# Learning Movement Policies in Bayesian Stackelberg Markov Games for Adaptive Moving Target Defense

*Sailik Sengupta and Subbarao Kambhampati*

Arizona State University

## Moving Target Defense

Dynamically move between deployment configurations to make it hard for an attacker to succeed by attacking a particular sensor.

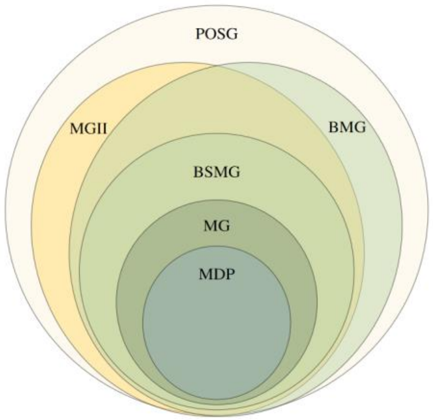
Often modeled as a leader-follower game between attacker and defender.

sailiks@asu.edu

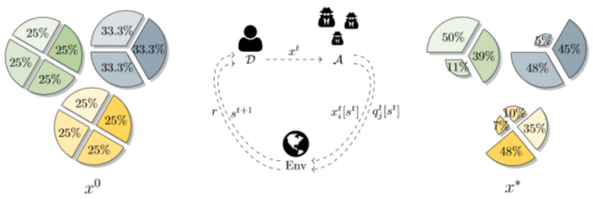
## Challenges and Modeling

Difficult to obtain parameters of the game up-front but interaction with a system may be possible.

Uncertainty over attacker types (imperfect information)
*script kiddie vs. nation state*



## Algorithms



Bayesian Strong Stackelberg Q-Learning:

Samples an adversary type, an actions from a policy and plays it.

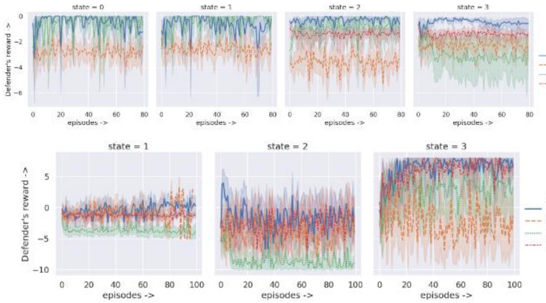
Resulting rewards updates Q-values for the defender and the adversary type.

Updates policy based on SSE of the Q-value Bayesian game.

## Results

Guaranteed to converge to the SSE of the BSMG!

Yields higher rewards for the defender compared to existing strategy inference and learning based approaches.
(Open AI style gym MARL envs.)



URS [Zhang *et. al.* 2014] BSG-MTD [Sengupta et. al. 2017]
[B-]EXP-Q [Klima *et. al.* 2016] Nash-Q [Hu *et. al.* 1998]