MS

GuardAttackPolicy

PEL:

model/MS-unseen-3oppo-ours-a1-l1-W5-K20-20240709060259/pel\_encoder\_iter\_199

BASELINE:

0.75

FINETUNE:

model/MS-6oppo-ours-a1-l1-W5-K20-20240709062744

0.912

GuardEvadePolicy

Base： 0.692

PEL：

model/MS-unseen-3oppo-ours-a1-l1-W5-K20-20240709063516/pel\_encoder\_iter\_199

Trained：

model/MS-6oppo-ours-a1-l1-W5-K20-20240709065011

0.972 0.986

**('RLPolicy', 'opponent\_policy\_models/PPO\_100000.pt', '4')**

**Baseline：-0.098**

**-0.162**

**PEL**

**model/MS-unseen-3oppo-ours-a1-l1-W5-K20-20240710014138/pel\_encoder\_iter\_199**

**model/MS-6oppo-ours-a1-l1-W5-K20-20240710015153/res\_encoder\_iter\_1999**