MACTA: A Multi-agent Reinforcement Learning Approach for Cache Timing Attacks and Detection

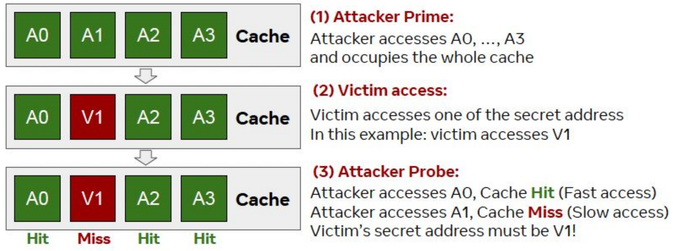
Jiaxun Cui, Xiaomeng Yang*, Mulong Luo*, Geunbae Lee*, Peter Stone, Hsien-Hsin S. Lee, Benjamin Lee, G. Edward Suh, Wenjie Xiong^, Yuandong Tian^

ICLR International Conference On Learning Representations

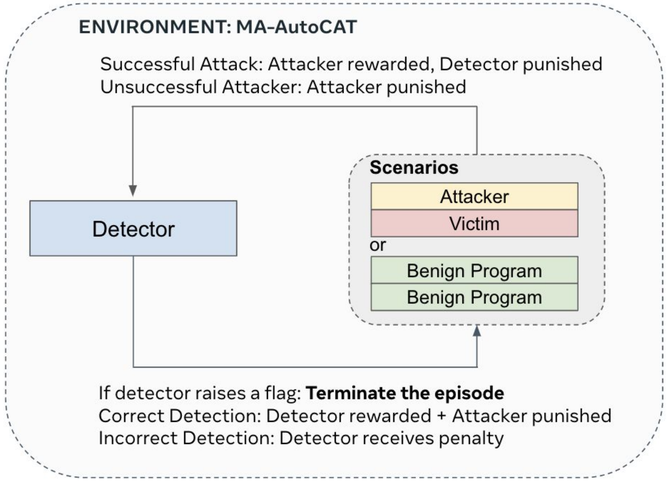
Cache Timing Attack Challenge

- Cache timing attacks forms when attacker and victim share the same cache.
- Attackers can infer the secret victim access of cache by observing its own cache access latencies.

Example Attack



Environment



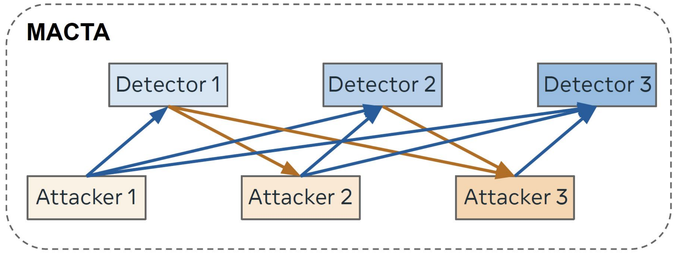
MACTA

MACTA optimizes Attacker policy and Detector policy jointly

1. **Transformer** observation encoder
2. Maintain a policy **pool** for each agent and increase the pool size with policy checkpoints during training
3. Approximate Best Responses to a **uniform mixture** of opponents using (Dual-Clip) Proximal Policy Optimization (PPO)

Generalizability

- MACTA detector generalizes to unseen attackers, with low False Alarm rate
- MACTA attacker mimics benign programs

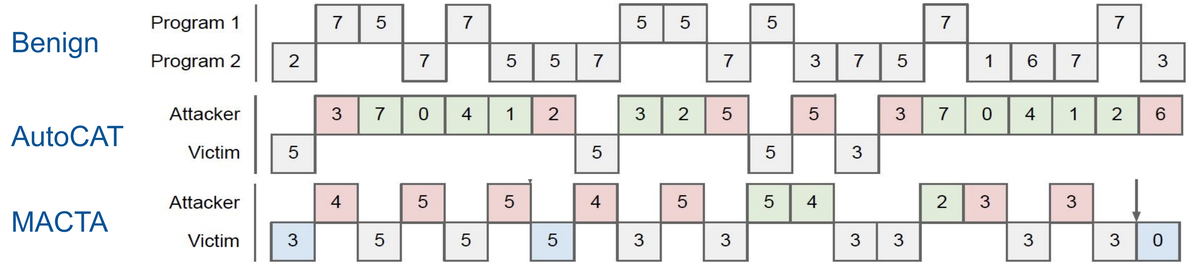


Detection Rate / False Alarm Rate

P → Q Q learns policy against P using PPO

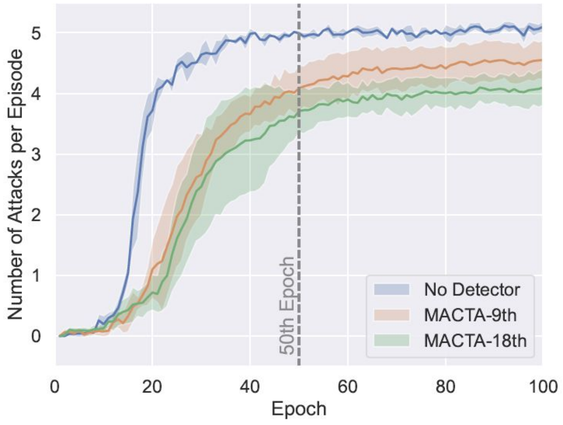
Detectors \ Opponents	Prime+Probe ↑	AutoCAT ↑	IBR-PPO Attacker ↑	MACTA Attacker ↑	Benign ↓
CC-Hunter (thold=0.45)	37.7 ± 0.6	13.7 ± 1.3	12.1± 0.4	16.4 ± 2.3	27.6 ± 0.9
Cyclone (One-Class SVM)	0.0 ± 0.0	55.8 ± 4.3	33.6 ± 12.8	9.0 ± 5.3	19.3 ± 0.9
Cyclone (SVM)	(**99.5** ± 0.1)	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	1.4 ± 0.2
IBR-PPO Detector	0.9 ± 0.7	7.3 ± 20.5	6.4 ± 15.6	8.4 ± 21.9	**0.4** ± 0.5
MACTA Detector	**97.8** ± 0.9	**99.9** ± 0.2	**99.6** ± 0.4	**31.2** ± 18.5	1.1 ± 0.2

Trajectories



Robustness

- MACTA detector reduces information leakage against adaptive attackers



Paper

Code

TEXAS The University of Texas at Austin Meta

VIRGINIA TECH intel Penn University of Pennsylvania

Cornell University Sony AI