

Stochastic Models for blockchain analysis

Blockchain fundamentals and applications

Pierre-O. Goffard

Université de Strasbourg
goffard@unistra.fr

18 octobre 2023

Outline

- 1 Introduction
- 2 Consensus protocol
- 3 Cryptocurrencies and Decentralized Finance (and insurance)

Plan

Introduction

1 Introduction

2 Consensus protocol

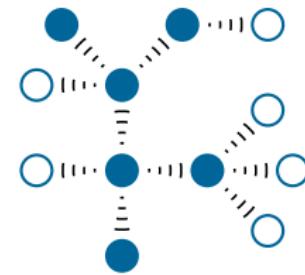
3 Cryptocurrencies and Decentralized Finance (and insurance)

Blockchain

Introduction

A data ledger made of a sequence of blocks maintained by a achieving consensus in a Peer-To-Peer network.

- Decentralized
- Public/private
- Permissionned/permissionless
- Immutable
- Incentive compatible



We will focus on public blockchain and their associated consensus protocol.

Blocks

Introduction

A block contains

- block height/ID
- Time stamp
- hash of the block
- hash of the previous block
- Set of transactions (data stored in the blockchain)

```
Block Height: 0
Block Hash: a52bea61a9f4131588cc101e8e1c731fa9f69f16934c5ab3a05a2134a42c13e0
Time:2021-07-12 10:03:04.812744
Block data: [{"sender": "Coinbase", "recipient": "Satoshi", "amount": 100, "fee": 1}]
Mined: False
Previous block hash: 0
-----
```

<https://www.blockchain.com/>

Consensus protocols

Introduction

The mechanism to make all the nodes agree on a common data history.

The three dimensions of blockchain systems analysis

1 Efficiency (Queueing theory)

- Throughputs
- Transaction confirmation time

2 Decentralization (Entropy)

- Fair distribution of the accounting right

3 Security (Insurance Risk Theory)

- Resistance to attacks



X. Fu, H. Wang, and P. Shi, "A survey of blockchain consensus algorithms : mechanism, design and applications," *Science China Information Sciences*, vol. 64, nov 2020.

Applications of blockchain : Cryptocurrency

Introduction

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



- Transaction anonymity
- No need for a trusted third party

Plan

Consensus protocol

1 Introduction

2 Consensus protocol

3 Cryptocurrencies and Decentralized Finance (and insurance)

Consensus protocol

Consensus protocol

Definition

Algorithm to allows the full nodes to agree on a common data history

It must rely on the scarce resources of the network

- bandwidth
- computational power
- storage (disk space)

Types of consensus protocols

Consensus protocol

1 Voting based

-  L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, pp. 382–401, July 1982.

-  Communication overhead
-  Denial of service

2 Leader based

- Proof-of-Work (computational power)
- Proof-of-Capacity and Proof-of-Spacetime (storage)
- Proof-of-Interaction (bandwidth)
- Proof-of-Stake (tokens)

Proof-of-Work

Consensus protocol

Objective

Elect a leader based on computational effort to append the next block.

What's inside a block ?

Consensus protocol

A block consists of

- a header
- a list of "transactions" that represents the information recorded through the blockchain.

The header usually includes

- the date and time of creation of the block,
- the block height which is the index inside the blockchain,
- the hash of the block
- the hash of the previous block.

Question

What is the hash of a block ?

Cryptographic Hash function

Consensus protocol

A function that maps data of arbitrary size (message) to a bit array of fixed size (hash value)

$$h : \{0,1\}^* \mapsto \{0,1\}^d.$$

A good hash function is

- deterministic
- quick to compute
- One way

→ For a given hash value \bar{h} it is hard to find a message m such that

$$h(m) = \bar{h}$$

- Collision resistant
 - Impossible to find m_1 and m_2 such that

$$h(m_1) = h(m_2)$$

- Chaotic

$$m_1 \approx m_2 \Rightarrow h(m_1) \neq h(m_2)$$

SHA-256

Consensus protocol

The SHA-256 function which converts any message into a hash value of 256 bits.

Example

The hexadecimal digest of the message

Happy new year

is

60a147c28568dc925c347bce20c910ef90f3774e2501ac63344f3411b6a6bf79

Hidden prediction

Consensus protocol



Matt Levine @matt_levine

Here is a SHA-256 hash of a prediction I am making:

64b70b0494580b278d7f1f551d482a3fb952a4b018b43090ffeb87b662d34847.



M. Levine, "The crypto story." Bloomberg business week, Oct. 2022.

Mining a block

Consensus protocol

```
Block Hash: 1fc23a429aa5aaf04d17e9057e03371f59ac8823b1441798940837fa2e318aaa
Block Height: 0
Time:2022-02-25 12:42:04.560217
Nonce:0
Block data: [ {'sender': 'Coinbase', 'recipient': 'Satoshi', 'amount': 100, 'fee': 0}, {'sender': 'Satoshi', 'recipient': 'Pierre-O', 'amount': 5, 'fee': 2}]
Previous block hash: 0
Mined: False
-----
```

Figure – A block that has not been mined yet.

Mining a block

Consensus protocol

The maximum value for a 256 bits number is

$$T_{\max} = 2^{256} - 1 \approx 1.16e^{77}.$$

Mining consists in drawing at random a nonce

$$\text{Nonce} \sim \text{Unif}(\{0, \dots, 2^{32} - 1\}),$$

until

$$h(\text{Nonce} | \text{Block info}) < T,$$

where T is referred to as the target.

Difficulty of the cryptopuzzle

$$D = \frac{T_{\max}}{T}.$$

Mining a block

Consensus protocol

If we set the difficulty to $D = 2^4$ then the hexadecimal digest must start with at least 1 leading 0

```
Block Hash: 0869032ad6b3e5b86a53f9dded5f7b09ab93b24cd5a79c1d8c81b0b3e748d226
Block Height: 0
Time:2022-02-25 13:41:48.039980
Nonce:2931734429
Block data: [{"sender": "Coinbase", "recipient": "Satoshi", "amount": 100, "fee": 0}, {"sender": "Satoshi", "recipient": "Pierre-O", "amount": 5, "fee": 2}]
Previous block hash: 0
Mined: True
-----
```

Figure – A mined block with a hash value having one leading zero.

The number of trials is geometrically distributed

- Exponential inter-block times
- Length of the blockchain = Poisson process

Bitcoin protocol

Consensus protocol

- One block every 10 minutes on average
- Depends on the hashrate of the network
- Difficulty adjustment every 2,016 blocks (\approx two weeks)
- Reward halving every 210,000 blocks

Check out <https://www.bitcoinblockhalf.com/>

Conflict resolution in blockchain

Consensus protocol

Fork

A fork arises when there is a disagreement between the nodes resulting in several branches in the blockchain.

LCR

The *Longest Chain Rule* states that if there exist several branches of the blockchain then the longest should be trusted.

In practice

- A branch can be considered legitimate if it is $k \in \mathbb{N}$ blocks ahead of its pursuers.
- Fork can be avoided when

$$\text{block appending time} > \text{propagation delay}$$

Bitcoin protocol

Consensus protocol

- One block every 10 minutes on average
- Depends on the hashrate of the network
- Difficulty adjustment every 2,016 blocks (\approx two weeks)
- Reward halving every 210,000 blocks

Check out <https://www.bitcoinblockhalf.com/>

Mining equipments

Consensus protocol

How it started

- CPU, GPU

How it is going

- Application Specific Integrated Chip (ASIC)
 - Network electricity consumption
 - E-Waste
 - Centralization issue

Proof of Stake

Consensus protocol

PoW is slow and ressource consuming. Let $\{1, \dots, N\}$ be a set of miners and $\{\pi_1, \dots, \pi_N\}$ be their share of cryptocoins.

PoS

- 1 Node $i \in \{1, \dots, N\}$ is selected with probability π_i to append the next block

Nodes are chosen according to what they own.

- Nothing at stake problem
- Rich gets richer ?
- <https://www.peercoin.net/>



F. Saleh, "Blockchain without waste : Proof-of-stake," *The Review of Financial Studies*, vol. 34, pp. 1156–1190, jul 2020.

Incentive mechanism

Consensus protocol

Participation to the network is costly to the nodes.

Native cryptocurrency

The nodes must be retributed for their hard work.

Plan

Cryptocurrencies and Decentralized Finance (and insurance)

1 Introduction

2 Consensus protocol

3 Cryptocurrencies and Decentralized Finance (and insurance)

TradFi Pain Points I

Cryptocurrencies and Decentralized Finance (and insurance)

1 Access barrier

TradFi Criteria ⇒ bank account

DeFi No Barrier

2 Centralization

TradFi Bank are recorder keeper ⇒ Cyber risk

DeFi Decentralized Ledger

3 High costs and intermediation

TradFi Transaction fees, account maintenance fees, wire transfer fees,...

DeFi Intermediaries ⇒ Smart contracts (Gas fees)

4 Slow transaction settlement

TradFi Cross border transactions ⇒ Takes day to be settled

DeFi Near instant settlement (30 min on the bitcoin blockchain)

5 Transparency and auditability

TradFi Difficulty to verify the accuracy of transactions and asset holding

DeFi Publicly available ledger and open source code

TradFi Pain Points II

Cryptocurrencies and Decentralized Finance (and insurance)

6 Censorship and restrictions

TradFi Asset in custody, censorship by governments

DeFi No interference of central authority

7 Global accessibility

TradFi Respect the operating hours

DeFi Operates 24/7

8 Fractional ownership

TradFi Real estate and art work are not divisible

DeFi Tokenized real world assets

9 Innovation and interoperability

TradFi Use outdated IT solutions

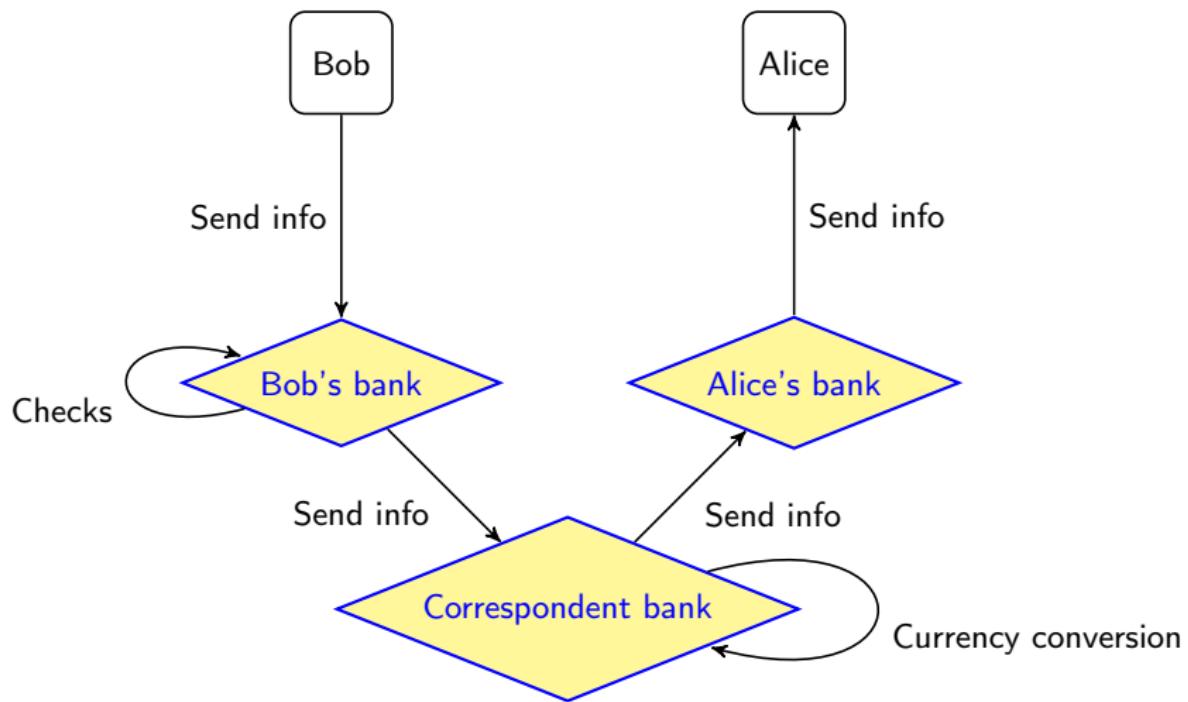
DeFi Interoperability between platforms and project



A. Lipton and A. Treccani, *Blockchain and Distributed Ledgers*.
WORLD SCIENTIFIC, apr 2021.

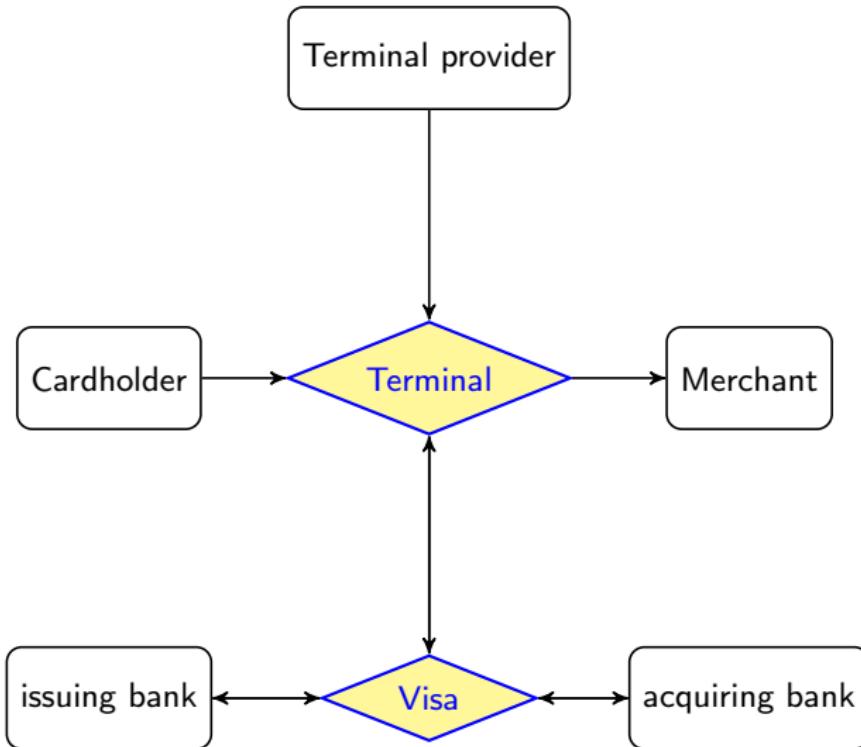
Wire transfer

Cryptocurrencies and Decentralized Finance (and insurance)



Credit card payments

Cryptocurrencies and Decentralized Finance (and insurance)



Cryptocurrencies

Cryptocurrencies and Decentralized Finance (and insurance)

- 1 No central authority (Decentralized network)
- 2 Ledger to record all the transactions and coin ownership (blockchain)
- 3 A coin generation process (block finding reward)
 - Incentive to the full nodes
- 4 Ownership can be proved cryptographically (wallet associated to a public/private key)
- 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 (double spending)



J. Lansky, "Possible state approaches to cryptocurrencies," *Journal of Systems Integration*, vol. 9, pp. 19–31, jan 2018.

Cryptocurrency implementation

Cryptocurrencies and Decentralized Finance (and insurance)

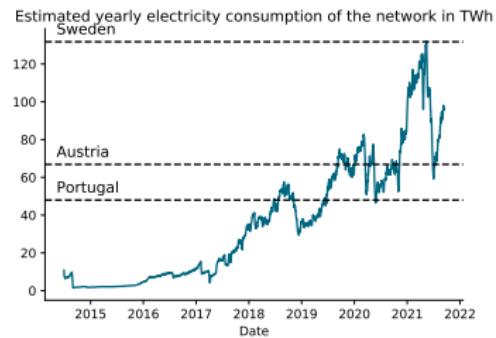
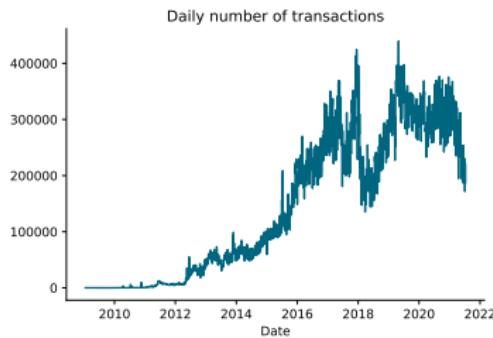
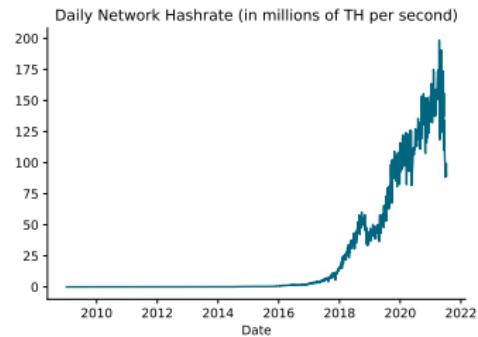
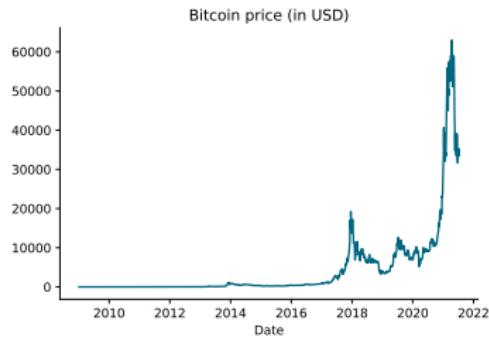
Blockchain parameters

- Consensus protocol (PoW or PoS)
 - ↳ Hash function (SHA-256 for Bitcoin and scrypt for LiteCoin)
 - ↳ Hybrid PoW/PoS (PeerCoin)
- Block generation time (<https://txstreet.com/v/eth-btc>)
 - ↳ every 10 minutes for Bitcoin
 - ↳ every 12 sec for Ethereum
- Block finding reward
 - ↳ Halved every 210,000 blocks in Bitcoin. It started at 50 BTC, is now 6.25 BTC
<https://www.bitcoinblockhalf.com/>
- Total coin supply
 - ↳ 21,000,000 in total for Bitcoin
- Transaction fees
 - ↳ GAS in Ethereum

These choices lead to the creation of multiple cryptocurrencies

Examples

Bitcoin and AltCoins (Ethereum, LiteCoin, DogeCoin, Ripple...), see https://en.wikipedia.org/wiki/List_of_cryptocurrencies



Decentralized finance

Cryptocurrencies and Decentralized Finance (and insurance)

DeFi extends the Bitcoin promises to more complex financial operations

- Collateralized lending
- Decentralized Exchange Platform
- Tokenized assets
- Fundraising vehicle (ICO, STO, ...)

Through the use of Smart Contract on the Ethereum blockchain.



S. M. Werner, D. Perez, L. Gudgeon, A. Klages-Mundt, D. Harz, and W. J. Knottenbelt,
“Sok : Decentralized finance (defi),” 2021.

Decentralized Exchange platforms

Cryptocurrencies and Decentralized Finance (and insurance)

Centralized Exchange

Binance, Coinbase, Kraken,...

Decentralized Exchange

Exchange where you trade one token for another through a smart contract.

Order book

Cryptocurrencies and Decentralized Finance (and insurance)

Table – Buy Orders

Price	Quantity	Total
9,950	10	99,500
9,900	5	49,500

Table – Sell Orders

Price	Quantity	Total
10,100	2	20,200
10,200	15	153,000

Lots of orders need to be issued

- Expensive
- Slow

Automated Market Makers (AMM)

Cryptocurrencies and Decentralized Finance (and insurance)

Instead of order book AMM relies on a simple mechanism based on

- liquidity pools
- Constant Function Market Makers

Decentralized insurance

Cryptocurrencies and Decentralized Finance (and insurance)

Parametric insurance

Compensation if a measurable quantity reaches a threshold

- Example : Flight delay insurance
 - [https://etherscan.io/address/
0xdc3d8fc2c41781b0259175bdc19516f7da11cba7](https://etherscan.io/address/0xdc3d8fc2c41781b0259175bdc19516f7da11cba7)
- Use smart contract and off-chain data through oracles
- Transparent and automatic

Blockchain as a research topic

Cryptocurrencies and Decentralized Finance (and insurance)

- Computer science
 - Peer-to-peer networks and consensus algorithm
 - Cryptography and security
- Economics
 - Game theory to study the incentive mechanism at play
 - Nature of the cryptoassets
- Operations research
 - Optimization of complex system
- Financial math
 - Valuation models for cryptoassets
- Machine learning and statistics
 - Open data
 - Interaction between blockchain users
 - (Social) network analysis
 - Clustering of public keys and addresses in the bitcoin blockchain.

Byzantine General problem

n generals must agree on a common battle plan, to either

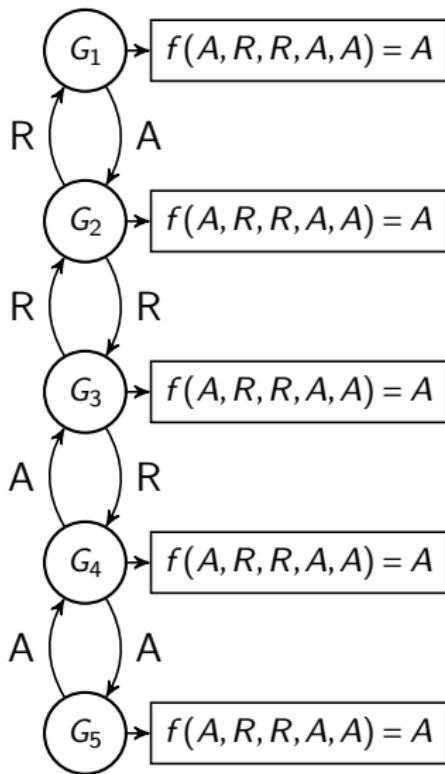
- Attack (A)
- Retreat (R)

Problem

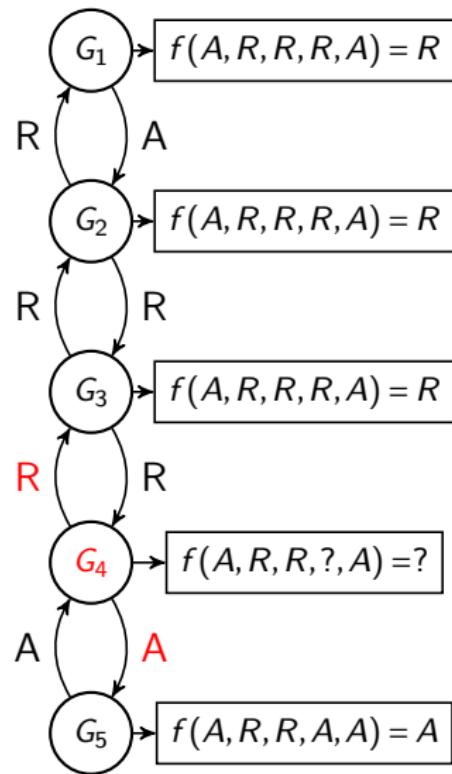
There are $m < n$ traitors among the generals

- 1 message $m(i,j)$ is sent to general j by general i
- 2 Consensus is reached as general j applies

$$f(\{m(i,j); i = 1, \dots, n\}) = \begin{cases} A, & \text{if } \sum_{i=1}^n \mathbb{I}_{m(i,j)=A} > n/2, \\ R, & \text{else.} \end{cases}$$



(e) No traitor



(f) One traitor

Figure – Majority vote with or without a traitor

Commanders and Lieutenants

One general is the commander while the others are the lieutenants

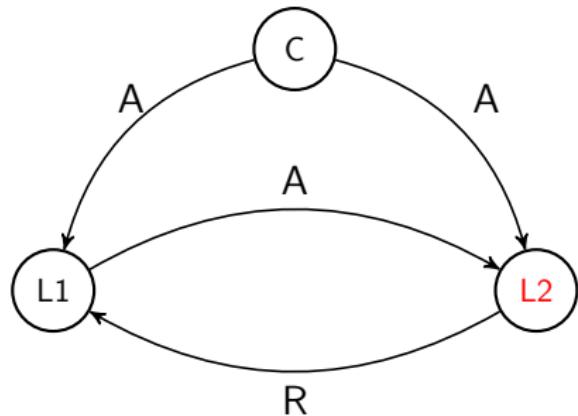
Objective

Design an algorithm so that the following conditions are met :

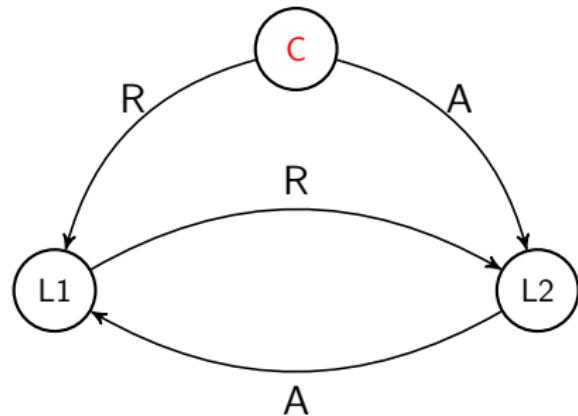
- C1 All the loyal lieutenants obey the same order
- C2 If the commanding general is loyal, then every loyal lieutenants obey the order he sends

Byzantine Fault Tolerance Theorem (Lamport et al.)

There are no solution to the Byzantine General problem for $n < 3m+1$ generals, where m is the number of traitors.



(a) Commander is loyal



(b) Commander is a traitor

Figure – Majority vote with or without a traitor

$n=4$ and $m=1$: Step 1

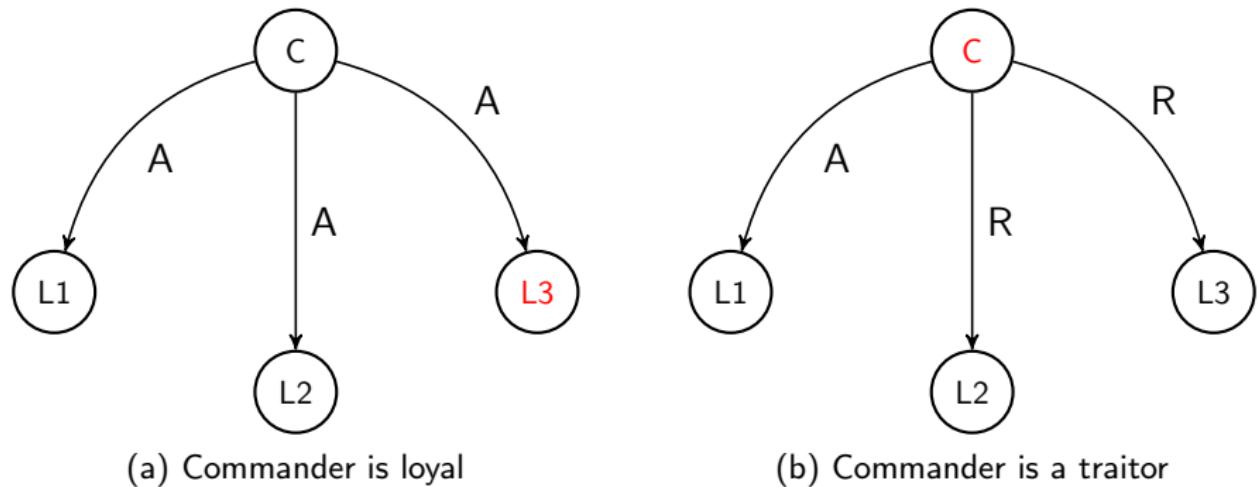
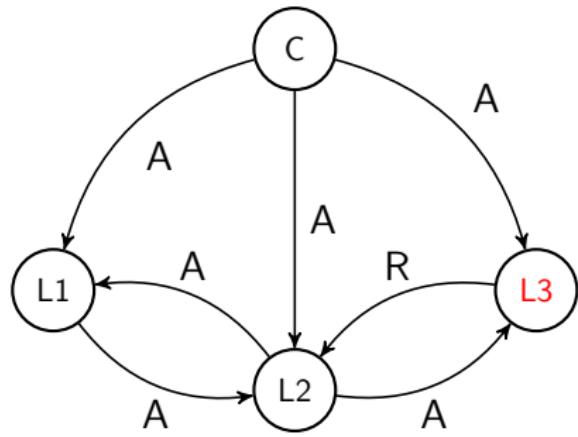
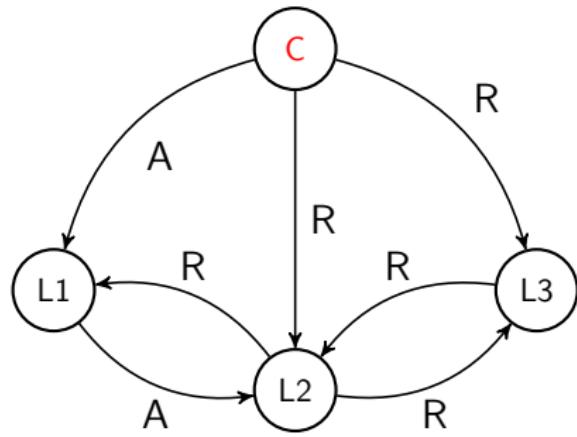


Figure – Illustration of the OM(m) algorithm in the case where $n=4$ and $m=1$.

$n=4$ and $m=1$: Step 2



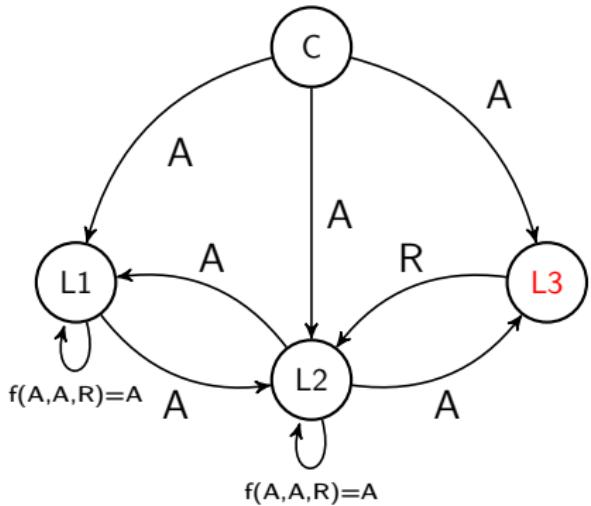
(a) Commander is loyal



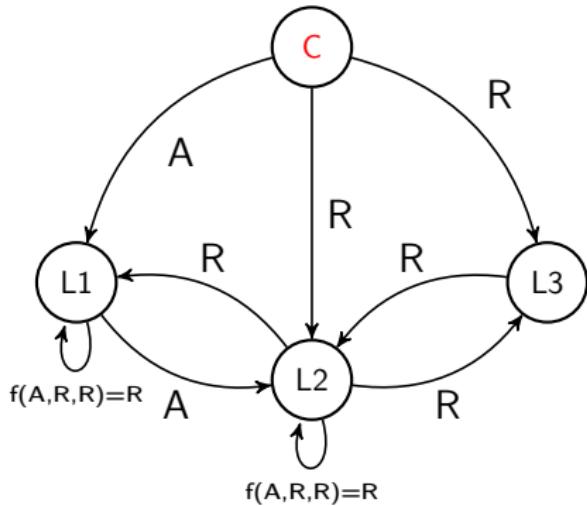
(b) Commander is a traitor

Figure – Illustration of the OM(m) algorithm in the case where $n=4$ and $m=1$.

n = 4 and *m = 1* : Step 3



(a) Commander is loyal, *C1* and *C2*



(b) Commander is a traitor, *C1*

Figure – Illustration of the OM(*m*) algorithm in the case where *n* = 4 and *m* = 1.

The problem with majority vote

The OM algorithm requires to send n^{m+1}

- ⚠ Communication overhead
- ⚠ Denial of service

Solution

Leader based protocols !

Proof-of-Work

Objective

Elect a leader based on computational effort to append the next block.