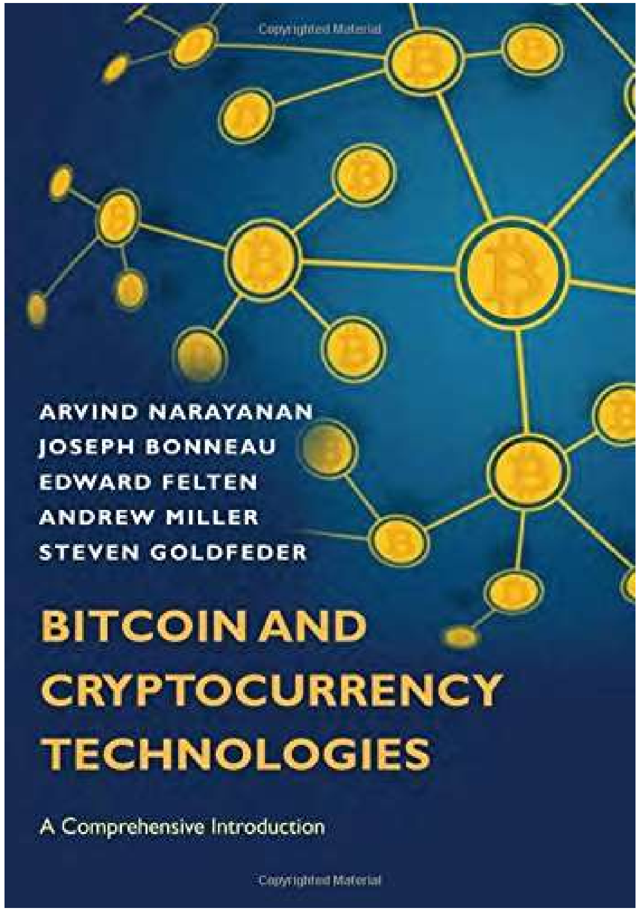
Blockchain
&
Business Application

Lecture:

Bitcoin & Anonymity

Chapter 6.1-6.4:Bitcoin & Anonymity



ARVIND NARAYANAN
JOSEPH BONNEAU
EDWARD FELTEN
ANDREW MILLER
STEVEN GOLDFEDER

BITCOIN AND CRYPTOCURRENCY TECHNOLOGIES

A Comprehensive Introduction

“Bitcoin is a secure and anonymous digital currency”

— WikiLeaks donations page

“Bitcoin won't hide you from the NSA's prying eyes”

— Wired UK

- Is Bitcoin anonymous?

- Do we want a cryptocurrency truly anonymous?

- Anonymity
  - Pros
  - Cons

Basic & Hard Questions

- Is having an anonymous cryptocurrency beneficial for the stakeholders?
- Is it good for society?
- Is there a way to isolate the positive aspects of anonymity while doing away with the negative parts?

Basic & Hard Questions

- Answers depend on one’s ethical values.
  - We won’t answer them
- We will examine arguments for & against anonymity
- We will study various technologies
  - some already present in Bitcoin
  - others that have been proposed to be added to it

- We’ll also look at proposals for altcoins that have different anonymity properties from Bitcoin.
- These technologies raise new questions:
  - How well do they work?
  - How difficult would they be to adopt?
  - What are the tradeoffs to be made in adopting them?

Anonymity Basics
Definition

- Literally: “Without a name”
- In Bitcoin, 2 possibilities: Interacting,
  - without using your real name, or
  - without using any name at all
- They lead to very different conclusions!

- Bitcoin addresses are hashes of public keys. You don't need to use your real name in order to interact with the system
    - Bitcoin is anonymous as you do not use your real name

- But you do use your public key hash as your identity
    - it is not; the address that you use is a pseudo-identity.

- Language of Comp. Science (CS), this middle ground: ***pseudonymity.***

- In CS, anonymity refers to
  - pseudonymity together with unlinkability .
- Unlinkability:
  - Defined wrt the capabilities of a specific adversary.
- Means
  - if a user interacts with the system repeatedly,
  - these different interactions should not be able to be tied to each other from the point of view of the adversary

- Bitcoin is pseudonymous, but
- Pseudonymity is not sufficient for privacy
- Blockchain is public:
  - If your address can be linked to your real ID
  - All your (ppf) transactions will be linked
- Linking is easy
  - Business apps usually ask your real ID
  - Credit card, shipping address.. At a Café?

- Side channels → Deanonymizing
  - Observing Bitcoin activities
  - Observing social media activities, linking them

- **Unlinkability**: It should be hard to link
  - Together different addresses of the same user
  - Together different transactions by the same user
  - The sender of a payment to its recipient (tricky!)

- **Anonymity set**
  - Sending bitcoins not directly but tru a long route?
  - No: Each appears on the chain and can be traced

- Don’t try full anonymity, try a limited set
- Given a particular adversary,
  - the anonymity set of your transaction:
  - the set of transactions which the adversary cannot distinguish from your transaction.
- Even if the adversary knows you made a transaction,
- they can only tell that it’s one of the transactions in the set,
- but not which one it is. → maximize size of the set

- Calculating the set is tricky
- Taint Analysis (popular)
  - How related two addresses are
  - If transactions always S→T, then a high taint score
- Not a good measure
  - Assumes adversary always uses same calculation
  - A more clever adversary; AI?

- Cryptocurrency anonymity problem worse than traditional banking
  - Banking: Revealing a single or few transactions
  - Bitcoin: Revealing all (ppf) transactions!

- Developing cryptocurrencies technically infeasible to reveal?

Ethical Issues

- We usually hide our salaries
  - What if we receive in Bitcoins?
- Business secrets
  - R&D
  - subcontractors
  - new products
- Money laundering; illegal business?
  - Transactions may be anonymous
  - Not the trade between coins&cash

Good, Bad?

- Can’t we design a technology
  - Taking the good
  - Leaving the bad?
- No!
  - Depends on moral viewpoint, what is good/bad
  - Can’t ask miners to include good/bad transactions

- Solution? Separate the technical anonymity from the legal principles

Anonymization vs. decentralization

- Perfect anonymity requires controls
  - blind-signature protocols with central authority

  →Destroying decentrality

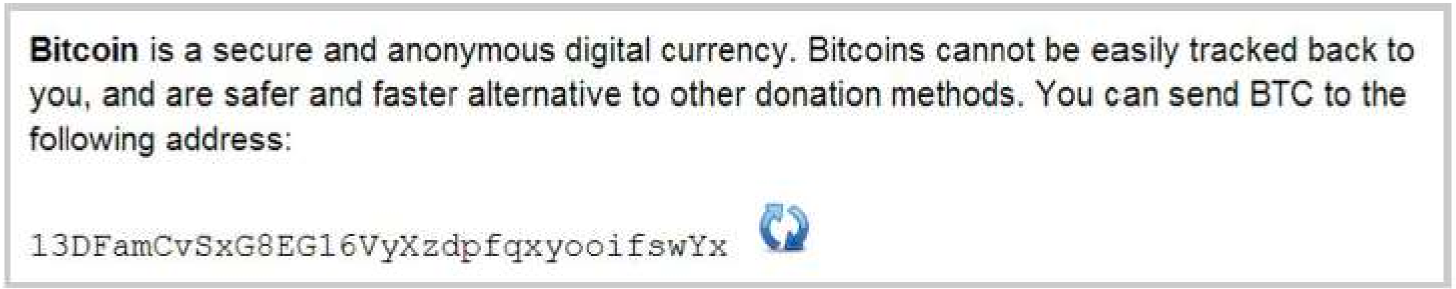
- Perfect decentralization requires controls
  - Preventing double-spend, public traceability

  →Destroying anonymity

- Anonymous decentralization?
  - Zerocoin, Zerocash (later)

How to De-anonymize Bitcoin

- Bitcoin is only pseudonymous,
  - all of your transactions or addresses
  - could ***potentially*** be linked together



Bitcoin is a secure and anonymous digital currency. Bitcoins cannot be easily tracked back to you, and are safer and faster alternative to other donation methods. You can send BTC to the following address:

```
13DFamCvSxG8EG16VyXzdpfqxyooifswYx
```

Figure 6.1: Snippet from Wikileaks donation page. Notice the refresh icon next to the Bitcoin address. Wikileaks follows the Bitcoin best practice of generating a new receiving address for every donation.

  - At the first glance, transactions look unlinkable
  - Each donation received (spent?) separately

shared spending is evidence of joint control

- Suppose Alice
  - has 5, 3 and 6 coins in different addresses
  - Wants to buy a teapot of 8 coins

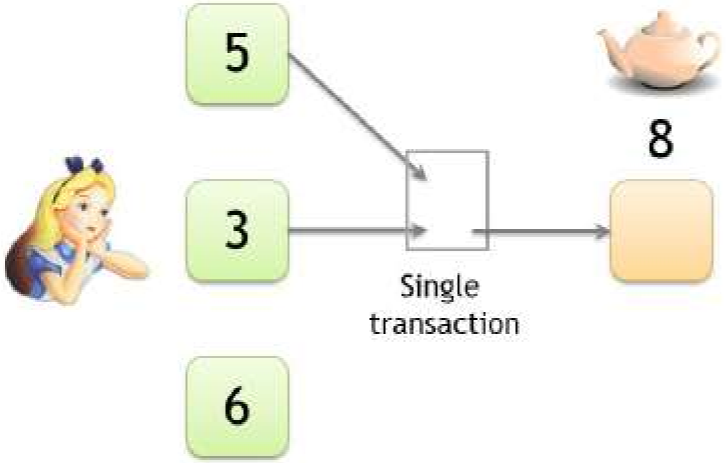


Figure 6.2 : To pay for the teapot, Alice has to create a single transaction having inputs that are at two different address. In doing so, Alice reveals that these two addresses are controlled by a single entity.

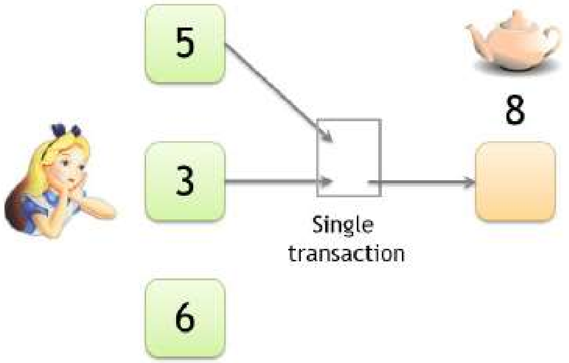


Figure 6.2 : To pay for the teapot, Alice has to create a single transaction having inputs that are at two different address. In doing so, Alice reveals that these two addresses are controlled by a single entity.

- Adversary can reveal that these two addresses are control by same person
- Adversary can continue observing transactions to link more

- Wallet software play a vital role in anonymity

- Suppose the pot was 8.5 coins

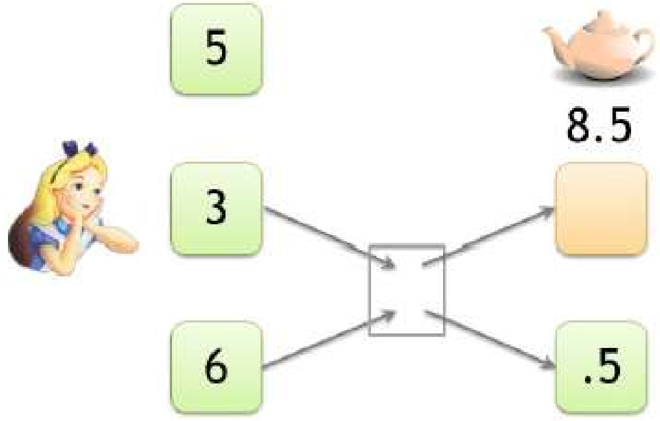


Figure 6.3: Change address. To pay for the teapot, Alice has to create a transaction with one output that goes to the merchant and another output that sends change back to herself.

- One output: store’s payment address
- The other: “change” address owned by herself
  - Heuristic!

Idioms of Use

- In 2013, researchers found in most wallets,
  - Change addresses are brand new
  - No change output addresses existing

→Providing a powerful heuristic on addresses

→Transaction Graph Analysis (TGA)

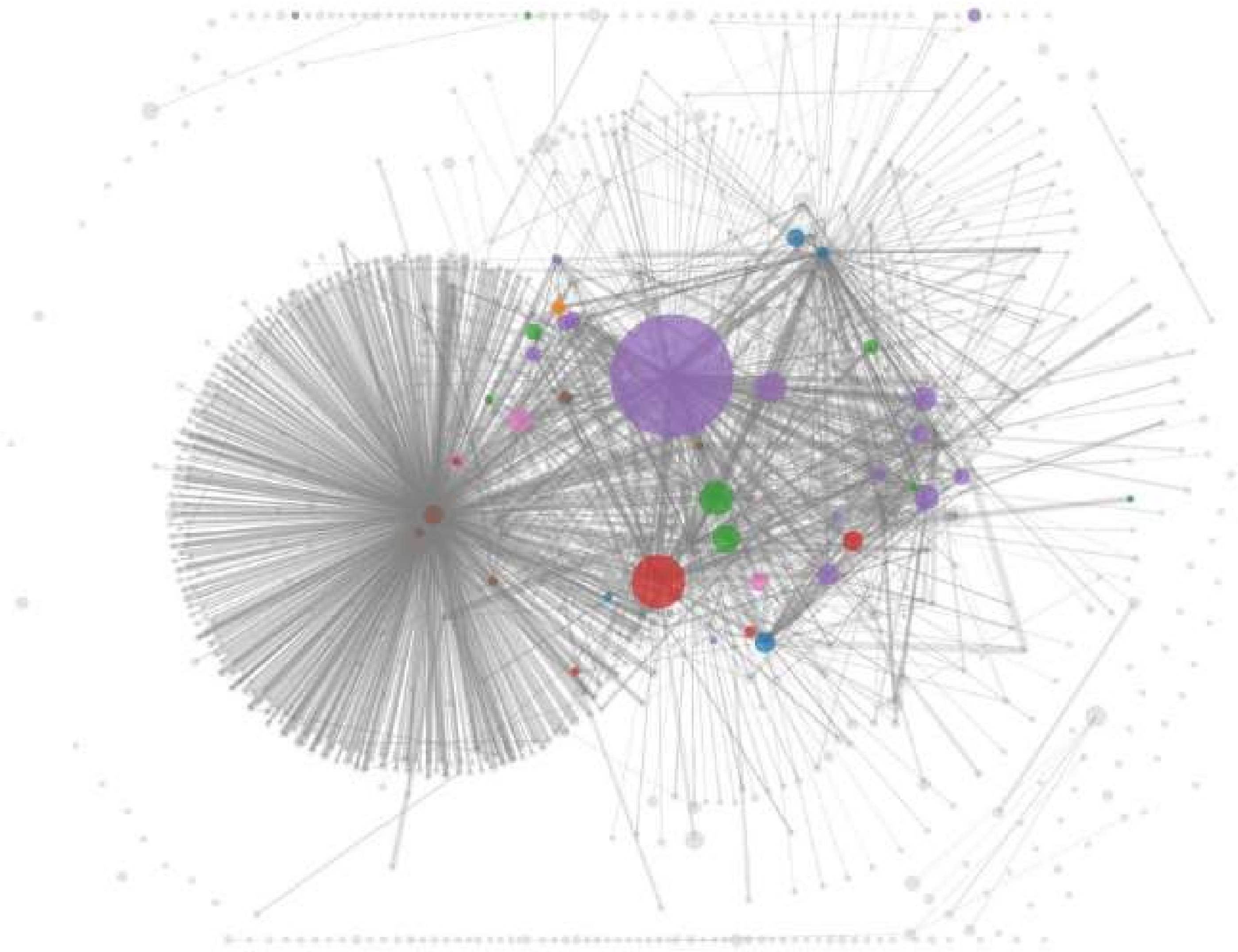


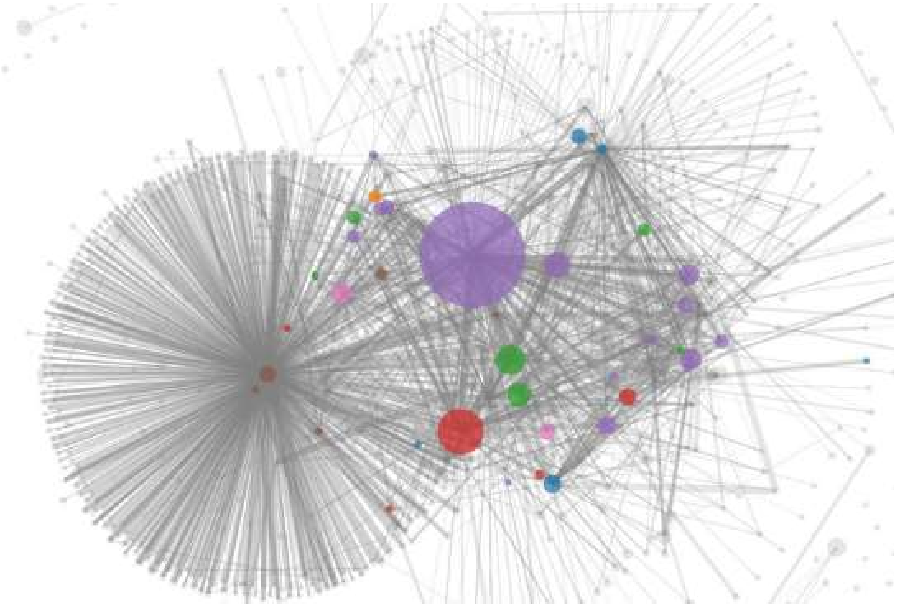
Figure 6.4: Clustering of addresses. In the 2013 paper *A Fistful of Bitcoins: Characterizing Payments Among Men with No Names,* researchers combined the shared-spending heuristic and the fresh-change-address heuristic to cluster Bitcoin addresses. The sizes of these circles represent the quantity of money flowing into those clusters, and each edge represents a transaction.

Attaching Real World IDs

- Mt. Gox was the largest exchange
  – Purple?

- Satoshi Dice (gambling)
  – Brown small?

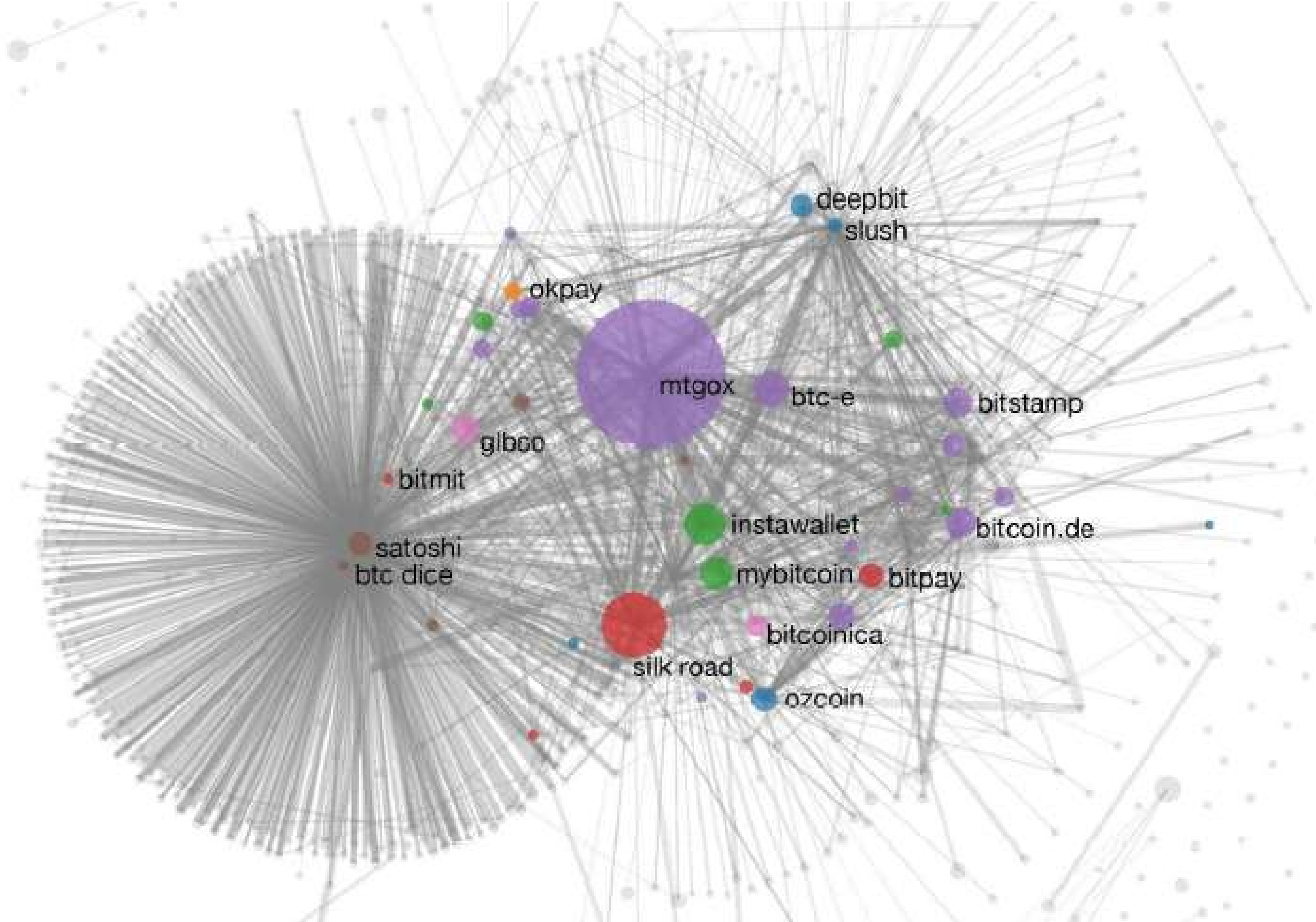
  – This method works only for a few dominant



Tagging by transacting?

- Go to their websites, learn the addresses?
- Does not work:
  - Each is a new address (remember wikileaks)
  - That address will never be added to the chain

- Make a real transaction: Send them coins ☺


deepbit
slush
okpay
mtgox
btc-e
bitstamp
glbco
bitmit
instawallet
bitcoin.de
satoshi
btc dice
mybitcoin
bitpay
bitcoinica
silk road
ozcoin

Identifying individuals

- Can we do the same thing for individuals
- Can we link addresses to real life IDs?

Directly transacting:

Anyone who transacts with an individual

  - an online or offline merchant,
  - an exchange,
  - a friend who splits a dinner bill using Bitcoin

- knows at least one address belonging to them

Via service providers:

- Almost everyone interacts with exhanges
- They have to ask the Real IDs

→Law enforcement can reveal Real IDs

Carelessness:

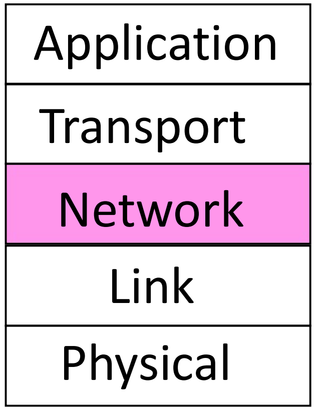
- People post their public addresses on forums

Things get worse over time
Transaction graph analysis improves

Network-layer deanonymization

- Alternative to transaction graph analysis
- Each transaction posted/announced
- Not necessarily added to the chain

- Network terminology
  - Blockchain: Application layer
  - P2P Network: Network layer



Application

Transport

Network

Link

Physical

- When a node creates a transaction, it connects to many nodes at once and broadcasts the transaction.
- If sufficiently many nodes on the network collude with each other (or are run by the same adversary),
- They could figure out the first node to broadcast any transaction.
- Presumably, that would be a node that’s run by the user who created the transaction → IP Address!

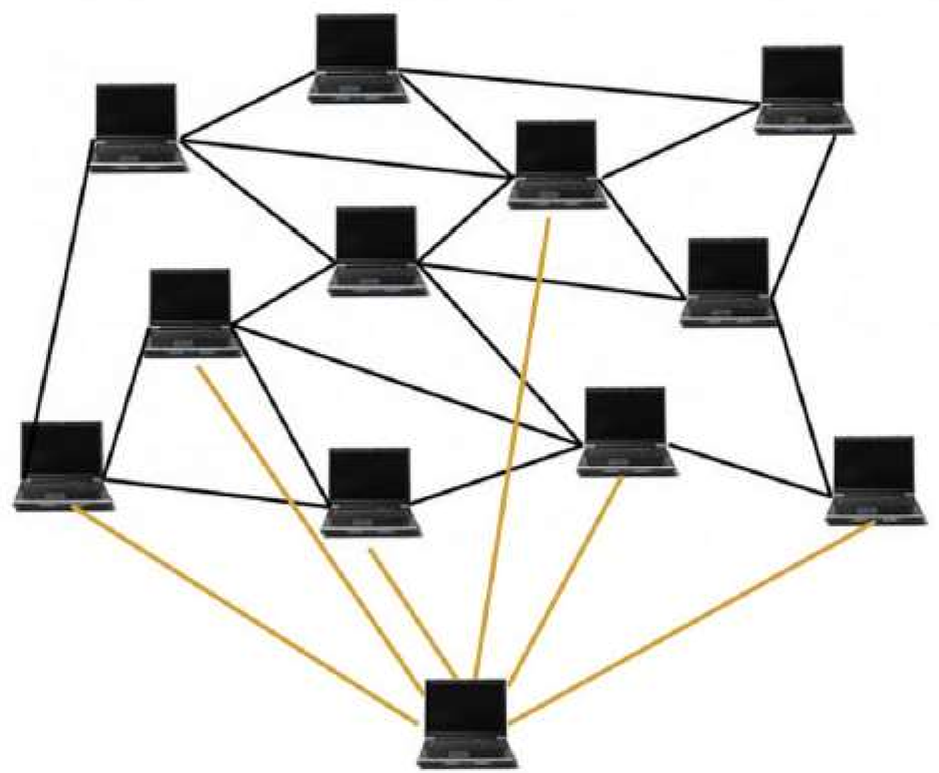
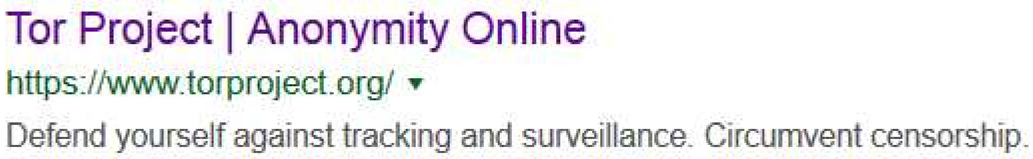


Figure 6.6*. *Network level deanonymization. As Dan Kaminsky pointed out in his 2011 Black Hat talk, “the first node to inform you of a transaction is probably the source of it.” This heuristic is amplified when multiple nodes cooperate and identify the same source.

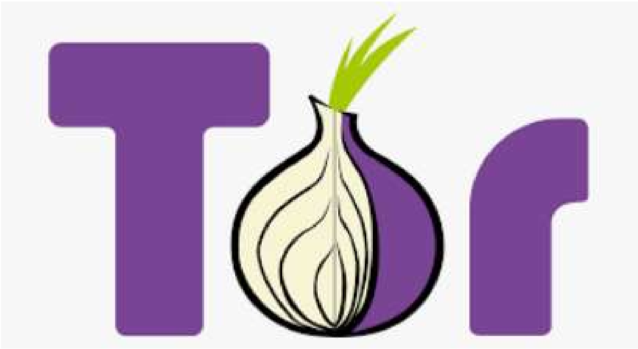
- But that is a problem of
  - “communication anonymity”
- There is ongoing research and solutions
- Such as Tor



Tor Project | Anonymity Online

https://www.torproject.org/ ▾

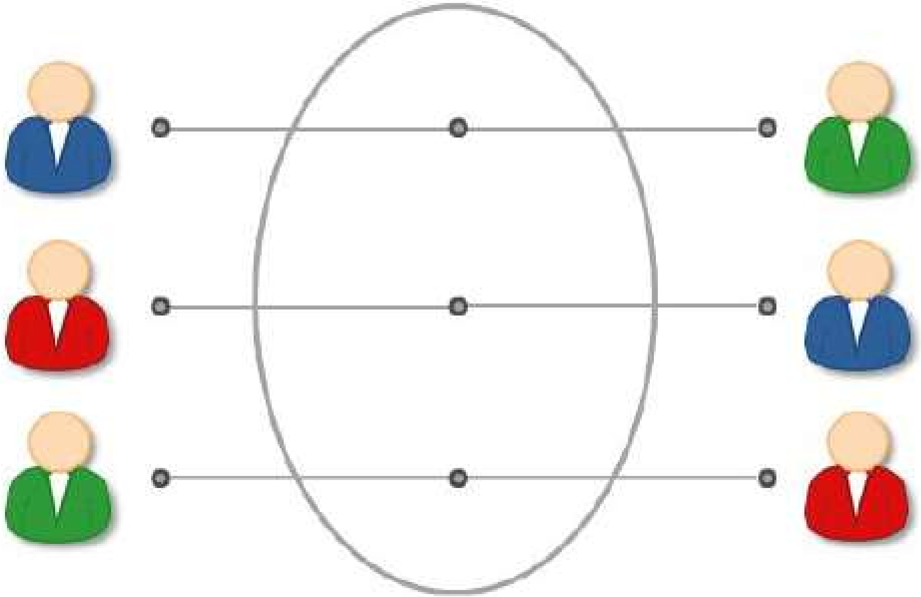
Defend yourself against tracking and surveillance. Circumvent censorship.



- Transaction Graph Analysis is more serious!

Mixing

- Attempts to make TGA less effective
- Not specific to cryptocurrencies (anonymity)
- If you want anonymity, use an ***intermediary***



Online wallets as mixes; as intermediaries.

- Services where you can
  - store your bitcoins online,
  - withdraw later.
  - Typically the coins that you withdraw won’t be the same as the coins you deposited.
- Support anonymity up to some degree
  - Anyways keep records, some ID info..

Dedicated mixing services.

- Promise not to keep records, nor require your ID
- No even need a username or other pseudonym

- You send your bitcoins to an mix's address
- You tell the mix a destination address to send bitcoins to.
- Mix will soon send you (other) bitcoins at
- It's essentially a swap

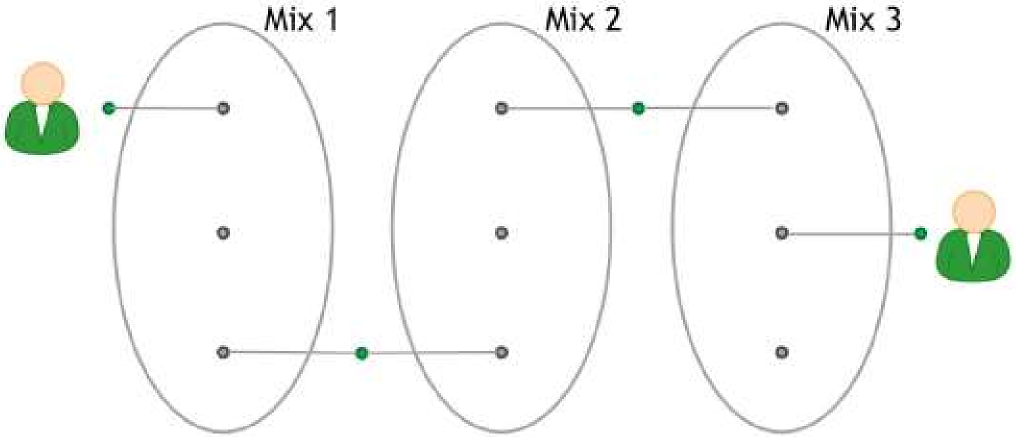
Dedicated mixing services.

- They “promise”
- You need to “trust”

- Mixing terminology
  - Mix vs. mixer? Mix ☺
  - Laundary? Nope!

Use a series of mixes

- Make it even more mixed
  - Graph algorithms are demanding!
- More difficult to trace on the graph
- At least one mix supposedly destroy records
- e.g. Tor uses a series of 3 routers



Uniform Transactions

- If transactions of
  - different users
  - different quantities
- Then it would be straightfw to trace
- Determine a ***chunk size***
- Huge size? Small transactions cannot
- Small size? Big transactions need to split a lot
- Multiple standard chunk sizes
- Tradeoff: Privacy vs. Efficiency

- ***Client side should be automated***
- Adversary observing attack
  - Not only amounts
  - But also timing
- Timing: Too complex by humans
- Interaction with mixes should be automatic

- ***Mixing Fees should be all-or-nothing.***
  - Not transaction fees but mixing fees
- Cutting a % from each transaction
  - destroy sizes of chunks
  - leads to vulnerabilities
- Instead of cutting ***0.1%*** of each transaction,
- Swallow ***1*** whole transactions among all ***1000***
- Mixer should convince that 0.1%, not 1%

- ***Mixing in practice***
  - No truly functioning mixing service
  - Many reported to steal coins
  - Users do not prefer mixing services
  - ***Anonymity loves company***
  - No strong anonymity with few users
- Mixers are all independent
- Not following the model

Decentralized Mixing

- Idea:
  - Replace mixers with a P2P network of users
  - Fits better to blockchain's philosophy
- Practical advantages
  - No need to wait for trusted service (bootsrapping)
  - Theft is impossible
- Several proposals, main: ***Coinjoin***

Coinjoin

- Different users jointly create single transaction
- Order of inputs & outputs randomized
  - – Participants check output addresses & amounts

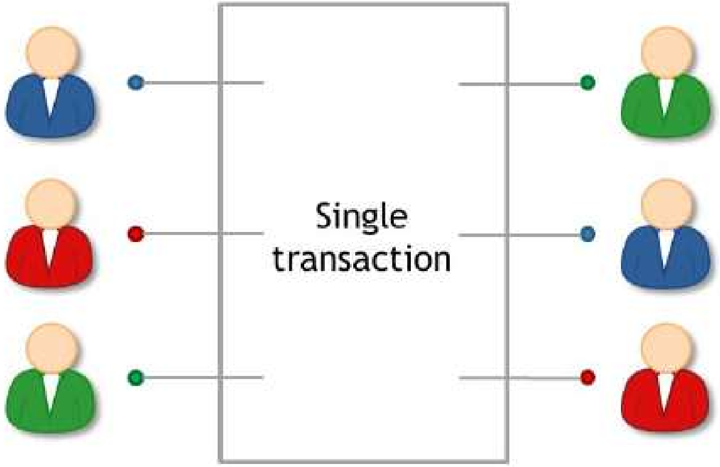


Figure 6.9. A Coinjoin transaction.

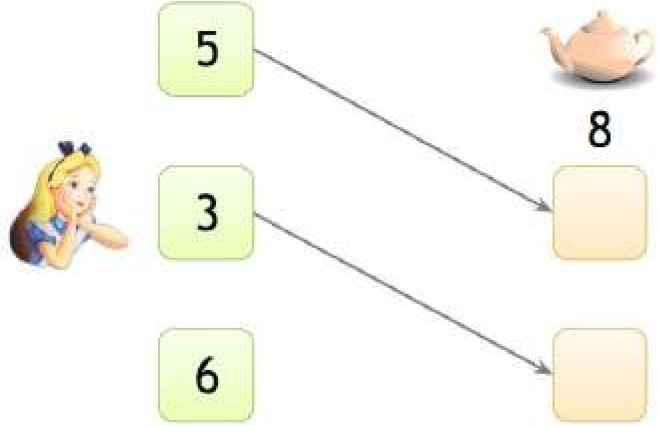
- Five steps of Coinjoin
  - – Find peers who want to mix
  - – Exchange input/output addresses
  - – Construct transaction
  - – Send the transaction around. Each peer signs after verifying their output is present. !!!
  - – Broadcast the transaction

High Level Flows

- Side channels can be very tricky
  - Suppose Alice receives 43.12312 BTC, each month
  - She immediately transfers 5% of that amount to her retirement account, which is another Bitcoin address.
- No mixing strategy can effectively hide the relationship between these two addresses
- The specific amounts and timing are extraordinarily unlikely to occur by chance.

Merge Avoidance

- Normally: Combine inputs for a single output
- Why not receiver provides many addresses?



- Alice makes payments at arbitrary times

Zerocoin and (its successor) Zerocash

- “Ingenious Cryptography”
- Others try to add anonymity on the protocol
- Zerocoin adds mixing at the protocol level
- Compatibility:
  - Zerocoin not directly compatible with Bitcoin
  - Can become by a soft forking
  - Zerocash cannot. It requires an altcoin

- Cryptographic guarantees
  - Protocol level mixing
  - No need to trust to anyone (mixes, intermediaries)

Zerocoin (extension of Basecoin)

- Basecoin: A Bitcoin like altcoin
- Convert Basecoins → Zerocoins → Basecoins
  - Link between original and new Basecoin is broken
- Basecoin acts like the currency
- Zerocoin provides a way for anonymity
- Your Zerocoins prove:
  - You owned Basecoins (not which)
  - You made them unspendable (Casino example)

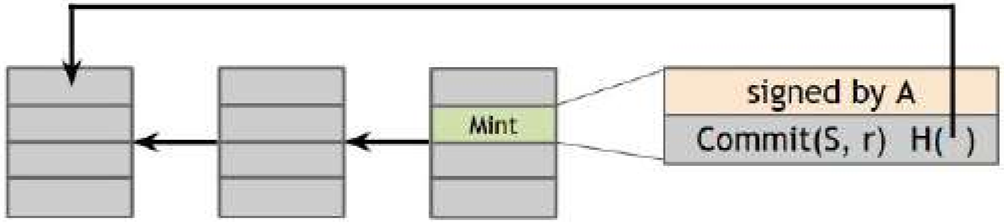
- **Zero-knowledge proofs.**
  - The key cryptographic tool
  - A way for somebody to prove a (mathematical) statement
  - Without revealing any other information that leads to that statement being true.
  - We'll treat ZKP as a black-box

"I know x such that H(x) belongs to the following set: {...}".

- Minting Zerocoins
  - Anyone can mint at any time, for any value
  - Suppose 1 Basecoin is worth 1 Zerocoin
  - You can mint 1 Zerocoin (free money? ☺)
  - To put in blockchain, have to give up 1 Basecoin
- Cryptographic commitments

- Three steps
  - Generate serial number ***S*** and a random secret ***r***
  - Compute ***Commit(S, r)*** ,
  - Publish the commitment onto the block chain

This burns a basecoin, making it unspendable, and creates a Zerocoin. Keep ***S*** and ***r*** secret for now.



To spend a Zerocoin

- To redeem a new basecoin,
  - need to prove that previously minted a zerocoin.
  - You could do this by opening your previous commitment, that is, revealing ***S*** and *r*.
  - But this makes the ***link*** between your old basecoin and your new basecoin apparent.
  - How can we break the link?
  - Zero-knowledge-proof!

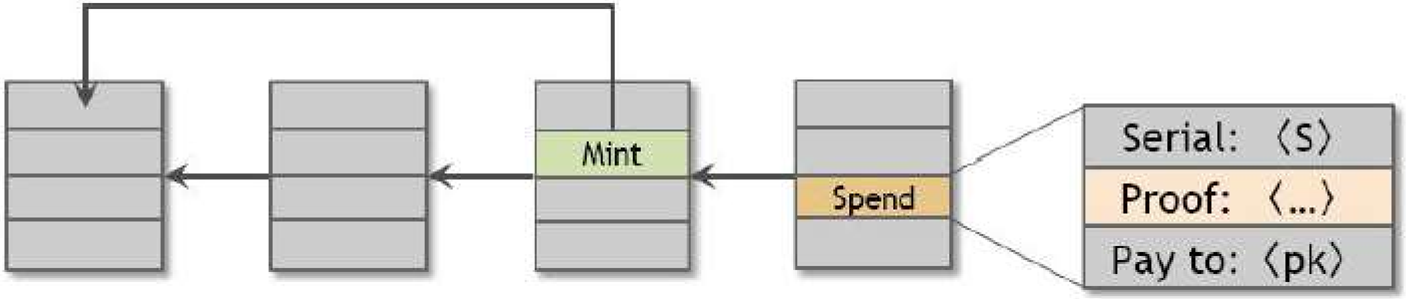
- Create a special “spend” transaction that contains S, along with a zero-knowledge proof of the statement:

“I know ***r*** such that ***Commit(S, r)*** is in the set ***{ c 1 ,c 2 ,...,c n }***”.

- Miners will verify your proof which establishes your ability
  - To open one of the zerocoin commitments on the block chain,
  - without actually opening it.
- Miners will also check that S has never been used
  - (double-spend)
- The output of your spend transaction → a new basecoin.

  For the output address, you should use an address that you own.

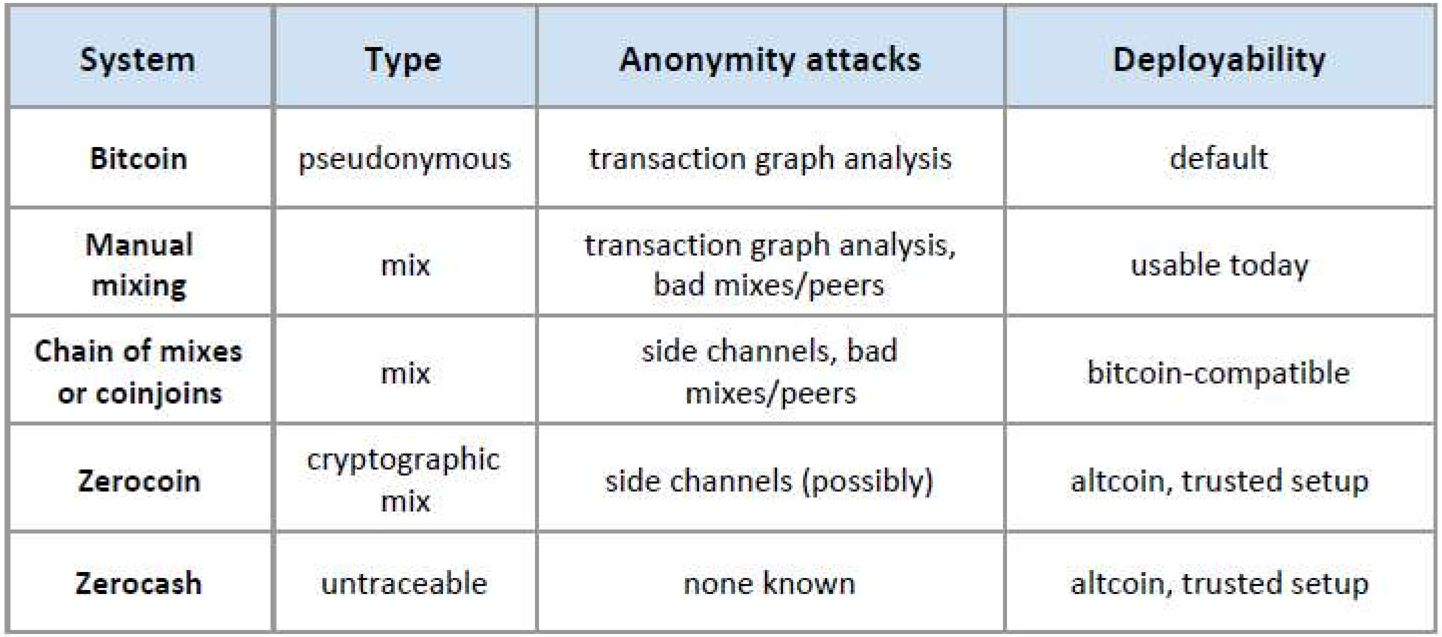
- Spending a Zerocoin



- **Anonymity**.
  - Observe that ***r*** is kept secret throughout;
  - Neither mint nor spend transaction reveals it.
  - Nobody knows which s.number ***S*** corresponds to which zerocoin.
  - This is the key concept behind Zerocoin's anonymity.

Zerocash

- A different anonymous cryptocurrency
- Builds on the concept of Zerocoin
- Takes the cryptography to the next level.
- It uses a zero-knowledge SNARKs (zk-SNARKS)
  - which are a way of making zero-knowledge proofs much more **compact** and ***efficient*** to verify.
- No need to a Basecoin



System	Type	Anonymity attacks	Deployability
Bitcoin	pseudonymous	transaction graph analysis	default
Manual mixing	mix	transaction graph analysis, bad mixes/peers	usable today
Chain of mixes or coinjoins	mix	side channels, bad mixes/peers	bitcoin-compatible
Zerocoin	cryptographic mix	side channels (possibly)	altcoin, trusted setup
Zerocash	untraceable	none known	altcoin, trusted setup

Table 6.14: A comparison of the anonymity technologies presented in this chapter