Thematic Track: Abstract number:

(for Office use only)

**TYPE THE TITLE OF THE ABSTRACT**

*Leave this space blank to facilitate blind review, once the abstract is accepted this space can be used to include the authors of the abstract*

*Keep sufficient space for author/s and affiliation/s of your abstract*

**Place your abstract here**

The Proof of Work (PoW) algorithm is the consensus mechanism used in blockchain networks, including Bitcoin, to achieve agreement on the state of the blockchain and to validate transactions. It requires participants, known as miners, to solve a complex and computing intensive cryptographic puzzle known as the PoW problem to create new blocks and add them to the blockchain. All the miners compete among themselves to solve the PoW puzzle and the miner to successfully solve the puzzle is granted the right to add the new block to the blockchain while also earning a reward for their efforts. Honest mining is the ethical and protocol-compliant approach to cryptocurrency mining which involves immediate broadcasting of discovered blocks across the network. This approach assumes that participating honestly in the network, by immediately sharing valid blocks, will yield the most profitable outcomes for miners in the long term. But it was proved later to gain more mining rewards by selfish mining by withholding the newly solved blocks without publishing them to the network. However, recent research shows that selfish mining attacks may not be optimal and proposed different versions of selfish mining allowing the attacker to earn potentially higher rewards. Later, Sapirshtein et al. formulated the mining problem as a Markov Decision Process (MDP) which can be solved to derive an optimal mining strategy [9]. However this approach required the knowledge of certain blockchain parameters such as the adversary hashrate and communication capability which is difficult to determine in real blockchain environments as they may change over time. Additionally, this model fails to account for various blockchain scenarios with differing stale block rates, eclipse attacks and network delays. Nevertheless, Gervais et al. [\*] proposed a new mining MDP model with a large state space capturing more detailed PoW blockchain instances. In a more recent study, Wang et al. [13] proposed a model free approach adopting the MDP model of [9] and employing Reinforcement Learning to achieve an optimal mining strategy. However, in their study, they used tabular Q-Learning which is not efficient for large state spaces.

**Keywords:** Keyword one, Keyword two, Keyword three, Keyword four, Keyword five

(maximum 5 keywords)

Provide acknowledgement of a grant, if any, for the work presented.

(Do not provide acknowledgement at the initial submission! Provide it when you are submitting the camera-ready version after receiving reviews)