The Tight Bound for Pure Price of Anarchy in an Extended Miner's Dilemma Game

北京大学前沿计算研究中心
Center on Frontiers of Computing Studies, Peking University

Qian Wang, Yurong Chen
CFCS, Computer Science Dept., Peking University

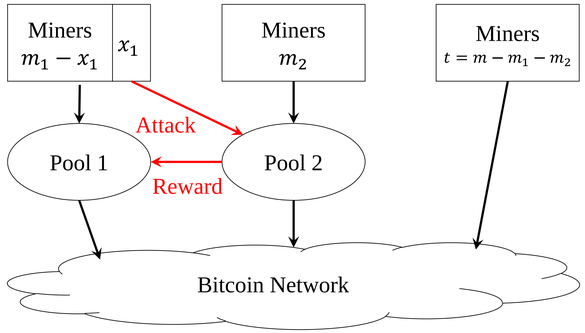
AAMAS2021 ONLINE, MAY 3–7

Problem

Block Withholding Attack

Instead of mining honestly, pools can be incentivized to infiltrate their own miners into other pools. These infiltrators report partial solutions but withhold full solutions, share block rewards but make no contribution to block mining.

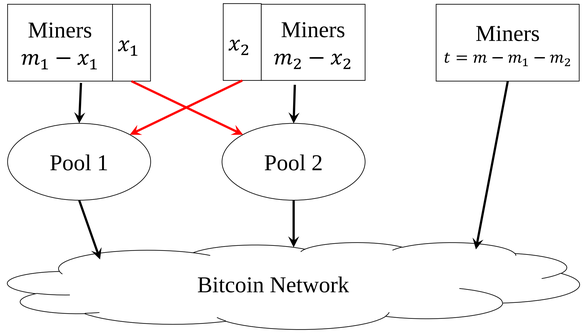
One-attacker scenario



Miner's Dilemma

The block withholding attack among mining pools can be modeled as a non-cooperative game called "the miner's dilemma", which reduces effective mining power in the system and leads to potential systemic instability in the blockchain.

Two pools attacking each other



Methodology

Total Rewards

$$R_1(x_1, x_2) = \frac{m_1 - x_1}{m - x_1 - x_2}$$

$$R_2(x_1, x_2) = \frac{m_2 - x_2}{m - x_1 - x_2}$$

Average Rewards

$$r_1(x_1, x_2) = \frac{R_1(x_1, x_2) + x_1 r_2(x_1, x_2)}{m_1 + x_2}$$

$$r_2(x_1, x_2) = \frac{R_2(x_1, x_2) + x_2 r_1(x_1, x_2)}{m_2 + x_1}$$

Pure NE

$$x_1^* = \underset{x_1 \in [0, m_1]}{\operatorname{argmax}}\ r_1(x_1, x_2^*)$$

$$x_2^* = \underset{x_2 \in [0, m_2]}{\operatorname{argmax}}\ r_2(x_1^*, x_2)$$

Theorem 1 ((Existence of NE)

Every 2-player miner's dilemma game admits at least one pure Nash equilibrium.

Two types of NE

Non-extreme pure NE: $(x_1^*, x_2^*) \in (0, m_1) \times (0, m_2)$ $\implies$ at most one
Extreme pure NE : $x_1^* \in \{0, m_1\}$ or $x_2^* \in \{0, m_2\}$ $\implies$ at most one
They **CANNOT** exist at the same time.

Theorem 2 ((Uniqueness of NE)

Every 2-player miner's dilemma game admits a unique pure Nash equilibrium.

Pure Price of Anarchy

$$PPoA = \frac{optimal\ social\ welfare}{worst\ social\ welfare\ in\ any\ pure\ NE} = \frac{m_1 + m_2}{m_1 + m_2 - x_1 - x_2}$$

Theorem 3 ((Tight Upper Bound of PPoA)

$PPoA \leq 2$, and equal to 2 if and only if $m_1 = m_2 = m/2$.

Extended Game

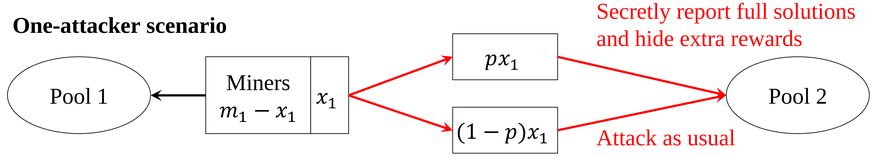
Betrayal Assumption

Since full solutions are also counted as shares, if an infiltrator secretly reports full solutions to the opponent pool, he can get more reward from the opponent mining pool and hide the extra reward for himself.

Here we introduce a betrayal parameter, $p \in [0, 1]$, to represent the percent of betrayal. Notice the standard model is a special case when $p = 0$.

The above results still hold.

One-attacker scenario

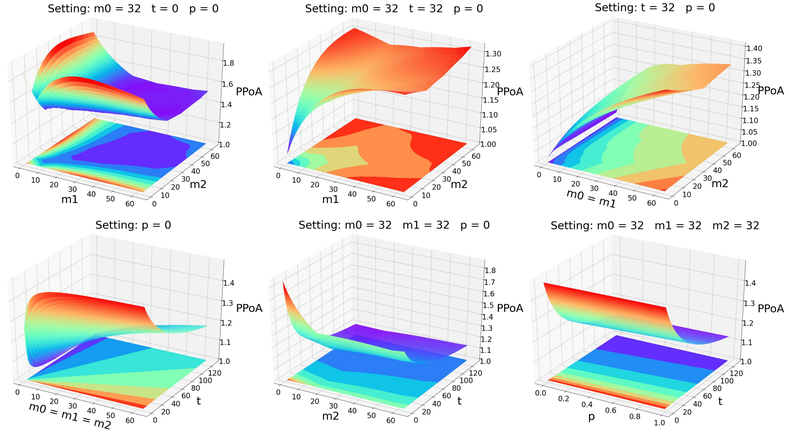


Summary

For every 2-player miner's dilemma game with betrayal assumption:
(1) The pure NE exists and is unique;
(2) The tight bound of $PPoA$ is $(1, 2]$.

Experiment

We focus on the PPoA of N-player miner's dilemma game and set $N = 3$ in our experiments.



Conjecture

For every N-player ($N \geq 2$) miner's dilemma game with betrayal assumption:
(1) The pure NE exists and is unique;
(2) The tight bound of $PPoA$ is $(1, 2]$.

Reference

[1] Colleen Alkalay-Houlihan and Nisarg Shah. 2019. The Pure Price of Anarchy of Pool Block Withholding Attacks in Bitcoin Mining. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33. 1724–1731.
[2] Ittay Eyal. 2015. The miner's dilemma. In 2015 IEEE Symposium on Security and Privacy. IEEE, 89–103.