Cryptocurrency Mining Strategies

CS4991 Capstone Report, 2022

Maximilian Dawkins
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA

Abstract
In Proof of Work (PoW) cryptocurrency systems such as Bitcoin, machine learning can be used to learn the most profitable mining strategy. This paper discusses past approaches in learning mining strategies, as well as future directions for research in cryptocurrency mining strategy. After simple optimization methods, Markov Decision Processes (MDPs) were first used to learn mining strategies. Then, reinforcement learning (RL) was later used as an improvement over the MDP method. These methods, however, assume that other miners are honest miners, and future research could improve upon the RL method to include adversaries that do not mine honestly.

1 INTRODUCTION

Bitcoin has held a value of over 30,000 US dollars in the past year, and the only way to generate new Bitcoin is by mining [6] [7]. Bitcoin is one of many PoW cryptocurrencies, meaning the mining process involves solving a computationally difficult hash puzzle [7]. The first one to solve the puzzle gets to successfully mine a block on the Bitcoin blockchain and receives 6.25 Bitcoin as a mining reward. This mining reward is very valuable, and Bitcoin mining efforts have progressed very far over the past decade. Miners have built specialized hardware, created group mining pools with split profits, and even run secretive scams and cyber-attacks to mine Bitcoin on other people's computers using their electricity [8]. Controlling more computational power leads to a greater chance of mining blocks and receiving rewards, so most mining efforts focus on computational power. However, miners do not have to submit a block that solves the mining puzzle as soon as it is found, which leads to new mining strategies that generate greater rewards.

2 RELATED WORKS

Bitcoin runs on a blockchain, which is a series of connected "blocks" that store information about transactions of Bitcoin that have occurred between Bitcoin accounts. The technical details of Bitcoin-like cryptocurrencies are described by Florian Tschorsch and Björn Scheuermann in [7].

The Bitcoin system is decentralized, so anyone can submit blocks to be added to the blockchain. For a block to be added, however, a majority of the "nodes," or running programs that keep track of the blockchain, must agree that the block is valid. For Bitcoin, this is done by ensuring that the SHA-256 hash of the block is less than a certain target value [7]. The SHA-256 hash converts some data, like a block for example, into a 256-bit number. With 256 bits, this number can be up to about 10^{77}, and if the target is small enough, then the chance that a block satisfies this hash condition can be absurdly low. This is the essence of the hash puzzle for PoW cryptocurrencies, and why it is computationally difficult to mine them.

As described by Satoshi Nakamoto in the Bitcoin whitepaper, a block is supposed to be mined about every 10 minutes, so there is a mining reward to both incentivize mining and create new cryptocurrency [9]. This reward is highly sought after, so people have used many methods to increase their mining capabilities. There are many ways that miners increase their computational power, but the focus of this paper is on different mining strategies. A mining strategy refers to the decisions that a miner makes on which block to mine and when to publish mined blocks. The most obvious mining strategy is called honest mining, and it was originally thought to be the most profitable strategy. Honest mining occurs when the miner always mines the latest block and publishes mined blocks as soon as they are found. This is also the strategy that is intended and expected by PoW cryptocurrencies as laid out in [9].

However, other strategies have been discovered that can yield higher rewards. A large concern with new strategies is that miners will be incentivized to work together with those mining more optimally, and a pool of miners will grow until they collectively control over 50% of mining capabilities. Since miners control which blocks get mined, having a majority group like this would allow them to control which blocks get included in the blockchain. This would allow them to exclude blocks outside the pool and gain all the mining rewards. More miners would join the pool, and the cryptocurrency would eventually become a centralized system controlled by the majority pool.

Ittay Eyal and Emin Gun Sirer discuss another mining strategy in [3] called selfish mining, which was later developed further by others. Ayelet Sapirshtein, Yonatan Sompolinsky, and Aviv Zohar first proposed a method for finding the optimal selfish mining strategy using MDPs in [1], and a better method that used RL was developed in [2] by Taotao Wang, Soung Chang Liew, and Shengli Zhang several years later.

3 SELFISH MINING

In 2014, [3] showed that another strategy called selfish mining can earn higher rewards by deviating from the honest mining protocol. The strategy relies on the fact that PoW cryptocurrencies always use the longest blockchain as the official blockchain [7]. Sometimes two miners mine a block at the same time, creating two new blocks A and B. In this case, both blocks are published, and miners can choose to mine off of either block. This creates a fork in the blockchain with miners mining on both sides of the fork. If a miner mines a block after block A first, then block A and the new block will become the longest chain and side A of the fork will become the official blockchain. If a miner instead mines a block after block B first, then side B of the fork would become the official blockchain.

When a selfish miner mines a new block, it takes advantage of the longest chain rule by choosing not to publish the block and starts keeping track of a private chain instead. The selfish miner will then mine on the private chain, giving the selfish miner a head start on mining the next block. If the honest miners catch up to the selfish miners by mining a block on the public chain, then the selfish miner will simply publish the private block that it mined, creating a fork. If the selfish miner can instead mine another block before the honest miners, they will again keep it hidden, and they will have a two-block lead on the honest miners.

With a lead of two or more blocks, the selfish miner will publish one block whenever the honest miners mine a block, keeping their lead while constantly causing forks to divide the honest miners' mining efforts. The head

start that the selfish miner gets from keeping some of their blocks private allows them to get credit for more of the mined blocks and increases their rewards. It was shown in [3] that if a miner has greater than 1/3 of the mining power, then they can use selfish mining to increase their revenue.

4 OPTIMAL MINING STRATEGIES

Following the discovery of selfish mining, Ayelet Sapirshtein, Yonatan Sompolinsky, and Aviv Zohar showed in [1] that even better strategies exist, and they came up with an algorithm to find the optimal mining strategy given certain parameters.

4.1 MDP Model

[1] first created a model for the mining environment which was similar to an MDP. Like the selfish mining algorithm in [3], the MDP method has a public chain that contains all the currently published blocks, and a private chain that contains the extra blocks that the selfish miner has not yet published. The MDP-like model in [1] has a set of "states", which represents the possible differences in length of the two chains, and a set of "actions", which represents what the miner can do at each state.

The model also defined the probabilities of going from one state to another given each action, and the mining rewards for transitioning to a state after following an action. This model is essentially what a MDP is, where there are states connected by actions, with probabilities and rewards corresponding to paths taken from one state to another after a given action is performed.

The difference between this model and a normal MDP is that in the objective function. Normal MDPs determine which actions are the best to choose in any state by trying to maximize some sum of rewards over many actions and transitions between states. This sum of rewards is called the objective function, and it is a linear sum of rewards in a normal MDP. In the mining model, the objective function is not a linear function because it is based on the ratio of the dishonest miner's rewards to the total rewards. To solve this problem, [1] converted the non-linear MDP into a set of linear MDPs, and then used a standard MDP solver to solve them while doing a search over the set of linear MDPs to find the optimal action policy.

4.2 RL Model

Although it was shown in [1] that optimal mining strategies can be found, the MDP model requires knowledge of certain parameters specific to the mining environment such as the true mining power of the miner. Since these parameters change over time and are difficult to know, [2] used a Q-learning RL model to find optimal strategies without knowing these parameters. Q-learning works by training an "agent" to perform actions in an environment. The possible actions in the RL model are the same as in the MDP model from [1]: Adopt the public blockchain and drop the private blockchain; Override the public blockchain by publishing one more block than the honest chain; Match the public blockchain by publishing an equal number of blocks and creating a fork; and Wait for the next block to be mined without changing anything. After the agent takes an action, it receives a reward based on the mining reward, and updates its strategy based on the reward.

The way a Q-learning strategy updates is by updating Q-values for each state-action pair. For example, if the miner becomes one block ahead of the honest miner and chooses to wait, then the Q-value for (1 block ahead, wait) might increase based on the reward for that action. Then, that action would be preferred because of the higher Q-value. Q-

learning does not require a model of the environment like the MDP because it just interacts directly with the environment and learns from the experiences. However, [2] still had to solve the problem of the non-linear objective function for RL that [1] solved in the MDP. To do this [2] modified the standard Q-learning algorithm to keep track of a pair of Q-values for each state-action pair, one each for the rewards of the honest and dishonest miners. Then, for a given state of the environment, the agent decides its action based on the Q-value pairs for the state-action pairs at that state. The simulation results from [2] show that the Q-learning model is still able to find optimal strategies.

5 OPTIMAL MINING IN MINING POOLS

The primary concern with dishonest mining is the development of a majority mining group, which would most likely occur in the form of a mining pool. Mining pools are groups of people who collectively mine a cryptocurrency and split the profits proportionally to mining contribution within the pool. In November 2021, there were multiple Bitcoin mining pools with over 15% of the mining capabilities [4]. However, [5] showed a countermeasure to be used against any mining pool that were to implement a selfish mining strategy. Mining pools distribute the mining task to its members, which includes the block that is being mined. If the mining pool is mining selfishly, the block sent out to the pool might not match the current block of the public chain. This would prove that the pool is mining selfishly, and the selfish mining strategy of the pool can be exploited.

6 CONCLUSIONS

The profitability of cryptocurrency mining has led people to develop numerous selfish mining strategies. These strategies are potentially dangerous to PoW crypto-currencies, and it is important to understand them and their limitations. So far, the Q-learning model from [2] is the strongest algorithm for finding optimal mining strategies because it can adapt to a dynamic environment where it does not have perfect information. However, mining pools, the most likely user of selfish mining strategies, reveal a lot of information that can be used to detect and defend against selfish mining.

7 FUTURE WORK

The mining strategies discussed in this paper assumed that other miners were honest. A future direction of research would be to observe how the RL model from [2] performs in environment simulations where other dishonest miners exist. Also, the real world differs from the assumptions of the discussed models in that there are many separate mining entities as opposed to only an honest and selfish group. This assumption is reasonable if all other miners are assumed to be honest, but more than two different mining entities should be simulated in an environment with other dishonest miners. Future research along these lines could provide valuable insights into potential limitations and countermeasures to the RL model's optimal mining strategies.

8 CS PROGRAM EVALUATION

Overall, the computer science program at UVA prepared me well for this research project. The Machine Learning, Artificial Intelligence, and Cryptocurrency classes gave me all the necessary conceptual knowledge required to provide analysis of the research in cryptocurrency mining strategies. The one enhancement I suggest is to require some reading of computer science research, which is something I have not done in any computer science classes at UVA. In many sectors of computer science, it is important to keep up with current research, and a struggle in

my research was getting used to reading technical computer science research papers.

REFERENCES

[1] Ayelet Sapirshtein, Yonatan Sompolinsky, and Aviv Zohar. 2015. Optimal Selfish Mining Strategies in Bitcoin. arXiv:1507.06183v2. Retrieved from https://arxiv.org/abs/1507.06183v2

[2] Taotao Wang, Soung Chang Liew, and Shengli Zhang. 2021. When Blockchain Meets AI: Optimal Mining Strategy Achieved By Machine Learning. arXiv:1911.12942. Retrieved from https://arxiv.org/abs/1911.12942v3

[3] Ittay Eyal and Emin Gun Sirer. Majority is not Enough: Bitcoin Mining is Vulnerable. arXiv:1311.0243v5. Retrieved from https://arxiv.org/abs/1311.0243

[4] Arijit Sarkar. 2021. Foundry USA becomes second-largest Bitcoin mining pool amid China ban. (November 2021). Retrieved from https://cointelegraph.com/news/foundry-usa-becomes-second-largest-bitcoin-mining-pool-amid-china-ban

[5] Suhyeon Lee and Seungjoo Kim. 2018. Pooled Mining Makes Selfish Mining Tricky. Retrieved from https://eprint.iacr.org/2018/1230.pdf

[6] CoinDesk. 2022. Consensus 2022. Retrieved from https://www.coindesk.com/price/bitcoin/

[7] Florian Tschorsch and Björn Scheuermann. 2015. Bitcoin and Beyond: A Technical Survey on Decentralized Digital Currencies. Retrieved from https://eprint.iacr.org/2015/464.pdf

[8] Wikipedia. 2017. Proof of Work. Retrieved from https://en.wikipedia.org/wiki/Proof_of_work

[9] Satoshi Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system. Retrieved from https://bitcoin.org/bitcoin.p