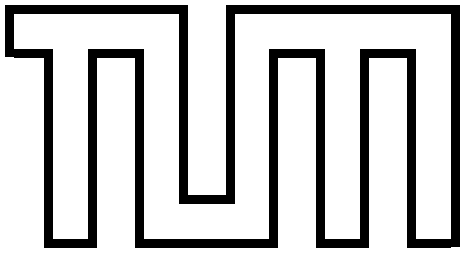


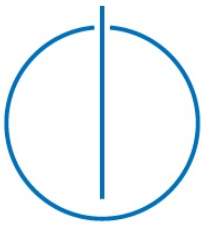
Technische Universität München

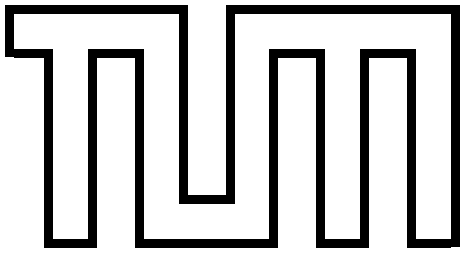
Fakultät für Information Systems

Master's Thesis in Informatik

Bitcoin-like Blockchain Simulation System

Fabian Schüssler





Technische Universität München

Fakultät für Informatik

Master's Thesis in Information Systems

Bitcoin-like Blockchain Simulation System

Bitcoin-ähnliches Blockchain Simulationssystem

Author: Fabian Schüssler

Supervisor: Prof. Dr. Hans-Arno Jacobsen

Advisor: Pezhman Nasirifard

Submission: xx.xx.20xx

I confirm that this master's thesis is my own work and I have documented all sources and material used.

München, xx.xx.20xx

(Fabian Schüssler)

Abstract

English Abstract

Inhaltsangabe

German Abstract

Acknowledgment

Acknowledgement

Contents

List of Figures

List of Tables

Abbreviations

tps	transactions per seconds.
VIBES	Visualizations of Interactive, Blockchain, Extended Simulations.

Chapter 1

Introduction

Currently and in the last years Blockchain technologies such as Bitcoin and Ethereum are a very hot topic. According to the Gartner Hype Cycle [1] Blockchain technology is undergoing the peak of inflated expectation in 2017. Price and market capitalization changes of cryptocurrencies are covered by the media.

Blockchain technology enables decentralized consensus and can be used for record keeping. Bitcoin is the first digital currency to solve the double-spending problem without the need of a trusted authority. One of the main problems of Bitcoin or in general of Blockchains is the low maximum amount of possible processed transactions per seconds. Additionally, the Bitcoin community disagrees about how to solve this scalability problem, already split about different approaches and created Bitcoin forks.

VIBES (Visualizations of Interactive, Blockchain, Extended Simulations) is a blockchain sim-ulator, which allows fast, scalable and configurable network simulations on a sin-gle computer without any additional resources. It was developed in a master thesis by Lyubomir Stoykov [2] and which is the foundation for this master thesis proposal. The goal of my master thesis is to improve VIBES to make more realistic simulations possible. In the future VIBES could be used by developers or heavy blockchain users to simulate changes to different Blockchains. So maybe in the future it can help the Bitcoin community to agree on change proposals.

There are other Simulators like Bitcoin-Simulator, but VIBES is the first simulator de-signed to be extended to blockchain systems beyond bitcoin and the first of its kind to be able to simu-late transactions in the network. Bitcoin-Simulator only simulates the network at block level and therefore does not consider transactions [3].

1.1 Motivation

Motivation of Thesis.

1.2 Problem Statement

There are lots of possibilities to improve VIBES, which were also already outlined in the mas-ter thesis of Stoykov. The portability to different kinds of Blockchains and the ease of use is very important.

Here are some possibilities to extend the current version of VIBES, that I would like to work on in my master thesis: • Add maximal block size and transaction incentives • Break propagation delay down into three components: network bandwidth, block size and distance between Nodes. • Improve speed by testing mutable variables • Differentiate between miners and full nodes • Calculate resources used by network: CPU, electricity and bandwidth. • Change transaction generation from coordinator to nodes. • Other improvements: code improvements, analysis and visualizing of information that is already captured by VIBES (probability of forks), ...

These extensions can improve the quality of the simulations and the use cases of VIBES. For example, the implications of Segwit2x could be analyzed. Maybe these extensions could also make it possible to realistically simulate the current Bitcoin blockchain.

1.3 Approach

The goal to make simulations more realistic. The approach from the original VIBES paper is taken over. Some designs and parts of the architecture have to be adjusted.

The configuration parameters (4.1.8) must be extended to break down propagation delay, to add maximal block size and transaction incentives.

Change transaction generation from coordinator to nodes (4.3.1).

Differentiate between miners and full nodes (4.3.2).

1.4 Organization

Organization of Thesis.

Chapter 2

Background

Background concepts required for understanding the thesis.

2.1 Bitcoin

2.1.1 Node/ Miner

in this work node = miner

2.1.2 Block

genesis block

Chapter 3

Related Work

Related work materials

Chapter 4

Approach

In this chapter the changes to the existing VIBES framework are presented one by one. First the reasons for each improvement are described. Then the prerequisites for every change and the design and architecture are shown. Finally the implementation is explained. Part of the implementation is the implementation of the front-end. The front-end shows the results of the back-end and can therefore also be used for the evaluation.

where to put bugfixes?

4.1 Bitcoin-like Blockchain Simulation

changes necessary to differentiate strategies especially in the front-end

4.2 Lazy Logging

Previously logging only occurred in the console. This made debugging for long simulations difficult. Especially for the evaluation of any implementations a log file is necessary. For this reason the Scala modules *logback* and *scala-logging* were integrated into the project. Every important event is logged into */logfile.log*.

4.3 Block size limit

One of the biggest unresolved issues of Bitcoin-like Blockchains is scalability. The main factor to measure scalability is transactions per seconds (tps). Previously VIBES had no block size limit. This means infinite transactions can be processed and changing input parameters has no effect on the scalability. To be able to investigate the effects of different input parameters on the scalability, the introduction of a block size limit is necessary. This allows a more accurate simulation of Bitcoin.

4.3.1 Prerequisites

configuration parameters

- *maxBlockSize*: the max block size in KB, the current default value is 1.000 KB

4.3.2 Design and Architecture

The main back-end change happens in the model VBlock. Depending if the simulation is a Bitcoin-like Blockchain Simulation all generated blocks obey the block size limit.

4.3.3 Implementation

```
object VBlock extends LazyLogging {
  def createWinnerBlock(node: VNode, timestamp: DateTime): VBlock = {
    var maxTransactionsPerBlock : Int = 0
    var processedTransactionsInBlock: Set[VTransaction] = Set.empty

    if (VConf.strategy == "BITCOIN_LIKE_BLOCKCHAIN") {
      // todo think about if to implement SegWit (maxBlockWeight vs
          ↪ maxBlockSize)
      maxTransactionsPerBlock = Math.floor(VConf.maxBlockSize /
          ↪ VConf.transactionSize).toInt

      // sorts the transaction pool by the transaction fee
      processedTransactionsInBlock =
          ↪ node.transactionPool.toSeq.sortWith(_.transactionFee >
          ↪ _.transactionFee).take(maxTransactionsPerBlock).toSet

      // sets confirmation status of transaction true
      processedTransactionsInBlock.foreach { _.confirmation = true }

      // sets confirmation level of transaction
      processedTransactionsInBlock.foreach { _.confirmationLevel =
          ↪ node.blockchain.size }
    } else {
      maxTransactionsPerBlock = node.transactionPool.size
      processedTransactionsInBlock = node.transactionPool
    }

    VBlock(
      id = UUID.randomUUID().toString,
      origin = node,
      transactions = processedTransactionsInBlock,
      level = node.blockchain.size,
      timestamp = timestamp,
      recipients = ListBuffer.empty,
      transactionPoolSize = node.transactionPool.size
    )
  }
}
```

4.4 Transaction incentives

4.5 Block time

4.6 Alternative history attack

4.6.1 Prerequisites

To simulate an alternative history attack additional **configuration parameters** are necessary. These parameters are used for the actual simulation of the attack, the calculation of the success probability of the attack and the maximum safe transaction value.

- *isAlternativeHistoryAttack*: if an alternative history attack is simulated as a boolean
- *hashRate*: attacker's hashRate as a percentage of the total hashRate of the Bitcoin Network
- *confirmations*: the amount of confirmations the attacked merchants are waiting for to accept a transaction
- *attackDuration*: the attacker gives up after mining a certain amount of blocks and not succeeding or if it is not possible any more to surpass the level of the honest blockchain
- *discountOnStolenGoods*: discount of the stolen goods by the attacker, a value from 0 (= full discount) to 1 (= no discount)
- *amountOfAttackedMerchants*: the attack is carried out against a certain amount of merchants at the same time
- *blockReward*: current block reward in BTC

4.6.2 Design and Architecture

Simulating the attack

In the following the attacker's nodes, blockchain or blocks are interchangeably described as *evil* and the honest networks' nodes as *good*.

The solution for the simulation of an alternative history attack selects nodes as attacking nodes according to the attacker's hashRate as a percentage of the total Bitcoin Network. The good and the evil nodes both can mine the genesis block. The genesis block is then the first block in both the good and the evil blockchain. For simplicity we assume that the attacker successfully sent the transactions to the attacked merchants in the second block of the honest blockchain. Immediately after the genesis block is mined, the evil nodes start mining together on their own evil blockchain. It is necessary for all nodes to update their neighbour nodes to only have their corresponding nodes as neighbours. For example in the case of a low amount of evil nodes and a low amount of neighbours... Since the attacker of course doesn't want his...

Finally the success of the simulated attack is decided if the attacker's blockchain level can surpass the honest's blockchain level after waiting for the Merchants confirmation and before the attack duration ends.

Calculating the success probability [4]

The ... is ... Before the formula to calculate the success probability of an alternative history attack is shown, the variables need to be explained. q is *hashRate*, the attacker's percentage of the hash rate of the total network. p is $1 - q$ and the percentage of the honest network.

$$p + q = 1 \tag{4.1}$$

It is the goal to calculate the success probability r. If the attacker's hash rate q is equal or bigger than p, then the success probability of the attacker is 100%.
Due to the implementation the behaviour of the implementation can deviate from the 100%. For example the variables *attackDuration*, *confirmations* or the simulation duration can have an impact.
If $q < p$, then the upper complex formula with binomial coefficients needs to be calculated.

$$r = \begin{cases} 1 - \sum_{m=0}^{n} \binom{m+n-1}{m}(p^n q^m - p^m q^n), & \text{if } q < p \\ 1, & \text{if } q \geq p \end{cases} \tag{4.2}$$

The formula for $q < p$ is transformed for the implementation. This allows the usage of factorial functions instead of binomial coefficients.

$$r = \begin{cases} 1 - \sum_{m=0}^{n} \frac{(m+n-1)!}{m!\,(n-1)!}(p^n q^m - p^m q^n), & \text{if } q < p \\ 1, & \text{if } q \geq p \end{cases} \tag{4.3}$$

Calculating the maximal safe transaction value [4]

(4.3)

$$\frac{(1-r)oB}{k(\alpha + r - 1)} \tag{4.4}$$

4.6.3 Implementation

Chapter 5

Evaluation

The evaluation chapter.

5.1 Correctness

The evaluation of the correctness of the Simulator's front-end output is essential. By validating the output of the front-end we also validate the output of the back-end.

Some parts of the evaluations are sample testing for consistency between the input parameters, expected and simulated or calculated output. Other parts are empirical tested.

For the calculation of the probability of a successful double spend and the maximal safe transaction value sample testing is used. For validating the success probability of simulated double spends empirical testing is applied.

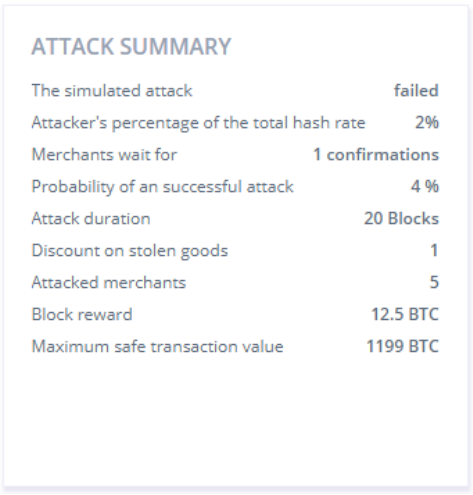
5.1.1 Calculation of the probability of a successful double spend

The probability of a successful double spend is tested with five samples and the results are compared to the figure A.1 from the original paper [4].

The edge cases q = 2% and n = 1, q = 50% and n = 1, q = 2% and n = 10 and q = 2% and n = 10 are tested as well as the common case of q = 30% and n = 6.

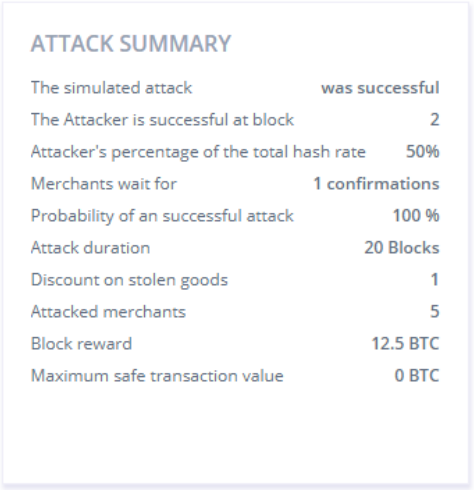
All mentioned test cases of figure 5.1 match the expected result.

... table comparing both

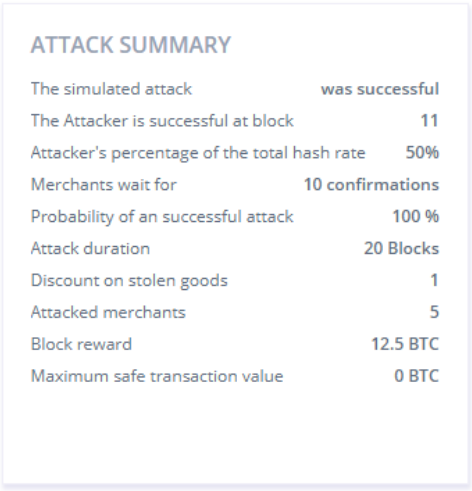


(a) Test case: q = 2% and n = 1

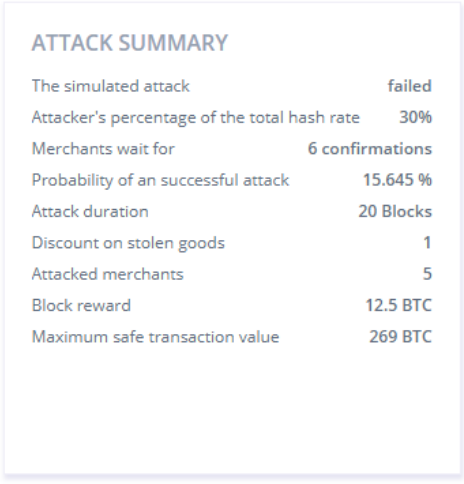
(b) Test case: q = 2% and n = 10



(c) Test case: q = 50% and n = 1



(d) Test case: q = 50% and n = 10



(e) Test case: q = 30% and n = 6

Figure 5.1: Test cases: probability of a successful double spend

The probability for the test case (e) with q = 30% and n = 6 is used in Chapter 5.1.3.

5.1.2 Calculation of the maximal safe transaction value

For testing the maximal safe transaction value the samples from the evaluation of the successful double spend probability are reused and compared to the figure A.2 from the ... paper [4].

The additional input parameters were:

- Attack duration: 20 Blocks
- Discount on stolen goods: 1
- Attacked merchants: 5
- Block reward: 12.5 BTC

Compared to the original paper the block reward was updated from 25 BTC to the current block reward of 12.5 BTC. This means the maximal safe transaction values need to be doubled to compare them with the correct values.

All test cases of figure 5.1 match the expected result, except for (a) and (b). (a) is off by one due to rounding. The correct result of (b) is infinite BTC, the front-end shows the maximal 32-bit signed integer value of 2,147,483,647 BTC.

... table comparing both

5.1.3 Simulation of double-spending: success probability

For the empirical testing of double-spending the following script was used. It starts the simulation, waits for a certain time to let the simulation finish and repeats this for a total of 100 times.

```
ECHO Start of Loop

FOR /L %%i IN (1,1,100) DO (
  ECHO %%i
  start chrome "http://localhost:8082/vibe?blockTime=567&
     ↪ numberOfNeighbours=4&numberOfNodes=20&simulateUntil
     ↪ =1531411943382&transactionSize=1&throughput=105&latency=900&
     ↪ neighboursDiscoveryInterval=3000&maxBlockSize=100&maxBlockWeight
     ↪ =4000&networkBandwidth=1&strategy=BITCOIN_LIKE_BLOCKCHAIN&
     ↪ transactionPropagationDelay=150&hashRate=30&confirmations=6"
  timeout /t 40
)
```

Information for replication: New features might make additional parameters in the URL necessary.

Outcome	Occurrences	Simulated probability	Expected probability
ATTACK NEITHER SUCCESSFUL NOR FAILED	45	-	-
ATTACK SUCCESSFUL	8	15.09%	15.645%
ATTACK FAILED	47	84.91%	84.355%
TOTAL	**100**	**100%**	**100%**

Table 5.1: Double-spending outcomes and their simulated and expected probabilities

After counting the occurrences of the outcomes in the logfile, they were summarized in the table 5.1. All simulations were finished in the specified time. The outcome "ATTACK NEITHER SUCCESSFUL NOR FAILED" can happen if the simulation time was too short. This number is ignored for the calculation of the probabilities. It is assumed that the unfinished simulations have a similar probability distribution like the finished ones. The simulated probability of the double-spending attack is with 15.09% very close to the expected probability of 15.645%, which was calculated in Chapter 5.1.1. Hereby is shown that the simulation of double-spending has the correct success probability.

5.1.4 Simulation: transactions per seconds

5.1.5 Simulation: transaction incentives

5.1.6 Simulation of a double spend: time between blocks

Chapter 6

Summary

Summary

6.1 Status

Final Status of the Thesis

6.2 Conclusions

Concluding remarks of Thesis

6.3 Future Work

Future Work

for bitcoin related stuff

node != miner selfish mining mining pools

Appendices

Appendix A

Appendix

A.1 First

First Appendix

q	1	2	3	4	5	6	7	8	9	10
2%	4%	0.237%	0.016%	0.001%	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0
4%	8%	0.934%	0.120%	0.016%	0.002%	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0
6%	12%	2.074%	0.394%	0.078%	0.016%	0.003%	0.001%	≈ 0	≈ 0	≈ 0
8%	16%	3.635%	0.905%	0.235%	0.063%	0.017%	0.005%	0.001%	≈ 0	≈ 0
10%	20%	5.600%	1.712%	0.546%	0.178%	0.059%	0.020%	0.007%	0.002%	0.001%
12%	24%	7.949%	2.864%	1.074%	0.412%	0.161%	0.063%	0.025%	0.010%	0.004%
14%	28%	10.662%	4.400%	1.887%	0.828%	0.369%	0.166%	0.075%	0.034%	0.016%
16%	32%	13.722%	6.352%	3.050%	1.497%	0.745%	0.375%	0.190%	0.097%	0.050%
18%	36%	17.107%	8.741%	4.626%	2.499%	1.369%	0.758%	0.423%	0.237%	0.134%
20%	40%	20.800%	11.584%	6.669%	3.916%	2.331%	1.401%	0.848%	0.516%	0.316%
22%	44%	24.781%	14.887%	9.227%	5.828%	3.729%	2.407%	1.565%	1.023%	0.672%
24%	48%	29.030%	18.650%	12.339%	8.310%	5.664%	3.895%	2.696%	1.876%	1.311%
26%	52%	33.530%	22.868%	16.031%	11.427%	8.238%	5.988%	4.380%	3.220%	2.377%
28%	56%	38.259%	27.530%	20.319%	15.232%	11.539%	8.810%	6.766%	5.221%	4.044%
30%	60%	43.200%	32.616%	25.207%	19.762%	15.645%	12.475%	10.003%	8.055%	6.511%
32%	64%	48.333%	38.105%	30.687%	25.037%	20.611%	17.080%	14.226%	11.897%	9.983%
34%	68%	53.638%	43.970%	36.738%	31.058%	26.470%	22.695%	19.548%	16.900%	14.655%
36%	72%	59.098%	50.179%	43.330%	37.807%	33.226%	29.356%	26.044%	23.182%	20.692%
38%	76%	64.691%	56.698%	50.421%	45.245%	40.854%	37.062%	33.743%	30.811%	28.201%
40%	80%	70.400%	63.488%	57.958%	53.314%	49.300%	45.769%	42.621%	39.787%	37.218%
42%	84%	76.205%	70.508%	65.882%	61.938%	58.480%	55.390%	52.595%	50.042%	47.692%
44%	88%	82.086%	77.715%	74.125%	71.028%	68.282%	65.801%	63.530%	61.431%	59.478%
46%	92%	88.026%	85.064%	82.612%	80.480%	78.573%	76.836%	75.234%	73.742%	72.342%
48%	96%	94.003%	92.508%	91.264%	90.177%	89.201%	88.307%	87.478%	86.703%	85.972%
50%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Figure A.1: The probability of a successful double spend, as a function of the attacker's hashrate q and the number of confirmations n. The abscissa shows the confirmations n and the ordinate shows the attacker's hashrate q.

q	1	2	3	4	5	6	7	8	9	10
2%	2400	42K	644K	9370K	$\approx \infty$	$\approx \infty$	$\approx \infty$	$\approx \infty$	$\approx \infty$	$\approx \infty$
4%	1150	10K	82K	615K	4437K	$\approx \infty$	$\approx \infty$	$\approx \infty$	$\approx \infty$	$\approx \infty$
6%	733	4722	25K	127K	626K	3018K	14M	$\approx \infty$	$\approx \infty$	$\approx \infty$
8%	525	2650	10K	42K	159K	588K	2144K	7749K	$\approx \infty$	$\approx \infty$
10%	400	1685	5741	18K	56K	168K	503K	1486K	4361K	12M
12%	316	1158	3391	9212	24K	62K	157K	396K	990K	2460K
14%	257	837	2172	5200	11K	27K	60K	132K	290K	632K
16%	212	628	1474	3178	6580	13K	26K	52K	102K	200K
18%	177	484	1043	2061	3901	7202	13K	23K	42K	74K
20%	150	380	763	1399	2453	4190	7039	11K	19K	31K
22%	127	303	571	983	1615	2582	4053	6288	9671	14K
24%	108	244	436	710	1103	1665	2467	3608	5229	7525
26%	92	198	337	523	775	1113	1570	2182	3005	4106
28%	78	161	263	392	556	766	1035	1377	1815	2372
30%	66	131	206	296	406	539	701	899	1141	1435
32%	56	106	162	225	299	385	485	602	740	901
34%	47	86	127	172	221	277	340	411	491	582
36%	38	69	99	130	164	200	240	283	331	383
38%	31	54	76	98	121	144	169	196	224	254
40%	25	42	57	72	87	102	118	134	151	168
42%	19	31	41	51	61	70	80	90	99	109
44%	13	21	28	34	40	46	51	57	62	68
46%	8	13	17	21	24	27	30	32	35	38
48%	4	6	8	9	10	12	13	14	15	16
50%	0	0	0	0	0	0	0	0	0	0

Figure A.2: The maximal safe transaction value, in BTC, as a function of the attacker's hashrate q and the number of confirmations n. The abscissa shows the confirmations n and the ordinate shows the attacker's hashrate q.

Bibliography

[1] K. Panetta, "Top trends in the gartner hype cycle for emerging technologies, 2017." `https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/`, 2017. Accessed: 2018-04-15.

[2] L. Stoykov, "Vibes: Fast blockchain simulations for large-scale peer-to-peer networks," Master's thesis, Technische Universität München, 2018.

[3] *On the security and performance of proof of work blockchains*, 2016.

[4] M. Rosenfeld, "Analysis of hashrate-based double spending," 02 2014.