Deanonymization and linkability of cryptocurrency transactions based on network analysis

Alex Biryukov, Sergei Tikhomirov

University of Luxembourg

17 June 2019 Euro S&P Stockholm, Sweden



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

meroducer

Γx clusterin

Veighting timestamp

relation matrix

Measuring anonymity

periments

Juliating the Jource

scussion

Outline

Introduction

Transaction clustering

Parallel connections
Weighting timestamp vectors
Correlation matrix
Measuring anonymity

Experimental results

Estimating the source IP

Discussion

Conclusion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

x clustering

Parallel connections
Weighting timestamp

Correlation matrix

leasuring anonymity

stimating the source l

scussion

Outline

Introduction

Transaction clustering

Parallel connections

Weighting timestamp vectors

Correlation matrix

Measuring anonymity

Experimental results

Estimating the source If

Discussion

Conclusion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections
Weighting timestamp

orrelation matrix

. .

stimating the source

scussion

Conclusion

Bitcoin



- ▶ The first to solve double-spending with proof-of-work
- Senders broadcast transactions into a P2P network
- Miners construct blocks (thus confirming transactions)

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Parallel connections
Weighting timestam

Correlation matrix

Measuring anonymity

periments

Estimating the source IP

cussion

Privacy in Bitcoin

- ► Transactions not linked to "real-world" identity
- ▶ Users can generate as many key pairs as they wish
- ► False sense of privacy?

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections

vectors

Measuring anonymity

periments

eriments

stimating the source II

scussion

Taint analysis heuristics

- ▶ All transaction inputs *probably* belong to the sender
- One output probably also belongs to the sender

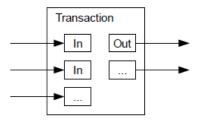


Figure: Bitcoin transaction structure

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

eighting timestamp

Measuring anonymity

periments

scussion

Privacy coins hinder blockchain analysis...



Dash: mixing by masternodes



Monero: ring signatures



Zcash: zk-SNARKs

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Parallel connections
Weighting timestamp
vectors

orrelation matrix leasuring anonymity

xperiments

stimating the source

scussion

...but what about network analysis?

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

x clustering

Parallel connections
Weighting timestam

Correlation matrix

Measuring anonymity

periments

Estimating the source IP

scussion

Conclusion

How do messages propagate through the network?

▶ What does a well-connected adversary learn?

Is it possible to link txs by the same user?

Our contributions

Deanonymization and linkability of cryptocurrency transactions based on network analysis

Biryukov, Tikhomirov

Introduction

× clustering

Parallel connections Weighting timestam

Correlation matrix

Measuring anonymit

periments

stimating the source IP

scussion

Conclusion

We introduce a new transaction clustering method based on weighted vectors of IP addresses

We validate our method with experiments on Bitcoin and three major privacy-focused cryptocurrencies

Outline

Introduction

Transaction clustering

Parallel connections
Weighting timestamp vector
Correlation matrix
Measuring anonymity

Experimental results
Estimating the source IF

Discussion

Conclusion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections
Weighting timestamp
vectors

Correlation matrix Measuring anonymity

xperiments

stimating the sourc

iscussion

Message propagation in Bitcoin

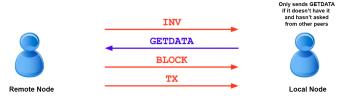


Figure: Bitcoin's 3-step message exchange

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections
Weighting timestamp
vectors

Lorrelation matrix

Measuring anonymity

xperiments

cussion

Broadcast randomization in Bitcoin and forks

- trickling: send to a random subset once every 100 ms
- diffusion: send to each neighbor after a random delay

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections
Weighting timestamp
vectors

orrelation matrix Measuring anonymity

neasuring anonymity

timating the source IP

cussion

Intuition

Transactions issued from the same node have correlated broadcast patterns.

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Tx clustering

Outline of our clustering method

- Establish parallel connections to many nodes
- ▶ Log timestamps of received tx announcements
- For each tx, consider IPs which announced it to us
- ► Cluster transactions with "similar" IP vectors
- ► Measure the decrease in anonymity

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections
Weighting timestam
vectors

Jorrelation matrix Measuring anonymity

xperiments

imating the source IP

SCUSSIOII

Parallel connections

▶ Default connections: 8 outgoing + up to 117 incoming

► We are unlikely to get a new tx quickly with only one connection per node

bcclient establishes parallel connections to nodes

Bitcoin and Zcash show similar distribution of free slots

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections

Weighting timestamp vectors

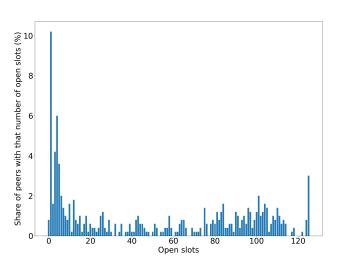
Measuring anonymity

periments

stimating the source IF

scussion

Bitcoin free slots



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

meroduction

Tx clustering

Parallel connections

Weighting timestamp

Correlation matrix

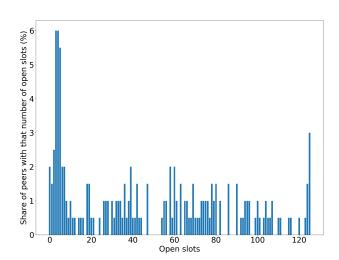
Measuring anonymity

kperiments

Estimating the source IP

iscussion

Zcash free slots



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introductio

Tx clustering

Parallel connections

Weighting timestamp

Correlation matrix

Measuring anonymity

stimating the source I

.

iscussion

Weighting timing vectors

IP addresses p_i announce a new tx to us at times t_i . We assign exponentially decreasing weights to p_i :

$$w(p_i) = e^{-(t_i/k)^2}$$

where the median IP gets weight 0.5:

$$k = \frac{t_{median}}{\sqrt{-\ln(0.5)}}$$

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Weighting timestamp vectors

Correlation matrix

vnorimente

Estimating the source IP

iscussion

Weighting timing vectors: example

High values indicate higher probability of an IP to be the sender or one of its entry nodes.

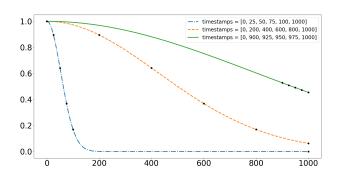


Figure: Weight functions for 3 timestamp vectors

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

x clustering

Weighting timestamp vectors

Correlation matrix
Measuring anonymity

xperiments

stimating the source IF

Clustering the correlation matrix

- ► For each pairwise correlations of weight vectors of txs
- ► Hypothesis: correlation matrix has a *block-diagonal* structure
- ► With a right permutation of rows and columns, related transactions will form clusters along the main diagonal

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Weighting timestamp vectors

Correlation matrix

Measuring anonymity

periments

Estimating the source IP

iscussion

Heatmap visualization

- Display correlations between weight vectors as matrix
- Darker color means higher correlation
- ▶ Matrix is symmetric by definition: corr(i, j) = corr(j, i)
- ▶ The main diagonal is black: correlation with oneself

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Weighting timestar

Correlation matrix

Measuring anonymity

xperiments

Estimating the source IP

iscussion

Measuring anonymity

We use anonymity degree proposed by Díaz et al. 1:

$$d = \frac{-\sum_{i=1}^{N} p_i log_2(p_i)}{log_2(N)}$$

where p_i is the probability of the i-th tx to originate from the given source.

- d = 1: users are equally likely to be the senders of a given message
- ightharpoonup d = 0: the attacker knows the senders of all messages

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clustering

Parallel connections

Weighting timestamp vectors

Correlation matrix

Measuring anonymity

periments

imating the source I

scussion

¹Díaz, Seys, Claessens, Preneel. Towards measuring anonymity. 2002

Putting the pieces together

- ► Connect to many nodes from servers on 3 continents
- ► Log transaction announcements
- Assign weights to vectors of timestamps
- Calculate pairwise correlations between weight vectors
- ▶ Apply the spectral co-clustering algorithm ²
- Calculate anonymity degree for our txs as ground truth
- ► Ethical considerations: mostly testnet, our own txs

Biryukov, Tikhomirov

Introduction

T× clusterin

arallel connections leighting timestamp

Measuring anonymity

neusuring unonymic

imating the source II

iscussion

Lonclusion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

²I.S.Dhillon. Co-clustering documents and words using bipartite spectral graph partitioning. 2001

Outline

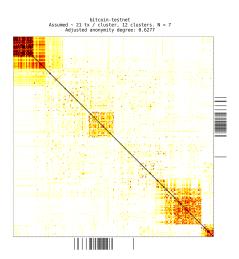
Experimental results

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Experiments

Bitcoin testnet: anonymity degree = 0.63



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Tx clusterin

Parallel connection

leighting timestamp

Correlation matrix

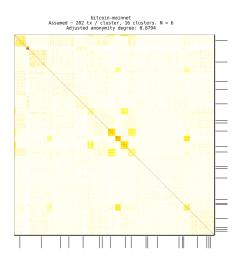
Measuring anon

Experiments

Estimating the source IF

scussion

Bitcoin mainnet: anonymity degree = 0.88



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

meroduction

x clustering

Parallel connection

ectors timestamp

Measuring anonymi

_ .

Experiments

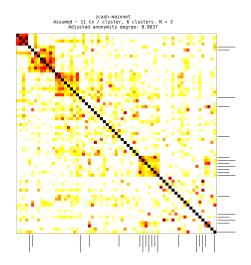
Estimating the source IF

cussion

Conclusion

Only connected to 1/10 of nodes, didn't occupy all slots.

Zcash: anonymity degree = 0.86



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

.....

Tx clusterin

Parallel connection

Veighting timestamp ectors

Correlation matr

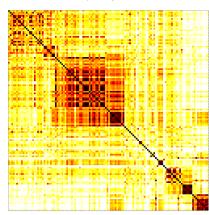
_

Experiments

_

Monero

monero-mainnet
Assumed ~ 12 tx / cluster, 10 clusters, N = 3



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Parallel connection

ectors

Measuring anonymit

Experiments

Estimating the source IP

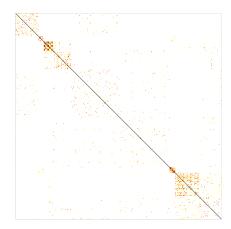
300331011

Conclusio

Experiment without our own transactions.

Dash

dash-mainnet. N = 4, 9 clusters



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Experiments

Estimating the source IP from ADDR messages

- ► A new node advertises its IP in ADDR messages
- ► We intersect the announced IPs from ADDRs with the highest-weighted IPs in tx clusters (Bitcoin testnet)
- ▶ In most experiments, the source IP appeared among top-5 highest weighted IPs in our transaction cluster

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Parallel connections
Weighting timestamp
vectors

Correlation matrix

neasuring anonymic

periments

Estimating the source IP

iscussion

Outline

Discussion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Discussion

Cost of attack

- Feasible for a moderately resourceful attacker
- Main cost components are bandwidth and storage
- ▶ We estimate the cost of a full-scale attack on Bitcoin mainnet at hundreds of US dollars
- Our experiments cost \$35 on AWS

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Weighting timestar

orrelation matrix

leasuring anonymity

periments

timating the source II

Discussion

Countermeasures

- ▶ Don't issue many txs in the same session
- Run nodes with increased number of connections
- Periodically drop and re-establish random connections
- ▶ Implement stronger broadcast randomization

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Parallel connections
Weighting timestamp

Correlation matrix

Measuring anonymit

periments

Discussion

300331011

Countermeasures (contd): new relay protocols

- ▶ Dandelion++: two-stage propagation for better anonymity. Only outgoing connections for first phase. Hard to force a remote node to connect to us
- ► Erlay (proposed 2019-05-28): "[A]nnouncements are only sent directly over a small number of connections (only 8 outgoing ones). [...] We [...] better withstand timing attacks"

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introduction

Tx clusterin

Weighting timestam

orrelation matrix

leasuring anonymity

stimating the source IP

Discussion

Outline

Conclusion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Conclusion

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Conclusion

Clustering works better on small networks

Announcement timings reveal related transactions

Randomization techniques are not very efficient

Future work: mobile wallets

- ▶ In our experiments, txs were issues from a full node
- How are mobile wallets different in terms of networking?
- Can we cluster transactions issued from mobile wallets?

Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Questions?

- cryptolux.org (we are hiring postdocs)
- s-tikhomirov.github.io



Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

Introductio

Tx clustering

eighting timestamp

rrelation matrix casuring anonymity

periments

scussion

Image credits

 Transaction structure: Andreas Antonopoulos. https://bit.ly/2MPDpba

▶ Data exchange: Samuel Omidiora. https://bit.ly/2MO8Mmo Deanonymization and linkability of cryptocurrency transactions based on network analysis

> Biryukov, Tikhomirov

minoduction

x clustering

Parallel connections

vectors

Measuring anonymity

xperiments

imating the source