Karlsruhe Institute of Technology

KASTEL

Simulation-based Evaluation of Coin Selection Strategies

Scaling Bitcoin: Retarget

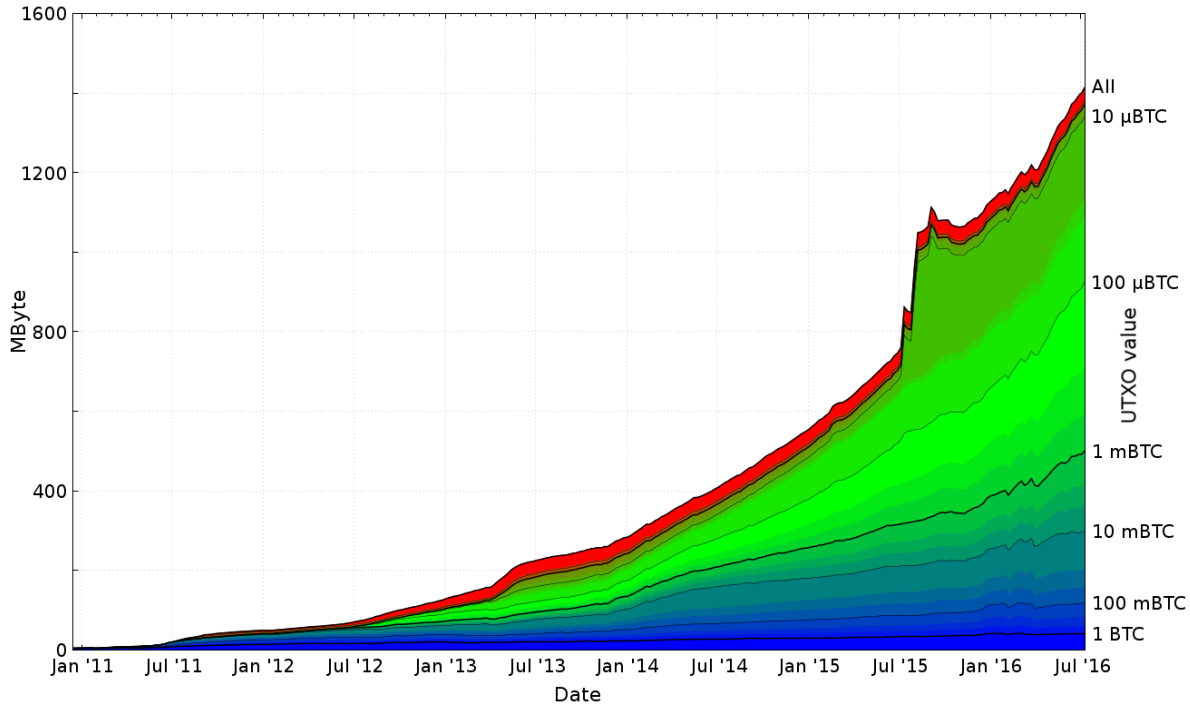
Mark Erhardt | October 9, 2016

INSTITUTE OF TELEMATICS

Outline

1. Motivation
2. Coin Selection
3. Framework
4. Simulation
5. Conclusion

UTXO Set Growth

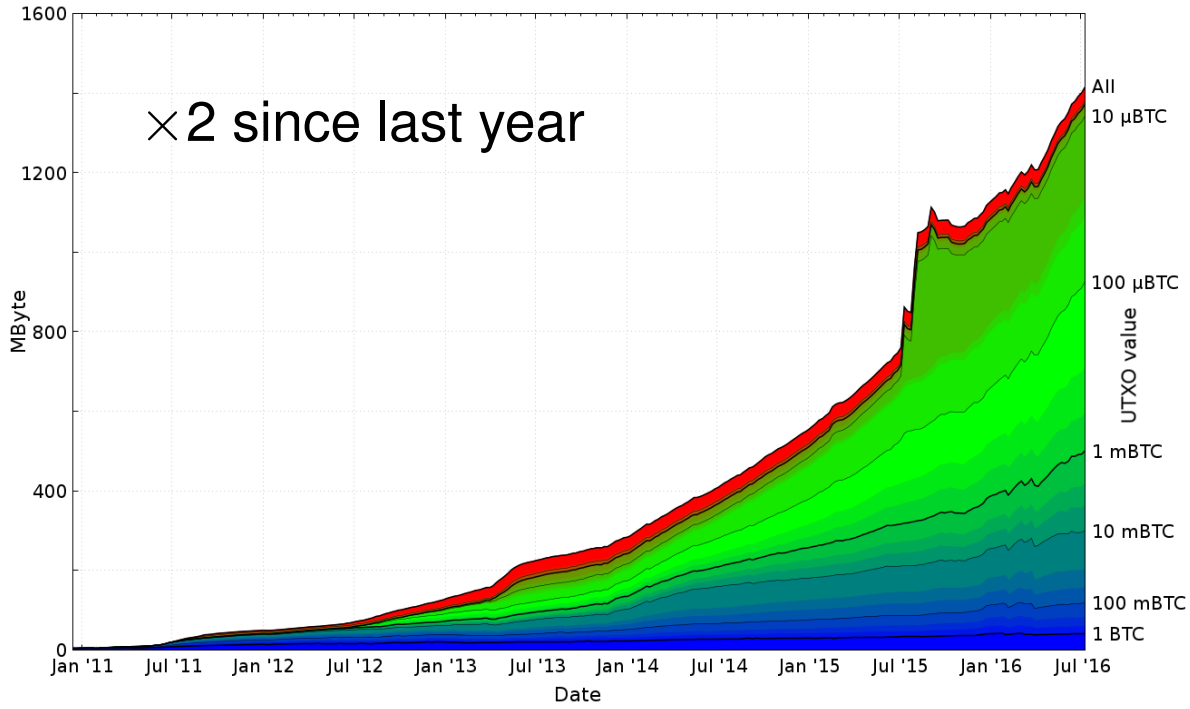


[Wuille, 2016]

What's the issue with the growth?

- UTXO set kept in RAM by miners
- average UTXO value:
  0.97 BTC (2014), 0.44 BTC (2015), 0.38 BTC (2016)

UTXO Set Growth

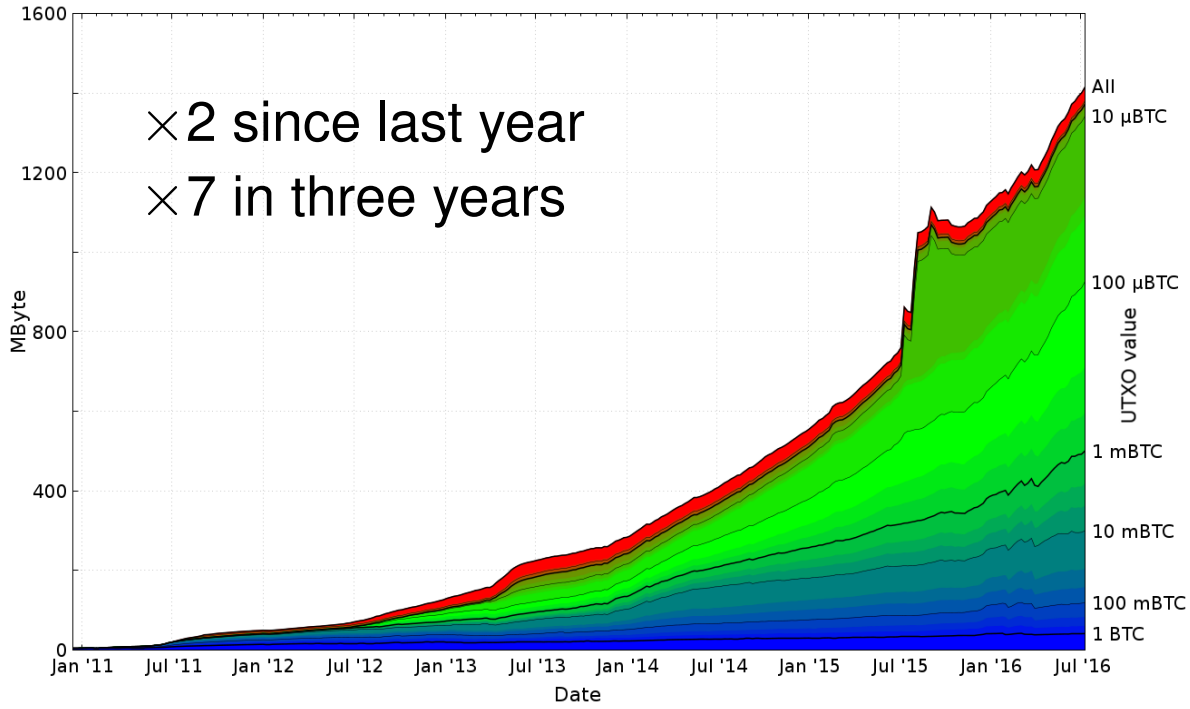


[Wuille, 2016]

What's the issue with the growth?

- UTXO set kept in RAM by miners
- average UTXO value:
  0.97 BTC (2014), 0.44 BTC (2015), 0.38 BTC (2016)

UTXO Set Growth

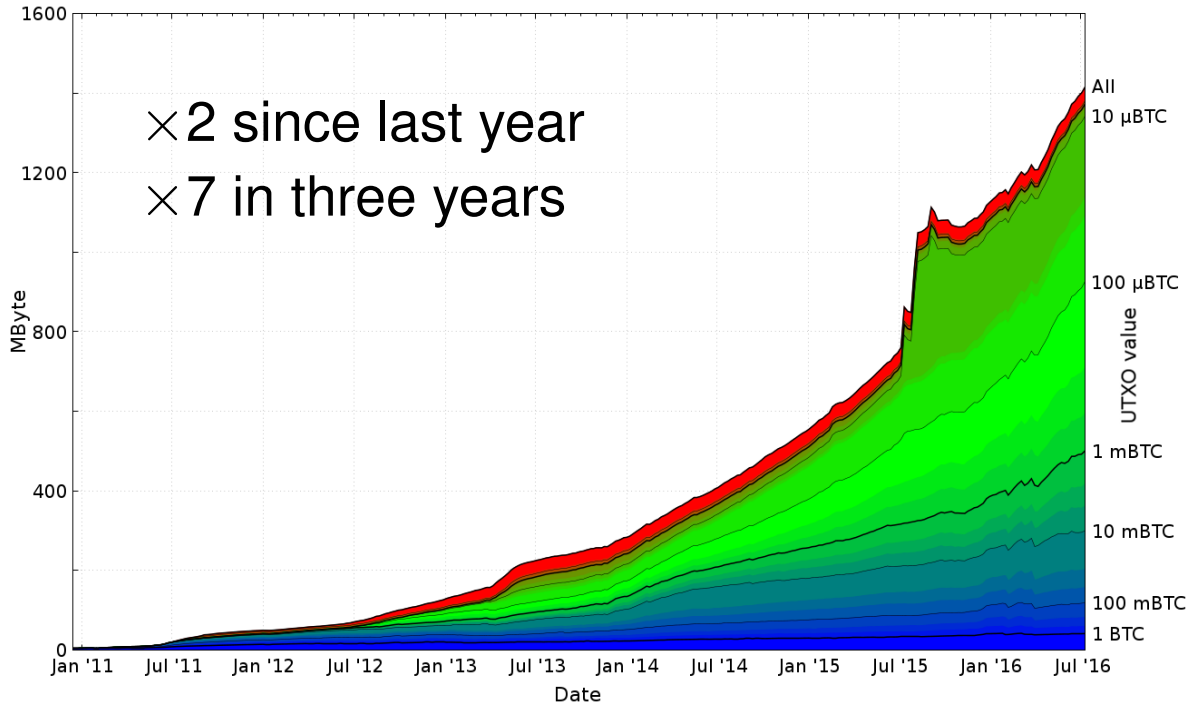


[Wuille, 2016]

What's the issue with the growth?

- UTXO set kept in RAM by miners
- average UTXO value:
  0.97 BTC (2014), 0.44 BTC (2015), 0.38 BTC (2016)

UTXO Set Growth

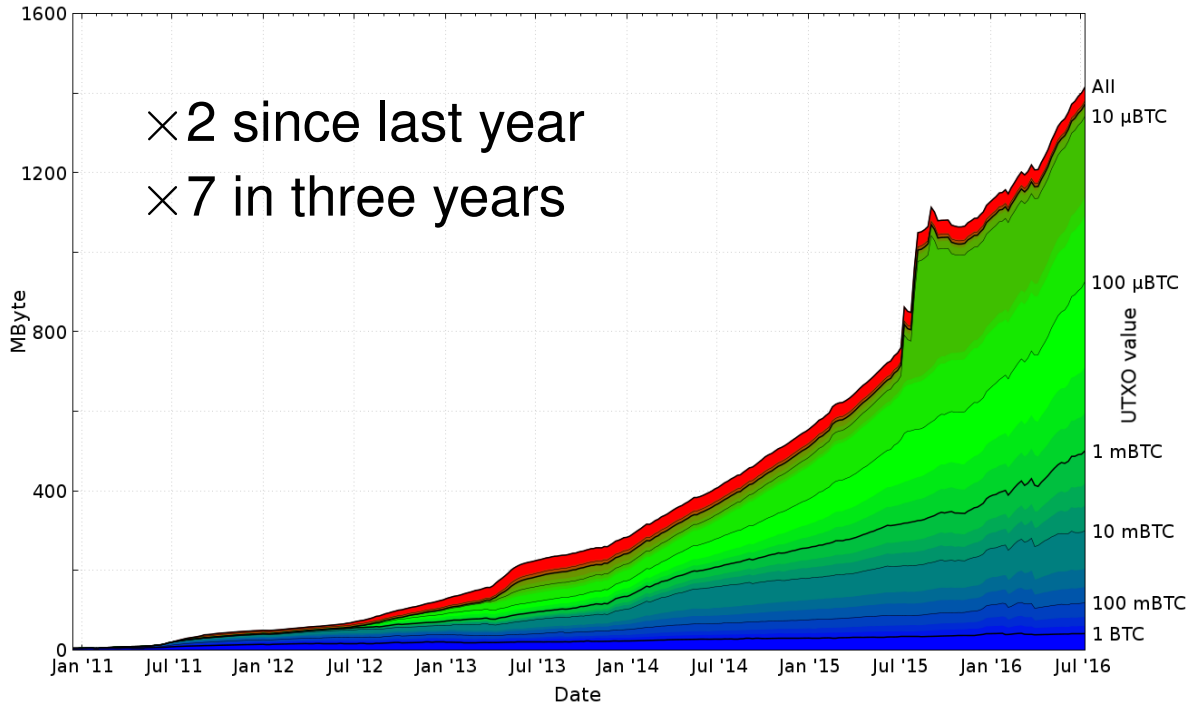


[Wuille, 2016]

What's the issue with the growth?

- UTXO set kept in RAM by miners
- average UTXO value:
  0.97 BTC (2014), 0.44 BTC (2015), 0.38 BTC (2016)

UTXO Set Growth



[Wuille, 2016]

What's the issue with the growth?

- UTXO set kept in RAM by miners
- average UTXO value:
  0.97 BTC (2014), 0.44 BTC (2015), 0.38 BTC (2016)

Coin Selection

How to choose which outputs to spend?

Coin Selection

How to choose which outputs to spend?

Hypothesis: Improved Coin Selection can reduce UTXO Set.

Hard Constraints and Goals

Constraints for Coin Selection

- available UTXO pool
- provide sufficient funds for payment and fee
- no dust outputs

Goals

- Minimize fees
- Reduce UTXO set
- Privacy

Hard Constraints and Goals

Constraints for Coin Selection

- available UTXO pool
- provide sufficient funds for payment and fee
- no dust outputs

Goals

- Minimize fees
- Reduce UTXO set
- Privacy

Conditions and Factors

Changing Conditions

- Priority $\rightarrow$ fee-market
- Blockspace demand

Influential Factors

- Payment sizes
- Short-term fees vs long-term fees
- Ratio of incoming and outgoing payments
- Size of newly generated changes

Conditions and Factors

Changing Conditions

- Priority $\rightarrow$ fee-market
- Blockspace demand

Influential Factors

- Payment sizes
- Short-term fees vs long-term fees
- Ratio of incoming and outgoing payments
- Size of newly generated changes

Some Ideas for Improvement

Strategy Ideas

- Changes of average target size (proposed by Luke-Jr in IRC)
- Add tiny change outputs to fee
- Target sized change outputs
- Random inputs

Some Ideas for Improvement

Strategy Ideas

- Changes of average target size (proposed by Luke-Jr in IRC)
- Add tiny change outputs to fee
- Target sized change outputs
- Random inputs

Some Ideas for Improvement

Strategy Ideas

- Changes of average target size (proposed by Luke-Jr in IRC)
- Add tiny change outputs to fee
- Target sized change outputs
- Random inputs

Some Ideas for Improvement

Strategy Ideas

- Changes of average target size (proposed by Luke-Jr in IRC)
- Add tiny change outputs to fee
- Target sized change outputs
- Random inputs

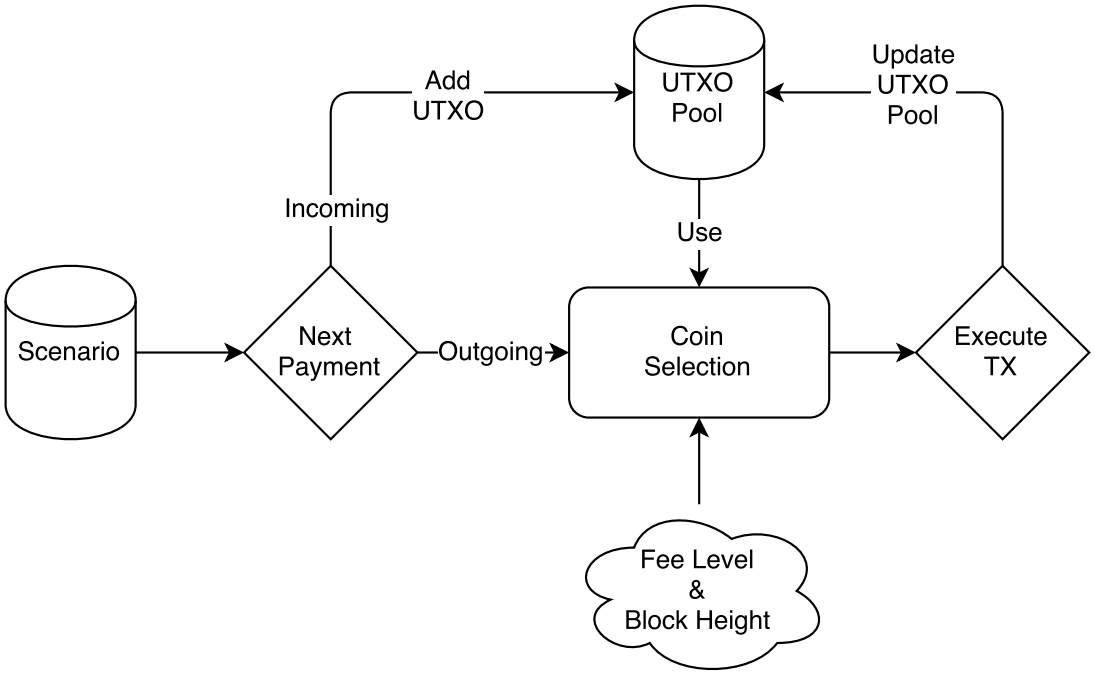
Some Ideas for Improvement

Strategy Ideas

- Changes of average target size (proposed by Luke-Jr in IRC)
- Add tiny change outputs to fee
- Target sized change outputs
- Random inputs

How to evaluate?

CoinSelectionSimulator



Considers:

- Selection policy
- Fees
- Transaction format (P2PKH or P2WPKH)
- Block height

Doesn't consider (yet):

- Addresses

Some Policies In Use (1/3)

Oldest First

- FIFO
- Add change outputs below *DustLimit* to fee
- *BreadWallet*  , *Electrum* 

Pruned Oldest First

- FIFO
- Post-selection pruning of smallest inputs
- Add change outputs below 5460 satoshis to fee
- *Mycelium* 

Some Policies In Use (1/3)

Oldest First

- FIFO
- Add change outputs below *DustLimit* to fee
- *BreadWallet*  , *Electrum* 

Pruned Oldest First

- FIFO
- Post-selection pruning of smallest inputs
- Add change outputs below `5460` satoshis to fee
- *Mycelium* 

More Policies in Use (2/3)

Highest Priority First

- Sorted by priority (age×value)
- *BitcoinJ* , *Bitcoin Wallet for Android*

Target Sized Change

- `n` randomly drawn buckets of inputs
- Select bucket that minimizes $\delta(change, target)$
- *Electrum Private Mode (not implemented yet in Simulation)*

More Policies in Use (2/3)

Highest Priority First

- Sorted by priority (age$\times$value)
- *BitcoinJ* bitcoinj , *Bitcoin Wallet for Android*

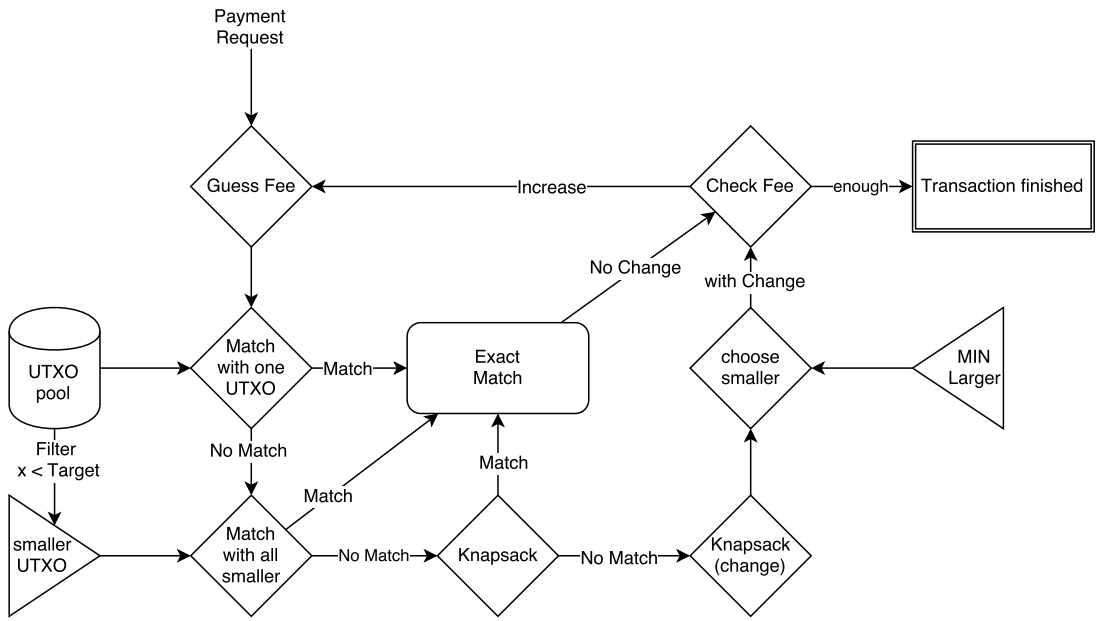
Target Sized Change

- `n` randomly drawn buckets of inputs
- Select bucket that minimizes $\delta(change, target)$
- *Electrum Private Mode (not implemented yet in Simulation)*

Bitcoin Core's Policy (3/3)

Avoid Change or Large Change

- attempts direct match
- pseudo-random knapsack algorithm
- minimum change of 0.01 BTC
- *Bitcoin Core* BitcoinCore



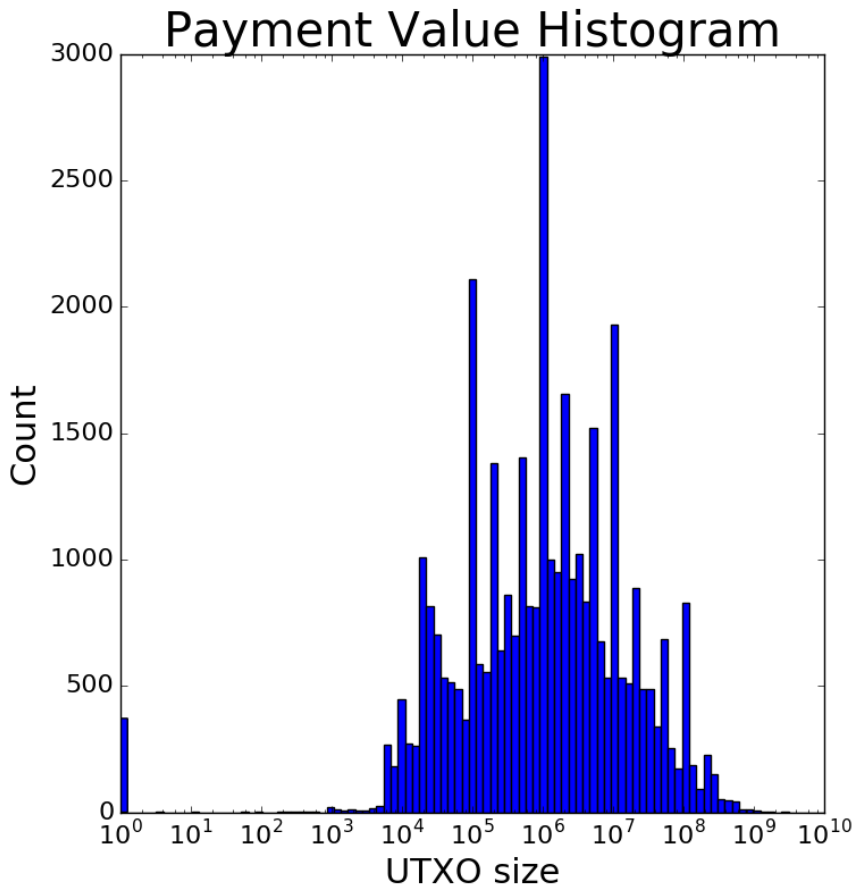
Simulation Scenario

Transaction data from moneypot.com [Havar, 2015]

- 24,388 incoming payments
- 11,860 outgoing payments

Other experiments:

- moneypot.com incoming condensed
- Gaussian generated payments



Simulation Results

Policy	∅ #UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
FIFO	182.87	399.62	629.07	3.03
Pruned FIFO	763.73	169.93	623.39	2.91
Highest Priority	2 551.52	789.52	629.05	2.50
"Core"	180.30	31.75	819.03	3.05

Results are highly scenario dependent!

Observations

- FIFO maintains almost as few UTXO as Core
- Pruned FIFO and Highest Priority accumulate small UTXO
- Bitcoin Core: overpays fees, computationally expensive, only $\approx 0.5\%$ Direct Matches (63 of 11860)

Simulation Results

Policy	∅ #UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
FIFO	182.87	399.62	629.07	3.03
Pruned FIFO	763.73	169.93	623.39	2.91
Highest Priority	2 551.52	789.52	629.05	2.50
"Core"	180.30	31.75	819.03	3.05

Results are highly scenario dependent!

Observations

- FIFO maintains almost as few UTXO as Core
- Pruned FIFO and Highest Priority accumulate small UTXO
- Bitcoin Core: overpays fees, computationally expensive, only $\approx 0.5\%$ Direct Matches (63 of 11860)

Simulation Results

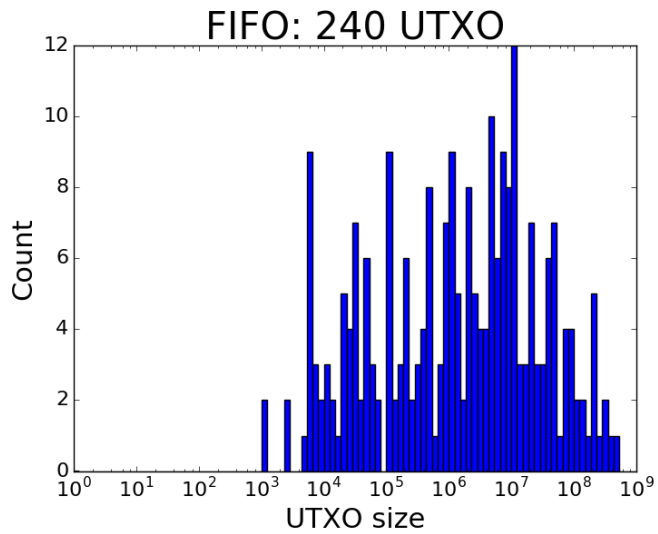
Policy	∅ #UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
FIFO	182.87	399.62	629.07	3.03
Pruned FIFO	763.73	169.93	623.39	2.91
Highest Priority	2 551.52	789.52	629.05	2.50
"Core"	180.30	31.75	819.03	3.05

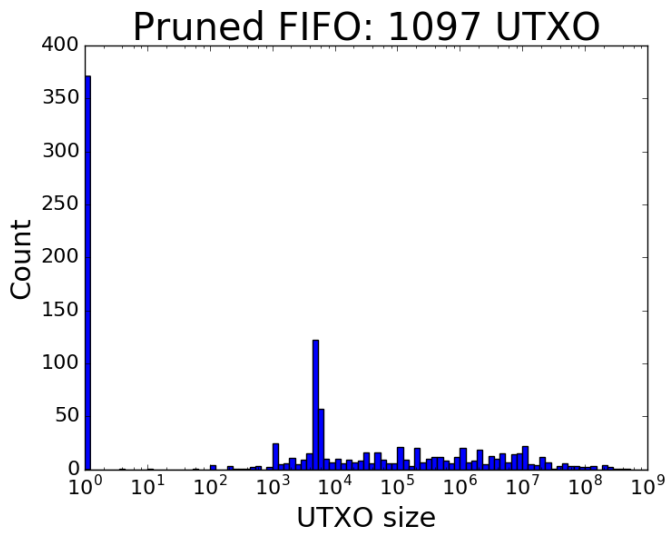
Results are highly scenario dependent!

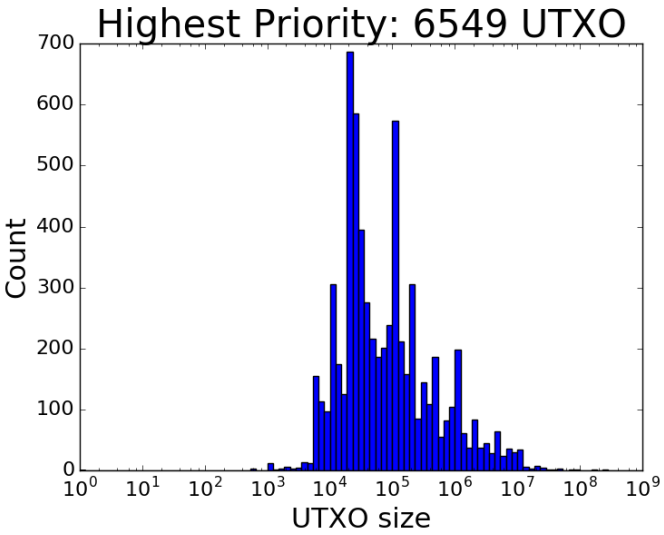
Observations

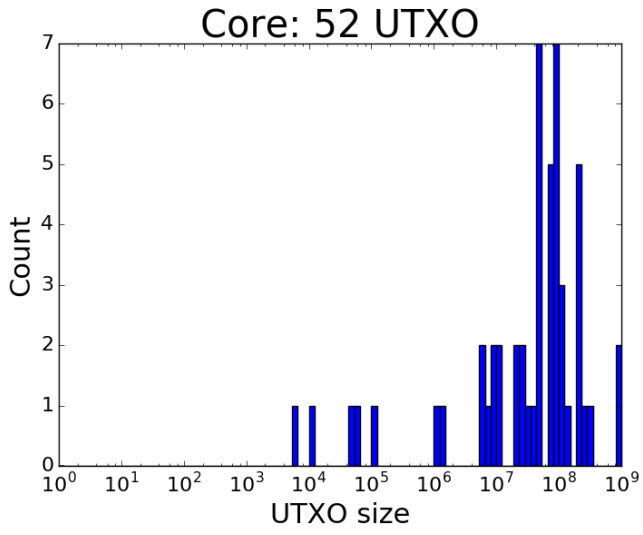
- FIFO maintains almost as few UTXO as Core
- Pruned FIFO and Highest Priority accumulate small UTXO
- Bitcoin Core: overpays fees, computationally expensive, only $\approx 0.5\%$ Direct Matches (63 of 11860)

Histogram of final UTXO pools









Simulation Results other strategies

Policy	∅ UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
Average Target	137.89	207.37	767.08	3.04
Wider Match Donation	165.24	32.95	829.38	3.02
Double Target	225.00	198.39	832.41	3.03
Single Random Draw (No MC)	185.16	384.43	629.13	3.03
Single Random Draw (0.01 BTC)	173.27	424.15	628.98	3.04
Core	180.30	31.75	819.03	3.05

Simulation Results other strategies

Policy	∅ UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
Average Target	137.89	207.37	767.08	3.04
Wider Match Donation	165.24	32.95	829.38	3.02
Double Target	225.00	198.39	832.41	3.03
Single Random Draw (No MC)	185.16	384.43	629.13	3.03
Single Random Draw (0.01 BTC)	173.27	424.15	628.98	3.04
Core	180.30	31.75	819.03	3.05

Average Target

- Bitcoin Core's selection
- MIN_CHANGE is mean target
- proposed by Luke-Jr

Simulation Results other strategies

Policy	∅ UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
Average Target	137.89	207.37	767.08	3.04
Wider Match Donation	165.24	32.95	829.38	3.02
Double Target	225.00	198.39	832.41	3.03
Single Random Draw (No MC)	185.16	384.43	629.13	3.03
Single Random Draw (0.01 BTC)	173.27	424.15	628.98	3.04
Core	180.30	31.75	819.03	3.05

Wider Match Donation

- Bitcoin Core's selection
- Range of $+(0, \mathit{DustLimit}) =$ Direct Match
- Add change up to Dust Limit to fee

Simulation Results other strategies

Policy	∅ UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
Average Target	137.89	207.37	767.08	3.04
Wider Match Donation	165.24	32.95	829.38	3.02
Double Target	225.00	198.39	832.41	3.03
Single Random Draw (No MC)	185.16	384.43	629.13	3.03
Single Random Draw (0.01 BTC)	173.27	424.15	628.98	3.04
Core	180.30	31.75	819.03	3.05

Double Target

- Bitcoin Core's selection
- MIN_CHANGE = target

Simulation Results other strategies

Policy	∅ UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
Average Target	137.89	207.37	767.08	3.04
Wider Match Donation	165.24	32.95	829.38	3.02
Double Target	225.00	198.39	832.41	3.03
Single Random Draw (No MC)	185.16	384.43	629.13	3.03
Single Random Draw (0.01 BTC)	173.27	424.15	628.98	3.04
Core	180.30	31.75	819.03	3.05

Single Random Draw

- Shuffle UTXO pool, pop front until sufficient once
- Equi-probably selection
- No MIN_CHANGE, MIN_CHANGE = 0.01 BTC

Conclusion

Presented:

- Examined Coin Selection Strategies
- Identified improvement opportunities for several prevalent strategies
- Simulation Framework: https://github.com/Xekyo/CoinSelectionSimulator (late October!)

Future Work:

- Addresses
- Privacy
- Only one Scenario: Additional scenario data welcome!

Conclusion

Presented:

- Examined Coin Selection Strategies
- Identified improvement opportunities for several prevalent strategies
- Simulation Framework: https://github.com/Xekyo/CoinSelectionSimulator (late October!)

Future Work:

- Addresses
- Privacy
- Only one Scenario: Additional scenario data welcome!

Thank you for your attention!

References I


Havar, R. (2015).
Issue#1643: Coinselection prunes extraneous inputs from ApproximateBestSubset.
[Online, retrieved on 2016-08-18].


Wuille, P. (2016).
UTXO breakdown per output amount.
[Online, retrieved on 2016-07-13].

Privacy

- Addresses easy to model

But:

- Taint or Value privacy?
- How to measure privacy?
- Simulation of Address behavior whole new problem

Framework Features and Limits

Features:

- Queuing of Outgoing Payments when insufficient Funds
- `nLockTime` for each Payment
- Fee estimation
- Multiple Wallets in parallel
- Extensive statistics:
  final value, mean #UTXO, final #UTXO, #received, #spent, #changes created, smallest change, biggest change, mean change, stDev of change, in transit ratio, total fees, average fees, fees to spend remaining UTXO, total cost, smallest input set, biggest input set, mean size of input set, stdev of input set size, final UTXO set

But:

- No Addresses
- `nLockTime` so far only simple Gaussian interval
- Results highly scenario dependent

Core Improvement Opportunities

- Skip useless iteration with zero fee.
- Estimate fee respective to selected set
- Direct Match only occurs in 63 of 11,860 cases. (Pruned FIFO: 725)
  $\rightarrow$ Less emphasis.

Mycelium DustLimit vs 5460

Policy	∅ #UTXO	Final #UTXO	∅ change [mBTC]	total cost [mBTC]	∅ #inputs
Pruned FIFO (5460)	763.73	1,013	169.93	623.39	2.91
Pruned FIFO (DustLimit)	774.30	1,097	170.44	633.73	2.91