Distributed Ledger Technology Workshop
1 - 02 - 2018

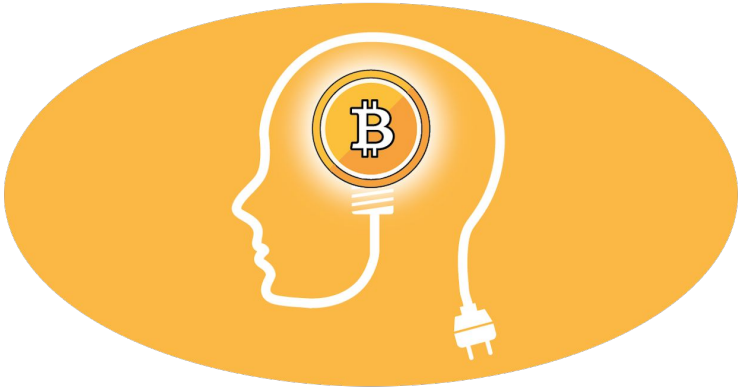
An analysis of the Bitcoin Users Graph: Detecting Artificial Behaviours

Damiano Di Francesco Maesa Andrea Marino Laura Ricci

UNIVERSITÀ DI PISA

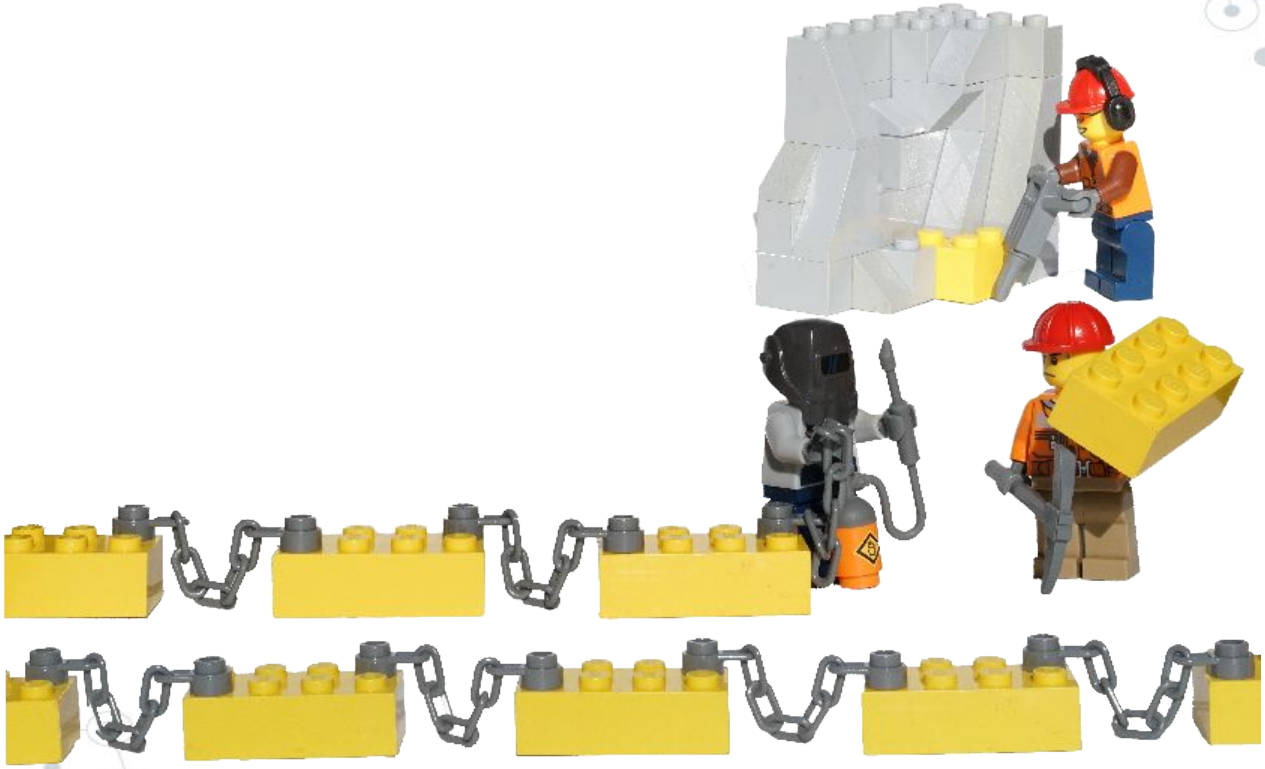
Key words

Bitcoin



Users Graph

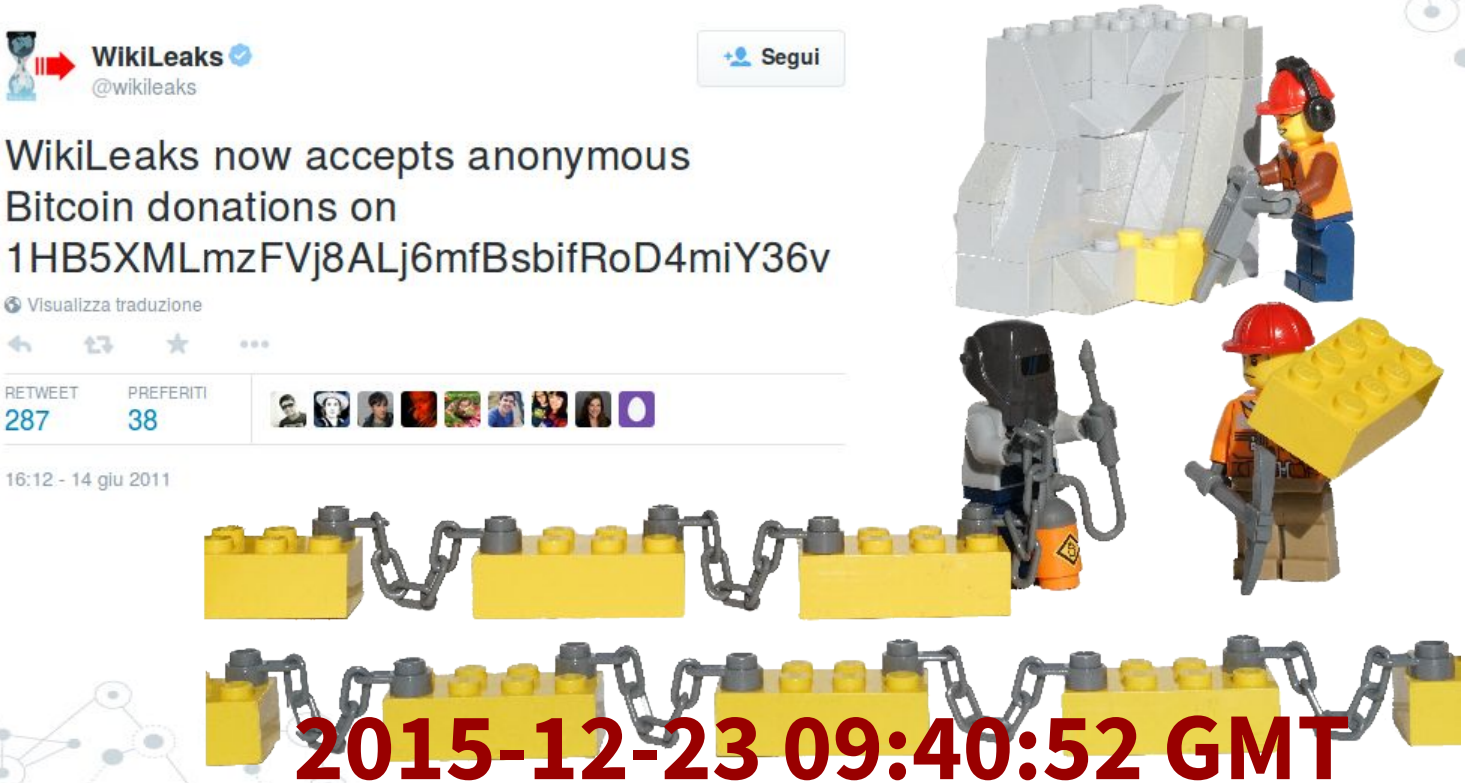
Data:



Data: Bitcoin Blockchain

2015-12-23 09:40:52 GMT

Data: Bitcoin Blockchain + External Informations



2015-12-23 09:40:52 GMT

Deanonymization Attack

blockchain

⇩

transactions graph

heuristic based clustering

⇩

users graph

external informations

⇩

identities graph

Pseudonymity

1Ez69SnzzmePmZX3WpEzMKTrcBF2gpNQ55

Pseudonymity

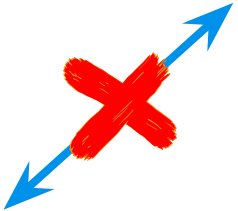
1A1zP1eP5QGefi2DMPTfTL5SLmv7DivfNa

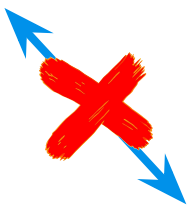
1Ez69SnzzmePmZX3WpEzMKTrcBF2gpNQ55

1XPTgDRhN8RFnzniWCddobD9iKZatrvH4

Pseudonymity

Pseudonymity



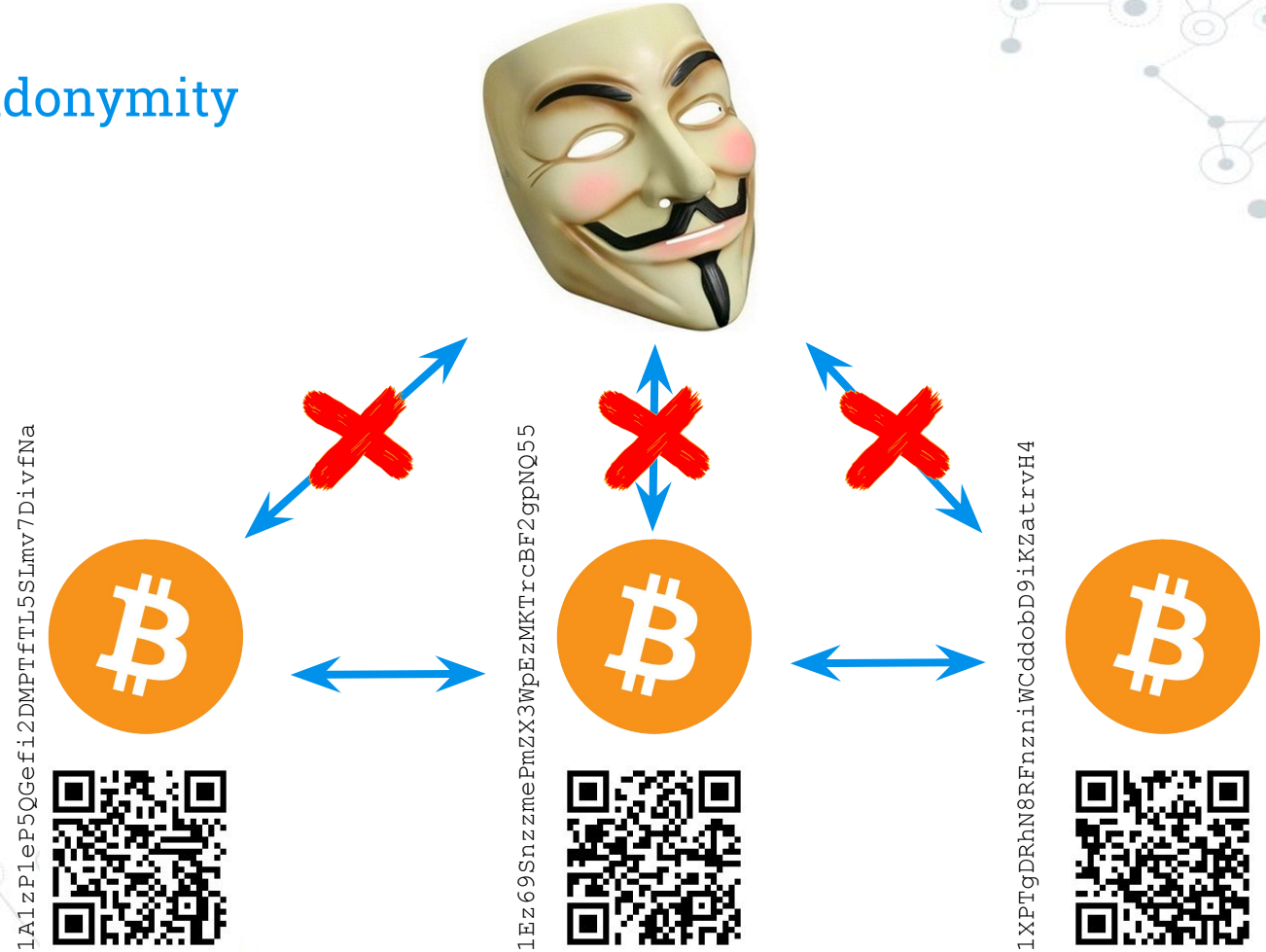


1A1zP1eP5QGefi2DMPTfTL5SLmv7DivfNa

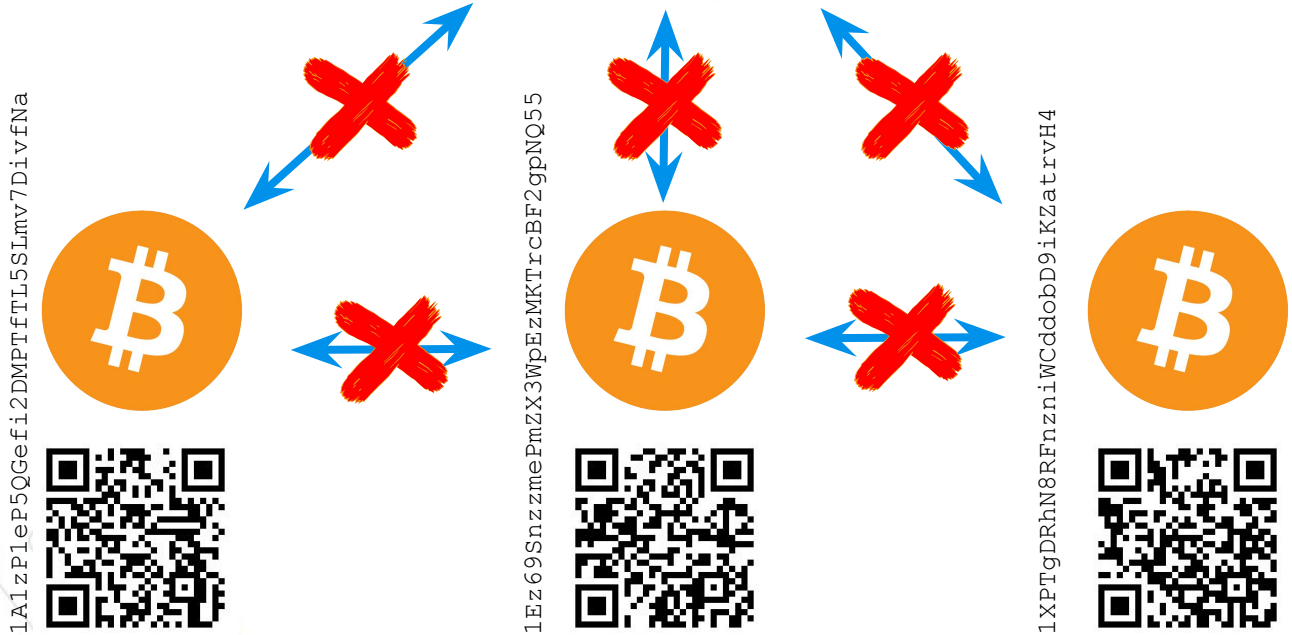
1Ez69SnzzmePmZX3WpEzMKTrcBF2gpNQ55

1XPTgDRhN8RFnzniWCddobD9iKZatrvH4

Pseudonymity



Pseudonymity



Deanonymization Attack

blockchain

transactions graph

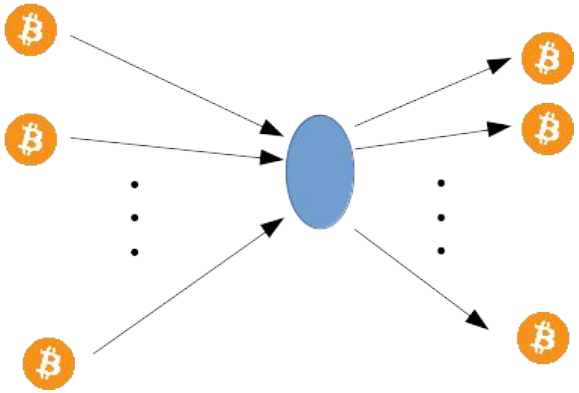
heuristic based clustering

users graph

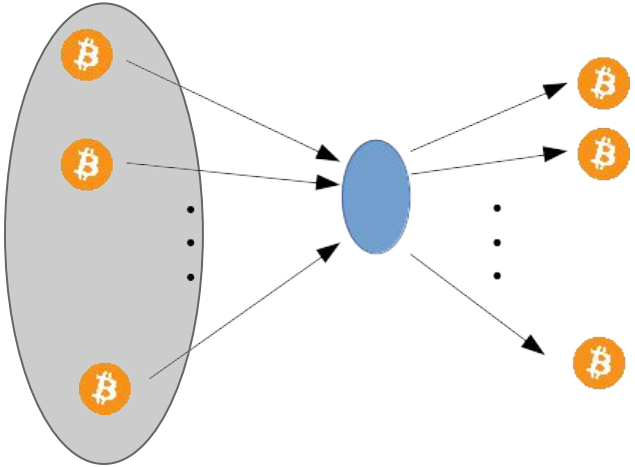
external informations

identities graph

Common inputs heuristic clustering



Common inputs heuristic clustering



$$O(|A| + \sum_{(A_1,A_2)\in T} (|A_1| + |A_2|))$$

Users Graph:

Weighted directed multigraph

- **46,144,246** Nodes
- **294,705,549** Edges

Densification, distance analysis, degree distribution, clustering coefficient and several centrality measures

WebGraph

Definition of richness:

$$d^t(u) = |\{(v,u) \ : \ (v,u) \in E^t\}|$$

Definition of richness:

$$b^t(u) = \sum_{(v,u)\in E^t} w(v,u) - \sum_{(u,v)\in E^t} w(u,v) - \phi^t(u) + \beta^t(u)$$

$$d^t(u) = |\{(v,u) \ : \ (v,u) \in E^t\}|$$

Definition of richness:

$$b^t(u) = \sum_{(v,u)\in E^t} w(v,u) - \sum_{(u,v)\in E^t} w(u,v) - \phi^t(u) + \beta^t(u)$$

$$d^t(u) = |\{(v,u) \ : \ (v,u) \in E^t\}|$$

Definition of richness:

$$b^t(u) = \sum_{(v,u)\in E^t} w(v,u) - \sum_{(u,v)\in E^t} w(u,v) - \phi^t(u) + \beta^t(u)$$

$$d^t(u) = |\{(v,u) \ : \ (v,u) \in E^t\}|$$

Definition of richness:

$$b^t(u) = \sum_{(v,u)\in E^t} w(v,u) - \sum_{(u,v)\in E^t} w(u,v) -$$

$$\phi^t(u) + \beta^t(u)$$

$$d^t(u) = |\{(v,u) \ : \ (v,u) \in E^t\}|$$

Definition of richness:

$$b^t(u) = \sum_{(v,u)\in E^t} w(v,u) - \sum_{(u,v)\in E^t} w(u,v) -$$

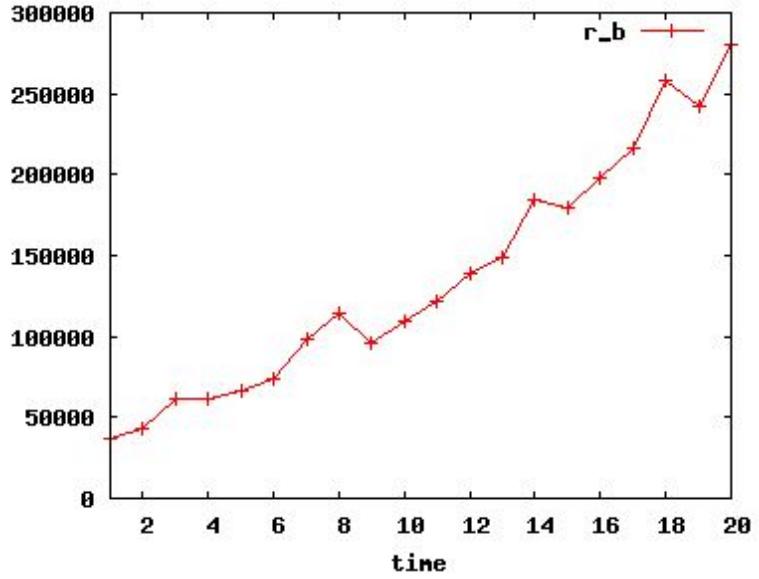
$$\phi^t(u) + \beta^t(u)$$

$$d^t(u) = |\{(v,u) \ : \ (v,u) \in E^t\}|$$

Rich get Richer:

- The richest users at time *t* are richer than the richest users at time *t'*<*t*.

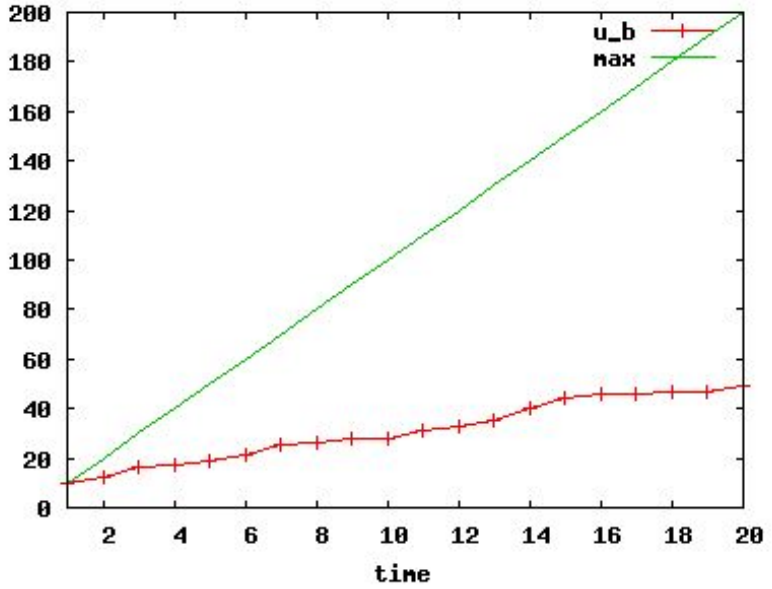
$$r_b^t = \frac{\sum_{u \in B_k^t} b^t(u)/k}{\sum_{u \in V^t} b^t(u)/|V^t|}$$



Rich get Richer:

- The richest users at a certain time *t* tend to remain the richest at time *t'*>*t*.

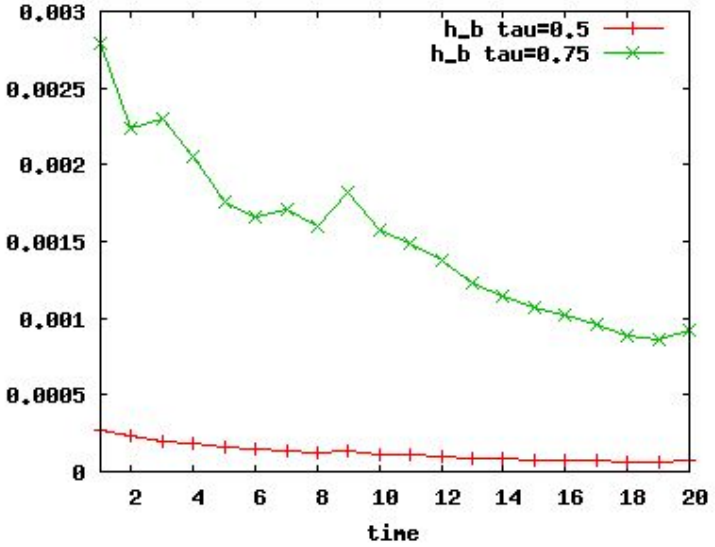
$$u_b^t = |\bigcup_{i=1}^{t} B_k^i|$$



Rich get Richer:

- The richness gets more concentrated with the progression of time.

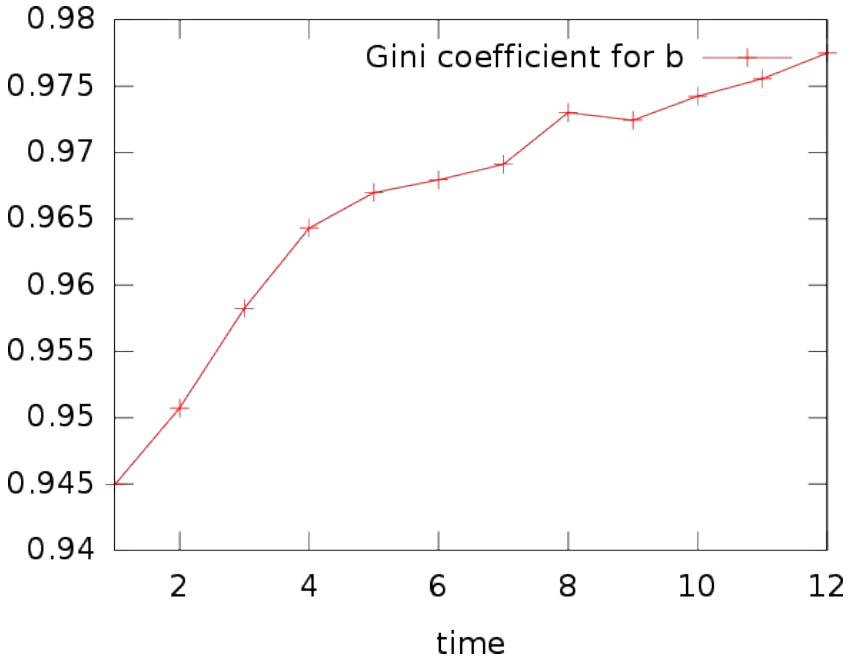
$$h_b^t = min\left\{ k \ : \ \frac{\sum_{u \in B_k^t} b^t(u)}{\sum_{u \in V^t} b^t(u)} > \tau \right\} / |V^t|$$

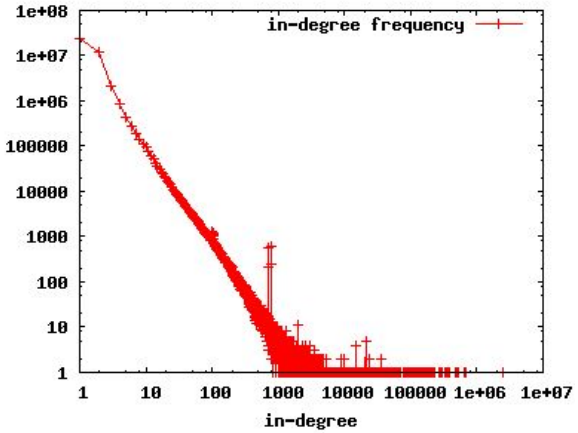


Rich get Richer:

- Gini coefficient.

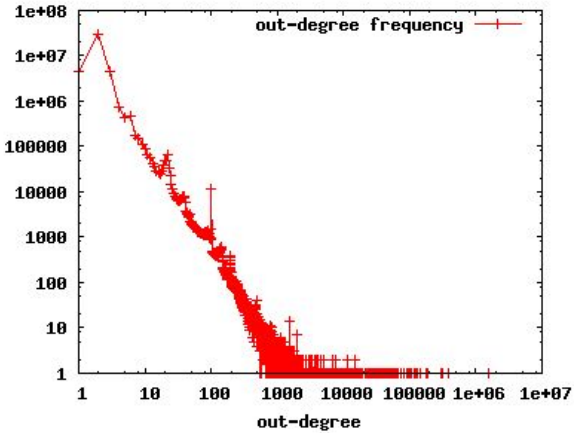
$$G = \frac{2\sum_{i=1}^{n} i x_i}{n\sum_{i=1}^{n} x_i} - \frac{n+1}{n}$$



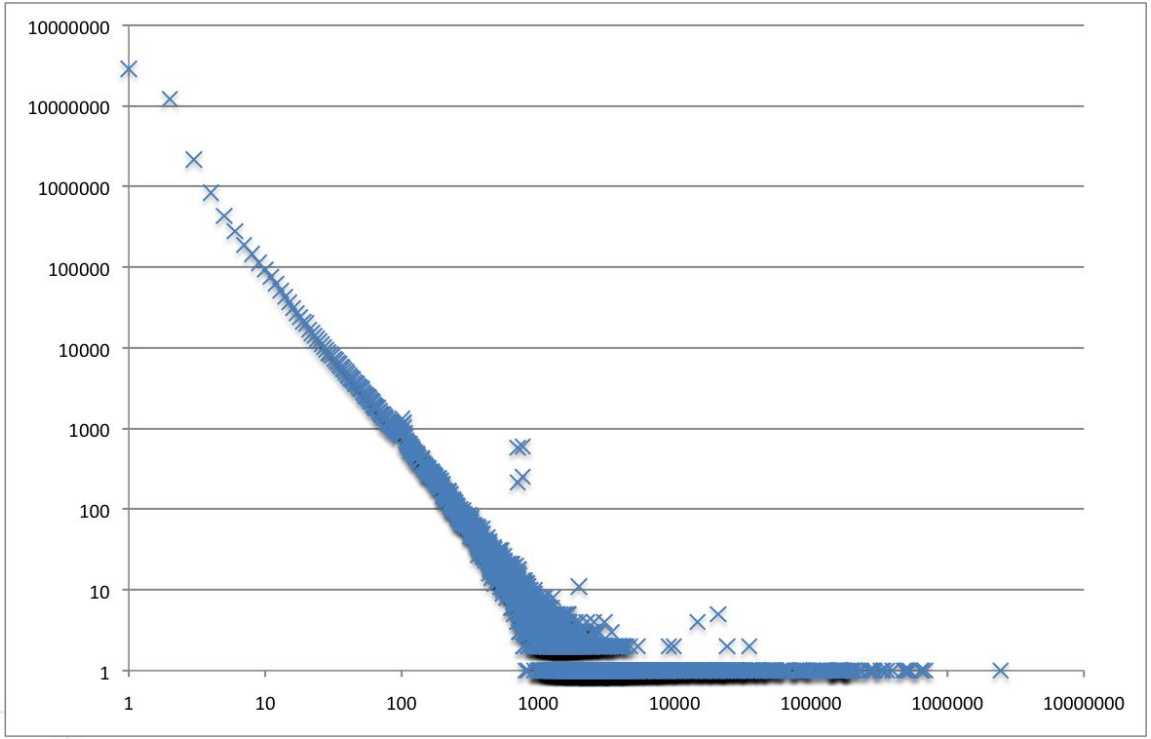


Small World Anomalies

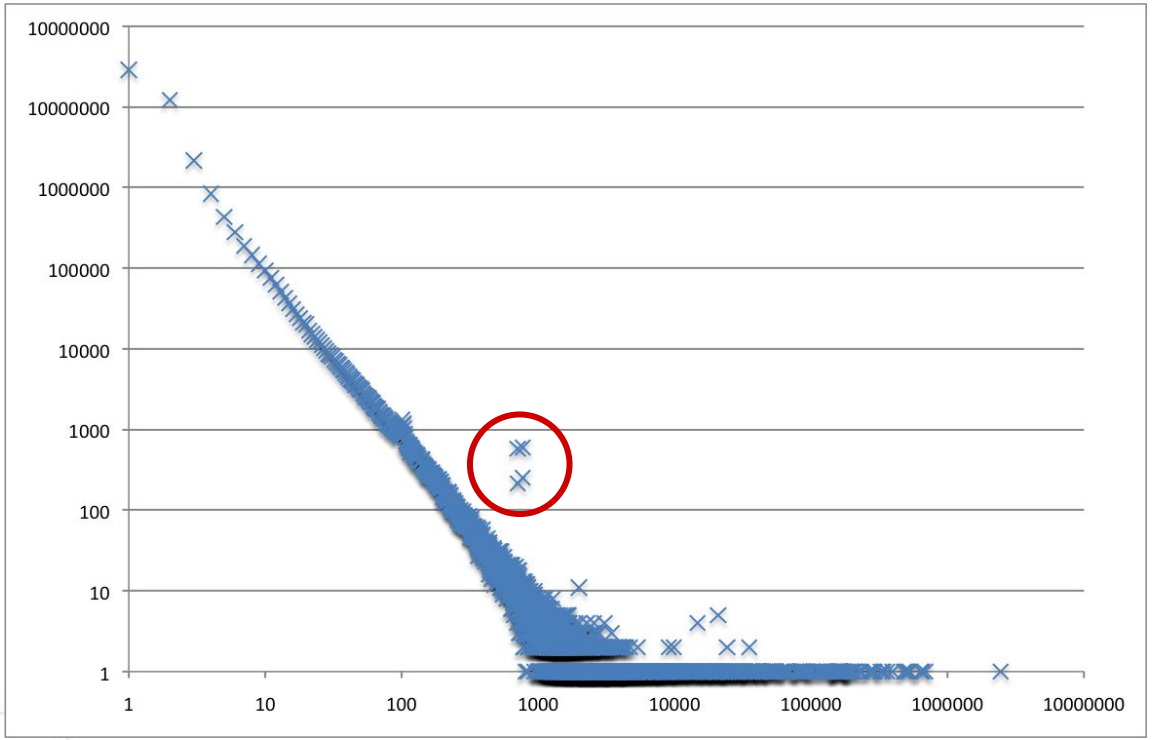
(whole graph)



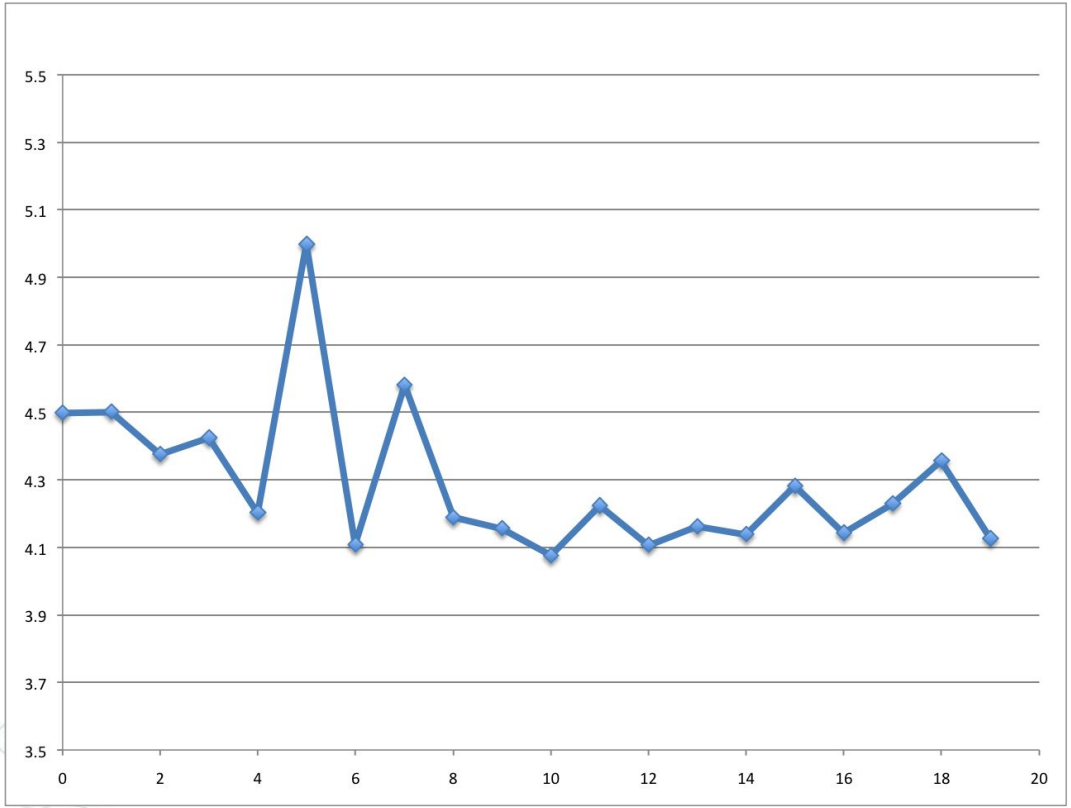
Indegree Distribution



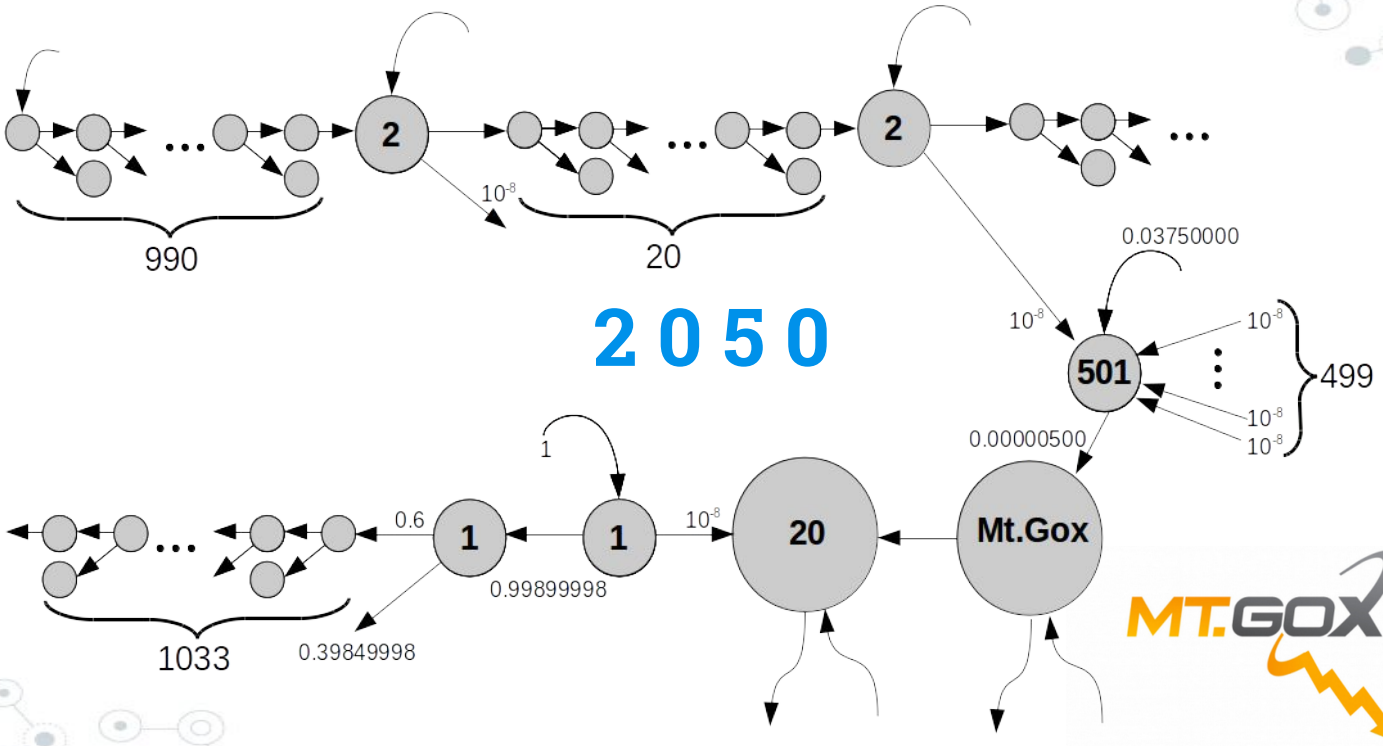
Indegree Distribution



Average Distance



Diameter


2
2
990
20
10^{-8}
0.03750000
10^{-8}
501
499
10^{-8}
10^{-8}
10^{-8}
0.00000500
2050
1
1
0.6
1
10^{-8}
20
Mt.Gox
0.99899998
1033
0.39849998
MT.GOX

An interesting transaction: 35dead89c059e846e2013a06a70cd84a7ba0f80da7741c283d6efd573e0a7319

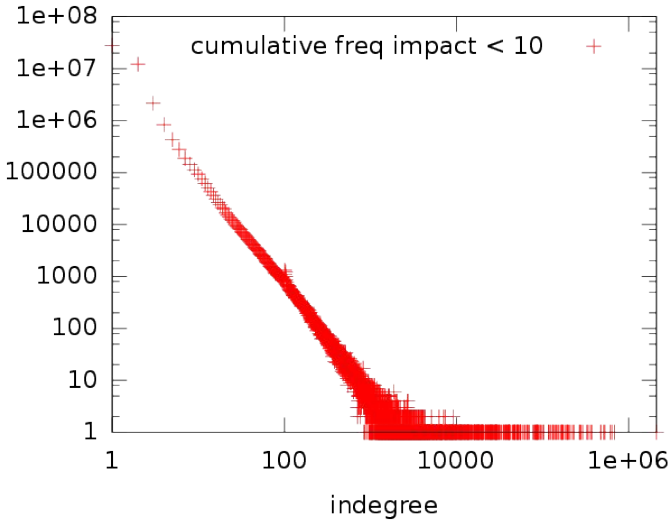
13hBbRdWGLCDNkDR8Pyf6jw3wM3q4EwbTF (0.0929742 BTC - Output)

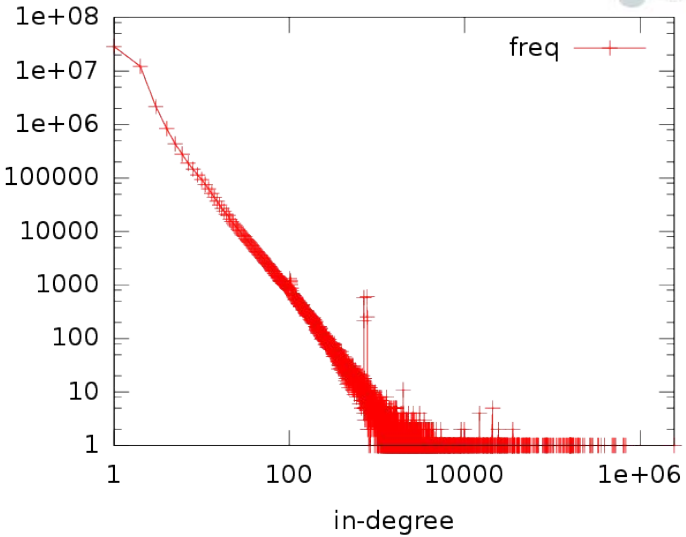
1G8zjUzeZBfJpeC...	(Jay_Pal) - (Unspent)	0.00001 BTC
14bFbRFJQXZTPG...	(jaronm) - (Unspent)	0.00001 BTC
1C3NNQ8Y7Qa7Q...	(Jambo) - (Spent)	0.00001 BTC
1DhFR2vR8w17BV...	(jarly) - (Unspent)	0.00001 BTC
14B3GFdtBZqxgNp...	(jarekb) - (Unspent)	0.00001 BTC
1MAVk5dcd3v9m8...	(Jameson) - (Unspent)	0.00001 BTC
19V4ITPAYUPT7Lx...	(jaredk) - (Unspent)	0.00001 BTC
143SikKpjzwhBy5Z...	(Jan) - (Unspent)	0.00001 BTC
12ZyGLt2ps2Agpy6...	(jak0b) - (Unspent)	0.00001 BTC
1PxBSHEZ3EbyZ4...	(jake262144) - (Unspent)	0.00001 BTC
1KseGadiZnuFoHo...	(jalidi) - (Unspent)	0.00001 BTC
1JadeSJWfvCuMA...	(Jade-) - (Unspent)	0.00001 BTC
1JALLENn5SAy4F...	(jallen	
1GsNNdHjYFSf5mBgHRwqJ	LEYiMcHYTA - (Spent)	0.0912558 BTC
1BgKdr9cZc2yFpX...	(jaken	
1EPY398Xk6cgqP...	(jabberwok) - (Unspent)	0.00001 BTC
19QkqAza7BHFTu...	(jackjack) - (Unspent)	0.00001 BTC
1jackkHHCLtxryJF...	(jackmanlnov) - (Unspent)	0.00001 BTC
17tqSDHshYLthDv...	(jacoder) - (Unspent)	0.00001 BTC
1CXwyYjrc9Vock5...	(jacroe) - (Unspent)	0.00001 BTC
16jVo7Vj7QdZ4Ap...	(jziHost.tk) - (Unspent)	0.00001 BTC
1FzcqruNwbXgkDk...	(ixne) - (Unspent)	0.00001 BTC
1MZVFA1VZmYTP...	(ixlpixl) - (Unspent)	0.00001 BTC
1BSGbFq4G8r3uck...	(JA37) - (Unspent)	0.00001 BTC
1aombYbEyggW4u...	(iyeman) - (Unspent)	0.00001 BTC
1FkN3XYsvuNdd2i...	(J.) - (Unspent)	0.00001 BTC
1DbQYvpXSVB7G...	(itsybitsy) - (Unspent)	0.00001 BTC
1CmUQsc87rByVi...	(itsmine) - (Unspent)	0.00001 BTC
19e3fcoLTu8YVFA...	(ivica) - (Unspent)	0.00001 BTC
1Lhifrdmq563mZxY...	(ivanish) - (Unspent)	0.00001 BTC
16Zo6werKP4akoTt...	(ItsASpork) - (Unspent)	0.00001 BTC

Classification:

- Pseudo Spam Transactions and Chains
- Generic Pseudo Spam Transactions and Chains
- Economical meaning

PSchains & GPSchains pruning

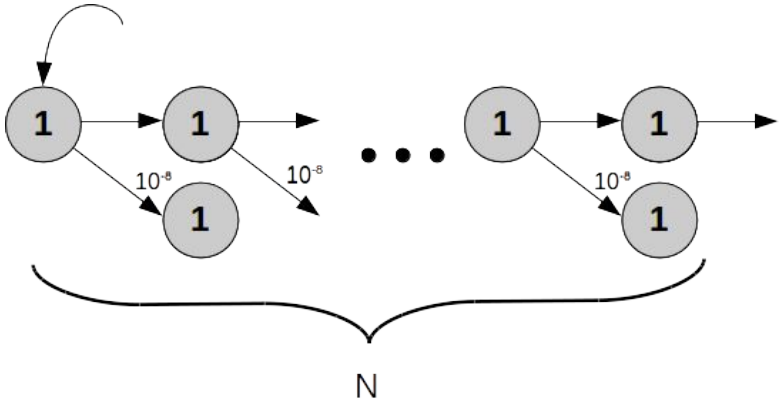




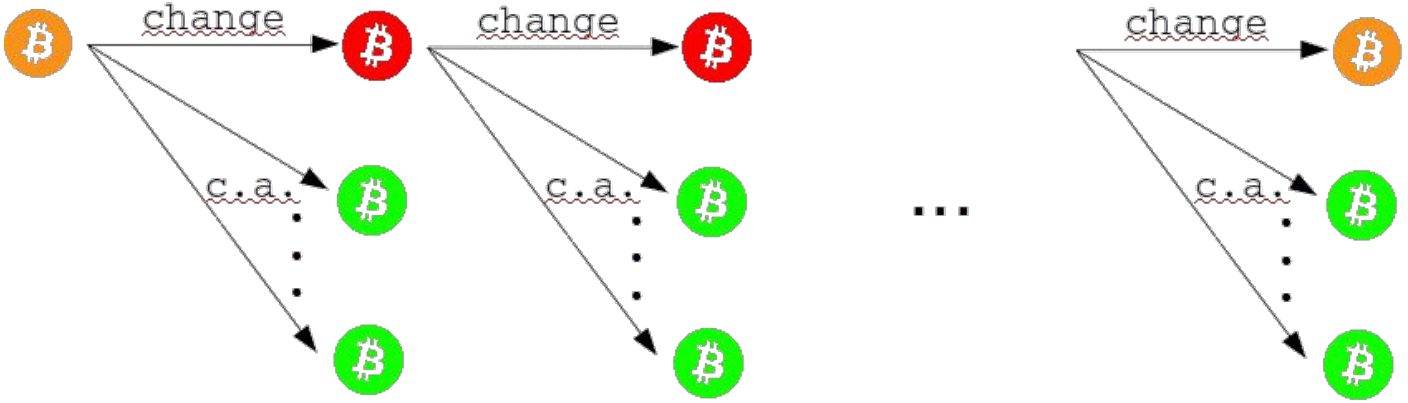
9 (575)

2050

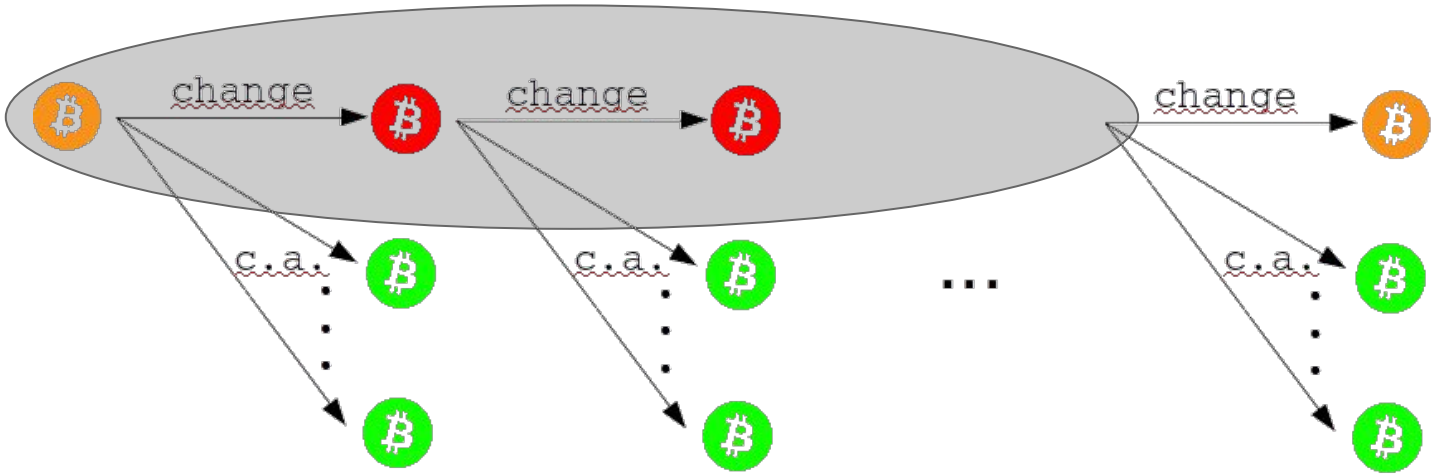
Detecting Patterns



Chain Heuristic:



Chain Heuristic:



Deanonymization Attack

blockchain

transactions graph

heuristic based clustering

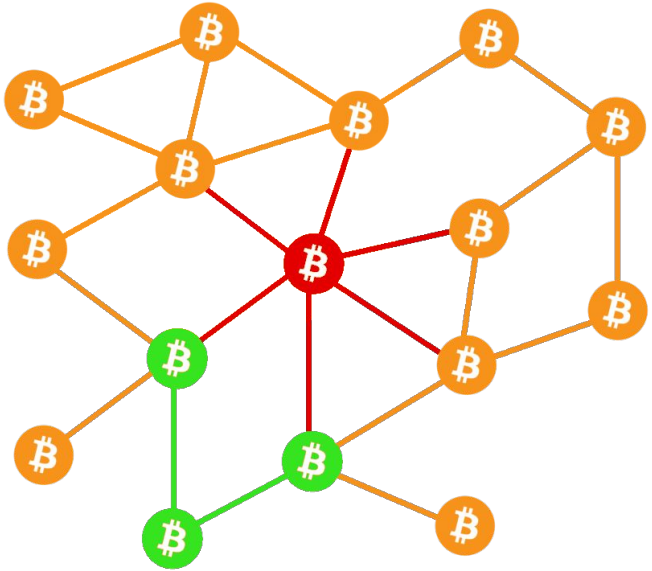
users graph

external informations

identities graph

External Informations

Network Listener (IP)



Eidgenössische Technische Hochschule Zürich