Report over

Bitcoin
A peer to peer Electronic Cash system

&

A Fistful of Bitcoins:
Characterizing Payments Among men with no names
Bitcoin: a peer to peer electronic cash system

The specific paper proposes a new electronic payment system, based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. The main problem that emerges in an electronic transaction is the possibility of double spending. Double spending is spending the same coin twice. In digital currency if the verification mechanism is missing it can lead to it. Anyone can copy that digital money and send them somewhere else. So, in this paper, a solution is proposed to the double-spending problem using a peer-to-peer distributed timestamp server to generate computational proof of the chronological order of transaction.

Implementation

An electronic coin is defined as a chain of digital signatures. Each owner transfers the coin to the next by digitally signing a hash of the previous transaction and the public key of the next owner and adding these to the end of the coin. A payee can verify the signatures to verify the chain of ownership.

So, how do we solve the double spending problem? In order to avoid a third party, we use a timestamp server. A timestamp server works by taking a hash of a block of items to be time-stamped and widely publishing the hash. The timestamp proves that the data must have existed at the time, obviously, in order to get into the hash. To implement a distributed timestamp server on a peer-to-peer basis, we will need to use a proof-of-work system. It involves scanning for a value that when hashed, such as with SHA-256 (NSA cryptographic hash algorithm). The hash begins with a number of zero bits. Block contains transactions to be validated and previous hash value. In Blockchain, this algorithm is used to confirm transactions and produce new blocks to the chain. This
is how mining works, and bitcoins are produced. With PoW, miners compete against each other to complete transactions on the network and get rewarded.

When a new transaction is about to happen:

1. Transaction is broadcasted to all nodes.
2. Each node collects new transactions into a block
3. Each node works on finding a difficult proof-of-work for its block
4. When a node finds a proof-of-work, it broadcasts the block to all nodes
5. Nodes accept the block only if all transactions in it are valid and not already spent
6. Nodes express their acceptance of the block by working on creating the next block in the chain, using the hash of the accepted block as the previous hash.

Nodes always consider the longest chain to be the correct one and will keep working on extending it. The tie will be broken when the next proof-of-work is found and one branch becomes longer; the nodes that were working on the other branch will then switch to the longer one.

Once the latest transaction in a coin is buried under enough blocks, the spent transactions before it can be discarded to save disk space. To facilitate this without breaking the block's hash, transactions are hashed in a Merkle Tree, with only the root included in the block's hash. Old blocks can then be compacted by stubbing off branches of the tree. The interior hashes do not need to be stored.

The user, in order to verify the payment, only needs to keep a copy of the block headers of the longest proof-of-work chain, which he can get by querying network nodes until he's convinced he has the longest chain. He can't check the transaction for himself, but by linking it to a place in the chain, he can see that a network node has accepted it, and blocks added after it further confirm the network has accepted it.

Although it would be possible to handle coins individually, it would be unwieldy to make a separate transaction for every cent in a transfer. To
allow value to be split and combined, transactions contain multiple inputs and outputs. Normally there will be either a single input from a larger previous transaction or multiple inputs combining smaller amounts, and at most two outputs: one for the payment, and one returning the change, if any, back to the sender.

E-cash systems, are famous because of their anonymity. Bitcoin manages this by making all transactions available to everyone but keeping the public keys anonymous. We can see that someone is sending money to another but we don’t know anything about him. As an additional firewall, a new key pair should be used for each transaction to keep them from being linked to a common owner. These transactions, apart from anonymous, they are very safe. It quickly becomes computationally impractical for an attacker to change transaction history, if honest nodes control a majority of CPU power. If a greedy attacker is able to assemble more CPU power than all the honest nodes, he would have to choose between using it to defraud people by stealing back his payments, or using it to generate new coins.
A fistful of Bitcoins: Characterizing payments among men with no names

As we showed in the paper above bitcoin is a peer to peer anonymous transaction system. This paper briefly introduces bitcoin and its characteristics. It analyzes its unique characteristic further, using heuristic clustering to group Bitcoin wallets based on evidence of shared authority, and then using re-identification attacks (i.e., empirical purchasing of goods and services) to classify the operators of those clusters. From this analysis, we characterize longitudinal changes in the Bitcoin market, the stresses these changes are placing on the system, and the challenges for those seeking to use Bitcoin for criminal or fraudulent purposes at scale. The goal is not to generally de-anonymize all Bitcoin users but rather to identify certain idioms of use present in concrete Bitcoin network implementations that erode the anonymity of the users who engage in them. The approach is based on the current availability of the Bitcoin blockchain.

There are many different participants in the Bitcoin ecosystem. We can identify them as:

**Miners**: The ones that work for generating the next block in the “Blockchain”. Generating a block is so computationally difficult that very few individual users attempt it on their own. Instead, users may join a mining pool.

**Purchasers**: buying bitcoins through one of the many exchanges

**Spenders**: eg.
○ gamble with one of the popular dice games such as Satoshi Dice
○ buy items from various online vendors, such as Bitmit ("the eBay of Bitcoin"), the notorious Tor-based service Silk Road, or with vendors, such as Wordpress

Investors: with firms such as Bitcoinica (shut down after a series of thefts) or Bitcoin Savings & Trust (later revealed as a major Ponzi scheme). Most of investment firms in bitcoin proved to be fraud.

Analyzing the transactions since the beginning of the bitcoin, proved 2 different eras:

- Adoption time: The first 15 months that Bitcoin was deployed almost all transactions involved exactly 50 bitcoins (the initial reward for mining a block)
- Second turning point in early 2012, in which the percentage of transactions carrying less than a single bitcoin in total value doubled abruptly (from 20% to 40%), while the percentage of transactions carrying less than 0.1 BTC tripled

Also by analyzing the bitcoin flow there are 2 clear turning points:

- In early 2011, represents a point at which users began meaningfully spending bitcoins, rather than just “hoarding” them
- In April 2012, the percentage of bitcoins being spent immediately doubled, and more generally half of all bitcoins are now spent within an hour of being received and 80% of bitcoins are spent within a day
In this paper, in order to identify public keys belonging to the types of services, they sought to “tag” as many addresses as possible. The predominant method for tagging users was simply transacting with them. These transactions happened between the researchers and:

- Mining Pools
- Wallets
- Bank exchanges (real time trading exchanges)
- Non-bank exchanges (fixed rate exchanges)
- Vendors (buying goods)
- Gambling (Satoshi Dice)
- Miscellaneous
- Other Source

<table>
<thead>
<tr>
<th>Mining</th>
<th>Wallets</th>
<th>Exchanges</th>
<th>Vendors</th>
<th>Gambling</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 BTC</td>
<td>BTC Guild</td>
<td>BTC-e</td>
<td>BTC Buy</td>
<td>Bit Elfin</td>
<td>Bit Visitor</td>
</tr>
<tr>
<td>ABC Pool</td>
<td>Deepbit</td>
<td>CA VinEx</td>
<td>BTCA</td>
<td>BitZino</td>
<td>Bifog</td>
</tr>
<tr>
<td>Bitclockers</td>
<td>EclipseMC</td>
<td>ICBit</td>
<td>BTC Gadgets</td>
<td>BTC Griffin</td>
<td>Bitfog</td>
</tr>
<tr>
<td>Bitminter</td>
<td>Eligius</td>
<td>Mercado Bitcoin</td>
<td>Bitcoin</td>
<td>BTC Griffin</td>
<td>Bitfog</td>
</tr>
<tr>
<td>My Wallet</td>
<td>Easywallet</td>
<td>Mt Gox</td>
<td>BittInstant</td>
<td>BTC Griffin</td>
<td>Bitfog</td>
</tr>
<tr>
<td>Coinbase</td>
<td>Instawallet</td>
<td>The Rock</td>
<td>Strongscoin</td>
<td>FastCash/Bitcoins</td>
<td>Bitfog</td>
</tr>
<tr>
<td>Easycoin</td>
<td>Paytuna</td>
<td>Naraimo Gold</td>
<td>WalletBit</td>
<td>Freelancer</td>
<td>Bitfog</td>
</tr>
<tr>
<td>BTC China</td>
<td>Viewex</td>
<td>OKPay</td>
<td>Vircurex</td>
<td>Satoshi Dice</td>
<td>Bitfog</td>
</tr>
</tbody>
</table>

- Mining Pools
- Wallets
- Bank exchanges (real time trading exchanges)
- Non-bank exchanges (fixed rate exchanges)
- Vendors (buying goods)
- Gambling (Satoshi Dice)
- Miscellaneous
- Other Source
They used to main heuristics in order to identify and classify users among these transactions. The aim was to link addresses controlled by the same user, with the goal of collapsing the many public keys seen in the block chain into larger entities.

**HEURISTIC 1**: If two (or more) addresses are inputs to the same transaction, they are controlled by the same user; i.e., for any transaction \( t \), all \( pk \in \text{inputs}(t) \) are controlled by the same user. (pk = public key)

- Using this heuristic, they partitioned the network into 5,579,176 clusters of users.
- some of them corresponded to the same user; e.g., there were 20 clusters that we tagged as being controlled by Mt. Gox.(big services spread their funds across many accounts)
- This cross-cluster naming was nevertheless not too common, and we thus ended up with 5,577,481 distinct clusters (recall we started with 12,056,684 public keys).

**HEURISTIC 2**: The one-time change address is controlled by the same user as the input addresses; i.e., for any transaction \( t \), the controller of \( \text{inputs}(t) \) also controls the one-time change address \( pk \in \text{outputs}(t) \) (if such an address exists).

Change address: When sending funds from your wallet, the remainder are sent to what is referred as a "change" bitcoin address. If we can identify change addresses, we can therefore potentially cluster not only the input addresses for a transaction (according to Heuristic 1) but also the change address and the input user.

We observed that major users tagged for anywhere from 20% to 40% of the total incoming value. Also, over 80% of the value mining pools receive is as change from themselves.
Although effective, Heuristic 2 is more challenging and significantly less safe than Heuristic 1. (we get false positives) This happens when a change address is used twice within a short period. Also happens when used as a self-change address, which is an advanced feature in some wallets.

This heuristic allowed us to name 1,600 times more addresses than manual Heuristic 1. The users we were able to name, capture an important and active slice of the Bitcoin network.

**Big Services in Bitcoin**

By monitoring the bitcoin flow, the demonstrated centrality of huge services prevents them from staying completely anonymous. For example, the satoshi dice (a huge gambling service) deals with 60% of the overall activity in the bitcoin network. This affects our metrics and the whole bitcoin flow, since as observed, the rise of micro valued transactions can be attributed just to satoshi dice. Also, it is the reason for the rise of immediate spending.

**Illicit activity**

Exchanges have essentially become chokepoints in the Bitcoin economy. To buy into or cash out of Bitcoin at scale, we argue that using an exchange is unavoidable. If a thief steals thousands of bitcoins, it is unavoidably visible within network, and thus the initial address of the thief is known he cannot simply transfer the bitcoins directly from the theft to a known exchange. A common technique for hiding the bitcoin flow is what we call Peeling. A single address begins with a relatively large amount of bitcoins. A smaller amount is then “peeled” off this larger amount, creating a transaction in which a small amount is transferred. This process is repeated until the larger amount is pared down, at which point the amount remaining might be aggregated with other such addresses to again yield a large amount in a single address.
Silk road

Silk Road was an online black market and the first modern darknet market, best known as a platform for selling illegal drugs. It is one of the most well-known and heavily scrutinized addresses in Bitcoin’s history (1DkyBEKt). It received 613,326 BTC in a period of eight months. The address was dissipated, the resulting funds were not sent en masse to any major services.

Tracking Thefts in Bitcoin

Thefts are in fact quite common within Bitcoin: almost every major service has been hacked and had bitcoins stolen, and some have shut down as a result. For each theft the set of transaction was found easily. The movement of the stolen money ranged from quite sophisticated layering and mixing to simple and easy to follow. For the thieves who used the more complex strategies, we saw little opportunity to track the flow of bitcoins (sophisticated strategies to hide the flow of money) but for the thieves that did not, there seemed to be ample opportunity to track the stolen money directly to an exchange.

Everyone had difficulty of cashing out at scale

Concluding

Using exchanges to cash out at scale is inevitable, and thus that Bitcoin does not provide a particularly easy or effective way to transact large volumes of illicitly obtained money. With Heuristics 1 & 2 we can monitor addresses and bitcoind flow. For large services or transactions full decentralizing and anonymity is difficult. There is a growing gap between the potential anonymity available in the Bitcoin protocol design and the actual anonymity that is currently achieved by users. Large bitcoin flow cant get unnoticed, but can get significantly masked.