Making Blockchain Validators Honest

Abiola Salau, Ram Dantu, Kirill Morozov, Syed Badruddoja, Kritagya Upadhyay
Department of Computer Science and Engineering
University of North Texas
Denton, TX, 76207, USA
E-mail: abiolasalau@my.unt.edu, Ram.Dantu@unt.edu, Kirill.Morozov@unt.edu, syedbadruddoja@my.unt.edu, kritagyaupadhyay@my.unt.edu

2022 Fourth International Conference on Blockchain Computing and Applications (BCCA) | 978-1-6654-9958-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/BCCA55292.2022.9921952

***Abstract*—The importance of honesty among blockchain validators can not be overemphasized, especially as blockchain is used by many as an underlying technology for the development of various Industry 4.0 solutions. In the blockchain consensus process, validators validate the correctness of transactions and propose new blocks for addition to the blockchain. They are rewarded for this task with the blockchain native token (e.g., ETH on the Ethereum blockchain). This reward is often distributed among validators with respect to their staked amount. An increasing number of validators joining the network leads to a decreasing chance of a validator being chosen for the validation task and thus a reduction in the validation reward. This situation results in some form of competition among the validators, leading them to carry out various malicious actions to influence the blockchain validator selection protocol in order to be chosen. In this paper, we examine the competitive interactions between validators in a blockchain consensus process and propose a model using an infinitely repeated game model that ensures that the validators are deterred from behaving maliciously while also encouraging a self-policing notion due to the extra incentive mechanism of an improved reputation score when a validator can verifiably report malicious activities by others. Further, we discuss the factors that can incentivize or disincentivize a validator to either continue to behave honestly or switch to dishonest behavior.**

***Index Terms*—Game theory, reputation, Blockchain, validation, cooperation, prisoner's dilemma.**

I. Introduction

Game theory has been well applied in scenarios where there is a need for multi-person decision making. The interactions among these people can be modeled as a game where each person, generally referred to as a player in the game based on a set of strategies they choose to play, tries to maximize their payoff regardless of what other player(s) does [34]. One of the fascinating research areas where such games can be utilized is the verification of transactions in the context of blockchain systems where miners (in Proof-of-Work) or validators (in Proof-of-Stake) attempt to maximize the reward they can get from the system. The competition among these validators (players) is very intense, since often not all players will get a share of the reward per round [31]. Therefore, some result in playing in a dishonest way to outsmart the other players and/or the system.

Different strategies are analyzed in the existing works of literature to address what issues this competition may cause. However, due to space limitation, in the related works section, we only discuss works closely related to ours and refer readers to *A Course in Game Theory* [1] for a more detailed introduction to game theory and [2] for a comprehensive read on applications of game theory in blockchain. Particularly of interest is the literature that applies game theory techniques to blockchain. For instance, the authors in [3], likewise [4], [5] examined the malicious activity of the players by comparing an honest approach with a dishonest strategy, i.e., players of a coalition (e.g., collaborating miners) can invest to acquire additional computing resources or launch distributed denial-of-service attacks against other competing coalitions.

A. Problem Definition and Motivation

Blockchain can be incorporated into small- and large-scale enterprise solutions to improve security, data transparency, and privacy of digital assets for Industry 4.0 [37]. However, for blockchain to effectively serve this purpose, the underlying consensus protocol that is used to verify the validity of a transaction has to be non-manipulable. In a blockchain consensus process, e.g., PoS, validators are chosen to verify the accuracy and legality of transactions before they are added to a block on the chain. The selection of the validators is often based on some system requirements, e.g., the biggest stake, reputation value, etc [26]. This can be seen as a competition since only selected validators will have the chance to participate in the block creation process and therefore earn a block creation reward.

Naturally, the players in a consensus protocol are rational, trying to maximize the payoff they can get from their participation in the consensus process [28]. For example, in the PoW-based blockchain systems, only the "winner" earns the mining reward. This results in the players' seeking devices that can do more computation per second, which leads to the use of GPUs, FPGAs and ASICs. The implication of this is that players who rely on the use of their CPU-powered machines will rarely get the chance to mine a new block, and a more critical problem this has for the blockchain system is that it pulls the system towards centralization (since nodes with better computing machines will always win the *mining race*) increasing the success probability of 51% attacks. To address this issue, there is a lot of work around ASIC-resistant designs for the PoW consensus protocol [12]–[15]. This means that a player may continuously play dishonestly in its interest if there is no security system to detect its maliciousness.

Hence, in this paper, we mainly address the problem of malicious validators in blockchain consensus protocols where validators behave dishonestly (e.g., selfish mining and behavior (see [26], [27], [32], [33], [36]) to get selected to add the next block to the blockchain. We consider any validator action deviating from the honest behavior of following the protocol as malicious, including but not limited to actions like approving (or disapproving) an invalid (or valid) transaction, delaying block proposals, not participating in the consensus process when selected, etc. Further, we study the incentives of malicious validators to deviate from playing according to the blockchain protocol, and through a game-theoretic approach and reward mechanism, we enforce honest behavior among the validators and achieve a *neighborhood watch* where the validators self-police themselves [6] [23].

B. Our Contributions

The contributions of this paper are summarized below.

- Firstly, we modeled the competition between transaction validators in a consensus protocol as a game in Sec. III.
- Secondly, we compared the short-term and long-term payoffs of the validators (players), showing that the (Dishonest, Dishonest) unique Nash equilibrium of the players can become a (Honest, Honest) unique sub-game perfect Nash Equilibrium (SPNE) with the appropriate discount factor (how well the player values the future), threat (in the form of penalty), and a reward mechanism (in the form of improved reputation).
- Thirdly, we established the value of the discount factor that can give the SPNE showing that it must be high (close to 1) to maintain SPNE which will deter the players from malicious acts and in turn ensure cooperative behavior with the protocol (Sec. IV-A).
- Fourthly, we evaluated our model by simulation with varying parameters and, from the result of our model, we discussed the factors that will ensure that the discount factor is achieved and maintainable (Sec. IV-B).
- Finally, with our approach, we will be able to deter the players from malicious actions and compel them to play honestly according to the protocol with the belief that a long-term gain is better than a short-term gain.

II. Related Works

Narang et. al in [25], proposed a framework that studies the interactions of agents on a B2B platform. In [5], [6], the authors proposed models to incentivize honest mining in PoW-based blockchain systems and prevent attackers from joining the mining pool. The authors in [7] studied the problem in the context of users and network providers, while in [8], the authors used reputation mechanisms and a Stackelberg game incentive model for crowdsensing networks. In paper [8], the authors applied an evolutionary game-theoretic approach to develop a reward mechanism for proof-of-stake blockchains.

The block withholding attack in mining pools introduced by [24] was further examined in [9] and the authors showed that the attack is more rewarding if a long-term is considered compared to a short-term. This form of attack was also considered in [10] where the authors tagged the interaction between mining pools as a form of prisoner's dilemma, which they called "miner's dilemma." In [11], the authors also studied the interaction among miners in Bitcoin and discussed the economics involved in whether rational players have enough incentives to deviate from the mining protocol. They showed that there is a Nash equilibrium for which all players can cooperate, but also showed that there are other equilibria where the players can behave maliciously.

Although, as seen thus far, there are existing works in the application of game theory to the blockchain. We emphasize that none of the above-mentioned works studies the interaction of validators with a focus on dishonest peers using an infinitely repeated game theoretic approach to tackle the problem. It is on this note that this brief further examines the interactions among validators and discusses possible scenarios of achieving long-term cooperation among the validators through a unique subgame perfect Nash equilibrium.

III. Model Requirements and Description

In this work, we apply a non-cooperative game model where players cannot form any agreement to coordinate their behavior, i.e., any form of cooperation among them must be self-enforcing in a manner like the prisoner's dilemma game. This is aimed at modeling the situation in real life where the validators are assumed to be scattered all over the world and do not know each other.

TABLE I
Payoff Matrix for Prisoner's Dilemma

P1, P2	Coopearate (C)	Defect (D)
Coopearate (C)	3, 3	1, 4
Defect (D)	4, 1	2, 2

The prisoner's dilemma, with payoff matrix as illustrated in Table I, is an example of non-cooperative games where two possible actions exist: (1.) C: Cooperation and (2.) D: Defection. The Nash equilibrium in this setting is (C, C) since the players are not able to coordinate their behavior. This is, however, not an optimal solution for them since they both can cooperate and get a lesser penalty playing (D, D). The concept of this game is important because many situations have similar settings such as in the blockchain consensus process where the validators compete to add a block to the blockchain which makes them decide on whether to be malicious or not to increase their reward [19].

A. Validator's dilemma

A prisoner's dilemma scenario exists in situations where two players act selfishly, pursuing their interest resulting in a Pareto inferior position, where they both end up worse than if they had cooperated and acted together [17], [18]. This is similar to the situation in the blockchain consensus process

where validators compete for the right to propose the next block on the chain. To further examine this issue, we revisit the works of [19], [20] and translate the models to the setting of blockchain consensus. In situations where the next validator is selected based on some resources like computational resources in PoW-based blockchain systems or stakes in certain PoS-based blockchain systems, a node with a better resource will have a better chance of being chosen. This approach however has limitations in the sense that it is easy to predict who the winner will be and therefore make it an easy target for an attack [29]. To neutralize this form of attack, there should be some form of entropy in the selection process which will make it less predictive for an adversary [30]. This, however, results in more competition among the validators which makes them result in some form of malicious activities to manipulate the system and get chosen.

The dilemma of a validator whether to be malicious or not is the basis of our model. To quantify this scenario for the formulation of our model, let $\mathbf{R}$ represent the reward a validator, V has to gain when it successfully validates transactions to be added to the blockchain. Recall that the probability of a validator being selected depends on the actions of other validators. For simplicity, but without loss of generality, let us assume that there are just two players, for instance, one validator, V deciding whether to be malicious or not against other validators, i.e., V_i where $i_{1\neq 2} \in \{1, 2\}$. Both players have a two-action set, either to be malicious (Dishonest (D)) or not (Honest (H)). If we consider that rewards earned for successfully adding a block to the chain are distributed based on a certain weight of the validator on the blockchain, e.g., its stake or reputation, we let $\alpha\mathbf{R}$ represent the reward for V_1 and $\beta\mathbf{R}$ represent the reward for V_2 where α and β are the weights (e.g., stake or reputation values) of V_1 and V_2 respectively. The utility (u) when both players play honestly (not maliciously) is given in equation 1.

$$u(V_1)_{H,H} = \alpha\mathbf{R} \; and \; u(V_2)_{H,H} = \beta\mathbf{R} \tag{1}$$

If one of the players decides to influence the process through any form of maliciousness while the other player plays honestly, we consider three cases.

Case 1. Increased selection odd for the malicious player (say V_1). To model this scenario, let Δ be the increase in odds for V_1 being selected due to its malicious action. Therefore, the odds of V_1 being chosen become $\alpha + \Delta$.

Case 2. Decreased selection odd for V_2. Although the impact of one player acting dishonestly directly affects the other player's chance, the selection weight for the other player remains unchanged but their odds of being chosen get reduced due to the maliciousness of the other player. Thus, the weight of V_2 will remain β.

Case 3. A punishment for V_1 if caught. Let the penalty for a malicious act be P. For our penalty scheme, we adopt the *Additive Increase and Multiplicative Decrease* (AIMD) approach to penalizing from [23] which means that an agent's utility grows steadily but decreases rapidly for every malicious act.

With the scenario described above, where we have one party being malicious and the other playing honestly, we can summarize the payoffs below and in Table II.

$$\begin{aligned} u(V_1)_{D,H} &= (\alpha + \Delta)\mathbf{R} - P \\ and \; u(V_2)_{D,H} &= \beta\mathbf{R} \end{aligned} \tag{2}$$

If the malicious party was V_2, the payoffs are symmetrical.

$$\begin{aligned} u(V_1)_{H,D} &= \alpha\mathbf{R} \; and \\ u(V_2)_{H,D} &= (\beta + \Delta)\mathbf{R} - P \end{aligned} \tag{3}$$

If both parties, however, play dishonestly D, D, for simplicity, let us assume the malicious acts have the same weight. Their payoffs will be identical to when they both played honestly, but this time there will be a penalty attached. The payoff is given below.

$$\begin{aligned} u(V_1)_{D,D} &= \alpha\mathbf{R} - P \; and \\ u(V_2)_{D,D} &= \beta\mathbf{R} - P \end{aligned} \tag{4}$$

From the set of payoffs above, intuitively, if the benefit from being dishonest is greater than the attached penalty, the players will have no motivation to play honestly since each player is interested in maximizing its payoff from the system. To ensure this is not the case and the players are forced to act honestly, the system must be designed in such a way that the validators self-police themselves, disclosing dishonest players and imposing a severe penalty on such dishonest player(s), while the honest players are adequately rewarded. To achieve this in our system, we designed a reward and penalty model using a game-theoretic approach such that honest behavior is enforced among the players and no player has an incentive to deviate from not being malicious.

TABLE II
Payoff Matrix for Validator's Dilemma

V1, V2	Honest (H)	Dishonest (D)
H	$\alpha\mathbf{R}$, $\beta\mathbf{R}$	$\alpha\mathbf{R}$, $(\beta + \Delta)\mathbf{R} - P$
D	$(\alpha + \Delta)\mathbf{R} - P$, $\beta\mathbf{R}$	$\alpha\mathbf{R} - P$, $\beta\mathbf{R} - P$

B. Game Description

Usually, in game theory, a game consists of a set of players, a set of actions, and a set of payoffs. The goal of each player is to maximize its eventual payoff through the set of actions it takes, regarded as strategies. The players arrive at a Nash equilibrium when neither of them can increase its payoff by altering its strategy while other players keep their strategies the same. The payoff for a player is the reward (or punishment) the player receives when they play a certain strategy in a game.

To achieve our goal, we will model a game such that we arrive at an *'honest-honest'* Nash Equilibrium where the players have no incentive to deviate given the strategy played by the other players.

Definition 1: A game, Γ in our context consist of a set of players (validators (Vs)), a set of actions, A and a utility function, u which determines the payoff of a player based on the strategy, σ. σ represents the combination of actions the player chooses to play.

$$\Gamma = (V_i, A_i, u_i)$$

where V_i are the players, A_i are the actions and u_i is the utility function $u_i : A \rightarrow R$ for each player V_i. R is the reward for the game.

Definition 2: The utility function u, illustrates the V_i's preference over different outcomes, if for any action $a \in A$ and $b \in A$, $u(a) > u(b)$ if and only if V_i prefers a to b and he weakly prefers outcome a over b if $u(a) \geq u(b)$.

Indefinitely Repeated Two-player Game Formulation and Components

Until now, the discussion has considered a setting where the game seems to be a one-shot game. However, in practice, the consensus process is continuous and the actions of a player in one game may influence the actions of the other players in the next game. For this reason, to study the interaction between the consensus nodes, we formulate the problem as a repeated Prisoner's Dilemma game.

Indefinitely repeated games give us insights into the structure of the behavior of individuals when they interact repeatedly, since the previous outcomes of their interactions affect their future behaviors, and repeated games can help to enforce cooperative behavior [22]. With this form of interaction, the players seek how they can receive the maximum short-term and long-term rewards. Such situations are modeled in game theory by repeated games [21]. For example, if player one betrays player two in a round, player two can betray player one in the next round as a punishment. However, if the threat of future penalties is strong enough, both players may choose to cooperate to avoid punishment. The players will cooperate to achieve Nash equilibrium.

With this, the notion of a discount factor is considered. The idea of a discount factor is that a player may be deterred from pursuing a short-term gain by the threat of a penalty that reduces its long-term gain, which correlates to what we saw in the following equation (4) that for the players to be discouraged from the unique Nash equilibrium of (D, D) to a unique Nash equilibrium of (H, H), the cost and penalty attached to malicious acts has to be greater than the reward possible from being dishonest. In a repeated game, the outcome is a sequence of "discounted" outcomes of a strategic game. Each player V_i has a payoff function u_i as seen in the previous section, but this time also includes a discount factor δ between 0 and 1 such that its long-term utility in the game is given by

$$u_i(a_1) + \delta u_i(a_2) + \delta^2 u_i(a_3) + \cdots + \delta^{N-1} u_i(a_N) = \sum_{n=1}^{N} \delta^{n-1} u_i(a_n) \quad (5)$$

where a_n represents a player's action in round n of the game.

IV. Analysis of Game Model with Incentive and Punishment Mechanism

Recall that the discount factor δ represents the probability of a player continuing in the next stage of the game. This means that a higher δ means a higher chance of surviving into the next period and the more patience the player has. For instance, in a reputation-based setting where the penalty of a malicious is a swift multiplicative decrease in the player's reputation value (say by a factor of 2), and to be chosen for the next round of consensus, the player requires a reputation value greater than a certain threshold. If the player puts more value in the long-term, i.e., being chosen in future rounds, they are going to avoid dishonesty and, in essence, cooperate and play according to the system protocol. Otherwise, if the player goes for a short-term gain, their reputation will be reduced and they will not be eligible for selection in the next period. This will mean that, for us to achieve an (H, H) sub-game perfect equilibrium situation, we must ensure that the penalty is a strong enough threat for the players and consequently raise the discount factor closer to 1.

In an infinitely repeated prisoner's dilemma, the *grim trigger* strategy is a strategy that has been identified to encourage cooperation between the players [18], [22]. In a grim strategy, a player starts with cooperation (playing honestly) and continues to cooperate unless the other player deviates (acting maliciously) at some point. This works even in the context of our work since a player has to start with cooperation and continue to cooperate to build their reputation to grow above the selection threshold. The task here then becomes, how do we ensure that this player does not attack the system (play maliciously) after being selected to participate in the consensus process? A good answer to this question is to ensure that the discount factor is as close to 1 as possible, which can be achieved by having a strong penalty as discussed earlier.

A. Mathematical Analysis

When playing the grim trigger strategy, two conditions should be checked. *(1.) when dishonesty is triggered, is it an equilibrium to continue to be malicious forever? and (2.) if there has been no dishonesty in the past, would any player want to be malicious?*

For the first condition, if a player is dishonest and grim strategy is triggered, both players will continue to play (D, D) indefinitely. Although as we saw earlier, (D, D) is a unique Nash equilibrium, it is however not what we want in our system design. Our system already neutralizes this setting since a player's reputation value is swiftly decreased and would be ineligible for subsequent periods the player will not be able to continue playing D.

For the second condition, we can take advantage of the *One-shot deviation principle* [35] which will enable us to ignore complex deviations and only check whether a player would not want to deviate to maliciousness in each period

other players play honestly. Now, taking our payoff matrix of Table II and factoring the discount factor δ into the setting, using equation (5), when V_1 plays the grim strategy its long-term utility is computed as below. If she continues to play honestly even after the first period, her payoff will be

$$u(H) = \alpha\mathbf{R} + \delta\alpha\mathbf{R} + \delta^2\alpha\mathbf{R} + \delta^3\alpha\mathbf{R} + \cdots = \frac{\alpha\mathbf{R}}{1-\delta}$$

However, if she decides to deviate after the first period, her payoff will be

$$u(D) = [(\alpha+\Delta)\mathbf{R} - P] + \delta(\alpha\mathbf{R} - P) + \delta^2(\alpha\mathbf{R} - P) + \cdots$$

$$= [(\alpha+\Delta)\mathbf{R} - P] + \frac{\delta(\alpha\mathbf{R} - P)}{1-\delta}$$

For cooperation to persist according to the requirement of our system, $u(H) \geq u(D)$ must hold. That is,

$$\frac{\alpha\mathbf{R}}{1-\delta} \geq [(\alpha+\Delta)\mathbf{R} - P] + \frac{\delta(\alpha\mathbf{R} - P)}{1-\delta} \tag{6}$$

From the inequality in equation 6, we can compute what the discount factor δ will be for V_1 to keep being honest,

$$\delta \geq 1 - \frac{P}{\Delta\mathbf{R}} \tag{7}$$

It is straight forward to observe that the payoff for V_2 in a similar setting will be identical with β replacing α since the payoff matrix in Table II is symmetrical.

$$u(H) = \beta\mathbf{R} + \delta\beta\mathbf{R} + \delta^2\beta\mathbf{R} + \delta^3\beta\mathbf{R} + \cdots = \frac{\beta\mathbf{R}}{1-\delta}$$

$$u(D) = [(\beta+\Delta)\mathbf{R} - P] + \delta(\beta\mathbf{R} - P) + \delta^2(\beta\mathbf{R} - P) + \cdots$$

$$= [(\beta+\Delta)\mathbf{R} - P] + \frac{\delta(\beta\mathbf{R} - P)}{1-\delta}$$

This will result to a discount factor of $\delta \geq 1 - \frac{P}{\Delta\mathbf{R}}$ which corresponds to what we have in equation (7). The implication of this is that the ratio $\frac{P}{\Delta\mathbf{R}}$ should be very small so that the discount factor δ is as high as possible (close to 1) in order to ensure that the players do not have an incentive to deviate from honest behavior.

B. Reducing the ratio $P/\Delta\mathbf{R}$

As we already saw from the previous subsection, the discount factor, which represents how much a player values participating in future periods has to be sufficiently large to deter a player from dishonesty. There are three variables in the ratio $P/\Delta\mathbf{R}$ which determine the discount factor. A change in any of them can affect the discount factor. P which is the penalty imposed on a malicious player, can be made stronger to deter the players from dishonesty. For instance, in a reputation-based system, the *Additive Increase and Multiplicative Decrease* approach can be adopted so that a player's reputation is drastically reduced when they are malicious, while players that report malicious behaviors can be rewarded with an additional increment in their reputation value. Such a system will also help to address Δ which is the increase in selection odds for a malicious player, since malicious activities are monitored and better reported by other players for a reward.

Finally, $\mathbf{R}$ which is the reward a player (i.e., validator) gets from adding a new block, can also be decreased to discourage the players from performing malicous acts. However, this may result in an unwanted effect since the motivation for participation in the consensus process is the reward itself.

V. Model Evaluation

In this section, we discuss the performance of our model in response to changes in the various parameters discussed in the previous section (see Sec. IV-B). The objectives of our evaluation are in two categories. *(1). To determine the optimal discount factor for our system by varying the parameters in equation (7). The optimal value of the discount factor is the value when all the agents have no incentive to deviate from the protocol, thereby ensuring cooperative and honest behavior of the agents. (2). To examine the impact of the AIMD penalty approach on the stake (e.g., reputation) and total payoff of an agent.*

TABLE III
Evaluation Parameters

Parameter	Value 1	Value 2
P	0.5	{0.5, 0.25, 0.125, 0.05}
δ	-	0.99
Δ	0.1 - 0.6	0.6
Weight (θ and ϕ)	-	0.01 - 0.99
$\mathbf{R}$	$1-100$	100

For the first objective, using the parameters in the second column of Table III, initially, we fix the value for penalty P at 0.5 since we are using the AIMD approach that steadily increases but swiftly decreases. This means that the stake of an agent drastically cuts by half whenever it is penalized for malicious action. We start with a 0.1 increment in selection odd Δ for a malicious player and steadily vary the reward to be earned by an agent when it successfully adds transactions to a block and adds a block to the chain, beginning with a reward of $\mathbf{R}$ of 1 unit and slowly increasing to 100 units while observing the changes in the discount factor. We repeat the experiment by increasing the Δ gradually to find the best value that returns an optimal discount factor.

From Figure 1, we can observe that Δ fixed at 0.1 through 0.5, the first value for the discount factor (δ) is ≤ 0, but at a value of 0.6, the δ has a minimum value of 0.167. This means that the least value for the increase in odd for our system is 0.6 since the discount factor—a probability—can not be negative. Also from the figure, looking at the curve for Δ fixed at 0.6, with a small increase in the reward, the discount factor sharply increases until about the 30 unit mark, where it saddles and becomes very close to 1 (0.99 to be specific). At a discount

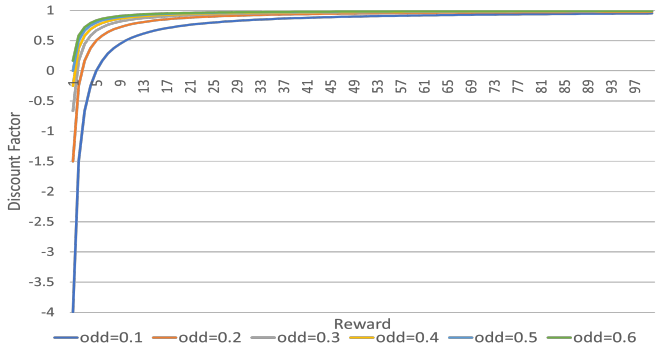

Fig. 1. Relationship between Reward and the Discount Factor with an Increase in Chance of Agent Selection Due to Maliciousness Held Constant

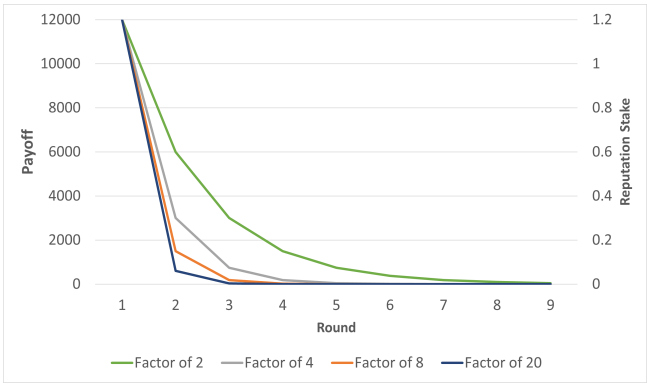

Fig. 3. The Drop in Reputation Stake of a Malicious Agent over Multiple Rounds with Varying Penalty Factors

factor of below 0.99, the agents may still have an incentive to deviate from honest behavior since a rational player would place more value on maximizing its immediate reward than waiting for the future.

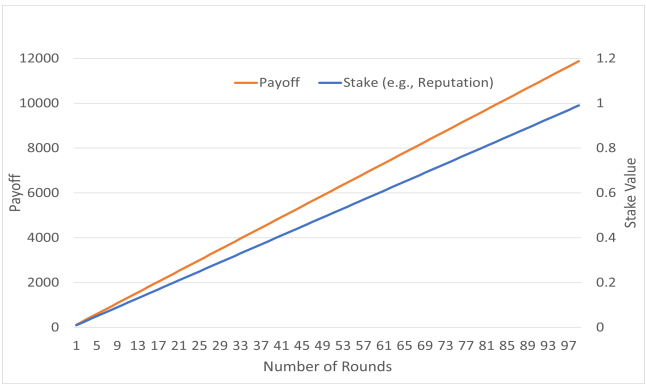

Fig. 2. The Variation in Agent Payoff and Reputation Over Multiple Rounds

However, with the discount factor set at ≥ 0.99 (i.e., $0.99 < \delta < 1$) and the other required parameters set as in the third column of the Table III, we forced the agents to play honestly or else get severely penalized and eventually removed from the system. In Figure 2, we can observe the steady growth of a consistently honest agent in terms of its total payoff and its stake in the form of reputation over multiple rounds. We, however, can observe from Figure 3 that a malicious agent will be swiftly penalized and its reputation drops sharply, contrary to how long it takes to grow to the same level as seen from Figure 2. With this drop in stake, once the agent falls below an acceptable threshold, it can no longer be selected for the consensus process. The Figure 3 shows the impact of different penalty factors on the drop in reputation for a malicious agent. We can see that the greater the factor, the more drastic the drop in reputation for the malicious agent.

VI. Conclusion and Future Work

Having a consensus protocol model where the participants can monitor themselves by receiving rewards for reporting attackers and malicious activities helps to improve the security of the system. In this paper, we have shown how honesty can be achieved in a competitive blockchain environment by analyzing the short and long-term payoffs for validators. We show that with a sufficient discount factor and the threat of a penalty for the future, a validator can be forced to act according to the consensus protocol.

As a follow-up to this paper, we plan to implement our model on a real blockchain platform, adjusting the parameters discussed in section IV-B while optimizing for what will be the optimum strategy to keep all validators honest in the long term. We also plan to investigate possible threat scenarios against our model, especially threat vectors related to reputation systems.

References

[1] Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press, Cambridge (1994)
[2] Liu, Z., Luong, N. C., Wang, W., Niyato, D., Wang, P., Liang, Y. C., and Kim, D. I. (2019). A survey on blockchain: A game theoretical perspective. IEEE Access, 7, 47615-47643.
[3] Johnson, B., Laszka, A., Grossklags, J., Vasek, M., Moore, T.: Game-theoretic analysis of DDoS attacks against bitcoin mining pools. In: International Conference on Financial Cryptography and Data Security, pp. 72–86. Springer (2014).
[4] Y. He, H. Li, X. Cheng, Y. Liu, C. Yang and L. Sun, "A Blockchain Based Truthful Incentive Mechanism for Distributed P2P Applications," in IEEE Access, vol. 6, pp. 27324-27335, 2018.
[5] C. Tang, L. Wu, G. Wen and Z. Zheng, "Incentivizing Honest Mining in Blockchain Networks: A Reputation Approach," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 1, pp. 117-121.
[6] Nojoumian, M., Golchubian, A., Njilla, L., Kwiat, K., and Kamhoua, C. (2018, July). Incentivizing blockchain miners to avoid dishonest mining strategies by a reputation-based paradigm. In Science and Information Conference (pp. 1118-1134). Springer, Cham.
[7] Trestian, Ramona , Ormond, Olga and Muntean, Gabriel-Miro (2011) Reputation-based network selection mechanism using game theory. Physical Communication, 4 (3) . pp. 156-171. ISSN 1874-4907 [Article].
[8] Zainib Noshad, Asad Ullah Khan, Shahid Abbas, Zain Abubaker, Nadeem Javaid, Muhammad Shafiq, Jin-Ghoo Choi, "An Incentive and Reputation Mechanism Based on Blockchain for Crowd Sensing Network", Journal of Sensors, vol. 2021, Article ID 1798256, 14 pages,
[9] Luu, L., Saha, R., Parameshwaran, I., Saxena, P., Hobor, A.: On power splitting games in distributed computation: the case of bitcoin pooled mining. In: Computer Security Foundations Symposium (CSF), 2015 IEEE 28th, pp. 397–411. IEEE (2015)
[10] Eyal, I.: The miner's dilemma. In: 2015 IEEE Symposium on Security and Privacy (SP), pp. 89–103. IEEE (2015)

[11] Kroll, J.A., Davey, I.C., Felten, E.W. (2013), "The economics of bitcoin mining, or bitcoin in the presence of adversaries".
[12] M. H. Ashik, M. M. Shahriar Maswood, A. G. Alharbi and D. Medhi, "FPoW: An ASIC-resistant Proof-of-Work for Blockchain Applications," 2020 IEEE Region 10 Symposium (TENSYMP), 2020, pp. 1608-1611,
[13] Y. Georghiades, S. Flolid and S. Vishwanath, "HashCore: Proof-of-Work Functions for General Purpose Processors," 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, pp. 1951-1959, doi: 10.1109/ICDCS.2019.00193.
[14] H. Cho, "ASIC-Resistance of Multi-Hash Proof-of-Work Mechanisms for Blockchain Consensus Protocols," in IEEE Access, vol. 6, pp. 66210-66222, 2018, doi: 10.1109/ACCESS.2018.2878895.
[15] H. Cho, " Proof-of-CAPTCHA: A True ASIC-Resistant Consensus" in International Journal of Advanced Research in Computer and Communication Engineering, Vol. 8, Issue 8, August 2019,
[16] Leslie, C. R. (2005). Antitrust amnesty, game theory, and cartel stability. J. Corp. L., 31, 453.
[17] Fulton, J. R., and Adamowicz, W. L. (1993). Factors that influence the commitment of members to their cooperative organization. Journal of Agricultural Cooperation, 8(1141-2016-92587), 39-53.
[18] Fan, C. P. (2000). Teaching children cooperation—An application of experimental game theory. Journal of Economic Behavior and Organization, 41(3), 191-209.
[19] Haugen, K. K. (2004). The performance-enhancing drug game. Journal of Sports Economics, 5(1), 67-86.
[20] Mohan, V. (2019). On the use of blockchain-based mechanisms to tackle academic misconduct. Research Policy, 48(9), 103805.
[21] A. Dixit, S. Skeath, Games of Strategy, W.W. Norton and Company, New York, 1999.
[22] Axelrod, R. The Evolution of Cooperation; Basic Books: New York, NY, USA, 1984. Available at https://ee.stanford.edu/ hellman/Breakthrough/book/pdfs/axelrod.pdf, Last Accessed: 23 July, 2022.
[23] A. Salau, R. Dantu and K. Upadhyay, (2021), "Data Cooperatives for Neighborhood Watch," 2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2021, pp. 1-9, doi: 10.1109/ICBC51069.2021.9461056.
[24] Rosenfeld, M.: Analysis of bitcoin pooled mining reward systems. arXiv preprint arXiv:1112.4980 (2011)
[25] S. Narang, M. Byali, P. Dayama, V. Pandit and Y. Narahari, "Design of Trusted B2B Market Platforms using Permissioned Blockchains and Game Theory," 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2019, pp. 385-393,
[26] Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017, June). An overview of blockchain technology: Architecture, consensus, and future trends. In 2017 IEEE international congress on big data (BigData congress) (pp. 557-564). IEEE.
[27] Dong, X., Wu, F., Faree, A., Guo, D., Shen, Y., & Ma, J. (2019). Selfholding: A combined attack model using selfish mining with block withholding attack. Computers & Security, 87, 101584.
[28] Leonardos, S., Reijsbergen, D., & Piliouras, G. (2020). Weighted voting on the blockchain: Improving consensus in proof of stake protocols. International Journal of Network Management, 30(5), e2093.
[29] Kiayias, A., Koutsoupias, E., Kyropoulou, M., & Tselekounis, Y. (2016, July). Blockchain mining games. In Proceedings of the 2016 ACM Conference on Economics and Computation (pp. 365-382).
[30] Alzahrani, N., & Bulusu, N. (2018, October). Towards true decentralization: A blockchain consensus protocol based on game theory and randomness. In International conference on decision and game theory for security (pp. 465-485). Springer, Cham.
[31] Motepalli, S., & Jacobsen, H. A. (2021, September). Reward mechanism for blockchains using evolutionary game theory. In 2021 3rd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS) (pp. 217-224). IEEE.
[32] Neuder, M., Moroz, D. J., Rao, R., & Parkes, D. C. (2019). Selfish behavior in the tezos proof-of-stake protocol. arXiv preprint arXiv:1912.02954.
[33] Fanti, G., Kogan, L., Oh, S., Ruan, K., Viswanath, P., & Wang, G. (2019, February). Compounding of wealth in proof-of-stake cryptocurrencies. In International conference on financial cryptography and data security (pp. 42-61). Springer, Cham.
[34] J. R. Marden and M. Effros, "The Price of Selfishness in Network Coding," in IEEE Transactions on Information Theory, vol. 58, no. 4, pp. 2349-2361, April 2012, doi: 10.1109/TIT.2011.2177576.
[35] Hendon, Ebbe, Hans Jørgen Jacobsen, and Birgitte Sloth. "The one-shot-deviation principle for sequential rationality." Games and Economic Behavior 12.2 (1996): 274-282.
[36] Eyal, I., & Sirer, E. G. (2014, March). Majority is not enough: Bitcoin mining is vulnerable. In International conference on financial cryptography and data security (pp. 436-454). Springer, Berlin, Heidelberg.
[37] Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Suman, R. (2021). Blockchain technology applications for Industry 4.0: A literature-based review. Blockchain: Research and Applications, 100027.