Learning Optimal Commitment to Overcome Insecurity

Avrim Blum, Nika Haghtalab, and Ariel Procaccia

Carnegie Mellon University

Security Game Model

A *Stackelberg* game between a defender and an attacker.

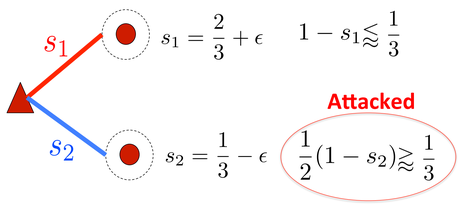
- Defender commits to a rand. deployment of resources.
- Attacker best-responds:

  → Attacks the target with highest expected payoff.

Goal: Find the optimal rand. deployment to commit to.

All target utilities

	Attacker / Defender	Covered	Not covered
1		0 / 0	1 / 0
2		0 / 1	0.5 / 0



Resources	Targets	Optimal	Attacker Util.	Defender Util.
s_1	s_1	$s_1 = \frac{2}{3} + \epsilon$	$1 - s_1 \lessapprox \frac{1}{3}$	$\lessapprox \frac{1}{3}$
s_2	s_2	$s_2 = \frac{1}{3} - \epsilon$	Attacked: $\frac{1}{2}(1 - s_2) \gtrapprox \frac{1}{3}$	

In general: # deployments could be exponential.

Eg. n resources with $2n$ targets, where each resource can defend either of targets i and $2i$ → 2^n deployments

Known vs. Unknown Attacker Utility

Standard assumption:

Attacker's utilities are known.

→ Solve LPs to find the optimal deployment.

What if it is not known?

Use *learning* to find a near optimal deployment.

→ Good for routine security tasks:

Our contribution wrt. Previous Results

Letchford et al.[2] addressed this question in Stackelberg games:

1. Assumed best-response regions are large enough to hit with random sampling.
2. # queries polynomial in # deployments.

→ In security games this is exp(# targets).

Our work

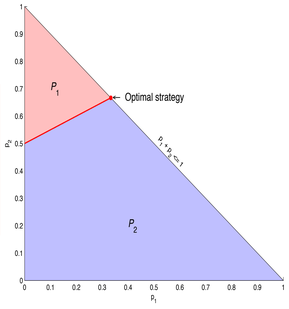
- **Relaxes the 1st assumption.**
- **Uses # queries polynomial in # targets.**

Convexity is Preserved in Low DImension

$$\max_{\vec{p}} \quad U_d(i, p_i) \qquad \text{(Coverage prob. of } i\text{; Defender's utility)}$$

$$\text{s.t.} \quad \text{Attacker best-responds by } i \text{ and } \vec{p} \text{ is implementable}$$

Convex Region



Theorem: In our projection of the strategy space (exponential in #targets) to the coverage probability space (linear # targets), the best-response regions stay convex.

Optimization Using Best-Response Queries

If we know which target is induced by the optimal strategy,

- Use an algorithm by Kalai & Vempala[1] to solve the above LP with that target using membership queries.

  → Attacker's best response is the membership oracle.

- Need well-centered initial point in that region.
  - Regions are not large enough to hit by samples.
  - What to do?

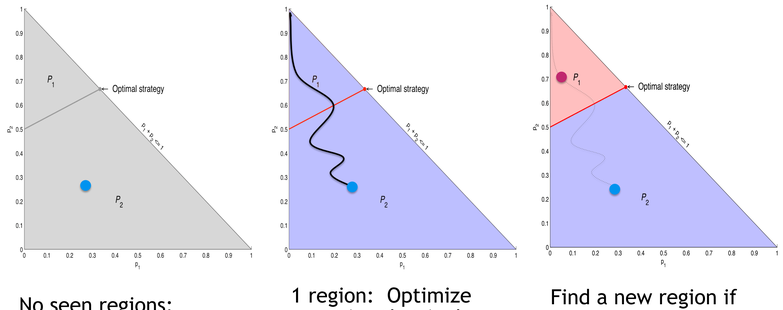
Recursively Finding New Regions

Pretend the responses seen so far are the only possible regions. Find the optimal strategy for these regions by the above argument.

Intuitively, if at all possible to induce a new response, this strategy should induce it.

Theorem: If the optimal strategy that only considers a subset of targets does not induce a new response, then this strategy is optimal over all responses.

Example



No seen regions: Pick at random

1 region: Optimize assuming that is the only region.

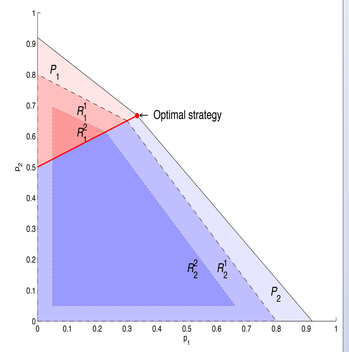
Find a new region if the current regions are not enough.

Challenge: Well-Centered Initial Points

How to ensure a well-centered initial point?

Layer the regions:

- Create margin between points from one layer to the next.



Results

We can (ϵ, δ)-learn the optimal strategy using only $\text{poly}(n \log(\frac{1}{\epsilon\delta}))$ best-response queries.

Acknowledgements

This material is based upon work supported by the NSF under grants CCF-1116892, CCF-1101215, CCF-1215883, and IIS-1350598.

References

1. A. Kalai and S. Vempala. Simulated annealing for convex optimization. Mathematics of Operations Research, 31(2):253-266, 2006
2. J. Letchford, V. Conitzer, and K. Munagala. Learning and approximating the optimal strategy to commit to. In Proc. of SAGT, pages 250-262, 2009.