Terrorist Financing Network From Blockchain Data

Report Data Forensic

Dinkla, Daan & Manca, Massimo - team 14

Introduction

Our idea is to analyze the Bitcoin ledger and make an ego network of criminal Bitcoin wallets. As the bitcoin ledger is public and has all transactions on there, it should also include malicious transactions. We found a press release by the United States Justice Department regarding the disruption of three terror finance cyber-enabled campaigns, in this release they mention several Bitcoin addresses owned by terrorist organizations. We start from these addresses and reconstruct a network of Bitcoin addresses that interacted with one of these terrorist wallets. There are several methods described in papers, like for example Lin et al. (2019) or Chang & Svetinovic (2018), that can 'identify' or cluster bitcoin addresses based on ownership. We applied these types of methods to make meaningful conclusions about the structure of the network of Bitcoin addresses. Eventually, we used an input and output address heuristic. With the input and output heuristic, 2500 accounts were linked to the same 135 persons. We developed a new heuristic based on forwarding paths that showed that 6 accounts acted as intermediaries and belonged to the same 3 persons.

This is a relevant issue because it is known that preventing terrorist groups' access to financial resources is an effective way to stop them (Kfir, 2020). Terrorist groups are known to use cryptocurrencies as a way to raise funding, be it not on a broad scale yet.

Crawler description

The data are retrieved from the Bitcoin ledger through the Blockchain Data API[1], which allows you to retrieve information about an account such as the total amount sent and received bitcoin and the list of the last 50 transactions. The crawler retrieves information about addresses the users specify with the API, applies the heuristics to find multiple addresses belonging to the

[1] See: https://www.blockchain.com/api/blockchain_api

same persons and returns a list of transactions between people, with their amount and date.

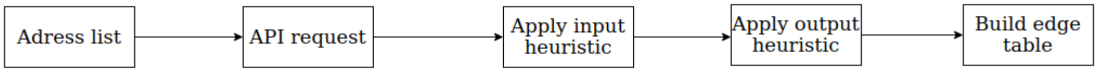


Figure 1: *crawler process illustration*

The users specify a set of addresses to inspect because there is a limit of requests that can be done per hour and there may be thousands of accounts appearing in the input or output even in only one transaction. Hence, it could take a considerable amount of time to explore all the accounts related to one address. The crawler waits 20 seconds between requests to avoid a block from the website.

The crawler applies two heuristics described in Zhang & Whang (2020). In Bitcoin transactions, there may be multiple input addresses and multiple output addresses, as shown in figure 10 in the appendix. The first heuristic assigns all the addresses that appear together as input in a transaction to the same identifier. It is necessary to have the key to the account to transfer money. This requires the sender to have the keys of all the addresses in the input of a transaction. Hence, it is safe to assume that he has control over them, and hence it is the owner. Although there may be some false positives with this method, because people can make transfers together, the authors explain that they do not consider them as such because they still have to know the key of each wallet and so they are co-owner. The algorithm detects if one or more than one addresses appear in the input of a transaction with at least another one that has been assigned to a person, and clusters them together. For example, if addresses a, b and c appear as input in transaction X and d, e and c appear as input in transaction Y, all addresses a, b, c, d, and e are connected to the same person since c is in common.

The second heuristic assigns ‘change addresses’ to the identifier of the input addresses. Change addresses are the addresses in a bitcoin transaction that are used to deposit the ‘change’ of a transaction. For a bitcoin transaction, one has to use the total amount of bitcoin present on the input address, this amount might be higher than the amount needed to be paid. Therefore the remaining bitcoin, i.e. the change, is deposited to a new address, the change address. This is also illustrated in figure 2. These change addresses have a different hash than the original input address but these addresses do belong to the same person.

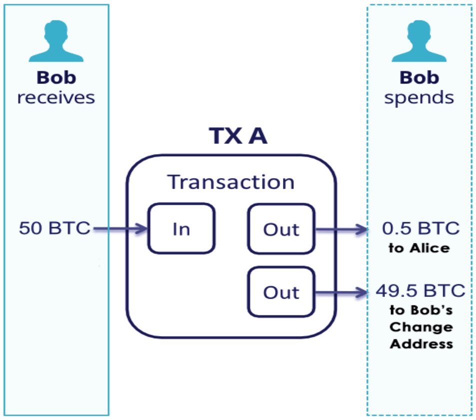


Figure 2: *A transaction with a change address*[2]

[2] Image retrieved from: https://coinguides.org/bitcoin-change-address-output/

Heuristic 2 aims to identify these change addresses by checking all output addresses of a transaction against four characteristics. An output address O in transaction t is a change address if:

1. This is the first appearance of address O in the ledger.
2. The transaction t is not coin generation.
3. There is no address among the output addresses that also appears as an input address in transaction t (self-change address).
4. The output addresses other than O do not satisfy condition (1).

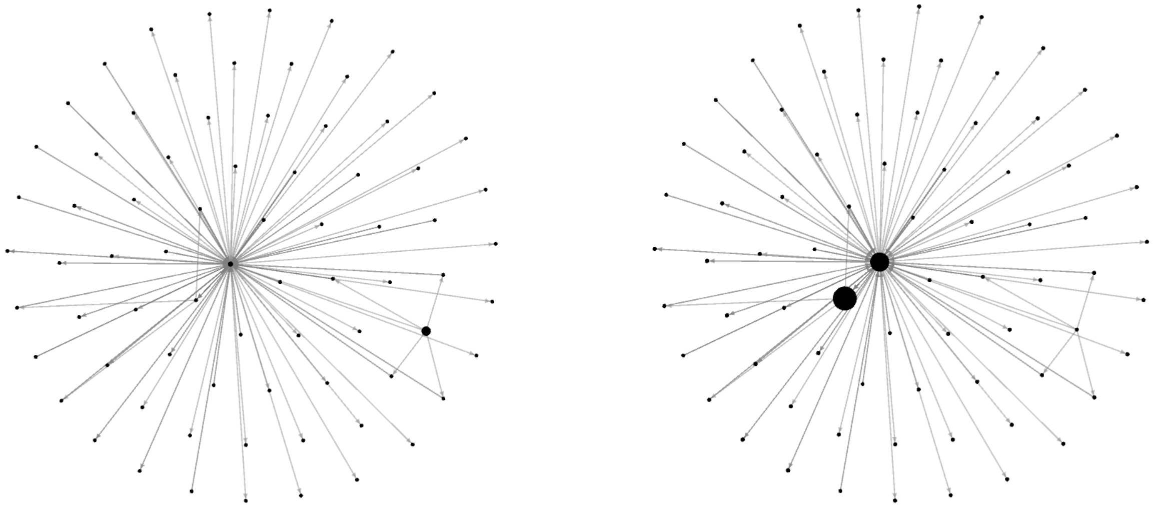
Just like the first heuristic this heuristic can lead to false positives and false negatives. For example, if a bitcoin transaction does not have any change we might falsely identify one of the recipient's addresses as a change address. There is however no completely reliable way of identifying change addresses and we have to accept some mistakes. The authors Zhang, Wang & Luo (2020) argue however that this method is more accurate than other heuristics aimed at detecting change addresses.

Methodology

Since older transactions involve several accounts that have been deactivated, only transactions that happened recently are of interest. The transaction history of the accounts related to the terrorist funding campaign, namely 17QAWGVpFV4gZ25NQug46e5mBho4uDP6MD and 15K9Zj1AU2hjT3ebZMtWqDsMv3fFxTNwpf show that the first account has 2 recent transactions, one in January 2021 and another at the end of March 2021, while the latter has no transactions in 2021. While these accounts seem to be in disuse, some of their recent contacts are still very active. Thus, we initially employed a methodology for building a network that started with iteratively crawling a list of most promising addresses and visualizing the result to find other interesting accounts. The criteria for identifying interesting accounts were i) the proximity to our target accounts (the two aforementioned), ii) the recency of their last activity (in general and with the targets), and iii) the position in the network (degree centrality and other measures). This methodology brought about the network in figure 3. The network so created was broad, with thousands of nodes, which made the analysis difficult under several aspects. For example, for applying the second heuristic and for having the relationships between nodes it was necessary to request information from thousands of accounts, which ended up in an amount of time that was not permissive. Hence, we decided to extract the accounts from the list of transactions of one suspicious account to build an ego network. The two addresses of the terrorist funding campaign would have been very interesting if they had a larger number of transactions that happened recently. We decided to choose an address that is active at this moment and is likely to be involved in criminal activity to understand more about it. The account 1Lm9BCDUKoBUk888DCXewM5p8bJyr83cEp satisfies our three conditions: it is one of the latest contacts of one of the accounts of the terrorists (17QAWGVpFV4gZ25NQug46e5mBho4uDP6MD), it was active in the last period and it is connected to an account that was reported to be in contact with scammers. That is account 1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s, which was one of the most central nodes in the network we initially inspected. It turned out that this account has a negative reputation since some users reported being scammed and their scammers transferred the money to this address. Both the accounts of the terrorist 17QAWGVpFV4gZ25NQug46e5mBho4uDP6MD and our next target, 1Lm9BCDUKoBUk888DCXewM5p8bJyr83cEp, were linked to this account. However, there are thousands of transactions with this account which raised problems to scrape it. These nodes and their relationships are shown in Figures 8 and 9 in the appendix. For the aforementioned reasons, we explored the ego network of the account 1Lm9BCDUKoBUk888DCXewM5p8bJyr83cEp. The analysis is presented in the next section.

Ego Network of criminal account

After selecting 1Lm9BCDUKoBUk888DCXewM5p8bJyr83cEp as our starting point, we scraped its transactions and applied the input heuristic. Before applying the heuristics, there were 2500 accounts. 2354 addresses were linked to 25 different persons with the input heuristic and 35 addresses were linked to 5 persons with the second heuristic. The total number of accounts after applying the two heuristics is 135. For each person, it was scraped the transactions of one of their accounts, since they appear together on the same transactions. The resulting ego network, shown in figure 3 is highly centralized. Almost every node has only one connection with the very two most central ones.



(a) Size represents out-degree

(b) size represents in-degree

Figure 3: *ego network after applying input output heuristics*

Clustering heuristics based on forwarding paths

Although the two heuristics significantly decreased the number of nodes linking multiple addresses to the same person, there may still be undetected clusters of addresses that belong to the same person. To detect them, we developed an algorithm that finds forwarding paths between addresses. A forwarding path is a sequence of directed links that forward the same amount of money to different addresses in a short amount of time. To make an example, node A transfers X amount of Bitcoin to address B, which in turn transfers X amount of bitcoin to node C. This sequence of events, when happening in a short period (i.e. a few minutes), suggests that node B is only an intermediary. Given the intent of remaining anonymous, it is likely that node A tried to disperse its traces making an intermediary step creating a fake account, which is node B. We implemented two heuristics that find two different types of forwarding paths. The first finds those transactions where the sender received an amount of bitcoin in the last 30 minutes. The second finds those transactions where the sender A transferred bitcoin to two different persons, B, and C, and after a short period B sent money to C. Here, again node B would be just an intermediary step, and A and B are linked to the same person. The two forwarding paths are shown in Figures 5 and 6.

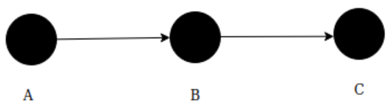


Figure 4: *forwarding path 1*

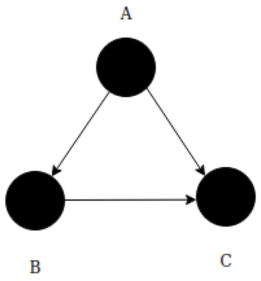
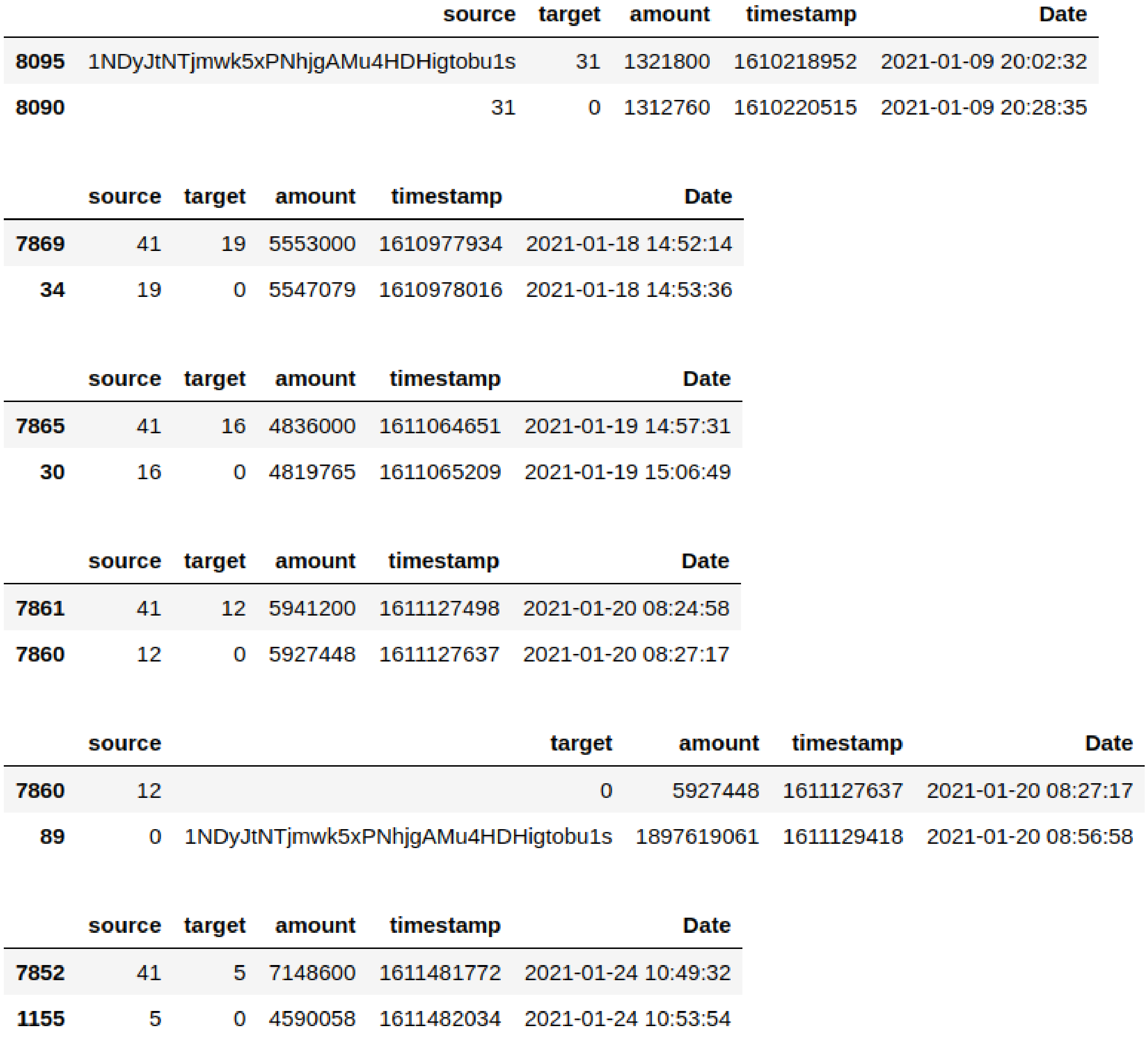


Figure 5: *forwarding path 2*

Results

The first heuristic found 6 accounts that acted as an intermediary, 4 of which connected to the same person with id 41. The second heuristic found 3 accounts that overlap entirely with the 4 belonging to 41 found with heuristic 1. The time elapsed between the forwarding from node to node was around 2-3 minutes in most cases and 28 minutes in only one case in heuristic 1. The amount forwarded was very similar to the one received in most cases, but not the same because of the cost of the commission in the transaction that was detracted.



	source	target	amount	timestamp	Date
8095	1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s	31	1321800	1610218952	2021-01-09 20:02:32
8090	31	0	1312760	1610220515	2021-01-09 20:28:35

	source	target	amount	timestamp	Date
7869	41	19	5553000	1610977934	2021-01-18 14:52:14
34	19	0	5547079	1610978016	2021-01-18 14:53:36

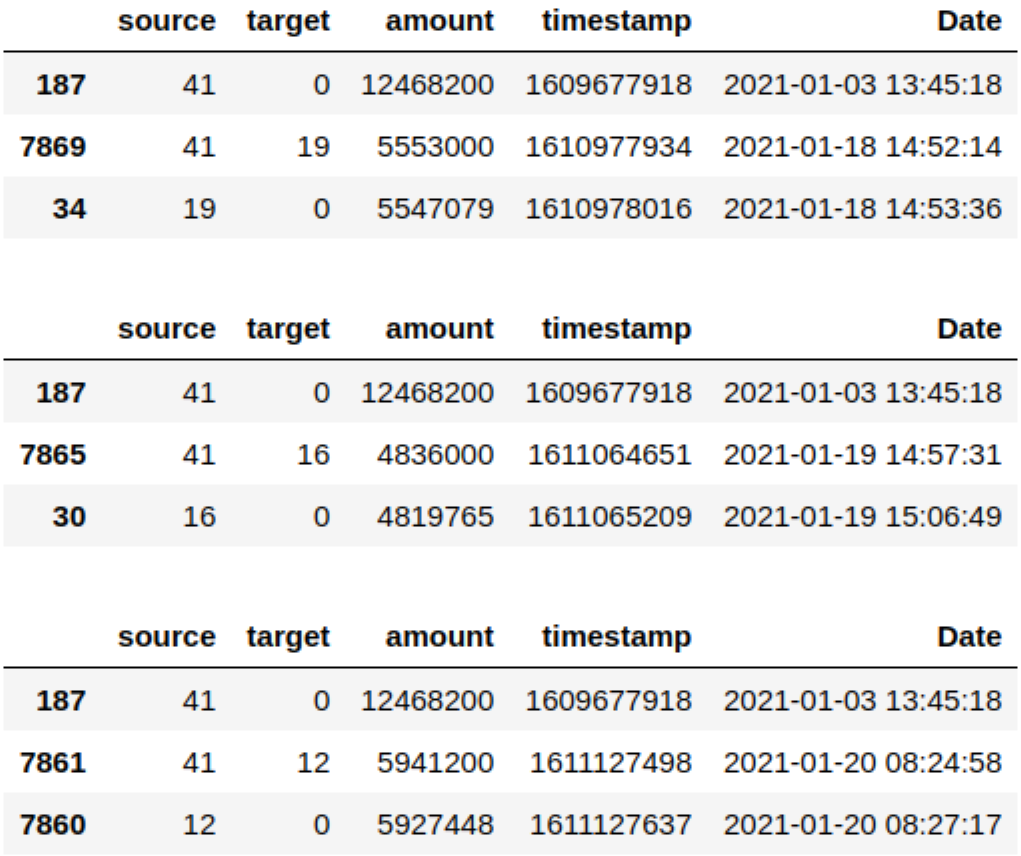
	source	target	amount	timestamp	Date
7865	41	16	4836000	1611064651	2021-01-19 14:57:31
30	16	0	4819765	1611065209	2021-01-19 15:06:49

	source	target	amount	timestamp	Date
7861	41	12	5941200	1611127498	2021-01-20 08:24:58
7860	12	0	5927448	1611127637	2021-01-20 08:27:17

	source	target	amount	timestamp	Date
7860	12	0	5927448	1611127637	2021-01-20 08:27:17
89	0	1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s	1897619061	1611129418	2021-01-20 08:56:58

	source	target	amount	timestamp	Date
7852	41	5	7148600	1611481772	2021-01-24 10:49:32
1155	5	0	4590058	1611482034	2021-01-24 10:53:54

Figure 6: *forwarding path 1 detected in the ego network after applying the input and output heuristics.*



	source	target	amount	timestamp	Date
187	41	0	12468200	1609677918	2021-01-03 13:45:18
7869	41	19	5553000	1610977934	2021-01-18 14:52:14
34	19	0	5547079	1610978016	2021-01-18 14:53:36

	source	target	amount	timestamp	Date
187	41	0	12468200	1609677918	2021-01-03 13:45:18
7865	41	16	4836000	1611064651	2021-01-19 14:57:31
30	16	0	4819765	1611065209	2021-01-19 15:06:49

	source	target	amount	timestamp	Date
187	41	0	12468200	1609677918	2021-01-03 13:45:18
7861	41	12	5941200	1611127498	2021-01-20 08:24:58
7860	12	0	5927448	1611127637	2021-01-20 08:27:17

Figure 7: *forwarding path 2 detected in the ego network after applying the input and output heuristics.*

Discussion and conclusion

The two heuristics based on forwarding paths converged to the same results, with the first one having a higher recall. The ego network, after applying the input and output heuristics, had a very centralized structure, in which two nodes attracted all the incoming edges from the others. We noticed that there was one node that had an exceptionally high out-degree value, whereas all the remaining nodes had only one connection with the most central nodes, as you can see in figure 3. We suspected that this node with several outgoing edges was anomalous. The heuristic based on forwarding paths found that that node (with id 41) uses multiple intermediary accounts to make a transfer of money to the most central nodes, as shown in Figures 6 and 7. Hence, it is likely that the heuristic found accounts that belonged to the same person who was trying to remain hidden. The scenario is realistic because all other nodes in this network send money to only two nodes (the big one in figure 3b) and because the transactions were forwarded in a very short time (2-3 minutes up to 10 minutes). The forward-path heuristic provides those nodes that are likely to belong to the same person, but a further manual inspection should be used to confirm this suspicion. For example, with node 41 it seems very likely that it used 19, 16,12, and 5 as an intermediary because it used this strategy systematically, the amount of time elapsed was very short and the amount of bitcoin forwarded was the same. However, it is unlikely that node 12 belongs to 0 as detected by heuristic forward path 1 and shown in figure 6. Node with id 12 was already detected to be an intermediary of 41 for interacting with 0. Hence, the connection of 12 with 0 found with heuristic forward path 1 can be a false positive.

The high centrality value of the two big nodes provides evidence that they were very popular and the others were supporting them. Also, the suspicious behavior of the supporters who tried to hide themself and the confirmed criminal nature of some of them, provide evidence that they operated in illicit activities.

For future research, it is important to understand the nature of the accounts that receive bitcoin from the most central nodes. One can imagine that those who send money to the big nodes are their supporters, but it is difficult to predict which mechanisms brought these two central nodes to send money to others. Furthermore, we limited our analysis to an ego-network of degree but it would be interesting to extend it to degree 2 for example, taking into consideration also the neighbors of the neighbors of the ego. Another topic of interest is the analysis of cyclic triads of type A -> B -> C -> A, which appeared several times in the ego network.

The analysis we performed on the ego network of one of the most interesting addresses the original terrorist address was connected to provides an interesting example case for the future. We show that it is possible to deduct an approximate number of individuals that one Bitcoin address has interacted with. Using heuristics presented in literature and our heuristics we were able to cluster several addresses together. This can be of use in the future. When law enforcement agencies find one or two addresses that belong to a specific person, using our method, it seems highly likely that they can identify other Bitcoin addresses that belong to these individuals. Additionally, they can construct a network of other individuals this address has interacted with, potentially increasing understanding of the financial networks these organizations form.

Appendix

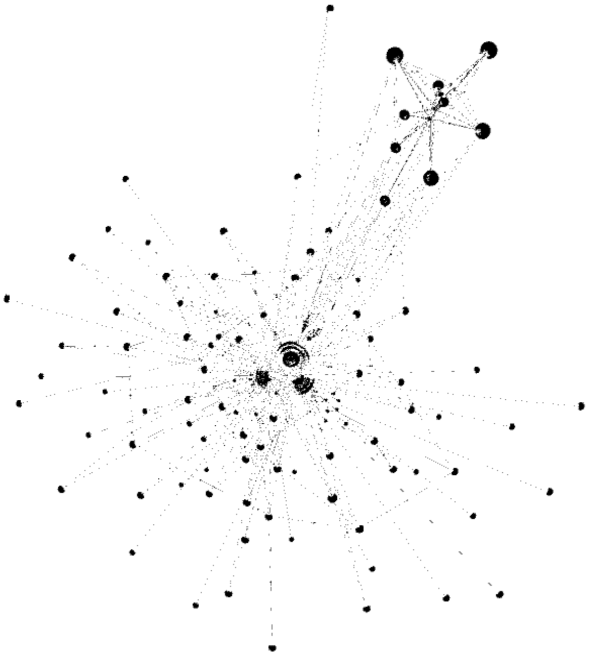


__Figure 8:__ visualization of network of transactions in 2021 of a list of addresses

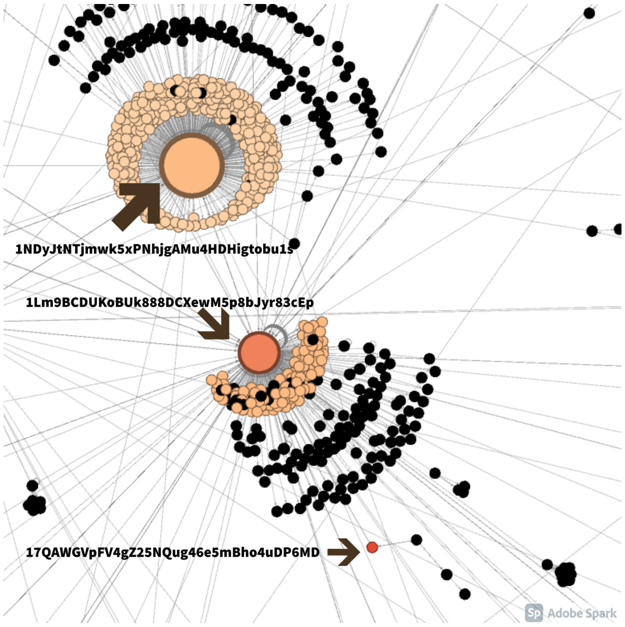
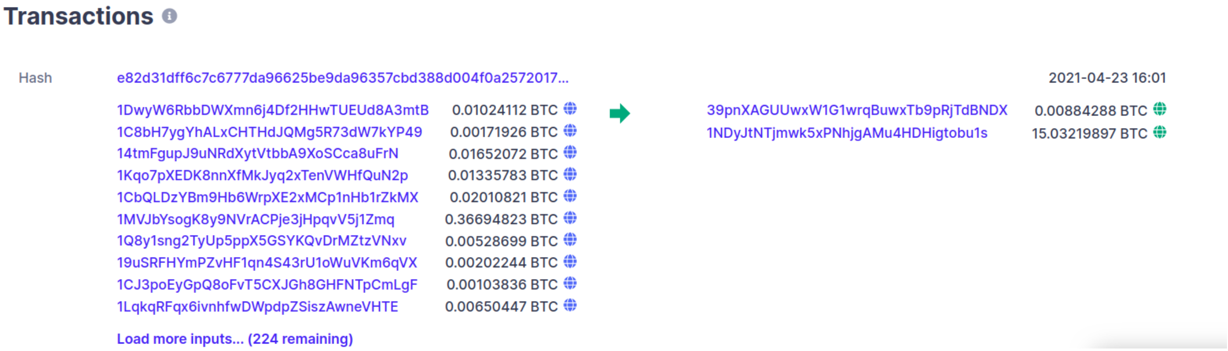

1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s
1Lm9BCDUKoBUk888DCXewM5p8bJyr83cEp
17QAWGVpFV4gZ25NQug46e5mBho4uDP6MD
Adobe Spark

__Figure 9:__ visualization of network of transactions in 2021 of interesting accounts



Transactions

Hash e82d31dff6c7c6777da96625be9da96357cbd388d004f0a2572017... 2021-04-23 16:01

1DwyW6RbbDWXmn6j4Df2HHwTUEUd8A3mtB	0.01024112 BTC	➡	39pnXAGUUwxW1G1wrqBuwxTb9pRjTdBNDX	0.00884288 BTC
1C8bH7ygYhALxCHTHdJQMg5R73dW7kYP49	0.00171926 BTC		1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s	15.03219897 BTC
14tmFgupJ9uNRdXytVtbbA9XoSCca8uFrN	0.01652072 BTC			
1Kqo7pXEDK8nnXfMkJyq2xTenVWHfQuN2p	0.01335783 BTC			
1CbQLDzYBm9Hb6WrpXE2xMCp1nHb1rZkMX	0.02010821 BTC			
1MVJbYsogK8y9NVrACPje3jHpqvV5j1Zmq	0.36694823 BTC			
1Q8y1sng2TyUp5ppX5GSYKQvDrMZtzVNxv	0.00528699 BTC			
19uSRFHYmPZvHF1qn4S43rU1oWuVKm6qVX	0.00202244 BTC			
1CJ3poEyGpQ8oFvT5CXJGh8GHFNTpCmLgF	0.00103836 BTC			
1LqkqRFqx6ivnhfwDWpdpZSiszAwneVHTE	0.00650447 BTC			

Load more inputs... (224 remaining)

***Figure 10:** example of transaction with multiple input and output addresses.*

Bitcoin Abuse Database

In the Bitcoin abuse dataset[3] we found additional information about some of the accounts we were interested in. In particular, account 1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s is reported to be the end point of different accounts which were involved in criminal activity. This account received a high number of transactions, making it the larger node in terms of in-degree. Regarding our first target account, the one shown on the image of the terrorists, the users of the website inform that “this address is being used to collect funds donation for terror organization”.

References

Aldridge, J., & Décary-Hétu, D. (2014). Not an'Ebay for Drugs': the Cryptomarket'Silk Road'as a paradigm shifting criminal innovation. Available at SSRN 2436643.

Campana, P. (2016). Explaining criminal networks: Strategies and potential pitfalls. Methodological Innovations, 9, 2059799115622748.

Irwin, A. S., & Milad, G. (2016). The use of crypto-currencies in funding violent jihad. Journal of Money Laundering Control.

Leukfeldt, E. R., Kleemans, E. R., & Stol, W. P. (2017). Cybercriminal networks, social ties and online forums: Social ties versus digital ties within phishing and malware networks. The British Journal of Criminology, 57(3), 704-722.

Zhang, Y., Wang, J., & Luo, J. (2020). Heuristic-Based Address Clustering in Bitcoin. IEEE Access, 8, 210582-210591.

[3] https://www.bitcoinabuse.com/reports/1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s