BLOCKASTICS

Stochastic models for blockchain analysis

Pierre-O Goffard

March 23, 2022

Chapter 1

Introduction

A blockchain is a distributed ledger made of a sequence of blocks maintained by achieving consensus among a number of nodes in a Peer-to-Peer network. The blockchain technology has attracted a lot of interest after the advent of the bitcoin cryptocurrency in 2008, see Nakamoto [2008]. Since then, the blockchain concept has been used to develop decentralized systems to store and maintain the integrity of time-stamped transaction data across peer-to-peer networks. Besides the creation of a digital currency, blockchain applications include the sharing of IT resources, the registration of authentication certificate or the implementation of smart contracts.

A blockchain is

- Decentralized as it is maintained by a network. Nodes can be light or full nodes. Light nodes are blockchain users that broadcast transactions, full nodes are in charge of verifying and recording the transactions, see Figure 1.1.

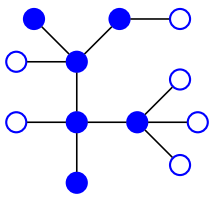


Figure 1.1: A network made of full nodes (blue) and light nodes (white)

  - A local copy is stored by each full node which grants security
  - The governance is not handled by a central authority

- Public or private. In public blockchain anyone can access the data, in private blockchain reading access is restricted.
- permissionned or permissionless. In permissionless blockchain, anyone can join the network as a full node.

- Immutable. Altering the information written in the blockchain is made difficult if not impossible.
- Incentive compatible. The process of reaching consensus is costly to the full nodes who must be compensated for their hard work.

The consensus protocols, at the core of the blockchain technologies, are the focus of these lecture notes. The goal is to evaluate consensus protocol according to three dimensions

1. Efficiency: The amount of data being processed per time unit
2. Decentralization: The fairness of the distribution of the decision power among the nodes
3. Security: The likelihood of a successful attack on the blockchain

Because consensus protocols involve random components, stochastic modelling is required to assess a blockchain system within the Efficiency/Decentralization/Security trilemma in Figure 1.2. As it is hard to improve one dimension without negatively impacting the other two, trade-offs

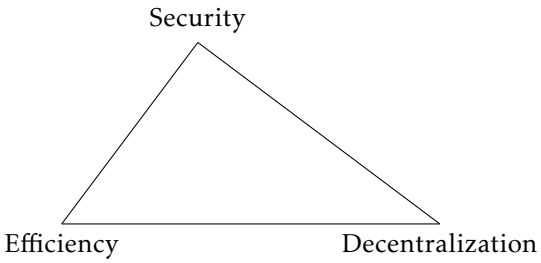


Figure 1.2: The blockchain trilemma

must be made. We will see how to use classical models of applied probability, including urn, epidemic, graph, queue and risk models, to provide numerically tractable indicators to quantify the efficiency, decentralization and security of blockchain systems. These indicators will then allow us to carry out sensitivity analysis with respect to the model parameters to optimize and improve blockchain implementations.

The main application of blockchain systems today is undoubtedly cryptocurrencies, the most well known of which being the bitcoin introduced by Nakamoto [2008]. Public and permissionless blockchain, like the bitcoin one, must be associated to a cryptocurrency. Indeed, to add a block to the bitcoin blockchains the full nodes compete to solve a cryptrographic puzzle using brute force search algorithm. The first node (referred to is miner) who finds a solution, appends the next block and collects a reward expressed in cryptocurrency. Assuming this reward is worth something, it offsets the operational cost, essentially the electricity consumed to run the computers 24/7. A cryptocurrency must be equipped with following features

1. No central authority (Decentralized network)

2. Ledger to record all the transactions and coin ownership (the blockchain)
3. A coin generation process (block finding reward)

   ↪ It creates an incentive compatible system to the full nodes
4. Ownership can be proved cryptographically, a wallet is secured with a public/private key system
5. Transactions can be issued by an entity proving ownership of the cryptographic unit through the private key
6. The system cannot process more than one transaction associated to the same cryptographic unit. It must be robust to double spending attack in which a fraudster is issuing two conflicting transactions to recover the funds she already spent

This charcaterization is given by Lansky [2018]. Cryptocurrencies draw their fundamental value from the fact that they

- provide transaction anonymity
- provide a reliable currency in certain regions of the world
- permit money transfer worldwide at low fare
- do not require a thrusted third party

Decentralized finance aims at extending the promises of bitcoin to An important implication of this architecture is disintermediation, it creates an environment where multiple parties can interact directly and transparently. Blockchain is therefore immediately relevant to banks and financial institutions which incur huge middlemen costs in settlements and other back office operations. Decentralized finance (DeFi) offers a new financial architecture that is non-custodial, permissionless, openly auditable, pseudo-anonymous and with potential new capital efficiencies. It extends the promise of the original bitcoin whitepaper (Nakamoto, 2008) of non-custodial transaction to more complex financial operations.

Blockchain is a research topic of interest to many communities outside of the applied mathematics one

The design of consensus protocol is

These notes are organized as follows. Chapter 2 presents the various consensus algorithms. Chapter 3 focuses on the security aspects. In Chapter 3, we take a look at decentralization in Chapter 4. We close on efficiency with Chapter 5.

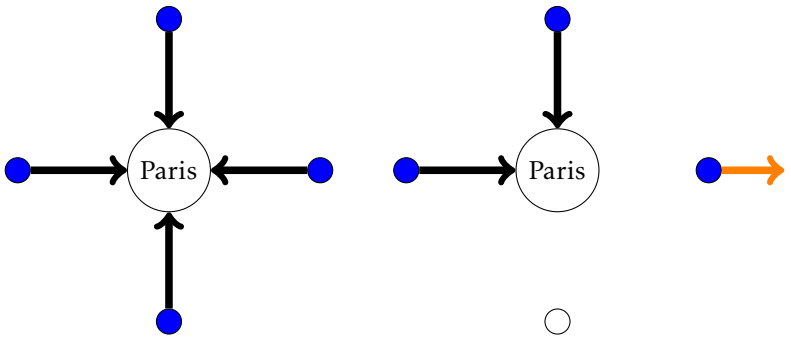
The topic of blockchain is of primary interest to computer scientists working on peer-to-peer networks and distributed algorithm. The problem of reaching consensus inside peer-to-peer networks is a classical problem in computer science. A consensus protocol just take advantage of the limited resources of the network which includes

- bandwidth
- computational power
- storage

The most natural solution is to proceed to a majority vote via a system of message exchange. It was proposed a long time ago to solve "The Byzantine general problem" as framed by Lamport et al. [1982] in an abstract way. The issue is that exchanging messages inside a peer-to-peer network that can grow very large is not a practical solution. The colossal number of messages is prohibitive, it leads to communication overhead and the failure of some nodes by denial of service.

The goal is to find an algorithm which allows the node of the networks to agree despite the presence of crashing nodes and faulty nodes (also referred to as Byzantine node)

A group of generals from the Byzantine army is surrounding an enemy city. Communicating only by messenger, they must agree on a common battle plan. There may be traitors who will attack instead of retreat or non responding generals who will do nothing. For the project to be successful a majority of the general must either retreat or attack. The problem then reduces to finding an algorithm to ensure that the loyal generals reach an agreement. The problem is illustrated on Figure 1.3.



(a) Coordinated attack

(b) Failed attack due to one traitor and one non-responding general

Figure 1.3: Illustration of the BYzantine general problem

In a blockchain system, we have a large network of nodes that broadcast transactions which corresponds to pieces of information that will be written in the blockchain. light nodes, full nodes consensus and write information.

- Computer science
- Economics
- Applied math and operations research

Chapter 2

Consensus protocol

The problem of reaching consensus within a Peer-to-peer network is a very old problem in computer science. An obvious solution is to rely on a majority vote. This is the solution proposed by SHostak and his co-author who, in passing, make a famous analogy with Byzantine generals trying to agree on a common battle plan. Here the battle plan corresponds to adding a new block with a set of transactions deemed valid and therefore agreeing on a common data history. A voting system inside a large network involves a colossal number of messages exchanged leading to the consumption of all the bandwidth, the failure of certain nodes by denial of service and delays in the synchronization of the network. Castro and Liskov's Practical Byzantine Fault Tolerance (PBFT) algorithm is the gold standard for practical implementation of a voting system within a peer-to-peer network. Despite these recent advances, such a system is not suitable for a network that can grow indefinitely. Bitcoin solved this scaling problem by proposing a system based on the election of a leader making unilateral decisions. The Proof of Work protocol appoints a leader based on its computing resources. Each node competes to solve a puzzle with a brute force search algorithm. The first node who is able to propose a solution append the next block. The search for a solution, referred to as mining, is associated with an operational cost borne by the nodes which is compensated by a reward expressed in the native blockchain cryptocurrency. The surge in cryptocurrency prices has led to a rush in block mining, leading to a major spike in the electricity consumption and electronic waste generation of blockchain networks. The blockchain network consumes as much electricity as countries the size of Thailand at the time of the writing. The need for a more restricted consensus protocol therefore becomes crucial. These remain based on the election of a leader but rely on other network resources. The Proof-of-Stake protocol sample the nodes with The proof of storage (no arm race, no electronic waste and useful like the file coin project)

2.1 Voting system

2.2 Leader system

- Public blockchain,
- operational cost,
- reward,
- incentive compatible
- Uses the scarce ressource of the network
  - Computational power (CPU, GPU)
  - Bandwidth
  - Storage space
  - Crypto coins

2.2.1 Proof-of-Work

2.2.2 Proof-of-Stake

Chapter 3

Security of blockchain systems

3.1 Double-spending in PoW

3.1.1 Random walk model

Double spending probability

Double spending time

3.1.2 Counting process model

Double spending probability

Double spending time

3.2 Blockwitholding in PoW

3.3 Nothing-at-stake in PoS

Chapter 4

Decentralization of blockchain system

4.1 Decentralization in PoS

Rich get richer? Polya's urn

4.1.1 Average stake own by each peer

4.1.2 Distribution of the stakes

4.2 Decentralization in PoW

4.2.1 Mining pools and reward systems

4.2.2 Mining pool risk analysis

Chapter 5

Efficiency of blockchain systems

5.1 A queueing model with bulk service

5.2 Latency and throughputs computation

Bibliography

Leslie Lamport, Robert Shostak, and Marshall Pease. The byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, pages 382–401, July 1982. URL https://www.microsoft.com/en-us/research/publication/byzantine-generals-problem/.

Jan Lansky. Possible state approaches to cryptocurrencies. *Journal of Systems Integration*, 9(1): 19–31, jan 2018. doi: 10.20470/jsi.v9i1.335.

S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Available at https://bitcoin.org/bitcoin.pdf, 2008. URL https://bitcoin.org/bitcoin.pdf.