THE BITCOIN TRANSACTION NETWORK

BITCOIN IN A NUTSHELL

HISTORY

- Invented by Satoshi Nakamoto (person or group of person), started in 2009
- The protocol is still evolving, the official *bitcoin core* is a GitHub repository, controlled by 5-10 individuals, on which anyone can propose contributions
  - Objectives: More efficient, faster, more secure, more anonymous,…

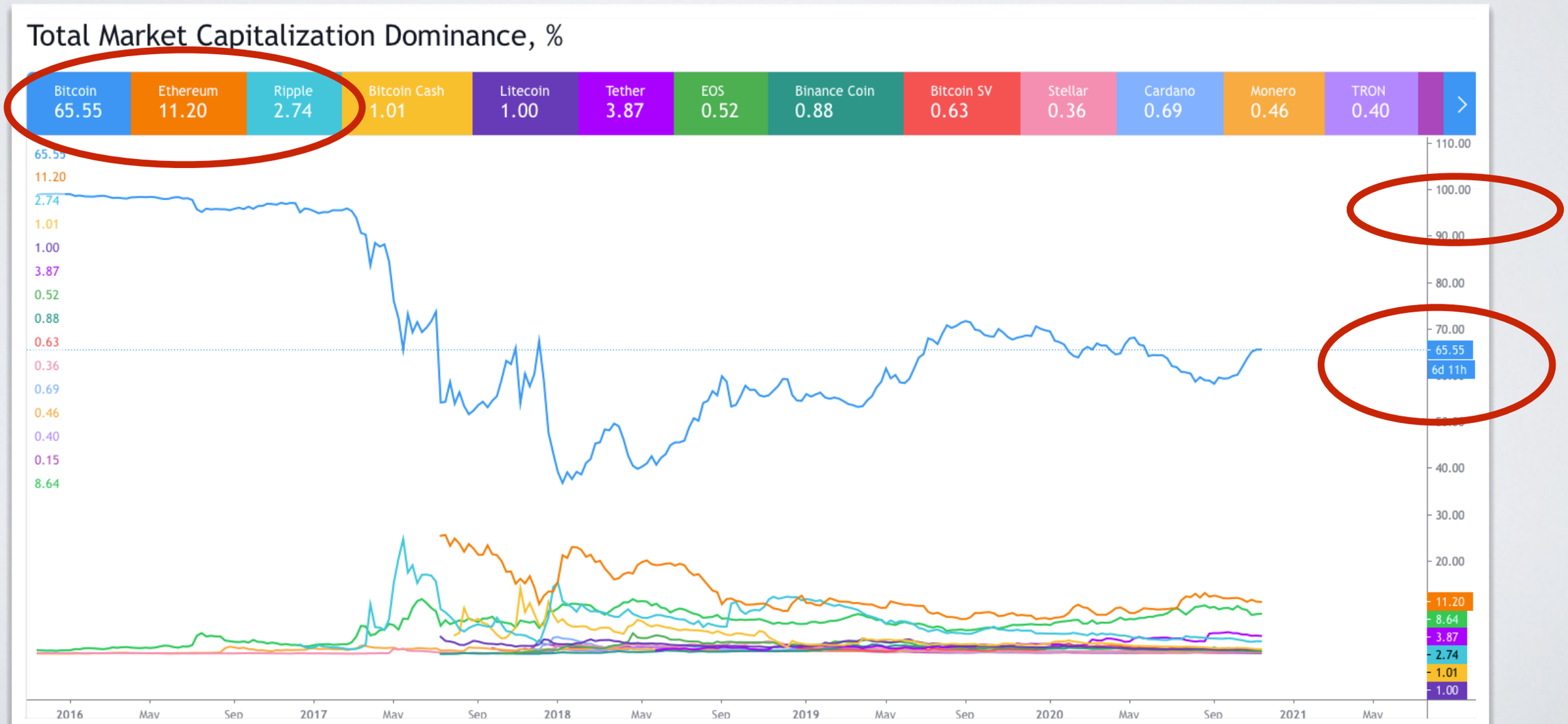
https://github.com/bitcoin/bitcoin

WHAT IS IT?

- A cryptocurrency
- A decentralized digital currency
  - No central authority (no central bank or state issue or guarantee the currency)
  - Cryptographic methods guarantee that no-one is cheating:
    - Issuing their own coins
    - Stealing coins
    - Etc.

IMPORTANCE

Bitcoin was the first cryptocurrency and is still has by far the highest Market capitalization. (Value of all existing coins)



SOME NUMBERS

- Bitcoins in existence(market cap) > $300 Billions
  - ≈ Samsung, intel, mastercard
- Transactions per day > 400,000
  - VISA: 150 million.
- Median transaction fee = $0,7
- Total value sent per day(blockchain) > $1 Billion
- Trading volume per day > Between $0.5 - 5 Billion
- Median transaction value = $400

DIGITAL LEDGER

- Bitcoin is based on a **blockchain**
  - Every transaction is stored in a *sequential database* (chain), a digital ledger.
    - Each new transaction is added at the end of the chain (in blocks)
    - Anyone can read everything in this chain
    - No-one can modify the older blocks in the chain
    - Adding a new element to the chain requires to solve a cryptographic problem

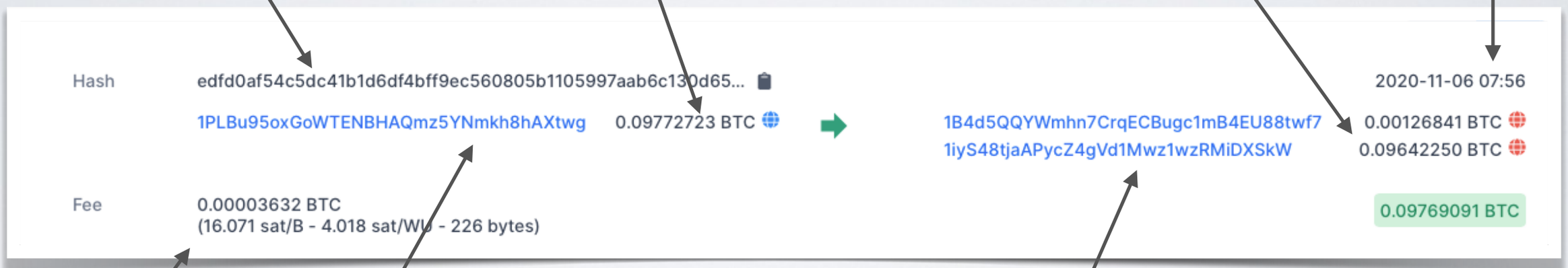
TYPICAL RECORD

Time of the transaction

Transaction identifier

Sum in input

Amounts sent



Transaction fee

Who is **sending**
(Bitcoin address)

Who is **receiving**
(Bitcoin addresses)

BITCOIN ANONYMITY

- Anyone can see all transactions=>We can study in details aggregated statistics
  - Evolutions of numbers, amount of transactions, fees, etc.

- So can we track user's activity?
  - Pseudonimity=>no way to link **bitcoin address** to **identity**
  - Users can create multiple addresses easily
  - Multiple addresses of a same person can sometimes be associated
  - In practice:
    - Large actors (companies, …) are not anonymous
    - Individual users can hide what they are doing

BITCOIN MARKETS

- Bitcoin value in $ is fixed based on exchange markets
  - Trading, much as any other currency
  - Trade operations are usually not written in the blockchain, the bank virtually exchange between counts of its customers
- Transaction fees are decided based on another market
  - **Miners** use computation power to solve cryptographic problems to include transactions in the blockchain
    - They are paid by 1)newly created coins 2)transaction fee
  - Anyone is free to propose any transaction fee
    - Miners choose in priority transactions with higher fees

BITCOIN MARKETS

- What are bitcoin transactions?
  - Mining
  - Exchange between users?
  - Users buying services/products?
  - Trading?
    - No, not directly. Trading is done on exchange platforms and mostly handled internally
  - Gambling?
  - Exchange between "banks", i.e., wallet managers?
  - Money laundering?
- Detail is not known(yet)

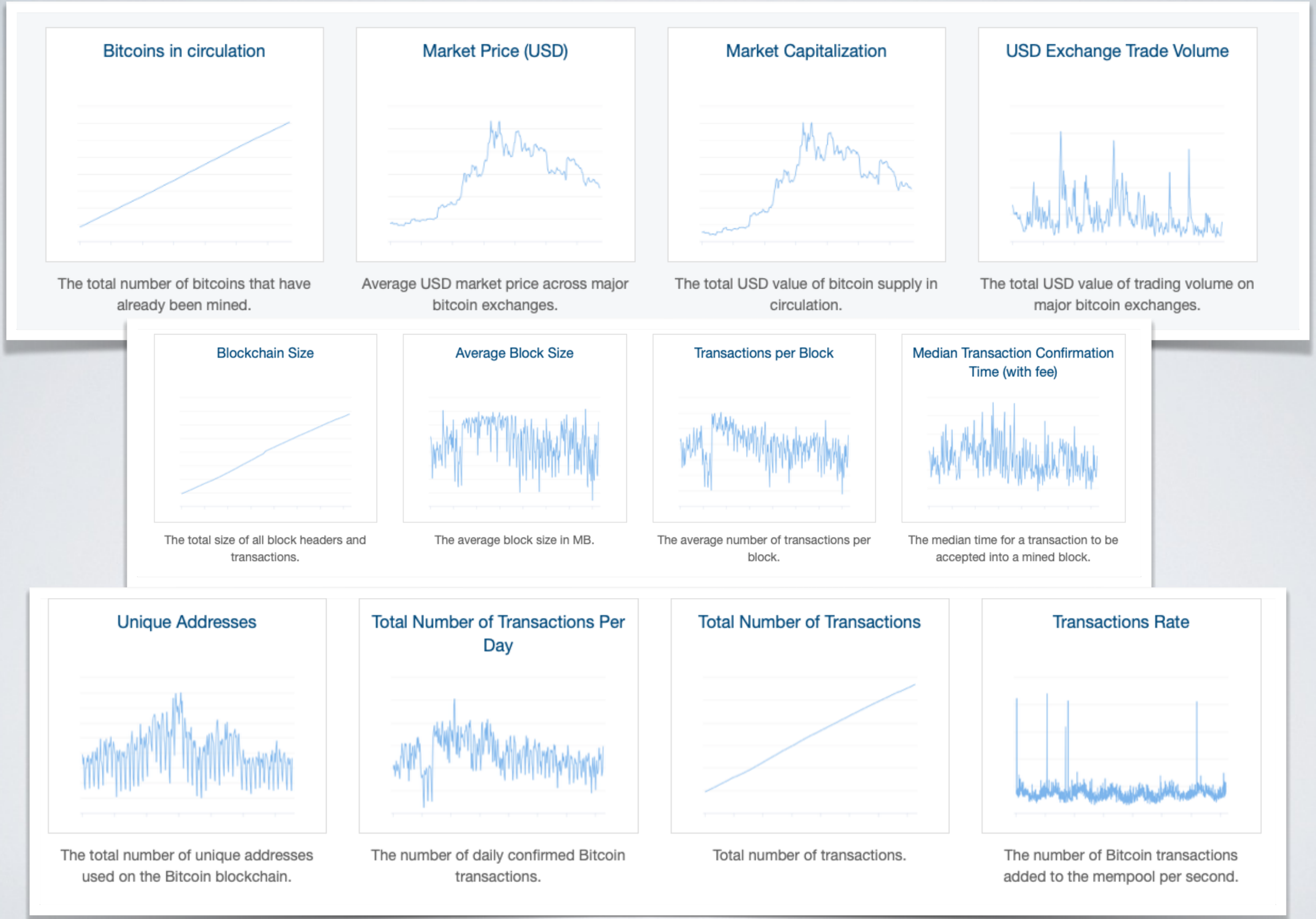
BITCOIN TRANSACTION NETWORK ANALYSIS

BITCOIN

- In this class, we are **not** interested in:
  - Cryptographic aspects
  - How the blockchain works
  - Governance of cryptocurrencies
  - Smart contracts
  - ICO
  - Macro-level analysis (transaction fee evolution, market price, etc.)

- What we are interested in:
  - Observing and understanding what is happening at the micro-level in one cryptocurrency (for this class, the largest one, Bitcoin) => **Look under the hood !**
  - How what is happening at the micro-level can be connected to what we observe at the macro-level (crisis, price fluctuation, macro-indicators…)

BITCOIN - MACRO LEVEL



https://www.blockchain.com/en/charts

BITCOIN - MACRO LEVEL

- This type of aggregated data is mostly identical to data you are used to in economy
- Can be studied with time series analysis (ARIMA, …)
- What is unique about Bitcoin:
  - We have all data about all transactions done using a given currency
  - We can use this information in relation with macro-level statistics
  - We can use it for new type of analysis

BITCOIN - DATA

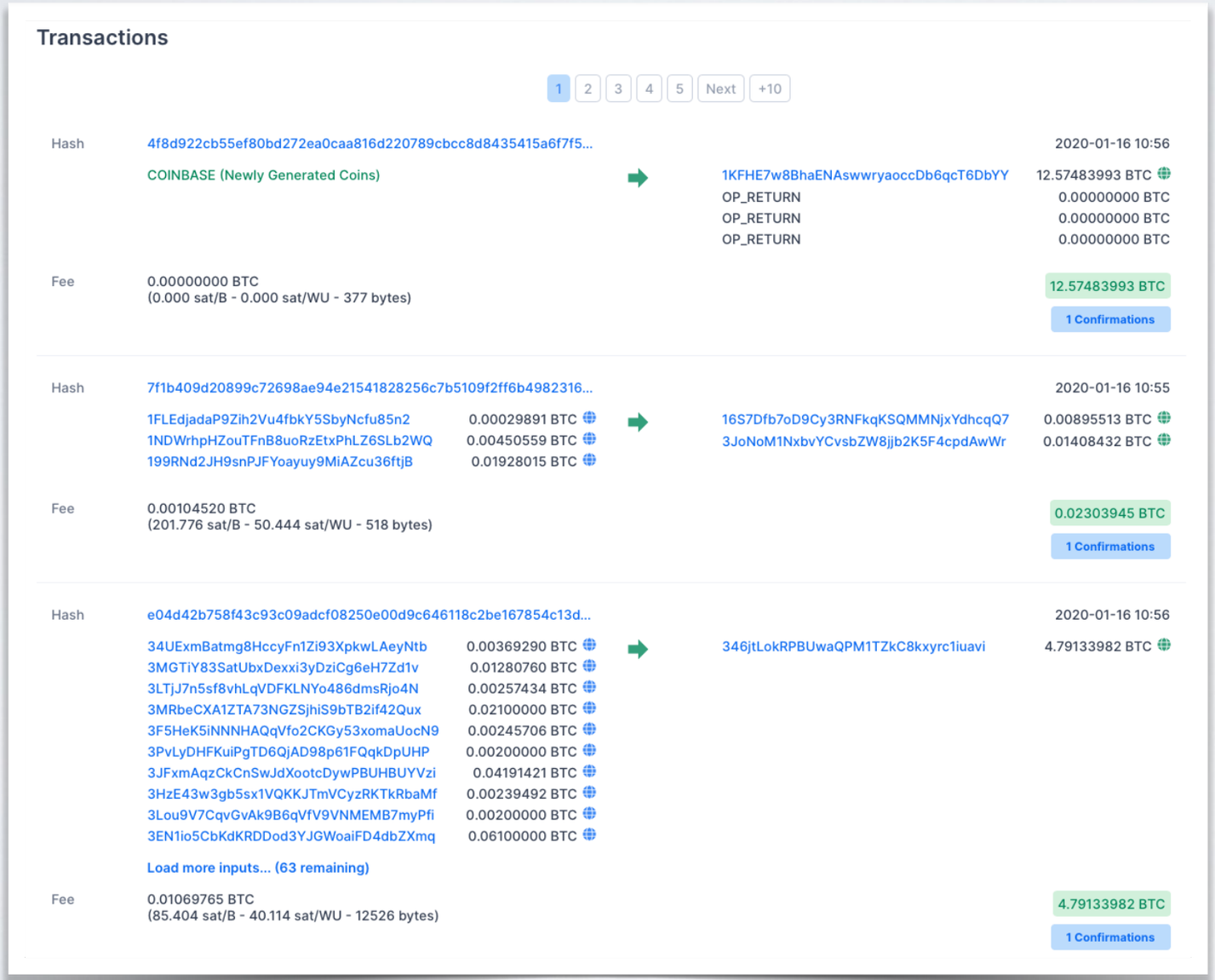
- The data we use: Content of the bitcoin blockchain
  - Seen as a simple list of transactions

Transaction	From	To	Value
t0	@1	@2	5
t1	@1	@3	2
…	…	…	…

- Bitcoin transactions are a little bit more complicated than that

BITCOIN - DATA

- You can explore it using tools such as a blockchain explorer
  - E.g.: https://www.blockchain.com/explorer



Transactions

1 2 3 4 5 Next +10

Hash	4f8d922cb55ef80bd272ea0caa816d220789cbcc8d8435415a6f7f5...			2020-01-16 10:56
	COINBASE (Newly Generated Coins)	→	1KFHE7w8BhaENAswwryaoccDb6qcT6DbYY	12.57483993 BTC
			OP_RETURN	0.00000000 BTC
			OP_RETURN	0.00000000 BTC
			OP_RETURN	0.00000000 BTC
Fee	0.00000000 BTC (0.000 sat/B - 0.000 sat/WU - 377 bytes)			12.57483993 BTC
				1 Confirmations
Hash	7f1b409d20899c72698ae94e21541828256c7b5109f2ff6b4982316...			2020-01-16 10:55
	1FLEdjadaP9Zih2Vu4fbkY5SbyNcfu85n2 0.00029891 BTC	→	16S7Dfb7oD9Cy3RNFkqKSQMMNjxYdhcqQ7	0.00895513 BTC
	1NDWrhpHZouTFnB8uoRzEtxPhLZ6SLb2WQ 0.00450559 BTC		3JoNoM1NxbvYCvsbZW8jjb2K5F4cpdAwWr	0.01408432 BTC
	199RNd2JH9snPJFYoayuy9MiAZcu36ftjB 0.01928015 BTC			
Fee	0.00104520 BTC (201.776 sat/B - 50.444 sat/WU - 518 bytes)			0.02303945 BTC
				1 Confirmations
Hash	e04d42b758f43c93c09adcf08250e00d9c646118c2be167854c13d...			2020-01-16 10:56
	34UExmBatmg8HccyFn1Zi93XpkwLAeyNtb 0.00369290 BTC	→	346jtLokRPBUwaQPM1TZkC8kxyrc1iuavi	4.79133982 BTC
	3MGTiY83SatUbxDexxi3yDziCg6eH7Zd1v 0.01280760 BTC			
	3LTjJ7n5sf8vhLqVDFKLNYo486dmsRjo4N 0.00257434 BTC			
	3MRbeCXA1ZTA73NGZSjhiS9bTB2if42Qux 0.02100000 BTC			
	3F5HeK5iNNNHAQqVfo2CKGy53xomaUocN9 0.00245706 BTC			
	3PvLyDHFKuiPgTD6QjAD98p61FQqkDpUHP 0.00200000 BTC			
	3JFxmAqzCkCnSwJdXootcDywPBUHBUYVzi 0.04191421 BTC			
	3HzE43w3gb5sx1VQKKJTmVCyzRKTkRbaMf 0.00239492 BTC			
	3Lou9V7CqvGvAk9B6qVfV9VNMEMB7myPfi 0.00200000 BTC			
	3EN1io5CbKdKRDDod3YJGWoaiFD4dbZXmq 0.06100000 BTC			
	Load more inputs... (63 remaining)			
Fee	0.01069765 BTC (85.404 sat/B - 40.114 sat/WU - 12526 bytes)			4.79133982 BTC
				1 Confirmations

Hash	7f1b409d20899c72698ae94e21541828256c7b5109f2ff6b4982316...			2020-01-16 10:55
	1FLEdjadaP9Zih2Vu4fbkY5SbyNcfu85n2	0.00029891 BTC	16S7Dfb7oD9Cy3RNFkqKSQMMNjxYdhcqQ7	0.00895513 BTC
	1NDWrhpHZouTFnB8uoRzEtxPhLZ6SLb2WQ	0.00450559 BTC	3JoNoM1NxbvYCvsbZW8jjb2K5F4cpdAwWr	0.01408432 BTC
	199RNd2JH9snPJFYoayuy9MiAZcu36ftjB	0.01928015 BTC		
Fee	0.00104520 BTC (201.776 sat/B - 50.444 sat/WU - 518 bytes)			0.02303945 BTC 1 Confirmations

Hash	e04d42b758f43c93c09adcf08250e00d9c646118c2be167854c13d...			2020-01-16 10:56
	34UExmBatmg8HccyFn1Zi93XpkwLAeyNtb	0.00369290 BTC	346jtLokRPBUwaQPM1TZkC8kxyrc1iuavi	4.79133982 BTC
	3MGTiY83SatUbxDexxi3yDziCg6eH7Zd1v	0.01280760 BTC		
	3LTjJ7n5sf8vhLqVDFKLNYo486dmsRjo4N	0.00257434 BTC		
	3MRbeCXA1ZTA73NGZSjhiS9bTB2if42Qux	0.02100000 BTC		
	3F5HeK5iNNNHAQqVfo2CKGy53xomaUocN9	0.00245706 BTC		
	3PvLyDHFKuiPgTD6QjAD98p61FQqkDpUHP	0.00200000 BTC		
	3JFxmAqzCkCnSwJdXootcDywPBUHBUYVzi	0.04191421 BTC		
	3HzE43w3gb5sx1VQKKJTmVCyzRKTkRbaMf	0.00239492 BTC		
	3Lou9V7CqvGvAk9B6qVfV9VNMEMB7myPfi	0.00200000 BTC		
	3EN1io5CbKdKRDDod3YJGWoaiFD4dbZXmq	0.06100000 BTC		
	Load more inputs... (63 remaining)			
Fee	0.01069765 BTC (85.404 sat/B - 40.114 sat/WU - 12526 bytes)			4.79133982 BTC 1 Confirmations

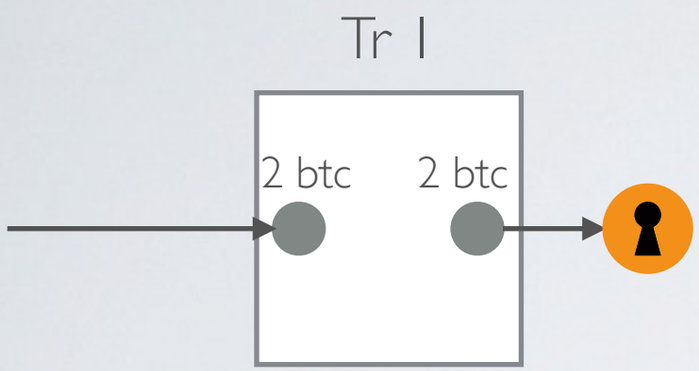
UNDERSTANDING BITCOIN TRANSACTIONS

- Transactions are between *m* "inputs" and *n* "outputs"
- Each *input* (resp. *output*) is a pair (value,bitcoin address)
- *inputs* are necessarily *outputs* of previous transactions
  - Unlocked by the private key of the payer

UNDERSTANDING BITCOIN TRANSACTIONS

- A user possess a **private key**
- A user can generate **public keys** (bitcoin adresses)
  - Instantaneously
  - At no cost
  - As often as wanted
- Public key ≈ lock that can be opened only by an associated private key

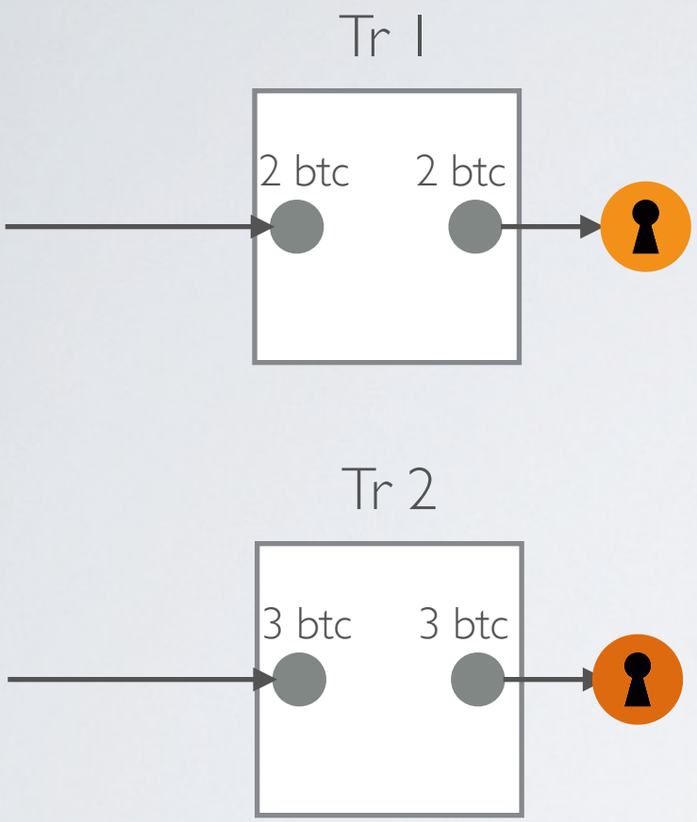
ILLUSTRATION



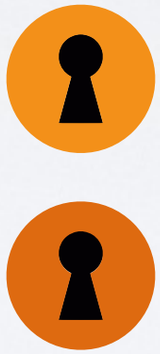
Public keys of user U1 :

1BusVkYQvbbGbSDZNo5DfhrFeQdgK1Y1VY

ILLUSTRATION



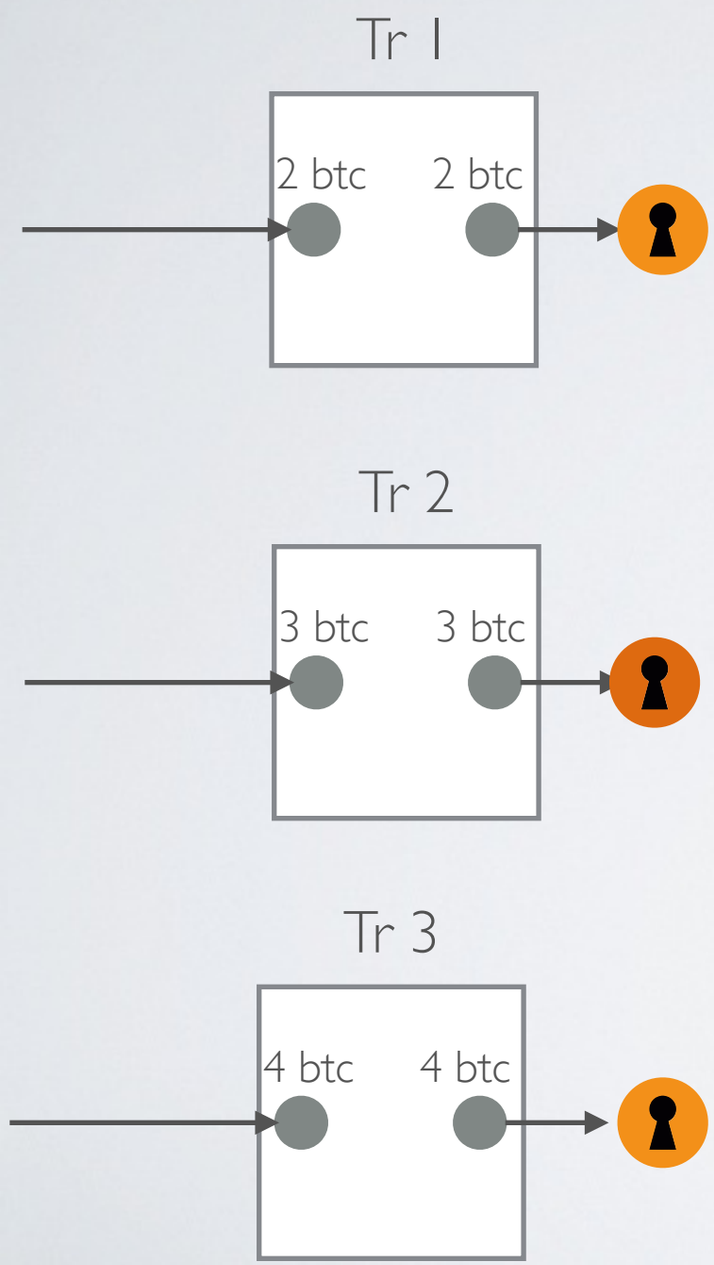
Public keys of user U1 :



1BusVkYQvbbGbSDZNo5DfhrFeQdgK1Y1VY

1QFdbGkhiCDFF45mBHgzWUdiqv55NJbd4u

ILLUSTRATION



Public keys of user U1 :

1BusVkYQvbbGbSDZNo5DfhrFeQdgK1Y1VY

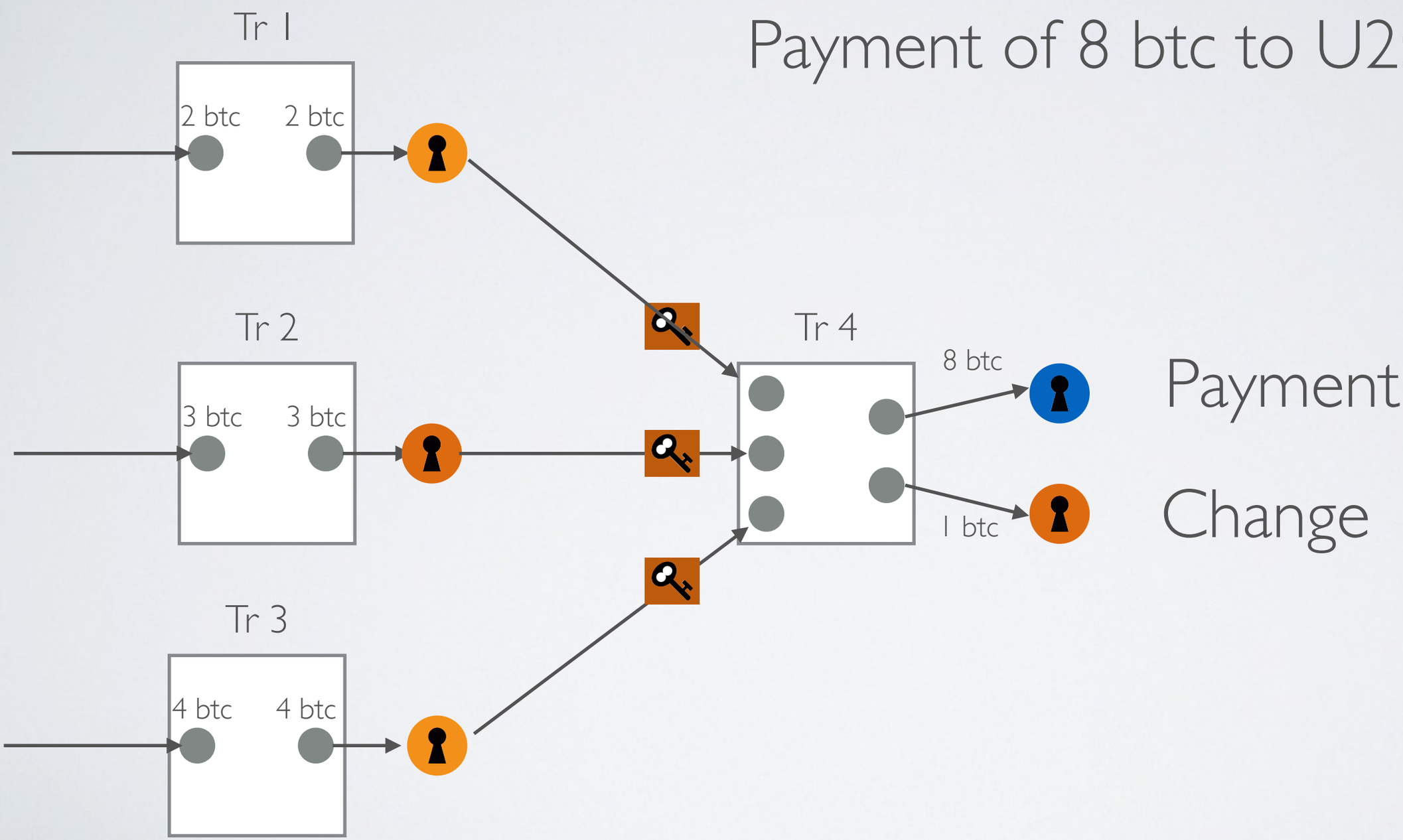
1QFdbGkhiCDFF45mBHgzWUdiqv55NJbd4u

“Wallet” of U1:

- 9 btc
- Divided in 3 “output”
- Locked by 2 different public keys

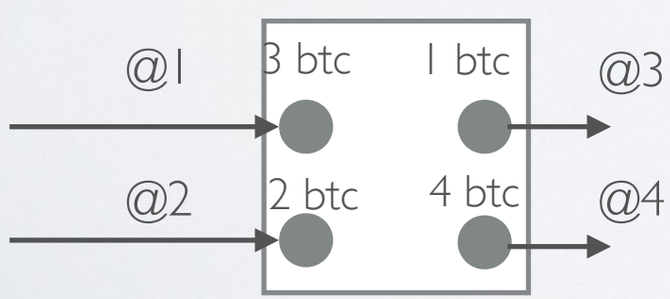
ILLUSTRATION

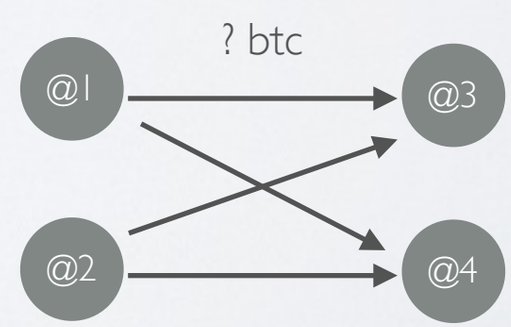
Payment of 8 btc to U2:



ADDRESS NETWORK

- First network, node=Address
  - Naive approach
  - One address ≠ one user!
- Node: bitcoin address (public key)
- Edge: input addresses to output addresses.
- Problem: most transactions have several inputs, several outputs
  - Values ?





ADDRESS NETWORK

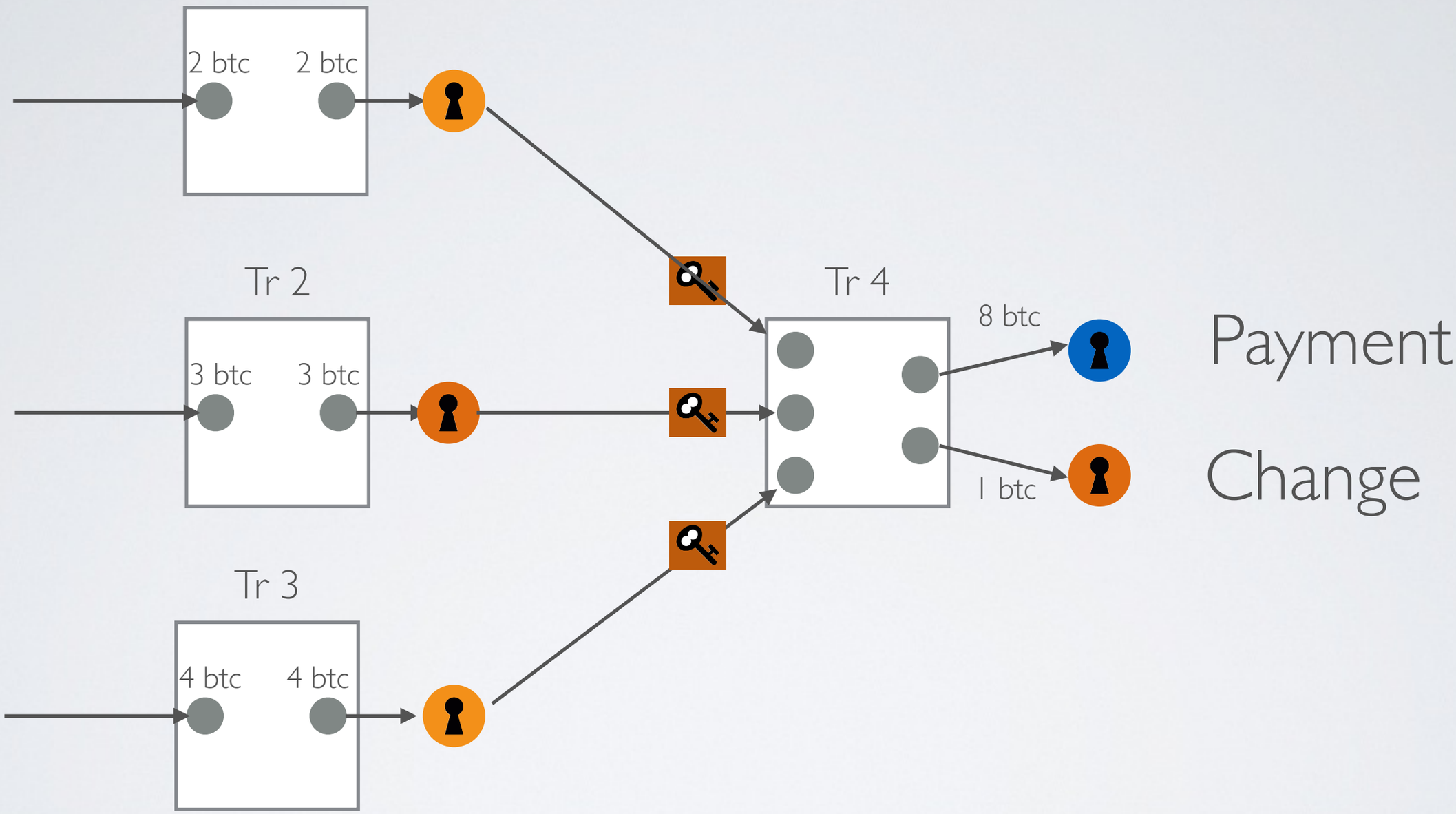
- # Transactions: 490 441
- # Transaction outputs: 1 210 004 (avg. 2,46)
- # Transaction inputs 1 211 790 (avg. 2.47)
- # Addresses: 933 645
- # @->@ Edges: 3 014 350

- Very big, hard to interpret

ACTOR NETWORK

- Transactions between "actors" of the bitcoin ecosystem
  - Individuals with their own private key (e.g., using BRD, Atomic Wallet, etc.)
  - Companies/organisations with their own private key
  - Exchanges (e.g., Binance, CoinBase, etc.)
  - Mining Pool
  - etc.
- An actor has **one** private key, but can have **many** public keys/addresses
- How to retrieve addresses belonging to the same actor?

ACTOR NETWORK

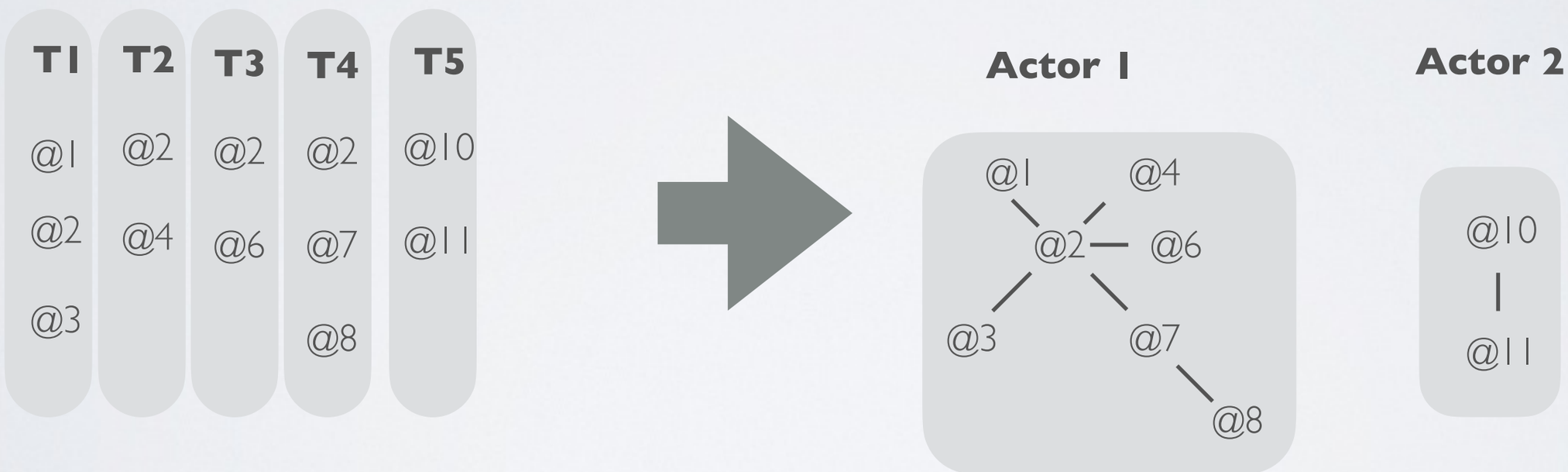


and are inputs of the same transaction
=> same actor

ACTOR NETWORK

- Actor identification: find all addresses of a same user
  - Currently a research question…
- Heuristics (input):
  - All addresses in input of a same transaction belongs to the same person

ACTOR NETWORK

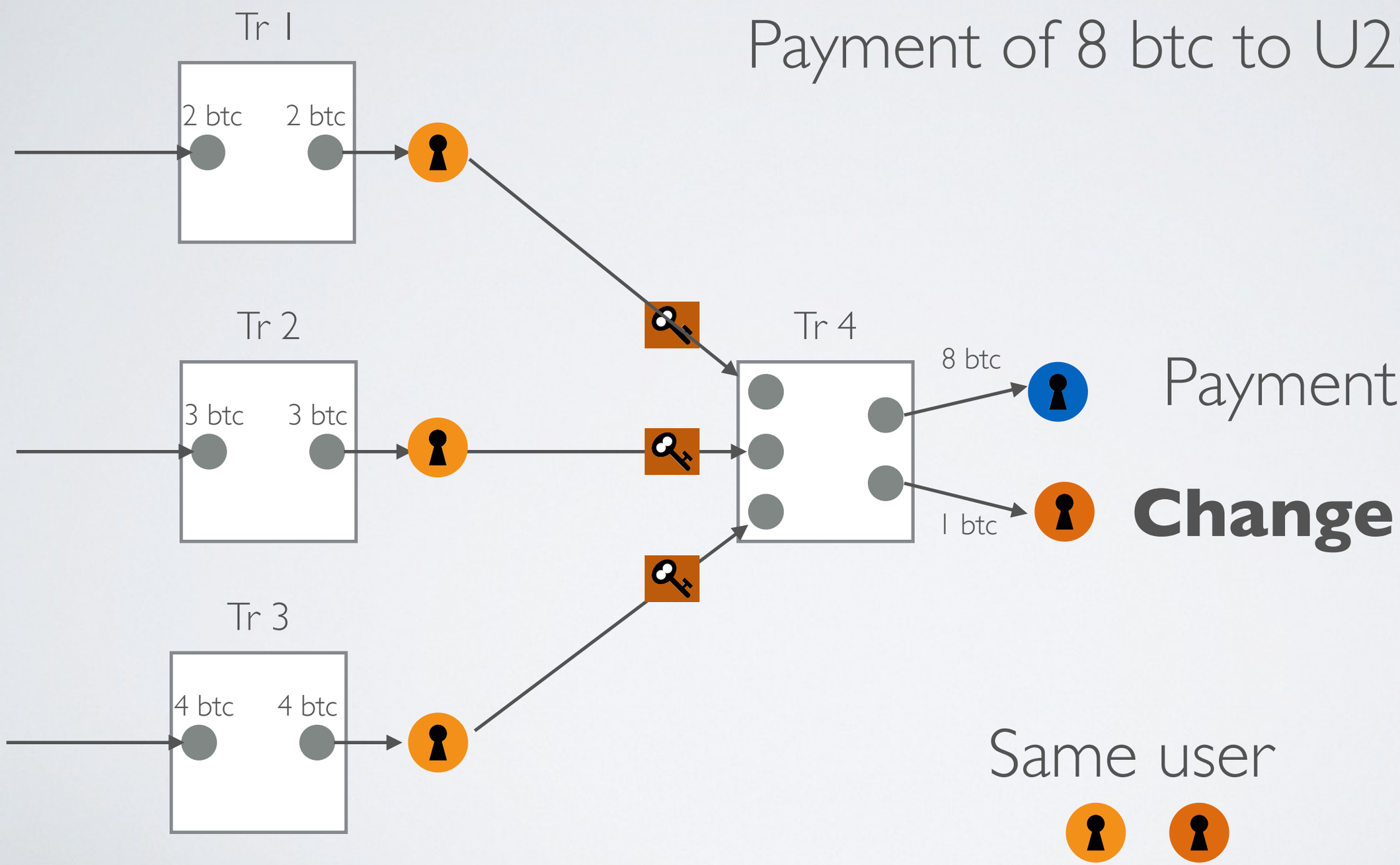


ACTOR NETWORK

- Actor identification: find all addresses of a same user
  - Currently a research question…
- Heuristics (input):
  - All addresses in input of a same transaction belongs to the same person
- Heuristics (output):
  - One of the addresses in output is probably a **change address**, thus an address of the same user as the one in input
  - But which one ?

ACTOR NETWORK

Payment of 8 btc to U2:



ACTOR NETWORK

- Heuristics (output):
  - One of the addresses in output is probably a **change address**, thus an address of the same user as the one in input
  - But which one ?
    - Lower value ?
    - Value with the same decimal as input?
    - Learn which one using machine learning and examples ?
    - …
    - => A research question, not in the scope of this class.

ACTOR NETWORK

- Group of addresses => Anonymous actor
  - Can we know who is this actor?
  - It is enough to identify *one* address
  - One transaction with a person/company => we know one of its addresses
  - On the internet, many company/individuals provide their addresses.
  - For some actors, we might infer their category
    - => Miners
    - => Large transactions profiles VS low transaction profiles
    - Has made transactions to identified money laundering services => suspicious
    - Machine learning => Automatically recognize profiles, identify similar actors, …
    - etc.

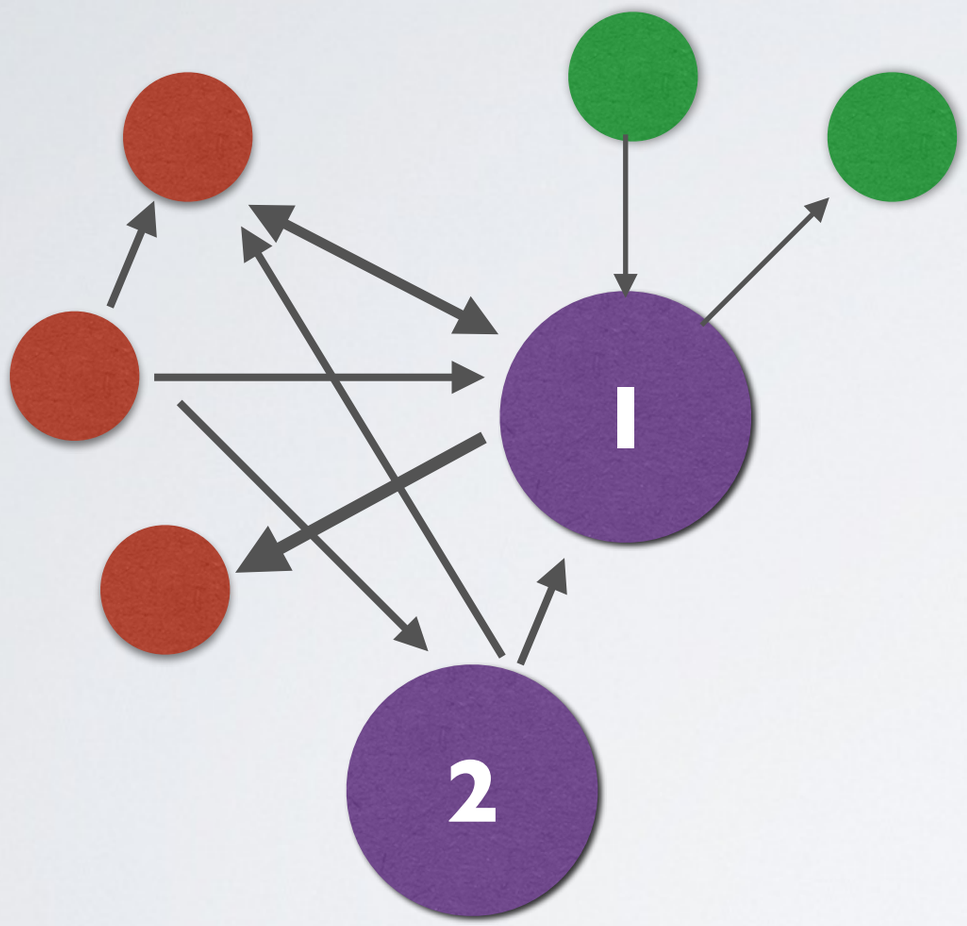
ACTOR NETWORK

List of actors addresses, for instance: https://www.walletexplorer.com

Top wallets

Exchanges:	Pools:	Services/others:	Gambling:	Old/historic:
Huobi.com (2)	BTCCPool	CoinPayments.net	SatoshiDice.com (original)	AgoraMarket
Bittrex.com	SlushPool.com (old) (old2)	Xapo.com	LuckyB.it (chatbot)	BetcoinDice.tm
Poloniex.com	GHash.io	Cubits.com	BitZillions.com	SilkRoadMarketplace
Luno.com	AntPool.com (old) (old2)	Cryptonator.com (old)	999Dice.com	DeepBit.net
BTC-e.com (output) (old)	BitMinter.com	BitPay.com (old) (old2) (old3)	CoinGaming.io	SilkRoad2Market
Kraken.com (old)	EclipseMC.com (old) (old2) (old3)	BitoEX.com	PrimeDice.com (old) (old2) (old3) (old4)	EvolutionMarket
LocalBitcoins.com (old)	KnCMiner.com	HaoBTC.com	CloudBet.com	Instawallet.org
Bitstamp.net (old)	Bitfury.org	Cryptopay.me (old)	SatoshiMines.com	UpDown.BT
MercadoBitcoin.com.br	BW.com	AlphaBayMarket (old)	NitrogenSports.eu	AbraxasMarket
BitZlato.com	Eligius.st	NucleusMarket	SecondsTrade.com	MintPal.com
Cryptsy.com (old)	Kano.is (old)	BitcoinFog	PocketDice.io	SealsWithClubs.eu
Bitcoin.de (old)	Telco214	CoinJar.com	FortuneJack.com	PandoraOpenMarket
Cex.io		BitcoinWallet.com	Rollin.io	MiddleEarthMarketplace
Binance.com (old)		HolyTransaction.com	BitZino.com	BtcDice.com
BtcTrade.com		HelixMixer (old) (old2) (old3) (old4) (old5) (old6) (old7) (old8) (old9) (old10) (old11) (old12) (old13) (old14) (old15) (old16) (old17) (old18) (old19) (old20) (old21) (old22) (old23) (old24) (old25) (old26) (old27) (old28) (old29) (old30) (old31) (old32) (old33) (old34)	BitcoinVideoCasino.com (old) (old2)	McxNOW.com
YoBit.net		BTCJam.com	Betcoin.ag (old)	SheepMarketplace
OKCoin.com (2)		VIP72.com	YABTCL.com	DiceOnCrack.com
BTCC.com (old) (old2)		MoonBit.co.in	SatoshiBet.com	BlackBankMarket
BX.in.th		CoinKite.com	SafeDice.com	BTCGuild.com
HitBtc.com (old)		FaucetBOX.com	Coinroll.com	Coin-Swap.net
MaiCoin.com		OkLink.com	Crypto-Games.net	BlueSkyMarketplace
Bter.com (old) (old2) (old3) (cold)		Purse.io	Betcoin.tm	Justcoin.com
CoinSpot.com.au		ePay.info	SwCPoker.eu	PinballCoin.com
Hashnest.com		Loanbase.com	SatoshiRoulette.com	Inputs.io
AnxPro.com		GermanPlazaMarket	BTCOracle.com	BitAces.me (old)
BitBay.net		Paymium.com	Peerbet.org	AllCoin.com
Bleutrade.com		Bitbond.com	AnoniBet.com	Bitcoin-24.com (old) (old-hotwallet)
Bitfinex.com (old) (old2)		CrimeNetwork.co (old)	Satoshi-Karoshi.com (old)	Betcoins.net
Matbea.com			777Coin.com	CrimeNetwork.biz
Bit-x.com			BitStarz.com	Bitcoin-Roulette.com
VirWoX.com			SatoshiCircle.com	Bitmit.net
Paxful.com			Coinichiwa.com	Cryptorush.in
BitBargain.co.uk				

OBTAINED NETWORK

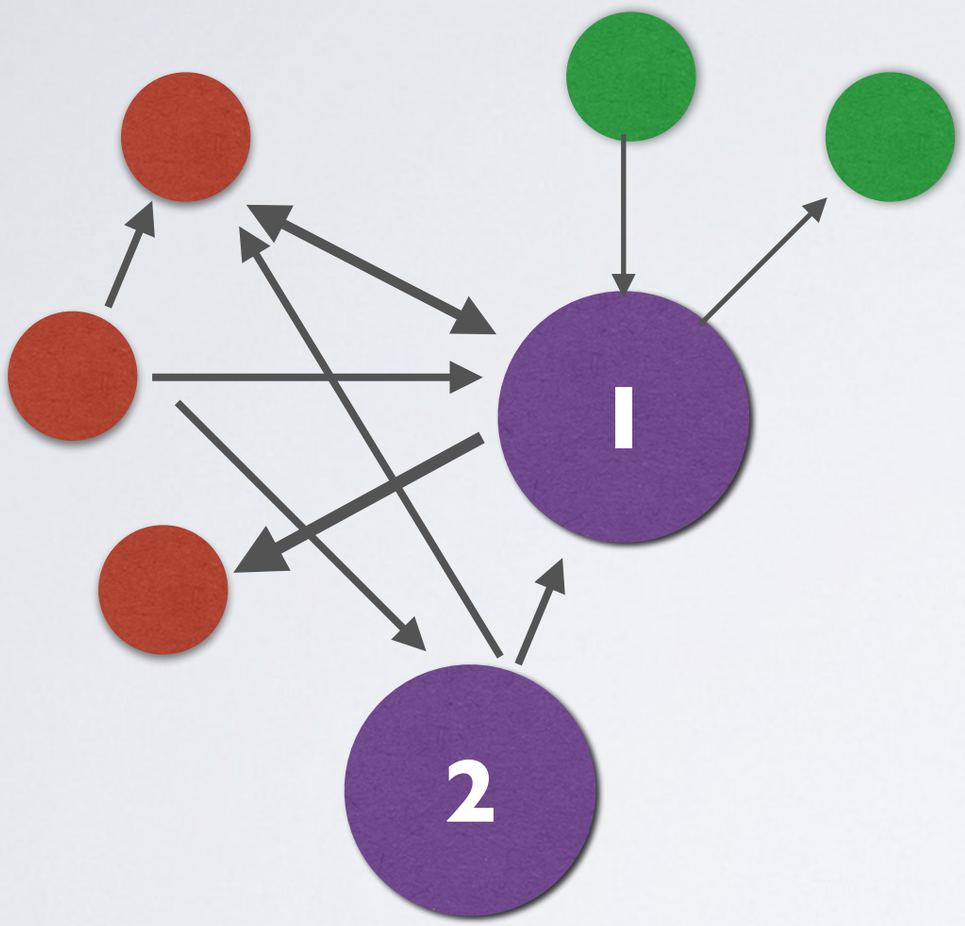


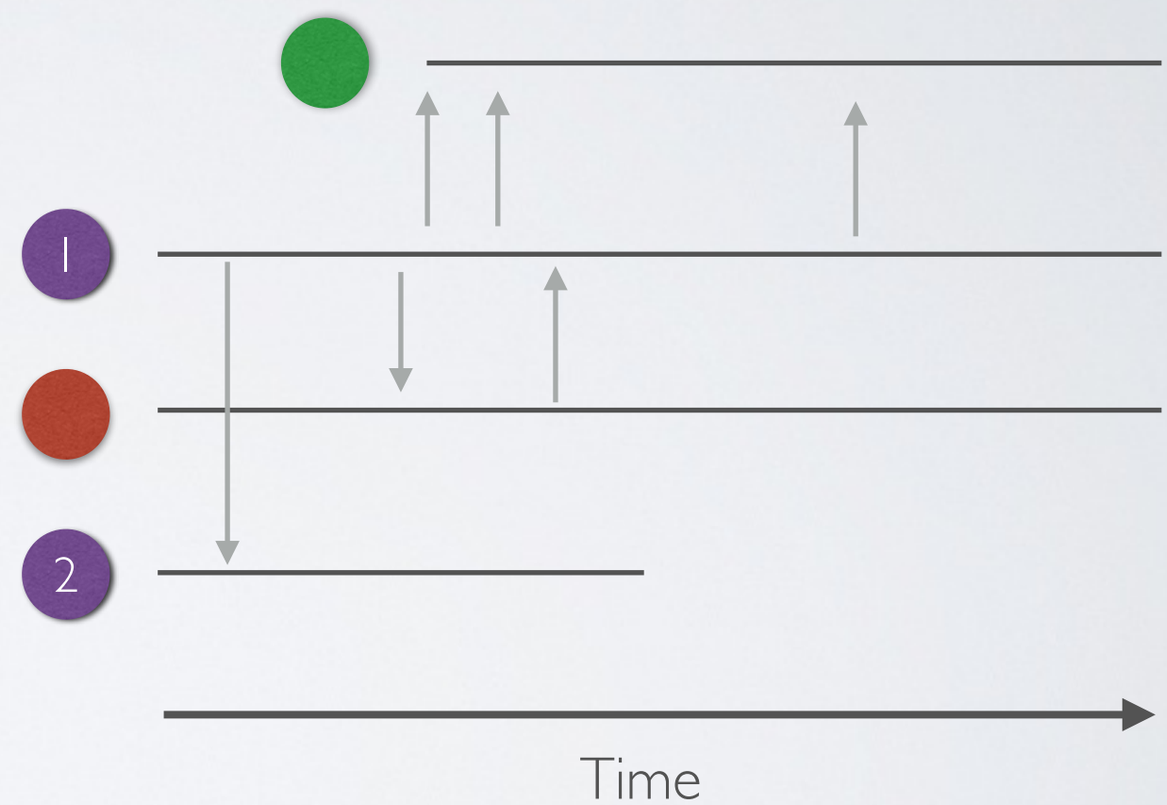
Identified nodes

Category 1

Category 2

OBTAINED NETWORK





Identified nodes

Category 1

Category 2

ACTOR NETWORK

- Example: 2 days (August 2&3 2016)
- Address network
  - # Transactions: 490 441
  - # Transaction outputs: 1 210 004 (avg. 2,46)
  - # Transaction inputs 1 211 790 (avg. 2.47)
  - # Addresses: 933 645
  - # @->@ Edges: 3 014 350
- Actor network
  - # Clusters: 456 012
  - Largest clusters sizes: 20 023, 19 381, 17 244
  - # Actor -> Actor Edges : 956 347