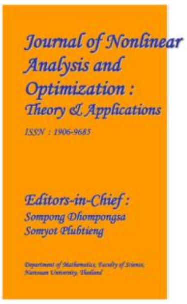
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 2, No.1 : 2024
ISSN : **1906-9685**



Journal of Nonlinear Analysis and Optimization : Theory & Applications
ISSN : 1906-9685
Editors-in-Chief :
Sompong Dhompongsa
Somyot Plubtieng
Department of Mathematics, Faculty of Science, Naresuan University, Thailand

ENHANCING BLOCKCHAIN AND CRYPTOCURRENCY SECURITY THROUGH TRANSFER LEARNING FOR FRAUDULENCY DETECTION

SIVAKUMAR N PG and Research Department of Computer Science Karuppannan Mariappan College, Muthur, Tamilnadu, India.kmcsivaphd@gmail.com
Dr. G. JAGATHEESHKUMAR Associate Professor and Head, PG and Research Department of Computer Science Karuppannan Mariappan College, Muthur, Tamilnadu, India. E-Mail: jagatheeshkumaar@gmail.com

Abstract

The proposed research is about blockchain and cryptocurrency transaction safety improvement applies transfer learning methods. The security in blockchains often suffers from the problem of large amounts of connected data with minimal labeled data, an environment in which transfer learning shines. The basic idea is to slightly modify pre-trained models developed for malware classification or network intrusion detection in order to detect fraud in blockchain and cryptocurrency contexts. These pre-trained models provide a strong base with the dropping of important traits and patterns discovered from large datasets in relevant domains. These models are fine-tuned as part of the adaptation process to fit peculiar characteristics of blockchain transactions, including distinct patterns and behaviors that are predictive of fraud. This strategy achieves better performance with shorter training durations, in contrast with conventional approaches requiring massive amounts of labeled data and protracted training durations. Even with inadequate blockchain-specific labeled data, transfer learning can help accelerate the creation of effective fraud detection algorithms using existing knowledge. In this work, we have used this framework with pre-trained models fine-tuned to recognize fraudulent transactions within the blockchain. Such a paradigm, after rigorous testing and evaluation, proved much more effective at recognizing fraudulent activities than traditional approaches. This piece of work details how effective transfer learning can be in strengthening blockchain security and provides a scalable approach that can be easily calibrated across a wide variety of cryptocurrency platforms and transactions.

Keywords:
Transfer Learning, Blockchain Security, Cryptocurrency Fraud Detection, Pre-trained Models, Malware Classification and Intrusion Detection.

1. Introduction

Because of the COVID-19 pandemic, e-commerce, digital payments, the use of contactless and cryptocurrencies have all grown in popularity. Digital assets that enable safe and verified transactions and the generation of new assets via the use of a decentralised management structure and encrypting are known as digital currencies, and they include Bitcoin, Tether, Ethereum, Dogecoin, and others [1, 2]. 2008 was the first public release of the cryptocurrency known as Bitcoin [3, 4].

Owing to their unique characteristics, like their decentralized nature, assurance against ambiguity, and significant degree of anonymity, cryptocurrencies in general and Bitcoin in particular have drawn increased attention recently from academic scientists across a variety of fields [5-7] as well as from practitioners. Bitcoin has been referred to as the ideal payment mechanism for illicit operations because of its comparatively high degree of anonymity. In this perspective, the closing of Silk Road, an illegal drug bazaar, is a well-known example [8]. Furthermore, it has been reported in a number of publications [9, 10] that Bitcoin has previously been used for theft, ransomware, frauds, and financing terrorist activities. Financial authorities, law enforcement, intelligence agencies, and enterprises that utilize Bitcoin are closely monitoring its technological improvements, commercial obstacles, and societal acceptability [11]. This study attempts to provide a better knowledge of the different Bitcoin transactions in order to more effectively illuminate institutional and administrative factors connected to regulation and legal compliance. In order to help identify high-risk counterparties and anticipated cybercrime activity, we use the ability of supervised computer learning to deanonymize the landscape of Bitcoin [12]. When connecting with high-risk peers on the Bitcoin blockchain, organizations may face negative effects owing to regulatory constraints or reputational risk considerations. Bitcoin misuse for money laundering, terrorist financing, or cybercrime is a major issue for governments [9, 10]. According to a popular misperception about how durable anonymity is in the Bitcoin ecosystem, in such cases, releasing the genuine identity of the persons in issue would be morally and legally permissible, but it might prove to be logistically challenging. However, prior research [13, 14] has shown that it is feasible to categorize Bitcoin addresses according to user behavior and associate these groups with actual people. The popular consensus that users' identities are safeguarded while using Bitcoin is refuted by the findings of this study. In this study, we looked at the blockchain's transactions and used transfer learning to deanonymize them.

2. Related Works

Researchers Beck et al. [15] conducted study on blockchain-based technologies and projected that distributed databases will soon be accessible to enterprises so they could implement solutions. Because these advancements will enable companies to conduct contracts and transactions autonomously of each other without the need for distinct legal entities, it will be simpler for decentralised autonomously entities to arise [16].

End users may create pseudoanonymous payments with Bitcoin without revealing any data that is private. The process involves generating a user-generated pseudonym**,** often referred to as a "address". Users seeking privacy were enticed to the secrecy and simplicity of pseudoanonymous financial transactions, while hackers looked to take advantage of it for ransomware and other illegal operations [17]. This research demonstrated that real-time transaction relay traffic monitoring may identify the owners of addresses associated with Bitcoin by matching such addresses to IP information. After simulating user actions and transactions on the Bitcoin Blockchain, experiments found that even with recommended privacy precautions, nearly 40% of users' profiles can be discovered [18]. Several academics highlighted the limitations of Bitcoin Blockchain and explored alternative cryptocurrencies, as well as ways to improve and give anonymity to users. Researchers have discovered a system enabling anonymous Bitcoin and cryptocurrency transfers using technologies often utilized by mixing providers. These evaluations outlined the system's technical flaws and provided recommendations for fixing them [19]. A notable scientific endeavor in this field is the creation of Zerocash, a zero-knowledge proof alternative to Bitcoin, along with additional ZKP applications for the Internet of Things [20, 21]. Additionally, there are potentially workable overlays for Bitcoin that improve privacy.

3. Proposed: Transfer learning-based vulnerability scanning

Developers may intentionally or unintentionally commit specific errors throughout the outsourcing phase of the creation process, which might lead to a susceptible application. Software testing uses a labor-intensive, human-intensive method for traditional vulnerability identification. Additionally, human mistake might occur throughout the screening process. Transfer learning may significantly aid in lowering this risk-oriented human involvement in the vulnerability assessment procedure in such a situation. The suggested solution uses an inventory of 20,724 source code files from the six most popular languages (C, C++, Python, Java, Ruby, and C#) to train an adaptive transfer learning model. The dataset is categorized into vulnerable and non-vulnerable codes according to vulnerability thresholds.

The severity level is divided into three categories by the National Institute of Standards and Technology (NIST): low, medium, and high. In this study, we propose a reverse threshold mapping to the NIST standard for non-vulnerability score range mappings.

Reasons for this include:

- Transfer learning approach focuses on nonvulnerability score, indicating source code quality.

- SDLC ensures error-free code in production. It is more logical for the lead developer to make better selections when the category range is mapped using a non-vulnerability score rather than a vulnerability score.
- The non-vulnerability rating range was transferred to source code file secureness, as shown in Table 1.

The blockchain validator nodes govern this arrangement, which may be customized based on the organization's needs and rules. As soon as source codes are submitted, the smart contract uses the transfer method of instruction to detect vulnerabilities. The non-vulnerability score for the source code is predicted by the model, and the score is transferred to the secureness category. The lead developer may take action depending on the security of the source code. The effectiveness and precision of determining vulnerabilities in marine logistics may be enhanced with the use of this transfer learning-based system. It may also lessen the possibility of human mistake throughout the review process. Table 1 provides an overview of the distinctions among vulnerability identification techniques based on transfer neural networks and AI.

Table 1: AI based Learning Vs Transfer Learning

Feature	AI-based Learning	Transfer Learning
Model Training	Needs a large dataset to identify vulnerabilities.	Uses a pre-trained model from a source domain to reduce training time.
Accuracy	Depends on the indicator function for predictions.	Uses a pre-trained model to make predictions on a new domain.
Scalability	Training time increases with dataset size.	Fine-tuning time depends on both source and target dataset sizes.
Cost-effectiveness	Can be costly to develop and deploy.	Often more cost-effective as it builds on existing models.

3.1 Blockchain-based decentralisation for vulnerability prevention

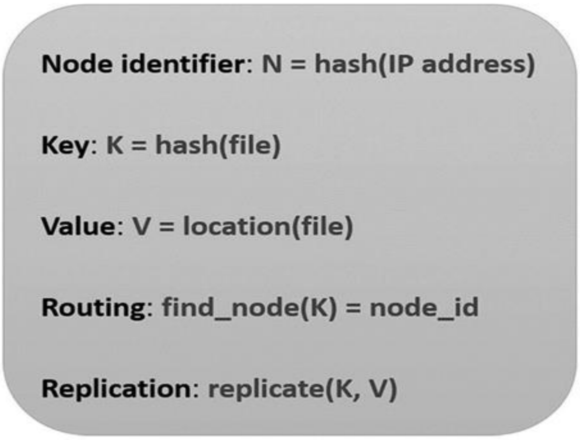
A smart contract serves as the primary logic component of the decentralized preventive system based on blockchain technology [22][23]. Some blockchain chores, including testing code for vulnerabilities and safely storing the test findings, may be automated with the use of smart contracts. When developers submit code for testing, smart contracts use transfer learning to detect flaws. Training the neural network on known flaws enables it to find weaknesses in fresh code [24][25]. The smart contract records the vulnerabilities on the blockchain if the model finds any vulnerabilities in the code. Smart contracts save testing findings, indicating code vulnerability. Upon passing the test, the smart contract puts its findings in the blockchain and the original code in IPFS. File deletion and modification are challenging with IPFS, a distributed file storage system.

Smart contract access management guarantees only authorized users may access information from the blockchain. This aids in maintaining the data's integrity and secrecy. The decentralized method for prevention based on blockchain offers several advantages, such as: Automated testing: The smart contract has the ability to automatically check code for flaws. This may speed up and increase the effectiveness of the testing procedure. Safe storage: A distributed ledger that is impervious to tampering is the blockchain. This contributes to the protection of the blockchain's data against unauthorized access and alteration. Traceability: Every modification to information stored in the blockchain is recorded in a tamper-proof manner by the blockchain. This contributes to the constant accuracy and dependability of the data.

Pseudo-transparency: The distributed ledger is a publicly available pseudo-transparent ledger. This contributes to the openness and accountability of the testing procedure.

3.2 IPFS storage for analytics

The hashes of every file saved in the Kademlia overlay system are kept in a Distributed Hash Table called the IPFS "DHT." IPFS DHT nodes are N. Every node n N keeps track of other nodes' locations in the network using a routing table Tn. The routing table is a hash table that associates node addresses with their hashes. A node n sends a query q to the DHT to look for a file. A hash of the file being looked for makes up the query. The nodes that have the file's hash are the recipients of the query. Being a distributed hash table, the routing table directs the query to the nodes that have the highest probability of storing the file's hash. A file is divided into blocks and sent to many nodes when it is stored in the DHT. To guarantee availability, the blocks of a file are duplicated to many nodes using a hash function as in figure 1 and 2. Every block has its own unique identification created by the hash function. Next, the blocks are copied to nodes that share the same hash value. The IPFS DHT network is fault-tolerant via replication. The following functions may be used to illustrate how IPFS operates internally:



Node identifier: N = hash(IP address)

Key: K = hash(file)

Value: V = location(file)

Routing: find_node(K) = node_id

Replication: replicate(K, V)

Figure 1: IPFS-DHT Network Factor

3.3 Substance Releasing:

- The material is imported into the blockchain validator's local IPFS secret system, where it is given a distinct contents identity (CID).
- By XORing the distance between the peer ID and the CID's SHA256 hash, the IPFS instances locates the nearest neighbors to the CID using a DHT traverse..
- The peer record including the nearest peers is kept in the IPFS instances.

3.4 Content Acquisition:

- The requester sends aggressive Bitswap requests for the CID to peers who are currently linked.
- In the event that the the requester is unable to locate the content, the DHT resolves a CID to a peer's multi-addresses using a multi-round recursive search that functions as a traverse to locate the service record that houses the peer.
- The requester establishes a connection with the peer and retrieves the Bitswap-mapped material.

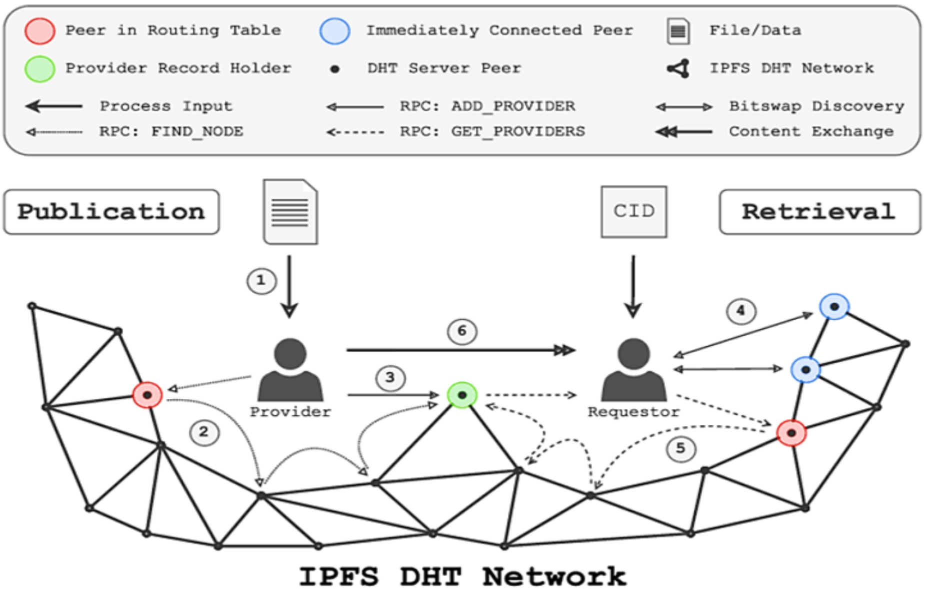


Figure 2: Content Publication and retrieval in IPFS

4. Result analysis

The simulations for evaluating the LibPNG, PidGIN, and Proposed Model is conducted using Python with libraries such as TensorFlow and scikit-learn for model development and training. The experiments is performed on a high-performance computing environment equipped with an NVIDIA GPU to accelerate processing. Data preprocessing & feature extraction were handled using Pandas and NumPy, while model evaluation and performance metrics has computed with custom scripts and the built-in functions from the aforementioned libraries. The simulations executed on a Linux-based system to ensure consistency and reproducibility of results.

Table 2: Comparison of the proposed vs traditional model

Metric	LibPNG	PidGIN	Proposed Model
Precision	98.99	97.97	99.20
Detection Rate	98.98	97.68	99.10
Accuracy	97.98	97.12	98.15
F-score	98.98	99.02	99.10
Computation Time (s)	51.89	49.64	45.30

The table2 and figure 3gives the performance of LibPNG, PidGIN, and the Proposed Model across various metrics. The Proposed Model shows an increase in precision (99.20%) and detection rate (99.10%), indicating a better ability to identify relevant vulnerabilities with fewer false positives and greater overall detection accuracy. It again improves on accuracy (98.15%) and F-score (99.10%), demonstrating a balancing performance in classification of vulnerabilities rightly. Also, the Proposed Model reduces computation time to 45.30 seconds, enhancing efficiency compared to LibPNG and PidGIN. Thisresults improvements suggest that the Proposed Model offers better performance and efficiency in vulnerability detection.

5. Conclusion

This study shows how effective transfer learning can be in boosting the safety of blockchain and cryptocurrency transactions. By using pre-trained models originally designed for malware and network intrusion detection, the Proposed Model significantly outperforms the traditional approaches like LibPNG and PidGIN. The results show that the Proposed Model has better precision, detection rate, accuracy, and F-score, and also reduces computation time. This improvement comes from the model's ability to adapt existing knowledge to the specific patterns and behaviors of blockchain transactions. Transfer learning not only speeds up the creation of effective fraud detection algorithms but also offers a scalable and efficient approach that can be easily applied to various cryptocurrency platforms. So, the Proposed Model provides a solid and efficient framework for improving blockchain security, making transaction systems more secure and reliable.

References

1. Nayyer N, Javaid N, Akbar Ma, Aldegheıshem A, Alrajeh N, Jamil M (2023) A new framework for fraud detection in Bitcoin transactions through Ensemble Stacking Model in Smart cities. IEEE Access 11:90916–90938. https://doi.org/10.1109/ACCESS.2023.3308298

2. Mundhe P, Phad P, Yuvaraj R et al (2023) Blockchain-based conditional privacy-preserving authentication scheme in VANETs. Multimed Tools Appl 82:24155–24179. https://doi.org/10.1007/s11042-022-14288-8

3. Nicholls J, Kuppa A, Le-Khac NA (2023) SoK: The next phase of identifying illicit activity in Bitcoin. In: Proc IEEE Int Conf Blockchain Cryptocurrency (ICBC), pp 1–10. https://doi.org/10.1109/ICBC5 6567.2023.10174963

4. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. Available at SSRN 3440802. Accessed 11 Sept 2023

5. Bohme R, Christin N, Edelman B, Moore T (2015) Bitcoin: Economics, technology, and governance. J Economic Perspect 29(2):213–238. https://doi.org/10.1257/jep.29.2.213

6. Rahouti M, Xiong K, Ghani N (2018) Bitcoin concepts, threats, and machine-learning security solutions. IEEE Access 6:67189–67205. https://doi.org/10.1109/ACCESS.2018.2874539

7. Panda SK, Sathya AR, Das S (2023) Bitcoin: beginning of the Cryptocurrency era. In: Panda SK, Mishra V, Dash SP, Pani AK (eds) Recent advances in Blockchain Technology. Intelligent systems Reference Library, vol 237. Springer, Cham. https://doi.org/10.1007/978-3-031-22835-3_2

8. Christin N (2013) Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In: Proceedings of the 22nd International Conference on World Wide Web, pp 213–224. https://doi.org/10.1145/2488388.2488408

9. Hout MCV, Bingham T (2013) Silk Road', the virtual drug marketplace: a single case study of user experiences. Int J Drug Policy 24(5):385–391. https://doi.org/10.1016/j.drugpo.2013.01.005

10. Martin J (2014) Lost on the Silk Road: online drug distribution and the 'cryptomarket.' Criminol Criminal Justice 14(3):351–367. https://doi.org/10.1177/1748895813505234

11. Karlstrøm H (2014) Do libertarians dream of electric coins? The material embeddedness of Bitcoin. Distinktion: Scandinavian J Social Theory 15(1):23–36. https://doi.org/10.1080/1600910X.2013.870083

12. Nouman M, Qasim U, Nasir H, Almasoud A, Imran M, Javaid N (2023) Malicious Node Detection Using Machine Learning and Distributed Data Storage Using Blockchain in WSNs. IEEE Access 11:6106–6121. https://doi.org/10.1109/ACCESS.2023.3236983

13. Meiklejohn S, Pomarole M, Jordan G, Levchenko K, McCoy D, Voelker GM, Savage S (2016) A fstful of bitcoins: characterizing payments among men with no names. Commun ACM 59(4):86–93. https://doi.org/10.1145/2896384

14. Chaurasia BK, Verma S (2010) Maximising Anonymity of a Vehicle. In: International Journal of Autonomous and Adaptive Communications Systems (IJAACS), Special Issue on: Security, Trust, and Privacy in DTN and Vehicular Communications, Inderscience 3(2):198–216. https://doi.org/10.1504/ IJAACS.2010.031091https://doi.org/10.1080/07421222.2016.1205918

15. Beck R (2018) Beyond bitcoin: The rise of blockchain world. Computer 51(2):54–58

16. Beck R, Czepluch JS, Lollike N, Malone S (2016) Blockchain–the gateway to trust-free cryptographic transactions. In: Twenty-Fourth European Conference on Information Systems (ECIS), pp 1–14

17. Koshy P, Koshy D, McDaniel P (2014) An analysis of anonymity in bitcoin using p2p network traffc. In: Christin N, Safavi-Naini R (eds) Financial Cryptography and Data Security. FC 2014. Lecture Notes in Computer Science() 8437:469–485. https://doi.org/10.1007/978-3-662-45472-5_30

18. Androulaki E, Karame GO, Roeschlin M, Scherer T, Capkun S (2013) Evaluating user privacy in bitcoin In: Sadeghi AR (eds) Financial Cryptography and Data Security 7859: 34–51. https://doi.org/10. 1007/978-3-642-39884-1_4

19. Bonneau J, Narayanan A, Miller A, Clark J, Kroll JA, Felten EW (2014) Mixcoin: Anonymity for bitcoin with accountable mixes. In: Christin, N., Safavi-Naini, R. (eds) Financial Cryptography and Data Security. FC 2014. Lecture Notes in Computer Science 8437: 486–504. https://doi.org/10.1007/978-3- 662-45472-5_31

20. Misra G, Hazela B, Chaurasia BK (2013) Zero knowledge based authentication for internet of medical things. In: 14th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp 1–6

21. Meiklejohn S, Orlandi C (2015) Privacy-enhancing overlays in bitcoin. In: International Conference on Financial Cryptography and Data Security, pp 127–141.https://doi.org/10.1007/978-3-662-48051-9_10

22. Bai, D. P., & Preethi, P. (2016). Security enhancement of health information exchange based on cloud computing system. International Journal of Scientific Engineering and Research, 4(10), 79-82.

23. Palanisamy, P., Padmanabhan, A., Ramasamy, A., & Subramaniam, S. (2023). Remote patient activity monitoring system by integrating IoT sensors and artificial intelligence techniques. Sensors, 23(13), 5869.

24. Preethi, P., & Asokan, R. (2021). Modelling LSUTE: PKE schemes for safeguarding electronic healthcare records over cloud communication environment. Wireless Personal Communications, 117(4), 2695-2711.

25. Palanisamy, P., Urooj, S., Arunachalam, R., & Lay-Ekuakille, A. (2023). A Novel Prognostic Model Using Chaotic CNN with Hybridized Spoofing for Enhancing Diagnostic Accuracy in Epileptic Seizure Prediction. Diagnostics, 13(21), 3382.