2020 International Conference on Emerging Technologies for Communications (ICETC 2020)

Reward and Penalty Mechanism in Proof-of-Stake Consensus Algorithm for Blockchain

Takeaki MATSUNAGA[†a)], **Yuanyu ZHANG**[†], **Masahiro SASABE**[†], ***and*** **Shoji KASAHARA**[†],

SUMMARY The Proof of Stake (PoS) is a consensus algorithm for blockchain, in which the integrity of a new block is validated according to voting of nodes called validators. Since a new generated block is confirmed according to validators' voting, it is important to motivate validators to vote correctly. One of incentive mechanisms for validators is a reward-penalty based incentivization, in which validators who contribute to correct consensuses are rewarded, while those who make incorrect block confirmation are penalized. In this paper, we consider a reward-penalty mechanism based on the voting profile of a validator. We quantify the voting profile of a validator with exponentially weighted moving average of voting results, and the reward/penalty of a current voting is derived with the voting profile value. We evaluate the performance of the proposed mechanism by computer simulation, investigating the impact of system parameters on the estimation accuracy of the validator profile and the amount of validator's stake. Numerical results show that the proposed algorithm can estimate the voting profile of a validator accurately even when the voting profile dynamically changes. It is also shown that the proposed mechanism gives more reward to validators with high voting profile.

key words: *Blockchain, Proof-of-Stake, Reward-Penalty Mechanism*

1. Introduction

Blockchain is a distributed ledger technology which is supporting cryptocurrencies used on the Internet. In a blockchain, transactions are stored in a block, and the new block is linked to a ledger database with a chain structure by consensus algorithms. One of the consensus algorithms for blockchains is Proof-of-Work (PoW), which has been adopted in Bitcoin [1] and Ethereum [2]. In PoW, the block-approving procedure is called mining. However, mining requires a huge amount of computing power, causing a huge amount of electricity consumption [3]. This energy waste is a drawback of PoW.

Proof-of-Stake (PoS) is a consensus algorithm developed for addressing the drawback of PoW. In PoS, the computing power is replaced by a deposit paid in the cryptocurrency, called stake. In PoW, a node with higher computing power obtains a higher chance to create a new block. In PoS, unlike PoW, a node that holds a higher amount of stake is more likely to create a new block.

In PoS, some participating nodes are selected as validators according to their amounts of stake. The main role of validators is to validate the integrity of a new block and to confirm it by voting. It is important for validators not only to verify generated blocks precisely but also to vote correctly. In order to make validators vote correctly, reward-penalty-based incentive mechanism plays an important role.

[†]The authors are with NAIST, Ikoma-shi, Nara, 630-0192, Japan.

a) E-mail: matsunaga.takeaki.mo9@is.naist.jp

In this paper, we consider a reward-penalty mechanism based on validators' voting profiles. The voting profile of a validator represents its reliability according to its voting history. A reward or penalty for the vote of a validator is calculated according to its reliability. Conducting simulation experiments, we investigate the impact of a validator's voting history on the amount of its stake.

This paper is organized as follows. Section 2 introduces the related work on the consensus algorithms for blockchains. In Section 3, we describe the details of our reward-penalty-based mechanism in the consensus algorithm of a blockchain with PoS. Numerical examples are shown in Section 4, and we conclude the paper and show future work in Section 5.

2. Related Work

In Bitcoin, a participating node obtains a reward if it solves a cryptographic puzzle associated with a new block, and the new block is appended to the blockchain [4]. Since the winning probability of mining depends on the computing power of participating nodes, a lot of high-performance special hardware is invested in mining. As a result, an enormous amount of electricity is consumed, leading to environmental damage [5]. In order to address the energy-consuming drawback of PoW, PoS has been proposed. PPcoin [6] is known as the first cryptocurrency with PoS consensus mechanism.

In Bitcoin and PPcoin, when more than one node solves a cryptographic puzzle, multiple new blocks are added to the latest block and the chain branches. This phenomenon is called a fork. Bitcoin and PPcoin solve the fork issue with the longest chain rule, in which the chain with the largest number of blocks is considered as a valid chain [1, 7]. However, the longest chain rule is not enough to guarantee the finality of transactions because of the possibility of another new long chains in the future.

One of PoS-based implementations guaranteeing the transaction finality is Tendermint [8]. In Tendermint, a participating node deposits tokens as a security deposit stake. A predefined number of participating nodes are selected as validators in descending order of stake. The validator responsible for generating a block is called a proposer. When the proposer generates a block, validators except the proposer validate the generated block.

The selected validators form a committee, and each validator votes based on its validation result of the generated block. If the number of votes for the new block is greater than or equal to the threshold prespecified by the network, the

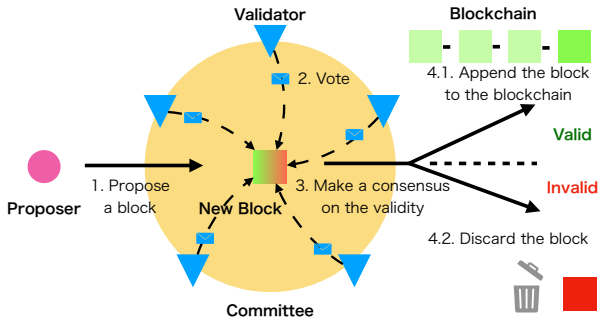

Fig. 1 Consensus procedure for PoS-based blockchain.

committee makes a consensus that the block is valid. The block approved by the committee is eventually appended to the blockchain. Due to the consensus with voting, fork never happens. In addition, the finality can be defined as the point at which the committee reaches the consensus on the generated block. (See Fig. 1 for details of PoS-based consensus algorithm.) The blockchain technologies such as Cosmos [9], Polkadot [10], and Ethereum 2.0 [11] adopt the PoS-based consensus algorithm, in which a consensus on the validity of a block is made by simple voting of dichotomous choices, an approval or not (including an abstention).

Leonardos et al. proposed the weighted voting for the PoS-based blockchains [12]. The weighted voting is a group-decision procedure with fixed population size [13, 14], in which the voting result is calculated with weights associated with voters' reliability such that the likelihood of the better choice of two alternatives is maximized. In [12], the voting reliability of a validator is quantified as a profile, and the quota of the weighted approval votes for a consensus is derived.

The reward-penalty based mechanisms for PoS-type blockchains have also been considered in [15, 16]. In those studies, the reward/penalty is considered for the voting result of a newly generated block, whereas the voting profile of a validator is not taken into consideration. In this paper, we consider a reward-penalty mechanism based on the voting profile of a validator, which is estimated from the voting history of the validator. A validator that votes correctly is given a reward. On the contrary, when the validator votes incorrectly, some of its stake are confiscated.

3. Reward-Penalty Mechanism for PoS

3.1 Reward and Penalty

Let $\mathcal{N} = \{1, 2, \ldots, V\}$ denote the set of validators that maintain the blockchain. A block is generated in every time slot $t \in \mathbb{N} \cup \{0\}$. The validators validate the integrity of the block in every block generation, and then vote. we define $s_{i,t} \in \mathbb{R}^+ \cup \{0\}$ as the amount of stake of validator i, and $r_{i,t} \in \mathbb{R}$ as the reward or penalty resulting from the voting outcome in time slot t. The amount of stake of validator i in time slot $t+1$, $s_{i,t+1}$, is defined by the following equation.

$$s_{i,t+1} = s_{i,t} + r_{i,t}. \tag{1}$$

When a new block is generated, all validators verify it and submit votes. If the majority of the votes is positive, the block is judged as "valid" and added to the blockchain. If the majority is negative, the block is determined as "invalid" and discarded. A vote of a validator is called "correct" if the validator judges a valid block as valid or an invalid one as invalid. On the contrary, if the validator judges a valid block as invalid or an invalid one as valid, the corresponding vote is called "incorrect". Let $x_{i,t} \in \{0, 1\}$ denote the voting result of validator i in time slot t. $x_{i,t} = 0$ represents that validator i voted for an incorrect block, while $x_{i,t} = 1$ implies that validator i voted for a correct block. Note that voting for an invalid block poses a serious threat to the integrity of the entire blockchain. There are two types of incorrect voting, one is intentional and the other is unintentional[†]. In this paper, we do not distinguish between them.

We define $r_{i,t}$ as the amount of reward or penalty for validator i at t, which is given by

$$r_{i,t} = \begin{cases} reward_{i,t}, & \text{if } x_{i,t} = 1, \\ penalty_{i,t}, & \text{otherwise}, \end{cases}$$

where $reward_{i,t} \in \mathbb{R}^+ \cup \{0\}$ is the reward and $penalty_{i,t} \in \mathbb{R}^-$ is the penalty. $reward_{i,t}$ is calculated by

$$reward_{i,t} = p_{i,t} \cdot \beta, \tag{2}$$

where $p_{i,t}$ is the profile of validator i at time slot t and is described in details in the next subsection. $\beta \geq 0$ is the basic reward at each time slot defined in the protocol. We can see from (2) that the higher $p_{i,t}$ of validator i is, the more $reward_{i,t}$ it can obtain.

Using the amount of stake $s_{i,t-1}$ of validator i in the last time slot $t-1$, we calculate $penalty_{i,t}$ as

$$penalty_{i,t} = -\gamma \cdot s_{i,t-1}, \quad 0 \leq \gamma \leq 1,$$

where γ is a weight parameter for penalty and determines how much of stake will be confiscated based on $s_{i,t-1}$.

3.2 Validator's Profile Update

In this subsection, we describe how to estimate the voting profile of a validator, which is the indicator of the accuracy for block verification. Let $p_{i,t} \in [0, 1]$ denote the profile of validator i at time slot t. The profile $p_{i,t}$ is updated by the following exponentially weighted moving average (EWMA)

$$p_{i,t} = \delta \cdot p_{i,t-1} + (1 - \delta) \cdot x_{i,t}, \quad 0 \leq \delta \leq 1. \tag{3}$$

Note that a large $p_{i,t}$ implies that validator i is likely to validate a new block correctly. A small δ results in $p_{i,t}$ that is dominantly affected by the current voting result $x_{i,t}$, while a larger δ gives a smoother estimate of $p_{i,t}$.

[†]An intentional vote can be regard as an attack to fail the correct consensus. On the other hand, the unintentional one can be caused by hardware failure or unstable network environments.

Table 1 Parameter Settings.

Parameter	Setting
δ	$\{0.9, 0.99, 0.999\}$
γ	$\{0, 0.01, 0.02\}$
ζ	$\{0.99, 0.999, 1\}$

4. Numerical Examples

4.1 Simulation Model

In our simulation, the unit of the simulation time $t \in \{0, 1, \ldots, T\}$ is the block-generation time, which is constant and equal to d [s]. The simulation starts at time slot 0 and ends at T. At the beginning of time slot t, a new block is generated. The new block is valid with probability $\xi \in [0, 1]$, independent of the other generated blocks.

Each validator votes for the integrity of the new block. We assume that all the validators vote correctly with probability ζ. That is, with probability ζ, the validator votes for a valid block or against an invalid block. On the contrary, with probability $1 - \zeta$, the validator votes for an invalid block or against a valid block.

Let S_t denote the total amount of stake on the network at time slot t, which is defined by $S_t = \sum_{i=1}^{n} s_{i,t}$. Based on this, we calculate the basic reward β by the following equation.

$$\beta = \frac{S_0 \cdot \epsilon}{T \cdot V},$$

where ϵ is the inflation rate for the period T. The amount of stake of the validator i in time slot t, $s_{i,t}$, is updated by (1) with the reward $r_{i,t}$. In terms of the performance measure, we consider the average amount of stake of validators in time slot t, $s_{avg,t}$, which is given by $s_{avg,t} = \sum_{i \in \mathcal{N}} s_{i,t}/V$.

In our simulation, the block generation time is set to $d = 60$ [s], and a block generation process for two weeks is simulated. From this assumption, we have $T = 20160$. The number of validators is set to $V = 1000$. For all the validators, the initial amount of stake is set to $s_{i,0} = 1000$, and the initial profile is set to $p_{i,0} = 0.5$. From these assumptions, we have $s_{avg,0} = 1000$ and $S_0 = 10^6$. In [17], the inflation rate for a year is set to 0.15. Using this value, the inflation rate for two weeks, ϵ, is set to

$$\epsilon = 0.15 \cdot \frac{14}{365} \simeq 5.7534 \times 10^{-3}.$$

We summarize the remaining parameter settings in Table 1.

4.2 Impact of EWMA Weight Parameter

First, we investigate how the EWMA weight parameter δ of the equation (3) affects the validator profile $p_{i,t}$. For simplicity, we consider the case of $\xi = 1$, that is, all the generated blocks are valid. In this experiment, we change the correct voting probability ζ, investigating how correctly $p_{i,t}$ follows ζ. ζ is changed according to the following scenarios. When $1 \le t < 2500$, $\zeta = 1$. When $2500 \le t < 7500$, $\zeta = 0.25$. When $7500 \le t < 12000$, $\zeta = 0.5$. When $12000 \le t < 17000$, $\zeta = 0.75$. When $17000 \le t$, $\zeta = 1$.

Fig. 2 shows the evolution of $p_{i,t}$ for one validator. Here, δ is set to 0.9. It is observed that $p_{i,t}$ takes the value of 1 until $t = 2500$. Then, $p_{i,t}$ suddenly decreases to the value lower than $\zeta = 0.25$ at $t = 2500$, and fluctuates greatly around 0.25. After $t = 7500$, $p_{i,t}$ fluctuates around 0.5, and we also observe the same tendency from $t = 12000$ to 17000. After $t = 17000$, $p_{i,t}$ returns to the value of 1.

Figs. 3 and 4 show the evolution of $p_{i,t}$ in cases of $\delta = 0.99$ and 0.999, respectively. It is found that $p_{i,t}$ exhibits less fluctuation with the increase of δ. This is because a large δ makes current samples less significant, smoothing the estimate of $p_{i,t}$. It is also observed that $p_{i,t}$ slowly approaches to the true value with the increase of ζ.

When designing the reward-penalty mechanism with $p_{i,t}$, it is important to consider the tradeoff between the fluctuation and convergence speed. In the following, we set $\delta = 0.99$.

4.3 Impact of Penalty Weight Parameter

Figs. 5 to 7 represent the average amount of stake of validators $s_{avg,t}$ in time slot t. Here, we set $\delta = 0.99$, and $s_{avg,t}$'s in three cases of $\gamma = 0$, 0.001, and 0.02 are plotted in each figure.

Fig. 5 shows the average amount of stake of validators in case of $\zeta = 0.99$. It is observed in this figure that when $\gamma = 0.02$, the average amount of stake significantly decreases. It is also observed that $s_{avg,t}$ for $\gamma = 0.01$ exhibits the same tendency as that for $\gamma = 0.02$, however, the former is greater than the latter. In case of $\gamma = 0$, on the contrary, the average amount of stake slightly increases.

Fig. 6 shows the average amount of stake in case of $\zeta = 0.999$. In this figure, we observe the same tendency as Fig. 5. However, decreasing rates of $s_{avg,t}$'s in Fig. 6 are smaller than those in Fig. 5.

Fig. 7 shows the average amount of stake for $\zeta = 1$, in which validators vote correctly. We observe a slight increase of stake, regardless of the penalty weights.

From the above results, it can be seen that as the penalty weight γ increases, more severe penalties are imposed on validators that don't vote correctly. If too much severe penalties are imposed, validators may leave the system, and the consensus mechanism with PoS is likely not to work. A careful design of the penalty weight parameters γ and ζ is required.

5. Conclusion

In this paper, we considered a reward-penalty-based mechanism for PoS. In our proposed mechanism, the reliability of a validator is characterized as a profile, and each validator is rewarded or penalized according to its voting, depending on its current profile. Numerical results showed that the proposed approach can estimate the profile of a validator accurately. In addition, it was shown that by adopting the profiles, validators with a high profile can obtain more rewards than validators with a low profile. This may motivate

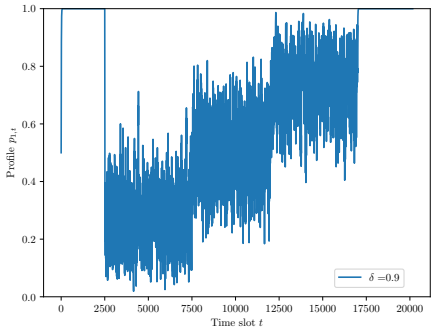

Fig. 2 Sample of $p_{i,t}$. ($\delta = 0.9$)

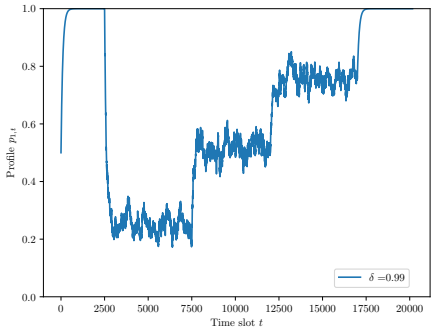

Fig. 3 Sample of $p_{i,t}$. ($\delta = 0.99$)

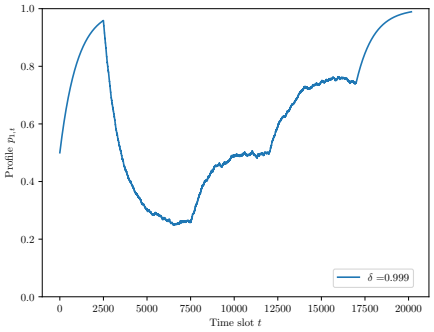

Fig. 4 Sample of $p_{i,t}$. ($\delta = 0.999$)

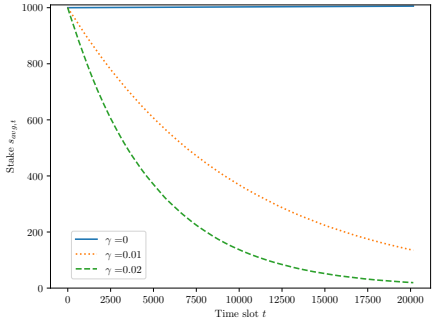

Fig. 5 Evolution of $s_{avg,t}$. ($\zeta = 0.99$)

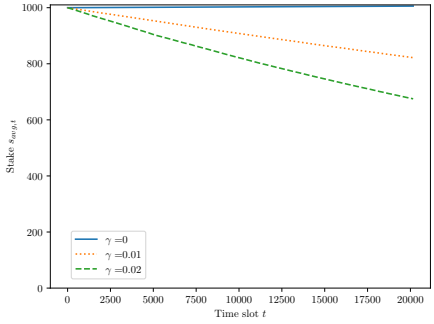

Fig. 6 Evolution of $s_{avg,t}$. ($\zeta = 0.999$)

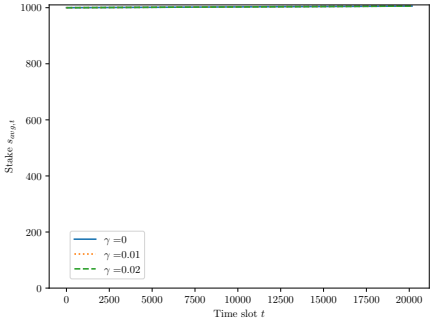

Fig. 7 Evolution of $s_{avg,t}$. ($\zeta = 1$)

validators to vote correctly.

In this proposed approach, the reuse of discounted rewards and confiscated stake by penalty is not taken into consideration. Further studies are needed to consider a reward-penalty-based mechanism with the reuse of stake in order to incentivize validators to vote more correctly.

Acknowledgements

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI (A) under Grant 19H01103, and the Support Center for Advanced Telecommunications (SCAT) Technology Research Foundation.

References

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System." `https://bitcoin.org/bitcoin.pdf`, 2008. (Accessed on 05/13/2020).

[2] V. Buterin *et al.*, "A Next-Generation Smart Contract and Decentralized Application Platform." `https://cryptorating.eu/whitepapers/Ethereum/Ethereum_white_paper.pdf`, 2014. (Accessed on 06/10/2020).

[3] K.J. O'Dwyer and D. Malone, "Bitcoin Mining and its Energy Footprint," *Proc. of ISSC 2014/CIICT 2014*, pp.280–285, IET, 2014.

[4] A.M. Antonopoulos, *Mastering Bitcoin*, O'Reilly, 2014.

[5] C. Mora, R.L. Rollins, K. Taladay, M.B. Kantar, M.K. Chock, M. Shimada, and E.C. Franklin, "Bitcoin Emissions Alone Could Push Global Warming above 2 C," *Nature Climate Change*, vol.8, no.11, pp.931–933, 2018.

[6] S. King and S. Nadal, "Ppcoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake." `https://decred.org/research/king2012.pdf`, 2012. (Accessed on 05/13/2020).

[7] StackExchange, "What Does the Term "Longest Chain" Mean?." `https://bitcoin.stackexchange.com/questions/5540/what-does-the-term-longest-chain-mean`, 2012. (Accessed on 06/05/2020).

[8] J. Kwon, "Tendermint: Consensus without Mining." `https://tendermint.com/static/docs/tendermint.pdf`, 2014. (Accessed on 05/30/2020).

[9] K. Jae and B. Ethan, "Cosmos A Network of Distributed Ledgers." `https://cosmos.network/cosmos-whitepaper.pdf`. (Accessed on 05/30/2020).

[10] G. Wood, "Polkadot: Vision for a Heterogeneous Multi-Chain Framework." `https://www.win.tue.nl/~mholende/seminar/references/ethereum_polkadot.pdf`, 2016. (Accessed on 05/13/2020).

[11] GitHub, "Ethereum 2.0 Specifications." `https://github.com/ethereum/eth2.0-specs`. (Accessed on 05/13/2020).

[12] S. Leonardos, D. Reijsbergen, and G. Piliouras, "Weighted Voting on the Blockchain: ImprovingConsensus in Proof of Stake Protocols," *Proc. of IEEE ICBC 2019*, pp.376–384, IEEE, 2019.

[13] R.C. Ben-Yashar and S.I. Nitzan, "The Optimal Decision Rule for Fixed-Size Committees in Dichotomous Choice Situations: The General Result," *International Economic Review*, vol.38, no.1, pp.175–186, 1997.

[14] L. Shapley and B. Grofman, "Optimizing Group Judgmental Accuracy in the Presence of Interdependencies," *Public Choice*, vol.43, no.3, pp.329–343, 1984.

[15] A. Ouaguid, N. Abghour, and M. Ouzzif, "Towards a New Reward and Punishment Approach for Blockchain-Based System," *Proc. of IEEE SysCoBIoTS 2019*, pp.1–7, IEEE, 2019.

[16] V. Buterin, "Slasher: A Punitive Proof-of-Stake Algorithm." `https://blog.ethereum.org/2014/01/15/slasher-a-punitive-proof-of-stake-algorithm`, 2014. (Accessed on 05/13/2020).

[17] GitHub, "Signal Non-Final Statu of Base Reward and Desired Issuance Goal #971." `https://github.com/ethereum/eth2.0-specs/pull/971`. (Accessed on 06/05/2020).