Computing Equilibria with Partial Commitment*

Vincent Conitzer

Duke University, Durham, NC, USA

Abstract. In security games, the solution concept commonly used is that of a Stackelberg equilibrium where the defender gets to commit to a mixed strategy. The motivation for this is that the attacker can repeatedly observe the defender's actions and learn her distribution over actions, before acting himself. If the actions were not observable, Nash (or perhaps correlated) equilibrium would arguably be a more natural solution concept. But what if some, but not all, aspects of the defender's actions are observable? In this paper, we introduce solution concepts corresponding to this case, both with and without correlation. We study their basic properties, whether these solutions can be efficiently computed, and the impact of additional observability on the utility obtained.

1 Introduction

Algorithms for computing game-theoretic solutions have long been of interest, but were for a long time not deployed in real-world applications (at least if we do not count, e.g., computer poker programs—for an overview of those, see Sandholm [21]—as real-world applications). This changed in 2007 with a series of deployed applications coming out of Milind Tambe's TEAMCORE research group at the University of Southern California. The games in question are what are now called security games, where a defender has to allocate limited resources to defend certain targets or patrol a certain area, and an attacker chooses a target to attack. The deployed applications include airport protection [20], assigning Federal Air Marshals to flights [22], patrolling in ports [2], fare inspection in transit systems [25], and patrolling to prevent wildlife poaching [11].

While most of the literature on computing game-theoretic solutions has focused on the computation of Nash equilibria—including the breakthrough result that even computing a single Nash equilibrium is PPAD-complete [10, 6]—in the security games applications the focus is instead on computing an optimal mixed strategy to commit to [8]. In this model, one player (in security games, the defender) chooses a mixed strategy, and the other (the attacker) observes this mixed strategy and best-responds to it. This sometimes helps, and never hurts, the former player [23]. Intriguingly, in two-player normal-form games, such a strategy can be computed in polynomial time via linear programming [8, 23].

* I dedicate this paper to my sister Jessica, her fiancé Jeremy, and their upcoming full commitment. I wish them a lifetime of happiness.

Another benefit of this model is that it sidesteps issues of equilibrium selection that the approach of computing (say) a Nash equilibrium might face.

Such technical conveniences aside, the standard motivation for assuming that the defender in security games can commit to a mixed strategy is as follows. The defender has to choose a course of action every day. The attacker, on the other hand, does not, and can observe the defender's actions over a period of time. Thus, the defender can establish a reputation for playing any particular mixed strategy. This can be beneficial for the defender: whereas in a simultaneous-move model (say, using Nash equilibrium as the solution concept), she can play only best responses to the attacker's strategy, in the commitment model she can commit to play something that is not a best response, which may incentivize the attacker to play something that is better for the defender. Of course, for this argument to work, it is crucial that the attacker observes over time which actions the defender takes before taking any action himself. Previous work has questioned this and considered models where there is uncertainty about whether the attacker observes the defender's actions at all [15, 14], as well as models where the attacker only gets a limited number of observations [19, 1].

In this paper, we consider a different setting where some defender actions are (externally) indistinguishable from each other. This captures, for example, the case where there are both observable and unobservable security measures, as is often the case. Here, two courses of action are indistinguishable if and only if they differ only in the unobservable component. It also captures the case where a guard can be assigned to a visible location (1), or to one of two invisible locations (2 or 3). In this case, the first action is distinguishable from the latter two, but the latter two are indistinguishable from each other. Indistinguishability is an equivalence relation that partitions the player's strategy space; we call one element of this partition a SIS (subset of indistinguishable strategies). Thus, the defender can establish a reputation for playing a particular distribution over the SISes. However, she cannot establish any reputation for how she plays *within* each SIS, because this is not externally observable. Thus, intuitively, when the defender plays from a particular SIS, she needs to play a strategy that, within that SIS, is a best response; however, if there is another strategy in a *different* SIS that is a better response, that is not a problem, because deviating to that strategy would be observable.

The specific contributions of this paper are as follows. We formalize solution concepts for these settings that generalize both Nash and correlated equilibrium, as well as the basic Stackelberg model with (full) commitment to mixed strategies. Further contributions include illustrative examples of these solutions, basic properties of the concepts, analysis of their computational complexity, and analysis of how the row player (defender)'s utility varies as a function of the amount of commitment power (as measured by observability).

2 Definitions and Basic Properties

We are now ready to define some basic concepts. Throughout, the row player (player 1) is the player with (some) commitment power, in the sense of being able to build a reputation. R denotes the set of rows, C the set of columns, and σ_1 and σ_2 denote mixed strategies over these, respectively.

Definition 1. *A* subset of indistinguishable strategies (SIS) S *is a maximal subset of* R *such that for any two rows* $r_1, r_2 \in S$*, the column player's observation is identical for* r_1 *and* r_2*. Let* $\mathcal{S}$ *denote the set of all SISes, constituting a partition of* R*. Given a mixed strategy* σ_1 *for the row player and a SIS* S*, let* $\sigma_1(S) = \sum_{r \in S} \sigma_1(r)$ *(where* $\sigma_1(r)$ *is the probability* σ_1 *puts on* r*).*

Since our focus is on games in which one player can build up a reputation and the other cannot, we do not consider SISes for the column player. Equivalently, we consider all the column player's strategies to be in the same SIS.

Definition 2. *Two mixed strategies* σ_1, σ_1' *are* indistinguishable *to the column player if for all* $S \in \mathcal{S}$*,* $\sigma_1(S) = \sigma_1'(S)$*.*

Example. Consider the following game:

	A	B
a	7,0	2,1
b	6,1	0,0
c	5,0	0,1
d	4,1	1,0

If the players move simultaneously, then a is a strictly dominant strategy and we obtain (a, B) as the iterated strict dominance solution (and hence the unique Nash equilibrium), with a utility of 2 for the row player. If the row player gets to commit to a mixed strategy, then she could commit to play a and b with probability 1/2 each, inducing the column player to play A,[1] resulting in a utility of 6.5 for the row player. (Even committing to a pure strategy—namely, b—would result in a utility of 6.) Now suppose $\mathcal{S} = \{\{a, b\}, \{c, d\}\}$, i.e., a and b are indistinguishable and so are c and d. In this case, playing a and b with probability 1/2 each (or playing b with probability 1) is indistinguishable from playing a with probability 1. Hence, it is not credible that the row player would ever play b, given that a is a strictly dominant strategy. But can the row player still do better than always playing a (and thereby inducing the column player to play B)?

We will return to this example shortly, but first we need to formalize the idea of a deviation that cannot be detected by the column player.

[1] As is commonly assumed in this model, ties for the column player are broken in the row player's favor; if not, the row player can simply commit to $1/2 - \epsilon$ on a and $1/2 + \epsilon$ on b.

Definition 3. *A profile* (σ_1, σ_2) *has* no undetectable beneficial deviations *if (1) for all* σ_2'*,* $u_2(\sigma_1, \sigma_2') \leq u_2(\sigma_1, \sigma_2)$*, and (2) for all* σ_1' *indistinguishable from* σ_1*,* $u_1(\sigma_1', \sigma_2) \leq u_1(\sigma_1, \sigma_2)$.

The following simple proposition points out that this is equivalent to the column player only putting probability on best responses, and the row player only putting probability on rows that *within their SIS* are best responses.

Proposition 1. *A profile* (σ_1, σ_2) *has no undetectable beneficial deviations if and only if (1) for all* $c, c' \in C$ *with* $\sigma_2(c) > 0$*,* $u_2(\sigma_1, c') \leq u_2(\sigma_1, c)$*, and (2) for all* $S \in \mathcal{S}$*, for all* $r, r' \in S$ *with* $\sigma_1(r) > 0$*,* $u_1(r', \sigma_2) \leq u_1(r, \sigma_2)$.

Example continued. In the game above, consider the profile

$$(((1/2)c, (1/2)d), ((1/2)A, (1/2)B))$$

This profile has no undetectable deviations: (1) the column player is playing a best response, and (2) the only undetectable deviations for the row player do not put any probability on $\{a, b\}$, and c and d are both equally good responses.

Note that a profile that has no undetectable beneficial deviations may still not be stable, in the sense that player 1 may prefer to deviate to a mixed strategy that is in fact distinguishable from σ_1, and build up a reputation for playing that strategy instead. But in a sense, these profiles are *feasible* solutions for the row player: *given* that the row player decides to build up a reputation for the distribution over SISes resulting from σ_1, the profile (σ_1, σ_2) is stable. This is similar to the sense in which in the regular Stackelberg model, any profile consisting of a mixed strategy for the row player and a best response for the column player is feasible: the row player may not have had good reason to commit to that particular mixed strategy, but *given* that she did, the profile is stable. In fact, this just corresponds to the special case of our model where all rows are distinguishable.

Proposition 2. *If* $|\mathcal{S}| = 1$ *(all rows are indistinguishable), then a profile has no undetectable beneficial deviations if and only if it is a Nash equilibrium of the game. If* $|\mathcal{S}| = |R|$ *(all rows are distinguishable), then a profile has no undetectable beneficial deviations if and only if the column player is best-responding.*

We can now define an optimal solution.

Definition 4. *A profile with no undetectable beneficial deviations is a* Stackelberg equilibrium with limited observation (SELO) *if among such profiles it maximizes the row player's utility.*

Example continued. In the game above, consider the profile

$$(((1/2)a, (1/2)d), ((1/2)A, (1/2)B))$$

This profile has no undetectable deviations: A and B are both best responses for the column player, and the row player strictly prefers a to b and is indifferent between c and d. It gives the row player utility 3.5. We now argue that it is in fact a SELO. First, note that a SELO must put at least probability $1/2$ on d: for, if it did not, then, because the row player would never play b, the column player would strictly prefer B, which would result in lower utility for the row player. Second, the column player must play B at least half the time, because otherwise, the row player would strictly prefer c to d—but if the row player only plays a and c, the column player would strictly prefer B. Under these two constraints, the row player would be best off having as much as possible of the remaining probabilities on a and A, and this results in the profile above.

Proposition 3. *If $|\mathcal{S}| = 1$ (all rows are indistinguishable), then a profile is a SELO if and only if it is a Nash equilibrium that maximizes the row player's utility among Nash equilibria. If $|\mathcal{S}| = |R|$ (all rows are distinguishable), then a profile is a SELO if and only if it is a Stackelberg equilibrium (with full observation).*

3 Computational Results

We now consider the complexity of computing a SELO. We immediately obtain:

Corollary 1. *When $|\mathcal{S}| = 1$, computing a SELO is NP-hard (and the maximum utility for the row player in a profile with no undetectable beneficial deviations is inapproximable unless P=NP).*

Proof. By Propositions 2 and 3, these problems are equivalent to maximizing the row player's utility in a Nash equilibrium, which is known to be NP-hard and inapproximable [13, 9].

This still leaves open the question of whether the problem becomes easier if the individual SISes have small size. Unfortunately, the next result shows that the problem remains NP-hard and inapproximable in this case. This motivates extending the model to one that allows correlation, as we will do in Section 4.

Theorem 1. *Computing a SELO remains NP-hard even when $|S| = 2$ for all $S \in \mathcal{S}$ (and in fact it is NP-hard to check whether there exists a profile with no undetectable beneficial deviations that gives the row player positive utility, even when all payoffs are nonnegative).*

Proof. We reduce from the EXACT-COVER-BY-3-SETS problem, in which we are given a set of elements T ($|T| = m$, with m divisible by 3) and subsets $T_j \subseteq T$ that each satisfy $|T_j| = 3$, and are asked whether there exist $m/3$ of these subsets that together cover all of T. For an arbitrary instance of this problem, we construct the following game. For each T_j, we add a SIS consisting of two rows, $\{T_j^+, T_j^-\}$, as well as a column T_j. For each element $t \in T$, we add a column t. The utility functions are as follows.

- $u_1(T_j^+, T_j) = m/3$ for any j

- $u_1(T_j^+, T_{j'}) = 0$ for any j, j' with $j \neq j'$
- $u_1(T_j^-, T_{j'}) = 1$ for any j, j'
- $u_1(r, t) = 0$ for any row r and element t
- $u_2(r, T_j) = m/3 - 1$ for any row r and any j
- $u_2(T_j^+, t) = 0$ for any j and $t \in T_j$
- $u_2(r, t) = m/3$ for any element t and row r that is not some T_j^+ with $t \in T_j$

First suppose the EXACT-COVER-BY-3-SETS instance has a solution. Let the row player play uniformly over the $m/3$ corresponding rows T_j^+, and the column player uniformly over the $m/3$ corresponding columns T_j. The row player's expected utility for any of the rows in her support is 1; deviating to the corresponding T_j^- would still only give her 1. The column player's expected utility is $m/3-1$ for any T_j; because the row player plays an exact cover, deviating to any t gives him expected utility $(m/3)(m/3-1)/(m/3) = m/3-1$. So this profile has no undetectable beneficial deviations (in fact it is a Nash equilibrium) and gives the row player an expected utility of 1.

Now suppose that the game has a SELO in which the row player gets positive utility, which implies that the column player puts total probability $p > 0$ on his T_j columns. It follows that for every $t \in T$, the total probability that the row player puts on rows T_j^+ with $t \in T_j$ is at least $3/m$, or otherwise the column player would strictly prefer playing t to playing any T_j. However, note that the row player can only put positive probability on rows T_j^+ where the corresponding column T_j receives probability at least $3p/m$ (thereby resulting in expected utility at least p for the row player for playing T_j^+), because otherwise the corresponding row T_j^- (which is indistinguishable) would be strictly preferable (resulting in expected utility p). But of course there can be at most $m/3$ such columns T_j, and these T_j must cover all the elements t by what we said before. Hence the EXACT-COVER-BY-3-SETS instance has a solution.

4 Adding Signaling

The notion of correlated equilibrium [4] results from augmenting a game with a trusted mediator that sends correlated signals to the agents. As is well known, without loss of generality, we can assume the signal that an agent receives is simply the action she is to take. This is for the following reason. If a correlated equilibrium relies on an agent randomizing among multiple actions conditional on receiving a particular signal, then we may as well have the mediator do this randomization on behalf of the agent before sending out the signal. It is well known that correlated equilibria can outperform Nash equilibria from all agents' perspectives. For example, consider Shapley's game, which is a version of rock-paper-scissors where choosing the same action as the other counts as a loss.

	A	B	C
a	0,0	1,0	0,1
b	0,1	0,0	1,0
c	1,0	0,1	0,0

Whereas the only Nash equilibrium of this game is for both players to randomize uniformly (resulting in $0, 0$ payoffs $1/3$ of the time), there is a correlated equilibrium that only results in the $1, 0$ and $0, 1$ outcomes, each $1/6$ of the time. That is, if the mediator is set up to draw one of these six entries uniformly at random, and then tell each agent what she is supposed to play (but not what the other is supposed to play), then each agent has an incentive to follow the recommendation: doing so will result in a win half the time, and it is not possible to do better given what the agent knows.

Correlated equilibria are easier to compute than Nash equilibria: given a game in normal form, there is a linear program formulation for computing even optimal correlated equilibria (say, ones that maximize the row player's utility). The linear program presented later in Figure 1 is closely related.

Similar signaling has received attention in the Stackelberg model. One may assume a more powerful leader in this model that can commit not only to taking actions in a particular way, but also to sending signals in a way that is correlated with how she takes actions. (Again, the motivation for using this in real applications might be that over time the leader develops a reputation for sending out signals according to a particular distribution, and playing particular distributions over actions conditional on those signals.) Because the leader can commit to sending signals in a particular way, there is no need to introduce an independent mediator entity in this context. As it turns out, in a two-player normal-form game this additional power does not buy the leader anything, but with more players it does [7]. Such signaling can also help in Bayesian games [24] and stochastic games [18], both from the perspective of increasing the leader's utility and from the perspective of making the computation easier.

It is straightforward to see that signaling can be useful in our limited commitment model as well. For example, if we just take Shapley's game with $|\mathcal{S}| = 1$, then by Proposition 3 without signaling we are stuck with the Nash equilibrium, but it seems we should be able to obtain the improved correlated equilibrium outcome with some form of signaling. But what is the right model of signaling here? We consider a very powerful model of signaling in this version of the paper. The full version of the paper also contains a discussion of weaker signaling models.

Definition 5. *In the* trusted mediator model*, the row player can design an independent trusted mediator that sends signals privately to each player according to a pre-specified joint distribution. After the round of play has completed, the mediator publicly reveals the signal sent to the row player.*

The after-the-fact public revelation of the signal sent to the row player allows the row player to commit to (i.e., in the long run develop a reputation for) responding to each signal with a particular distribution of play. Specifically, after each completed round, the column player learns the signal sent to, and the SIS played by, the row player.[2] Thus, if the row player according to the signal

[2] It is easy to get confused here—does the column player not learn more in a round purely by virtue of his own payoff from that round? It is important to remember

that she received was supposed to play an action from a particular SIS, then the column player can verify that she did. However, the row player may have an incentive to deviate *within* a SIS, because this is undetectable.

In the appendix of the full version of the paper, we show that under the trusted mediator model, without loss of generality a signal consists of just an action to play. With this in mind, we now define formally what it means for a correlated profile to have no undetectable beneficial deviations.

Definition 6. *A correlated profile σ has* no undetectable beneficial deviations *if (1) for all $c, c' \in C$ with $\sum_{r \in R} \sigma(r,c) > 0$, we have $\sum_{r \in R} \sigma(r,c)(u_2(r,c) - u_2(r,c')) \geq 0$, and (2) for all $S \in \mathcal{S}$, for all $r, r' \in S$ with $\sum_{c \in C} \sigma(r,c) > 0$, we have $\sum_{c \in C} \sigma(r,c)(u_1(r,c) - u_1(r',c)) \geq 0$.*

Note that, as is well known in the formulation of correlated equilibrium, in the first inequality, we can use $\sigma(r,c)$ rather than the more cumbersome $\sigma(r,c)/\sum_{r'' \in R} \sigma(r'',c)$, which would be the conditional probability of seeing r given a signal of c, because the denominator is a constant (similar for the second inequality). As a result, the condition that $\sum_{r \in R} \sigma(r,c) > 0$ is in fact not necessary because the inequality is vacuously true otherwise. This is what allows the standard linear program formulation of correlated equilibrium, as well as the linear program we present below in Figure 1.

Definition 7. *A correlated profile with no undetectable beneficial deviations is a* Stackelberg equilibrium with signaling and limited observation (SESLO) *if among such profiles it maximizes the row player's utility.*

Example. Consider the following game:

	A	B	C	D
a	0,0	12,0	0,1	0,0
b	0,1	0,0	12,0	0,0
c	12,0	0,1	0,0	0,0
d	5,0	5,0	5,0	0,1
e	7,0	7,0	7,0	1,1

Suppose $\mathcal{S} = \{\{a, b, c, d\}, \{e\}\}$. Then the following correlated profile (in which the signal an agent receives is which action to take) is a SESLO:

$$((1/9)(a, B), (1/9)(a, C), (1/9)(b, A), (1/9)(b, C), (1/9)(c, A), (1/9)(c, B),$$
$$(1/9)(e, A), (1/9)(e, B), (1/9)(e, C))$$

With this profile, for any signal the column player can receive, following the signal will give him utility 1/3, and so will any deviation. For any signal the row

that we are not considering repeated play by the column player. The idea is that the column player can observe over time the signals and how the row player acts *before* the column player ever acts. For discussion of security contexts in which certain types of players can receive messages that are inaccessible to other types, see Xu et al. [24].

player receives in SIS $\{a, b, c, d\}$, following the signal will give her 6; deviating to a, b, or c will give either 0 or 6, and deviating to d will give 5. The row player obtains utility 19/3 from this profile.[3] In contrast, without any commitment (if $|\mathcal{S}|$ had been 1), the outcome (e, D) would have been a SESLO, giving the row player utility only 1. Also, without signaling (but still with $\mathcal{S} = \{\{a, b, c, d\}, \{e\}\}$), the outcome (e, D) would have been a SELO. For consider a mixed-strategy profile without any undetectable beneficial deviations, and suppose it puts positive probability on at least one of A, B, and C. Then at least one of a, b, and c must get positive probability as well, for otherwise the column player would be better off playing D. Because a, b, and c are all in the same SIS and perform equally well against D, and because A, B, and C all perform equally well against d and e, if we condition on the players playing from a,b,c and A,B,C, the result must be a Nash equilibrium of that 3×3 game, which means that all of A, B, and C get the same probability. But in that case, d (which is in the same SIS) is a better response for the row player, and we have a contradiction. Hence any SELO involves the column player always playing D and the most the row player can obtain is 1.

We next have the following simple proposition that the ability to signal never hurts the row player.

Proposition 4. *The row player's utility from a SESLO is always at least that of a SELO.*

Proof. We show that an uncorrelated profile (σ_1, σ_2) that has no undetectable deviations (in the sense of Definition 3) also has no undetectable deviations (in the sense of Definition 6) when interpreted as a correlated profile σ (with $\sigma(r, c) = \sigma_1(r)\sigma_2(c)$); the result follows. First, for all $c, c' \in C$ with $\sum_{r \in R} \sigma(r, c) > 0$ (which is equivalent to $\sigma_2(c) > 0$), we have $\sum_{r \in R} \sigma(r, c)(u_2(r, c) - u_2(r, c')) = \sigma_2(c) \sum_{r \in R} \sigma_1(r)(u_2(r, c) - u_2(r, c')) = \sigma_2(c)(u_2(\sigma_1, c) - u_2(\sigma_1, c')) \geq 0$ by the best-response condition of Definition 3. Similarly, for all $S \in \mathcal{S}$, for all $r, r' \in S$ with $\sum_{c \in C} \sigma(r, c) > 0$ (which is equivalent to $\sigma_1(r) > 0$), we have $\sum_{c \in C} \sigma(r, c)(u_1(r, c) - u_1(r', c)) = \sigma_1(r) \sum_{c \in C} \sigma_2(c)(u_1(r, c) - u_1(r', c)) = \sigma_1(r)(u_1(r, \sigma_2) - u_1(r', \sigma_2)) \geq 0$ by the best-response-within-a-SIS condition of Definition 3.

Proposition 5. *If $|\mathcal{S}| = 1$ (all rows are indistinguishable), then a profile is a SESLO if and only if it is a correlated equilibrium that maximizes the row player's utility. If $|\mathcal{S}| = |R|$ (all rows are distinguishable), then a profile is a SESLO if and only if it is a Stackelberg equilibrium with signaling (which can do no better than a Stackelberg equilibrium without signaling).*

5 Computational Results

It turns out that with signaling, we do not face hardness. The linear program in Figure 1 can be used to compute a SESLO. It is a simple modification of the

[3] This was verified to be optimal using the linear program in Figure 1; same for the next case.

standard linear program for correlated equilibrium, the differences being that (1) for the row player, only deviations within a SIS are considered, and (2) there is an objective of maximizing the row player's utility. The special case where $|\mathcal{S}| = |R|$ has no constraints for the row player, and that special case of the linear program has previously been described by Conitzer and Korzhyk [7].

$$\begin{array}{rl} \textbf{maximize} & \sum_{r\in R, c\in C} u_1(r,c)p(r,c) \\ (\forall\ S\in\mathcal{S})\ (\forall\ r,r'\in S) & \sum_{c\in C}(u_1(r',c)-u_1(r,c))p(r,c)\le 0 \\ (\forall\ c,c'\in C) & \sum_{r\in R}(u_2(r,c')-u_2(r,c))p(r,c)\le 0 \\ & \sum_{r\in R, c\in C} p(r,c) = 1 \\ (\forall\ r\in R, c\in C) & p(r,c)\ge 0 \end{array}$$

Fig. 1. Linear program for computing a SESLO.

Theorem 2. *A SESLO can be computed in polynomial time.*

6 The Value of More Commitment Power

More strategies being distinguishable corresponds to more commitment power for the row player. As commitment power (in this particular sense) increases, does the utility the row player can obtain always increase gradually? (Note that it can never *decrease* the row player's utility, because all it will do is remove constraints in the optimization.) If she has close to full commitment power, does this guarantee her most of the benefit of full commitment power? Is some non-trivial minimal amount of commitment power necessary to obtain much benefit from it? The next two results demonstrate that the answer to all these questions is "no": there can be big jumps in the utility that the row player can obtain, both on the side close to full commitment power (Proposition 6) and on the side close to no commitment power (Proposition 7). (For an earlier study comparing the value of being able to commit completely to that of not being able to commit at all, see Letchford et al. [17]; for one assessing the value of correlation without commitment, see Ashlagi et al. [3].)

Proposition 6. *For any $\epsilon > 0$ and any $n > 1$, there exists an $n \times (n+1)$ game with all payoffs in $[0,1]$ such that if $|\mathcal{S}| = |R| = n$, the row player can obtain utility $1-\epsilon$ (even without signaling), but for any $\mathcal{S}$ with $|\mathcal{S}| < |R| = n$, the row player can obtain utility at most ϵ (even with signaling).*

Proof. Let $R = \{1, \ldots, n\}$ and $C = \{1, \ldots, n+1\}$. Let $u_1(i,j) = i\epsilon/n$ for $j \le n$, and let $u_1(i, n+1) = 1-(n-i)\epsilon/n$. Let $u_2(i,j) = (1+1/n)/2$ for $i \ne j$ and $j \le n$, let $u_2(i,i) = 0$ (for $i \le n$), and let $u_2(i, n+1) = 1/2$ for all i.

Suppose $|\mathcal{S}| = |R| = n$. Then, by Proposition 3, we are in the regular Stackelberg model, and the row player can commit to a uniform strategy, putting probability $1/n$ on each i. As a result the expected utility for the column player for playing some $j \leq n$ is $((n-1)/n)(1+1/n)/2 = (n-1)(n+1)/(2n^2) = (n^2-1)/(2n^2) < 1/2$, so to best-respond he needs to play $n+1$, resulting in a utility for the row player that is greater than $1-(n-1)\epsilon/n > 1-\epsilon$.

On the other hand, suppose that $|\mathcal{S}| < |R| = n$. Hence there exists some $S \in \mathcal{S}$ with $i, i' \in S$, $i < i'$. Note that i' strictly dominates i, so the row player will never play i in a SELO or even a SESLO. But then, the column player can obtain $(1+1/n)/2 > 1/2$ by playing i, and hence will not play $n+1$. As a result the row player obtains at most $n\epsilon/n = \epsilon$.

Proposition 7. *For any $\epsilon > 0$ and any $n > 1$, there exists an $n \times (n+1)$ game with all payoffs in $[0,1]$ such that for any $\mathcal{S}$ with $|\mathcal{S}| > 1$, the row player can obtain utility $1-\epsilon$ (even without signaling), but if $|\mathcal{S}| = 1$, the row player can only obtain utility 0 (even with signaling).*

Proof. Let $R = \{1, \ldots, n\}$ and $C = \{1, \ldots, n+1\}$. Let $u_1(i,j) = 1-\epsilon$ for $i \neq j$ and $j \leq n$, let $u_1(i,i) = 1$ (for $i \leq n$), and let $u_1(i, n+1) = 0$ for all i. Let $u_2(i,j) = 1$ for $i \neq j$ and $j \leq n$, let $u_2(i,i) = 0$ (for $i \leq n$), and let $u_2(i, n+1) = (n-1/2)/n$ for all i.

Suppose $|\mathcal{S}| > 1$. Then, the row player can commit to put 0 probability on some $S \in \mathcal{S}$, and therefore, 0 probability on some i. Hence, this i is a best response for the column player, and the row player obtains $1-\epsilon$. (The row player may be able to do better yet, but this is a feasible solution.)

On the other hand, suppose $|\mathcal{S}| = 1$. By Proposition 3, the row player can only obtain the utility of the best Nash equilibrium of the game for her (or, in the case with signaling, the utility of the best correlated equilibrium, by Proposition 5). We now show that in every Nash equilibrium (or even correlated equilibrium) of the game, the column player puts all his probability on $n+1$, from which the result follows immediately. For suppose the column player sometimes plays some $j \leq n$. Then, for the row player to best-respond, she has to maximize the probability of choosing the same j (conditional on the column player playing some $j \leq n$). (Or, more precisely in the case of correlated equilibrium, conditional on receiving any signal that leaves open the possibility that the column player plays some $j \leq n$, the row player has to maximize the probability of picking the same j.) She can always make this probability at least $1/n$ by choosing uniformly at random. Hence, the column player's expected utility (conditional on playing $j \leq n$) is at most $(n-1)/n$. But then $n+1$ is a strictly better response, so we do not have a Nash (or correlated) equilibrium.

Of course, the above two results are extreme cases. Can we say anything about what happens "typically"? To illustrate this, we present the results for randomly generated games in Figure 2. For each data point, 1000 games of size $m \times n$ were generated by choosing utilities uniformly at random. The rows were then evenly (round-robin) spread over a given number of SISes, and the game was solved using the GNU Linear Programming Kit (GLPK) with the linear program from Figure 1. The leftmost points (1 SIS) correspond to no commitment power (best correlated equilibrium), and the rightmost points (at least when the number of SISes is at least m) correspond to full commitment power (best Stackelberg mixed strategy). From this experiment, we can observe that most of the value of commitment is already obtained when moving from one SIS to two.

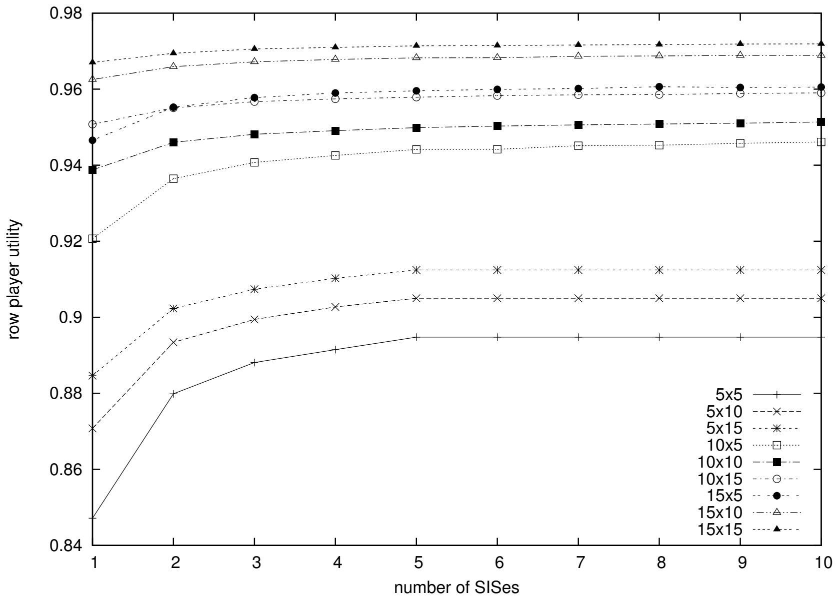

Fig. 2. Utility obtained by the row player as a function of commitment power (number of SISes), for various sizes of $m \times n$ games.

7 Conclusion

The model of the defender being able to commit to a mixed strategy has been popular in security games, motivated by the idea that the attacker can learn the distribution over time. This model has previously been questioned, and limited observability has previously been studied in various senses, including the attacker obtaining only a limited number of observations [19, 1] as well as the attacker observing (perfectly) only with some probability [15, 14]. Here, we considered a different type of limited observability, where certain courses of action are distinguishable from each other, but others are not. As a result, the row player's pure strategies partition into SISes, and she can commit to a distribution over SISes but not to how she plays within each SIS. We showed that it is NP-hard to compute a Stackelberg equilibrium with limited observation in this context, even when the SISes are small (Theorem 1). We then introduced a modified model with signaling and showed that in it, Stackelberg equilibria can be computed in polynomial time (Theorem 2). Finally, we showed that the cost of introducing a bit of additional unobservability can be large both when close to full observability (Proposition 6) and close to no observability (Proposition 7); however, in simulations, introducing a little bit of observability already gives most of the value of full observability.

Future research may be devoted to the following questions. Are there algorithms for computing a SELO that are efficient for special cases of the problem or that run fast on "typical" games? Another direction for future work concerns learning in games, which is a topic that has been thoroughly studied in the simultaneous-move case (see, e.g., Fudenberg and Levine [12]), but also already to some extent in the mixed-strategy commitment case [16, 5]. A model of learning in games with partial commitment needs to generalize models for both of these cases. Finally, can we mathematically prove what is suggested by the experiment in Figure 2, namely that in random games most of the value of commitment is already obtained with only two SISes?

Acknowledgments

I am thankful for support from ARO under grants W911NF-12-1-0550 and W911NF-11-1-0332, NSF under awards IIS-1527434, IIS-0953756, CCF-1101659, and CCF-1337215, and a Guggenheim Fellowship.

References

1. An, B., Kempe, D., Kiekintveld, C., Shieh, E., Singh, S.P., Tambe, M., Vorobeychik, Y.: Security games with limited surveillance. In: Proceedings of the National Conference on Artificial Intelligence (AAAI). Toronto, ON, Canada (2012)
2. An, B., Shieh, E., Tambe, M., Yang, R., Baldwin, C., DiRenzo, J., Maule, B., Meyer, G.: PROTECT - A deployed game theoretic system for strategic security allocation for the United States Coast Guard. AI Magazine 33(4), 96–110 (2012)
3. Ashlagi, I., Monderer, D., Tennenholtz, M.: On the value of correlation. Journal of Artificial Intelligence Research 33, 575–613 (2008)
4. Aumann, R.: Subjectivity and correlation in randomized strategies. Journal of Mathematical Economics 1, 67–96 (1974)
5. Balcan, M.F., Blum, A., Haghtalab, N., Procaccia, A.D.: Commitment without regrets: Online learning in Stackelberg security games. In: Proceedings of the ACM Conference on Economics and Computation (EC). pp. 61–78. Portland, OR, USA (2015)
6. Chen, X., Deng, X., Teng, S.H.: Settling the complexity of computing two-player Nash equilibria. Journal of the ACM 56(3) (2009)
7. Conitzer, V., Korzhyk, D.: Commitment to correlated strategies. In: Proceedings of the National Conference on Artificial Intelligence (AAAI). pp. 632–637. San Francisco, CA, USA (2011)
8. Conitzer, V., Sandholm, T.: Computing the optimal strategy to commit to. In: Proceedings of the ACM Conference on Electronic Commerce (EC). pp. 82–90. Ann Arbor, MI, USA (2006)
9. Conitzer, V., Sandholm, T.: New complexity results about Nash equilibria. Games and Economic Behavior 63(2), 621–641 (2008)

10. Daskalakis, C., Goldberg, P., Papadimitriou, C.H.: The complexity of computing a Nash equilibrium. SIAM Journal on Computing 39(1), 195–259 (2009)
11. Fang, F., Nguyen, T.H., Pickles, R., Lam, W.Y., Clements, G.R., An, B., Singh, A., Tambe, M., Lemieux, A.: Deploying PAWS: Field optimization of the protection assistant for wildlife security. In: Proceedings of the Twenty-Eighth Innovative Applications of Artificial Intelligence Conference (IAAI). Phoenix, AZ, USA (2016)
12. Fudenberg, D., Levine, D.: The Theory of Learning in Games. MIT Press (1998)
13. Gilboa, I., Zemel, E.: Nash and correlated equilibria: Some complexity considerations. Games and Economic Behavior 1, 80–93 (1989)
14. Korzhyk, D., Conitzer, V., Parr, R.: Solving Stackelberg games with uncertain observability. In: Proceedings of the Tenth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). pp. 1013–1020. Taipei, Taiwan (2011)
15. Korzhyk, D., Yin, Z., Kiekintveld, C., Conitzer, V., Tambe, M.: Stackelberg vs. Nash in security games: An extended investigation of interchangeability, equivalence, and uniqueness. Journal of Artificial Intelligence Research 41(2), 297–327 (2011)
16. Letchford, J., Conitzer, V., Munagala, K.: Learning and approximating the optimal strategy to commit to. In: Proceedings of the Second Symposium on Algorithmic Game Theory (SAGT). pp. 250–262. Paphos, Cyprus (2009)
17. Letchford, J., Korzhyk, D., Conitzer, V.: On the value of commitment. Autonomous Agents and Multi-Agent Systems 28(6), 986–1016 (2014)
18. Letchford, J., MacDermed, L., Conitzer, V., Parr, R., Isbell, C.: Computing optimal strategies to commit to in stochastic games. In: Proceedings of the National Conference on Artificial Intelligence (AAAI). pp. 1380–1386. Toronto, ON, Canada (2012)
19. Pita, J., Jain, M., Ordóñez, F., Tambe, M., Kraus, S.: Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. Artificial Intelligence 174(15), 1142–1171 (2010)
20. Pita, J., Jain, M., Western, C., Portway, C., Tambe, M., Ordóñez, F., Kraus, S., Parachuri, P.: Deployed ARMOR protection: The application of a game-theoretic model for security at the Los Angeles International Airport. In: Proceedings of the Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). pp. 125–132. Estoril, Portugal (2008)
21. Sandholm, T.: The state of solving large incomplete-information games, and application to poker. AI Magazine 31(4), 13–32 (2010), Special Issue on Algorithmic Game Theory
22. Tsai, J., Rathi, S., Kiekintveld, C., Ordóñez, F., Tambe, M.: IRIS - A tool for strategic security allocation in transportation networks. In: Proceedings of the Eighth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). pp. 37–44. Budapest, Hungary (2009)
23. von Stengel, B., Zamir, S.: Leadership games with convex strategy sets. Games and Economic Behavior 69, 446–457 (2010)
24. Xu, H., Freeman, R., Conitzer, V., Dughmi, S., Tambe, M.: Signaling in Bayesian Stackelberg games. In: Proceedings of the Fifteenth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). pp. 150–158. Singapore (2016)
25. Yin, Z., Jiang, A.X., Tambe, M., Kiekintveld, C., Leyton-Brown, K., Sandholm, T., Sullivan, J.P.: TRUSTS: Scheduling randomized patrols for fare inspection in transit systems using game theory. AI Magazine 33(4), 59–72 (2012)