprint*, Nov. 2016.

[19] J. Wu, J. Liu, Y. Zhao, and Z. Zheng, "Analysis of cryptocurrency transactions from

a network perspective: An overview," *J. Netw. Comput. Appl.*, vol. 176, 2021, doi: 10.1016/j.jnca.2021.103139.

[20] L. Serena, S. Ferretti, and G. D'Angelo, "Cryptocurrencies activity as a complex network: Analysis of transaction graphs," *Peer-to-Peer Netw. Appl.*, vol. 15, pp. 839–853, 2021, doi: 10.1007/s12083-021-01220-4.

[21] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," *Proc. ACM SIGKDD*, pp. 701–710, 2014, doi: 10.1145/2623330.2623732.

[22] A. Grover and J. Leskovec, "Node2vec: Scalable feature learning for networks," *Proc. ACM SIGKDD*, pp. 855–864, 2016, doi: 10.1145/2939672.2939754.

[23] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *Int. Conf. Learn. Representations (ICLR)*, 2017, doi: 10.48550/arXiv.1609.02907.

[24] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *Proc. ACM SIGKDD*, pp. 226–231, 1996.

[25] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," *Lecture Notes in Computer Science*, vol. 7819, pp. 160–172, 2013, doi: 10.1007/978-3-642-37456-2_14.

[26] S. Kim and T. Sangsawang, "Automated identification of gait anomalies using deep autoencoder and isolation forest for hybrid anomaly detection," *Int. J. Res. Metaverse*, vol. 3, no. 1, pp. 29–45, 2026, doi: 10.47738/ijrm.v3i1.44.

[27] M. Alkhoze and M. Almasre, "Sentiment analysis of Mobile Legends Play Store reviews using support vector machine and naïve Bayes," *J. Digit. Mark. Digit. Curr.*, vol. 2, no. 4, pp. 368–389, 2025, doi: 10.47738/jdmdc.v2i4.44.

[28] R. A. M. Aljohani and A. A. Alnahdi, "Exploring football player salary prediction using random forest: Leveraging player demographics and team associations," *Int. J. Appl. Inf. Manag.*, vol. 5, no. 4, pp. 203–213, 2025, doi: 10.47738/ijaim.v5i4.115.

[29] G. Toer and G. Kim, "An empirical study on the impact of feature scaling and encoding strategies on machine learning regression pipelines," *Int. J. Informatics Inf. Syst.*, vol. 9, no. 1, pp. 242–256, 2026, doi: 10.47738/ijiis.v9i1.293.

[30] A. Iskandar, A. Suroso, I. Hermadi, and L. Prasetyo, "Time series forecasting of environmental dynamics in urban ecotourism forest using deep learning," *J. Appl. Data Sci.*, vol. 7, no. 1, pp. 101–115, 2025, doi: 10.47738/jads.v7i1.1029.

[31] M. Pratama, F. El Hakim, D. A. Syahputra, D. Dermawan, A. Asmunin, S. Nudin, and A. Nurhidayat, "Hybrid transformer–XGBoost model optimized with ant colony algorithm for early heart disease detection: A risk factor-driven and interpretable method," *J. Appl. Data Sci.*, vol. 7, no. 1, pp. 148–164, 2025, doi: 10.47738/jads.v7i1.969.