Abstract No: SO-02

Optimizing selfish mining strategies through deep reinforcement learning

Wijewardhana W. T. R. N. D. K.[1*], Vidanagamachchi S. M.[1] and Arachchilage N. A. G.[2]

[1]Department of Computer Science, Faculty of Science, University of Ruhuna, Sri Lanka.
[2]School of Computing Technologies, RMIT University, Australia.
dini2040gh@gmail.com*

Selfish mining is a type of mining attack where miners strategically release blocks to create forks in the main branch with the intention of acquiring a large portion of the mining reward. Traditional strategies use a Markov Decision Process (MDP) with a non-linear objective function that requires variable blockchain parameters, which are hard to determine, while model-free approaches like multi-dimensional Q-learning overcome this by learning optimal policies without prior blockchain information. Despite this, existing algorithms remain largely impractical for real blockchain networks, as they fail to account for realistic blockchain features, exhibit inefficient learning in large state spaces, and suffer from slow convergence rates. In this work, we propose a novel model-free Deep Reinforcement Learning (DRL) algorithm for optimal selfish mining, enabling dynamic learning without requiring prior knowledge of network parameters. The study aims to leverage deep neural networks along with advanced exploration and experience replay mechanisms to achieve faster convergence and improved learning efficiency in large state spaces which are inherent in real-world blockchain instances. The non-linearity of the objective function is addressed by incorporating two Double DQNs (DDQNs), one for adversary and one for honest network, which work together to effectively optimize the non-linear objective function. The proposed model is evaluated by constructing a Bitcoin-like Proof-of-Work blockchain simulator which takes into account various real-world blockchain parameters such as stale block rates, propagation delays, and eclipse attacks. Our simulations indicate that the proposed model achieves optimal gains while enhancing the robustness and convergence of the algorithm in large state spaces and dynamically adjusting the mining policy as the blockchain environment evolves.

Keywords: Blockchain, Bitcoin, Selfish mining, Deep reinforcement learning