Coin Selection

@DG Lab - Karl-Johan Alm

- We receive bitcoin in transactions.
- We refer to them using outpoints, which are txid + output index.
- Even if we use the same address (which we shouldn't) to receive multiple times, each one will have a different outpoint. Even if transaction is the same.

Our "balance" is really just the sum of the coins in each of the outpoints that we are able to spend (usually by having the private key).

To send coins, we have to (1) pick enough coins to cover the amount, and (2) send us back the change, if any.

Ideally, we want to minimize the *number* of coins we spend at any one time, for several reasons:

- 1. Every time we "aggregate", we connect the histories of the two outpoints to one owner.
- 2. We pay fees based on the tx size; more coins means bigger size.

We are also incentivized to try to find coins so that the input is about the same as the output. If we manage, we can skip the change output, which means a smaller transaction (= lower fees).

This process is called "Coin Selection".

Coins are selected as follows:

- Shuffle coins
- Iterate over coins;
 - if coin value == target, return coin
 - \circ $\,$ if coin value < target, put it in "value" list and add its value to the total
 - otherwise, if coin value < smallest previous coin larger than target, mark it as smallestHigher
- If the sum of the total == target, return the "value" list
- If the sum is less, and there is a marked coin, return the marked coin
- ...

Coins are selected as follows:

- Sort value list in reversed order
- Approximate the best subset (see next slide)
- Return the approximated solution, or (if bad), the marked coin

Subset approximation:

- Repeat the following a large number of times (1000):
 - Loop over coins twice, and include the coins if a random boolean is true, or if this is the second pass and the coin has not been included yet.
 - If target was reached, see if the total beats the high score (by being smaller), and keep the solution if so.

The entire process is executed up to 7 times, with different parameters. It prefers *not* to use

- recently confirmed or unconfirmed coins
- coins in a long unconfirmed chain

among other things.





@DG Lab - Karl-Johan Alm

Twitter: @kallewoof