

A Deep Dive into Bitcoin Mining Pools

An Empirical Analysis of Mining Shares

Matteo Romiti¹ Aljosha Judmayer²
Alexei Zamyatin^{2, 3} Bernhard Haslhofer¹

May 31, 2019

¹Austrian Institute of Technology

²SBA Research

³Imperial College London

Why do we care about miners?

- Miners decide which transactions to include in a block
- Miners decide which blocks to include in the chain
- Miners are rewarded with new coins and transaction fees
- Miners secure the network (Proof-of-Work algorithm)
- Miners can attack the network (e.g., double-spend)

What about Mining Pools?

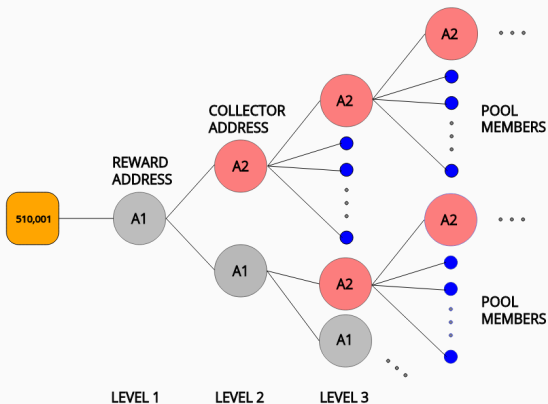
- Solo mining is not profitable anymore
- Miners join pools for steadier revenues
- Pools compete to create blocks and claim the rewards
- Pool managers coordinate work and rewards among members

Research Questions

1. How did the mining distribution evolve over time? How can we attribute blocks to mining pools?
2. How do mining pools distribute rewards? How can we detect payments to pool members?
3. How decentralized is a mining pool? Who are the pool members? What's their behavior?

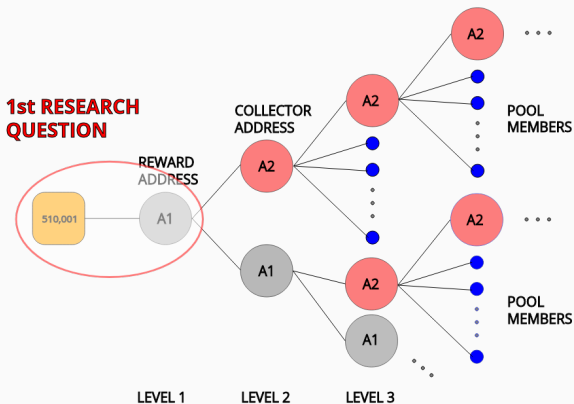
Background

Background — Coinbase Flow



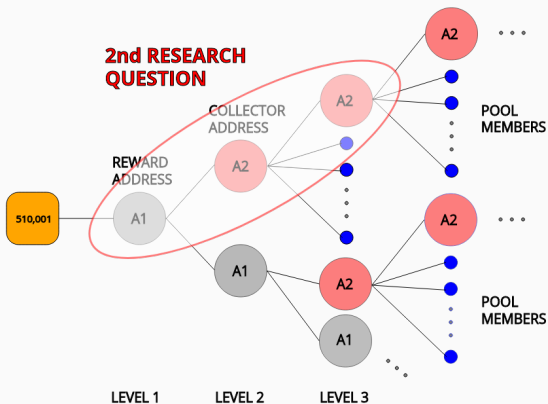
Example of a coinbase flow

Background — Coinbase Flow



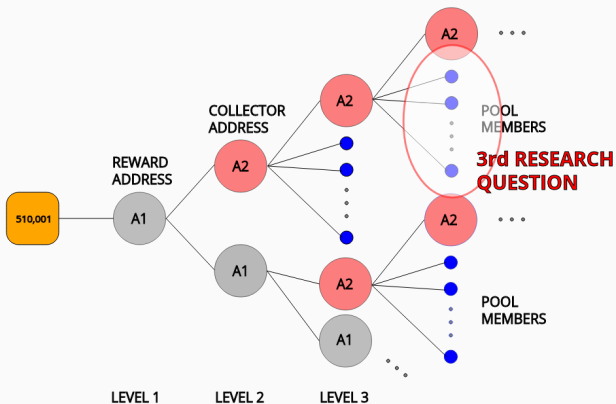
Example of a coinbase flow

Background — Coinbase Flow



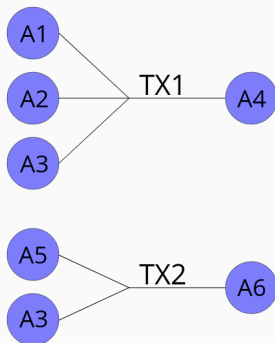
Example of a coinbase flow

Background — Coinbase Flow



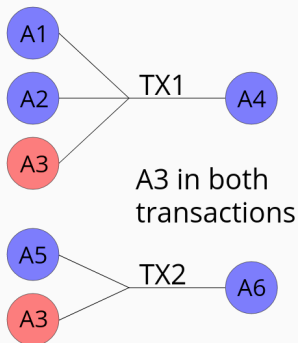
Example of a coinbase flow

Background — Multiple-Input Clustering Heuristic



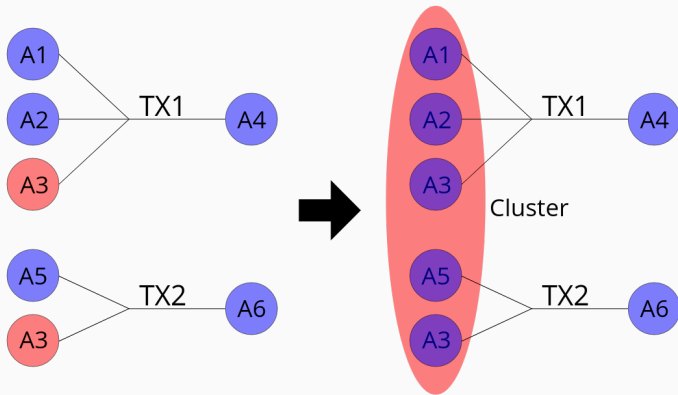
Multiple-input clustering heuristic

Background — Multiple-Input Clustering Heuristic



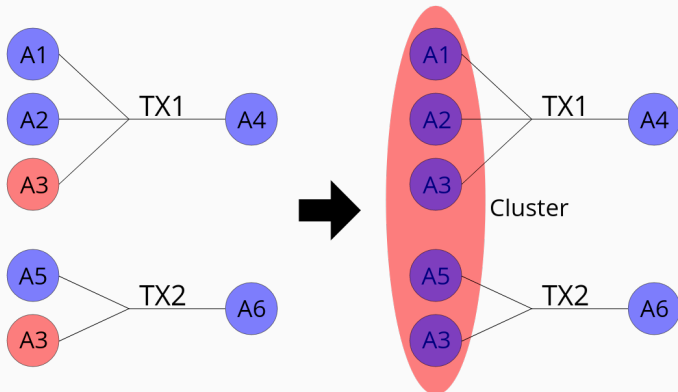
Multiple-input clustering heuristic

Background — Multiple-Input Clustering Heuristic



Multiple-input clustering heuristic

Background — Multiple-Input Clustering Heuristic



GRAPHSENSE

Block Attribution

Block Attribution — Data Sources

- Blocktrail API (till block 514239, March 2018)
- Blockchain.info Github repository
- BTC.com Github repository
- Walletexplorer.com's tags with multiple-input clustering heuristic (GraphSense)
- Coinbase markers manually retrieved

Block Attribution — Methodology

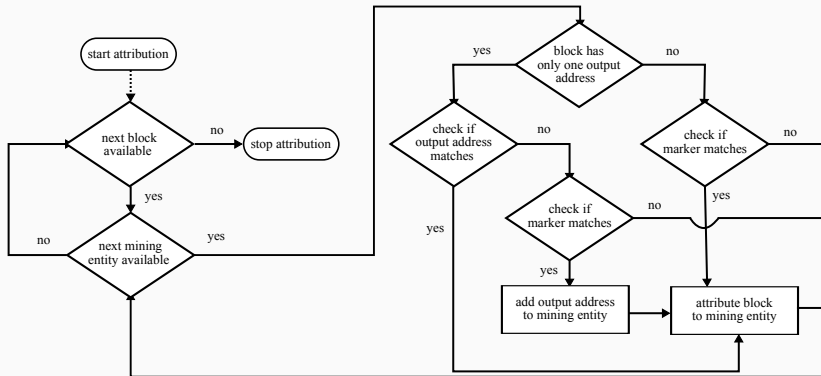


Figure 1: High level flow chart representing our attribution scheme.

Block Attribution — Results: Conflicts

Table 1: 684 attribution conflicts out of 556400 blocks (0.0012%). From 500,000 to 556,400, we attributed 96.5% of the blocks (blockchain.info 92%, ~ 32,100 BTC difference)

Miner 1	Miner 2	Number of conflicts	Example blocks
BTC.TOP	CANOE	338	516210, 516275, ...
Bixin	TangPool	142	339210, 339284, ...
BTC.com	Waterhole	113	478230, 478328, ...
BTC.TOP	WAYI.CN	81	509073, 509100, ...
ViaBTC	Okminer	5	510279, 523217,
Yourbtc	OzCoin	3	159846, 159929, 159964
BitcoinRussia	Bitcoin-Ukraine	1	524045
F2Pool	BTCC Pool	1	482886

Block Attribution — Results: Mining Shares

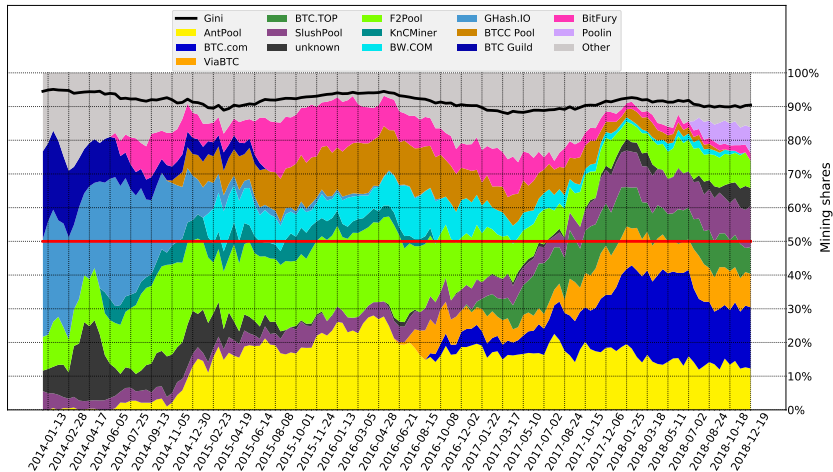


Figure 2: Evolution of mining pools market shares (2013-12-21 to 2018-12-19).

Payout Patterns

- Output of the block attribution procedure to select BTC.com, AntPool and ViaBTC blocks
- Blockchain.info API to retrieve coinbase flows
- GraphSense to detect exchanges

Payout Patterns — Methodology: BTC.com

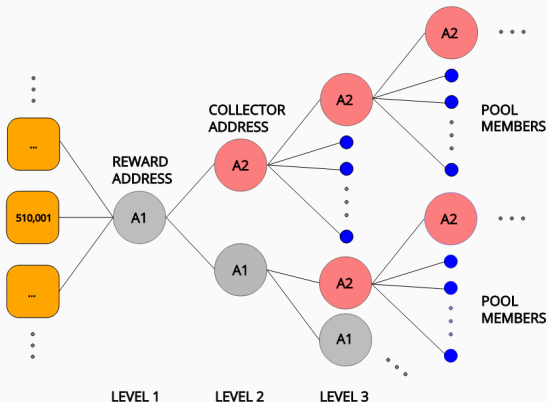


Figure 3: BTC.com payout pattern observed between block 510,000 and 514,032. In gray: reward addresses, in red: addresses performing payout transactions, in blue: pool members, in green: change addresses. Rounded squares are coinbases of blocks mined by the pool.

Payout Patterns — Methodology: AntPool

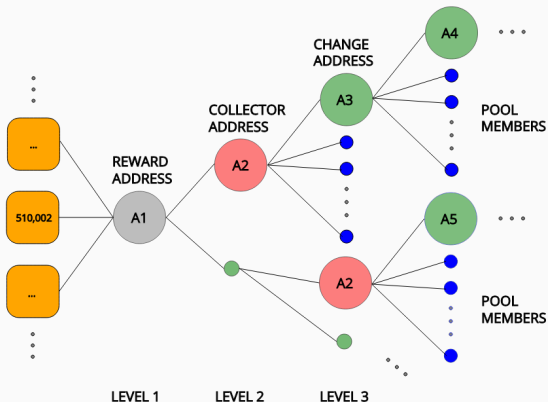


Figure 4: AntPool payout pattern observed between block 510,000 and 514,032. In gray: reward addresses, in red: addresses performing payout transactions, in blue: pool members, in green: change addresses. Rounded squares are coinbases of blocks mined by the pool.

Payout Patterns — Methodology: ViaBTC

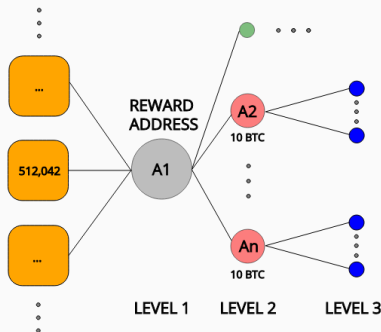


Figure 5: ViaBTC payout pattern observed between block 510,000 and 514,032. In gray: reward addresses, in red: addresses performing payout transactions, in blue: pool members, in green: change addresses. Rounded squares are coinbases of blocks mined by the pool.

Table 2: Statistics of retrieved data between block 510,000 and 514,032 (~ 4 weeks). N_B : number of blocks mined by the pool, N_{TX} : number of identified payout transactions, N_A : number of identified members' addresses, N_C : number of identified clusters, BTC_M : BTC mined by the pool, BTC_P : BTC paid to pool members (addresses), μ : median value of address reuse.

Pool Name	N_B	N_{TX}	N_A	N_C	BTC_M	BTC_P	$\frac{BTC_P}{BTC_M}$	μ	$\frac{\mu}{N_A}$
BTC.com	1,020	225	20,444	8,900	13,059	12,057	92%	20	9.8×10^{-4}
AntPool	617	408	14,166	5,082	7,887	2,333	30%	2	1.4×10^{-4}
ViaBTC	457	104	7,171	3,121	5,841	4,284	75%	5	7.0×10^{-4}

Pools Members

- Output of the block attribution procedure
- Output of the payout patterns
- Walletexplorer.com + multiple-input clustering heuristic (GraphSense)

Pools Members — Results: Cross-Pool Mining

Table 3: Cross-pool mining between block 510,000 and 514,032 (~ 4 weeks), including how much BTC from each pool has been received by those common addresses.

Pool 1	Pool 2	Addresses in common	Clusters in common	BTC from Pool 1	BTC from Pool 2
BTC.com	AntPool	537 (1.58%)	434 (3.2%)	664.3 (5.5%)	176.8 (7.6%)
AntPool	ViaBTC	115 (0.54%)	196 (2.4%)	11.1 (0.47%)	102.6 (2.4%)
ViaBTC	BTC.com	250 (0.91%)	267 (2.3%)	175.4 (4.1%)	174.1 (1.4%)

Pools Members — Results: Main Entities

Table 4: Cross-pool mining at a cluster level in the time period between block 510,000 and 514,032 (~ 4 weeks). W: wallet provider, E: exchange service, P: known mining pool, M: unknown mining entity.

Entity/Actor	Service	BTC.com			AntPool			ViaBTC			Total BTC
		BTC	%BTC	#Addr.	BTC	%BTC	#Addr.	BTC	%BTC	#Addr.	
Unknown	?	8930.39	74.07	13286	1682.25	72.09	8888	2877.02	67.17	4845	13489.67
Bixin	W+E+P	1663.75	13.80	1061	241.28	10.34	546	795.36	18.57	476	2700.39
Huobi.com	E	808.64	6.71	964	142.04	6.09	759	225.50	5.27	322	1176.19
Bittrex.com	E	83.71	0.69	348	29.56	1.27	251	43.36	1.01	177	156.63
Xapo.com	W	26.96	0.22	94	70.75	3.03	64	5.79	0.14	33	103.50
Poloniex.com	E	42.65	0.35	381	11.52	0.49	268	19.97	0.47	139	74.15
Luno.com	W+E	36.59	0.30	258	4.06	0.17	104	4.39	0.10	60	45.04
Bitstamp.net	E	8.94	0.07	57	3.55	0.15	38	3.91	0.09	22	16.39
Cryptonator.com	W+E	5.75	0.05	80	0.70	0.03	41	2.70	0.06	33	9.15
BitoEX.com	W	5.09	0.04	23	1.12	0.05	35	2.19	0.05	4	8.39
CoinHako.com	W+E	3.59	0.03	4	0.29	0.01	3	0.24	0.01	2	4.12
Bitcoin.de	E	1.86	0.02	26	0.76	0.03	13	0.58	0.01	7	3.19

Pools Members — Results: Pool Centralization

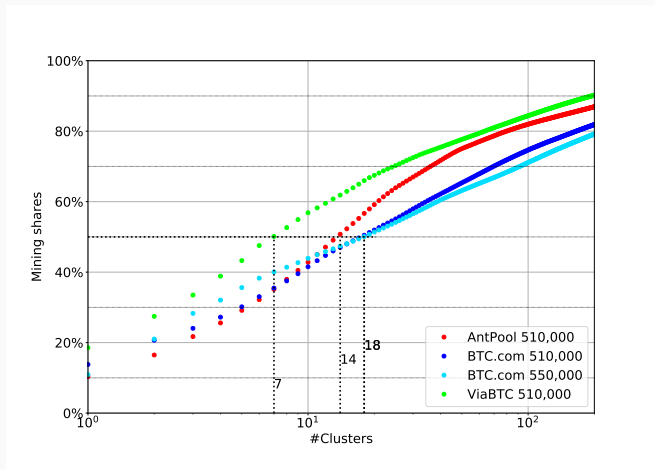


Figure 6: Cumulative sum of mining shares over clusters for each pool (log-scale). Black-dotted lines highlight the number of clusters controlling 50% of each pool.

Future work

- Further research on unknown entities (entity classification)
- Improve heuristics to get more payout patterns
- Extend analysis to other pools
- IP-network traffic measurement

Conclusion

Conclusion

- 3 pools can easily reach 50% of the hash rate
- It is possible to find pool members because pools follow payout patterns
- High hash-rate concentration within major pools
- Cross-pool mining occurs
- Unknown mining prevails

Appendix

Block Attribution — Results

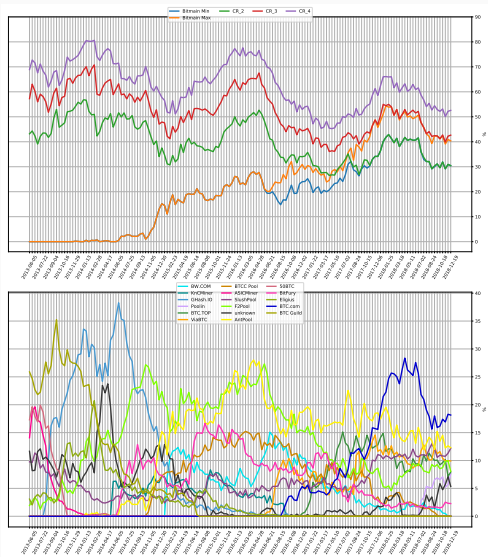


Figure 7: Concentration indexes and mining shares.

Pools Members — Results: Cross-Pool Mining

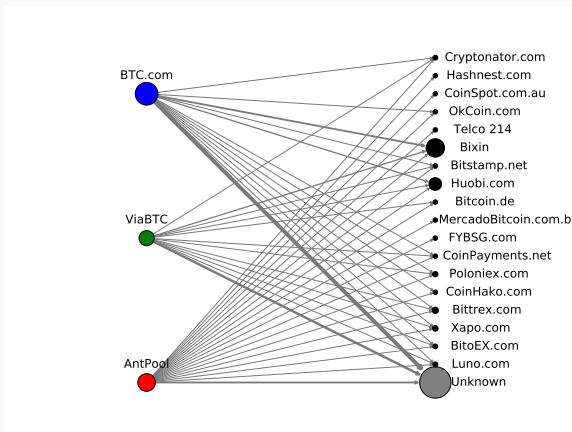


Figure 8: Flow of mining rewards from mining pools to their members. In black: wallet services and exchanges, in gray: unknown entities. This plot covers the top 400 clusters from each mining pool sorted by received BTC. *Unknown* entities (1118) were combined into one node.

Pools Members — Results: Cross-Pool Unknown Mining

Table 5: Cross-pool mining of the ten largest unknown mining clusters sorted by total amount of BTC received by the three pools in the time period between block 510,000 and 514,032 (~ 4 weeks).

Cluster ID	BTC.com		AntPool		ViaBTC		Mined BTC	Total BTC Received
	BTC	%BTC	BTC	%BTC	BTC	%BTC		
327539880	409.34	3.40	122.10	5.23	258.55	6.04	789.99	521,939
324067473	295.02	2.45	90.44	3.88	189.15	4.42	574.61	3,756,583
350822682	244.77	2.03	9.29	0.40	182.92	4.27	436.98	110,566
350824718	244.67	2.03	65.65	2.81	46.20	1.08	356.52	112,680
333653856	153.02	1.27	54.02	2.31	83.60	1.95	290.63	130,680
372448840	181.10	1.50	33.64	1.44	55.73	1.30	270.48	882,713
234254928	93.31	0.77	27.18	1.16	58.68	1.37	179.17	905,101
249123673	15.63	0.13	0.40	0.02	107.23	2.50	123.26	6,812,938
349962609	8.67	0.07	39.01	1.67	19.74	0.46	67.41	1,173,892
311503667	38.94	0.32	7.47	0.32	7.77	0.18	54.18	486,338

Algorithm 1 Find payout patterns in BTC.com, AntPool and ViaBTC.

```
1: for each pool do  
2:   for each mined block do  
3:     get coinbase flow of block  
4:     compute number of addresses at each step of the coinbase flow  
5:     save results  
6:   plot data and look for common patterns among flows
```

Payout Patterns — Methodology

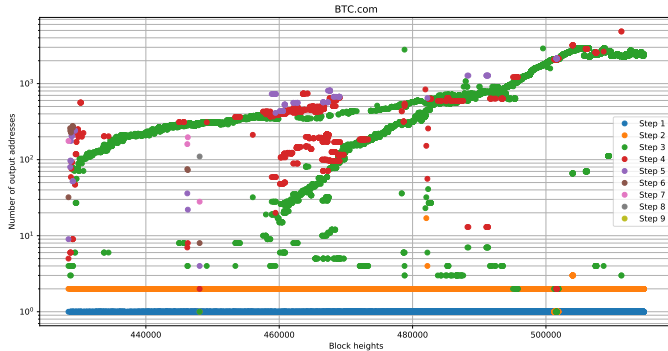


Figure 9: Payout trend for BTC.com