

Article

Not peer-reviewed version

Survey for Exploring Blockchain with Graph Neural Network

[Amanda Hasebe](#) *

Posted Date: 30 September 2024

doi: 10.20944/preprints202409.2362.v1

Keywords: Graph Neural Network; Graph Convolutional Networks; Graph Attention Networks; blockchain; cryptocurrency



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Survey for Exploring Blockchain with Graph Neural Network

Amanda Hasebe

Affiliation; amanda.hasebe@outlook.com

Abstract: This paper explores the use of GNN architectures, including Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), and GraphSAGE, in modeling blockchain data. A specific focus is given to their applications in cryptocurrency transaction networks, where GNNs can detect fraudulent activities and track illicit fund flows by learning from transaction patterns and network structure. Additionally, future directions for GNN research in blockchain are discussed, including scalability challenges, privacy-preserving models, and multi-chain graph analysis. Through these advancements, GNNs offer significant potential for enhancing the security, transparency, and efficiency of blockchain and cryptocurrency ecosystems. This paper also shows an example of fraud detection for Ethereum cryptocurrencies.

Keywords: Graph Neural Network; Graph Convolutional Networks; Graph Attention Networks; blockchain; cryptocurrency

Introduction

Graph Neural Networks (GNNs) offer a novel approach to analyzing data structures with inherent graph-like properties. This makes them particularly suitable for blockchain applications. Blockchain, by its nature, is a distributed ledger that can be modeled as a graph, with nodes representing transactions or blocks and edges symbolizing relationships, such as dependencies or flows between transactions. GNNs, which generalize traditional neural networks to graph-structured data, can potentially address many of the challenges faced by blockchain technology, including issues related to security, scalability, and efficiency.

Blockchain technology, which underpins cryptocurrencies like Bitcoin and Ethereum, works by maintaining a decentralized, immutable ledger. However, blockchain faces several limitations, including scalability, energy consumption, and security risks. Current consensus mechanisms, such as Proof of Work (PoW) and Proof of Stake (PoS), are resource-intensive and limit the speed and scalability of the system. Moreover, the rapid growth of blockchain networks adds complexity, making transaction verification and fraud detection more challenging. These challenges have led researchers to explore the use of GNNs, which can capture the dynamic and relational nature of blockchain data, to enhance various aspects of blockchain technology.

Review on Existing Literatures

GNN has been successful in representing complex relationships among data to be learned. Graph Neural Networks, such as Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), and GraphSAGE, enable neural network architectures to learn from graph-structured data by leveraging message-passing mechanisms. This ability to model complex relationships makes GNNs a powerful tool for blockchain applications. For instance, GNNs can be used to detect patterns of fraudulent activity, such as Sybil attacks or double spending, by learning from the structure of transaction networks [1]. Additionally, GNNs have the potential to optimize blockchain consensus mechanisms by learning from patterns in node interactions, thereby improving

the speed and efficiency of consensus protocols. Thus, GNN was able to investigate sports analytics like what Wang et al. [2] conducted.

Beyond security, GNNs also offer promising solutions to the scalability challenges of blockchain. By employing GNNs for transaction validation and block verification, blockchain systems could potentially reduce computational overhead, allowing for faster processing and greater throughput. Furthermore, GNNs can improve the execution of smart contracts by predicting outcomes based on the relationships between various contract calls [1]. This application is particularly relevant in decentralized finance (DeFi) ecosystems, where complex financial instruments are often layered on top of blockchain networks [3]. GNNs can help model and analyze these networks, enabling better prediction of financial risks and returns [4].

Methodology of GNN Learning from Blockchain

Graph Neural Networks (GNNs) learn from graph-structured data through the process of message passing, where information is propagated between neighboring nodes over the edges in a graph. In the context of blockchain, this is particularly useful because a blockchain network can be represented as a graph, where nodes correspond to entities such as transactions, blocks, or addresses, and edges represent relationships like ownership, transaction history, or dependencies between blocks. The core idea behind message passing in GNNs is to iteratively aggregate and transform information from a node's neighbors to update its representation. In models like Graph Convolutional Networks (GCNs), each node's representation is updated based on the features of its neighbors through a series of weighted aggregations, allowing the model to capture local structural patterns in the graph [5].

When applied to blockchain, GNNs can learn patterns from the connections between transactions or the flow of cryptocurrency between addresses. For instance, in a network of transactions, each transaction (node) can be connected to other transactions by edges that represent flows of assets. GNNs can iteratively learn from these connections by analyzing how transactions influence one another, detecting anomalies such as fraudulent behavior, or recognizing patterns associated with illicit activities like money laundering. For example, in decentralized finance (DeFi), GNNs can model relationships between smart contracts and their interactions, identifying potentially risky or unstable financial products by recognizing unusual patterns in contract dependencies.

Graph Attention Networks (GATs) further improve this learning process by applying attention mechanisms to weigh the importance of different neighbors. In a blockchain context, not all neighbors are equally important. For example, some transactions may be more influential than others due to higher values or certain temporal dependencies. GATs allow the network to assign different importance to each neighbor during the aggregation process, leading to more refined node embeddings [6]. This mechanism is particularly useful for filtering out noise in complex blockchain networks [7], where a transaction may be connected to many irrelevant or less informative transactions.

GraphSAGE introduces an alternative method to learn on large graphs by sampling and aggregating features from a fixed number of neighbors instead of all neighbors [8]. This makes it computationally efficient to train GNNs on large blockchain networks, which can consist of thousands or millions of nodes and edges [9]. By learning localized patterns in the network, GraphSAGE can generalize to unseen parts of the blockchain and help predict future transaction behaviors, optimize block verification processes, or even improve the scalability of consensus mechanisms [10].

Overall, GNNs allow for learning from both local and global structures in blockchain networks, making them powerful tools for a variety of tasks, such as transaction classification, fraud detection, and consensus optimization. By iteratively aggregating and transforming information from connected entities in the blockchain, GNNs can uncover hidden patterns and relationships that traditional machine learning models might miss.

Despite these potential benefits, the integration of GNNs into blockchain systems poses several challenges. The decentralized nature of blockchain, where nodes operate independently, can make it

difficult to apply traditional GNN training methods, which typically rely on centralized data. Additionally, the computational complexity of GNNs may strain the already resource-intensive blockchain infrastructure [5]. Privacy concerns also arise, as the use of GNNs might require the sharing of sensitive transaction data between nodes. Nonetheless, these challenges present opportunities for future research. For instance, advancements in distributed GNN training or the use of privacy-preserving techniques like federated learning could overcome these barriers.

In conclusion, GNNs offer significant potential for improving blockchain technology by enhancing security [11], optimizing consensus mechanisms, and addressing scalability issues. However, further research is needed to fully realize these benefits and address the challenges associated with implementing GNNs in decentralized environments. The combination of GNNs and blockchain could pave the way for more efficient, secure, and scalable distributed ledger systems in the future.

Blockchain and Fraud Detection

Graph Neural Networks (GNNs) can learn from blockchain data by leveraging the underlying graph structure of the blockchain, which captures both individual transaction features and their relationships. For example, blockchain transactions can be represented as a directed graph, where each node corresponds to a transaction, and edges represent the flow of assets or dependencies between transactions. GNNs apply a message-passing mechanism that iteratively updates each node's representation by aggregating information from its neighboring nodes. Through this process, GNNs can detect patterns and relationships across the blockchain network, which are often critical for identifying fraud, optimizing consensus mechanisms, and improving scalability.

A specific application of GNN learning in blockchain is fraud detection in cryptocurrency transaction networks. In this case, each transaction is represented as a node with features such as transaction amount, sender, and receiver. Transactions are connected by edges that represent the flow of cryptocurrency between them. The GNN can be trained on historical blockchain data labeled as either legitimate or fraudulent [12]. During training, the GNN updates its node representations based on the relationships between transactions and their surrounding context in the graph. This allows the network to capture patterns that are commonly associated with fraudulent activities, such as certain transaction chains leading to known fraudulent addresses or cyclical transaction flows often linked to money laundering schemes. Once trained, the GNN can predict the likelihood of fraud for new, unseen transactions by analyzing their position and connections within the blockchain graph.

For example, consider a Ponzi scheme on a blockchain network where funds are funneled through a series of transactions to an exit scam. A GNN can learn from labeled data by analyzing how transactions interact with flagged fraudulent wallets and by identifying suspicious transaction paths, even when individual transactions may not have obvious red flags on their own. The relational aspect of GNNs, which allows them to take into account both the local and global structure of the blockchain network, makes them particularly suitable for uncovering such hidden patterns that traditional methods might miss.

Cryptocurrency transaction networks provide a particularly rich environment for the application of Graph Neural Networks (GNNs). These networks are inherently structured as directed graphs, where nodes represent entities such as wallet addresses or transactions, and edges signify the transfer of cryptocurrency between these entities. The transactional relationships in these networks exhibit complex, dynamic patterns influenced by user behavior, economic factors, and illicit activities such as money laundering or fraud. GNNs are well-suited to model these complex relationships because they can learn from both the transactional features and the topological structure of the network.

In cryptocurrency transaction networks, GNNs can be used to detect various patterns of malicious behavior. For instance, transactions associated with phishing schemes, Ponzi schemes, or money laundering typically follow identifiable patterns that can be captured by GNNs. One common approach is to model the network such that wallet addresses or transactions are nodes, and edges represent cryptocurrency transfers between them. By applying GNNs to this structure, the model can

learn patterns such as repeated interactions between a set of addresses or cyclical transaction flows, which are often indicative of fraudulent activities. This is an advantage over traditional methods that may only consider the attributes of individual transactions without factoring in their relational context within the network. Tan et al. [13] presented an example of GNN's fraud detection ability on Ethereum cryptocurrency.

GNNs are also valuable for tracking the flow of funds across multiple wallets. A specific challenge in cryptocurrency networks is the existence of mixers or tumblers, which are designed to obfuscate the trail of cryptocurrency by splitting funds into smaller parts and distributing them across many wallets. GNNs can assist in identifying such laundering schemes by analyzing the graph of transactions and recognizing the spread of funds across many wallets that eventually reconverge into one or a few key addresses. This capability is especially relevant for law enforcement and regulatory agencies that monitor cryptocurrency transactions for compliance with anti-money laundering (AML) regulations.

Moreover, cryptocurrency transaction networks evolve rapidly as new transactions are continuously added. This dynamic nature can be effectively modeled with GNN architectures such as GraphSAGE, which are designed to handle large, evolving graphs by sampling a fixed number of neighbors during each message-passing iteration. As new transactions occur, these models can continuously learn from the growing network, making them particularly suited for real-time applications such as fraud detection in live cryptocurrency systems. Cryptocurrency transactions require speedy detection, which gives opportunity for GNN's inception to fraud detection.

Future Directions and Conclusion

In conclusion, the integration of Graph Neural Networks (GNNs) into blockchain technology represents a promising frontier for addressing many of the challenges inherent to decentralized systems. GNNs' ability to model complex relationships within graph-structured data allows them to capture both local and global patterns in blockchain networks, which can be critical for applications like fraud detection, consensus optimization, and scalability improvements. By iteratively aggregating information from nodes (e.g., transactions or blocks) and their neighboring connections, GNNs provide insights that traditional machine learning models cannot achieve when working with non-graph data. This capability makes GNNs particularly powerful for identifying hidden, large-scale patterns in blockchain networks, such as detecting fraudulent transactions, optimizing transaction validation, and improving the performance of decentralized finance (DeFi) platforms.

Looking toward future directions, several challenges and research opportunities arise. First, the computational complexity of training GNNs on large, decentralized blockchain networks remains a significant hurdle. Blockchain networks like Bitcoin and Ethereum consist of millions of nodes and edges, making GNN training computationally expensive. Research into more efficient GNN architectures, such as GraphSAGE or scalable attention mechanisms, could help address this issue. Another area of research is the development of privacy-preserving GNN models. Blockchain's decentralized nature often means that full access to data across nodes is not possible, which complicates GNN training. Techniques such as federated learning or homomorphic encryption could be explored to enable GNNs to learn from blockchain data while preserving the privacy of individual participants.

Additionally, more work is needed to apply GNNs beyond fraud detection, such as in optimizing consensus mechanisms, improving smart contract execution, and analyzing decentralized finance systems. Combining GNNs with other types of machine learning models, such as reinforcement learning or graph-based reinforcement learning, might unlock new possibilities for more intelligent and adaptive blockchain systems. Finally, as blockchain ecosystems become more complex and multi-chain systems (i.e., interoperability between different blockchains) become common, GNNs can be extended to analyze multi-graph environments. These advancements can lead to more secure, efficient, and scalable blockchain architectures in the future.

References

- [1] S. Thudi, P. Sharma and A. Dutta, "Implementing Machine Learning on Blockchains for Transactional Privacy," *Journal of Distributed Computing*, 2023.
- [2] Z. Wang, Y. Zhu, Z. Li, Z. Wang, H. Qin and X. Liu, "Graph neural network recommendation system for football formation," *Applied Science and Biotechnology Journal for Advanced Research*, vol. 3, p. 33–39, 2024.
- [3] J. Shen, J. Zhou, Y. Xie, S. Yu and Q. Xuan, "Identity inference on blockchain using graph neural network," in *Blockchain and Trustworthy Systems: Third International Conference, BlockSys 2021, Guangzhou, China, August 5–6, 2021, Revised Selected Papers 3*, 2021.
- [4] A. Sharma, P. K. Singh, E. Podoplelova, V. Gavrilenko, A. Tselykh and A. Bozhenyuk, "Graph Neural Network-Based Anomaly Detection in Blockchain Network," in *International Conference on Computing, Communications, and Cyber-Security*, 2022.
- [5] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.
- [6] P. Veličković, "Message passing all the way up," *arXiv preprint arXiv:2202.11097*, 2022.
- [7] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio and Y. Bengio, "Graph Attention Networks," *arXiv preprint arXiv:1710.10903*, 2018.
- [8] W. Hamilton, Z. Ying and J. Leskovec, "Inductive Representation Learning on Large Graphs," in *Advances in Neural Information Processing Systems*, 2017.
- [9] X. Song, R. Ma, J. Li, M. Zhang and D. P. Wipf, "Network in graph neural network," *arXiv preprint arXiv:2111.11638*, 2021.
- [10] W. W. Lo, S. Layeghy, M. Sarhan, M. Gallagher and M. Portmann, "E-graphsage: A graph neural network based intrusion detection system for iot," in *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*, 2022.
- [11] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," *Manuscript*, 2008, 2008.

- [12] S. Zhang, L. Qian, Z. Zhang and Y. Guo, "Deep Learning for Blockchain: A Data-Driven Approach to Blockchain Transaction Analysis," *arXiv preprint arXiv:1906.08280*, 2019.
- [13] R. Tan, Q. Tan, P. Zhang and Z. Li, "Graph neural network for ethereum fraud detection," in *2021 IEEE international conference on big knowledge (ICBK)*, 2021.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.