Bitcoin 网络、匿名、监管

吴嘉婧 副教授

中山大学

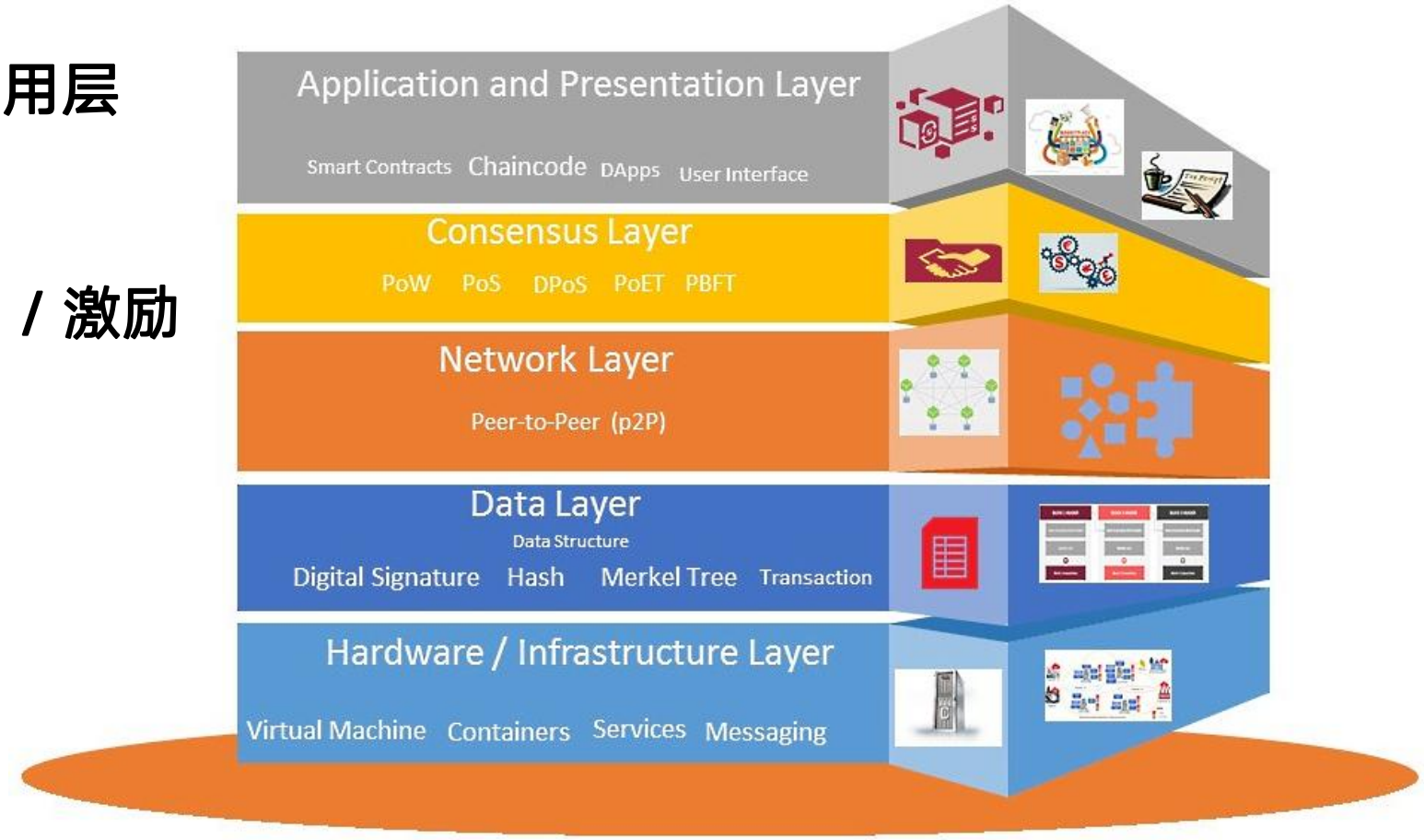
计算机学院

Outline & Keywords of this Class

- Part 1: 比特币 网络
- Part 2: 匿名
- Part 3: 监管

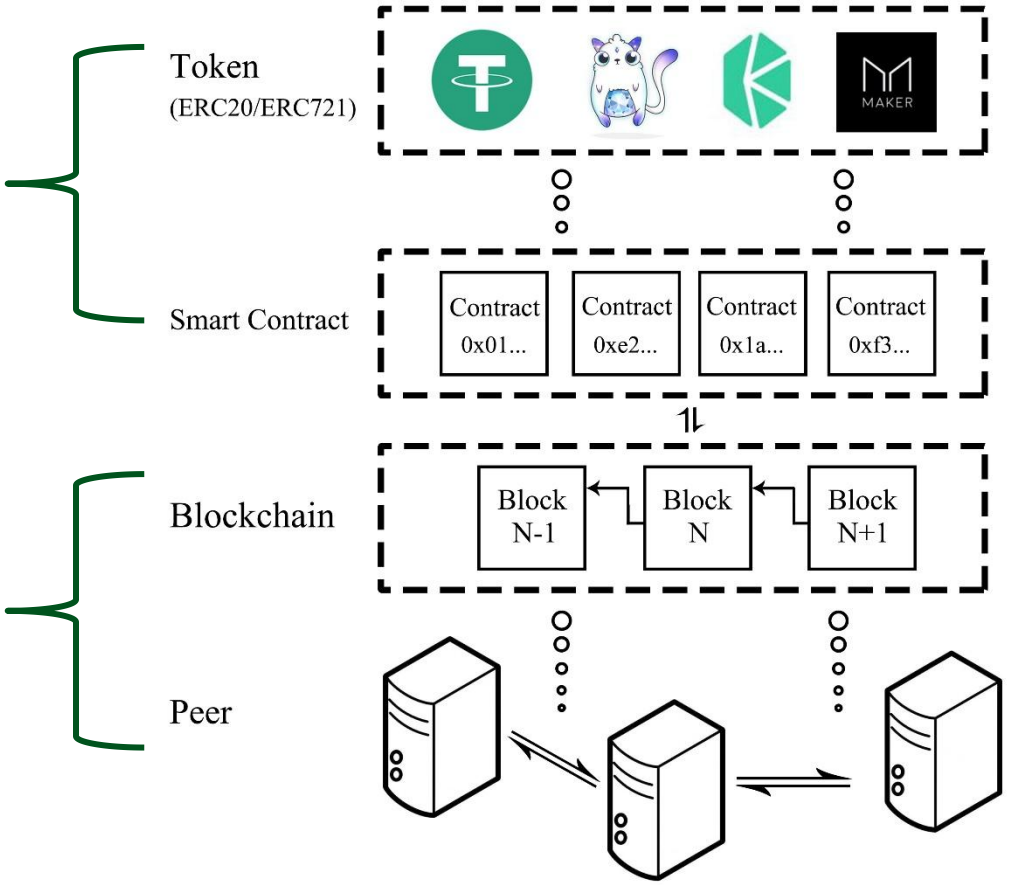
区块链的分层

- 业务应用层
- 合约
- 共识层 / 激励
- 网络层
- 数据层
- 硬件层



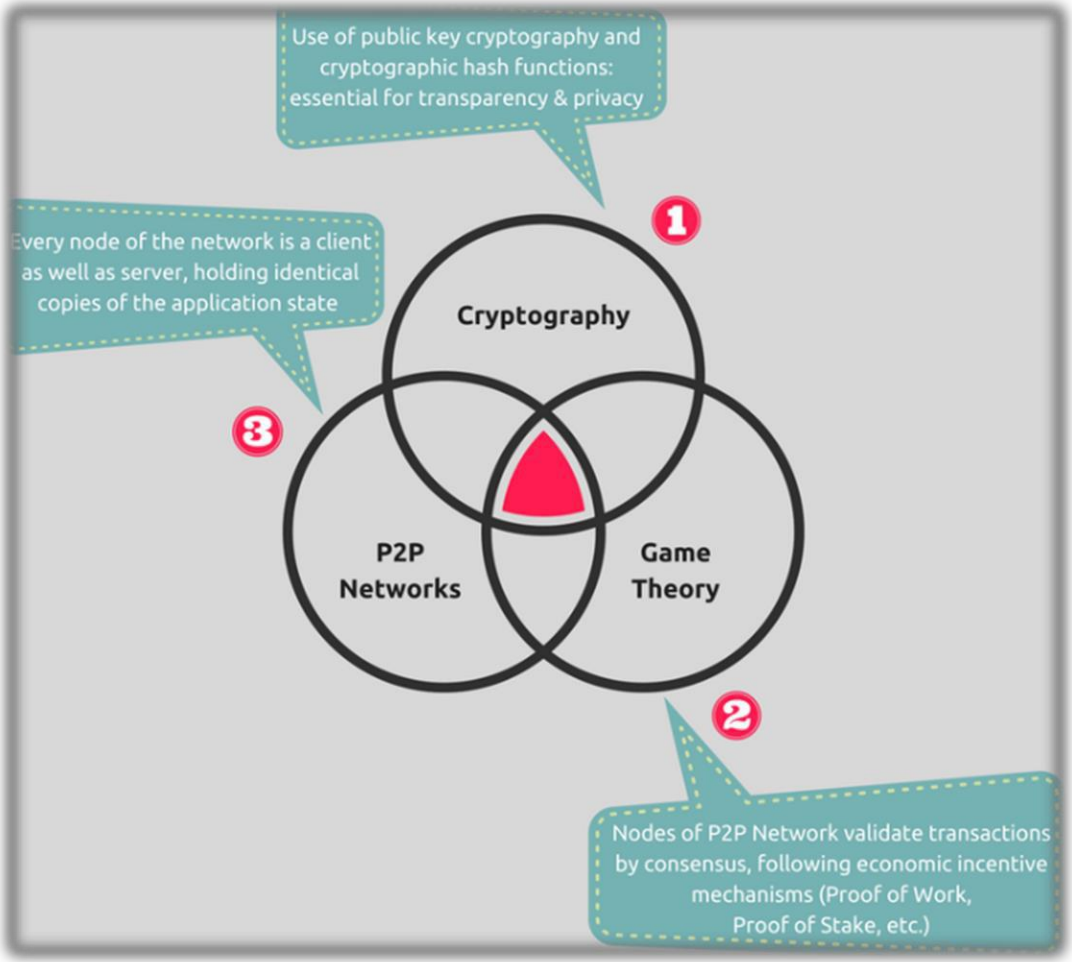
区块链的分层

- 业务/合约/激励
- 节点/网络/共识



区块链关键技术

- Blockchain is built on top of three key technologies

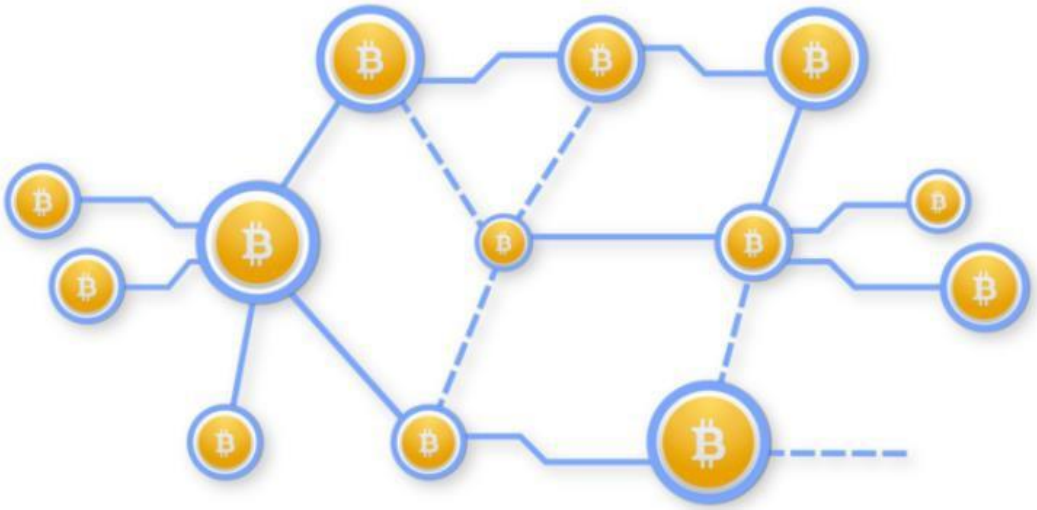


5

比特币底层 P2P 网络运行机制

The Bitcoin P2P Network

- Bitcoin Network
  - pure Peer-to-Peer principle
  - Bitcoin clients have to agree on account balances
  - Goal: Consistent view in the whole network



7

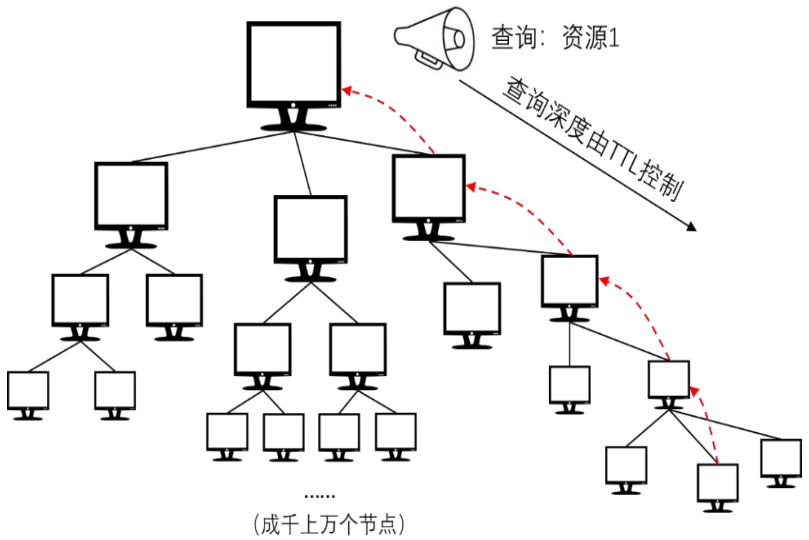
P2P Network Structures

- Bitcoin: Unstructured Peer-to-Peer network
  - Structured Network 结构化拓扑
    - Main advantage of structured networks – quick search of the specific information
    - Structured P2P networks overcome the limitations of unstructured networks by maintaining a Distributed Hash Table (DHT)
    - and by allowing each peer to be responsible for a specific part of the content in the network.
    - Not applicable to Bitcoin: It requires all nodes need (more or less) complete information
  - Unstructured 非结构化拓扑: No overhead for maintaining the structure
    - There is no guarantee that flooding will find a peer that has the desired data
    - Flooding also causes a high amount of signaling traffic in the network and hence such networks typically have a very poor search efficiency

8

拓扑形式

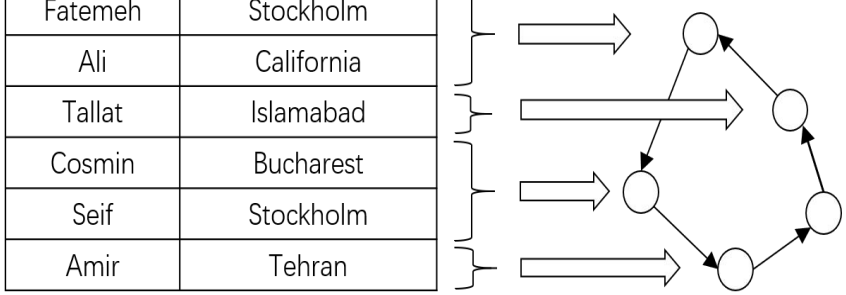
- 全分布式非结构化拓扑
  - 没有使用中心索引服务器，其节点拥有真正的对等关系
  - 洪泛（Flooding）数据广播，即节点会将接收到的消息向邻居节点转发，直到所有节点都接收到了这个消息或消息传播的深度到达一定的限制
- 特点
  - 可能会出现广播风暴
  - 实现快速的消息传播和资源查找



拓扑形式

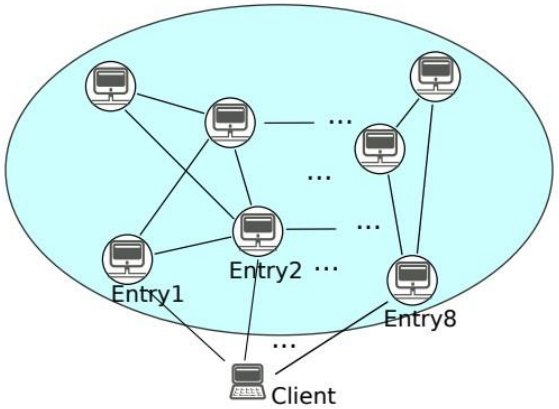
- 全分布式结构化拓扑
  - 采用分布式散列表（Distributed Hash Tables，简称DHT）来实现整个网络的寻址和存储，从而结构化地址管理
  - 分布式散列表将存储着网络中所有资源信息的散列表划分成很多不连续的小块，分散地存储在多个节点上
- 特点
  - 维护机制较为复杂
  - 良好的健壮性、可扩展性和动态适应性

Key	Value
Fatemeh	Stockholm
Ali	California
Tallat	Islamabad
Cosmin	Bucharest
Seif	Stockholm
Amir	Tehran



Joining the Peer-to-Peer network

- First step: Finding some other peers
  - Requires "cheating": Finding peers without some central system is difficult
  - Bitcoin's approaches
    - Use pre-configured IP addresses
    - Get IP addresses from an IRC channel (no longer used in the default setting)
    - Get IP addresses via the Domain Name System (DNS servers run by volunteers)



Connections in the Bitcoin network

- Node knows some IP addresses of other nodes
  - Node connects to a certain number (default: 8) of these nodes
  - Node accepts incoming connections beyond that limit
  - On average: About 30 connections per node that accepts incoming connections
- Inactive nodes are deleted from lists after timeout (several hours)

Bitcoin transactions and blockchain

- Individual transaction from A to B
  - A signs the transaction using the private key of his address
  - A broadcasts the transaction to the whole Bitcoin network
- Confirmation of transactions: through a block
  - Nodes (miners) collect TX to form a "**block**"
  - Miners append block to blockchain and compute a PoW
  - Successful miner broadcasts the block to the whole Bitcoin network

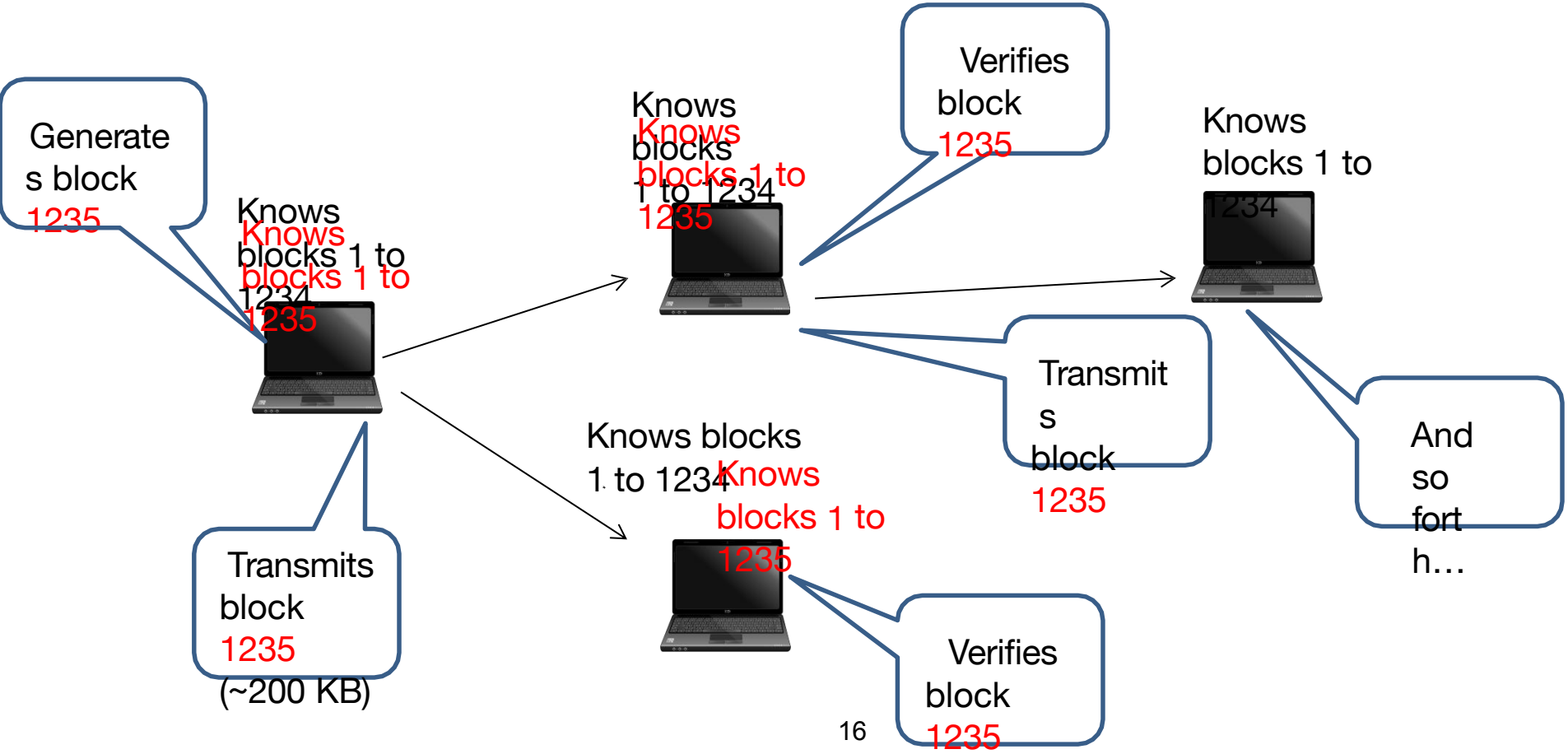
Broadcasting

- Sender informs all connected Bitcoin nodes about availability of a new TX / new block
  - *Invite message*
- On receipt of an *invite message*
  - Node requests the TX / block if it does not know it
  - Node verifies the TX / block based on local blockchain copy
  - Node informs all connected Bitcoin nodes about availability of a new TX / block

比特币 P2P 网络的一致性要求

How to reach a consistent view?

- Goal of the P2P network: Consistent view
  - Network becomes inconsistent once a new block is generated

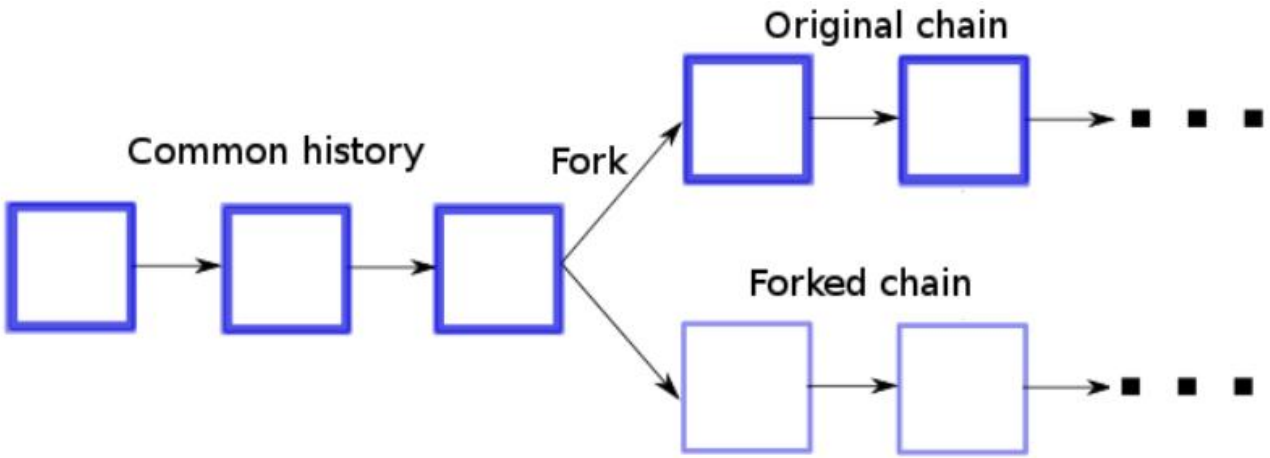


Information Propagation

- Investigation by *Decker* and *Wattenhofer* (Proc. IEEE P2P'13)
  - Connection to a large number of nodes, observation of information propagation
  - 结论1. Average time till a node receives a new block: 12.6 seconds
  - 结论2. Long tail: 5% of nodes do not have the new block after 40 seconds

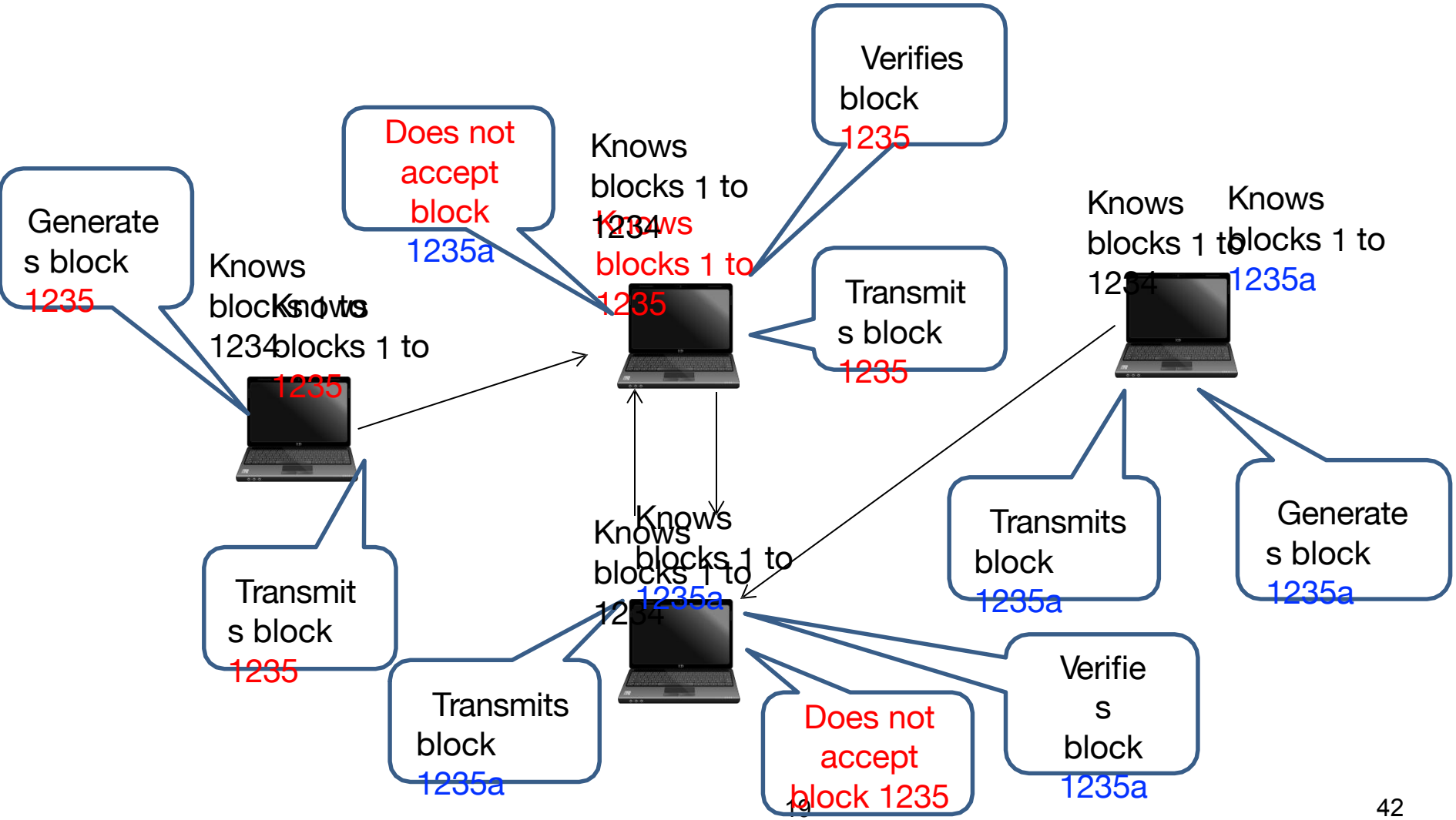
Information Propagation

- Problem of propagation time: Other miner may find a new block within that time
  - **blockchain fork:** two inconsistent versions of the blockchain
  - *Decker* and *Wattenhofer* observe 169 blockchain forks during a period of 10,000 generated blocks



18

An example of inconsistent view



Dealing with inconsistency

- Each miner continues with one version of the blockchain
  - First newly generated block leads to longest chain
  - All nodes switch to longest chain once that block has been received
- Transactions present in the shorter chain:
  - Not lost,
  - but integrated into the next block

Outline of this Class

- Part 1: 比特币 网络
- Part 2: 匿名
- Part 3: 监管

比特币网络的“匿名”特性

匿名 (Anonymity)

- 匿名：不使用名字
- 化名：不用真实姓名

- 比特币账户的地址：公钥哈希值

- 计算机科学中，匿名是
  - 无关联 (unlinkability) 的化名

- Questions: 匿名对加密数字货币
  - 有什么好处？
  - 有什么负面作用？

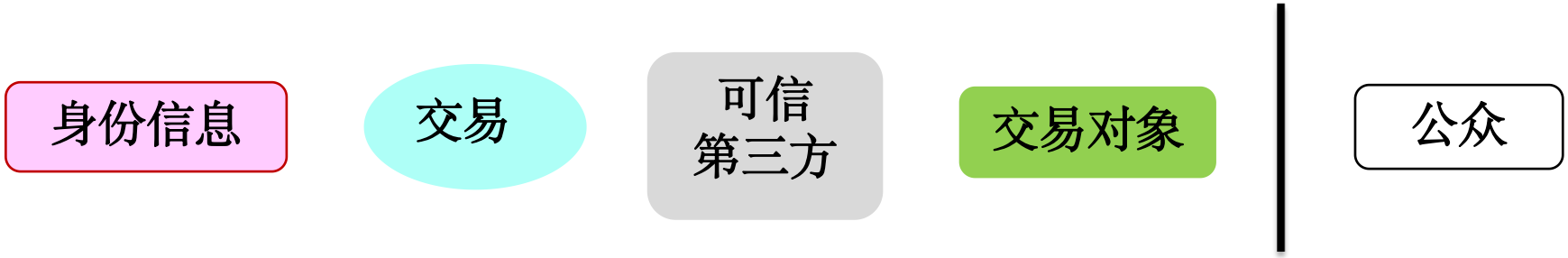
匿名性的必要性

- 为何人们需要匿名性?
  - 比特币是一个公链系统：Open
  - 一旦暴露身份，所有隐私不保

- 主要的担心：隐私问题

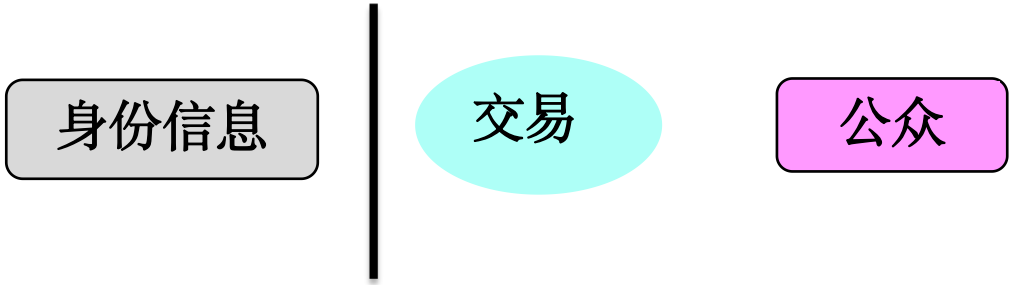
区块链隐私模型

古典互联网隐私模型



隔离： 巨头私有

区块链隐私模型



数据公开

隔离： 区块链匿名

Privacy issues in P2P Blockchain Networks

- Bitcoin privacy research concerning the TX graph (交易图谱)
  - linking different Bitcoin addresses of a user
  - 那么，你所有的TX (过去的，现在的，未来的) 都可以关联到你的身份

- **Concerns** in such the P2P network
  - Threat: Figure out origin (IP address) of a TX by finding the first node that broadcasts it
  - To Overcome:
    - Try to get connections to as many nodes as possible
    - Join the network under many fake identities to get many other nodes to connect to you

Privacy issues in Real-World Networks

- Bitcoin Exchange or BTC Wallet
  - They need your ID, & Credit Card
  - 那么，某些比特币业务可以关联到你的身份
- **Concerns** in such the Real-World network
  - Threat:
    - 比特币支付：使你暴露, they don't even have to know your name
    - 旁敲侧击：交易活跃时间 与 社交账号活跃时间有关联
    - 污点分析：推算两个地址相关性， Sender, Receiver 相对固定
  - To Overcome:
    - 需要更强的无关联性属性

27

交易的“无关联性”

- 几个关键属性
  - 同一个用户的不同地址应该不易关联
  - 同一个用户的不同交易应该不易关联
  - 一个交易的交易双方应该不易关联

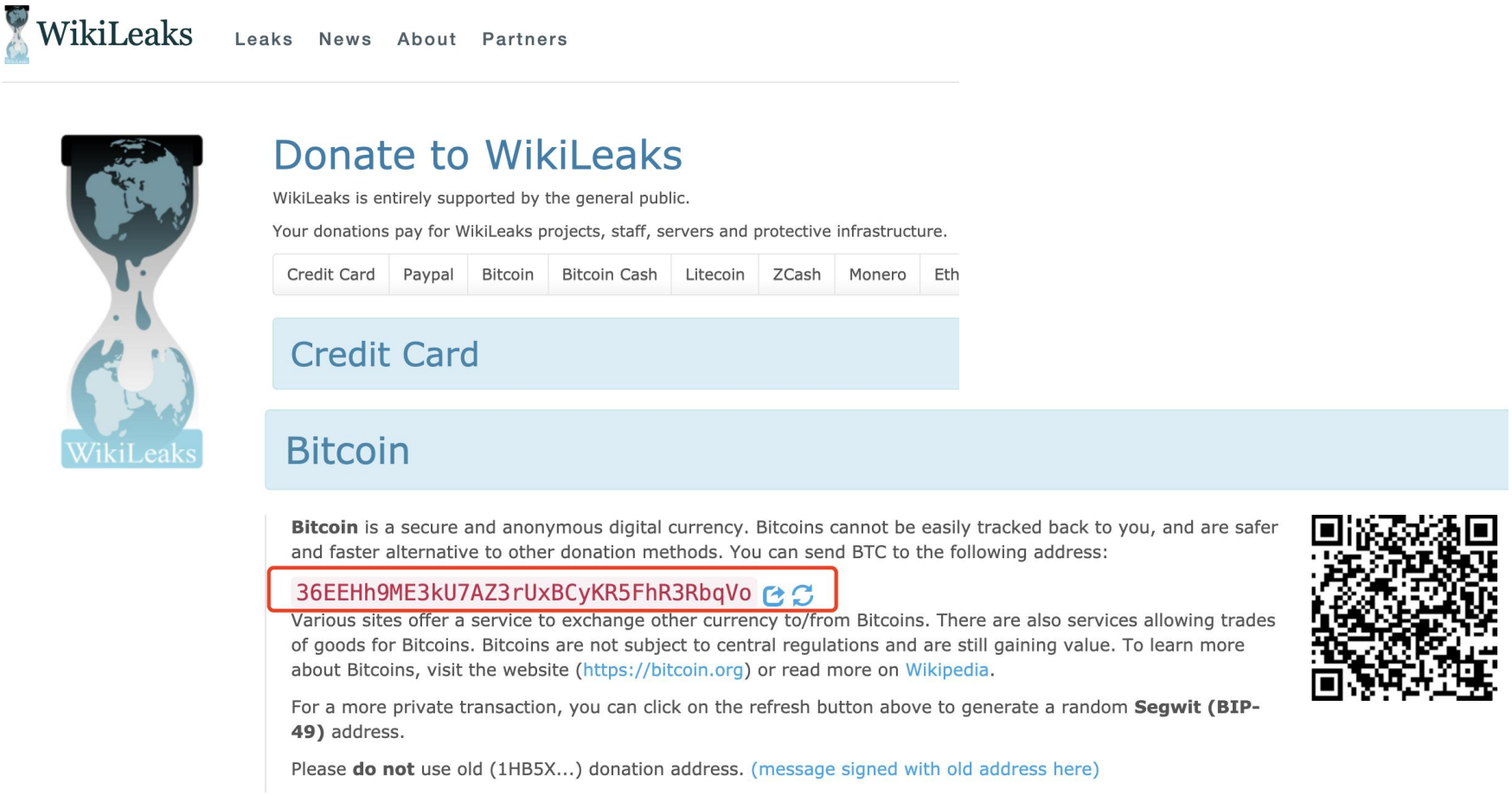
对比特币“匿名”的讨论

匿名性的讨论

- 匿名化 与 去中心化
  - 乔姆 (Chaum) 发明的 e-cash 系统：中心化的匿名方案，依赖于一个中央权威机构——银行的盲签协议
  - 如果强制去中心化，需要有一种能够追踪交易并且能防止双花的机制——对匿名化的威胁

如何 对比特币的使用 匿名化

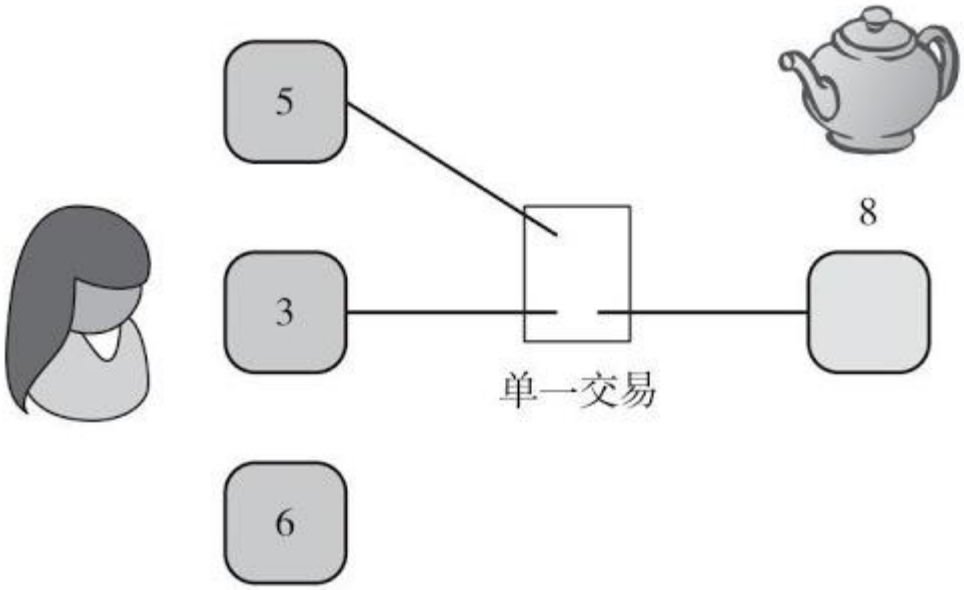
- 维基解密的例子—— 经常更换自己的收款地址



如何对比特币 去匿名化?

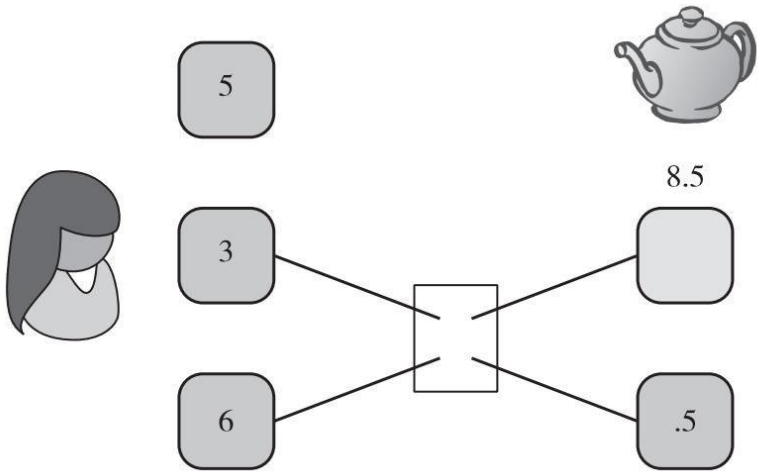
- 维基解密的例子
  - 更换不同的收款地址
  - 这些地址之间一定是无法关联的吗?

- 如何推测出不同地址之间的关联性?
  - 多地址输入交易
  - 共同输入
  - 共同控制



通过“零钱地址”来合理推测

- 如果水壶涨价到8.5 BTC：找零
- 可以推断
  - 两个输入地址属于同一个用户
  - 甚至其中一个输出地址也属于该用户
  - 惯用法则：
    - 零钱地址通常是被钱包软件新创造出的地址，和输入地址关联



其他 advanced“去匿名化”方法

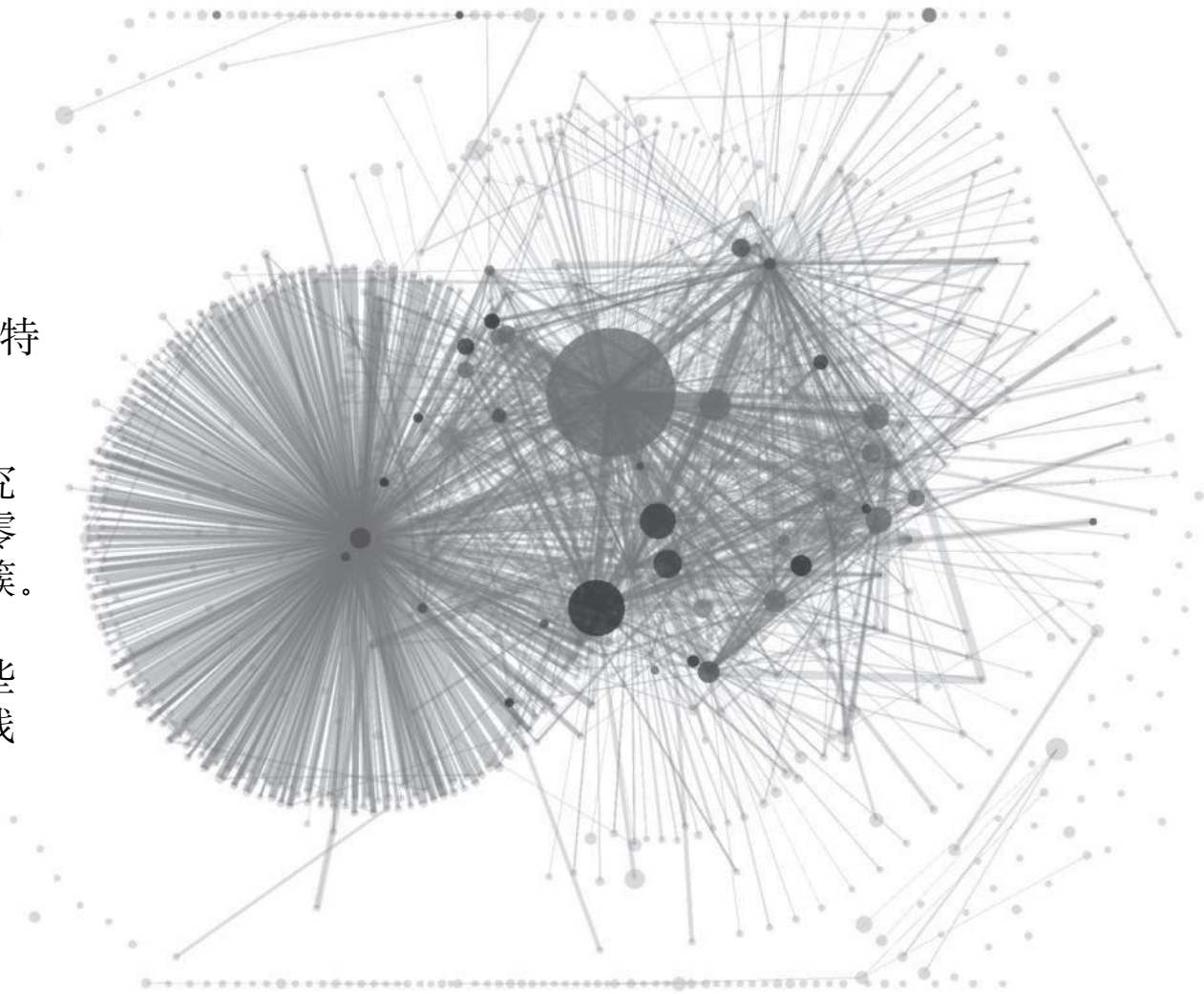
关联真实世界的身份到地址簇

地址簇

摘自2013年的一篇论文“一把比特币：寻找支付特征”。

在一组没有姓名的用户中，研究者将联合支付的地址和全新的零钱地址归类到一个比特币地址簇。

图中，圆形的大小表示流入这些地址簇里的货币数量，每一条线则代表一个交易。



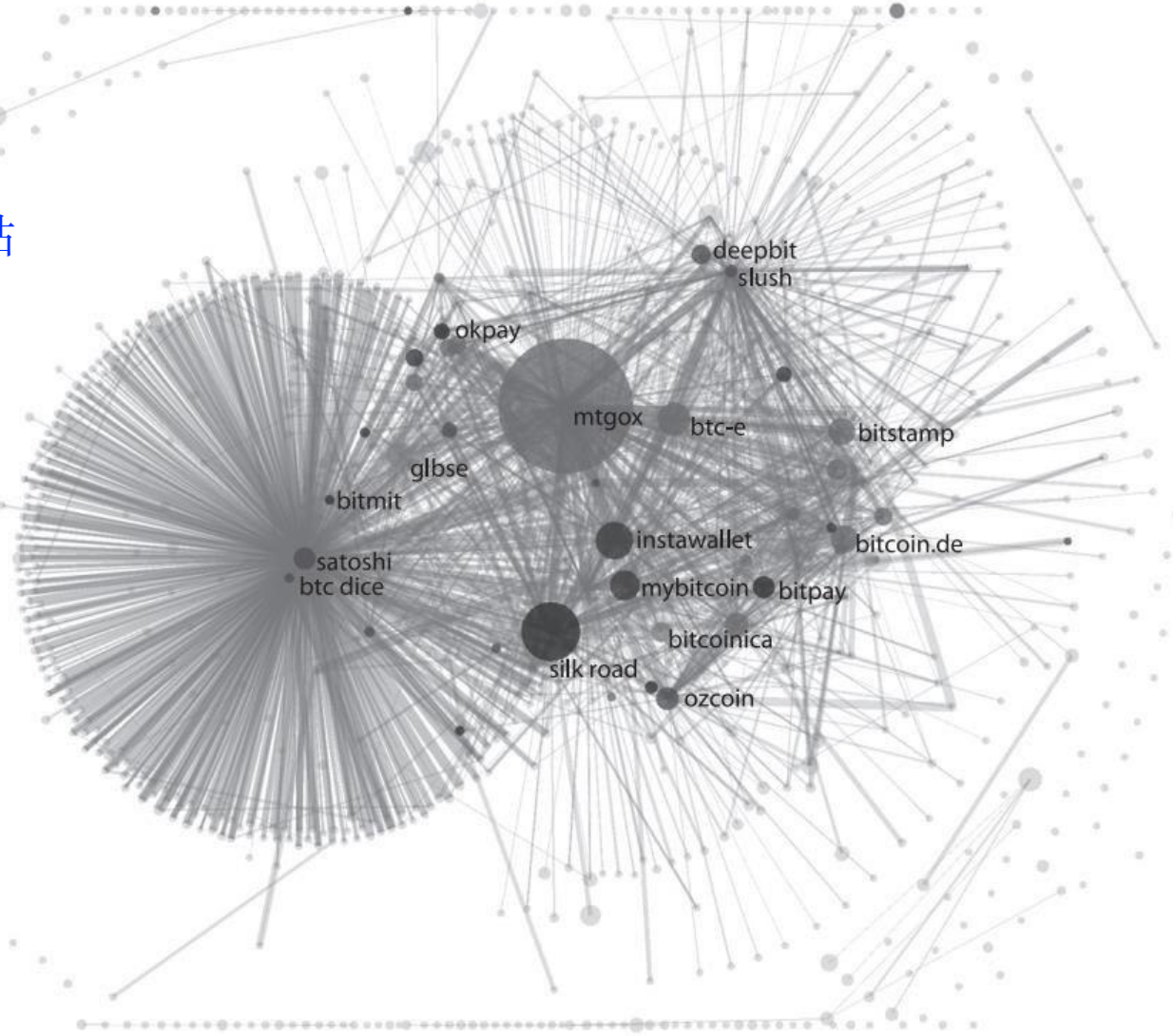
关联真实世界的身份到地址簇-交易图谱分析

标签簇

利用交易进行标记:
交易所、钱包服务、博彩网站

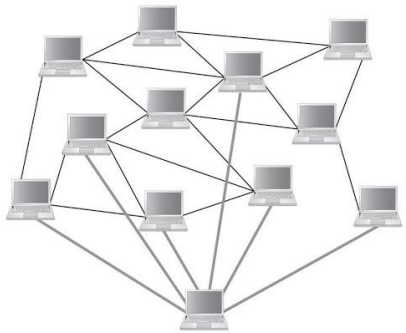
通过和不同的比特币服务提供商进行交易,
米克尔·约翰等人得以辨识并且标记这些簇在真实世界中的身份

去匿名化方法:
TX graph analysis
交易图谱分析



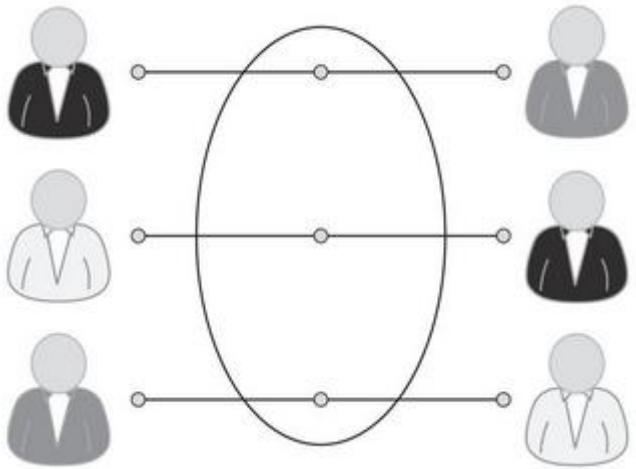
网络层的 去匿名化

- 2011年Dan Kaminsky 在 Black Hat 大会上提出网络层去匿名化的概念
- 基于一个观察：第一个通知交易的节点很有可能就是交易源头
- 当有多个节点配合并且对同一个交易源头进行识别的时候，这种方法的实际效果会更加明显。
- Tor （洋葱路由）协议可以应对：但是要求low latency
  - Tor是一个世界范围的计算机网络，在请求开始点以加密方式转发请求，直到到达网络中最后一台称为出口节点（exit node）的计算机。出口节点会对请求进行解密并传输到目标服务器。出口节点是专门用于流量离开Tor网络的最后一跳，也是用于返回流量的第一跳。使用Tor时，与您通信的系统将所有传入的流量视为来自出口节点。他们不知道您的位置，也不知道您的真实IP地址。此外，Tor网络中其他系统也不能确定您的位置，因为实质上只是转发流量，并不知道流量的实际来。请求的响应将返回系统，但对于Tor网络而言，流量源仅充当了路径中的一跳。本质上，您是匿名的。
  - 洋葱路由器 能够匿名化你的 Web 浏览与发布、即时通讯、IRC、SSH 和其他使用 TCP 协议的应用。
- 混币网络 (Mix Net)

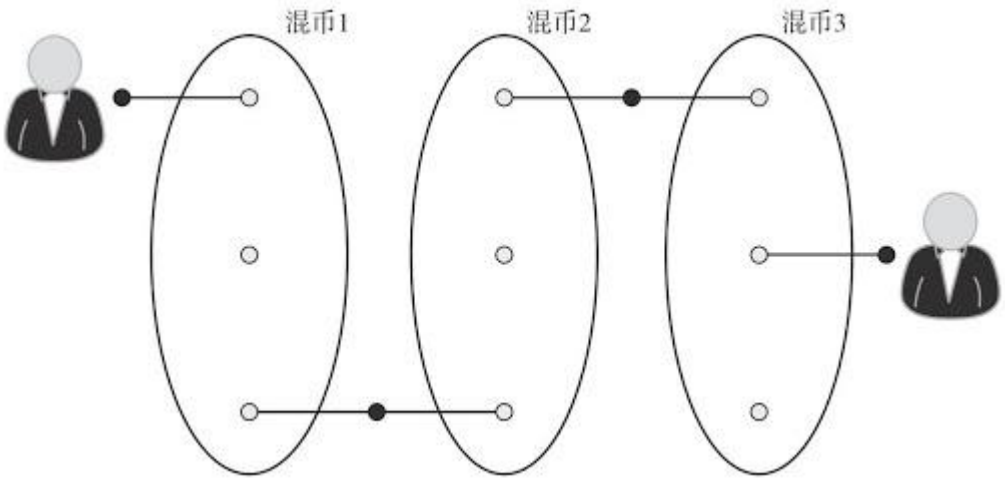


混币 -- 让 交易图谱分析 变得无效

- 混币模式
  - 中介
  - 混币在线钱包
  - 专项混币服务

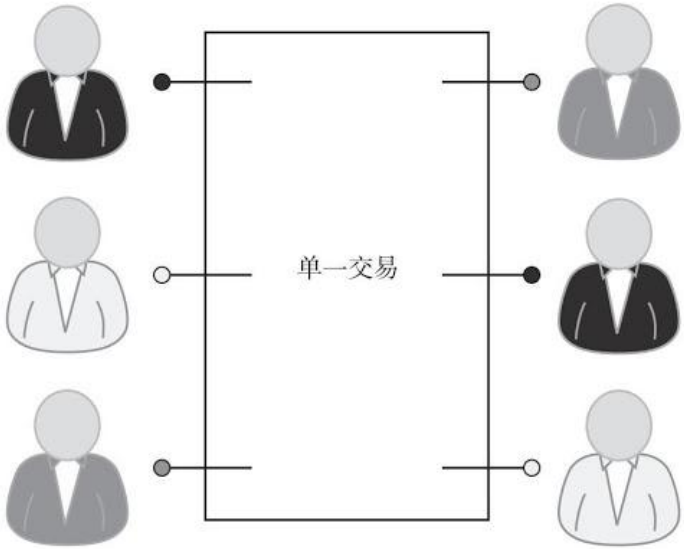


- 多重混币
  - 不能将用户最初发送的BTC关联到最终接受的BTC
  - 风险:
    - Service Provider 跑路



分布式混币

- 分布式混币，特点包括
  - 用户之间的 P2P 模式市实现混币交易
  - 没有自举过程：用户不需要等待一个有公信力的 Service Provider
  - 盗币行为在分布式混币模式下 impossible
  - 提供更好的匿名性

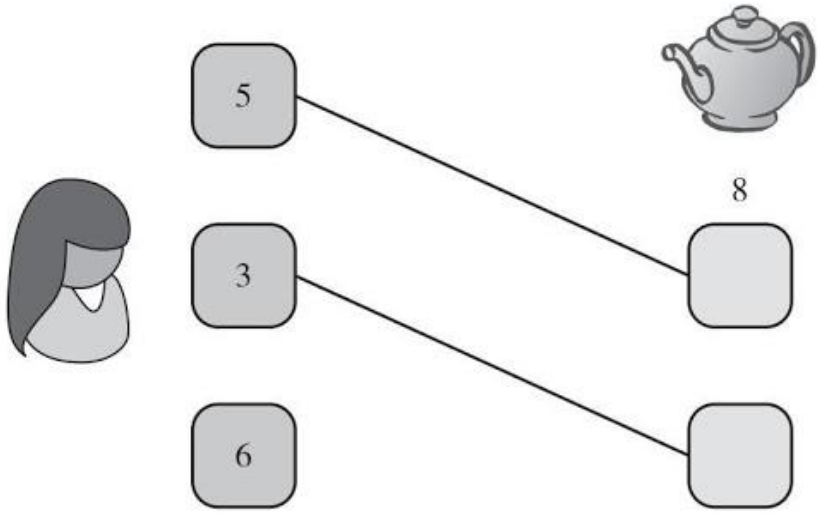


- 主要方案：合币 (Coinjoin)
  - 不同用户共同创建一个单一TX
  - 每人独立签名
  - 输入与输出Addr的顺序都是随机的
  - 首先需要发现彼此; 交换 input/output; 构造TX; 轮流签名; 广播TX

分布式混币 应对 高风险交易流

- 高风险交易流:
  - ex: Alice 每月将固定的薪水的5%存入退休基金账号
- 合并规避 (merge avoidance)
  - 帮助存在高风险交易流的用户重获无关联性
  - 允许使用多个 Output 地址进行接收BTC

- 爱丽丝想要用8个比特币去购买一只茶壶,
- 店铺提供了两个地址给她,
- 她可以支付5个比特币到其中一个地址而支付3个比特币到另外一个地址,
- 与她的可用输入资金匹配了,这样就可以避免暴露两个地址都是属于爱丽丝的事实。

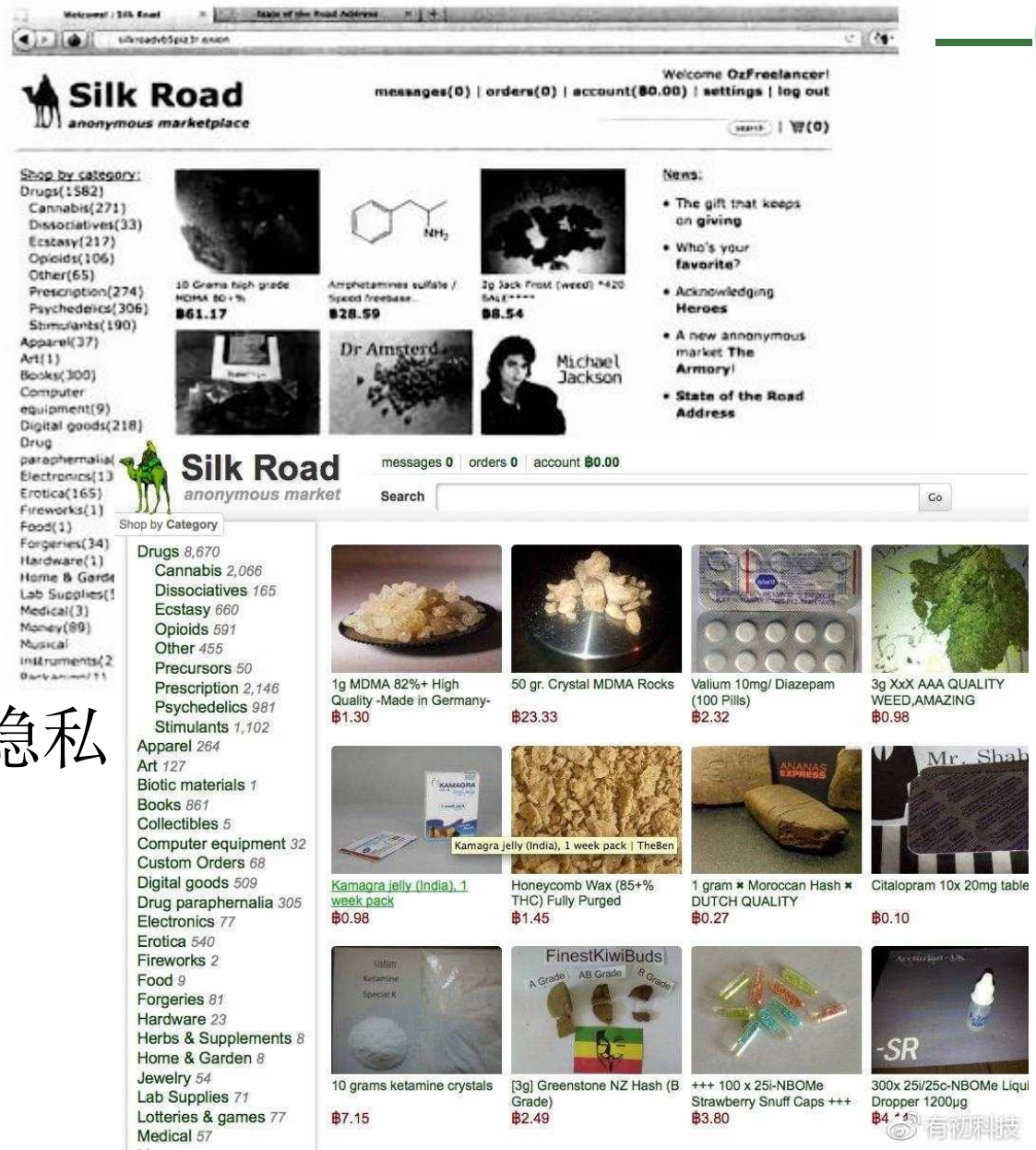


Outline & Keywords of this Class

- Part 1: 比特币 网络
- Part 2: 匿名
- Part 3: 监管

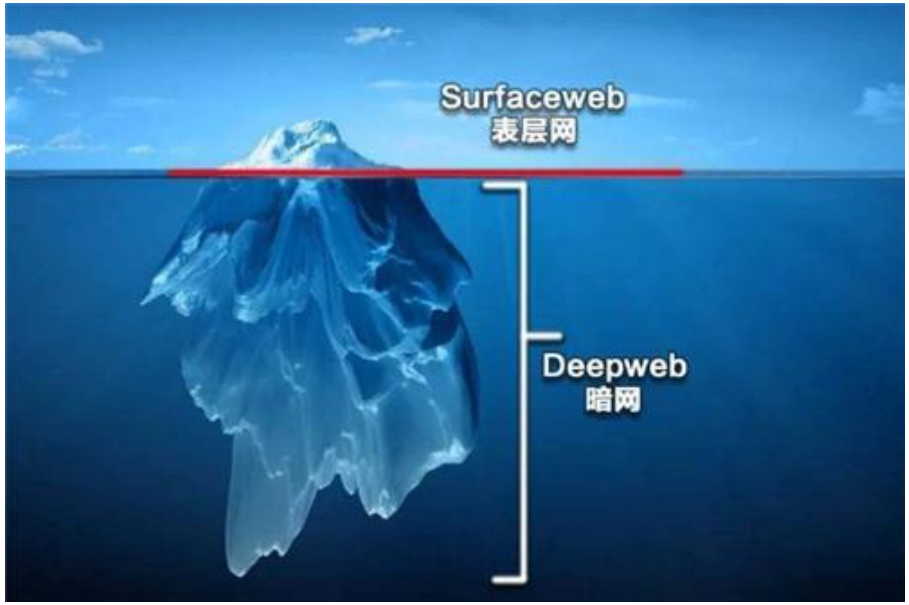
政府对比特币的关注

- 资本管制
- 犯罪
  - 丝绸之路
  - 暗网
  - 毒品、枪
  - 人和人的隐私



政府对比特币的关注

- 暗网：Dark Web
  - 黑暗网络，通常它们被人进行了**加密传输**、P2P对等网络、多点中继混淆等，普通用户是无法自己进入的，甚至根本不知道它的存在。
- 互联网：表层网（4%）+深层网（暗网96%）



政府对比特币的关注

- 反洗钱
  - KYC: Know Your Customer 原则
  - 识别并验证客户
  - 评估客户风险
  - 监控异常举动

监管与若干手段

- 监管是可取的吗?
  - 自由市场并不总是给出最有效的结果，所以监管是有益于社会的
- 反串谋和反垄断
- 发放加密货币牌照
- 发布强制政策，限制加密货币的流通与交易