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


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

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

Robustness

- MACTA detector reduces information leakage against adaptive attackers

Example Attack

(1) Attacker Prime: Attacker accesses A0, A1, A2 and occupies the whole cache
(2) Victim access: Victim accesses one of the secret address in this example: victim accesses V1
(3) Attacker Probe: Attacker accesses A0, Cache Hit (Fast access) Attacker accesses A1, Cache Miss (Slow access) Victim's secret address must be V1

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.

Environment

ENVIRONMENT: MA-AutoCAT
Successful Attack: Attacker rewarded, Detector punished
Unsuccessful Attack: Attacker punished

Scenarios

- Attacker
- Victim

If detector raises a flag: Terminate the episode
Correct Detection: Detector rewarded + Attacker punished
Incorrect Detection: Detector receives penalty

Trajectories

Benign

Program 1

Program 2

AutoCAT

Attacker

Victim

MACTA

Attacker

Victim

Detection Rate / False Alarm Rate

P ⟷ Q
Q learns policy against P using PPO