Introduction

Real-time information such as footprints and agents’ subsequent actions upon receiving the information, e.g., rangers following the footprints to chase the poacher, are neglected in previous work of Green Security Games. Our paper fills the gap. First, we propose a new game model GSG-I which augments GSGs with sequential movement and the vital element of real-time information. Second, we design a novel deep reinforcement learning-based algorithm, DeDOL, to compute a patrolling strategy that adapts to the real-time information against a best-responding attacker.

GSG-I Game Model

Attacker: red circle. He tries to put some attacking tools in the world to maximize the damage, meantime avoid defender.
Defender: blue circle. She tries to catch the attacker and remove the attacking tools as soon as possible.
Sequential Interaction: at each time step, attacker and defender both chooses a direction to move. Attacker also decides whether to put an attacking tool. The game environment decides whether an attacking tool will successfully launch an attack.
Footprints: red/blue arrows. Players leave footprints as they move.
Local observation: Players only observe opponent’s footprints in the current cell.

Approximating best response strategy against a fixed opponent by DQN

DeDOL: Computing Optimal Patrol Strategy by Enhancing PSRO

- Vanilla PSRO
- Domain-specific initial strategy
  Attacker: parameterized heuristic
  \[ p_0 (a_t \mid s_t) = \frac{\exp(\langle o_s \rangle + \langle o_2 \rangle)}{\sum_{a_t} \exp(\langle o_s \rangle + \langle o_2 \rangle)} \]
  Defender: random sweeping. She moves along the boundary to find outgoing attacker footprints to follow. If multiple footprints, she randomly chooses one direction.

- Local Mode
  Ease training
  Provide good building block strategy
  Global mode with 4 entry points
  Local mode with top-left entry point

- DeDOL workflow
  DeDOL (Local, Entry Point 1, Go)
  DeDOL (Local, Entry Point 2, Go)
  DeDOL (Global, Go)
  Combine all subgames to G_f
  DeDOL (Global, Go)
  DeDOL (Local, Go)

- Experimental Results
  - Best Response Approximation
  - Expected Utility against a best-responding poacher

- DeDOL Overview
  DeDOL builds upon the double oracle (DO) framework and the policy-space response oracle (PSRO). It starts with a restricted game and iteratively adds best response strategies to it, which is approximated by training DQN. Exploring the game structure, DeDOL uses domain-specific heuristic strategies as initial strategies in PSRO, and constructs several local modes for efficient and parallelized training.

- Acknowledgment
  The Azure resources are provided by Microsoft for Research AI for Earth award program.