Blockchain analysis

Heuristics

- **“Clustering addresses”** refers to the process of attributing numerous addresses to the same wallet/controlling entity through the use of transaction behaviour heuristics.

- There are a number of factors which go into these heuristics. We will cover some of the well known ones in the following slide. It is important to note however that none of the heuristics are definite. They can be wrong and as investigators it is necessary to corroborate the results.

- Cryptocurrencies focused on principles of self sovereignty recognise blockchain analysis as an attack on the network. **Small minorities within these communities are working to break the heuristics used and minimise the ability to undermine privacy within the protocol**.

- The key point here is that the methods being used to provide Blockchain Forensic Tools is likely to evolve in line with efforts to break the heuristics. It could become more difficult to identify how results are being provided and this makes it important to keep informed on the subject.

Heuristic 1 – Common Input Ownership

Assumption: All inputs in a transaction belong to the same entity as they reside in the same wallet.

- Vast majority of bitcoin transactions are simple in nature. Very few collaborative transactions.
- As a result one wallet controlled by one entity will have provided all of the transaction inputs to send funds.

Heuristic 2 – Change address detection:

- Change amounts are linked to addresses never previously seen in the blockchain.
- If an output address is the same as an input address it is the change.
- Wallet fingerprinting can be used to detect change outputs because a change output is the one spent with the same wallet fingerprint.
- Round numbers as an output are payments not change.
- If the values of the inputs are more than one of the outputs but less than another, the lower figure output is change (Unnecessary input heuristic) e.g.

Inputs	Outputs	Assumption
1BTC	3.5BTC →	Payment
2BTC	0.5BTC →	Change
1BTC		

Assumptions:

- **If an output address has been reused it is very likely to be a payment output**, not a change output. This is because change addresses are created automatically by wallet software but payment addresses are manually sent between humans.

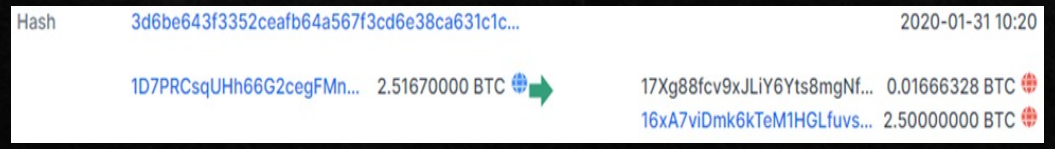
- Entities utilise wallet defaults for coin selection and fee payments.

- **Many payment amounts are round numbers**, for example 1 BTC or 0.1 BTC. The leftover change amount would then be a non-round number (e.g. 1.78213974 BTC). This potentially useful for finding the change address. The amount may be a round number in another currency. The amount 2.24159873 BTC isn't round in bitcoin but when converted to USD it may be close to an exact dollar value.

Heuristic examples

Common input heuristic

Change heuristic: Outputs are different script types.



Change heuristic: The address 16xA7 was active prior to this transaction. The address 17Xg88 was first active as part of this transaction. Round payment made to 16xA7 address

Private intelligence

Further to these heuristics Blockchain Forensic Tools will **utilise industry intelligence and covert surveillance tactics to attribute entities to addresses and build clusters**. This will mean it is often **opaque as to how an identification or cluster has been developed**.

It is possible to try and find connections. You can manually check for the heuristics and carry out open source research. This may however become unpractical (if significant amounts of data) or turn out to be inconclusive. In these instances it would be worth noting down the efforts made and the negative result. Such process will at least show efforts have been made to understand the intelligence and quantify it's origins.

Practical examples

1

The licence can be used by anyone, just log out and the credentials can then be used by another. The best practice is actually getting hands on so please try to have a go.

2

Complete an audit trail using the spreadsheet or whatever other mechanism works for you.

3

Think about how you would look to progress the enquiries you identify and record this in the strategy.

Practical 1

- Upload the .grf file in Chainalysis
- For each cluster identify the theft transaction.
- Plot the inputs which funded the victim wallet
- Highlight the services contributing to the victim wallet (provide brief over view of what the service is/where based).

Practical 1

- Plot the wallets used by the offenders
- Looking at 17vDMhe1Ym9XQMyaa6ahW4fvU5ibVCniBi, can you think of why it's transaction history does not seem to fit with the intelligence?
- Identify any opportunities to cluster addresses together and merge them into one cluster
- Complete a strategy outlining the parameters used for the investigation and the opportunities identified for pursuing the offenders/proceeds of crime.

Practical 2 – Kucoin hack

- 1NRsEQRg5EjmJHbPUX7YADVPcPzCQBkyU7
  12FACbewf5Fy9nmeaLQtm6Ugo5WS8g2Hay
  1TYyommJW3uhjhcnHhUSuTQFqSBAxBDPV
- Search blockchair.com using the above addresses, what can you identify in terms of assets held? What's is interesting in regards to this?
- Look at these addresses on open source websites, can you find information to parallel the attribution of these addresses

Practical 2 – Kucoin hack

- 0xeb31973e0febf3e3d7058234a5ebbae1ab4b8c23
- Search this address on https://etherscan.io, what can you find out on this website?
- TB3j1gUXaLXXq2bstiSMfjQ9R7Yh9DdDgK
- Search this address on Tronscan.io and https://trx.tokenview.com/en, what can you find out on this website?
- Prepare a brief report of your findings

Practical 2 – Kucoin hack

- Open Chainalysis and plot the three BTC addresses
- Look at the funds coming into the suspect wallet, can you identify any lines of enquiry relating to this?
- Look at the funds coming out of the suspect wallet, what lines of enquiry can you identify in respect of this?
- Complete a strategy outlining the parameters used for the investigation and the opportunities identified for pursuing the offenders/proceeds of crime.