Semantic Graph Analysis to Combat Cryptocurrency Misinformation on the Web*

Daniel Kazenoff[1 [0000−0002−8327−4853]], Oshani Seneviratne[1 [0000−0001−8518−917X]], and Deborah L. McGuinness[1 [0000−0001−7037−4567]]

Rensselaer Polytechnic Institute, Troy, NY, USA
kazendrpi@gmail.com, senevo@rpi.edu, dlm@cs.rpi.edu

Abstract. With the hype around blockchain technologies, misinformation on 'get rich quick' scams are becoming rampant. In this work, we describe a solution that puts in the groundwork to identify fraudulent users and track them across multiple blockchains using semantic modeling. The application of Semantic Web and Linked Data technologies provides a well-grounded solution to connecting fragmented but conceptually linked resources. This paper focuses on showing that through the integration of ontology-driven knowledge graphs and a queryable graph database, a novel off-chain protocol utilizing comprehensive cross-chain integration techniques can be used to link an identity across multiple blockchains, and provide a significantly enhanced foundation for provenance data analysis for scam activity detection. This foundation could help reduce the challenges users face as they try to safely and effectively navigate the decentralized cryptocurrency financial ecosystem.

Keywords: provenance, resource linkage, blockchain, scams

1 Introduction

As it exists today, blockchain still carries much mystique and uncertainty even though the first large-scale implementation of the technology is over a decade old, with Bitcoin introducing itself to the world in 2009. Undoubtedly, blockchain as a technology aims to be the harbinger of a new era of secure, trustless transactions that do not require a third-party facilitator or arbitrator. In a nutshell, a global ledger of immutable, digital records stored in virtual "blocks" is maintained by distributed, independent computers. These independent nodes work together to maintain the global ledger of transactions, which is uniquely fault-tolerant and decentralized to establish trust among all participants in the network. The digital currencies we have analyzed as part of this work are Bitcoin (BTC), Ethereum (ETH), and Bitcoin Cash (BCH).

While major blockchain technologies pride themselves on their fault-tolerant, secure nature at their core, the structures built around or on top of these blockchains are exploited daily and prone to a multitude of attack vectors. The entry points to these decentralized ledger technologies are usually via the Web. Therefore, the same misinformation challenges that exist on the Web apply to users who want to understand and invest in these nascent technologies.

For a technology still in its infancy with such vast potential, its key drawbacks must be identified, and methods for extracting more information about blockchain activities need to be developed. Such disadvantages include high fragmentation across blockchains, and investigative difficulties when dealing with increasing instances of crypto exchange hacks, and other fraudulent activities that are fueled by misinformation on the Web. We will introduce some of the blockchain's growing pains, particularly fraudulent scams, and fragmentation in the blockchain ecosystem.

Phishing scams are nothing new, but have unique advantages when being operated on decentralized currencies. Digitally transacted US currency must be tied to a bank account, which is tied to one's identity, and is certain to pass under the watchful eye of a regulatory agency. Conversely, many digital currencies are quite easy to operate anonymously. Many were intentionally built to make tracking the flow of currency nearly impossible. The inputs and resulting outputs from any given Bitcoin (BTC)[1] transaction can consist of multiple wallet addresses not necessarily related to the account of the original sender of the BTC. The effects of this are that malicious actors can pose as whatever convincing entity they want and lure people into traps that continually evolve in complexity, causing people to send large amounts of currency (often BTC) to malicious addresses. Video re-runs of a famous tech celebrity giving a speech on cryptocurrencies, with QR codes of malicious receiver addresses and enticing rewards or other social media posts, are used as baits. More recently, there was a Twitter hack where the hackers took over many high profile accounts in a cryptocurrency scam [1]. This is a newer type of operation identified in Phillips and Wilder's discussion of cryptocurrency scams [2]. Their discussion lays bare the need for fiat-accepting exchanges, i.e., exchanges that let users buy cryptocurrencies with central, government-controlled currencies, such as the US dollar, euro, to have preventative tools to identify these scam addresses (and potentially all accounts directly associated accounts), given that exchanges are the most popular destination for these illicitly obtained funds.

No widespread or easily accessible solutions exist for maintaining or viewing interoperability between disparate blockchains. For example, if Alice held one BTC and wanted to complete an expensive purchase from Bob using an Ethereum smart contract that governs the terms of the transactions in a computable manner, she would not have an easy way to use her BTC for this purchase without going through an exchange. Such a centralized exchange requires the trust of its users, and do not align with the decentralized ecosystems that lets mutually distrusting parties to interact. The primary solution now is to rely

[1] `https://bitcoin.org`

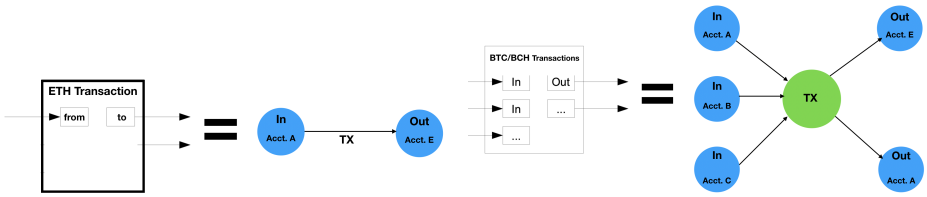
on a crypto exchange on the Web to take Alice's BTC, and swap in an equivalent BTC to ETH value from the exchanges' store of multiple currencies. Alice would have no way of viewing how the flow of her BTC eventually made it to Bob's wallet that has ETH tokens. Moreover, the exchange rate value is the key metric that could be manipulated by the third party exchange to its advantage, reducing the real value of ETH that Alice's BTC could be worth. There is no way to prove if indeed a given exchange rate was entirely fair, and worst of all, a third-party intermediary controlled the transaction between Alice and Bob. In other cases, accounts operated by scammers can target vulnerabilities in the underlying blockchains themselves. These situations are reminiscent of the DAO hack [3], which caused the original hard-fork of the Ethereum network, resulting in two disparate Ethereum ecosystems. While rare, it is still prudent to develop a resource that can adequately mark these suspicious addresses in all cases for future whitelisting, such that no legitimate exchange of value could allow them to realize their illicitly obtained values.

2 Integrative Blockchain Provenance Analyzer

Given the blockchain's growing pains identified above, we borrow techniques from decades of semantic web research to link the fraudulent accounts on various platforms to better inform users. Specifically, we utilize the provenance ontology (Prov-O) in our work. Prov-O's strict, organized, yet broad schema makes it a well-suited candidate for converting data from different blockchains, as well as from the Web, toward a common representation that can be interpreted by any system.

The function of the "Integrative Blockchain Provenance Analyzer" (IBPA) is to graphically define cross-chain information from the perspective of any given user (or potential user) of a set of cryptographic currencies. Most crucially, the IBPA runs tests on those inputted wallet-addresses to assign a Pass/Fail score to addresses that match fraudulent criteria, which enables any user suspecting a crypto scam to ascertain if it infact is a scam. Additionally, the IBPA is chain-agnostic - meaning that should a scam operation accept ETH, BTC, BCH, or other currencies, the data abstraction performed by the IBPA would not be impacted in any way and would yield similar results. However, the diversification of crypto holdings poses a challenge for cohesive user analysis. Conventionally, this would be managed by a centralized exchange behind the scenes (with or without graphical representation). Therefore, identifying trends and interactions of users of these different currencies is left up to the exchange itself, with the near-zero ability for everyday users to analyze other cryptocurrency accounts themselves. The IBPA puts in the groundwork needed to "link" addresses, accounts, and transaction nodes together on a graph instance by associating provided wallet addresses to a single "identifier" property. The identifier could be a URI for a user ("Alice," as an example), and the wallet addresses Alice has (or presumed to be Alice's) across her different cryptocurrency holdings are then associated to her user node on the graph using Prov-O constructs such

as `prov:wasAssociatedWith`. These public wallet addresses are all that is required to access the transactional data that will be formatted and converted to the queryable, targeted graph instance of cross-chain interactions by the IBPA. The graph instance can be targeted and analyzed with all the built-in querying capabilities of Neo4j and its external Awesome Procedures On Cypher (APOC) library.[2] Anyone can run these queries themselves, so the accessibility of cross-chain analysis outside of private exchanges is achieved. Additionally, the IBPA interface is extensible, such that additional functionalities could be quickly built with more 1-click queries that display analytical results to the interface. Multiple cross-chain currencies could be added to the interface as well, allowing for a comprehensive transactional profile to be built from a series of crypto addresses connected to a single user.



(a) Ethereum Transactions (b) Bitcoin Transactions

Fig. 1: Representing Blockchain Transactions as a Graph

2.1 Directed Graph Network of Blockchain Transactions

The IBPA intends to represent user nodes in a fashion that isolates raw transaction (TX) data into its simplest components necessary for straightforward provenance analysis. Not all cryptocurrencies use a standard method of representing transactions in their respective systems. The usage of Prov-O in IBPA fills this void as it will now enable a common vocabulary to analyze any scam activities either within or across various cryptocurrencies. Our implementation currently supports Ethereum and Bitcoin, and could be extended to support other cryptocurrency schemes. For instance, the Ethereum network provides "to" and "from" addresses for each transaction and contract call posted on the network, that can be modeled using `prov:hadRole`. It is easy to draw a 1-degree connection from one Ethereum user account to the next because the raw TX data is available due to Ethereum's "account-based scheme." Despite every Ethereum transaction occurring as its own discrete unit, consisting of an input and output account address, schema can be transformed into a less complicated graphical format by removing the "TX" intermediary node for this cross-chain provenance analysis (see fig. 1a). A similar method can be applied to the Bitcoin network

[2] `https://neo4j.com/labs/apoc`

with its UTXO scheme (Unspent Transaction Output schema), but it is a bit more complicated. Fig. 1b shows how BTC and BCH transactions are split and interpreted in our graph. In both figures, the right side is the resulting graph representation for each respective currency. Additionally, the directed-acyclic nature is desired to reduce the inevitable redundancies seen in blockchain transactions. Instead of creating hundreds of new user nodes over the lifetime of an account, a single node and relationship can be drawn between repeated patterns of transactions, which dramatically reduces the need for extraneous data ingestion, thus enhancing the speed of the queries in question. The IBPA can be run successfully to determine a suspicious wallet address on an average laptop due to the high efficiency of how the raw chain data is processed.

2.2 Determination Classifier

Several tests are performed on each address entered into the IBPA to classify if a particular address is a scam address or not. Upon various criteria being met, an integer score variable is automatically incremented, using a combination of visual and quantitative classification methods. Normal crypto accounts (chosen at random from a block explorer) when processed by the IBPA into a graph instance resemble a structure similar to fig 2. Here, a single user "Alice", is passing currency back and forth between other users and herself over a long period. However, accounts that are carrying out phishing scams resemble, with a striking similarity, to the graphs shown in fig 3. In this figure, now Alice is the fraudulent actor, luring in unsuspecting holders of crypto as part of the scam. Please note that this is a real depiction of one of the accounts that carried out a BTC scam on the Web. Specifically, the number of incoming relationships, outgoing relationships, and the relationship between those two values are analyzed by the IBPA. Additionally, the values of cryptocurrency sent to the address of interest provide important clues as to its ostensible purpose.

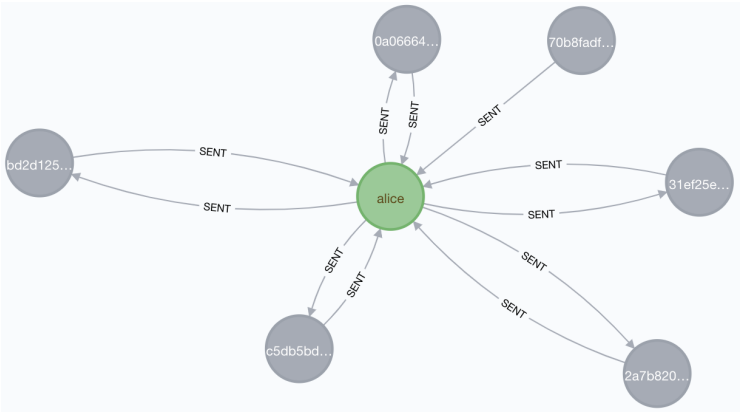


Fig. 2: Example of a Typical Control Crypto Address

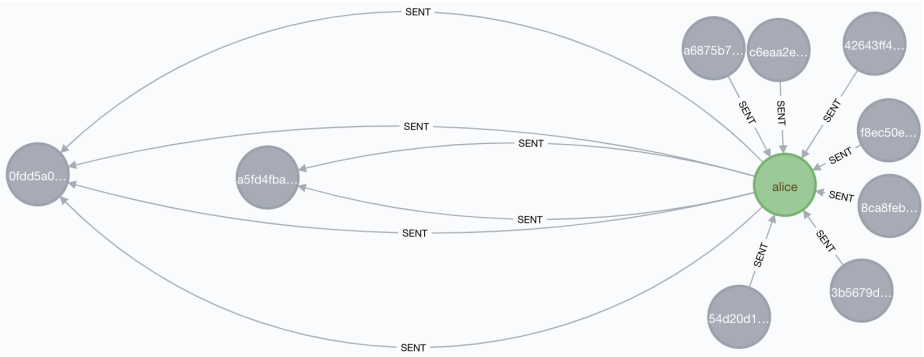


Fig. 3: Example of a Typical Scam BTC Address

Compared with typical addresses (often sending a variable number of currency tokens back and forth between a few addresses), wallets behind phishing operations often request of their victims provide a specified amount of currency in increments that are easy to remember, like 0.1, 0.2, 0.5. If a significant number of transactions are sent to an address that never returns tokens to the senders, especially in those increments, we deduce this as extremely likely to be a scam address. Additional analytics are run on the timing of the transactions. Typically, the interquartile (or non-outlier) range for the highest rate of transactions per hour occurs within a several-hour period. Compared with a standard wallet or receiver address - which will periodically receive and transfer out currency as deemed appropriate - phishing accounts' activity is often abrupt and absolute. Over a few hours, enough transactions are sent to the receiver address at a rate of 1-3 per hour, then the scammer promptly shuts down the operation, and empties the balance out into several different addresses.

3 Results

As a result of the graphs produced via the raw data processing by the IBPA, pattern matching tests can be run to determine whether or not specific addresses exhibit similarities those carrying out scam operations. These tests take into account a variety of parameters outlined in table 1. Parameters, like the average number of incoming transactions over a unit of time, the quantity of incoming/outgoing transactions, and the quantity of incoming/outgoing nodes, tell an essential story. Upon processing those various quantities with the help of the automated queries, the IBPA will return a "PASS" response to the interface if the ranked address score is less than a certain threshold. Conversely, a"FAIL" response is sent to the interface in cases where the score is greater than or equal to the threshold. As can be in table 1, this threshold value is taken as 3, which is highly arbitrary, and we note that this is only based on the preliminary tests for this iteration of the IBPA for a small set of accounts (n=16).

Avg Tx/Hr	Median Tokens	Incoming Nodes	Outgoing Nodes	Incoming Rels	Outgoing Rels	IBPA Score & Classification
0.015	0.0133	7	6	7	6	1 (regular)
0.015	0.6035	5	5	5	5	1 (regular)
<0.01	0.0003	9	0	9	0	1 (regular)
<0.01	0.7922	1	1	1	1	0 (regular)
<0.01	0	0	5	5	5	1 (regular)
0.07	0.01	7	2	13	11	5 (scam)
3.00	0.01	7	3	6	5	9 (scam)
8.73	0.01	11	1	11	5	11 (scam)
0.41	0.005	6	1	6	1	7 (scam)
4.00	0.022	6	3	11	24	3 (scam)

Table 1: IBPA results for sample of regular (non-scam) and scam accounts. Any score >= 3 results in a failing score and the corresponding address is flagged as a suspicious address meeting fraudulent criteria.

The standard BTC transactions selected at random did not exhibit graph shapes or structures close to those associated with the fraudulent operation. Of course, not all fraudulent cryptocurrency operations could be identified this way. Scams like false crypto exchanges operate under very different, much more long-term circumstances when compared with short-term live-streaming scams. However, the standard operations involving even such scams that garner tens of thousands of views can be identified by the IBPA. Whether or not the blockchain is UTXO, account-based, or predicated on another scheme, degree-1 transactions (providing standard public tx data) are all that the IBPA requires to assert relevant predictions. While some existing methods may propose highly resource-intensive operations that look at entire blockchain networks to identify long-term suspicious behavior patterns, this IBPA identifies this instantly, referencing a highly localized portion of a network to produce a highly accurate determination of how it applies to a regular user.

4 Related Work

The IBPA is by no means the first attempt at identifying fraudulent addresses on decentralized networks, and it certainly will not be the last. Chen et al [4] crucially lay out the need for more significant research into detecting crypto fraud and propose a graphical analysis API to aid in fraudulent Ethereum address. Vasek and Moore [5] put forth a granular investigation into the lifetime of crypto scams and their impacts. Bartoletti et al [6] investigated data mining and intensive classifiers to identify BTC Ponzi schemes. These works and others illustrate the importance of identifying bad actors who put innocent users at significant financial risk. There have already been some investigations by developers into detecting cryptocurrency fraud with Neo4j [7]. While their investigation was only focused on the Bitcoin blockchain, the investigation done

by this team proves that it is possible to use the tool to upload blockchain data and convert it into a useful, Cypher-queryable format to analyze suspicious addresses. Our solution differs from all these works because we utilize semantic web ontologies in representing the relationships between the various actors. We plan to incorporate these scam activities on the blockchain with the real-world events in a semantically meaningful manner that could lead to even more inferences in ascertaining fraudulent activities.

5 Conclusion

Blockchain platforms, currencies, and networks carry the immense potential to enhance our daily lives by empowering individuals to make their peer-to-peer transactions in almost any context, but also revolutionizing the way even centralized banking entities perform transactions across the globe. The growth of cryptocurrency phishing scams has become widely known, and available resources to protect the decentralized community are minimal. The IBPA described in the paper shows initial successes in presenting a method for chain-agnostic determination of whether or not a particular address is involved in a phishing operation by analyzing the patterns observed. However, for operations that operate as fraudulent crypto exchanges, more investigation would be required to identify addresses involved on decentralized networks robustly. These exchanges can assign individual addresses to each unique user, and will promptly siphon out funds in no unique or "flaggable" manner that would be particularly unlike control addresses. Ideally, future implementations of the IBPA would incorporate machine learning. While the current process for analyzing addresses is available, additional processing requirements would be necessary to ingest the massive pool of fraudulent transactions throughout the various blockchain ecosystems for large scale data analysis. Additional sweeping throughout the Web for instances of scam accounts would prove useful for this task. No singular attribute is the be-all-end-all for a scam account - they are merely pieces of evidence contributing to an arbitrary score. The accumulated relative score of these activities combined gives away the nature of addresses involved in these illegal operations. As time progresses, the classifier may require tweaks and neural net integrations as more suitable training datasets become available.

In conclusion, this preliminary evaluation proves that this minimally resource-intensive platform can accurately discover addresses that match the criteria for fraudulent phishing scams posing as legitimate live-stream giveaways. The IBPA can also be used to build comprehensive graphs enabling access to interoperable provenance queries that span across several blockchains for wallet addresses that are known to be linked to a single user. Consequently, the IBPA can be used as a plug-in module for future projects seeking enhanced knowledge of "bad actor" provenance across a variety of blockchain networks.

References

1. N. Statt, "Twitter's massive attack: What we know after Apple, Biden, Obama, Musk, and others tweeted a bitcoin scam (Jul 16, 2020)," online; accessed 09 August 2020. [Online]. Available: https://www.theverge.com/2020/7/15/21326200/elon-musk-bill-gates-twitter-hack-bitcoin-scam-compromised
2. R. Phillips and H. Wilder, "Tracing Cryptocurrency Scams: Clustering Replicated Advance-Fee and Phishing Websites," *arXiv preprint arXiv:2005.14440*, 2020.
3. M. I. Mehar, C. L. Shier, A. Giambattista, E. Gong, G. Fletcher, R. Sanayhie, H. M. Kim, and M. Laskowski, "Understanding a revolutionary and flawed grand experiment in blockchain: the DAO attack," *Journal of Cases on Information Technology (JCIT)*, vol. 21, no. 1, pp. 19–32, 2019.
4. W. Chen, X. Guo, Z. Chen, Z. Zheng, and Y. Lu, "Phishing Scam Detection on Ethereum: Towards Financial Security for Blockchain Ecosystem," *International Joint Conferences on Artificial Intelligence Organization*, pp. 4506–4512, 7 2020, special Track on AI in FinTech. [Online]. Available: https://doi.org/10.24963/ijcai.2020/621
5. M. Vasek and T. Moore, "Analyzing the Bitcoin Ponzi Scheme Ecosystem," *Financial Cryptography Workshops*, 2018.
6. M. Bartoletti, B. Pes, and S. Serusi, "Data mining for detecting Bitcoin Ponzi schemes," *CoRR*, vol. abs/1803.00646, 2018. [Online]. Available: http://arxiv.org/abs/1803.00646
7. C. Miles, "Detecting Cryptocurrency Fraud with Neo4j (Mar 30, 2020)," online; accessed 25 July 2020. [Online]. Available: https://neo4j.com/blog/detecting-cryptocurrency-fraud-with-neo4j