Lecture Note 16: Stackelberg equilibrium and Security Games

Christian Kroer*

April 17, 2022

1 Introduction

In this lecture we introduce *Stackelberg equilibrium*. Stackelberg equilibrium is an equilibrium notion for two-player general-sum games where one player is a *leader* and the other player is a *follower* (it can also be generalized to multiple leaders and/or followers). This model is appropriate for example when modeling competing firms and first-mover advantage, or as we will see in this lecture, security settings centered around asset protection.

2 Stackelberg Equilibrium

We will consider a two-player normal-form game where there is a leader ℓ and a follower f. The leader has a finite set of actions A_ℓ and the follower has a finite set of actions A_f. We let Δ^ℓ, Δ^f denote the set of probability distributions over the leader and follower actions. We will consider a general-sum game with utilities $u_i(a_\ell, a_f)$ for $i \in \{\ell, f\}$. We abuse notation slightly and let

$$u_i(x, y) = \mathbb{E}_{a_\ell \sim x, a_f \sim y} \left[u_i(a_\ell, a_f) \right],$$

where $x \in \Delta^\ell, y \in \Delta^f$ are probability distributions over A_ℓ and A_f respectively. In general we assume that the leader is able to commit to a strategy $x \in X$, and given such an x, the follower chooses their strategy from the best-response set

$$BR(x) = \text{argmax}_{y \in \Delta^f} u_f(x, y).$$

The goal of the leader is to choose a strategy a strategy x maximizing their utility subject to the follower best responding. Formally, they wish to solve

$$\max_{x \in \Delta^\ell} u_\ell(x, y) \ s.t. \ y \in BR(x). \tag{1}$$

However, this optimization problem has a problem currently. Can you see what it is?

The issue is that $BR(x)$ may be set valued, and $u_\ell(x, y)$ generally would differ defending on which $y \in BR(x)$ is chosen. In that case we need a rule for how to choose among the set of best responses. In a *strong Stackelberg equilibrium* (SSE) we assume that the follower breaks ties in favor of the leader. In that case the optimization problem is

$$\max_{x \in \Delta^\ell, y \in BR(x)} u_\ell(x, y). \tag{2}$$

*Department of Industrial Engineering and Operations Research, Columbia University. Email: christian.kroer@columbia.edu.

SSE is, in a sense, the most optimistic variant. Conversely, we may consider the most pessimistic assumption, that ties are broken adversarially. This yields the *weak Stackelberg equilibrium* (WSE)

$$\max_{x \in \Delta^\ell} \min_{y \in BR(x)} u_\ell(x, y). \tag{3}$$

In practice SSE has been by far the most popular. One major advantage of SSE is that it is always guaranteed to exist, whereas WSE is not.

A first question we might ask ourselves is whether it always helps or hurts to be able to first commit to a strategy, as compared to playing a Nash equilibrium.

First, let us consider the zero-sum case. If we are in a zero-sum game, then we already saw from von Neumann's minimax theorem that we can represent the Nash equilibrium problem as

$$\min_{x \in \Delta^\ell} \max_{y \in \Delta^f} \langle x, Ay \rangle = \max_{y \in \Delta^f} \min_{x \in \Delta^\ell} \langle x, Ay \rangle.$$

It follows that Nash equilibrium and Stackelberg equilibrium are equivalent in this setting.

Second, consider the case where we restrict the leader to only committing to pure actions $a \in A_\ell$, then committing to a strategy first may hurt the leader (consider rock-paper-scissors). On the other hand, if we allow commitment to any $x \in \Delta^\ell$, then it turns out that committing to a strategy only helps.

Theorem 1. *In a general-sum game, the leader achieves weakly more utility in SSE than in any Nash equilibrium.*

Proof. Consider the Nash equilibrium (x, y) that yields the highest utility for the leader. Since the follower breaks ties in favor of the leader, we get that if the leader commits to x then the follower can at worst pick y from $BR(x)$. If they don't pick y, then they must pick something that yields even better utility for the leader. □

Similarly, it can be shown that the WSE solution is at least as good as *some* Nash equilibrium payoff for the leader (see Von Stengel and Zamir [9] for a proof). Thus, if we consider the range of payoffs $[L, H]$ from the lowest to highest in Stackelberg equilibrium, then that range lies above the range that we would get for Nash equilibrium.

A classic example of the difference between Nash equilibrium and Stackelberg equilibrium is in the context of *inspection games*. In an inspection game, an inspector chooses whether to inspect or not, and the inspectee chooses whether to cheat or not. An example game is shown below

	cheat	no cheat
inspect	-6, -9	-1,0
no inspection	-10, 1	0, 0

The goal of the inspector is to deter cheating, and inspecting incurs a cost of -1. When cheating occurs the inspector incurs a heavy negative cost, whether detected or not (so the goal is *not* to catch cheaters, but rather to deter cheating). The inspectee gains utility from cheating undetected (-10, 1), but incurs a heavy fine if they cheat and are inspected (-6,-9).

There is a single unique Nash equilibrium in this game, where the inspector inspects with probability $\frac{1}{10}$, and the inspectee cheats with probability $\frac{1}{5}$. This yields expected utilities of (-2, 0) for the two players.

Now consider the same game, but where we allow the inspector to be the leader in a Stackelberg game. Any strategy that inspects with probability at least $\frac{1}{10}$ will make not cheating a best response for the follower. The SSE of the game is for the inspector to inspect with probability $\frac{1}{10}$ and the

inspectee to not cheat. This yields expected utilities $(-\frac{1}{10}, 0)$, which is much better for the inspector. Note furthermore that if we consider the WSE solution concept, then the inspector must inspect with probability *strictly* greater than $\frac{1}{10}$ in order to make not cheating the only best response. But this means that a WSE does not exist, since for every leader strategy that inspects with probability $p > \frac{1}{10}$, the leader can improve their utility by inspecting with any probability in the open interval $(\frac{1}{10}, p)$.

In the normal-form game setup given above, an SSE can be computed in polynomial time. In particular, say that we wanted to maximize leader utility while getting the follower to commit to a particular action $a_f \in A_f$. We may solve this problem using the following LP:

$$\max_{x \in \Delta^\ell} \sum_{a \in A_\ell} x_a u_\ell(a, a_f)$$
$$s.t. \sum_{a \in A_\ell} x_a u_f(a, a_f) \geq \sum_{a \in A_\ell} x_a u_f(a, a'_f), \ \forall a'_f \in A_f$$

Now, in order to find the optimal strategy to commit to, we may iterate over all $a_f \in A_f$, solve the LP for each, and pick the optimal solution x^* associated to the LP with the highest value.

Once we have the optimal strategy x^*, we may find the associated follower strategy simply by picking the pure strategy a_f for which x^* was the LP solution. Generally, it is easy to see that it is always enough to consider only pure strategies when choosing the follower strategy in an SSE (why?). The same holds true for WSE.

This LP-based algorithm also proves that an SSE is always guaranteed to exist.

3 Security Games

In the security games model (SGM) a defender (the leader) is interested in protecting a set of targets using limited resource, while an attacker (the follower) is able to observe the strategy of the leader, and best respond to it. A classical example would be that of protecting an airport: say we have 5 vulnerable locations at the airport, but only 2 patrol units. How can we schedule the patrols so as to provide maximum coverage across the 5 vulnerable locations, while taking into account the fact that an attacker would prefer certain locations over others?

The basic security games model has a set of T targets (note that we could have a single target appear twice in T, representing multiple time steps). The defender controls a set of resources R that can be assigned to a *schedule* from a set $S \subseteq 2^T$ of possible schedules. A schedule is a subset of targets that are simultaneously covered if a resource is assigned that given schedule (for example in the airport example, a resource would be a patrol, and schedules would be the set of feasible patrols across targets). We say that a target is "covered" if the defender assigns a resource to a schedule that covers it. The action space for the attacker consists of choosing which single target to attack. In the basic SSG model, the utility function of both the defender and attacker depends only on which target is attacked, and whether it is covered or not. Formally, we say that the defender receives utility $u_d^c(t)$ if target t is attacked and covered, and utility $u_d^u(t)$ if target t is attacked an not covered. Similarly, the attacker gains utility $u_a^c(t)$ if target t is attacked and covered, and $u_a^u(t)$ if target t is attacked and not covered. If the resources R are not homogenous then there may be an *assignment function* $A : R \to S$ denoting the set of schedules s that resource r can be assigned.

For security games we will restrict our attention to SSE. Given a strategy x for the defender, we get a deployment of resources to targets for the defender, with an induced probability distribution $p_c(t|x)$ of whether each target is covered. A strategy for the attacker simply specifies a single target

t to attack. Thus for a strategy pair x, t the expected utility for the defender is $p_c(t|x)u_d^c(t) + (1 - p_c(t|x)u_d^u(t))$, with attacker utility defined analogously.

3.1 Algorithms for Security Games

So now that we have a game model for security games, can we just apply our LP result on computing SSE in order to get an SSE for security games? Not quite: in order to convert the SGM into a standard normal-form game we get a combinatorial blow-up: consider that a pure strategy would be a deployment of resources to targets. But now let's say that we just have d patrols as our resources and k targets, and a simple model where each patrol can cover exactly one target. In that case we have $\binom{d}{k}$ pure strategies for the leader. A similar blow-up happens for other natural setups such as when each resource can cover two targets (e.g. air marshals that protect an outgoing and then ingoing flight as their daily routine).

In the special case where each resource covers exactly one target (equivalently, schedules have size 1) there is an LP-based approach that can still construct an SSE in polynomial time. This LP still allows heterogeneous resources; below we let $A(r)$ be the set of targets that resource r is allowed to cover. The key idea in the LP is to use the marginal coverage probability $p_c(t|x)$ as our decision variable. We will have an LP where the variable c_t is the coverage probability on target t, and the variable $c_{r,t}$ is the probability that resource r provides coverage for $t \in A(r)$. The goal is to maximize the defender utility subject to making some target t^* a best response for the attacker. We can then solve for each $t^* \in T$ as before, and pick the best. In this LP, we let $u_a(t|c) = c_t u_a^c(t) + (1 - c_t)u_a^u(t)$, with $u_d(t|c)$ defined analogously.

$$\begin{aligned} \max_{c \geq 0} \quad & u_d(t^*|c) \\ s.t. \quad & c_t = \sum_{r \in R \text{ s.t. } t \in A(r)} c_{r,t} && \leq 1, && \forall t \in T \\ & \sum_{t \in A(r)} c_{r,t} && \leq 1, && \forall r \in R \\ & u_a(t|c) && \leq u_a(t^*|c), && \forall t \in T. \end{aligned} \tag{4}$$

This LP is polynomial in size, even though the set of pure strategies is exponential in size. It is however not immediately obvious whether the given coverage probabilities are implementable. It turns out that they are, and this can be shown via the famous Birkhoff-von Neumann theorem. Before stating the theorem, we need the definition of a *doubly substochastic matrix*, which is a matrix $M \in \mathbb{R}^{m \times n}$ such that all entries are nonnegative, each row sums to at most 1, and each column sums to at most 1.

Theorem 2 (Birkhoff-von Neumann theorem)**.** *If M is doubly substochastic, then there exist matrices $M_1, M_2, \ldots, M_q$, and weights $w_1, w_2, \ldots, w_q \in (0, 1]$, such that:*

1. $\sum_k w_k = 1$
2. $\sum_k w_k M_k = M$
3. *For all k, M_k is doubly substochastic, and all entries are in $\{0, 1\}$*

Informally, the theorem states that if we have a doubly substochastic matrix, then it is possible to express it as a convex combination of "pure" or $\{0, 1\}$ doubly substochastic matrices.

The coverage probabilities $c_{r,t}$ from our LP can be viewed as a matrix with rows corresponding to resources and columns corresponding to targets. By the constraints in our LP, that matrix is

doubly substochastic. It follows from the Birkhoff-von Neumann theorem that there exists pure-strategy matrices M_k (they are pure strategies by the 3rd condition of the theorem) such that their convex combination under the weight vector w adds up the correct coverage probabilities (by the 2nd condition of the theorem). By the first condition, the vector w defines a mixed strategy.

One final worry is that we don't know how large q will be in our application of the Birkhoff-von Neumann theorem. Luckily, it turns out one can show (in general), that q is $O((m+n)^2)$, and the corresponding M_k, w_k can be found in $O((m+n)^{4.5})$ time using the Dulmage-Halperin algorithm.

Unfortunately, in the more general case where schedules may cover more than one target the trick using marginal coverage probabilities turns out to fail. In that case, computing an SSE turns out to be NP-hard.

Still, it turns out we can formulate this problem as the following *mixed-integer program* (MIP):

$$\begin{array}{lll} \max\limits_{c \geq 0} & u_d(t^*|c) & \\ s.t. & c_t = \sum\limits_{r \in R \text{ s.t. } t \in A(r)} c_{r,t} & \leq 1, \quad \forall t \in T \\ & \sum\limits_{t \in A(r)} c_{r,t} & \leq 1, \quad \forall r \in R \\ & u_a(t|c) & \leq u_a(t^*|c), \quad \forall t \in T. \end{array} \tag{5}$$

4 Criticisms of Security Games

In security games we make a number of assumptions that can easily be critiqued: first, we assume that the attacker perfectly observes the defender strategy. Secondly, the defender knows exactly what the utility function of the attacker is (and SSE relies heavily on this). Thirdly, we assume that the attacker is perfectly rational. There are ways to address these assumptions. For example, a lot of work has gone into modeling adversaries in a way that is robust either to misspecification of the utility functions or followers not being perfectly rational.

5 Bayesian Games

One way to deal with uncertainty around utility is to assume that each player has a parameterized utility function $u_i(\cdot, \cdot|\theta_i)$, where $\theta_i \in \Theta_i$ is the *type* of player i. In Bayesian games, we assume each player draws their type from a pair of known distributions over Θ_ℓ, Θ_f. The player observes their own type before choosing an action, but not the type of the follower.

It turns out that in the special case where the follower has a single type θ_f and the leader has a probability mass $p_\ell(\theta)$ over a finite set Θ_ℓ, the LP approach for normal-form games can easily be extended to this setting and yields an optimal strategy for the leader. However, the more interesting case where the follower has multiple types is unfortunately NP-hard.

6 Historical Notes

The Stackelberg game model was introduced by Von Stackelberg [8] in order to analyze competing firms and first-mover advantage.

The foundations for the use of Stackelberg equilibrium in security games were laid by Von Stengel and Zamir [9] who showed that it always helps to commit to a strategy, as long as mixed strategies are allowed, and Conitzer and Sandholm [2] who gave efficient algorithms and complexity results around computing Stackelberg equilibrium for various game models.

In the context of security, Stackelberg equilibrium was first applied to airport security at Los Angeles International Airport [6], and has since been applied to problems such as preventing poaching and illegal fishing [3] and airport security screening [1]. An overview of deployed systems and new directions can be found in Sinha et al. [7].

The approach based on representing strategies in terms of the marginal probability of coverage was introduced by Kiekintveld et al. [4], and the results on polynomial-time algorithms and computational complexity in this model were given by Korzhyk et al. [5].

References

[1] Matthew Brown, Arunesh Sinha, Aaron Schlenker, and Milind Tambe. One size does not fit all: A game-theoretic approach for dynamically and effectively screening for threats. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[2] Vincent Conitzer and Tuomas Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 82–90, 2006.

[3] Fei Fang, Peter Stone, and Milind Tambe. When security games go green: Designing defender strategies to prevent poaching and illegal fishing. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[4] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordóñez, and Milind Tambe. Computing optimal randomized resource allocations for massive security games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696, 2009.

[5] Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr. Complexity of computing optimal stackelberg strategies in security resource allocation games. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.

[6] James Pita, Manish Jain, Janusz Marecki, Fernando Ordóñez, Christopher Portway, Milind Tambe, Craig Western, Praveen Paruchuri, and Sarit Kraus. Deployed armor protection: the application of a game theoretic model for security at the los angeles international airport. 2008.

[7] Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, and Milind Tambe. Stackelberg security games: Looking beyond a decade of success. IJCAI, 2018.

[8] Heinrich Von Stackelberg. *Marktform und gleichgewicht*. J. springer, 1934.

[9] Bernhard Von Stengel and Shmuel Zamir. Leadership games with convex strategy sets. *Games and Economic Behavior*, 69(2):446–457, 2010.