Cryptocurrencies and the Block Chain

Martin Weber
Johannes Kepler University Linz
napster2202@gmail.com

Abstract. This review article presents the system architecture behind most of nowadays cryptocurrencies. Thereby I will focus on the application scenario of a decentralized cryptocurrency, but also point out the potential for generalization. The architecture was mainly defined by the Bitcoin network, the first and still largest implementation of its kind. I will discuss the components that compile this structure and mention security weaknesses as well as the opportunities for attackers. Furthermore, I outline improvements and detail solutions that had been implemented to cope with the scaling network. Additionally, secondary aspects like privacy and usability are covered.

1 Introduction

With the spread of computing and the internet many things we use for our daily lives have transferred from the physical to the virtual world and many more are expected to do so in the next couple of years. With this transition and an internet connection the physical location of our goods is no longer important for accessing them. Financial transactions are carried out electronically, people and companies store and process their data in the cloud. With all the comfort of handing over these things we also hand over controls of money, documents and other data. All this is built on trust. The society trusts banks to handle their money correctly and third party services to protect data against unauthorized access.

Besides the change from paper to electronic ledgers the way in which transactions are done has not changed much. Today financial transactions between two parties still rely on a number of banks that trust each other in handling this transaction correctly, like back in the days of paper ledgers. The passing on of this transaction record between the trusting banks introduces a lot of work and possible points of failure. Whereas the transaction handling between the banks seems to have gotten out of focus of attackers, the last mile communication between parties and their banks is still an interesting target for attackers.

Because of the costs for each transaction this system does not scale well to transactions with small or no financial values. The need for a system that can be trusted to handle transactions has become omnipresent with the rise of cloud computing. Hitoshi Yamamoto published a decentralized currency system called Bitcoin in 2008. Instead of trusting a small number of globally operating banks, or other service providers, the parties in the Bitcoin system do not need to trust anybody. Each party can cryptographically verify all transactions ever made in

the Bitcoin system. This enables the agreement on a public ledger with no trust between parties.

Whereas Bitcoin was invented as a decentralized currency system, the systems architecture tackles a much more general problem, namely the agreement of untrusted parties to a common public ledger. One key component of this ledger is called block chain which is secured by a cryptographic riddle. When public media refers to "Block Chain Technology" they usually mean the whole transaction-handling-system of Bitcoin, whereas the actual block chain is only one component of that system.

This survey will explain how the technical implementations of cryptocurrencies and the block chain work and it will point out the critical aspects. I will illustrate possible attacks together with the potential impact on the security of transactions. As the block chain was introduced with the Bitcoin system and this is still the most prevalent application, it will be used as an example throughout this work. Additionally, I will also mention applications besides cryptocurrencies and give an outlook over the potential applications for block chain technology in the near future.

2 Cryptocurrencies

There is no clear definition of what properties a potential cryptocurrency has to fulfill. In 2008 Bitcoin was published by Hitoshi Yamamoto, what nowadays is referred to as the first cryptocurrency system. Today Bitcoin is seen as a reference implementation and thereby defines what we call a cryptocurrency today. A cryptocurrency system is a peer-to-peer system where the control over the currency is defined by publicly known mathematical properties and not by a centralized trusted authority. Furthermore, such a system does not assume any premature level of trust between its participants, instead trust for each transaction is guaranteed by a set of mathematical functions. Because trust is ensured, there is no need to reveal the own identity when conducting a transaction.

Addresses in cryptocurrencies are representations of public keys, instead of account numbers in traditional banking. In cryptocurrency systems there is no explicit record of the account balances, rather there is a list of all unspent incoming transactions. With this list each user can calculate his balance. Whereas in traditional banking banks guarantee the spenders solvency, in the cryptocurrency system the spender can prove this cryptographically without a third party.

2.1 Addresses

To create an address for Bitcoin the user first needs to generate a random private key. The Elliptic Curve Digital Signature Algorithm (ECDSA) using curve secp256k1 can then generate a corresponding public key. This public key is then double hashed firstly with SHA256 and secondly with RIPEDMD-160. A byte for version specification is added at the front and a checksum at the end. The whole address is encoded using base58 as shown in figure 1.

Users are strongly recommended to generate a different address for each transaction to protect their anonymity. Although, through the publicly known transaction chain it is possible to trace back transactions and thereby link together certain addresses. Therefore, the addresses do not offer complete anonymity but allow obfuscation of the owner [2].

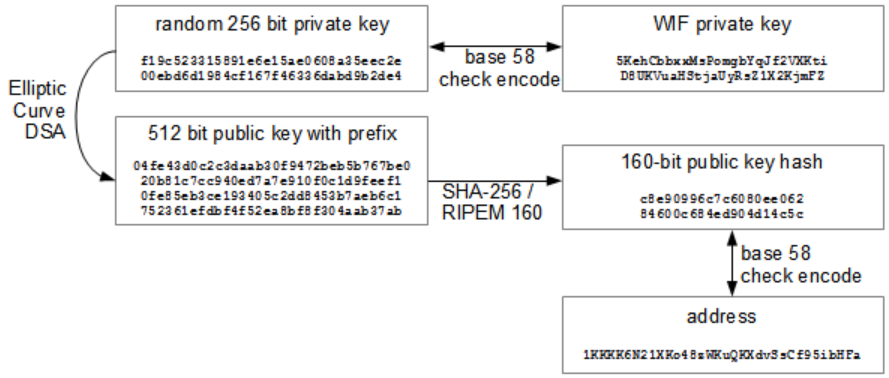

Fig. 1. Address generation process as currently implemented in Bitcoin [9].

2.2 Transactions

Because transactions replace the account balances it is necessary that already spent transactions cannot be spent a second time, and new transactions cannot be created without verifying their origin. This is ensured by linking earnings and spendings together to form a transaction chain. Transactions can have multiple inputs and outputs.

To keep the number of unspent transactions low, all input transaction must be spent entirely. Due to the multiple outputs the spender can simply redirect the change to an address he owns. Because all transactions are known to all members of the network, it can be verified that the spender is the owner of the input transactions. This is done with cryptographic signature that proves that the spender has control over the private key that belongs to the public key in the input transactions.

Each transaction has an attached public key script, which is generated by the sender. This script is versatile and can specify complex recipient-structures. For example, 2 out of 3 receivers need to agree when spending the transactions value. For most transactions it only contains two amounts and addresses, one for the actual receiver and one for the change to an address controlled by the spender. When using a transaction as input the spending transaction must supply a signature for each of its inputs which fulfills the instructions specified in the

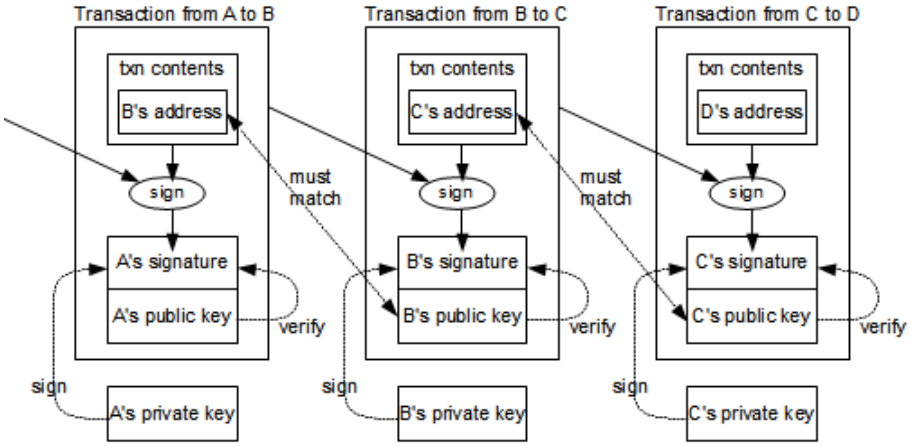

Fig. 2. How transactions are chained and verified with their inputs in Bitcoin [13]

public key scripts. For the most common transactions, like in figure 2, this means including the full public key and a cryptographic signature of the transactions content with the secret key.

3 Block Chain

Because of the decentralized structure and different propagation times through the network, a user would be able to submit two transactions spending the same inputs and there would be no legitimate way of proving which transaction was submitted first. There is the need for a technology that achieves an agreement on a sorted list of transactions across a network of untrusted parties. The technology employed in most cryptocurrencies is called block chain.

3.1 Manner of functioning

A block chain stores arbitrary blocks of data while securing their order by cryptographic methods. Each block of the block chain contains the hash of the previous block, to secure the order and prevent parties from precomputing blocks. The only way to append a block at the end of the block chain is by solving a cryptographic puzzle. The solving process of these cryptographic puzzle is called mining. The solution, which depends on the content of the block, must be part of the block as well. This enables any other party in the network to counter-check the correctness of the suggested solution. The puzzle must also safeguard the block content against manipulation.

The block chain keeps the transactions in their order and therefore prevents double-spending of inputs. A new block is always appended to the end of the chain.

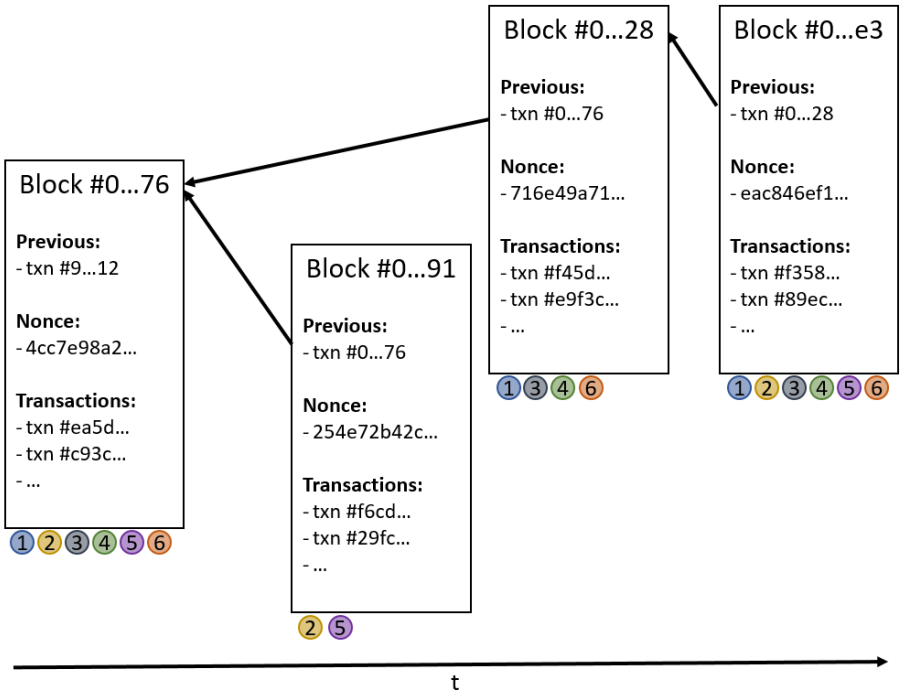

Fig. 3. How the block chain is extended over time. Block 2 and 3 illustrates a fork of the chain, where there is disagreement on the longest branch. As block 4 is solved on the top of block 3, all nodes adapt to the longer branch and the fork is resolved.

When two blocks are submitted at almost the same time, each node in the network keeps mining on top of the block they received first, but they also keep track of the changes in the other chains. As shown in figure 3, at some point in time one chain will be longer than all the others. This will then be considered as the main chain. Because of their own interest miners will mine a block on top of the chain that is more likely to survive, and that is by statistical means the currently longest one [11].

Because parallel running ends are assumed to be temporary, Bitcoin introduced checkpoints which fixate blocks by a hard-coded list after a certain time [1]. This is not relevant for the normal case, but if a problem would invalidate the current block chain there is something like a snapshot, that could be restored without losing all transactions and Bitcoins ever mined.

3.2 Simplified Payment Verification in Bitcoin

To check whether a transaction was included in a block or not, a node would need to download and hash all transactions of the corresponding block. With the growing number of transactions this is a major concern for devices with limited

computational power or network bandwidth, such as mobile phones. For this purpose, Bitcoin introduced a simplified payment verification method with an adaptable privacy parameter.

To ease this process the network computes a merkle-tree for the transactions included in each block. Like illustrated in figure 4, the hash values for each transaction are computed separately at first. Then two hash values are concatenated together and hashed again. This process is carried out in tree fashion until only one root hash value is left. The values of the intermediate nodes are stored additionally.

A client that now wants to check whether a transaction was included only needs to follow the path from the root to the transaction it is interested in. All other branches can be considered by their intermediate values when checking the overall hash value. This allows a check with logarithmic complexity.

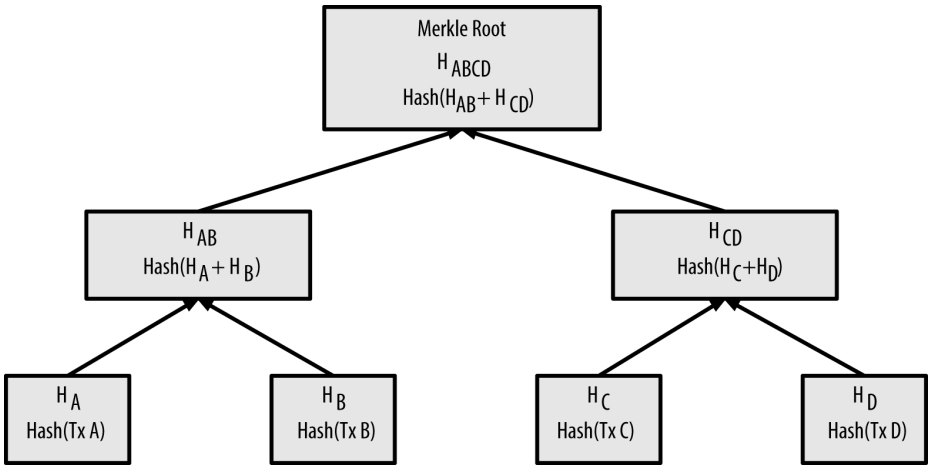

Fig. 4. Merkle-tree structure computed for all the transactions included in a solved block in Bitcoin [14].

The drawback of this method is that other nodes in the network are able to link you to the transaction you are tracing. To defend this, Bitcoin uses bloom filters which are probabilistic data structures that are able to asses if data has already occurred in a stream by using hashes. These filters are commonly used in database or for routing in networks. Bloom filters can have an adjustable false positive rate but a false negative rate of zero. The bloom filter needs to fetch all transactions associated with the addresses of the user and additionals to preserve privacy. Otherwise it is possible to infer the users identity from the transactions retrieved in the verification process. Depending on the users bandwidth and the privacy level the user seeks, he or she can adjust the bloom filter's false positive parameter.

3.3 Coin Selection

Transactions not only serve as a mechanism form transferring money from one entity to another but also handle the transfer of its own transaction reward. The reward is assigned to whichever node solves the block in which the transaction gets included. Blocks have a maximum size, therefore miners will choose to include the transactions which result in their highest reward. In Bitcoin there is a complex fee policy which determines if a transaction requires a fee and its amount. Some transactions are free, which is possible because miners receive an additional reward for each block they solve.

Because of the various factors determining the cost for a transaction, like the value of change, the age of the inputs or the total value of the transaction, selecting the best inputs is a hard problem. Furthermore, the change should either be avoided completely or otherwise have a reasonable size to be used as an input for a later transaction. In fact, the problem is similar to the knapsack problem which is known to be np-hard [6].

3.4 Side Chains

These are block chains which can exchange trading units (proof of work units, like Bitcoins) with other block chains, but cannot instantiate new trading units. Because of its decentralized structure for Bitcoin it is not easy to evolve over time. Each change has to be accepted and realized by each participant. Side chains would allow users to switch technologies when they are ready for. Because side chains cannot create their own units, a fixed, always possible exchange/transfer procedure needs to be implemented and maintained [4].

4 Mining

Mining is essential for securing the block chain, but because of its costs it needs to be incentivized which makes it a possible target for fraud [10]. The goal of mining is solving a cryptographic riddle which allows appending a block to the end of a block chain.

4.1 Essential puzzle properties

- Scalable difficulty:
  The ability to adapt to rising and falling block generation rates in the network.
- Fair distribution of wealth:
  None of the miners should have an advantage through knowledge gained by previously solved blocks. Each miner must have the same likelihood generating the next block, independent of previous solutions.
- Easily verifiable results:
  Verification should be carried out fast, to keep the advantage of the previously winning node (it does not need to check its own block) minimal.

– Sensitivity to tempering block data:
  It needs to prevent attackers from modifying data that is protected by the block chain.

Currently most spent on cryptocurrency puzzles is only for the reason of solving the next block, without gaining any additional real world benefit or knowledge [7].

4.2 The puzzle in Bitcoin

As can be seen in the figure 3 of the block chain, each block consists of multiple properties. The puzzle in Bitcoin is to modify the block in a way that its hash value is below a certain numeric threshold. The modification is done by appending nonce values, which do not carry any information but influence the hash-code. Finding these values can only be done by trial and error (brute force).

In the Bitcoin network a puzzle is solved roughly every 10 minutes. To keep this rate constant despite the increase in computing power, the complexity of the puzzle is adapted weekly. Bitcoin riddles can be solved using standard CPUs and GPUs or specialized hardware like FPGAs and ASICs. As the revenue in Bitcoins for a solved block decreases and solving needs more computational effort over time, energy efficiency is an important issue. FPGAs and ASICs achieve much higher hash-rates per energy than CPUs or GPUs, which makes it hard to compete for individuals [3].

4.3 Useful and ASIC-resistant puzzles

Other cryptocurrencies like Litecoin or Monero have puzzles that use functions which require a significant amount of memory. Litecoin for example uses scrypt. This makes them suitable for running on a CPU or GPU but prevents them from being solved lucratively on nowadays FPGAs and ASICs. These currencies therefore offer individual miners cost revenue ratios that are much closer to those that big mining pools can achieve than in the Bitcoin network [8].

Bonneau et al. [7] reported on ASICs solving scrypt, but because scrypt is designed with a parameterizable memory consumption which directly affects the generated hash values, it is not completely clear if those can achieve superior performance when used to solve blocks for cryptocurrencies using scrypt.

4.4 Mining Pools

Because solving a block for a single miner is very unlikely, miners have formed pools. Mining pools encapsulate multiple machines working on the problem and split the revenue in the case of success. Some mining pools are open to the public whereas others are commercial assemblies.

The communication protocols between miners within a mining pool defer from pool to pool and are not standardized. For public pools it is essential that participants prove that they are actually working on the problem. Miners have to

report either the successful solution, or another proof of work, to the authority of the pool. In Bitcoin this is often the solution which gave the lowest hash-value [5].

Furthermore, there are pools where the authority takes care of the verification of the last block and others where each compute node does the verification on its own. Some mining pools do not even care about verification and therefore risk mining on the top of an invalid block. This already led to the loss of computational work with a value of about $50,000 [15].

As mentioned above, mining pools have authorities that need to be trusted premature, something that cryptocurrencies aim to avoid.

4.5 Discussion on average solving time

Shorter solving time results in more blocks which result in more validation work. Because checking the validity of a block is a constant amount of computational work, miners with high computing resources benefit from higher block rates. Longer solving time implies that it takes longer until a transaction can be seen as surely included in the succeeding chain.

5 Selfish Mining Attack

The Bitcoin system always aims at setting the attacker in a race against the whole net rather than a single other party. This works well in a network of many small parties but has its weaknesses when a few number of parties have control over large ratios of the computational power. Often majority is considered as the critical size of a pool that threatens security in the network, but Eyal and Sirer [10] have shown that this threshold is considerably lower.

Selfish mining is a strategy that allows a pool of a sufficient size to obtain a revenue larger than its ratio of mining power. The strategy forces the honest miners into performing wasted computations. Selfish miners achieve this by not revealing their mined blocks immediately but with some delay. This extra time increases the probability of finding the next block for the pool of selfish miners. Because the selfish miners only control a relatively small part of the computational power, their hidden branch will not remain leading infinitely long. Therefore, the selfish miners will eventually publish their hidden blocks with some delay.

5.1 The Selfish Mining Algorithm

With this intuition, we can illustrate the operations of the strategy through sample scenarios involving the different public and private chain lengths. Whenever the public branch is longer than the selfish branch, the selfish miners adapt to the public one.

When the selfish miners solve a new block they keep their block private. Now there are two possibilities, either the selfish miners are able to get a second block ahead of the public chain, or more likely the public chain catches up.

In the first case where the selfish miners get aware of a new block on the public chain, they immediately publish their block and fork the chain. Because of different propagation times through the network some part of the honest miners will receive the selfish mined block first, and therefore mine on top of that chain. If the next block is solved on top of the honest chain the selfish miners work is invalidated, but if it is solved on top of the selfish chain the selfish pool takes the profit from the lately released block.

In the other cases where the selfish miners find the next solution, they develop a beneficial lead of two blocks. With this benefit the selfish miners continue mining on the top of their chain. For every block the honest miners find, the selfish miners publish one of their chain. Because the selfish miners are a minority, their lead will eventually shrink to one leading block. When this happens the selfish pool immediately publishes their private branch. Since this branch is still one block longer, the honest miners will adapt to that and the selfish miners collect the revenue of all their blocks.

5.2 Boundaries

For enjoying a guaranteed beneficial revenue per effort ratio through the selfish mining strategy in the Bitcoin network, the selfish mining-pool needs to have control of at least 1/3 of the computational power. This is due to the expected value how often the honest miners will invalidate the hidden chain of the selfish miners compared to how likely it is that the selfish miners take the lead with 2 or more blocks. Additionally, the number of controlled nodes is important when the selfish miners need to compete with the honest miners in broadcasting their blocks because the honest miners have caught up.

5.3 Threat

The selfish mining strategy does not prevent others from submitting transactions, nor can it be used for double spending attacks, but it gives one party a higher revenue per computational effort.

Yet, it needs to be stated that a pool of this size has invested considerable money in hardware and takes a high risk for its own when attacking the Bitcoin system.

6 Other Block Chain Applications

6.1 Block Chain based Accounting System

Researchers and the Linux foundation are working together with private companies like R3 and IBM on a block chain based accounting system. The block chain architecture could enable a decentralized storage witch would increase the difficulty for subsequent manipulation and result in higher reliability [12].

6.2 Decentralizing Privacy

The work of the researchers Zyskind et al. [16] addresses a use case beyond the financial space. They model an example how a block chain can be turned into an access-control manager for personal data in an untrusted environment. The authors also give an idea how block chains can be used to tackle trusted computing problems in society. The technology would also allow the implementation of laws and regulation concerning data processing. The block chain can be used as a legal evidence because of being computationally tamper-proof. Through the protection with the secret key it explicitly allows the owner to precisely choose which parts of the stored information he wants to reveal.

7 Conclusion

As currently Fintech is a booming market and many startups have no smaller target than revolutionizing the banking sector, there is definitely more to come. The technology of cryptocurrencies assures a level of security that the current, much slower and costly system of banks is not able to guarantee.

Big mining pools certainly pose a threat to the democracy in such systems, but compared to the influence governments have on the current financial systems, that does not seem that frightening anymore.

In the developed world banks supply a quite secure way of monetary transfer, but when thinking of developing countries controlled by corrupt governments, totalitarian regimes and with no register of residents, cryptocurrencies have some very appealing benefits.

On the contrary though, in case of somebody being threatened to transfer money to the threatener's account, there is no way to either seize the money or rollback the transaction without the corresponding private key. Additionally, if something goes wrong there is nobody to blame or sue. Another drawback is that because of its high degree of privacy it can be a good tool to hide any kind of illegal financing, like bribery or funds for terrorism.

References

1. Alqassem, I., Svetinovic, D.: Towards reference architecture for cryptocurrencies: Bitcoin architectural analysis. In: IEEE International Conference on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing. pp. 436–443 (2014)
2. Androulaki, E., Karame, G., Roeschlin, M., Scherer, T., Capkun, S.: Evaluating user privacy in bitcoin. In: Financial Cryptography and Data Security, Lecture Notes in Computer Science, vol. 7859, pp. 34–51 (2013)
3. Anish Dev, J.: Bitcoin mining acceleration and performance quantification. In: IEEE 27th Canadian Conference on Electrical and Computer Engineering. pp. 1–6 (2014)
4. Back, A., Corallo, M., Dashjr, L., Friedenbach, M., Maxwell, G., Miller, A., Poelstra, A., Timón, J., Wuille, P.: Enabling blockchain innovations with pegged sidechains. https://blockstream.com/sidechains.pdf (Accessed: 2015-10-30)

5. Beikverdi, A., Song, J.: Trend of centralization in bitcoin's distributed network. In: 16th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. pp. 1–6 (2015)
6. Bonneau: https://freedom-to-tinker.com/blog/jbonneau/bitcoin-mining-is-np-hard accessed on 2016-01-10
7. Bonneau, J., Miller, A., Clark, J., Narayanan, A., Kroll, J., Felten, E.: Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. In: IEEE Symposium on Security and Privacy. pp. 104–121 (2015)
8. Coindesk: http://www.coindesk.com/information/how-to-mine-litecoin accessed on 2016-01-10
9. Eckel: https://bhelx.simst.im/articles/generating-bitcoin-keys-from-scratch-with-ruby accessed on 2016-01-10
10. Eyal, I., Sirer, E.: Majority is not enough: Bitcoin mining is vulnerable. In: Financial Cryptography and Data Security, Lecture Notes in Computer Science, vol. 8437, pp. 436–454 (2014)
11. Garay, J., Kiayias, A., Leonardos, N.: The bitcoin backbone protocol: Analysis and applications. In: Advances in Cryptology - EUROCRYPT, Lecture Notes in Computer Science, vol. 9057, pp. 281–310 (2015)
12. Grüner: http://www.golem.de/news/finanzindustrie-linux-foundation-startet-blockchain-projekt-1512-118103.html accessed on 2016-01-10
13. Shirriff: http://www.righto.com/2014/02/bitcoins-hard-way-using-raw-bitcoin.html accessed on 2016-01-10
14. StackExchange, B.: http://bitcoin.stackexchange.com/questions/10479/what-is-the-merkle-root accessed on 2016-01-10
15. Wiki, B.: https://en.bitcoin.it/wiki/Comparison_of_mining_pools accessed on 2016-01-10
16. Zyskind, G., Nathan, O., Pentland, A.: Decentralizing privacy: Using blockchain to protect personal data. In: IEEE Security and Privacy Workshops. pp. 180–184 (2015)