

With the advancement of modern technologies, the security battle between a legitimate system (LS) and an adversary is becoming increasingly sophisticated, involving complex interactions in unknown dynamic environments. Stochastic game (SG), together with multi-agent reinforcement learning (MARL), offers a systematic framework for the study of information warfare in current and emerging cyber-physical systems. In practical security games, each player usually has only incomplete information about the opponent, which induces information asymmetry. This paper exploits information asymmetry from a new angle, considering how to exploit information unknown to the opponent to the player's advantage. Two new MARL algorithms, termed minimax post-decision state (minimax-PDS) and Win-or-Learn Fast post-decision state (WoLF-PDS), are proposed, which enable the LS to learn and adapt faster in dynamic environments by exploiting its information advantage. The proposed algorithms are provably convergent and rational, respectively. Also, numerical results are presented to show their effectiveness through three important applications.