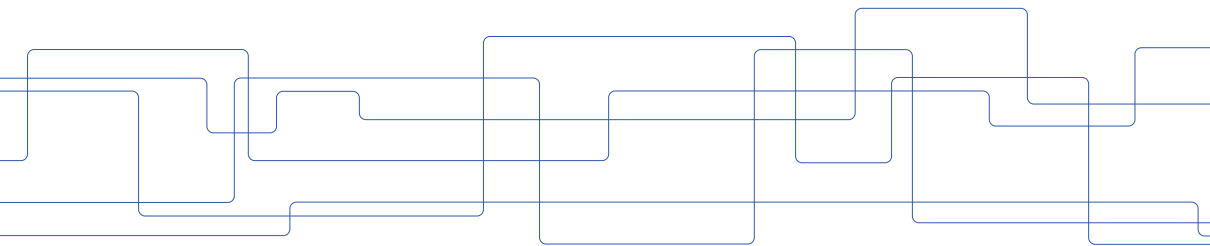
SENTIENCE

Simulation-based reinforcement-learning security operations center

Jakob Nyberg Pontus Johnson Rolf Stadler Teodor Sommestad (FOI)

KTH Royal School of Technology

26-01-2023



Me

- Jakob Nyberg.
- Civilingenjörsexamen from Uppsala University.
  - Embedded systems.
  - Machine learning.

SENTIENCE

- Project started in spring 2021.
- Project within CDIS
  - Centre for Cyber Defence and Information Security
- Funded by MSB.
- Long-term goal is to develop a *semi-autonomous (cyber) security operations center*.

Myndigheten för samhällsskydd och beredskap

Semi-autonomous Security Operations Center

- Use autonomous agents to filter and analyze the current threat situation.
- Produce policies to plan ahead and suggest suitable actions for defense.

Semi-autonomous Security Operations Center

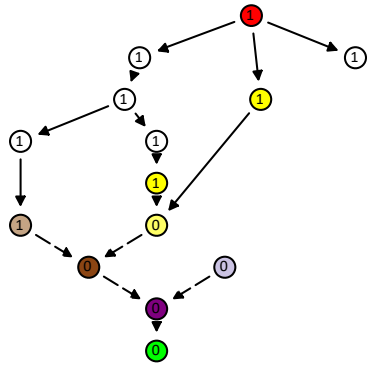
- Use autonomous agents to filter and analyze the current threat situation.
- Produce policies to plan ahead and suggest suitable actions for defense.
- Aid *human* defenders in their decision process.
- The system should be a *resource* to humans, not a replacement.

Approach

- Intrusion response.
- Autonomous agents that learn from data.
- Train in simulation, evaluate in real system.

Approach

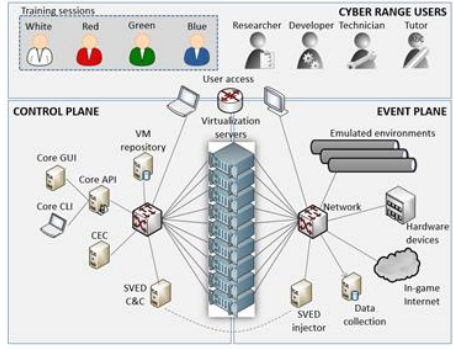
- Intrusion response.
- Autonomous agents that learn from data.
- Train in simulation, evaluate in real system.
- Defense exercise formulated as a Markov game or decision process.
- Defender and attacker agents.
- Find policies using *reinforcement learning*.



Beyond the Simulation

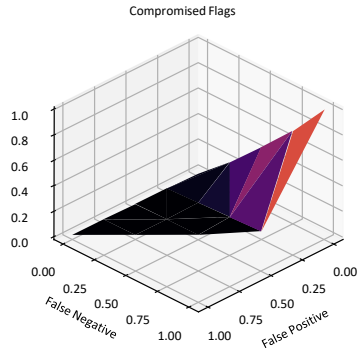
CRATE

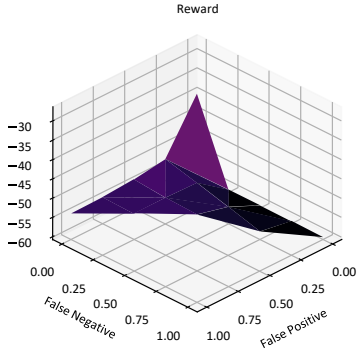
- Cyber Range And Training Environment developed and operated by FOI.
- Used for cyber defense exercises.
- Can emulate different network configurations.
- Red-team emulation using LORE.



Preliminary Results

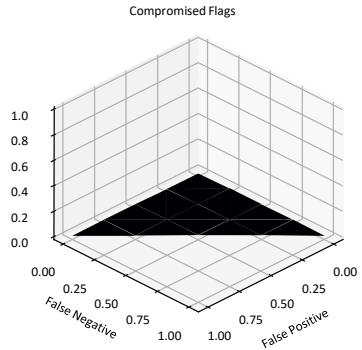
False alerts and missed alerts — Heuristic policy

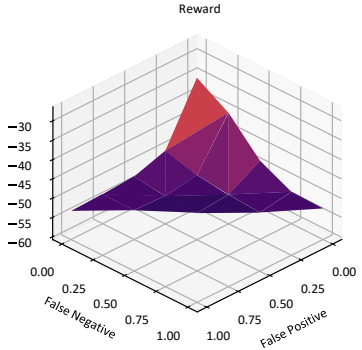




Preliminary Results

False alerts and missed alerts — RL policy





https://www.kth.se/profile/jaknyb

Thank you for listening.
Email queries to **jaknyb@kth.se**