Learning Adversary Behaviour in a Stackelberg Security Game

Manasa Jagadeesh, Mitali Seth

Introduction

In real life scenarios, security has become an increasingly important aspect.
Security scenarios can be modeled as a strategic interaction between an attacker and a defender.
In this project, we aim to learn the adversary's utilities so as to better predict their behaviour.

Security Games

- A security game is a model for resource allocation in adversarial environments. There are two players, one **attacker** and one **defender**.
- The defender seeks to protect her targets with a **limited set of resources** available, and the attacker aims to attack the most favourable target.

Stackelberg Security Game

- The **Stackelberg model** can be used to model a security game – the leader is the defender, and the follower is the attacker.
- The defender picks her mixed strategy of which targets to protect, and the defender observes this, and then picks her strategy.

Defender \ Attacker	c	d
a	3, 1	5, 0
b	2, 0	4, 2

Payoff matrix for a SSG

Bounded Rationality

Rationality of the decision maker is affected by some factors:

- Finite amount of time
- Information available
- Limited cognitive abilities

Quantal Response

To tackle bounded rationality, a **stochastic distribution** of the adversary's response is used- the higher the expected utility of a target, the more likely the adversary will attack it.
QR model [1] introduces a parameter called λ

- If λ=0, the adversary is completely irrational
- If λ-> ∞, the predicted response converges to optimal.

Now, the probability that a given target is chosen is,

$$p(t) = \frac{e^{\lambda * x_t}}{\Sigma_t e^{\lambda * x_t}}$$

Subjective Utility

A **Subjective Utility function** is introduced in [2], which says:
The utility of a target for the attacker is dependent on a linear combination of the coverage probability, the reward and the penalty to the attacker on attacking that target.

$$U_t^a = w_1 X_t + w_2 R_t^a + w_3 P_t^a$$

Parameters

X_t : Coverage probability

R_t^a : Reward for the adversary if he attacks the target when not covered by the defender

P_t^a : Penalty for the adversary if he attacks the target when covered by the defender

Learning parameters

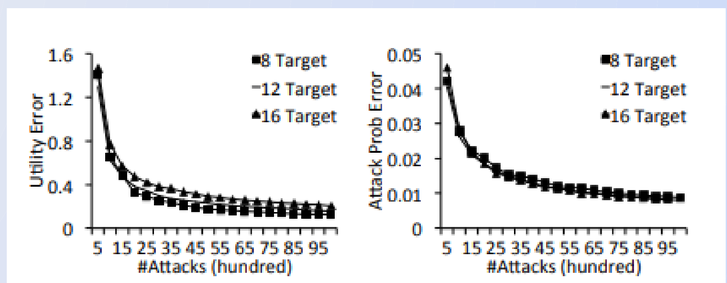
When the SU function is used in the QR model instead of the coverage probability directly, we get the **SUQR model** [2]

- The **Maximum Likelihood Estimator** (MLE) is used to learn the model parameters (w_1, w_2, w_3).
- The log-likelihood function was solved by a convex optimization solver in [2].
- The parameters are said to be **uniformly learnt.**
- This is called **Closed Form Estimation.**

Problem definition

The goal is to learn the utility functions $u_t(.)$ of the attacker for all targets.

Given three defender strategies that are sufficiently different from each other, and a polynomial number of adversary responses to these strategies, the model parameters can be learnt with high accuracy. [3]
This result holds for **linear utility functions.**



(a) Error in $u_t(\cdot)$ (b) Error in predicted attack distribution

Future Work

- **Synthetic data** to be generated by modeling the game.
- **Learning of adversary parameters** using the methods described above.
- Testing **statistical validity** of the results so obtained.

Examples of Security Games

The US Coast Guard which is tasked with protecting ports and ferries has only a limited set of resources, and hence needs to deploy them based on a strategy. Terrorist attacks are one-off instances.

The size of wildlife reserves, and a general lack of funds make the curbing of poaching a difficult issue. As poaching happens repeatedly, past data can be used to learn adversary strategies.

References

1. Richard D McKelvey et. al., Quantal Response Equilibria for Normal Form Games, 1996
2. Thanh H. Nguyen et. al., Analyzing the Effectiveness of Adversary Modeling in Security Games, AAAI 2013
3. Nika Haghtalab et. al., Three Strategies to Success: Learning Adversary Models in Security Games, IJCAI-2016.

Images source: Google Images