Reorgs in Ethereum 2.0 and Multi-Agent Selfish Mining

Michael Neuder Yonatan Sompolinsky
Daniel J. Moroz Rithvik Rao David C. Parkes

School of Engineering and Applied Sciences
Harvard University

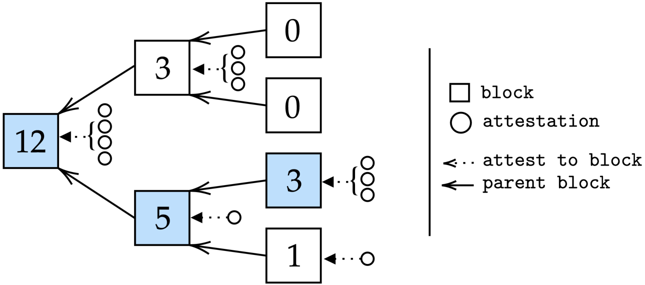
MEV Roast: Reorg Edition
August 2021

Introduction

- Outline Eth 2.0 reorg paper, which appeared in the "Game Theory in Blockchain" workshop at WINE 2020.
- Present our current work on multi-agent selfish mining.

Ethereum 2.0

Fork-Choice Rule

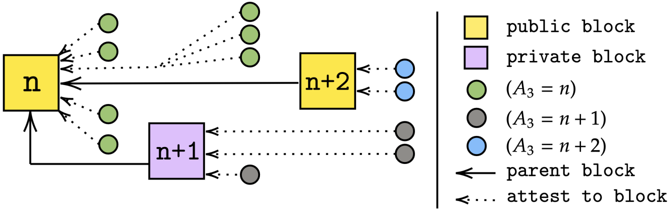


- HLMD-GHOST[1] uses *weight* to determine the head of the canonical chain.
- Each block annotated with its weight.
- Blue blocks are heaviest branch at each fork and thus part of canonical chain.

[1] Hybrid Latest Message Driven Greedy Heaviest Observed SubTree

Malicious Reorgs

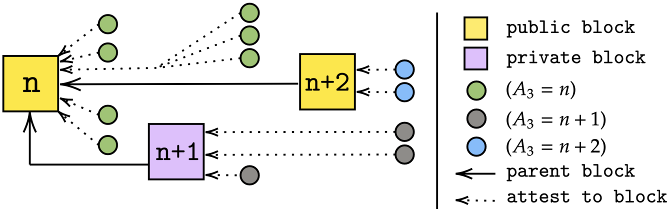
Strategy

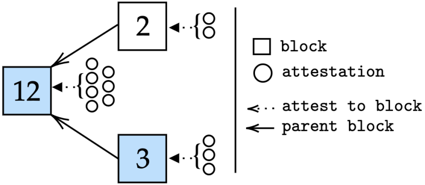


- The attacker privately proposes block $n+1$ and attests with $(A_3 = n+1)$. Honest validators instead attest with $(A_3 = n)$.
- At slot $n+2$, an honest validator will propose a block whose parent is the slot n block.
- The attacker then releases private attestations and block $n+1$, which is seen as the head of the chain by HLMD-GHOST.

Malicious Reorgs

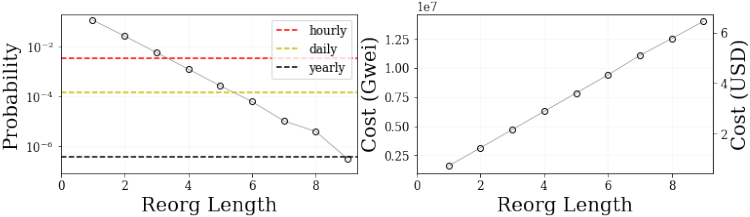
Strategy





Malicious Reorgs

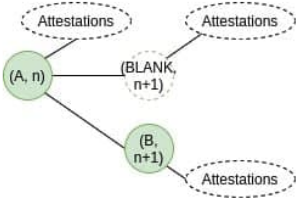
Probability for 30% Attacker



- Use Monte Carlo simulation of 10^7 randomly generated epochs.
- In this case, we only consider reorgs that occur *within a single epoch*.
- Cost is the amount of reward lost, or the opportunity cost of playing this dishonest strategy (no slashing occurs).

Malicious Reorgs

Deterrent



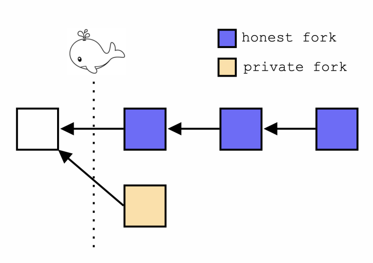
- As proposed in HF1: *Allow attesters to vote for an empty slot.*[2]
- Requires the attacker to control a simple majority of attestations for subsequent blocks.[3]

[2] https://notes.ethereum.org/@vbuterin/HF1_proposal

[3] https://www.paradigm.xyz/2021/07/ethereum-reorgs-after-the-merge/

Selfish Mining

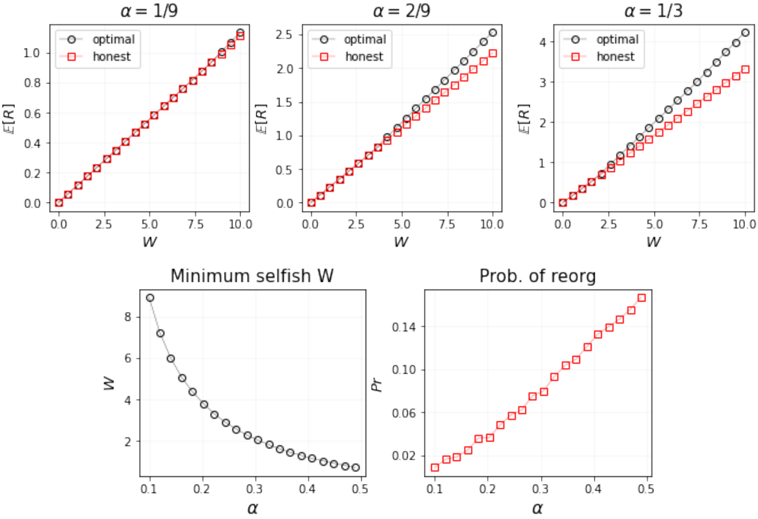
Single Agent with MEV Reward



- A single rational agent with mining power $\alpha \in (0, 0.5)$.
- Remaining $1 - \alpha$ mining power mines honestly.
- Winner gets block rewards and a bonus MEV reward, W.
- *Objective* — find W such that selfish mining is optimal.

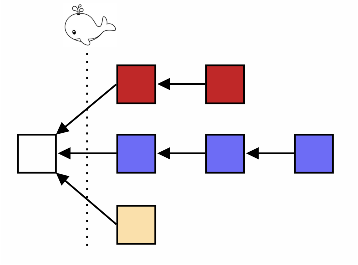
Selfish Mining

Single Agent with MEV Reward



Multi-Agent Selfish Mining

Construction



- A $n-$player stochastic game, where each agent decides to be honest or selfish at each stage.
- Cost for mining for each period is $1/n$, and block reward is 1.
- Winner gets a bonus MEV reward, W.
- Terminates when one fork reaches a fixed length, ℓ.

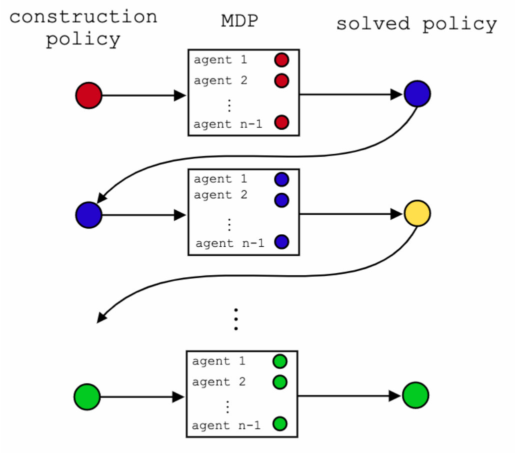
Multi-Agent Selfish Mining

Expectations

- Miners must decide when it is optimal to give up.
- For $W = 0$, honest mining should be optimal
  $\implies$ all-honest equilibrium.
- For large W, selfish mining should be optimal
  $\implies$ all-selfish equilibrium.
- Looking for values of W, where all-selfish and all-honest are simultaneous equilibria.
  - "It is only worth attacking if everyone else is also attacking."
  - *Intuition* — The network is easier to attack if it is fragmented.

Multi-Agent Selfish Mining

Algorithm



Multi-Agent Selfish Mining

Results: 3-player, length-4

- For all $W <= 2.3$, any starting policy converges to all-honest equilibrium.
- For all $W >= 2.8$, any starting policy converges to all-mostly-selfish equilibrium.
  - "If my fork is behind by 2 or more blocks, mine honestly. Otherwise, mine selfishly."
- For all $W \in (2.3, 2.8)$, no convergence.
  - Enters a 2-cycle of slightly different, mostly-selfish policies.

Wrap up

- Need to generalize to non-symmetric case.
- *Goal:* Understand what values of W potentially lead to instability.

Thanks!