Sensor-Reveal Games

João P. Hespanha

Denis Garagić

Abstract—We consider two-player nonzero-sum partial information games, called sensor-reveal games, in which one of the players (which we call the attacker) decides whether or not to engage in an illegal activity and the other player (which we call the defender) wants to detect the attacker’s action based on noisy sensor measurements. The partial information character of the game arises from the fact that the attacker controls which sensor measurements will be revealed to the defender; with the understanding that it may be costly to the attacker to reveal non-informative sensors, rather than sensors that carry useful information about the attack. Such games arise in several areas including computer security and law enforcement. We show that, for a very general sensor model, this game admits a closed form solution and provide explicit formulas for the Nash policies for both players. For scenarios in which the defender may not know the parameters that determine the cost function of the attacker, we provide a data-driven approach for the defender to compute an optimal policy based on fictitious play. The resulting algorithm is guaranteed to converge to a Nash equilibrium when both players rely on fictitious play. A brief numerical example illustrates the use of fictitious play.

Index Terms—Game Theory, Partial Information, Estimation, Cyber Security

I. INTRODUCTION

This paper considers a two-player sensor-reveal game in which an attacker decides whether or not to engage in an illegal activity and a defender wants to determine whether or not the attacker is indeed engaged in an illegal activity. Several noisy sensor may be used by the defender to make its decision, but it is the attacker who controls which sensor measurements will be revealed to the defender, with the understanding that it may be costly to the attacker to reveal uninformative sensors, rather than sensors that would give away its activity.

This problem is formalized as a non-cooperative game between the attacker and the defender, where the defender wants to minimize a cost associated with making bad decisions, whereas the attacker wants to minimize a cost associated with being caught. The attacker’s cost also includes rewards associated with pursuing the illegal activity and hiding informative sensors.

The sensor-reveal game provides a useful model for multiple problems in areas ranging from environmental conservation to computer security. The detection of illegal, unreported and unregulated (IUU) fishing is a national priority to many nations including the USA and recent reports on IUU fishing estimate that one in five fish in global markets are caught by vessels illegally, amounting to about $23.5B per year [2]. Tracking vessels through their on-board automatic identification system (AIS) is a cost effective mechanism to detect IUU fishing, which is currently in use by global non-profit conservation organizations like Global Fishing Watch [1]. However, vessels can turn off their AIS to evade detection, which complicates the detection of IUU through AIS monitoring. The decision by a vessel to engage in IUU fishing and turn off/on its AIS system can be modeled as a sensor-reveal game, as defined in this paper. Current systems detect and flag AIS anomalies [6], but do not reason about behaviors revealed by the AIS analysis in an adversarial context. A key novelty of using a game theoretical framework to interpret AIS in IUU fishing detection is that it does take into account adversary behavior.

Sensor-reveal games arise in computer security when cyber-defense system must make decisions, such as opening/closing firewalls, starting and stopping services, authorizing/deauthorizing users, and killing processes, based on reports from sensors that may have been tampered with by an attacker [17]. Such sensors include processes that log events like user authentication, network traffic, email activity, and access to services or files [8]. Especially relevant for this paper are scenarios where an attacker has infiltrated a system and gained privileges that would enabled her to turn off one or more sensors, with the understanding that taking such an action could be inferred from reports by other sensors that have not been compromised.

The sensor-reveal game falls into the category of partial-information games with a non-nested information structure because none of the players has strictly more information than the other. Note that, while the attacker decides which sensor to reveal, the sensors are stochastic and the attacker must make its decision of which sensor to reveal without knowing what measurement (realization) the sensor will actually reveal to the defender.

The mismatch between the information available to the players typically leads to a significant increase in complexity. This is because, as players plan their actions, they must hypothesize over all policies of the other player, as well as over all possible observations of the opponent, regardless of whether those are past or futures observations. This generally breaks down solutions that rely on some form of dynamic programming [3] to reduce complexity. Because of this, partial information games are poorly understood and...
the literature is much sparser than that for full information games. Notable exceptions are games with lack of information for one of the players [11, 16] and games with particular structures such as the Duel game [9], the Rabbit and Hunter game [5], the Searchlight game [12, 13], etc. Games of perfect recall can be expanded into sequence form, which can limit their overall computational complexity [10].

While it is typically computationally very difficult to solve partial information non-cooperative games, we will see that it is possible to compute Nash equilibria for the sensor reveal game by considering a representation in extensive form that limits the computational cost by analyzing independently different branches of the game’s decision tree. Our results provide explicit Nash policies for the attacker and the defender, in terms of the key game parameters (see Section III).

One challenge to implementing a Nash equilibrium for the defender is that its Nash policy depends on the values of several parameters that appear in the attacker’s cost, which is problematic because the defender may not know the attacker’s precise goals. To overcome this difficulty, in Section IV we consider fictitious play, which is a data-driven approach to compute the defender’s policy. This learning mechanism does not require knowledge of the attacker’s cost function and converges to the best response against any fixed policy used by the attacker (Nash or not). We use results by Berger for 2 × n bimatrix games to show that in the sensor reveal game we have convergence to a Nash equilibrium even when both players use fictitious play, which is generally not true [15]. A simulation example illustrates how fictitious play can adapt to an attacker that changes its policy, enabling the defender to maintain optimality without knowledge of the opponents intent.

II. THE SENSOR REVEAL GAME

Consider a two-player game in which one of the players (which we call the attacker) decides whether or not to engage in an illegal activity and the other player (which we call the defender) wants to detect that activity based on noisy sensor measurements. In the problem considered here, the attacker controls which sensor measurements will be revealed to the defender, with the understanding that it may be costly to the attacker to reveal “bad” (i.e., uninformative) sensors, rather than “good” (i.e., activity revealing) sensors.

Formally, we denote by \( \theta \) the decision by the attacker regarding whether or not to engage in the illegal activity, with the understanding that

\[
\theta = \begin{cases} 
1 & \text{attacker engages in illegal activity} \\
0 & \text{attacker does not engage in illegal activity.}
\end{cases}
\]

and we denote by \( \hat{\theta} \) the defender’s estimate of the value of \( \theta \). The defender wants to minimize a cost of the form

\[
J_{\text{def}} := AP(\hat{\theta} = 1, \theta = 0) + B P(\hat{\theta} = 0, \theta = 1),
\]

where \( A \geq 0 \) and \( B \geq 0 \) are parameters that establish the cost of a false detection and of a missed detection, respectively.

We assume that the attacker can choose to reveal the measurement of one of \( N \) noisy sensors. Each sensor \( i \in \{1, 2, \ldots, N\} \) produces a measurement \( y_i \in \mathcal{Y} \) and the attacker selects which sensor to reveal to minimize a cost of the form

\[
J_{\text{att}} := -R \theta + C P(\hat{\theta} = 1, \theta = 1)
- F P(\hat{\theta} = 1, \theta = 0) + S_\sigma,
\]

where \( R \geq 0 \) is a reward associated with engaging in the illegal activity (i.e., \( \theta = 1 \)), \( C \geq 0 \) is the cost of being caught, \( \sigma \) is the sensor that the attacker decides to reveal, \( S_\sigma \) the cost of revealing sensor \( \sigma \), and \( F \geq 0 \) is a reward to the attacker for generating a false alarm. The following assumption is introduced to exclude a trivial solution to the problem:

**Assumption 1:** The reward \( R - C \) associated with setting \( \theta = 1 \) and being caught is smaller than the reward \( F \) of generating a false alarm, i.e.,

\[
R - C < F.
\]

This assumption excludes the trivial solution for the attacker to always select \( \theta = 1 \).

Formally, we have a nonzero sum game where the attacker has two decision variables \( (\theta, \sigma) \) and the defender has a single decision variable \( \hat{\theta} \), which must be selected based on knowledge of which sensor \( \sigma \) was revealed and the value \( y_\sigma \) reported by the sensor. Deterministic estimation policies for the defender are thus decisions rules of the form

\[
\hat{\theta} = f(\sigma, y_\sigma),
\]

where \( \sigma \) is the sensor revealed by the attacker, \( y_\sigma \) the actual sensor measurement that the defender received to make its decision, and \( f(\cdot) \) a binary function from \( \{1, 2, \ldots, N\} \times \mathcal{Y} \) to \{0, 1\}. For the purposes of our analysis, it is convenient to express (3) as

\[
\hat{\theta} = \begin{cases} 
1 & y_\sigma \in Y_\sigma, \\
0 & y_\sigma \notin Y_\sigma,
\end{cases}
\]

where each \( Y_i, i \in \{1, 2, \ldots, N\} \) denotes the subset of elements \( y_i \in \mathcal{Y} \) for which \( f(i, y_i) = 1 \). Associated with each set \( Y_i \) we define the parameters

\[
p_{\text{fp}}^i := P(y_i \in Y_i | \theta = 0), \quad p_{\text{fn}}^i := P(y_i \notin Y_i | \theta = 1),
\]

which can be regarded as the corresponding sensor’s probabilities of a false positive and a false negative, respectively, and provide a measure of the sensor’s reliability. The problem becomes nontrivial when the different sensors have different levels of “reliability” and it is costly for the attacker to reveal sensors that convey to the defender very little information about the true value of \( \theta \), i.e., the problem is mostly interesting when \( S_i \) is larger (costly to reveal sensors) for sensors with large values for \( p_{\text{fp}}^i \) and \( p_{\text{fn}}^i \) (not very informative).
A. Nature of Sensor Measurements

The notion of “sensor” and “measurement” considered in this paper is kept very general to make sure that our results cover a wide range of problems. Specifically, we allow each measurement $y_i$ to range from a simple real-valued random variable to a vector-valued stochastic process, defined either in continuous or discrete-time; with the understanding that the defender’s policy (3) must be measurable in the appropriate sense.

Consider, for example, the IUU fishing detection problem mentioned in the introduction, where the attacker is a vessel potentially engaged in IUU fishing. In this problem, a particular sensor measurement $y_i$ typically includes satellite imagery and AIS measurements collected over a given period of time. The attacker has little control over the satellite imagery being collected, so all sensors $i \in \{1, 2, \ldots, N\}$ will include those data, but it does control when to turn on/off its AIS. This could be modeled by associating with $y_1$ a measurement that contains only satellite imagery (AIS always off) and with $y_2$ a measurement that contains satellite imagery and full AIS data (AIS always on). Our formulation also permits intermediate scenarios where the AIS is turns on/off intermittently, corresponding to other forms of measurements $y_i$, that may all still include satellite imagery, but differ by how many times the attacker exposed the AIS data over the interval of time of interest. In practice, the number $N$ of possible ways the AIS data can be revealed to the defender is very large, so we are mostly interested in solutions that scale well with the number $N$ of sensors that the attacker reveals. Analogous situations arise in the computer security domain, where an attacker may choose to turn on/off different combinations of cyber-security sensors.

III. Computation of Nash Equilibria

For the purpose of computing a Nash equilibrium for this game, it is convenient to consider the extensive form decision tree depicted in Figure 1, where the branches represent the player’s decisions and the dashed ellipses represent the information sets for the defender, i.e., sets of decision points that are indistinguishable based on the information available to the defender [7]. This representation of the game permits the independent analysis of each subtree concerning to a particular choice for $\sigma$ by the attacker. To this effect, suppose that the attacker selected a particular sensor $\sigma = i \in \{1, 2, \ldots, N\}$ and consider the pure (i.e., deterministic) choices that each player needs to consider on the subtree corresponding to $\sigma = i$:

1) The attacker must select either $\hat{\theta} = 0$ or $\hat{\theta} = 1$.
2) The defender must select a policy $\mu$ that maps $y_i \in \{0, 1\}$ to the estimate $\hat{\theta} \in \{0, 1\}$. We consider 3 possible policies: the defender ignores the measurement $y_i$ and always selects $\hat{\theta} = 0$, it ignores the measurement and always selects $\hat{\theta} = 1$, or it follows the sensor recommendation and selects $\hat{\theta}$ based on whether or not $y_i \in Y_i$. Formally, these 3 options correspond the following 3 policies:

$$
\hat{\theta} = \begin{cases} 
\mu_1(y_i) = 0, & \forall y_i \in \mathcal{Y} \\
\mu_2(y_i) = 1, & \forall y_i \in \mathcal{Y} \\
\mu_3(y_i) = \begin{cases} 
1 & \forall y_i \in Y_i \\
0 & \forall y_i \notin Y_i 
\end{cases}
\end{cases}
$$

(6)

For the time being we assume that the set $Y_i$ associated with the decision $\hat{\theta} = 1$ for sensor $i$ is given, but we will later discuss how to “optimize” this set, which is really a decision variable for the defender. In principle, there is a forth possible policy $\mu_4(y_i) = 1 - \mu_3(y_i)$ where the defender chooses the opposite of the sensor representation, but that policy will not be needed to compute a Nash equilibrium.

Straightforward computations can be used to show that the problem corresponding to the subtree $\sigma = i$ in Figure 1 can be represented by the following bi-matrix game

$$
A^{i}_{\text{att}} := S_i + \begin{bmatrix} \theta = 0 & \theta = 1 & \hat{\theta} = 1 \text{ if } y_i \in Y_i \\
\theta = 0 & 0 & -F \\
\theta = 1 & -R & -R + C' \\
\hat{\theta} = 1 & -R + C(1 - p_{i\text{in}}) & -R + C(1 - p_{i\text{in}}) \end{bmatrix}
$$

(7a)

$$
B^{i}_{\text{def}} := \begin{bmatrix} \theta = 0 & \theta = 1 & \hat{\theta} = 1 \text{ if } y_i \in Y_i \\
\theta = 0 & 0 & A p_{i\text{in}} \\
\theta = 1 & B & 0 \end{bmatrix}
$$

(7b)

where $A^{i}_{\text{att}}$ and $B^{i}_{\text{def}}$ should be viewed as cost matrices for the attacker and defender, respectively, and each row and column was labeled with the corresponding policies for the attacker and defender, respectively.

The following result, proved in Section III-A, provides explicit formulas for a mixed Nash equilibrium for the bi-matrix game (7) associated with the attacker’s decision to reveal sensor $\sigma = i$.

**Theorem 1**: Under Assumption 1 and further assuming that

$$
\hat{p}_{i\text{in}} + \hat{p}_{i\text{in}} \leq 1,
$$

the bi-matrix game (7) has a mixed Nash equilibrium of the form

$$
y^{*}_{\text{att}} = \begin{bmatrix} \frac{B(1 - \hat{p}_{i\text{in}})}{B(1 - \hat{p}_{i\text{in}}) + A\hat{p}_{i\text{in}}} & \hat{C} \geq R \\
\frac{B(1 - \hat{p}_{i\text{in}}) + A\hat{p}_{i\text{in}}}{B(1 - \hat{p}_{i\text{in}}) + A\hat{p}_{i\text{in}}} \end{bmatrix}
$$

(8a)

$$
y^{*}_{\text{def}} = \begin{bmatrix} \frac{\hat{C} - R}{C'} \\
\frac{0}{C'} \end{bmatrix} \hat{C} < R
$$

(8b)

with values

$$
J^{*}_{\text{att}} = S_i - F \begin{bmatrix} \frac{R\hat{p}_{i\text{in}}}{C'} & \hat{C} \geq R \\
\frac{R\hat{p}_{i\text{in}}}{C'} \frac{(R - C)(1 - \hat{p}_{i\text{in}}) + C\hat{p}_{i\text{in}}}{C + F - C'} & \hat{C} < R
\end{bmatrix}
$$
\[ J_{\text{def}}^* = \begin{cases} \frac{ABp_{\text{fp}}^i}{B(1-p_{\text{in}}^i)+Ap_{\text{ip}}^i}, & \tilde{C}^i \geq R \\ \frac{ABp_{\text{fn}}^i}{A(1-p_{\text{in}}^i)+Bp_{\text{in}}^i}, & \tilde{C}^i < R \end{cases} \]

where \( \tilde{C}^i := C(1-p_{\text{in}}^i) + Fp_{\text{ip}}^i \).

Two conclusions can be drawn from Theorem 1:

1) For each sensor \( i \), the defender should select the set \( Y_i \) in (4) that minimizes its cost

\[ J_{\text{def}}^* = \begin{cases} \frac{ABp_{\text{fp}}^i}{B(1-p_{\text{in}}^i)+Ap_{\text{ip}}^i}, & \tilde{C}^i \geq R \\ \frac{ABp_{\text{fn}}^i}{A(1-p_{\text{in}}^i)+Bp_{\text{in}}^i}, & \tilde{C}^i < R \end{cases} \quad (9) \]

2) The attacker should choose to reveal the sensor \( i \) that leads to the smallest value of its cost

\[ J_{\text{att}}^* = S_i - F \left[ \frac{Rp_{\text{ip}}^i}{C(1-p_{\text{in}}^i)+Cp_{\text{in}}^i}, \tilde{C}^i \geq R \right] \left[ \frac{C}{F-C}, \tilde{C}^i < R \right]. \]

When the attacker gets no reward for generating false alarms \( (F = 0) \), the choice of which sensor to use is purely based on the costs \( S_i \) of revealing the sensors and the attacker will always select the least costly sensor.

Remark 1 (Complex Sensors): As noted in Section II-A, each sensor measurement \( y_i \) may be an object with a complex stochastic characterization. However, Theorem 1 shows that while the different measurement models can be quite complex, the key parameters that affect the Nash equilibrium are the reliability parameters \( p_{\text{ip}}^i \) and \( p_{\text{in}}^i \) in (5). For simple models, these parameters may be computed analytically, but in realistic scenarios they may need to be learned from data. Either way these parameters do not depend on the attacker’s intentions and can be computed/learned using standard Bayesian techniques to minimize (9).

A. Proof of Theorem 1

We consider separately two types of mixed Nash equilibria:

1) We start by considering mixed Nash equilibria of the form

\[ y := \left[ y_1 \quad 1 - y_1 \right]^T, \quad z := \left[ z_1 \quad 0 \quad 1 - z_1 \right]^T, \]

\[ y_1, z_1 \in [0, 1], \]

\[ A_{\text{att}}^z = \frac{p}{p}, \quad y' B_{\text{def}}^z = \left[ \tilde{q} \quad q \quad q \right], \quad \tilde{q} \geq q, \]

that would satisfy the usual quadratic program for mixed Nash equilibria [7, Chapter 10]. These equations lead to

\[ p = S_i z_1 + (-Fp_{\text{ip}}^i + S_i)(1 - z_1) \]

\[ = (-R + S_i) z_1 + (-R + C(1-p_{\text{in}}^i) + S_i)(1 - z_1) \]

\[ q = B(1 - y_1) = Ap_{\text{ip}}^i y_1 + Bp_{\text{in}}^i (1 - y_1) \]

\[ \tilde{q} = A y_1 \]

from which we get

\[ z_1 = \frac{Fp_{\text{ip}}^i + C(1-p_{\text{in}}^i) - R}{Fp_{\text{ip}}^i + C(1-p_{\text{in}}^i)}, \]

\[ y_1 = \frac{B(1-p_{\text{in}}^i) + Ap_{\text{ip}}^i}{B(1-p_{\text{in}}^i) + Ap_{\text{ip}}^i}, \]

\[ p = \frac{S_i(Fp_{\text{ip}}^i + C(1-p_{\text{in}}^i)) - RFp_{\text{ip}}^i}{Fp_{\text{ip}}^i + C(1-p_{\text{in}}^i)}, \]

\[ q = \frac{ABp_{\text{in}}^i}{B(1-p_{\text{in}}^i) + Ap_{\text{ip}}^i}, \]

\[ \tilde{q} = \frac{AB(1-p_{\text{in}}^i)}{B(1-p_{\text{in}}^i) + Ap_{\text{ip}}^i}. \]

So we have a Nash equilibrium as long as

\[ Fp_{\text{ip}}^i + C(1-p_{\text{in}}^i) \geq R, \]

which guarantees that \( z_1, y_1 \in [0, 1] \) and

\[ p_{\text{ip}}^i + p_{\text{in}}^i \leq 1 \]

which guarantees that \( \tilde{q} \geq q \).

2) We next consider mixed Nash equilibria of the form

\[ y := \left[ y_1 \quad 1 - y_1 \right]^T, \quad z := \left[ 0 \quad z_2 \quad 1 - z_2 \right]^T, \]

\[ y_1, z_2 \in [0, 1], \]

\[ A_{\text{att}}^z = \frac{p}{p}, \quad y' B_{\text{def}}^z = \left[ \tilde{q} \quad q \quad q \right], \quad \tilde{q} \geq q. \]

These equations lead to

\[ p = (-F + S_i) z_2 + (-Fp_{\text{ip}}^i + S_i)(1 - z_2) \]

\[ = (-R + C + S_i) z_2 \]

\[ + (-R + C(1-p_{\text{in}}^i) + S_i)(1 - z_2) \]

\[ q = Ay_1 = Ap_{\text{ip}}^i y_1 + Bp_{\text{in}}^i (1 - y_1) \]

\[ \tilde{q} = B(1 - y_1) \]
from which we get
\[ z_2 = \frac{R - C(1 - p_{i_{in}}^i) - Fp_{i_{ip}}^i}{F(1 - p_{i_{in}}^i) + Cp_{i_{in}}^i}, \]
\[ y_1 = \frac{Bp_{i_{in}}^i}{A(1 - p_{i_{ip}}^i) + Bp_{i_{in}}^i}, \]
\[ p = \frac{F(S_t - R + C)(1 - p_{i_{ip}}^i) + C(S_t - F)p_{i_{in}}^i}{F(1 - p_{i_{ip}}^i) + Cp_{i_{in}}^i}, \]
\[ q = \frac{ABp_{i_{in}}^i}{A(1 - p_{i_{ip}}^i) + Bp_{i_{in}}^i}, \]
\[ \bar{q} = \frac{AB(1 - p_{i_{ip}}^i)}{A(1 - p_{i_{ip}}^i) + Bp_{i_{in}}^i}. \]

So we now have a Nash equilibrium as long as
\[ Fp_{i_{ip}}^i + C(1 - p_{i_{in}}^i) \leq R, \quad R - C \leq F, \]
\[ p_{i_{in}}^i + p_{i_{ip}}^i \leq 1. \]

IV. FICTITIOUS PLAY

In fictitious play, the defender constructs a running average \( \bar{y}^i(t) \) of the mixed policy \( y^i(t) \) used by the attacker when the sensor \( i \) is selected:
\[ \bar{y}^i(t) = \frac{1}{t} \sum_{k=1}^{t} y^i(k) \] (10a)

and uses a mixed policy
\[ z^i(t) \in \beta^i_{def}(\bar{y}^i(t)) := \arg\min_{z \in \mathbb{S}^2} \bar{y}^i(t)'B_{def}^i z, \] (10b)

where \( \mathbb{S}^2 \) denotes the simplex of probability distributions in \( \mathbb{R}^2 \) and \( \beta^i_{def}(\bar{y}^i(t)) \) the set of the defender’s best response against the attacker’s policy \( \bar{y}^i(t) \). The inclusion in (10b) means that the defender may choose any of the (possibly many) best responses against \( \bar{y}^i(t) \). Fictitious play has the following desirable features:

(i) The defender’s dynamics only depend on parameters that appear in its own cost matrix \( B_{def}^i \) and on a running average of the attackers policy, which can be computed based solely on observing the attacker’s actions. Crucially, the defender can use fictitious play even when the attacker’s intentions (which are encoded in the parameters of its cost matrix \( A_{att}^i \)) are unknown to the defender.

(ii) If the attacker is using a constant policy \( y^i(t) = y^i, \forall t \), the defender’s policy is guaranteed to converge to the best response against \( y^i \). In particular, if the attacker is playing a Nash equilibrium policy, then the defender’s policy converges to the Nash equilibrium.

An additional desirable property of fictitious play that does not hold for every game, but that does hold for the game considered in this paper, is that if the attacker is also playing fictitious play, i.e.,
\[ z^i(t) = \frac{1}{t} \sum_{k=1}^{t} z^i(k) \] (11a)
\[ y^i(t) \in \beta^i_{att}(\bar{z}^i(t)) := \arg\min_{y \in \mathbb{S}^2} y'B_{att}^i \bar{z}^i(t), \] (11b)

then both players are guaranteed to converge to a Nash equilibrium. This is stated in the following theorem that follows directly from results in [4].

**Theorem 2:** Assume that
\[ A \neq 0, B \neq 0, R \neq 0, R - C \neq F, \]
\[ C(1 - p_{i_{in}}^i) + Fp_{i_{ip}}^i \neq R, \]
\[ p_{i_{ip}}^i \in (0, 1), p_{i_{in}}^i \in (0, 1), \] (12)
then (10)–(11) converges to a Nash equilibrium of the bimatrix game (7).

The key assumptions needed to apply the results in [4] are that one of the players only has 2 actions (in our case the attacker) and that the game is not degenerate, i.e., that for every pure strategy of a player there is a unique best response. In our bimatrix game, non-degeneracy means that all columns of \( A_{att}^i \) and all rows \( B_{def}^i \) have no repeated entries, which is guaranteed by (12). This property holds generically.

Figure 2 illustrates the use of fictitious play by the defender in a scenario where the attacker starts by using the fixed (non-Nash) policy \( y^i = [0 1]' \) that corresponds to always pursuing the illegal activity (\( \theta = 1 \)) and, at time \( t = 10^5 \), switches to also using fictitious play. We can see the defender’s policy first adjusting to the best response to \( y^i = [0 1]' \), which not surprisingly turns out to be always selecting \( \hat{\theta} = 1 \). When the attacker switches to fictitious play, then both policies converge to the Nash equilibrium predicted by Theorem 1, as expected in view of Theorem 2. The figure also shows the evolution of the rewards for the both players, where we can see the cost for the defender initially decreasing as its policies adjusts to the best response to \( y^i = [0 1]' \). When the attacker switches to fictitious play, its cost decreases and the defender’s cost increases, which is consistent with the fact that the attacker benefits by abandoning its original non-Nash policy. In these simulations, we have used averaging with fading memory, i.e.,
\[ \bar{y}^i(t + 1) = (1 - \gamma)\bar{y}^i(t) + \gamma y^i(t), \]
\[ \bar{z}(t + 1) = (1 - \gamma)\bar{z}(t) + \gamma z(t), \gamma \in (0, 1), \]

for some \( \gamma \in (0, 1) \), which is more robust to changes in the opponents policy and is akin to continuous-time best response dynamics for which [4] also provides convergence results under the same assumptions on the bimatrix game.

V. FUTURE WORK

An action for the defender not considered in the current version of the sensor-reveal game is to postpone the decision to declare a specific value for \( \hat{\theta} \) and, instead, request additional sensor data. In a computer security application this could correspond to asking for more detailed logging and in an IUU fishing scenario this could correspond to requesting additional satellite imagery. In practice, this would mean an additional action for the defender and therefore additional columns for the matrices \( A_{att}^i \) and \( B_{def}^i \) in (7). It would also
mean additional terms in the defender’s cost function (1) to penalize delaying a decision and to consider the cost involved in getting the additional measurements. This would not fundamentally change the methodology that we have used to compute the Nash equilibrium but may change the type of equilibrium found and its dependence on the game parameters.

Another important variation of this problem arises when multiple attackers act in a cooperative fashion, either because the illegal activity is a cooperative endeavor that requires multiple agents or because the defender has limited sensor resources that must be allocated to survey the different attackers.

**REFERENCES**