Sharding-Based Proof-of-Stake Blockchain Protocol: Security Analysis *

Abdelatif Hafid[1][0000−0002−1377−1387], Abdelhakim Hafid[1][0000−0001−8597−7344], and Adil Senhaji[2][0000−0002−0542−1360]

[1] University of Montreal, Montréal, QC H3T 1N8, CA
abdelatif.hafid@umontreal.ca, ahafid@iro.umontreal.ca
[2] Mizuho securities, New York, NY 10022, USA
adil.senhaji@mizuhogroup.com

Abstract. Blockchain technology has been gaining great interest from a variety of sectors including healthcare, supply chain, and cryptocurrencies. However, Blockchain suffers from its limited ability to scale (i.e., low throughput and high latency). Several solutions have been proposed to tackle this issue. In particular, sharding proved that it is one of the most promising solutions to Blockchain scalability. Sharding can be divided into two major categories: (1) Sharding-based Proof-of-Work (PoW) Blockchain protocols, and (2) Sharding-based Proof-of-Stake (PoS) Blockchain protocols. The two categories achieve a good performance (i.e., good throughput with a reasonable latency), but raise security issues. This article focuses on the second category. In this paper, we provide a probabilistic model to analyze the security of these protocols. More specifically, we compute the probability of committing a faulty block and measure the security by computing the number of years to fail. Finally, we evaluate the effectiveness of the proposed model via a numerical analysis.

Keywords: Blockchain scalability · Sharding · Security analysis · Proof-of-Stake · Practical Byzantine fault tolerance

1 Introduction

With the rise of Bitcoin [8], Blockchain has attracted significant attention from both industry and academia. More specifically, it has been adopted in different industry segments including healthcare [7], finance [9], and public sector [1]. However, the capacity of Blockchain to scale is very limited [4]. For example, in the case of cryptocurrencies, Bitcoin [8] handles between 3-7 transactions per second (tx/s), which is very limited compared to traditional payment systems (e.g., PayPal [10]). Several solutions were proposed to scale Blockchain. In particular, sharding has emerged as a promising solution [4]. Sharding consists of

* Supported by CIRRELT – Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation.

partitioning the network into sub-networks, called shards; all shards work in parallel to enhance the performance of the network. More specifically, each shard processes a sub-set of transactions instead of the entire network processing all the transactions. While sharding considerably improves scalability, it decreases the level of Blockchain security. More specifically, in sharding-based Blockchains, it is easy for a malicious user (aka, malicious/Byzantine node) to conquer and attack a single shard compared to the whole network; this attack is well-known as a shard takeover attack (aka, 1% attack) [5].

Blockchain networks are susceptible to sybil attacks by malicious nodes (called sybil nodes). Several consensus mechanisms (e.g., **P**roof-**o**f-**W**ork (PoW), **P**roof-**o**f-**S**take (PoS), and **p**ractical **B**yzantine **F**ault **T**olerance (pBFT)) have been proposed to defend against these sybil nodes. Sharding-based Blockchain protocols [12, 4] can be classified into two classes: sharding-based PoW and sharding-based PoS Blockchain protocols.

Recently, Hafid et al. [2, 5, 3] proposed mathematical models to analyze the security of sharding-based PoW Blockchain protocols. In this paper, we focus on the sharding-based PoS Blockchain protocols. We propose a probabilistic model to analyze the security of these protocols by computing the probability of committing a faulty block. Based on these probabilities, we calculate the number of years to fail for the purpose of quantifying and measuring the security of the network.

The rest of the paper is organized as follows. Section 2 presents the proposed probabilistic model. Section 3 presents numerical results and evaluates the proposed model. Section 4 concludes the paper.

2 Probabilistic Model

In this section, we propose a probabilistic model to analyze the security of sharding-based PoS Blockchain protocols. Generally, a sharding-based PoS Blockchain protocol (e.g., Incognito [6]) consists of one specific chain, called the beacon chain and many shard chains. The beacon chain synchronizes all shard chains in the network.

First, we start by computing the probability of a shard to commit a faulty block. Second, we calculate the probability of the beacon chain to commit a faulty block. Third, we compute the probability of all shards committing a faulty block. Finally, based on all these probabilities, we compute the probability of committing/adding a faulty block to the blockchain.

2.1 Notations & Definitions

Table 1 shows the list of symbols and variables that are used to describe the proposed probabilistic model.

Definition 1 (Faulty Block). *A faulty block is a block that contains fraudulent transactions.*

Table 1: Notations & Symbols.

Notation	Description
$\mathcal{N}$	Number of users
n	Committee size of a shard
n'	Committee size of the beacon chain
H	Number of honest validators in a shard
M	Number of malicious validators in a shard
$\mathcal{V}$	Number of validators in a shard ($\mathcal{V} = H + M$)
ζ	Number of shards
X	Random variable that computes the number of malicious nodes in the committee of a shard
H'	Number of honest validators in the beacon chain
M'	Number of malicious validators in the beacon chain
$\mathcal{V}'$	Number of validators in the beacon chain ($\mathcal{V}' = H' + M'$)
X'	Random variable that computes the number of malicious nodes in the committee of the beacon chain
r	Resiliency of the shard committee
r'	Resiliency of the beacon committee
R	Percentage of malicious validators in a shard chain
R'	Percentage of malicious validators in the beacon chain
$\mathcal{P}_f$	Probability of conquering the protocol
$\mathcal{P}$	Probability of a shard to commit a faulty block
$\mathcal{P}^{'}$	Probability of the beacon chain to commit a faulty block
$\mathcal{P}^{''}$	Probability of all shards committing a faulty block
$\mathcal{Y}_f$	Number of years to fail

Definition 2 (Conquering the Protocol). *A protocol is said to be conquered if the malicious nodes success to add a faulty block to the blockchain.*

Definition 3 (Committee Resiliency of a Shard). *The maximum percentage of malicious nodes that the committee of the shard chain can support whereas still being secure.*

Definition 4 (Committee Resiliency of the Beacon Chain). *The maximum percentage of malicious nodes that the committee of the beacon chain can support whereas still being secure.*

2.2 Architecture

In this section, we present a sample architecture of sharding-based PoS Blockchain protocols. This scheme is similar to that of Incognito [6].

Figure 1 shows a sample sharding-based PoS Blockchain protocol, which contains a single beacon chain and ζ shard chains. Shard chains produce blocks in parallel. All shard chains are synchronized by the beacon chain. More specifically, each shard has its own committee (i.e., a subset of the network nodes), which is randomly assigned by the beacon chain. Each shard chain processes a subset of

the transactions submitted to the network. When a shard block is created, the beacon committee verifies the block; if it is valid, it adds the block header to the beacon chain. Otherwise, it drops it and sends the proof to other shards for a vote to slash the misbehaving shard committee. Furthermore, in each epoch, the beacon chain shuffles committees, of the shards, to increase the security of the blockchain. For Incognito [6], when a new random number is generated, the beacon chain shuffles the committees; one epoch, for Incognito, corresponds to generating a new random number. This number is generated periodically in a round-robin fashion [6], [11].

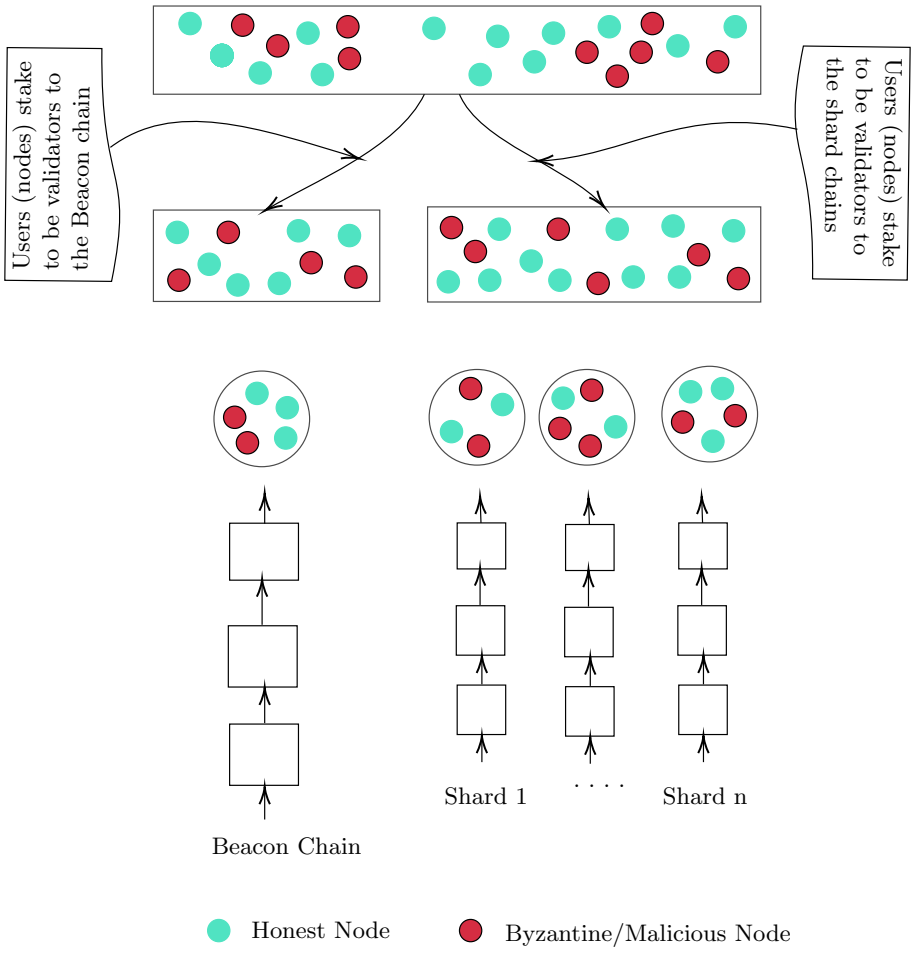


Fig. 1: A sharding-based PoS and pBFT Blockchain protocol.

2.3 Probability distributions

Generally, to add a faulty block to a sharding-based PoS Blockchain protocol (e.g., Incognito [6]), it must be confirmed by at least β ($0 < \beta < 1$; $\beta = r$) of the shard committee members, by at least β of the beacon committee members ($\beta = r'$), and by at least β of all shards' committees. For *Incognito* [6], $\beta = r = r' = \frac{2}{3}$.

Lemma 1. *The probability of a shard to commit a faulty block ($\mathcal{P}$) can be expressed as follows:*

$$P(X \geq \beta n) = \sum_{j=\beta n}^{n} \frac{\binom{M}{j}\binom{H}{n-j}}{\binom{\mathcal{V}}{n}} \tag{1}$$

Proof of **Lemma** 1 results directly from the cumulative hypergeometric distribution [3, 5].

Lemma 2. *The probability of at least β of all shards committees committing a faulty block ($\mathcal{P}^{'}$) can be computed as follows:*

$$\sum_{i=\frac{2\zeta}{3}}^{\zeta} \Big(P(X \geq \beta n)\Big)^{i} = \sum_{i=\beta\zeta}^{\zeta} \sum_{\alpha=\beta n}^{n} \left(\frac{\binom{M}{\alpha}\binom{H}{n-\alpha}}{\binom{\mathcal{V}}{n}}\right)^{i} \tag{2}$$

Proof. The minimum number of committees to commit a faulty block is $\beta\zeta$, where ζ is the number of shards. The probability of exactly $\beta\zeta$ committees confirm/agree to add a faulty block can be expressed as follows:

$$P_{\beta\zeta} = \Big(P(X \geq \beta n)\Big)^{\beta\zeta} \tag{3}$$

The probability to commit a faulty block by exactly $\beta\zeta + 1$ committees can be expressed as follows.

$$P_{\beta\zeta+1} = \Big(P(X \geq \beta n)\Big)^{\beta\zeta+1} \tag{4}$$

Similarly, the probability of exactly ζ committees (the entire number of shards in this case) agreeing to add a faulty block can be expressed as follows:

$$P_{\zeta} = \Big(P(X \geq \beta n)\Big)^{\zeta} \tag{5}$$

A faulty block can be committed if $\beta\zeta$ or $\beta\zeta + 1$ or $\beta\zeta + 2, \cdots$, or ζ committees agree to add this block. This can be mathematically computed by the sum over all these probabilities and can be expressed as follows:

$$\mathcal{P}^{''} = P_{\beta\zeta} + P_{\beta\zeta+1} + \cdots + P_{\zeta} \tag{6}$$

Lemma 3. *The probability of the beacon's committee committing a faulty block ($\mathcal{P}^{'}$) can be expressed as follows:*

$$P(X' \geq \beta n') = \sum_{j=\beta n'}^{n'} \frac{\binom{M'}{j}\binom{H'}{n'-j}}{\binom{\mathcal{V}'}{n'}} \tag{7}$$

Proof of **Lemma** 3 results directly from the cumulative hypergeometric distribution [3, 5].

Theorem 1 (Committing a Faulty Block). *The probability of committing a faulty block ($\mathcal{P}_f$) by a given shard can be expressed as follows:*

$$\mathcal{P}_f = \sum_{k=\beta n}^{n} \sum_{i=\beta\zeta}^{\zeta} \sum_{\alpha=\beta n}^{n} \sum_{j=\beta n'}^{n'} \frac{\binom{M}{k}\binom{H}{n-k}\binom{M}{\alpha}^{i}\binom{H}{n-\alpha}^{i}\binom{M'}{j}\binom{H'}{n'-j}}{\binom{\mathcal{V}}{n}\binom{H}{n-\alpha}^{i}\binom{\mathcal{V}'}{n'}} \tag{8}$$

Proof. To commit a faulty block, it must be confirmed/verified by at least β of the shard committee members, by at least β of the beacon committee members, and by at least β of all shards' committees. This can be expressed by the product over the three probabilities (the calculated probabilities in **Lemmas** 1, 2, and 3).

2.4 Years to Fail

To make the measurement of the security more readable, we propose to compute the number of years to fail ($\mathcal{Y}_f$) based on the calculated failure probability (i.e., the probability of conquering the protocol). This number can be expressed as follows:

$$\mathcal{Y}_f = 1/\mathcal{P}_f/\mathcal{N}_s \tag{9}$$

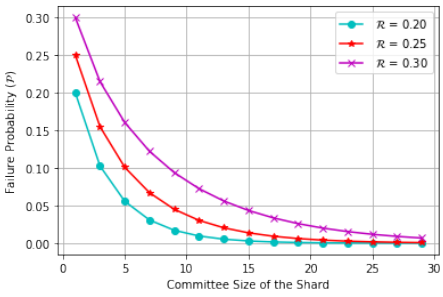
where $\mathcal{P}_f$ is the probability of committing (adding) a faulty block to the blockchain and $\mathcal{N}_s$ is the number of sharding rounds per year (aka, number of epochs per year).

3 Results & Evaluation

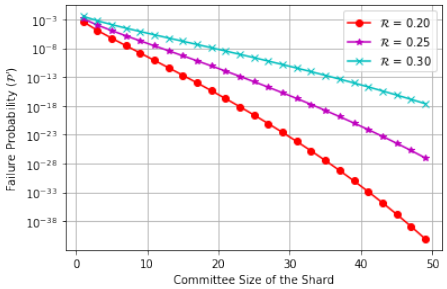
In this section, we evaluate the effectiveness of the proposed probabilistic model via numerical simulations.

3.1 Simulation Setup

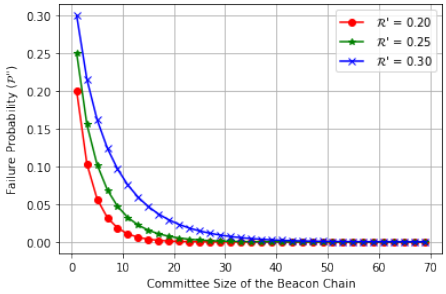
In order to implement the proposed model, we make use of a built-in Python library called **SciPy**. Particularly, we import **hypergeom** from **scipy.stats** model.



(a) Shard Failure Probability



(b) Failure Probability of all Shards



(c) Beacon Failure Probability

Fig. 2: (a) Probability of a shard to commit a faulty block ($\mathcal{P}$) versus the committee size of the shard (n), (b) Log-scale plot of the probability of all shards committing a faulty block ($\mathcal{P}^{'}$) versus the size of the committee (n), and (c) Probability of the beacon chain to commit a faulty block ($\mathcal{P}''$) versus the size committee of the beacon chain (n').

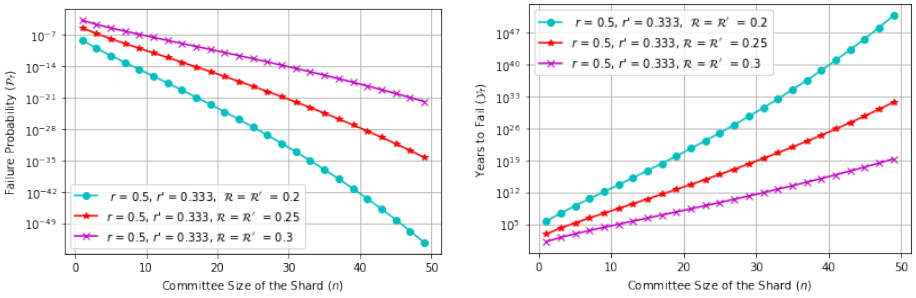
3.2 Results and Analysis

In Figure 2, we assume a network with $\mathcal{N}$= 2000 nodes, $\mathcal{V}$ = 200, $\mathcal{V}'$ = 400, ζ = 8, $r = r' = 0.5$.

Figure 2a shows the probability of a shard to commit a faulty block versus the size of the committee. We observe that the probability $\mathcal{P}$ decreases when the size of the committee increases. More specifically, we observe that the probability corresponding to $\mathcal{R} = 0.2$ (i.e., 20% of malicious nodes in each shard) decreases rapidly compared to those of $\mathcal{R} = 0.25$ and $\mathcal{R} = 0.3$; this can be explained by the small percentage of malicious nodes. In other words, as the percentage of malicious nodes gets smaller the probability decreases and vice versa.

Figure 2b shows the probability of all shards committing a faulty block versus the size of the committee. We observe that the probability $\mathcal{P}'$ decreases when the size of the committee increases. Similarly, as the percentage of malicious nodes slightly increases in the shard, the probability of committing a faulty block increases.

Figure 2c shows the probability of the beacon chain to commit a faulty block ($\mathcal{P}''$) versus the size committee of the beacon chain (n'). We also observe that the probability $\mathcal{P}''$ decreases when the size of the committee increases. More specifically, we observe that the probability corresponding to $\mathcal{R} = 0.2$ (i.e. 20% of malicious nodes in the beacon chain) decreases sharply compared to those of $\mathcal{R} = 0.25$ and $\mathcal{R} = 0.3$.



(a) Probability of Conquering the Protocol (b) Years to Fail

Fig. 3: Log-scale plot: (a) Probability of conquering the protocol ($\mathcal{P}_f$) versus the committee size of the shard (n), (b) Number of years to fail ($\mathcal{Y}_f$) versus the committee size of the shard (n).

In Figure 3, we assume a network with 2000 nodes, $\mathcal{V}$ = 200, $\mathcal{V}'$ = 400, ζ = 8, n' = 100. Figure 3a shows the probability of the conquering the protocol when varying the committee size of the shard. We observe that as the committee size of the shard increases the probability of conquering the protocol decreases. Figure 3b shows the number of years to fail ($\mathcal{Y}_f$) versus the committee size of

the shard. We observe that when the committee size of the shard increases the number of years to fail increases.

Table 2: Probability of conquering the protocol.

$\mathcal{R} = \mathcal{R}'$	10 %	15 %	20 %	30 %
$\mathcal{P}_f^{\text{a}}$	3.63E-66	2.10E-34	1.58E-18	1.70E-04
Y_f^{a}	7.56E+62	1.30E+31	1.74E+17	16.12
$\mathcal{P}_f^{\text{b}}$	0.0	5.14E-80	2.01E-41	5.30E-07
Y_f^{b}	inf	5.33E+76	1.36E+38	5171.32

[a]Scenario 1; [b]Scenario 2.

In Table 2, we assume two scenarios to show the effectiveness and the feasibility of the proposed model: Scenario 1 proposes a network with $\mathcal{N} = 2000$, $\zeta = 8$, $\mathcal{V} = 200$, $\mathcal{V}' = 400$, and $r = r' = 0.333$ whereas Scenario 2 proposes a network with $\mathcal{N} = 4000$, $\zeta = 8$, $\mathcal{V} = 400$, $\mathcal{V}' = 800$, and $r = r' = 0.333$. It is noteworthy that the proposed model can be adopted to any scenario.

Table 2 shows the probability of conquering the chain (i.e., the probability of committing a faulty block; it is calculated based on **Theorem** 1) for different percentages of malicious nodes in the shards as well as in the beacon chain. Moreover, Table 2 shows the number of years to fail corresponding to these probabilities. We observe that as the percentage of malicious nodes increases the number of years to fail decreases. More specifically, we observe that the probability of conquering the chain is extremely low even with 20% of malicious nodes in each shard as well as in the beacon chain. This achieves a good security, which is about $1.74E + 17$ years to fail.

Finally, we conclude that by adjusting the committee size of the shard as well as the committee size of the beacon chain, we could protect sharded Blockchain systems (based on PoS) against malicious nodes (e.g., Sybil nodes).

4 Conclusion

In this paper, we address the security of sharding-based PoS Blockchain protocols. In particular, we provide a probabilistic model to compute the probability of committing a faulty block. Based on this probability, we compute the number of years to fail. Furthermore, this article depicts that we can control the number of years to fail by adjusting the committee size of the shard as well as the committee size of the beacon chain. Our future work includes the computation of the failure probability across-shard transaction.

References

1. Abou Jaoude, J., George Saade, R.: Blockchain applications - usage in different domains. IEEE Access **7**, 45360–45381 (2019). https://doi.org/10.1109/ACCESS.2019.2902501

2. Hafid, A., Hafid, A.S., Samih, M.: New mathematical model to analyze security of sharding-based blockchain protocols. IEEE Access **7**(19232899), 185447–185457 (2019). https://doi.org/10.1109/ACCESS.2019.2961065
3. Hafid, A., Hafid, A.S., Samih, M.: A novel methodology-based joint hypergeometric distribution to analyze the security of sharded blockchains. IEEE Access **8**(20000968), 179389–179399 (2020). https://doi.org/10.1109/ACCESS.2020.3027952
4. Hafid, A., Hafid, A.S., Samih, M.: Scaling blockchains: A comprehensive survey. IEEE Access **8**(19800223), 125244–125262 (2020). https://doi.org/10.1109/ACCESS.2020.3007251
5. Hafid, A., Hafid, A.S., Samih, M.: A tractable probabilistic approach to analyze sybil attacks in sharding-based blockchain protocols. arXiv (2020), https://arxiv.org/abs/2104.07215
6. Incognito: (2021), https://we.incognito.org/t/incognito-whitepaper-incognito-mode-for-cryptonetworks/168
7. Kassab, M.H., DeFranco, J., Malas, T., Laplante, P., destefanis, g., Graciano Neto, V.V.: Exploring research in blockchain for healthcare and a roadmap for the future. IEEE Transactions on Emerging Topics in Computing pp. 1–1 (2019). https://doi.org/10.1109/TETC.2019.2936881
8. Nakamoto, S.: Bitcoin whitepaper. URL: https://bitcoin.org/bitcoin.pdf (2008)
9. Nordgren, A., Weckström, E., Martikainen, M., Lehner, O.M.: Blockchain in the fields of finance and accounting: a disruptive technology or an overhyped phenomenon. ACRN Oxford Journal of Finance and Risk Perspectives **8**(1), 47–58 (2019)
10. PayPal: (2021), https://www.paypal.com/fr/webapps/mpp/home
11. Rasmussen, R.V., Trick, M.A.: Round robin scheduling–a survey. European Journal of Operational Research **188**(3), 617–636 (2008)
12. Wang, G., Shi, Z.J., Nixon, M., Han, S.: Sok: Sharding on blockchain. In: Proceedings of the 1st ACM Conference on Advances in Financial Technologies. pp. 41–61 (2019)