

Trading Network Performance for Cash in the Bitcoin Blockchain

subtitle

Enrico Tedeschi

Master Thesis in Computer Science



Abstract

Nowadays blockchain systems are emerging and they are spreading each day more. Cryptocurrencies are the biggest example of a physical implementation of this protocol, having in 2012 more than 50 thousands transactions per day and reaching now in 2017 more than 400 thousands of transactions approved every day. In this thesis we evaluate the most famous blockchain system, the *Bitcoin blockchain*. Public blockchains have emerged as a plausible messaging substrate for applications that require highly reliable communication. However, sending messages over existing blockchains can be cumbersome and costly as miners require payment to establish consensus on the sequence of messages. The blockchain protocol requires an always growing size of the informations stored in it so its *scalability* is the biggest problem. For that reason we collected data to be analyzed and stored in our own dataframe, saving up to $x10$ space for the analysis.

This thesis will consider the network performance of the Bitcoin public ledger when used as a messaging substrate. From 2009 to 2017 a lot of analysis has been done on Bitcoin blockchain and meanwhile its block size limit changed multiple times, from 256 bytes up to 1 Mb, the Bitcoin price raised from $\sim 0.7 \$$ to more than 4.000 \\$ and different papers were published discussing whether changing or not the block size limit or talking about the fees a miner could get from clients. We read and considered previous analysis on Bitcoin blockchain, presented our own dataset, which contains a significant portion of the Bitcoin blockchain, updated at 09-2017, discussed our results and compared them with other evaluations obtained during the years, then we also discussed how the fee paid to miners evolves during time and how much a client could pay for a faster approval time, plus we take into consideration transaction visibility, blockchain growth and fees paid to miners. From this we propose and evaluate, using machine learning techniques, three different cost prediction models for predicting bandwidth per Bitcoin cost of upcoming transaction. The models can be used by application to throttle network traffic to optimize message delivery. We also discuss and consider, according to the data obtained, whether the block size limit should be increased for an higher *throughput* or not.

People are using Bitcoin because it has a lower fee rate and no central authority,

we aim to find any possible relation between the fee paid from a transaction to a miner and the approval time of this transaction, plus we also noticed that the bigger is the blockchain size the more the system become centralized, since only few members, or nodes, of the p2p network can support and use the full blockchain.

Bitcoin blockchain has been analyzed with a blockchain analytics system, developed using Bitcoin's API and data were collected both by using the API and parsing *blockchain.info* HTML pages. A total of # transactions has been evaluated, more transactions than ever were considered before and useful informations about the Bitcoin blockchain emerged. This thesis gives also a measurement about accuracy of data provided from *blockchain.info*.

Contents

Abstract	i
List of Figures	v
My list of definitions	vii
1 Introduction	1
1.1 Problem Statement	2
1.2 Method / Context	2
1.3 Outline	2
2 Related Works - SotA	3
2.1 A Transaction Fee Market Exists Without a Block Size Limit	3
2.1.1 Miner's Profit Equation	4
2.1.2 The Mempool Demand Curve	5
2.1.3 The Block Space Supply Curve	5
2.1.4 Maximizing the Miner's Profit	6
3 Technical Background	7
4 Blockchain Analytics System	9
4.1 Blockchain Data Sources	9
4.2 System Architecture	9
4.2.1 Data Retrieval	9
4.2.2 Data Manipulations	9
4.2.3 Methods	9
4.3 Version Control	9
5 Blockchain Observations	11
5.1 Blockchain Growth	11
5.2 Retrieval Block Time	11
5.3 Block Analysis	11
5.4 Bandwidth	11
5.5 Block Fee	11

5.6 Models	11
6 Conclusions	13
6.1 Discussion	13
6.2 Future Implementation	13
6.3 Comments	13
References	15
A Terminology	19
B List of Symbols	21
C Listing	25

List of Figures

2.1 Maximizing the profit of a miner evaluating the unconfirmed transactions from the Bitcoin blockchain. Mempool demand, $M_{demand}(b)$, and space supply, $M_{demand}(Q)$, curve are represented with our analytics system, and the block space Q^* for a maximum profit is around 1 Mb. There the gap between the two curves is bigger.	6
---	---

My list of definitions

/ 1

Introduction

Blockchain technology opens up to usages in several important sectors such as trading, file storage, and identity management. This technology is widely used in decentralized digital currencies, and at the present, the most common are Bitcoin (2008) and Ethereum (2013), but many more are emerging.

Blockchains essentially implements a distributed consensus protocols that enable a set of untrusted processes to agree on the content of an append only data structures. These ledgers are divided into blocks and linked together in sequence by hashes. They facilitate transactions between consenting individuals who would otherwise have no means to trust each other and deal with geographical separation and interfacing difficulties. This technology promises a highly resilient and communication substrate where messages are kept potentially for a long time.

Nonetheless, decentralized digital currencies also have some side effects. The most relevant is *scalability*, due to the steady growth of the blockchain. It should be also considered that decentralized cryptocurrencies operate in open (or permissionless) networks in which the ledger of data could be manipulated from arbitrary adversaries and according also to the paper from University of Singapore [38] security of smart contracts has not received much attention yet. And since the only part not protected from cryptography is the *order of transactions* [10], an attacker would try to convince the network that a transaction occurred earlier than another one to gain money. The security bugs in smart contracts are classified as *Transaction-Ordering Dependence, Timestamp*

Dependence, Mishandled Exceptions and Reentrancy Vulnerability [38].

Scalability and network performances are urgent concern in existing Blockchain-based cryptocurrencies [25]. It is true that Bitcoin has lower fees than centralized currency schemes, but these properties come at a performance and scalability cost. According to [10], if Bitcoin would have the same amount of transactions of a VISA circuit, its blockchain would grow about 1 MB per 3 seconds. Centralized schemes, like VISA are immediate, while having a throughput of 2,000 transactions/sec up to 56,000 transactions/sec [25]. In this thesis we presents some of our observation of the performance of the Bitcoin blockchain. We provide detailed insights and analysis on how Bitcoin's characteristics have changed over time, and provide an updated model describing how the Bitcoin blockchain will grow. We analyse the correlation between the fee paid to the miner and block creation time, the time it takes for a transaction to be visible in the whole network. Three different models are proposed to describe how applications best can spend money to improve network characteritics, this affects average bandwidth available to an application.

1.1 Problem Statement

1.2 Method / Context

1.3 Outline

/2

Related Works - SotA

Scalability and analysis on the blockchain has been taken into consideration by many researchers in the past years. This chapter summarizes the most relevant papers or works that talks about Bitcoin, decentralized cryptocurrencies, concepts behind them and statistical analysis on the blockchain. In a previous paper that we wrote [46], we enhanced the importance of paying for having a certain bandwidth in the Bitcoin network. A paper from Peter R. Rizun [42], explains how a rational Bitcoin miner should select transactions from his node mempool, when creating a new block, in order to maximize his profit. Analysis on the blockchain in matter of scalability is showed in the Position Paper of Kyle Croman [25], they analyze how fundamental bottlenecks in Bitcoin limit the ability of its current peer-to-peer overlay network to support substantially higher throughputs and lower latencies. We are going to test the throughput as well, comparing it with the one showed in this paper. Furthermore, to fully understand how is possible to make money out of the blockchain and mining, is necessary to have a view of how VISA [11] makes money as well.

2.1 A Transaction Fee Market Exists Without a Block Size Limit

This paper shows how a Bitcoin miner should select transactions from his node's mempool, when creating a new block, in order to maximize his profit

in the absence of a block size limit. *Block space supply curve* and *mempool demand curve* are explained, and the paper shows how the supply and demand curves from classical economics are related to the derivatives of these two curves.

They claim that the block-size limit determines the transaction throughput and one of their concern regards whether, in the absence of such a limit or if that limit is far above the transactional demand, a *healthy transaction fee market* would develop which charges users the full cost to post transactions. The object of this paper is indeed to consider whether or not such a fee market is likely to emerge if miners, rather than the protocol, limit the block size. In this paper they derive the *miner's profit equation* and then they introduce two novel concepts called the *mempool demand curve* and the *block space subbly curve*.

2.1.1 Miner's Profit Equation

Every time a block is mined, the miner expects to generate a revenue $\langle V \rangle$ at hashing cost $\langle C \rangle$ to earn profit per block

$$\langle \Pi \rangle = \langle V \rangle - \langle C \rangle. \quad (2.1)$$

Miner's profit equation in 2.1 shows the gain of a miner $\langle \Pi \rangle$, where the hashing cost is represented as follows:

$$\langle C \rangle = \eta h T. \quad (2.2)$$

So the hashing cost $\langle C \rangle$ is directly dependent from the miner's individual hash rate, h , the cost per hash, η , and the creation time, T . Moreover, is important to consider the expectation value of a miner's revenue per block, this value is represented with $\langle V \rangle$ and is equal to the amount he would earn if he won the block multiplied by his probability of winning. So the expected revenue would be: $\langle V \rangle = (R + M)h/H$, where the amount he would earn is the sum of the block reward, R , and the transaction fees, M . His probability of winning, assuming all blocks propagating instantly, is equal to the ration of his hash rate, h , to the total hash rate of the Bitcoin network, H . The problem with this equation is that it does not reflect the miner's diminished chances of winning if he chooses to publish a block that propagates slowly to the other miners. If a miner finds first a valid block, but his solution is received after most miners are working on another, then his block will likely be discarded. This effect is called *orphaning*. The equation, considering the orphaning factor, $\mathbb{P}_{\text{orphan}}$, is the following:

$$\langle V \rangle = (R + M) \frac{h}{H} (1 - \mathbb{P}_{\text{orphan}}). \quad (2.3)$$

Where the chance that a block gets orphaned increases with the amount of time it takes the block to propagate to the other miners. Indeed, if τ is the block propagation time, the probability of orphaning is defined as:

$$\mathbb{P}_{\text{orphan}} = 1 - e^{-\frac{\tau}{T}}. \quad (2.4)$$

In conclusion the *miner's profit equation* is defined as:

$$\langle \Pi \rangle = (R + M) \frac{h}{H} e^{-\frac{\tau}{T}} - \eta h T \quad (2.5)$$

A *rational miner* selects which transactions to include in his block in a manner that maximizes the expectation value of his profit. This selection is explained with the *mempool demand curve* and the *block space supply curve*.

2.1.2 The Mempool Demand Curve

The set of transactions that still need to be approved and included in a block is called *mempool*. The mempool set is denoted with \mathcal{N} and the number of transactions contained within it as n . According to the size limit, a block can select a $b \leq n$ transactions from \mathcal{N} to create a new block $\mathcal{B} \subset \mathcal{N}$.

A block first includes transactions with a higher *fee density*. This last, is a ratio between the *transaction fee* and the *transaction size*. To construct the mempool demand curve, is necessary first sorting the mempool from greatest fee density to least and then associating an index $\{i : 1, 2, \dots, n - 1, n\}$ with each transaction in the resulting list.

The mempool demand curve will be then a graphical representation of the sum of the fees offered by each transaction in this sorted list:

$$M_{\text{demand}}(b) \equiv \sum_{i=1}^b \text{fee}_i, \quad (2.6)$$

and the sum of each transaction's size in bytes:

$$Q(b) \equiv \sum_{i=1}^b \text{size}_i. \quad (2.7)$$

2.1.3 The Block Space Supply Curve

The size of the block a miner elects to produce controls the fees he attempts to claim, $M(Q)$, and the propagation time he chooses to risk, $\tau(Q)$. The block space supply curve represents the fees a miner requires to cover the additional cost of

supplying block space Q . This cost grows exponentially with the propagation time. The equation which represents this curve is the following:

$$M_{\text{Supply}}(Q) = R \left(e^{\frac{\Delta\tau(Q)}{T}} - 1 \right), \quad (2.8)$$

where $\Delta\tau(Q) \equiv \tau(Q) - \tau(0)$. The propagation time τ , is just an esteem from the propagation delay versus the block size.

2.1.4 Maximizing the Miner's Profit

To maximize his profit, the miner construct a mempool demand curve and a space supply curve. The block size Q^* where the miner's surplus, $M_{\text{demand}} - M_{\text{Supply}}$, is largest represents the point of maximum profit. In the Figure 2.1 are represented the mempool demand curve and the space supply curve, calculated using our analitic system.

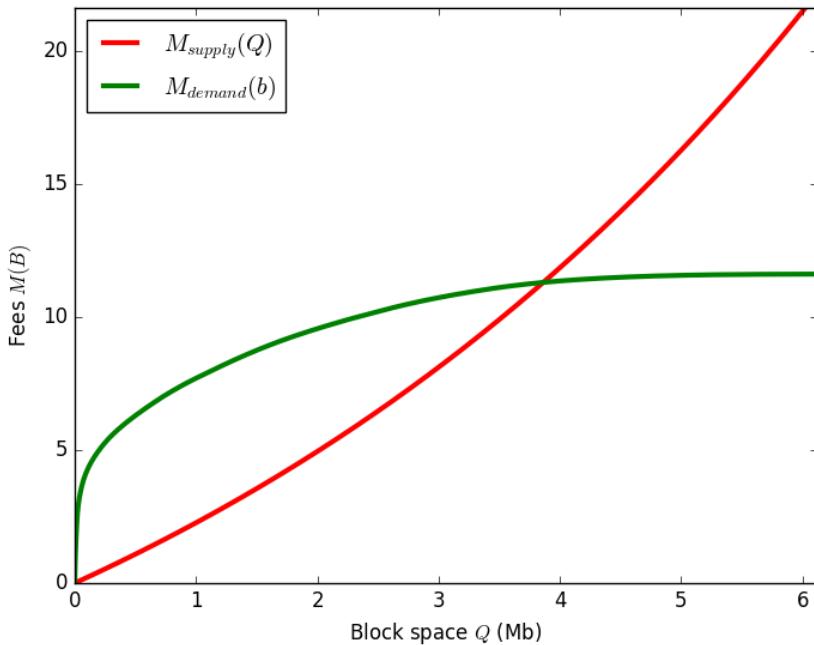


Figure 2.1: Maximizing the profit of a miner evaluating the unconfirmed transactions from the Bitcoin blockchain. Mempool demand, $M_{\text{demand}}(b)$, and space supply, $M_{\text{Supply}}(Q)$, curve are represented with our analytics system, and the block space Q^* for a maximum profit is around 1 Mb. There the gap between the two curves is bigger.



3

Technical Background

/4

Blockchain Analytics System

4.1 Blockchain Data Sources

4.2 System Architecture

4.2.1 Data Retrieval

4.2.2 Data Manipulations

4.2.3 Methods

4.3 Version Control

/5

Blockchain Observations

- 5.1 Blockchain Growth**
- 5.2 Retrieval Block Time**
- 5.3 Block Analysis**
- 5.4 Bandwidth**
- 5.5 Block Fee**
- 5.6 Models**

/6

Conclusions

6.1 Discussion

6.2 Future Implementation

6.3 Comments

show bibliography [41], [49], [21], [29], [38], [27], [10], [15], [39], [43], [33], [35], [17], [31], [47], [37], [1], [6], [24], [7], [45], [22], [44], [4], [8], [9], [16], [19], [18], [36], [20], [12], [25], [30], [32], [13], [2], [3], [42] [5], [46], [28], [34], [11], [23], [40], [14], [48].

References

- [1] Bitcoin api's, api-v1-client-python. <https://github.com/blockchain/api-v1-client-python>.
- [2] Bitcoin mining hashing rate. <https://blockchain.info/charts/hash-rate>.
- [3] Bitcoin nodes. <https://bitnodes.21.co/>.
- [4] Bitocoin's blockchain website. <https://blockchain.info>.
- [5] Construct a linear, no-fork, best version of the bitcoin blockchain. <https://github.com/bitcoin/bitcoin/tree/master/contrib/linearize>.
- [6] Ethereum api's, pyethereum. <https://github.com/ethereum/pyethereum>.
- [7] Ethereum wiki/patricia tree. <https://github.com/ethereum/wiki/wiki/Patricia-Tree>.
- [8] Ethereum's blockchain analysis website. <https://etherscan.io>.
- [9] Ethereum's blockchain website. <https://etherchain.org/>.
- [10] Ethereum's white paper. <https://github.com/ethereum/wiki/wiki/White-Paper>.
- [11] How visa make money. <https://revenuesandprofits.com/how-visa-makes-money-understanding-visa-business-model/>.
- [12] Matplotlib for data plotting. <https://matplotlib.org>.
- [13] Mining hardware comparison. https://en.bitcoin.it/wiki/Mining_hardware_comparison.

- [14] pandas: Python Data Analysis Library. <http://pandas.pydata.org/>, 2012.
- [15] Ethereum foundation - the solidity contract-oriented programming language. Technical report, <https://solidity.readthedocs.io/en/develop/>, 2014.
- [16] Bitcoin mining process. <http://bitcoinminer.com/>, 2015.
- [17] Bitcoin website – mining. <https://www.bitcoinmining.com>, 2016.
- [18] Ethereum project – website. <https://www.ethereum.org>, 2016.
- [19] tradeblock.com – analysis on the blockchain. <https://tradeblock.com>, 2016.
- [20] Alfred V. Aho and Jeffrey D. Ullman. *Foundations of Computer Science*. Computer Science Press, Inc., New York, NY, USA, 1992.
- [21] Adam Back. Hashcash - a denial of service counter-measure. Technical report, 2002.
- [22] Paul Baran. *On Distributed Communications: Introduction to Distributed Communication Networks*. The Rand Corporation, 1964.
- [23] Massimo Bartoletti, Andrea Bracciali, Stefano Lande, and Livio Pompianu. A general framework for bitcoin analytics. *CoRR*, abs/1707.01021, 2017.
- [24] Georg Becker. *Merkle Signature Schemes, Merkle Trees and Their Cryptanalysis*. 2008.
- [25] Kyle Croman, Christian Decker, Ittay Eyal, Adem Efe Gencer, Ari Juels, Ahmed Kosba, Andrew Miller, Prateek Saxena, Elaine Shi, Emin Gün Sirer, Dawn Song, and Roger Wattenhofer. *On Scaling Decentralized Blockchains*, pages 106–125. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.
- [26] C. Decker and R. Wattenhofer. Information propagation in the bitcoin network. In *IEEE P2P 2013 Proceedings*, pages 1–10, Sept 2013.
- [27] Kevin Delmolino, Mitchell Arnett, Ahmed Kosba, Andrew Miller, and Elaine Shi. *Step by Step Towards Creating a Safe Smart Contract: Lessons and Insights from a Cryptocurrency Lab*, pages 79–94. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.
- [28] Jacob Donnelly. Bitcoin network still backlogged with tens of thousands

- of unconfirmed transactions, causing delays. *Bitcoin Magazine*, July 2015.
- [29] Cynthia Dwork and Moni Naor. Pricing via processing or combatting junk mail. In *Proceedings of the 12th Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO '92, pages 139–147, London, UK, UK, 1993. Springer-Verlag.
 - [30] Ittay Eyal, Adem Efe Gencer, Emin Gün Sirer, and Robbert van Renesse. Bitcoin-ng: A scalable blockchain protocol. *CoRR*, abs/1510.02037, 2015.
 - [31] Rui Garcia, Rodrigo Rodrigues, and Nuno Preguiça. Efficient middleware for byzantine fault tolerant database replication. In *Proceedings of the Sixth Conference on Computer Systems*, EuroSys '11, pages 107–122, New York, NY, USA, 2011. ACM.
 - [32] Begnaud Francis Hildebrand. *Introduction to Numerical Analysis: 2Nd Edition*. Dover Publications, Inc., New York, NY, USA, 1987.
 - [33] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages, and Computation (3rd Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2006.
 - [34] Nicolas Houy. The economics of bitcoin transaction fees. Working Papers 1407, Groupe d'Analyse et de Théorie Economique (GATE), Centre national de la recherche scientifique (CNRS), Université Lyon 2, Ecole Normale Supérieure, 2014.
 - [35] Håvard D Johansen, Robbert van Renesse, Ymir Vigfusson, and Dag Johansen. Fireflies: A secure and scalable membership and gossip service. *ACM Transactions on Computer Systems (TOCS)*, 33(2):5:1–5:32, May 2015.
 - [36] Craig Larman. *Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development (3rd Edition)*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2004.
 - [37] Aldelir Fernando Luiz, Lau Cheuk Lung, and Miguel Correia. Mitra: Byzantine fault-tolerant middleware for transaction processing on replicated databases. *SIGMOD Rec.*, 43(1):32–38, May 2014.
 - [38] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, and Aquinas Hobor. Making smart contracts smarter. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, pages 254–269, New York, NY, USA, 2016. ACM.

- [39] Loi Luu, Jason Teutsch, Raghav Kulkarni, and Prateek Saxena. Demystifying incentives in the consensus computer. In *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security, CCS '15*, pages 706–719, New York, NY, USA, 2015. ACM.
- [40] Malte Möser and Rainer Böhme. *Trends, Tips, Tolls: A Longitudinal Study of Bitcoin Transaction Fees*, pages 19–33. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.
- [41] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system,” <http://bitcoin.org/bitcoin.pdf>, 2008.
- [42] Peter R. Rizun. A transaction fee market exists without a block size limit. Technical report, 2015.
- [43] Chaitya B. Shah and Drashti R. Panchal. Secured hash algorithm-1: Review paper. Technical report, Indus Institute of Technology and Engineering, Gujarat Technological University, 2014.
- [44] William Stallings. *Cryptography and Network Security: Principles and Practice*. Pearson Education, 3rd edition, 2002.
- [45] M. Swan. *Blockchain: Blueprint for a New Economy*. O'Reilly Media, 2015.
- [46] Enrico Tedeschi. Paying for bandwidth in blockchain internet applications. Technical report, UiT Arctic University of Norway, 2016.
- [47] Ben Vandiver, Hari Balakrishnan, Barbara Liskov, and Samuel Madden. Tolerating Byzantine Faults in Transaction Processing Systems Using Commit Barrier Scheduling. In *ACM SOSP*, Stevenson, WA, October 2007.
- [48] Michael Waskom, Olga Botvinnik, drewokane, Paul Hobson, David Yaroslav Halchenko, Saulius Lukauskas, John B. Cole, Jordi Warmenhoven, Julian de Ruiter, Stephan Hoyer, Jake Vanderplas, Santi Villalba, Gero Kunter, Eric Quintero, Marcel Martin, Alistair Miles, Kyle Meyer, Tom Augspurger, Tal Yarkoni, Pete Bachant, Mike Williams, Constantine Evans, Clark Fitzgerald, Brian, Daniel Wehner, Gregory Hitz, Erik Ziegler, Adel Qalieh, and Antony Lee. *seaborn: v0.7.1* (june 2016), June 2016.
- [49] Dr. Gavin Wood. Ethereum: A secure decentralised generalised transaction ledger. Technical report, 2014.



Terminology

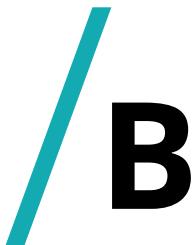
RLP: Stands for recursive length prefix. It is a serialization method for encoding arbitrary structured binary data (byte arrays).

KEC-256: Another serialization method generating a 256-bit hash.

full node: A full node in a decentralized digital currency peer-to-peer network, is a node that stores and processes the entirety of every block, storing locally the entire size of the blockchain.

light node: A light node in a decentralized digital currency peer-to-peer network, is a node that only stores the part of the blockchain it needs.

satoshi: Unit of the Bitcoin currency. 100,000,000 satoshi are 1 BTC (Bitcoin).



List of Symbols

t_B	transaction approved in a block B .
t_{in}	transaction input in bitcoin (\mathbb{B}). All the money sent.
t_{ou}	transaction output (\mathbb{B}). All the money received.
t_f	transaction fee (\mathbb{B}). $t_{in} - t_{ou}$.
t_q	transaction size, in bytes.
t_l	commit latency of a single transaction. $B_{epoch} - t_{epoch}$.
\mathcal{T}	expected block interval time (~ 10 min)
\mathbb{P}_{orphan}	probability that given a block is orphaned.
τ	block solution propagation time, we consider a $\tau = 10$ seconds

according to Decker [26].

η	cost per hash.
$\langle \Pi \rangle$	expectation value of a miner's profit per block.
$\langle V \rangle$	expectation value of a miner's revenue per block.
$\langle C \rangle$	expectation value of a miner's hashing cost per block.
R	block reward, currently at 12.5 B.
h	miner's individual hash rate.
H	total hash rate of Bitcoin network.
Q	block size or block space in bytes.
Q^*	the block size that maximizes the miner's expected profit.
ρ	fee density, or the price per byte for block space.
M	money, bitcoin (B).
$M_{\text{demand}}(b)$	partial sum of the b transaction fees in mempool in order of descending fee density.
$M_{\text{supply}}(Q)$	miner's cost due to orphaning to produce a certain block size Q .
\mathcal{N}	the set of transactions in a miner's mempool.
n	number of transactions in a miner's mempool.

B single block.

B_t transaction root that links to every transaction in a block B .



List^{ing}

Listing C.1: Smart contract transaction code in Ethereum.

```
1 function transfer(address _to, uint256 _value) {
2     /* Add and subtract new balances */
3     balanceOf[msg.sender] -= _value;
4     balanceOf[_to] += _value;
5 }
```

Listing C.2: Example of a smart contract that rewards users who solve a computational puzzle [?].

```
1 contract Puzzle {
2     address public owner ;
3     bool public locked ;
4     uint public reward ;
5     bytes32 public diff ;
6     bytes public solution ;
7
8     function Puzzle () // constructor {
9         owner = msg. sender ;
10        reward = msg . value ;
11        locked = false ;
12        diff = bytes32 (11111); // pre - defined
13            difficulty
14    }
15
16    function (){ // main code , runs at every
17        invocation
18        if ( msg. sender == owner ){ // update reward
19            if ( locked )
20                throw ;
21            owner . send ( reward );
22            reward = msg . value ;
23        }
24        else
25            if ( msg . data . length > 0){ // submit a
26                solution
27            if ( locked ) throw ;
28            if ( sha256 (msg. data ) < diff ){
29                msg. sender . send ( reward ); // send
                    reward
                solution = msg. data ;
                locked
            }}}
```

Listing C.3: Application usage.

```

1 Usage: observ.py -t number
2   -h | --help      : usage
3   -i              : gives info of the blockchain in
4     the file .txt
5   -t number       : add on top a number of blocks.
6     The blocks retrieved will be the most recent
7     ones. If the blockchain growth more than the
8     block requested do -u (update)
9   -e number       : append blocks at the end of the .
10    txt file. Fetch older blocks starting from the
11    last retrieved
12   -P              : plot all
13   -p start [end] : plot data in .txt file in a
14     certain period of time, from start to end. If
15     only start then consider from start to the end
16     of the .txt file
17   -R              : plot the regression and the
18     models that predict the blockchain
19   -r start [end] : plot the regression and the
20     models in a certain period of time, from start
21     to end. If only start then consider from start
22     to the end of the .txt file
23   -u              : update the local blockchain to
24     the last block created

```

Listing C.4: Block object represented in Python according to api-v1-client-python to retrieve data on the blockchain. The function get_block() will return an object of this type [1].

```
1 class Block:
2     def __init__(self, b):
3         self.hash = b['hash']
4         self.version = b['ver']
5         self.previous_block = b['prev_block']
6         self.merkle_root = b['mrkl_root']
7         self.time = b['time']
8         self.bits = b['bits']
9         self.fee = b['fee']
10        self.nonce = b['nonce']
11        self.n_tx = b['n_tx']
12        self.size = b['size']
13        self.block_index = b['block_index']
14        self.main_chain = b['main_chain']
15        self.height = b['height']
16        self.received_time = b.get('received_time', b['time'])
17        self.relayed_by = b.get('relayed_by')
18        self.transactions = [Transaction(t) for t in b['tx']]
19        for tx in self.transactions:
20            tx.block_height = self.height
```

Listing C.5: Transaction object represented in Python according to api-v1-client-python to retrieve data on the blockchain. The function get_transaction() will return an object of this type [1].

```

1  class Transaction:
2      def __init__(self, t):
3          self.double_spend = t.get('double_spend', False)
4          self.block_height = t.get('block_height')
5          self.time = t['time']
6          self.relayed_by = t[' relayed_by ']
7          self.hash = t['hash']
8          self.tx_index = t['tx_index']
9          self.version = t['ver']
10         self.size = t['size']
11         self.inputs = [Input(i) for i in t['inputs']]
12         self.outputs = [Output(o) for o in t['out']]
13
14         if self.block_height is None:
15             self.block_height = -1

```

Listing C.6: Json object returned from the method *get_block()* in the Bitcoin API class blockexplorer.py [1]

```

hash : str
version : int
previous_block : str
merkle_root : str
time : int
bits : int
fee : int
nonce : int
n_tx : int
size : int
block_index : int
main_chain : bool
height : int
received_time : int
relayed_by : string
transactions : array of Transaction objects

```

Listing C.7: Structure of the latest block retrieved. The function `get_latest_block()` will return an object with this structure.

```

1 class LatestBlock:
2     def __init__(self, b):
3         self.hash = b['hash']
4         self.time = b['time']
5         self.block_index = b['block_index']
6         self.height = b['height']
7         self.tx_indexes = [i for i in b['txIndexes']]
```

Listing C.8: Collecting data starting from the last element in the `blockchain.txt` file.

```

1 earliest_hash = get_earliest_hash()
2 get_blockchain(n, earliest_hash)
3
4 def get_blockchain(n, hash = None):
5     [...]
6     if (hash): # start the retrieval from the hash
7         append_end = True # in that way the
8             write_blockchain method knows that has to
9                 append blocks and not write them at the
10                  beginning
11     last_block = blockexplorer.get_block(hash)
12     [...]
13
14 def get_earliest_hash():
15     hash_list = get_list_from_file("hash") # method to
16         collect data from blockchain.txt file having
17             as attribute "hash"
18     length = len(hash_list)
19     earliest_hash = hash_list[length - 1]
20     return earliest_hash
```

Listing C.9: Calling blockchain.info through python API and retrieving part of the blockchain.

```

1 from blockchain import blockexplorer
2 # get the last block
3 last_block = blockexplorer.get_latest_block()
4 hash_last_block = last_block.hash
5
6 # current block now is the last block
7 current_block = blockexplorer.get_block(
    hash_last_block)

```

Listing C.10: How read bandwidth is calculated, using the function `datetime.now()` before and after the API call.

```

1 start_time = datetime.datetime.now() # -----
2 current_block = blockexplorer.get_block(
    current_block.previous_block)
3 end_time = datetime.datetime.now() # -----
4 time_to_fetch = end_time - start_time
5 time_in_seconds = get_time_in_seconds(time_to_fetch)
6
7 #latency
8 fetch_time_list.append(time_in_seconds)
9
10 # calculate Bandwidth with MB/s
11 block_size = float(current_block.size) / 1000000
12 bandwidth = block_size / time_in_seconds
13 bandwidth_list.append(bandwidth)

```

Listing C.11: Function that get the average write bandwidth of the block, calculating the time for each transaction to be visible in the public ledger of data.

```
1 def get_avg_transaction_time(block):
2     # take transactions the block
3     transactions = block.transactions
4
5     # get block time -- when it is visible in the
6     # blockchain, so when it was created
7     block_time = block.time
8
9     # list of the creation time for all the
10    # transaction in the block
11    transactions_time_list = []
12
13    # list of the time that each transaction needs
14    # before being visible in the blockchain
15    time_to_be_visible = []
16
17    for t in transactions:
18        transactions_time_list.append(float(t.time))
19
20        for t_time in transactions_time_list:
21            time_to_be_visible.append(float(block_time -
22                t_time))
23
24    average_per_block = sum(time_to_be_visible) / len(
25        time_to_be_visible)
26
27    return average_per_block
```

Listing C.12: How the file.write is performed in the blockchain analytics system. Differences with add and append.

```

1 if (append):
2     for i in range(n):
3         file.write("block_informations")
4
5 else: # add on top
6     hash_list_in_file = get_list_from_file("hash")
7     first_hash = hash_list_in_file[0]
8     elements = len(hash_list_in_file)
9     last_hash = hash_list_in_file[elements - 1]
10    met_first = False
11    with io.FileIO(file_name, "a+") as file:
12        file.seek(0) # place at the beginning of the
13        file
14        existing_lines = file.readlines() # read the
15        already existing lines
16        file.seek(0)
17        file.truncate() # delete all the file
18        file.seek(0)
19        i = 0
20        while (i < n):
21            if (first_hash == hash[i]):
22                met_first = True
23            while((met_first == False) and (i < n)):
24                # append on top
25                file.write("block_informations")
26                i = i + 1
27                if ((i < n) and (first_hash == hash[i])):
28                    met_first = True
# when the block retrieved meets the one
# already in the blockchain write the old file
file.writelines(existing_lines)

```

Listing C.13: Changing all the values inside a Python list in one code line.

```

1 # size_list from byte to MB
2 size_list[:] = [x / 1000000 for x in size_list]
3 # time_list from seconds in minutes
4 time_list[:] = [x / 60 for x in time_list]

```

Listing C.14: Method that allows, given an attribute present in the blockchain.txt file, to create a list containing informations only about this quality using *regular expressions*.

```
1 hash_list = get_list_from_file("hash")
2
3 def get_list_from_file(attribute):
4     list_to_return = []
5     if (os.path.isfile("blockchain.txt")):
6         # open the file and read in it --
7         blockchain_file
8         with open("blockchain.txt", "r") as
9             blockchain_file:
10                for line in blockchain_file:
11                    # regular expression that puts in a list the
12                    # line just read: eg. ['hash', '<block_hash
13                    >']
14                    list = re.findall(r"\w'+'\w'", line)
15                    # list[0] --> contains the attribute
16                    # list[1] --> contains the value
17                    if ((list) and (list[0] == attribute)):
18                        list_to_return.append(list[1])
19
20 return list_to_return
```

Listing C.15: Generation of the growing size and time lists. Calculated following the Equations ??, ??.

```

1 def create_growing_time_list(time_list):
2     reversed_time_list = time_list[::-1]
3     time_to_append = 0
4     previous_time = 0
5     growing_time_list = []
6     growing_time_list.append(previous_time)
7     for time_el in reversed_time_list:
8         time_to_append = (float(time_el) / (60 * 60)) +
9             previous_time # time in hours
10        growing_time_list.append(time_to_append)
11        previous_time = time_to_append
12    return growing_time_list
13
14 def create_growing_size_list(size_list):
15     reversed_size_list = size_list[::-1]
16     growing_size_list = []
17     value_to_append = 0
18     size_back = 0
19     growing_size_list.append(value_to_append)
20     for size_el in reversed_size_list:
21         value_to_append = size_el + size_back
22         growing_size_list.append(value_to_append)
23         size_back = value_to_append
24     return growing_size_list

```

Listing C.16: Example of a polynomial interpolation of two lists using NumPy libraries.

```

1 import numpy as np
2
3 model = np.polyfit(list1, list2, deg) # polynomial
4             interpolation between list1 and list2 with the
5             degree of the return polynomial
6 # deg = 1 --> linear interpolation
7 # deg = 2 --> quadratic
8 # deg = 3 --> cubic
9 # ...
8 predicted = np.polyval(model, list1)
9 plt.plot(list1, predicted, 'b-', label="pol_interp")

```

Listing C.17: Check the status of the local blockchain, True if valid, False if not.

```
1 def check_blockchain():
2     check = True
3     list = get_list_from_file("height")
4     number = int(list[0])
5     length_list = len(list)
6     for i in range(length_list):
7         if (number != int(list[i])):
8             check = False
9         number = number - 1
10    return check
```