Blockchain

an introduction

Research paper

A. Shanti Bruyn

August 26, 2017

supervisor: Rob van der Mei

Contents

Management summary iii

Introduction iv

1 Introduction to blockchain technology 1
1.1 What is blockchain? . . . 1
1.1.1 What problem does blockchain solve? . . . 4
1.2 History of blockchain technology . . . 5
1.3 Challenges of blockchain . . . 5

2 Building blocks of blockchain 8
2.1 The Database . . . 8
2.2 A Block . . . 10
2.3 The Hash . . . 10
2.4 A Miner/Node . . . 12
2.5 A Transaction . . . 13
2.6 A Fork . . . 14
2.7 Blockchain Safety . . . 14
2.8 Incentive . . . 15

3 Literature survey 16
3.1 Applications of blockchain . . . 16
3.1.1 Bitcoin . . . 17
3.1.2 Bitnation . . . 18
3.1.3 Energy reserve supply market . . . 18

4 Variable exploration 19
4.1 Nakamoto Consensus . . . 19
4.2 Indicators/Metrics . . . 20
4.3 How is everything connected? . . . 22
4.4 Example Calculation . . . 25
4.4.1 The given parameters . . . 25
4.4.1.1 Calculated given parameter . . . 25
4.4.2 Chosen parameters . . . 26
4.4.3 Indirectly influenced parameters . . . 26
4.4.3.1 Inter nodes time . . . 27
4.4.3.2 $\mathbb{P}(\text{fork})$ . . . 28

5 Relations between variables 30
5.1 Blocksize . . . 31
5.2 Headersize . . . 33
5.3 Transaction size . . . 34

5.4 Mining power node A 36
5.5 Difficulty Cryptopuzzle 38

6 Discussion/further research **40**

7 Conclusion **41**

A Glossary

Chapter

Management summary

A summary concerning an introduction into blockchain should start by stating that it is complex and combines many new concepts. It really is a totally different game compared to what is currently known. Therefor one cannot explain it briefly nor clarify blockchain by comparing it to something familiar. To some extent too, the pro's, con's, risks and opportunities are still to be found.
Having said that, an effort will be made to give a general peek into the world of blockchain in this summary.

Blockchain is a new type of database. The reason why there is such a call for this new type of database is because it solves the previously unsolvable *double spending problem* without a middleman, opening up a range of new possibilities.
In this database the data is saved in a block, Which in turn is linked to other blocks in a chain creating the blockchain. To secure the blockchain a system called proof-of-work is used. In short this means there is so much work (i.e. processing power) needed to find a block, it is virtually impossible to alter the blockchain afterwards. This work is done by so called miners who -when they find a block- get a small payment for their effort.

Blockchain does have some important aspects to keep in mind. For instance what is saved in blockchain can never be removed or altered. Depending on the cause, this can either be a major advantage or disadvantage.
Blockchain can even be damaging to the environment because the security system used demands extreme amounts of energy.
These two aspects are mentioned only as examples of possible considerations when wanting to use blockchain. In the paper itself many more are presented.

To further explain the workings of blockchain, this paper focusses on the Nakamoto blockchain, the original and the most commonly known for its use in Bitcoin.

Chapter

Introduction

The objective of this paper is a general exploration of blockchain, with the major questions to be answered being:

- What is blockchain?
- How does it work?
- What are it's possibilities?
- Finally, and more precisely: how do the different variables within blockchain work together?

In order to be able to get to answer all of these questions first (and for most) this paper will be a survey of what other people have already written on blockchain, research done on blockchain or even done/achieved with blockchain. For the last question this paper will try to give a concise overview of the different variables that all have something to do with blockchain and how they interact with each other. And it will give some insights into how the creator of the blockchain can influence certain aspects of the blockchain.

There have been rapid developments within blockchain and this is still ongoing. There are already many different types of blockchain that surfaced, each slightly different from one another in many different respects. This makes it impossible to cover the whole range in unambiguous terms and without losing overview. For this reason this paper focusses on the first blockchain, the Nakamoto blockchain which is the most well known blockchain due to its connection to Bitcoin.

Even with this restriction, there are many new terms and concepts which are all intertwined with each other. As these can not all be clarified at the beginning of this paper a Glossary is provided in an Annex.

Chapter 1

Introduction to blockchain technology

Lately there is a lot to do about a new sort of technology called 'Blockchain'. This chapter will give an explanation of the concept. Starting in section 1.1 with a general idea of Blockchain. In chapter 2 Blockchain will be 'dissected' and each section will be thoroughly analyzed. For a brief background on the technological origins see section 1.2. A lot of uses have already been found for Blockchain, in section 3.1 some of these will be discussed. To conclude this chapter some challenges Blockchain faces will be discussed in section 1.3, giving a complete picture of Blockchain.

1.1 What is blockchain?

A blockchain is a decentralized, distributed database that is used to maintain a continuously growing list of records, called blocks. Each block contains a timestamp and a link to a previous block. By design and by purpose blockchains are inherently resistant to modification of the data. Functionally, a blockchain can serve as 'an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way.' [9]

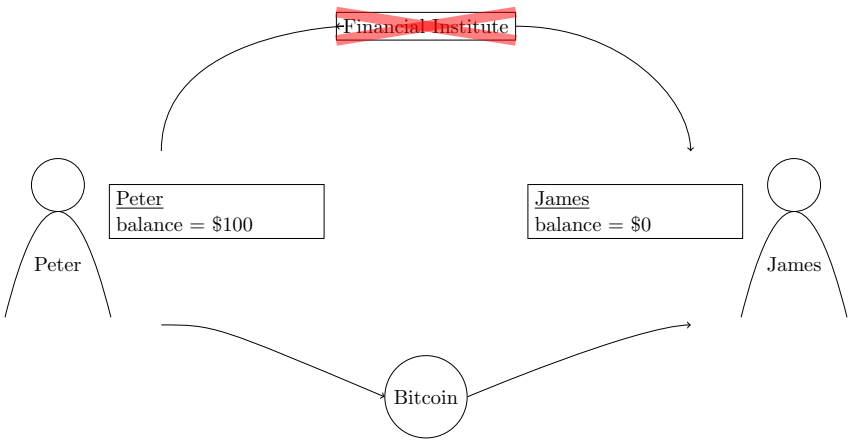


Figure 1.1: With blockchain a centralized third party is no longer needed

What does this mean and how does this work?
Well, lets take a crypto valuta like Bitcoin (for more information on Bitcoin see subsection 3.1.1). The crypto valuta is built on blockchain so it doesn't need a third 'authority' (such as a bank) anymore.

Running Example 1

Peter wants to give $100 to James (see Figure 1.1).

Without blockchain
Peter would send his bank a request to send $100 of his account to his friends' acount. The bank would check a few things like whether Peter actually has the $100. If everything checks out the bank will send Peters $100 to James' account.

With blockchain
Peter creates a transaction of $100 to James and sends this transaction over the internet. This transaction is included in a block. All miners check whether this is a valid transaction. If it is, James has the $100 of Peter.

The Running Example on page 2 illustrates the use of blockchain. Instead of the transaction being checked by a third authority like a bank, it is being 'checked' by everyone who takes part in the system and everyone who will join the system in the future. Thus releasing the need of a centralized third party, this has several advantages like less transfer costs (after all, there are no more man hours needed to check everything). It has the potential of being more anonymous while making it both easier to pay globally and nearly impossible to 'reverse' transaction (which the third 'trusted' party could decide to do).
All of this while maintaining the promise of the same certainty of getting your money as one would get from a Financial Institute.

Because blockchain is decentralized and distributed all current nodes and all future nodes to come can check whether every transaction follows some given rules. This makes sure someone can't promise money to two people at the same time.
Removing the need of a third authority who has the monopoly on all the information and can make decisions which are very hard to check.

How can blockchain achieve this certainty of one getting their promised money, without a third party which checks this?
To answer this: let's continue the Running Example on page 2.

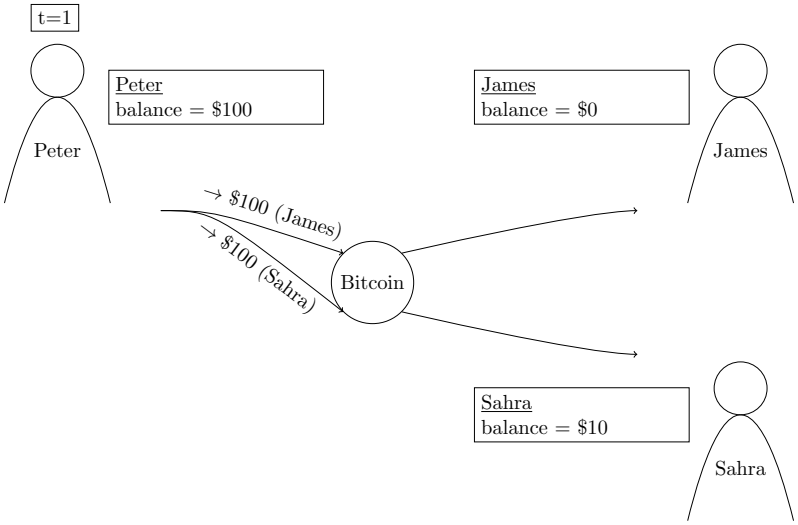


Figure 1.2: Peter promised his $100 to both James and Sahra!

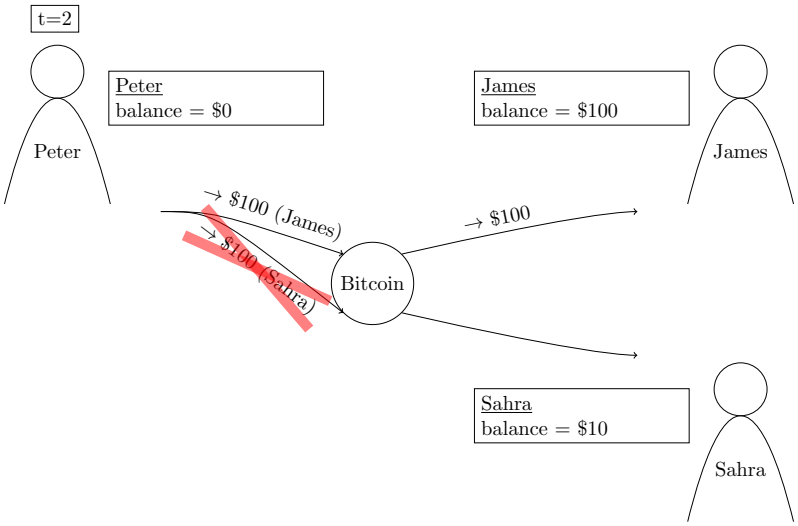


Figure 1.3: Blockchain checks this, and only the first promised, will receive the money

Running Example 2

In Figure 1.2 Peter promised both James and Sahra his $100!

Since both promises are done in between the same two blocks, arbitrarily one of them is chosen to be included into the next block (say block number 433). This happens to be the promise toward James.
Now when block 434 is created (at time t=2 Figure 1.3) every node verifies that Peter does not have the required $100 anymore to give to Sahra. And thus Peters promise towards Sahra is not kept.
When James checks his balance, he sees that the promis Peter made to him has been kept.

This example shows that the order of blocks in the chain conclusively determines the order in which the transactions take place.

Blockchain is named that way because it is basically an endless chain of blocks. The order of this chain defines the order in which the transaction in the blocks took place (as shown by Running Example on page 4).
In Figure 1.4 an example of such a chain is given. The different colors indicate different 'types' of blocks. The chain 'grows' so to speak from bottom up. The first block is a special block, because it is the only block in the whole chain which has no preceding block. That's why it is the only one colored green.
The black blocks are the 'normal' blocks. These are the blocks which form the longest (and thus official) chain. The purple blocks form so called 'forks'. These form when two blocks are found at exactly the same moment. For a short moment in time there are two chains of equal length. Until for one of the chains a new block is found quicker than for the other chain it is undecided which of the two is the 'official' chain.

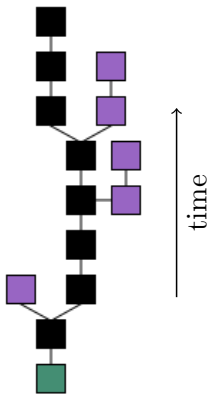


Figure 1.4: An example of blocks in a chain [1]

1.1.1 What problem does blockchain solve?

The creator of blockchain created blockchain with a very specific problem in mind to solve: the *double spending problem.*

> "What is needed is an electronic payment system based on cryptograhpic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. Transactions that are computationally impractical to reverse would protect sellers from fraud, and routine escrow mechanisms could easily be implemented to protect buyers. In this paper, we propose a solution to the double-spending problem using a peer-to-peer distributed timestamp server to generate computational proof of the chronological order of transactions."
>
> — Satoshi Nakamoto [6]

[1] Figure obtained from [9]

In other words: a growing problem in an ever more globalizing world is trust. People do business with people they have never met - and probably will never even meet face to face. How can you be sure the other person will pay you what he/she promised? And if he/she doesn't, is there anything you could do against this? Or are you just gambling all your merchant wares?
All sorts of authorities (like Pay-pall) have been brought into life to help solve these type of problems. But they do have the power to for instance reverse the payments made.

Blockchain solves all of these problems. Because the database is distributed, it is extremely transparent: everyone on this planet could check for him- or herself if all the rules have been followed until now.
There is no longer one party that can decide to reverse a payment: once a payment is done, it's done.
Added bonus is that the blocks can't be changed anymore because of the proof-of-work (section 2.3).

1.2 History of blockchain technology

Blockchain seemingly came up out of no where together with Bitcoin in 2013. Ever since it has been of interest to an increasing number of people. Currently a momentum around blockchain has been formed now the 'big four' are investing in it [2]. Chances are blockchain is going to be of growing importance in the future. Dubai is even planning on being "the first blockchain powered government in the world by 2020" [22].

What path did blockchain follow before it so spectacularly appeared to the wider public?
Both blockchain and Bitcoin are a creation of 'Satoshi Nakamoto'. Until now it is unclear who this is, it could theoretically even be a group of people. He himself claimed to be a man living in Japan, born on 5 April 1975. However, there is still some doubt and quite some names have already passed as possible real identities [5].
In Nakamoto's paper '*Bitcoin: A Peer-to-Peer Electronic Cash System*' from 2008 he introduces Bitcoin to the world and explains how it works [6].
As with most inventions he used and combined many already present theories / techniques. Especially his encryption methods have been around for a while. For instance the way blockchain works with public and private keys stems from a paper from 1980 by R.C. Merkle "*Protocols for public key cryptosystems*". A lot of the cryptology, and techniques that make blockchain so secure date from the 90-ies, as can be deduced from Nakamoto's literature list [6].
According to some the only new part that sets blockchain apart is that every transaction is being hashed and carefully 'braided' together with every new transaction [17].

1.3 Challenges of blockchain

Despite it's many useful properties, there is a need to make some cautionary remarks on Blockchain.
First of all 'There is a tradeoff between performance and security with blockchain: faster blocks mean more forks mean less security." [15]
A relatively easy Example on page 6 already shows that faster blocks indeed mean more forks. But why do more forks mean less security? To answer this question take a look again at Figure 1.4. As said before all the purple blocks are forks: whenever a fork comes into existence this is not known straight away. Every node just thinks they got the last block and continue

[2] See http://www.coindesk.com/big-four-accounting-firms-meet-to-weigh-benefits-of-blockchain-consortium/

mining. Only when they receive a new block where the previous block is not the block they were working on will they know there was a fork. This means that for as long as it is unresolved which part of the fork will be the 'final' (black) chain, it is still not certain in what order the transactions will have taken place officially.

Example 3

Let's say there are two blockchains: chain A and chain B, that find a new block every 10 minutes, 15 seconds resp.
A fork occurs when two blocks are found within lets say a second. Furthermore the assumption is made that the distribution of the time between two blocks found is exponential. Some simple math will show that chain B has many more forks:

Chain A
X = *the number of blocks found in one second*
$X \sim Poisson\left(\frac{1}{10 \times 60}\right)$

$$\begin{aligned}\mathbb{P}(X > 1) &= 1 - \mathbb{P}(X = 0 \text{ or } X = 1) \\ &= 1 - \Big(\mathbb{P}(X = 0) + \mathbb{P}(X = 1)\Big) \\ &\approx 1 - \Big(0,999998613\Big) \\ &\approx 0,00000138735\end{aligned}$$

Chain B
Y = *the number of blocks found in one second*
$Y \sim Poisson\left(\frac{1}{15}\right)$

$$\begin{aligned}\mathbb{P}(Y > 1) &= 1 - \mathbb{P}(Y = 0 \text{ or } X = 1) \\ &= 1 - \Big(\mathbb{P}(Y = 0) + \mathbb{P}(Y = 1)\Big) \\ &\approx 1 - \Big(0,997874117\Big) \\ &\approx 0,002125883\end{aligned}$$

For both chains the probability of a fork is very small. However chain B has rougly 1532 times more probability of a fork than chain A.

Another issue with Blockchain is the block size. For instance with Bitcoin the "block size is 1 MB, yielding only 1 to 3.5 transactions per second (...) for typical transaction sizes" [15]. Combining these two points (i.e. probability of a fork and blocksize) creates a situation where there is a certain ceiling which any Blockchain will eventually reach given the makers do want to keep the chain safe.
This may result in possibly unwanted effects, like for instance with Bitcoin. In Bitcoin it is

probably a matter of time before this ceiling, combined with the miners preference for higher income, will result in rising transaction costs.

Another completely different point of concern is the growing size of the Blockchain (in Bitcoins case the chain has outgrown the 2 TB already). Because of this a growing number of computers don't have the space needed to store the complete chain and thus can't check the chain for themselves or think it takes to long to check it completely. A solution to this problem has already been found: there are servers which do store and check the complete chain and tell others whether or not the chain is valid. However, this results (again) in a more centralized system. This soulution undermines 'decentralization', one of the core principles of blockchain, and can thus hardly be seens as a satisfactory solution.

A completely different type of challenge is that blockchain, to remain safe, has no 'back door'. This means that for instance if someone loses her private key, there is no way she can reach her Bitcoins again. Essentially she just lost her wallet with all the money in it. But this money can never be 'found' to be used again. From now on it will be in the system, but never used again.

The blockchain is vulnerable to so called 'selfish mining attacks' of size up to almost $\frac{1}{3}$ [15]. This means as much as that if one entity owns $\frac{1}{3}$ of all the mining power or more, it could alter the blockchain as it sees fit. It would still need to obey the blockchain rules, for instance with Bitcoin no more than 21 million Bitcoins can exist. So even a selfish mining attack, cannot 'print' more Bitcoins out of nowhere.

Due to the proof-of-work A LOT of mining power is needed, and with that a lot of energy, to keep the blockchain up and running. This implies serious pressure on the environment and a link with many problems surrounding the environment, like global warming.

Another issue that arises due to the proof-of-work, is that the blockchain is very susceptible to sudden mining power changes. If the mining power drops suddenly, it will take ages before the next block is found. On the other hand, if the mining power rises all of a sudden, blocks will be found much faster and thus many more forks will occur.

Chapter 2

Building blocks of blockchain

Until now blockchain has been explained on a very general level, looking from a birdsview. With some examples as to how blockchain can be used and why one would want to. Some of these parts were a bit of an over-simplification in order to help place everything in the right perspective. In this chapter the birdsview will be dropped. Instead blockchain will be looked at from nearby: each separate piece will be dissected and examined closely. With the hope to be able to better understand the general oversight given in chapter 1.

2.1 The Database

Blockchain is not a 'normal' database. Meaning that it does not exist of tables with rows and columns. Rather it exists of a ledger of past transactions.
How this database works is easiest explained via an example with valuta:

- In the first block there is one transaction: some person1 gets $100.

Block 1
$100 → person1

- In the second block person1 gives $20 to person2.

Block 2	
$20	person1 → person2

  Now all the miners check whether person1 has $20. Looking back at block 1, they all see person1 got $100 so he has the $20.

- Now the next block person1 gives $50 to person3 and person2 gives $10 to person4.

Block 3	
$50	person1 → person3
$10	person2 → person4

  Now all the miners check whether person 1 has $50: combining block 1 and 2 one can see that person1 still has $100−$20=$80.
  The second check: whether person2 has $10 is answered by block2 alone: he has.

So instead of keeping track of a persons account balance in a neat table

person	**balance**
person1	$30
person2	$10
person3	$50
person4	$10

A Blockchain keeps track of the transactions and 'recalculates the balance' whenever necessary. To be more precise, a blockchain doesn't 'recalculate' the balance, it rather 'spends transactions'. For this means it is necesary to keep track of 'already spent' transactions (expired) and those which haven't been spent yet (active). Looking at the example above.

- Transaction 1 is active. Furthermore, no transaction was needed for this transaction to take place.

Block 1

transaction(s)			
used	new	$	from → to
-	1	$100	→ person1

- This block, transaction 1 is 'used' in order to be able to pay person2. Payment is made in transaction 2. The 'change' of $80 is set in a transaction (transaction 3) back to person1.

Block 2

transaction(s)			
used	new	$	from → to
1	2	$20	person1 → person2
	3	$80	person1 → person1

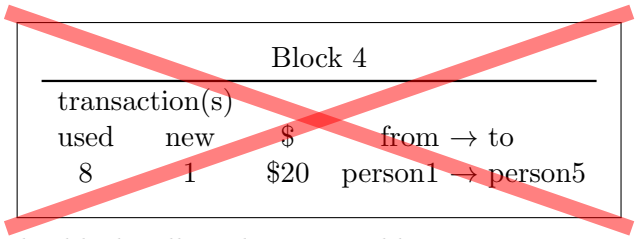
Now all the miners check that transaction 1 was not used before and thus is still active. And all of the $100 is spent.

- Now the next block person1 gives $50 to person3 and person2 gives $10 to person4.

Block 3

transaction(s)			
used	new	$	from → to
3	4	$50	person1 → person3
	5	$30	person1 → person1
2	6	$10	person2 → person4
	7	$10	person2 → person2

Both transaction 2 and 3 are now used in new transactions and can't be used again. Meaning that if a block were to come up using an already used transaction it would be disregarded and won't make it to the chain.



This block will not be accepted because transaction 1 has already been used. Even though person1 does still have $20 ($30 in fact), which can be seen in block 3 transaction 5.

2.2 A Block

Every block in a blockchain consists of the same components (see Figure 2.1):

- A block number
- The hash of the previous block (via this means the 'chain' is being formed)
- Nonce, a random number, see below for more information
- Data: the transactions
- Timestamp with the time the block is created / found
- The hash of the current block

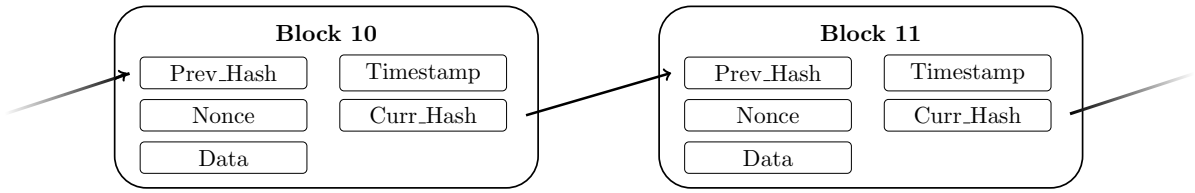


Figure 2.1: Two subsequent blocks in blockchain with their attributes

In other words, every block has a block number which, combined with the previous hash and the current hash, determine the order in which the transactions took place. The accompanying timestamp determine the time at which the transactions are recorded to have taken place. The Data contains the transactions, which could consist of nearly everything, not solely valuta. Last, but not least, there is the Nonce. This is used for the proof of work, and will be described in more detail in section 2.3.
There is one exception to this: the first block. It is the only block which differs slightly from all the other blocks. This makes it very special. It is the only block with no previous block, and thus no previous hash.

2.3 The Hash

The hash is the difficult mathematical problem (or the 'proof-of-work') the miner has to solve in order to 'find' a block. This goes as follows:

1. A random number is guessed (this is called the 'Nonce'[10])
2. The Nonce is added at the end of all the data in the block

3. This is all hashed according to SHA256 method [1] [11]

4. If this hash starts with a predetermined number of zero's a new block is found. If not, the miner has to start again at step 1 with guessing another Nonce. [21]

Example 4

Node number 825 is trying to find a new block, it is trying to create an accepted hash for the following block:

Block 598
Nonce: 1 *Transactions*

An accepted hash for this blockchain starts with 4 zero's.
The current hash would be
"bac6d67daf63c7a06bab569adeadab130d332ed4c870da314d87f6f1b4c8a409"
So this would not be an accepted hash. The node now tries Nonce 2:

Block 598
Nonce: 2 *Transactions*

Resulting in hash:
"6ee6b0c4aa8e6aaa369dfaee1379a59cec8797e7fc8ad1a358d57e1a87a1466d"
Not good either...

The node continues the trial and error until it finds a good hash, at Nonce 11316:

Block 598
Nonce: 11316 *Transactions*

Results in:
"000015783b764259d382017d91a36d206d0600e2cbb3567748f46a33fe9297cf"

Now it is said that the node found block 598.

This is the most 'expensive' part of the blockchain: the energy required to look for a block. This is the part which solves the 'double spending problem' and the part which makes the trust -which was always necessary before- unnecessary. Because of this, as long as 50% [2] or more of

[1] For this research it suffices to note that a SHA256 hash always has the same length, regardless the amount of data hashed. And that it is a top-notch hashing type, for now there is only one type which might be considered 'stronger'

[2] There is some disagreement about the exact number, some say the blockchain is safe as long as attackers have less than 51%, some say the blockchain is only safe as long as attackers have $\frac{1}{3}$ of the total mining power or less.

the nodes are 'honest' (meaning: not trying to harm the system), the blockchain is safe.

2.4 A Miner/Node

A CPU that tries to solve a difficult math problem in order to be able to find a new block is called a miner or a node. The process of finding a new block is as follows.
The steps of how a blockchain operates in a network as given by Satoshi Nakamoto [5, 6]:

1. New transactions are broadcast 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.

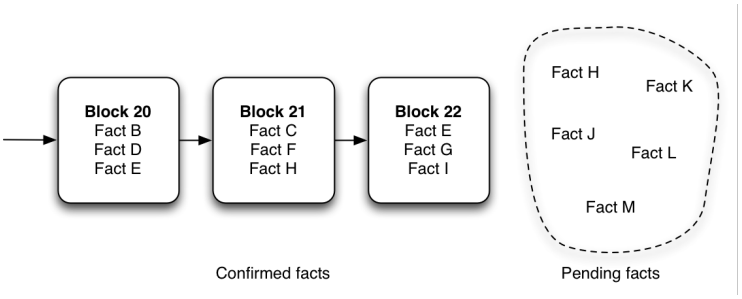


Figure 2.2: The new transaction of step 1 is added to the Pending facts [21]

If there are too many pending facts to fit into one block, the miner chooses the pending facts (Figure 2.2) it wants to include in the block. In the case of a crypto-valuta like Bitcoin, miners often choose for the transactions which yield the highest transaction costs since the miner who finds the block 'earns' all the transaction costs in that block.

From the '*pending facts*' in Figure 2.2 the nodes take the transactions which have to come into the next block. If there are to many pending facts, a selection will be made. This selection can be different depending on the node as shown in Figure 2.3.

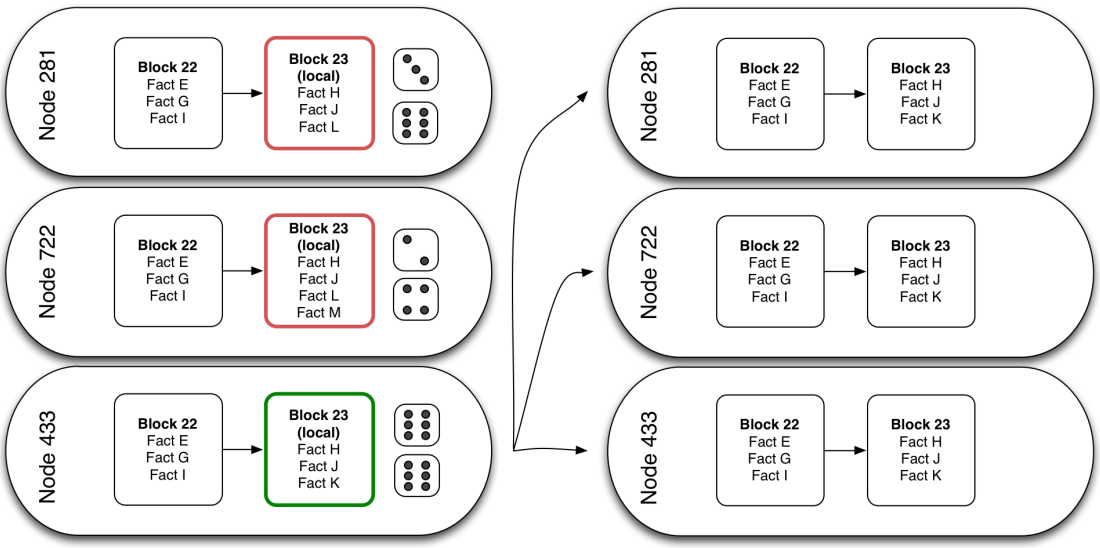


Figure 2.3: All nodes try to find the new block, in this case node 433 managed! [21]

The node that first completes the 'proof-of-work' (in Figure 2.3 throwing double six with two dices) is said to have 'found' the block. It then continues to broadcast this block to all other nodes. They then continue working on the next block.

2.5 A Transaction

A transaction in Blockchain certainly can include, but is not limited to valuta. In case of a bit-valuta it is limited to that. However for instance Etherium is a 'crypto-equity', "meaning that Ethers are completely programmable. One needs Ethers in case of smart contracts and applications." Applications like for instance ownershiprights and voting rights. [7] "Blockchain-based smart contracts are contracts that can be partially or fully executed or enforced without human interaction. (...) Some Blockchain implementations could enable the coding of contracts that wil execute when specified conditions are met." [9].

With regard to crypto-valuta, a big difference with current banking is that the blockchain will not keep track of the balance. Instead, it keeps track of all the transactions and whether or not they have already been 'spent'. So at any given time the balance can be recalculated, but it is in itself not stored.

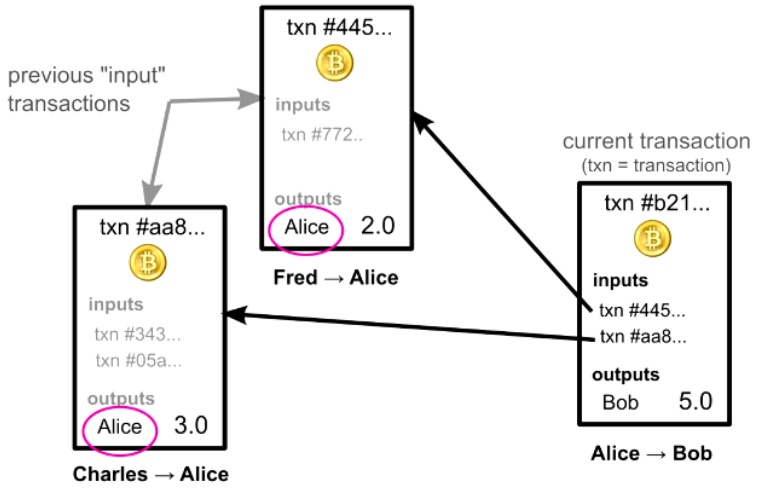


Figure 2.4: Paying with transactions instead of with balance [3]

Example 5

In Figure 2.4 Alice wants to pay Bob 5 Bitcoin. The blockchain decides she can do this because the two transactions 'txn #445...' and 'txn #aa8...' have not yet been spent and together make the needed 5 Bitcoin.

2.6 A Fork

A fork is formed when two miners find a block at the same time. Both blocks will be destributed in the network and all miners will continue working on the block they received first. Until the next block is found, there are 2 'last blocks' and it hasn't been decided yet which one of these blocks will eventually be included in the chain and which one will be discarded. An example of such forks can be seen in Figure 1.4: the purple blocks are forks which didn't make the cut. Meaning that the next block was found quicker for the counter-part block. Which is in this figure black, because it is now part of the official chain.

2.7 Blockchain Safety

One of the reasons blockchain is considered so safe is because it is completely transparent. Anyone who has a computer and access to the internet can check the chain [4] and see for themselves whether everything up till now is legitimate. Another reason is because of the *proof-of-work* concept blockchain works with (see section 2.3). Meaning that if one wants to change a former

[3] Figure obtained from https://www.slideshare.net/tsasaa12/bitcoin-cryptocurrency-48132020 slide 19/65

[4] There are several different types of chains, a company could for instance decide that it only allows miners from within the company. For this research paper one assumes a completely open chain accessible by everyone, like Bitcoin

block, he or she has to repeat the proof-of-work for all the coming blocks as well. Making it as good as impossible to alter the blockchain. The proof-of-work concept also solves the problem of one-IP-address-one-vote problem, since in this system a vote can be seen as 'working on a block' it is one-CPU-one-vote.

Next to these safety-perks because of the setup, there are some additional 'safety-perks' because blockchain uses top-notch hashing and encryption technologies. Making it nearly impossible to figure out things that are supposed to stay anonymous, like finding the person behind an anonymous transaction (hence it is still uncertain who 'Satoshi Nakamoto' is - see section 1.2).

2.8 Incentive

According to some the new aspect to blockchain is that mr. Nakamoto inserted an incentive 'for nodes to support the network'. As 'the first transaction in a block is a special transaction that starts a new coin owned by the creator of the block' [6].

This incentive works as follows: the hashing is the part of blockchain that actually costs energy. It is simultaneously the part that keeps the blockchain safe. But since it costs energy - and thus money - without the incentive there one would not want to help keep the blockchain safe. However, because of this incentive one gets payed to do this work. The account belonging to the node who finds the next block, gets a payment big enough to flip the incentive so that people want to mine blocks.
For instance with Bitcoin, the account linked to the node that found the block gets all the transaction costs of that block, and -until the total number of Bitcoins reaches 21 million- some Bitcoin that is newly introduced into the blockchain with the creation of a new block. Making it a lucrative business.

> "The incentive may help encourage nodes to stay honest. 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. He ought to find it more profitable to play by the rules, such rules that favour him with more new coins than everyone else combined, than to undermine the system and the validity of his own wealth."
>
> — Satoshi Nakamoto [6]

This seems to be what makes blockchain so much more resilient to people with bad intentions. Nakamoto found a way for the people with a bad intention, that it would be more profitable for them to keep the system clean.

Before blockchain it was required that all participants were known. Because of the proof-of-work this is no longer required, and the proof-of-work only works because of this incentive.

Chapter 3

Literature survey

Since blockchain is such a hot topic right now, a lot of people have written about it, done research or tried things out. This chapter will try to give a small view into some of the aspects other people looked into.

3.1 Applications of blockchain

Since blockchain made its appearence with Bitcoin (see subsection 3.1.1) many new uses have already been found. Way to many for all of them to be discussed in this one paper. To give an impression of all blockchains possibilities this chapter will mention a few examples and some of them will be discussed in more detail.

Some noteworthy (possible) applications of blockchain next to the widely-known application of payments:

1. A bitmortgage, some sort of market for mortgages [3]
   The University TU-Delft created a working prototype for a blockchain based 'mortgage market'.

2. Loans [16]
   Quite similar to item 1 but slightly broader. According to Microsoft blockchain can be used for all sorts of loans.

3. Obligations [16]
   As an obligation can be viewed as a special type of loan, it comes naturally that with item 2 in mind, obligations would go very well with blockchain as well.

4. Voting system [7]
   One of the things smart contracts on blockchain can be used for. Ethereum already built-in smart contracts which can be easily used for this purpose.

5. Less overhead [14]
   Mostly banks - but this could definitely be interesting to more companies - are looking for ways to reduce their overhead with blockchain.
   Which is quite imagineable, since blockchain can be seen as some sort of 'irrevertable ledger', it makes sense to try and reduce overhead created to a large extent to reduce possibilities for fraud.

6. Supply chains [16]
   If all parties involved in supply chains can join in one blockchain, this can ease the communication. Everything will be visible for all parties at all times, making the whole process run smoother.

7. Identity management and verification [1, 16, 19] [1]
   'Ownership of data will shift back from the central parties to the individual". The idea being that since all the data is already being collected and sold - sometimes wrong information - it is better that the individual at least knows what is being said about him/her. And in case of wrong information, that it is correctable.

8. Conduct a majority of the Emirate's business using blockchain [22]
   Dubai wants all it's government services and transactions on blockchain, and nearly all it's businesses done via blockchain by 2020.

The final three examples will be worked out in a little more detail below.

9. Bitcoin
   The very first crypto-valuta, launched by Satoshi Nakamoto himself.

10. Bitnation [13]
    A 'Decentralized Borderless Voluntary Nation, a blockchain powered jurisdiction', an example of what can be done politically / socially with blockchain.

11. Energy reserve supply market [20]
    Creating a supply and demand market for energy where smaller parties can join just as easily as big ones. Hopefully to ease the strain on the already existing network on peak moments for demand of energy.

3.1.1 Bitcoin

By far the most well-known application of blockchain is Bitcoin. It could be argued that blockchain got so well known *because* of Bitcoin.
Just to make sure everyone is on the same page: Bitcoin is a crypto valuta. The first valuta with no bank and/or nation behind it to organize and maintain it.

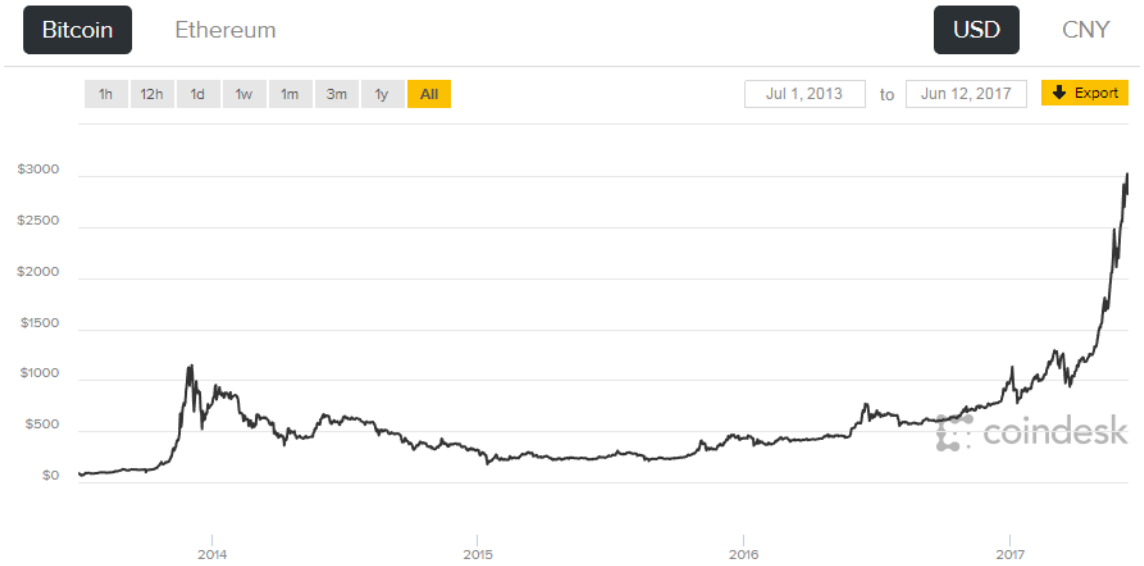


Figure 3.1: The course of Bitcoin since it's launch in 2013 till 12 June 2017 [23]

In Figure 3.1 the course of Bitcoin in USD is shown since it's launch is 2013. Because of the lack of interference it is prone to a lot of fluctuation. At the same time it can be seen that it rises to new hights.

[1] A lot of causion seems to be in place here. Since once something is in a blockchain, it can never be removed or altered again.

This link to a more conventional currency concludes this subsection as much in the preceding paragraphs already shed more light on the workings of Bitcoin.

3.1.2 Bitnation

Called into life by Susanne Tarkowski Tempelhof on July 14, 2014. Tempelhof's father was stateless for a decade, which brought her interest into this field. She has spent her life till now studying the topic of 'non-geographically contingent governance service aggregators'. With the launch of Bitcoin, she realized she possibly could make her dream true, of 'providing more services than just health insurance, but also things like education and security, through networks of local subcontractors'.

> "Bitnation offers the same services as those provided by traditional governments, but in a geographically unbound way. Any individual from around the world can become a citizen of Bitnation by signing on to the constitution".

After being launched on July 14 2014 Bitnation took off. On October 5 the first blockchain marriage was a fact. December 2014 marked the moment when Bitnation announced working on a Basic Income protocol. In April 2015 the first birth certificate came about. In November 2015 Estonia announced it's partnership with Bitnation. This gave access to all e-residence of Estonia to the blockchain of Bitnation with their physical ID cards. February 2016 was a big month for Bitnation, they moved their equity to the ethereum blockchain ánd coded their virtual nation constitution called Pangea. In April 2016 Liberland (a micronation in former Joegoslavia) also partnered up with Bitnation, giving their citizens the same easy services as the Estonians already had.

According to International Business Times "Bitnation is doing for identity and statehood, what Bitcoin is doing for money" [2].

For more details see [13].

3.1.3 Energy reserve supply market

As reported in Computable [20], power Grid Operator Tennet started a pilot in 2017, to investigate whether blockchain technology can help increase the number of decentralized energy parties on the balancing market. This pilot is a cooperation between Tennet, energy market Vandebron and technology parner IBM.

Tennet bears responsibility for the balance on the Dutch energy network. In order to be able to deliver on a continuous base, supply and demand have to be in balance 24/7. Generally the conventional ways of gaining power are easier to balance everything out. However, with an ever increasing part of the energy originating from sustainable energy, it is getting harder and harder to maintain the balance. The idea behind this pilot is that Vandebron delivers energy from recharging electric cars. IBM created a solution based on blockchain which links small batteries to the larger system of Tennet.

By combining this all together, the different parties can regulate the electricity in a safe, transparent way while garanteeing the necessary flexibility.
If the pilot succeeds Tennet wants to invite other decentralized electricity sources to join the project.

[2] http://www.ibtimes.co.uk/bitnation-estonian-government-start-spreading-sovereign-jurisdiction-blockchain-15

Chapter 4

Variable exploration

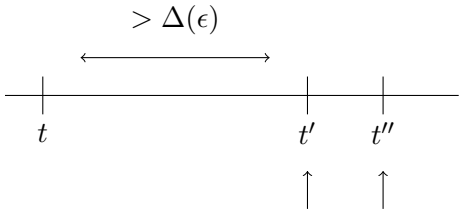
Note: this whole chapter has been built on [15]. Starting off with how they define the 'Nakamoto Consensus' followed by the 'Metrics' they found in order to be able to 'measure' the pre-defined Nakamoto Consensus. After which in section 4.3 this paper will make an effort to pry out some base variables and show how these are all intertwined with each other. To end with an elaborate calculation example with these variables in section 4.4.

4.1 Nakamoto Consensus

Nakamoto described his idea on the blockchain in his paper [6]. But he did not define it mathematically. Now in the paper [15] it has been tried to define the so called 'Nakamoto Consensus' mathematically. Here their definitions will be given and accompanied with a small illustration.

Termination:

> "There exists a time difference function $\Delta(.)$ such that, given a time t and a value $0 < \epsilon < 1$, the probability is smaller than ϵ that at times $t', t'' > t + \Delta(\epsilon)$ a node returns two different states for the machine at time t."



On both times the state in t is
the same according to this node:

$$\mathbb{P}\Big(\Big\{X(t') \neq X(t'')\Big\}\Big) < \epsilon$$

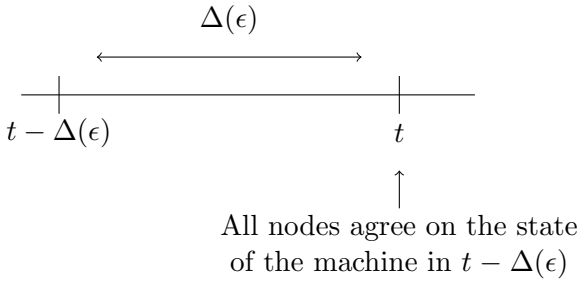
With:
$X(s)$ = the state the machine is in at time
t according to the same node in moment s.

Figure 4.1: Illustration Termination

In other words: "One node's agreement through time with itself."

Agreement:

"There eixsts a time difference function $\Delta(.)$ such that, given a $0 < \epsilon < 1$, the probability that at time t two nodes return different states for $t - \Delta(\epsilon)$ is smaller than ϵ."



$Y(k)$ = state machine at time $t - \Delta(\epsilon)$ according to node k, at current time t.
$\mathbb{P}\left(\left\{Y(k) \neq Y(j) \mid k \neq j\right\}\right) < \epsilon$

Figure 4.2: Illustration Agreement

In other words: "Agreement among different nodes"

Validity:

"If the fraction of mining power by Byzantine nodes is bounded by f, i.e. $\forall t : \frac{\sum_{b \in B(t)} m(b)}{\sum_{n \in \mathcal{N}} m(n)} < f$, then the average fractin of state machine transitions that are not inputs of honest nodes is smaller than f."
"At any time t, a subset of nodes $B(t) \subset \mathcal{N}$ are Byzantine.
$m(i)$ = mining power of node i".

In other words: "A big enough part of the system should be fair"
Note that in reality it will be very hard to determine which nodes are 'fair'.

4.2 Indicators/Metrics

These metrics (or indicators) are designed by [15] to evaluate the unique properties of Nakamoto consensus. Because of the possible verbal confusion which might accompany the word 'metrics' (which sounds quite similar to 'matrix') this paper will use the term 'indicator'.

1. Consensus Delay:
   = the time it takes a system to reach agreement.
   (ϵ, δ) consesnsus delay = $\delta\%$ of the time $\epsilon\%$ of the nodes agree on the state (ϵ, δ) seconds ago.
   e.g. $(95\%, 90\%)$ consesnsus delay = 10 seconds
   means: 90% of the time, 95% of the nodes agree on the state of the machine 10 seconds ago.

2. Fairness:
   = optimally the largest miner and the non-largest miners' representation in the transitions set should be the same as their respective mining powers.

   1. $\dfrac{\text{transitions not coming from largest miner}}{\text{all transactions}}$
   2. $\dfrac{\text{mining power not owned by the largest miner}}{\text{all mining power}}$

   $$\text{fairness} = \frac{1.}{2.} \qquad \text{optimally fairness} = 1.0$$

   Two notes of caution:

   1. ratio 1. is invisible. There are companies by now for Bitcoin whose job it is to keep this even more invisible: you can tell them your transactions, and they will 'mix them up' with transactions of other people. Where in the end everyone gets their own money back, but it will be nearly impossible to trace the original transactions back to each other. Hence how much one 'account' owns.
   2. This only holds when one can see miners as individuals, which is for instance with Bitcoin already tricky due to cartel formation.

3. Mining power:
   $$= \frac{\text{mining power that secures the system}}{\text{total mining power}}$$

4. Time to Prune:
   This implies what time a user has to wait to be confident a transaction has occured.
   δ time to prune =
   $$\delta \text{ percentile} \left(\begin{array}{l}\text{time a node learns this trans-}\\ \text{action has never taken place}\end{array} - \begin{array}{l}\text{time a node learns about a}\\ \text{transaction}\end{array} \right)$$

5. Time to Win: Average time wasted due to forks
   δ time to win =
   $$\delta \text{ percentile} \left(\begin{array}{l}\text{last time a (different) node}\\ \text{disagrees}\end{array} - \begin{array}{l}\text{the first time a node blieves a}\\ \text{never-to-be-pruned-transition}\\ \text{has occured}\end{array} \right)$$

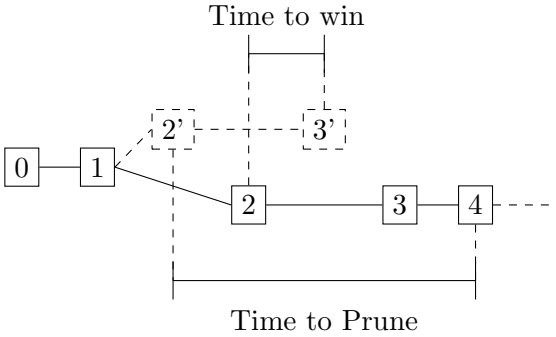


Figure 4.3: Example 'Time to Prune' and 'Time to win' [15]

These indicators turn out to be very difficult -if not impossible- to calculate for a blockchain-in-use (like the Bitcoin blockchain). For each one there is at least one variable uncertain. For instance with the indicator 'Mining power' there is no way of knowing which mining power is

used to secure the system. Probably for this reason the paper [15] worked with simulations. For the present paper there will not be done any simulations. There will only be looked at variables which can be calculated with more or less certainty on a blockchain-in-use.

4.3 How is everything connected?

This section will try to give a useful overview of all the variables and indicators and the relation amongst each other. Followed by an example calculation in section 4.4.

Table 4.1: Notes on Figure 4.4:

label	formula
(1) [1]	$\underset{\text{(MB)}}{\text{blocksize}} = \underset{\text{(bytes)}}{\text{headersize}} + \underset{\text{(bytes)}}{\text{transaction size}} \times \text{\# transactions in a block}$
(2)	$\underset{\text{(\# hashes / second)}}{\text{total mining power}} = \underset{\text{nodes} \in \text{system}}{\sum} \underset{\text{(\# hashes / second)}}{m(i)}$ with $m(i)$ = mining power of node i
(3)	$\underset{\text{(seconds)}}{\text{inter nodes times}} = \frac{\text{blocksize}}{b(i,j)} \quad \frac{(MB)}{(MB/\text{second})}$ with $b(i,j)$ = bandwidth between nodes i and j, $i \neq j$
(4)	$\underset{\text{(seconds)}}{\text{system width}} = \max \left\{ \underset{\text{(seconds)}}{\text{inter nodes times}} \right\}$
(5)	$\underset{\text{(blocks/minute)}}{\text{blockfrequency}} = \frac{\text{total mining power}}{\text{difficulty cryptopuzzel}} \quad \frac{\text{(\# hashes / second)}}{\text{(expected \# hashes needed)}}$
(6) [2]	$\text{\# transactions per second} = \underset{\text{(\# blocks / minute)}}{\text{blockfrequency}} \times \text{\# transactions per block}$
(7) [3]	$\underset{(\in \{0,1\})}{\mathbb{P}(\text{fork})} = 1 - (1+\text{blockfrequency} \times \text{system width}) \times e^{-\text{blockfrequency} \times \text{system width}}$

[1] Keeping all units in mind the actual formula would be:

$$\text{blocksize} = \frac{\text{headersize} + \text{transaction size} \times \text{\# transactions in a block}}{1048576}$$

[2] Keeping all units in mind the actual formula would be:

$$\text{\# transactions per second} = \frac{\text{blockfrequency}}{60} \times \text{\# transactions per block}$$

[3] Formula derived :

$$\begin{aligned}
\mathbb{P}(\text{fork}) &= \mathbb{P}\Big(N(X + \underbrace{\Delta t}_{\geq \text{ system width}}) - N(X) > 1\Big) && \text{with } N(X) = \text{\# blocks found in time } (0, X) \\
&= 1 - e^{-\text{blockfrequency} \times \text{system width}} && N(X) \sim \text{Poisson}\left(\frac{1}{\text{blockfrequency}}\right) \\
&\quad - (\text{blockfrequency} \times \text{system width}) \\
&\quad \times e^{-\text{blockfrequency} \times \text{system width}} \\
&= 1 - (1+\text{blockfrequency} \times \text{system width}) \\
&\quad \times e^{-\text{blockfrequency} \times \text{system width}}
\end{aligned}$$

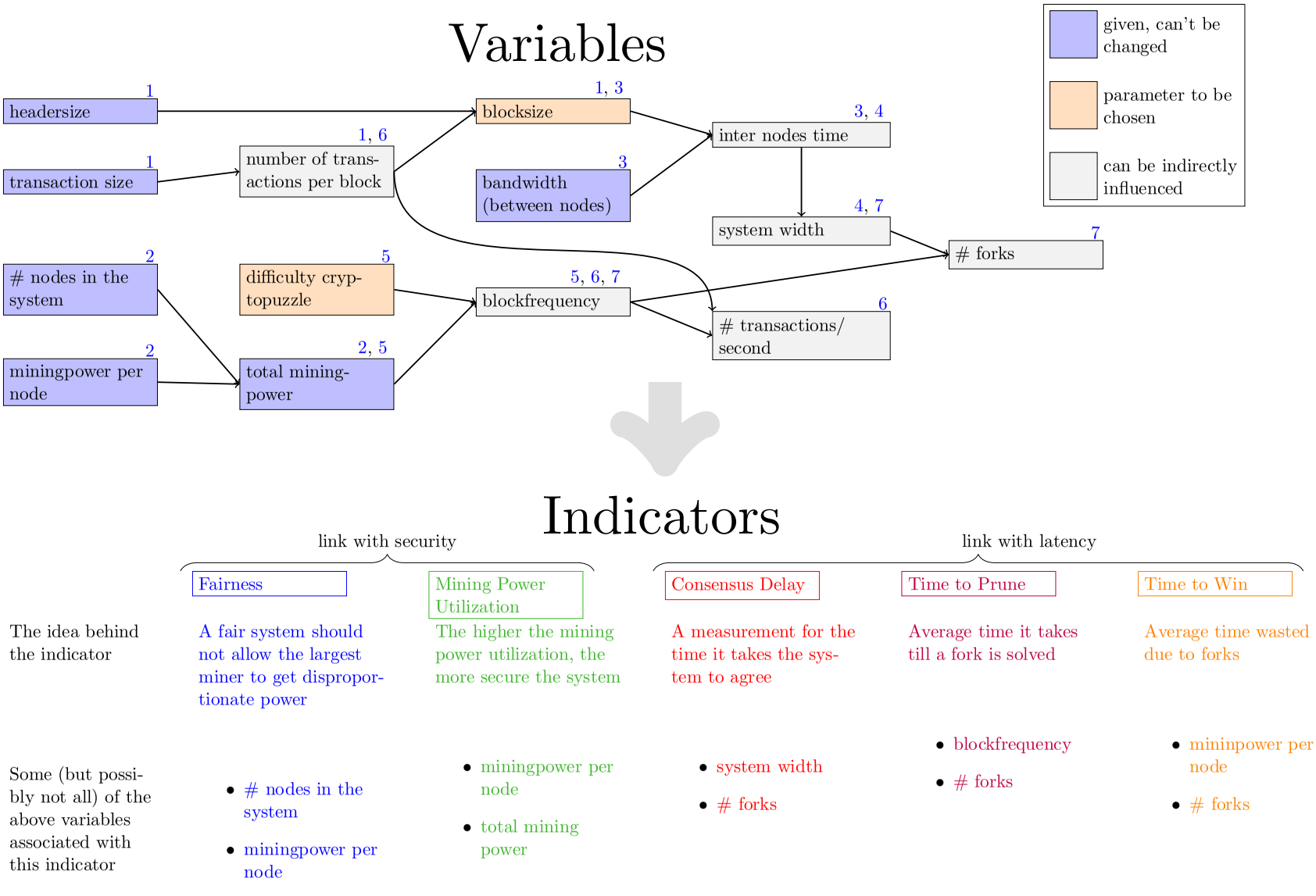


Figure 4.4: The connection between all the variables

4.4 Example Calculation

The former is all very theoretical. This part will give an idea as of how to actually use the formula's given above. This will be done in the form of an elaborate example. Everything will be calculated for one blockchain system. Where the node-graph looks as in Figure 4.5. Hopefully this will help the reader get some feeling for blockchain.

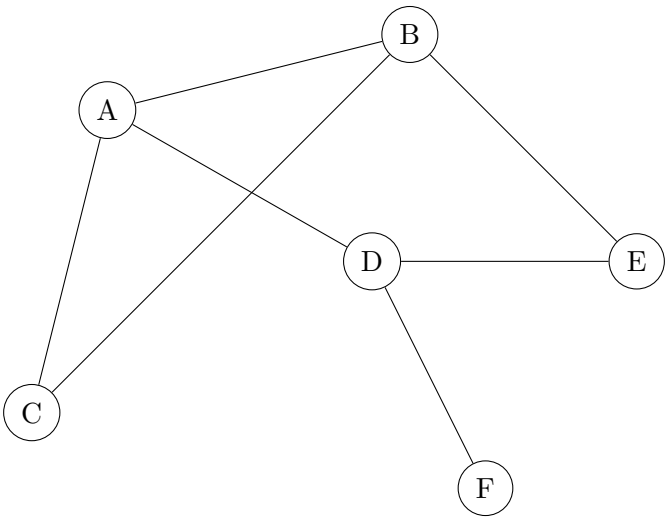


Figure 4.5: The node network

4.4.1 The given parameters

The given variables (as the name suggests) from Figure 4.4 have to be given. There is no way the creator of the blockchain can change these. Many of them are dictated by the users/nodes of the blockchain causing them to change regularly. This example will work with the following:

name	**value**
headersize	80 bytes
transaction size	500 bytes
# nodes in the system	6

	A	B	C	D	E	F
Mining power per node	1000	1500	500	750	900	1100

Bandwidth

	A	B	C	D	E	F
A	×	9	10	7		
B		×	6		10	
C			×			
D				×	7	10
E					×	
F						×

4.4.1.1 Calculated given parameter

This is quite a special 'type' of parameter: there is only one of this type! Usually when a parameter depends on other parameters the creator of the blockchain can influence it at least

a little. But this parameter is different. It is the only one which does need other parameters in order to be able to calculate it, but the creator cannot do anything to change it. Making it both a part of the 'given' parameters and the 'calculated' parameters.

name	**value**
total mining power	$= \sum_{i \in \text{ nodes}} m(i)$ $= 1000 + 1500 + 500 + 750 + 900 + 1100$ $= 5750$

4.4.2 Chosen parameters

This the only type of parameter the creator of the blockchain can choose. All other parameters that the blockchain creator has any influence on one way or the other depend on these two variables. Because these two influence so much, and partly have trade-offs it is often a matter of the creators personal preference and insights in to what is most important for this particular use of blockchain which dictate how they are chosen.

name	**value**	
blocksize	= 1 MB	(= 1 048 576 bytes)
difficulty cryptopuzzle	= 57 500	

4.4.3 Indirectly influenced parameters

Last but definitely not least! This is the most interesting part of the calculation: how do the chosen parameters influence the blockchain? How far does their influence reach? Can a creator make up for some 'bad-luck' when the blockchain has to deal with suboptimal given parameters?

name	value
Number of transactions per block	$= \frac{\text{blocksize - headersize}}{\text{transaction size}}$
	$= \frac{1,048,576 - 80}{500}$
	$= 2{,}096.992$
blockfrequency	$= \frac{\text{total mining power}}{\text{difficulty cryptopuzzel}}$
	$= \frac{5,750}{57,500}$
	$= 0.1$ blocks per minute
inter nodes time	$= \frac{\text{blocksize}}{b(i,j)}$
	$=$

	A	B	C	D	E	F
A	×	$^1/_9$	$^1/_{10}$	$^1/_7$	$^{19}/_{90}$	$^{19}/_{90}$
B		×	$^1/_6$	$^{17}/_{70}$	$^1/_{10}$	$^{12}/_{35}$
C			×	$^{17}/_{70}$	$^{14}/_{45}$	$^{12}/_{35}$
D				×	$^1/_7$	$^1/_{10}$
E					×	$^{17}/_{70}$
F						×

	(for the □ see subsubsection 4.4.3.1)
system width	$= {}^{12}/_{35}$ second
# transactions/second	$= \frac{\text{blockfrequency}}{60} \times$ # transactions per block
	$= \frac{0.1}{60} \times 2{,}096.992$
	≈ 3.5
$\mathbb{P}$(forks)	$\approx 0.057\%$ for calculation see subsubsection 4.4.3.2.

4.4.3.1 Inter nodes time

The inter nodes time is a bit tricky to calculate. This is because one needs the bandwidth between all nodes, even those whithout a direct line (□). In order to get these, one first needs the 'quickest' route between the not directly linked nodes. Resulting in a *Shortest path problem* [4]. Now since this is such a small graph it is solved relatively easy. This paper will not go into how to solve such a problem and instead, will just give the answers to be able to continue this example:

$$\begin{aligned} AE &= ABE \\ &= \frac{1}{9} + \frac{1}{10} \\ &= \frac{19}{90} \end{aligned} \qquad \begin{aligned} BD &= BED \\ &= \frac{1}{10} + \frac{1}{7} \\ &= \frac{17}{70} \end{aligned}$$

$$\begin{aligned} AF &= ADF \\ &= \frac{1}{9} + \frac{1}{10} \\ &= \frac{19}{90} \end{aligned} \qquad \begin{aligned} BF &= BEDF \\ &= \frac{1}{10} + \frac{1}{7} + \frac{1}{10} \\ &= \frac{12}{35} \end{aligned}$$

[4] https://en.wikipedia.org/wiki/Shortest_path_problem

$$
\begin{aligned}
CD &= CAD \\
&= \frac{1}{10} + \frac{1}{7} \\
&= \frac{17}{70}
\end{aligned}
\qquad
\begin{aligned}
EF &= EDF \\
&= \frac{1}{7} + \frac{1}{10} \\
&= \frac{17}{70}
\end{aligned}
$$

$$
\begin{aligned}
CE &= CABE \\
&= \frac{1}{10} + \frac{1}{9} + \frac{1}{10} \\
&= \frac{14}{45}
\end{aligned}
$$

$$
\begin{aligned}
CF &= CADF \\
&= \frac{1}{10} + \frac{1}{7} + \frac{1}{10} \\
&= \frac{12}{35}
\end{aligned}
$$

All inter node times are shown in Figure 4.6.

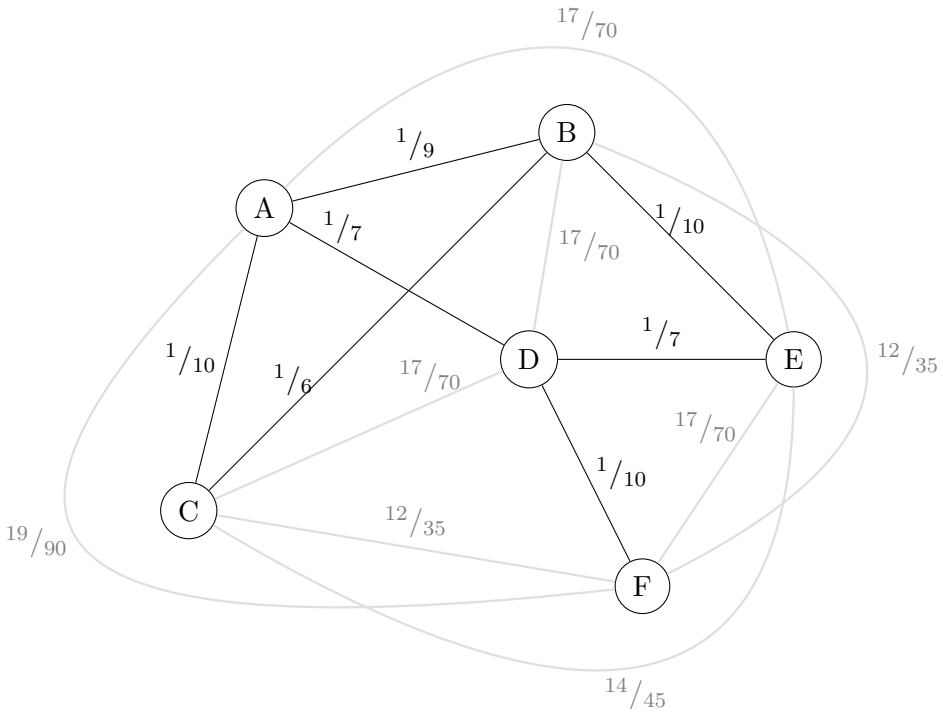


Figure 4.6: All inter nodes times

4.4.3.2 $\mathbb{P}(\textbf{fork})$

The general formula has already been given in Table 4.1 (7). However here an effort will be made to take the reader through the whole process to get to this formula for this specific example.

This is the least straight forward calculation this report will contain. A small example has already been given in the Example on page 6. Here it will be done for the current elaborate example.

To start off, the assumption is made that the number of blocks found in the system in a given time interval follows a Poisson process. This is not a very weird assumption to make. Let's look at it this way: there are LOTS of nodes (big n) and a very small chance of finding a block (small p) creating a more or less 'stable' λ. In this case λ represents the number of blocks found in a unit of time. In this example being blockfrequency = 0.1 per minute.

Now up for some declaring:
$N\big(t\big)$ = the number of blocks found in the time (0,t)
$\quad \sim Poisson(0.1t)$
$\Delta t \geq$ system width

$$
\begin{aligned}
\mathbb{P}\big(\text{fork}\big) &= \mathbb{P}\Big(\big\{ \text{ more than one block found in the time interval } (0, \Delta t) \big\}\Big) \\
&= \mathbb{P}\Big(N\big(\Delta t\big) > 1\Big) \\
&\text{take the complementairy probability} \\
&= 1 - \mathbb{P}\Big(N\big(\Delta t\big) \leq 1\Big) \\
&= 1 - \mathbb{P}\Big(N\big(\Delta t\big) = 0\Big) - \mathbb{P}\Big(N\big(\Delta t\big) = 1\Big) \\
&= 1 - \frac{e^{-0.1 \cdot \frac{12}{35}} (0.1 \cdot \frac{12}{35})^0}{0!} - \frac{e^{-0.1 \cdot \frac{12}{35}} (0.1 \cdot \frac{12}{35})^1}{1!} \\
&= 1 - e^{-0.1 \cdot \frac{12}{35}} - 0.1 \cdot \frac{12}{35} \cdot e^{-0.1 \cdot \frac{12}{35}} \\
&\approx 1 - 0.966295 - 0.03313013 \\
&\approx 0.00057449 \\
&\approx 0.057\%
\end{aligned}
$$

In other words the probability of a fork with these given parameters would be about 0.057%.

An other noteworthy finding from the above calculation is that the chance on $\mathbb{P}\Big(N\big(\Delta t\big) = 0\Big)$ is by far the biggest chance. In other words, the chance of a block found in a time interval of $^{12}/_{35}$ seconds is not even 6%! Which is what would be expected with a low blockfrequency.

Chapter 5

Relations between variables

Having defined all the relations between the variables and found their formulas it was time to see whether some of them might have more interesting relations than expected.
A program written in *R* allowed for some exploration of the relationships between some of the variables. The variables which have been altered in order to look for a change are:

- headersize
- transaction size
- mining power node A
- blocksize
- difficulty cryptopuzzel

In order to see the real difference a variable does, one time all the figures have been made where all variables are kept at one value.
There are 16 variables which are checked to see whether any change is visible:

- headersize
- transaction size
- mining power node A
- mining power node B
- mining power node C
- mining power node D
- mining power node E
- mining power node F
- total mining power
- blocksize
- difficulty cryptopuzzel
- number of transactions per block
- blockfrequency
- system width
- transactions per second
- P(fork)

Obviously it is expected that 'mining power node {A,B,C,D,E,F}' all react similarly. But since they together make 'total mining power' the choice has been made to keep them all in the figures.

For comparison purposes, the 'dull' figure where every variable stays stable is added as well, see Figure 5.1.

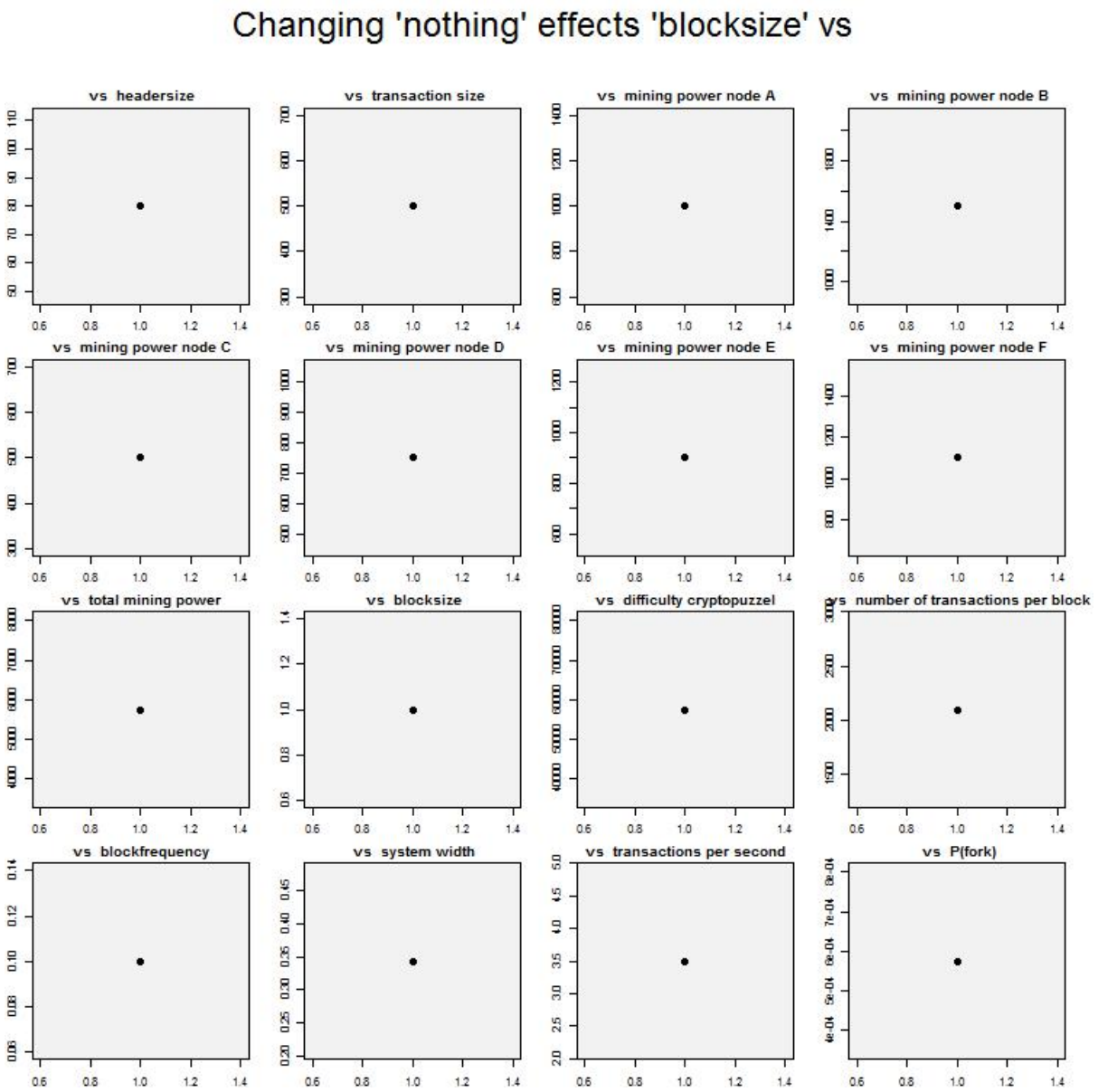


Figure 5.1: The 'dull' figure where nothing is changed

Below some of the results are shown.

5.1 Blocksize

Looking at Figure 4.4 it can be expected that the following variables will show some degree of change:

- number of transactions per block
- system width
- transactions per second
- P(fork)

The blocksize is gradually changed from 0 MB to 150 MB. Both numbers are extreme: no blocksize is ever going to be set at 0 MB for the simple reason that no data could be stored at

all. Similarly 150 MB is really very big. For instance Bitcoin works with blocksize 1MB. The general result is shown in Figure 5.2.

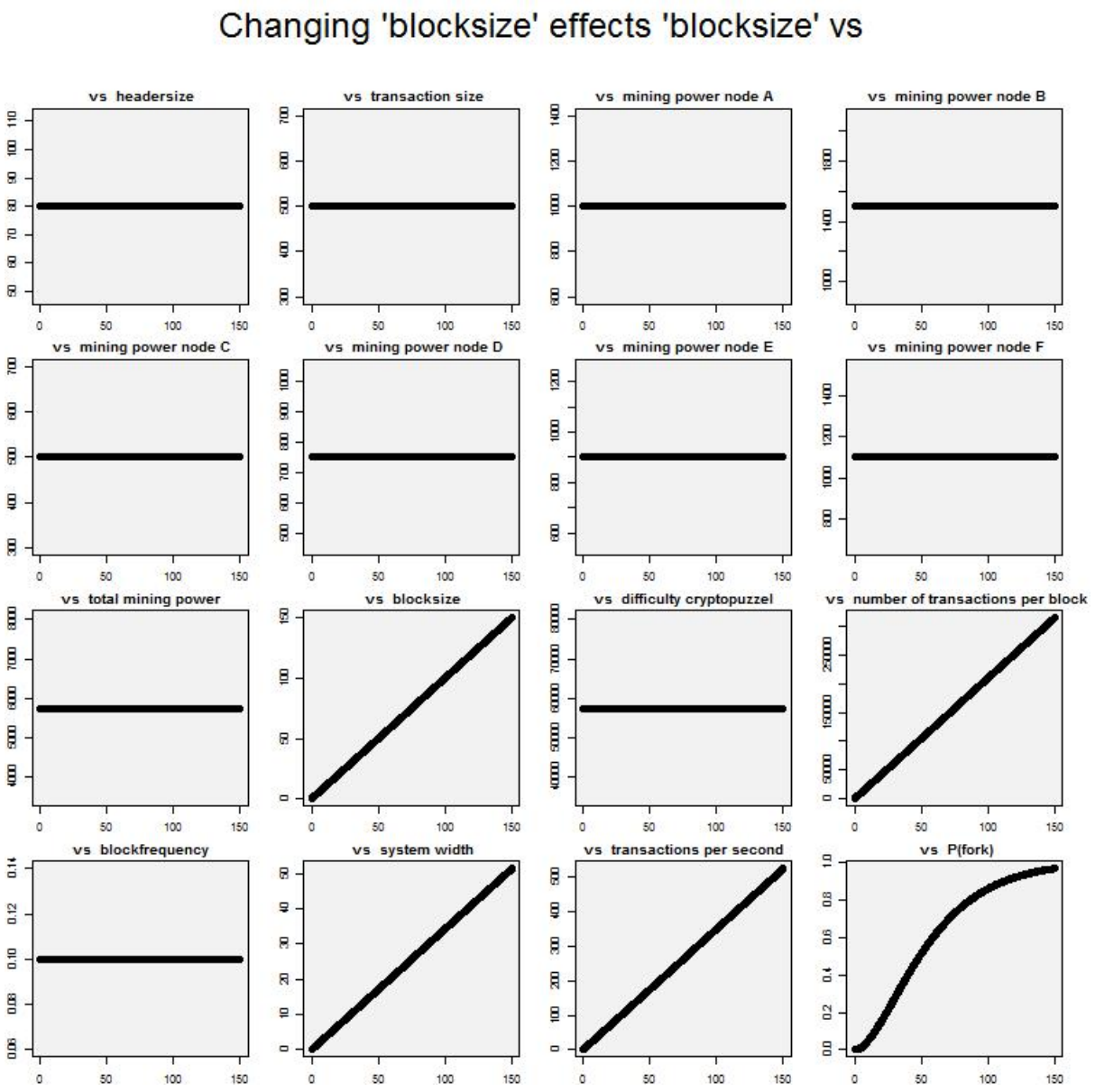


Figure 5.2: The effect blocksize has on the different variables

As expected on most variables there is no real influence. The 'number of transactions per block', 'system width' and 'transactions per second' all have a linear relation with blocksize. 'P(fork)' however has a more 'S-like' relation. In Figure 5.3 this relation is shown a bit better.

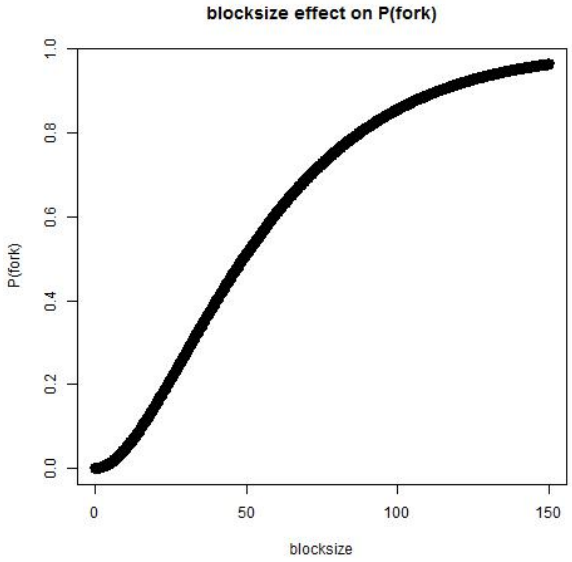


Figure 5.3: The effect of blocksize on P(fork)

It can be seen that when the blockchain is very small, the probability of a fork occuring is extremely small and doesn't increase very quickly. This soon changes and the probability of a fork occuring rises quickly. Until it gets close to certainty with a limit of probability 1.

5.2 Headersize

Looking at Figure 4.4 it can be expected that the following variables will show some degree of change:

- number of transactions per block
- transactions per second

And indeed as Figure 5.4 shows, only those two variables are affected. Both in the expected linear way. Since the blocksize is taken up partly by the headersize, increasing the headersize leaves less space for the transactions. Resulting in the negative angle of the line.

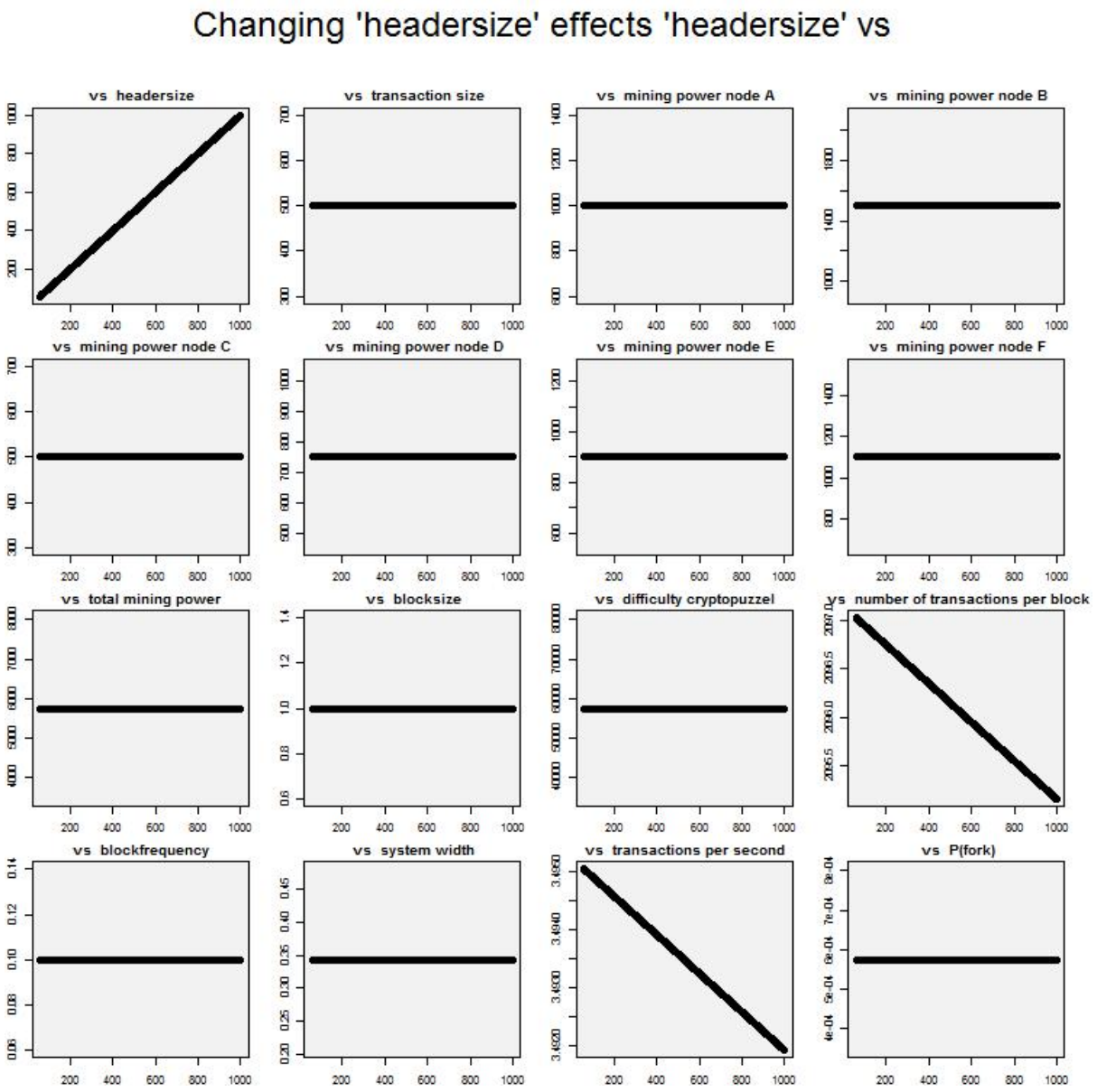


Figure 5.4: The effect of adjusting the headersize

5.3 Transaction size

Looking at Figure 4.4 it can be expected that the following variables will show some degree of change:

- number of transactions per block
- transactions per second

And indeed as Figure 5.5 shows, only those two variables are affected. Since the blocksize is taken up mostly by the transactions, increasing the transaction size leaves exponentially less space for more transactions. Resulting in the quick downward slope of both graphs (see Figure 5.6 and Figure 5.7).

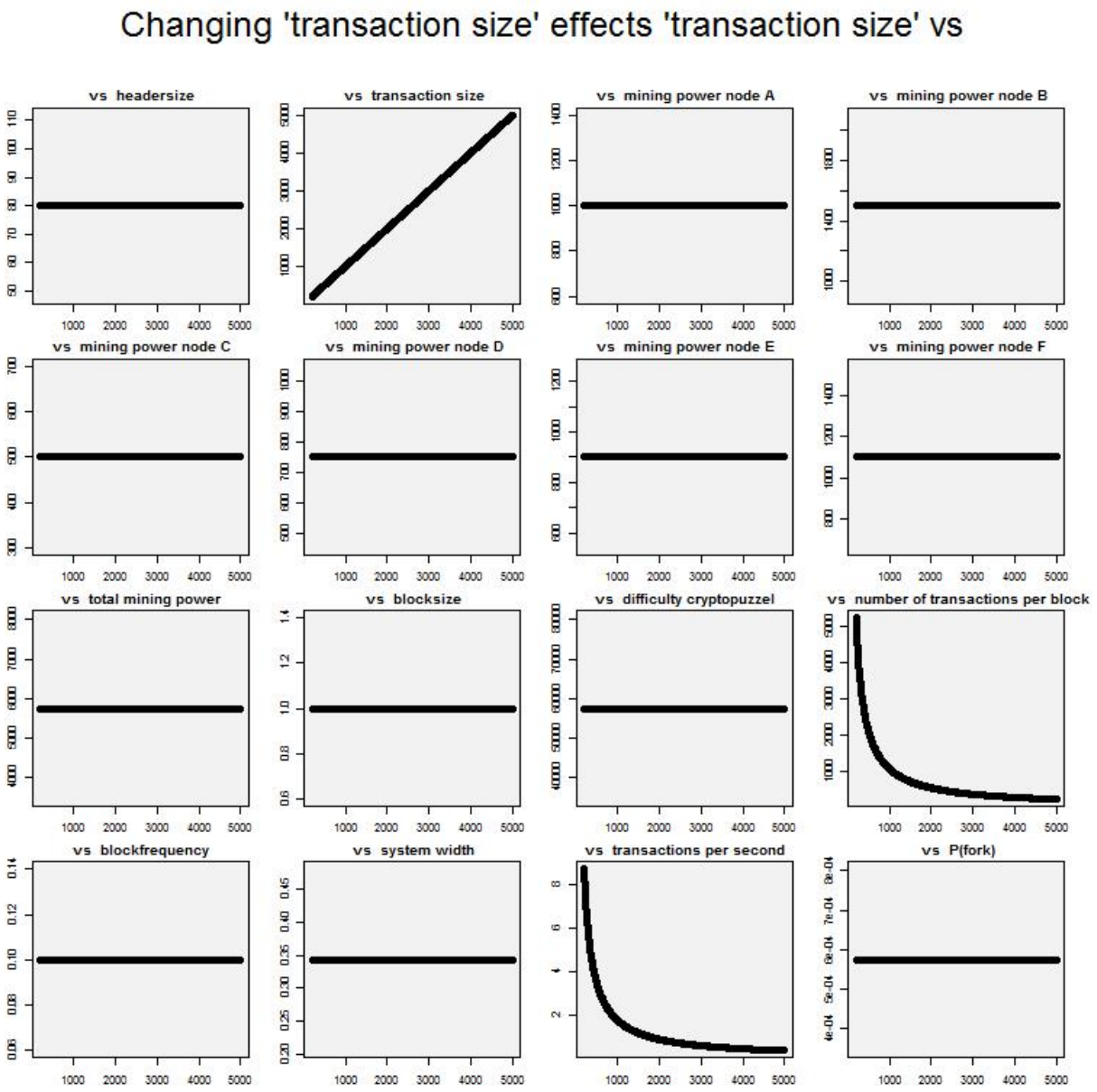


Figure 5.5: The effect of adjusting the transaction size

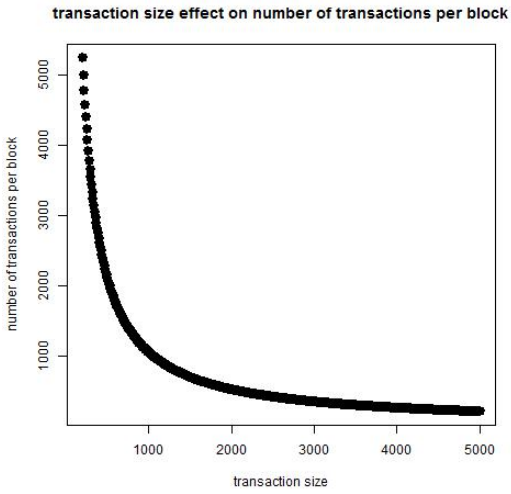


Figure 5.6: The effect of adjusting the transaction size on the number of transactions per block

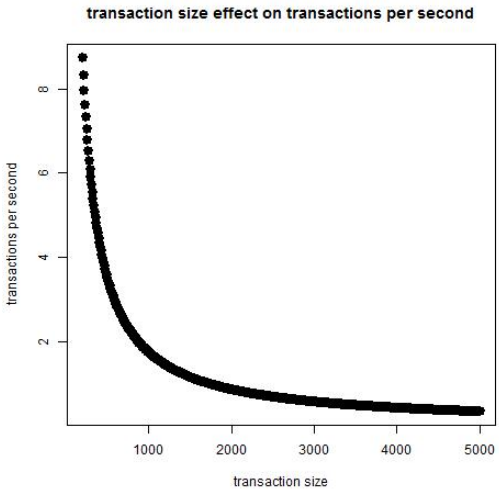


Figure 5.7: The effect of adjusting the transaction size on the number of transactions per second

5.4 Mining power node A

Looking at Figure 4.4 it can be expected that the following variables will show some degree of change:

- total mining power
- blockfrequency
- number of transactions per second
- P(fork)

And indeed as Figure 5.8 shows, those variables are affected. Most of them in the expected linear way. Again the P(fork) shows a slight 'S-form', just as with section 5.1.

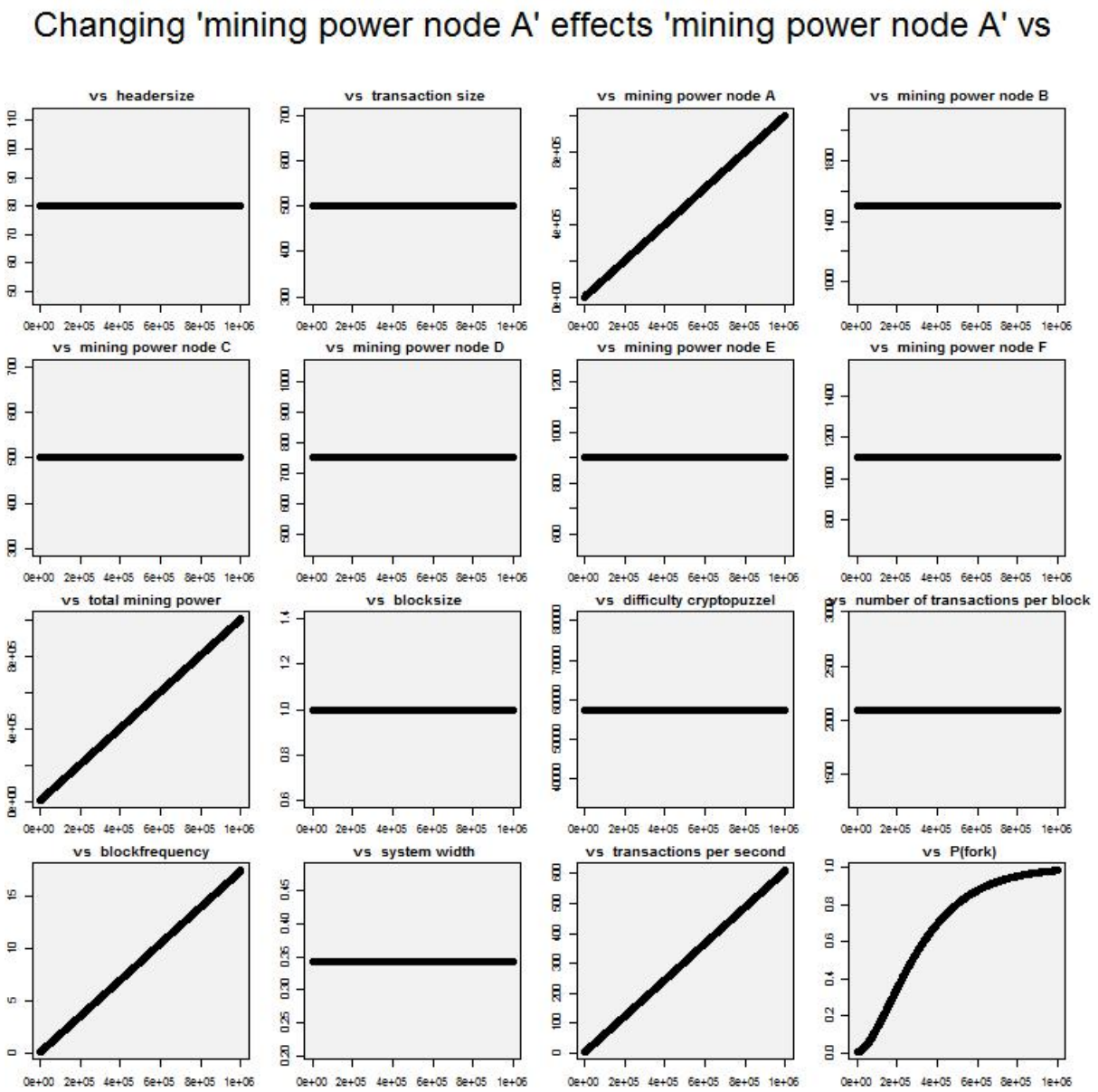


Figure 5.8: The effect of adjusting the mining power node A

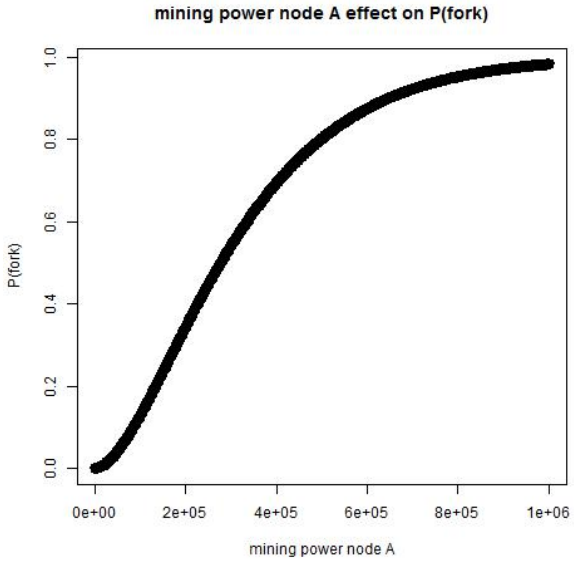


Figure 5.9: The effect of adjusting the mining Power of a Node

This S-graph is shown more clearly in Figure 5.9. It is easily explained why this S-form occurs: Since the mining power goes up, there are more frequent blocks. If block frequency goes up, so does the probability of getting a fork. Starting at nearly no chance of a fork, going up to near certainty (probability of 1).

5.5 Difficulty Cryptopuzzle

Looking at Figure 4.4 it can be expected that the following variables will show some degree of change:

- blockfrequency
- number of transactions per second
- P(fork)

And indeed as Figure 5.10 shows, only those three variables are affected. All in the same way: a higher difficulty, results in less blocks, thus less transactions per second and thus a smaller probability on forks.

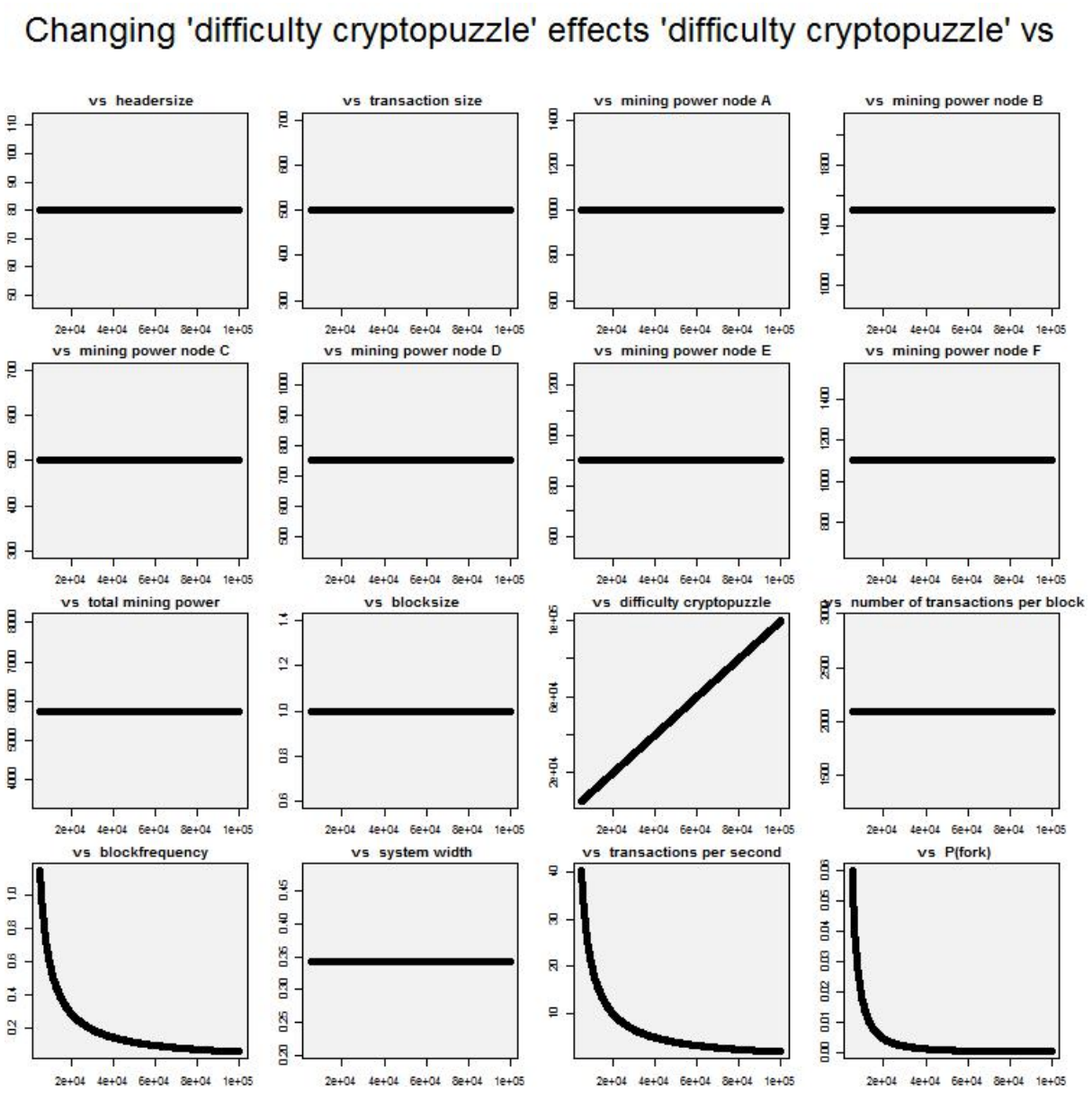


Figure 5.10: The effect of adjusting the difficulty of the cryptopuzzle

This time the graphs seem clear enough from the general overview. For that reason no 'close up' have been added.

Chapter 6

Discussion/further research

The initial research has answered some interesting questions. But during the research some new questions arose. This chapter presents some of the questions which arose, possible interesting thoughts which belong nowhere else in the paper and suggestions for further research.

1. Many people are looking into blockchain. Some people are finding smart solutions to some of blockchains challenges. For instance a '*bitcoin New Generation*' works with microblocks to solve the issue where blockchain can only contain a limited number of transactions per second.

2. Another '*blockchain new generation*' restricts the number of neighbouring nodes, increasing the unknown part of the blockchain and making it even harder for a Byzantine node to plan a double spending attack.

3. A major issue with blockchain is the energy consumption necessary for the proof-of-work. Another way of validation is proposed already, called '*proof-of-stake*'. As of yet there is no scientific proof for its safety, but some cryptovaluta have been using it for a while now (e.g. Blockcoin) without any known hacks.

All three aspects present good follow-up research questions.

Chapter 7

Conclusion

This paper sets out to provide a general exploration of blockchain. After having done the research, it became clear that explaining blockchain in a concise manner to people who do not know it already is nearly impossible. Blockchain involves too many new concepts which cannot easily be compared to familiar concepts. But taken together, the Management Summary and the Glossary presented with this paper may provide a good general idea of this new system.

Blockchain is a new type of database which solves the double spending problem without a middleman, opening up a whole range of new possibilities. In this database the data is saved in blocks arranged as links in a chain. To secure this block-chain a system called proof-of work is used. In this system so much work (i.e. processing power) is needed to find a block that it is virtually impossible to alter the blockchain afterwards. The work is done by so called miners who get a payment for their effort and the system is set up in such a way that its financially more advantageous for the miners to keep the system in good order than to try and subvert it.

To further explore and explain the workings of blockchain this paper focusses in more detail on the Nakamoto blockchain, the original and the most commonly known for its use in Bitcoin.

The conclusion can be drawn that Blockchain is a very useful new type of database which probably harbors many still-to-be-found solutions. It allows for solutions previously unthinkable. Each sector should take a close look at blockchain, and see whether it has in the past put questions aside which -at that time- were unsolvable, but might very well be solvable with blockchain.

On the other hand there are limits too to the possibilities and uses of blockchain, e.g. it is not suitable in any case where data has to be able to be removed. Also, the research indicates some points of caution, such as use of energy.

References

[1] Wordt Nederland de nieuwe Silicon Valley met Blockchain?
https://www.nrc.nl/advertentie/deloitte/wordt-nederland-de-nieuwe-silicon-valley-met-bl

[2] Blockchain-Coalitie presenteert actieagenda
http://agconnect.nl/artikel/blockchain-coalitie-presenteert-actieagenda?utm_source=nb_agc_20170401&utm_medium=email&utm_term=&utm_content=&utm_campaign=1-04-2017

[3] TU-Delft heeft een werkend prototype blockchain hypotheekmarktplaats
https://agconnect.nl/artikel/tu-delft-heeft-blockchain-hypotheek-werkend

[4] 13 vragen over blockchain
https://agconnect.nl/artikel/13-vragen-over-de-blockchain

[5] Satoshi Nakamoto, the creator(s) of bitcoin
https://en.wikipedia.org/wiki/Satoshi_Nakamoto

[6] Bitcoin: A Peer-to-Peer Electronic Cash System
https://bitcoin.org/bitcoin.pdf

[7] Algemene info van pro-Ethereum
http://www.bestebank.org/ethereum/

[8] Programmig in Solidity video
https://youtu.be/VllP64YvFDY

[9] Blockchain Wikipedia
https://en.wikipedia.org/wiki/Blockchain

[10] ANDERS BROWNWORTH, 'Blockchain 101 - A Visual Demo'
https://www.youtube.com/watch?v=_160oMzblY8

[11] ANDERS BROWNWORTH, 'Blockchain Demo'
https://anders.com/blockchain/

[12] Byzantine fault tolerance
https://en.wikipedia.org/wiki/Byzantine_fault_tolerance

[13] Bitnation
https://en.wikipedia.org/wiki/Bitnation

[14] Wordt 2017 het blockchain-jaar voor Nederlandse banken?
https://agconnect.nl/artikel/wordt-2017-het-blockchain-jaar-voor-nederlandse-banken?utm_source=nb_agc_20170403&utm_medium=email&utm_term=&utm_content=&utm_campaign=3-04-2017

[15] ITTAY EYAL, ADAM EFE GENCER, EMIN GÜN SIRER, ROBBERT VAN RENESSE, CORNELL UNIVERSITY, "Bitcoin-NG: A Scalable Blockchain Protocol", 16/18-03-2016
https://www.usenix.org/system/files/conference/nsdi16/nsdi16-paper-eyal.pdf

[16] COMPUTABLE, 'Microsoft levert gratis blockchain-testomgeving', 10-02-2017
https://www.computable.nl/artikel/nieuws/finance/5951836/250449/microsoft-levert-gratis-blockchain-testomgeving.html

[17] CONNECT, "Blockchain is hype", 18-04-2017
https://agconnect.nl/blog/blockchain-hype?utm_source=nb_agc_20170418&utm_medium=email&utm_term=&utm_content=&utm_campaign=18-04-2017

[18] WHITEFIELD DIFFIE and MARTIN E. HELLMAN, 'New Directions in Cryptography', 6-11-1976
https://www-ee.stanford.edu/~hellman/publications/24.pdf

[19] CONNECT, 'Blockchain geeft burger autonomie terug', 24-04-2017
https://agconnect.nl/artikel/blockchain-geeft-burger-autonomie-terug?utm_source=nb_agc_20170424&utm_medium=email&utm_term=&utm_content=&utm_campaign=24-04-2017

[20] COMPUTABLE, 'Tennet test blockchain voor energienet', 4-5-2017
https://www.computable.nl/artikel/nieuws/security/6014258/250449/tennet-test-blockchain-voor-energienet.html?utm_source=nieuwsbrief&utm_medium=email&utm_campaign=Dagelijks_04_05_2017&utm_content=topartikelen

[21] FRANCOIS ZANINOTTO, 'The Blockchain Explained to Web Developers, Part 1: The Theory', 28-4-2016
https://marmelab.com/blog/2016/04/28/blockchain-for-web-developers-the-theory.html

[22] NIKHIL LOHADE, 'Dubai Aims to Be a City Build on Blockchain', 24-4-2017
https://www.wsj.com/articles/dubai-aims-to-be-a-city-built-on-blockchain-1493086080?__prclt=DyD2kRRG

[23] COINDESK
http://www.coindesk.com/price/

[24] ADAM BACK, 'Hashcash - A Denial of Service Counter-Measure'
http://www.hashcash.org/papers/hashcash.pdf

Disclaimer:
During the research a lot of sources have been used. Great effort has been made to do right to their efforts and reference all the sources. However if a reference is missing, I sincerely apologize.

Appendix A

Glossary

Blockchain is at the start a difficult concept. This is mostly because it needs a different view and involves new ideas on several different important issues. Which need each other in order to be able to describe them. This Glossary is therefore not so much a completely correct description of used terms, but rather is ment for the reader in the first part of the report, to help him form a basic idea of the broad concept of blockchain. Hopefully the remainder of the report will correct and deepen the initial understanding of the given concepts.

Agree
Once a block has become a link in the chain of the blockchain, it means that all miners/nodes have agreed on the validity of the data in this block.

Block
A block is where the blockchain saves its data. The blockchain saves data in bunches at a time, instead of a continuous stream. Such a bunch of data saved at the same time is called a block. Blocks are arranged as links in a chain , hence the name "blockchain".

Chain
The blocks in the blockchain are linked together in order of time. Since all the blocks are linked together, this forms a chain.

Miner / node
A CPU which joined the blockchain, and which works very hard to find the next block in order to reap some benefit (generally in the form of crypto-valuta or tokens).

Nonce
A random figure (generated by the computer) needed to find a block. The Nonce influences the Hash which in turn makes enormous amounts of processing power necessary to find a block. This in turn is key to the security system.

Proof-of-work
This is the security system of many blockchain systems. Because there is so much work (processing power) needed to find a block, it is virtually impossible to alter the blockchain afterwards. This is what keeps the blockchain safe.