Sensor-Reveal Games

João P. Hespanha Denis Garagić

***Abstract*— We introduce 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 provides the measurements that 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 introduces 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 provides the measurements that will be available 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 decision, 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 solution concept pursued here is that of a Nash equilibria, which captures the goals of both players and guarantees that no player will regret its decision, once the decision of the other player becomes known [7].

This material is based upon work supported by DARPA under the contract HT0011990016 and by the U.S. Office of Naval Research under the MURI grant No. N00014-16-1-2710.

J. Hespanha is with the University of California, Santa Barbara, USA.
D. Garagić is with BAE Systems, FAST Labs, Burlington, MA, USA.

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, because the path of a vessel engaged in fishing is significantly different from that of a cargo vessel. AIS signals are thus currently in use by global non-profit conservation organizations like Global Fishing Watch [1]. However, vessels can turn off their AIS to hide their paths, which complicates the detection of IUU fishing through AIS monitoring. The decision by a vessel to engage in IUU fishing and turn off/on the global position reports provided by AIS can be modeled as a sensor-reveal game, as defined in this paper. With AIS off, the vessel hides a path that may reveal IUU fishing, but AIS is primarily a safety system used by rescue operations to predict the location of ships in trouble, so turning AIS off will increase risk. In addition, and more importantly from the perspective of adversarial reasoning, turning off AIS can be a strong indication of IUU activity. 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 [19]. 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, we consider stochastic sensors and the attacker must make the decision of which sensors to reveal without knowing which measurements (realizations) the

sensors will actually reveal to the defender, which justifies the terminology "sensor-reveal" rather than "measurement-reveal." Games in which an attacker selects specific measurements from compromised stochastic sensors have been considered in [18].

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, 17] and games with particular structures such as the Duel game [9], the Rabbit and Hunter game [5], the Searchlight game [13, 14], etc. Games of perfect recall can be expanded into sequence form, which can limit their overall computational complexity [10]. Some reduction in complexity is also possible by exploring the availability of information that is common to both players [12].

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 the 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 [4] for $2 \times n$ 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 [16]. 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 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 θ 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 θ. The defender wants to minimize a cost of the form

$$J_{\text{def}} \coloneqq A\,\mathrm{P}(\hat{\theta} = 1, \theta = 0) + B\,\mathrm{P}(\hat{\theta} = 0, \theta = 1), \quad (1)$$

where $A \geqslant 0$ and $B \geqslant 0$ 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

$$J_{\text{att}} \coloneqq -R\theta + C\,\mathrm{P}(\hat{\theta} = 1, \theta = 1) - F\,\mathrm{P}(\hat{\theta} = 1, \theta = 0) + S_\sigma, \quad (2)$$

where $R \geqslant 0$ is a reward associated with engaging in the illegal activity (i.e., choosing $\theta = 1$), $C \geqslant 0$ is the cost of being caught, σ is the sensor that the attacker decides to reveal, S_σ the cost of revealing sensor σ, and $F \geqslant 0$ is a rewards 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 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 (θ and σ) and the defender has a single decision variable $\hat{\theta}$, which must be selected based on knowledge of which sensor σ was revealed and the value y_σ reported by the sensor. Deterministic estimation policies for the defender are thus decisions rules of the form

$$\hat{\theta} = \delta(\sigma, y_\sigma), \quad (3)$$

where σ is the sensor revealed by the attacker, y_σ the actual sensor measurement that the defender received to make its decision, and $\delta(\cdot)$ a 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

$$\hat{\theta} = \begin{cases} 1 & y_\sigma \in Y_\sigma, \\ 0 & y_\sigma \notin Y_\sigma, \end{cases} \quad (4)$$

where each $Y_i \subset \mathcal{Y}$, $i \in \{1, 2, \ldots, N\}$ denotes the subset of elements y_i in $\mathcal{Y}$ for which $\delta(i, y_i) = 1$. In practice, the

representation in (4) for the defender policy in (3), means that we can regard the sets Y_i, $i \in \{1, 2, \ldots, N\}$ as the defender's policy. Associated with each set Y_i, we define the parameters

$$p^i_{\text{fp}}(Y_i) \coloneqq \text{P}(y_i \in Y_i \,|\, \theta = 0),$$
$$p^i_{\text{fn}}(Y_i) \coloneqq \text{P}(y_i \notin Y_i \,|\, \theta = 1),$$

that can be regarded as the corresponding probabilities of a false positive and a false negative, respectively, and provide a measure of the sensor's reliability for the given set Y_i. The problem becomes especially interesting, 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 θ, i.e., the problem is especially interesting when S_i is larger (costly to reveal sensor i) for sensors with large values for $p^i_{\text{fp}}(Y_i)$ and $p^i_{\text{fn}}(Y_i)$ (sensor i not very informative).

A *pure Nash equilibrium* for this game is thus a collection of (deterministic) sets $\left\{Y_i^* : i \in \{1, 2, \ldots, N\}\right\}$ that define the defender's policy and a (deterministic) choice (θ^*, σ^*) for the attacker such that

1) when the defender uses the decision rule (4) based on the sets $\{Y_i^*\}$, the attacker's cost J_{att} in (2) is minimized for the pair (θ^*, σ^*), over all possible $(\theta, \sigma) \in \{0, 1\} \times \{1, 2, \ldots, N\}$; and
2) when the attacker selects (θ^*, σ^*), the defender's cost J_{def} in (1) is minimized by using the sets $\{Y_i^*\}$ in (4), over all possible sets $\{Y_i \subset \mathcal{Y}\}$.

Following standard terminology, a *mixed Nash equilibrium* follows a similar definition, but with the deterministic choices replaced by distributions over the sets of all deterministic policies [7]. Specifically, a mix Nash equilibrium consists of a probability distribution over all possible sets $\{Y_i \subset \mathcal{Y}\}$ for the defender and a probability distribution over all possible $(\theta, \sigma) \in \{0, 1\} \times \{1, 2, \ldots, N\}$ for the attacker. We shall see, however, that there will be no need to randomize over the sensor selection σ, just $\{Y_i\}$ and θ.

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. This means that when we say that the attacker chooses to reveal "one sensor out of N", this may actually correspond to selecting "one combination of sensors out of N possible combinations."

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 AIS measurements and satellite imagery 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 turned 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 chose 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 corresponding to a particular choice for σ 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 $\theta = 0$ or $\theta = 1$.
2) The defender must select the set Y_i that defines the estimate $\hat{\theta}$ in (4), with $Y_i \subset \mathcal{Y}$ selected among all possible subsets of $\mathcal{Y}$. In general, the number of options for each Y_i may be infinite and even uncountable. However, in this paper we restrict the defender's choice to a finite set of possibilities for each Y_i, that we enumerate as follows

$$\varnothing, \mathcal{Y}, \mathcal{Y}_i^{(1)}, \mathcal{Y}_i^{(2)}, \ldots, \mathcal{Y}_i^{(M_i)}, \tag{5}$$

where $\varnothing$ denotes the empty set (corresponding to always setting $\hat{\theta} = 0$), $\mathcal{Y}$ the set of all possible measurements (corresponding to always setting $\hat{\theta} = 1$), and the remaining $\mathcal{Y}_i^{(j)}$ correspond to intermediate policies.

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 $2 \times (2 + M_i)$ bi-matrix game

$$A^i_{\text{att}} \coloneqq \begin{bmatrix} & \varnothing\,(\hat{\theta}{=}0) & \mathcal{Y}\,(\hat{\theta}{=}1) & \mathcal{Y}_i^{(1)} & \cdots \\ \theta{=}0 & S_i & S_i{-}F & S_i{-}Fp^i_{\text{fp}}(\mathcal{Y}_i^{(1)}) & \cdots \\ \theta{=}1 & S_i{-}R & S_i{-}R{+}C & S_i{-}R{+}C\left(1{-}p^i_{\text{fn}}(\mathcal{Y}_i^{(1)})\right) & \cdots \end{bmatrix} \tag{6a}$$

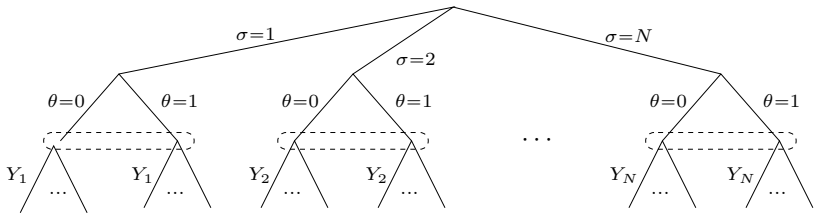

Fig. 1. Extensive form representation of the sensor-reveal game, with the top two branches corresponding to attacker decisions (which sensor to select and whether or not to engage in the illegal activity) and the bottom branch to the defender's decision [selection of the sets $\{Y_i\}$ that define the estimate in (4)].

$$B^i_{\text{def}} \coloneqq \begin{bmatrix} & \varnothing\,(\hat\theta{=}0) & \mathcal{Y}\,(\hat\theta{=}1) & \mathcal{Y}_i^{(1)} & \mathcal{Y}_i^{(2)} & \cdots \\ \theta{=}0 & 0 & A & Ap^i_{\text{fp}}(\mathcal{Y}_i^{(1)}) & Ap^i_{\text{fp}}(\mathcal{Y}_i^{(2)}) & \cdots \\ \theta{=}1 & B & 0 & Bp^i_{\text{fn}}(\mathcal{Y}_i^{(1)}) & Bp^i_{\text{fn}}(\mathcal{Y}_i^{(2)}) & \cdots \end{bmatrix}, \tag{6b}$$

where A^i_{att} and B^i_{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 (as enumerated above). In the sequel, we shall compute mixed policies for the attacker and defender that correspond to probability distributions over the columns and rows of these matrices, respectively.

The remaining of this section is focused on the case $M_i = 1$ of only 3 sets in (5). This enable us to simplify the notation by dropping the argument $(\mathcal{Y}_i^{(1)})$ from $p^i_{\text{fp}}(\mathcal{Y}_i^{(1)})$ and $p^i_{\text{fn}}(\mathcal{Y}_i^{(1)})$. While considering only one "non-trivial" set in (5) may seem restrictive, we shall see that it is possible to select the set $\mathcal{Y}_i^{(1)}$ to minimize the defender's cost, which significantly decreases the conservativeness of this simplification.

The following result, proved in Section III-A, provides explicit formulas for a mixed Nash equilibrium for the bimatrix game (6) associated with the attacker's decision to reveal sensor $\sigma = i$. The results of this theorem will subsequently be used to determine which sensor the attacker should reveal.

Theorem 1: Consider the case of $M_i = 1$ of only 3 sets in (5) with

$$p^i_{\text{fp}} + p^i_{\text{fn}} \leqslant 1. \tag{7}$$

Under Assumption 1, the bimatrix game (6) has a mixed Nash equilibrium of the form

$$y^i_{\text{att}}{}^* = \begin{cases} \left[\, \frac{B(1-p^i_{\text{fn}})}{Ap^i_{\text{fp}}+B(1-p^i_{\text{fn}})} \;\; \frac{Ap^i_{\text{fp}}}{Ap^i_{\text{fp}}+B(1-p^i_{\text{fn}})} \,\right]' & \bar C^i \geqslant R \\ \left[\, \frac{Bp^i_{\text{fn}}}{A(1-p^i_{\text{fp}})+Bp^i_{\text{fn}}} \;\; \frac{A(1-p^i_{\text{fp}})}{A(1-p^i_{\text{fp}})+Bp^i_{\text{fn}}} \,\right]' & \bar C^i < R \end{cases} \tag{8a}$$

$$z^i_{\text{def}}{}^* = \begin{cases} \left[\, \frac{\bar C^i - R}{\bar C^i} \;\; 0 \;\; \frac{R}{\bar C^i} \,\right]' & \bar C^i \geqslant R \\ \left[\, 0 \;\; \frac{R-\bar C^i}{C+F-\bar C^i} \;\; \frac{C+F-R}{C+F-\bar C^i} \,\right]' & \bar C^i < R \end{cases} \tag{8b}$$

with values

$$J^i_{\text{att}}{}^* = S_i - F \begin{cases} \frac{Rp^i_{\text{fp}}}{\bar C^i}, & \bar C^i \geqslant R \\ \frac{(R-C)(1-p^i_{\text{fp}})+Cp^i_{\text{fn}}}{C+F-\bar C^i}, & \bar C^i < R \end{cases} \tag{9a}$$

$$J^i_{\text{def}}{}^* = \begin{cases} \frac{ABp^i_{\text{fp}}}{Ap^i_{\text{fp}}+B(1-p^i_{\text{fn}})}, & \bar C^i \geqslant R \\ \frac{ABp^i_{\text{fn}}}{A(1-p^i_{\text{fp}})+Bp^i_{\text{fn}}}, & \bar C^i < R \end{cases} \tag{9b}$$

where $\bar C^i \coloneqq C(1-p^i_{\text{fn}}) + Fp^i_{\text{fp}}$ and all the probabilities p^i_{fp}, p^i_{fn} that appear above correspond to the set $\mathcal{Y}_i^{(1)}$. □

For any informative sensor, one should expect p^i_{fp} and p^i_{fn} to be no larger than 0.5, so (7) can be assumed without loss of generality.

We conclude from Theorem 1 that the attacker minimizes its cost by revealing the sensor i that leads to the smallest value of its cost $J^i_{\text{att}}{}^*$ in (9a). This provides the last piece of the policy for the attacker:

$$\sigma = \arg\min_i S_i - F \begin{cases} \frac{Rp^i_{\text{fp}}}{\bar C^i} & \bar C^i \geqslant R \\ \frac{(R-C)(1-p^i_{\text{fp}})+Cp^i_{\text{fn}}}{C+F-\bar C^i} & \bar C^i < R, \end{cases} \tag{10}$$

which corresponds to the top branch of the decision tree depicted in Figure 1. We thus conclude that the attacker's selection of the sensor index σ can be deterministic (pure), but the selection of θ will typically be mixed and given by the distribution in (8a) for $i = \sigma$. The defender's selection of Y_i will typically also be mixed and given by the distribution in (8b).

Remark 1 (Selection of $\mathcal{Y}_i^{(1)}$): In view of (9b), the defender will minimize its cost by selecting the sets $\mathcal{Y}_i^{(1)} \subset \mathcal{Y}$, $i \in \{1, 2, \ldots, N\}$ to minimize

$$\begin{cases} \frac{ABp^i_{\text{fp}}}{Ap^i_{\text{fp}}+B(1-p^i_{\text{fn}})}, & \bar C^i \geqslant R \\ \frac{ABp^i_{\text{fn}}}{A(1-p^i_{\text{fp}})+Bp^i_{\text{fn}}}, & \bar C^i < R, \end{cases} \tag{11}$$

which depends on $\mathcal{Y}_i^{(1)} \subset \mathcal{Y}$ through the parameters p^i_{fp} and p^i_{fn}. □

Remark 2 (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 only parameters that affect the Nash equilibrium are the reliability parameters p^i_{fp} and p^i_{fn}. For simple sensor models, these parameters can be computed analytically, but in realistic scenarios they may need to be learned from data. □

A. *Proof of Theorem 1*

As before, the probabilities p^i_{fp} and p^i_{fn} that appear below correspond to the set $\mathcal{Y}^{(1)}_i$. We consider separately two types of Nash equilibria, which will correspond to the branches $\bar{C}^i \geqslant R$ and $\bar{C}^i < R$ in the theorem's formulas.

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

$$y \coloneqq \begin{bmatrix}\eta & 1-\eta\end{bmatrix}', \qquad z \coloneqq \begin{bmatrix}\zeta & 0 & 1-\zeta\end{bmatrix}',$$

with $\eta, \zeta \in [0,1]$ and

$$A^i_{\text{att}} z = \begin{bmatrix} p \\ p \end{bmatrix}, \tag{12a}$$

$$y' B^i_{\text{def}} = \begin{bmatrix} q_1 & q_0 & q_1 \end{bmatrix}, \tag{12b}$$

$$q_0 \geqslant q_1, \tag{12c}$$

that would satisfy the usual quadratic program for mixed Nash equilibria for bi-matrix games [7, Chapter 10]. Equation (12a) leads to

$$\zeta = \frac{C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}} - R}{C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}}},$$

$$p = S_i - F \frac{R p^i_{\text{fp}}}{C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}}}$$

and equation (12b) to

$$\eta = \frac{B(1-p^i_{\text{fn}})}{A p^i_{\text{fp}} + B(1-p^i_{\text{fn}})}$$

$$q_1 = \frac{AB p^i_{\text{fp}}}{A p^i_{\text{fp}} + B(1-p^i_{\text{fn}})}$$

$$q_0 = q_1 + AB \frac{1 - p^i_{\text{fn}} - p^i_{\text{fp}}}{A p^i_{\text{fp}} + B(1-p^i_{\text{fn}})}.$$

So we have a Nash equilibrium as long as

$$\zeta \in [0,1] \quad \Leftrightarrow \quad C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}} \geqslant R, \tag{13a}$$

$$q_0 \geqslant q_1 \quad \Leftrightarrow \quad p^i_{\text{fn}} + p^i_{\text{fp}} \leqslant 1. \tag{13b}$$

Since the condition (13b) always holds by assumption (7), we have obtained a mixed Nash equilibrium that holds as long as $\bar{C}^i \coloneqq C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}} \geqslant R$.

2) We next consider mixed Nash equilibria of the form

$$y \coloneqq \begin{bmatrix}\eta & 1-\eta\end{bmatrix}', \qquad z \coloneqq \begin{bmatrix}0 & \zeta & 1-\zeta\end{bmatrix}',$$

with $\eta, \zeta \in [0,1]$ and

$$A^i_{\text{att}} z = \begin{bmatrix} p \\ p \end{bmatrix}, \tag{14a}$$

$$y' B^i_{\text{def}} = \begin{bmatrix} q_0 & q_1 & q_1 \end{bmatrix}, \tag{14b}$$

$$q_0 \geqslant q_1. \tag{14c}$$

Equation (14a) leads to

$$\zeta = \frac{R - C(1-p^i_{\text{fn}}) - F p^i_{\text{fp}}}{C p^i_{\text{fn}} + F(1-p^i_{\text{fp}})},$$

$$p = S_i - F \frac{(R-C)(1-p^i_{\text{fp}}) + C p^i_{\text{fn}}}{C p^i_{\text{fn}} + F(1-p^i_{\text{fp}})}$$

and equation (14b) to

$$\eta = \frac{B p^i_{\text{fn}}}{A(1-p^i_{\text{fp}}) + B p^i_{\text{fn}}}$$

$$q_1 = \frac{AB p^i_{\text{fn}}}{A(1-p^i_{\text{fp}}) + B p^i_{\text{fn}}}$$

$$q_0 = q_1 + AB \frac{1 - p^i_{\text{fn}} - p^i_{\text{fp}}}{A(1-p^i_{\text{fp}}) + B p^i_{\text{fn}}}.$$

So we now have a Nash equilibrium as long as

$$\zeta \in [0,1] \;\Leftrightarrow\; C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}} \leqslant R, \; R - C \leqslant F, \tag{15a}$$

$$q_0 \geqslant q_1 \;\Leftrightarrow\; p^i_{\text{fn}} + p^i_{\text{fp}} \leqslant 1. \tag{15b}$$

Since the right-most condition in (15a) always hold due to Assumption 1 and condition (15b) always holds by assumption (7), we have obtained a mixed Nash equilibrium that holds as long as $\bar{C}^i \coloneqq C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}} \leqslant R$. ∎

IV. Fictitious Play

In fictitious play, for each sensor i the defender constructs a running average $\bar{y}_i(t)$ of the mixed policy $y_i(t) \in \mathbb{S}^2$ used by the attacker when each sensor i is selected:

$$\bar{y}_i(t) = \frac{1}{t} \sum_{k=1}^{t} y_i(k) \in \mathbb{S}^2, \quad \forall \in i\{1,2,\dots,N\} \tag{16a}$$

and uses a mixed policy

$$z^\sigma(t) \in \beta^\sigma_{\text{def}}(\bar{y}^\sigma(t)) \coloneqq \arg\min_{z \in \mathbb{S}^{2+M_i}} \bar{y}^\sigma(t)' B^\sigma_{\text{def}} z, \tag{16b}$$

where $\mathbb{S}^\ell$ denotes the simplex of probability distributions in $\mathbb{R}^\ell$ and $\beta^i_{\text{def}}(\bar{y}_i)$ the set of the defender's best response against the attacker's policy $\bar{y}_i$. The inclusion in (16b) 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^i_{def} 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^i_{att}) are unknown to the defender.

(ii) If the attacker is using a constant policy $y_i(t) = y^\dagger$, $\forall t$, the defender's policy is guarantees to converge to the best response against $y^\dagger$. In particular, if the attacker is playing a Nash equilibrium policy, then the defenders' 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.,

$$\bar{z}_i(t) = \frac{1}{t} \sum_{k=1}^{t} z_i(k) \in \mathbb{S}^{2+M_i}, \quad \forall \in i\{1,2,\dots,N\} \tag{17a}$$

$$y^\sigma(t) \in \beta^\sigma_{\text{att}}(\bar{z}^\sigma(t)) \coloneqq \arg\min_{y\in\mathbb{S}^2} y' A^\sigma_{\text{att}} \bar{z}^\sigma(t), \tag{17b}$$

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

Theorem 2: Consider the general case $M_i \geqslant 1$ in (5) and assume that

$$\begin{aligned} &A \neq 0,\ B \neq 0,\ R \neq 0,\ R - C \neq F, \\ &C(1-p^i_{\text{fn}}) + F p^i_{\text{fp}} \neq R,\ p^i_{\text{fp}} \in (0,1),\ p^i_{\text{fn}} \in (0,1), \end{aligned} \tag{18}$$

then (16)–(17) converges to a Nash equilibrium of the bimatrix game (6) with $i = \sigma$. □

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} and all rows B_{def} have no repeated entries, which is guaranteed by (18). This property holds generically and excludes trivial solutions.

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 $y^\dagger = [0\ 1]'$ 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 $y^\dagger = [0\ 1]'$, 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 $y^\dagger = [0\ 1]'$. 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.,

$$\begin{aligned} \bar{y}_i(t+1) &= (1-\gamma)\bar{y}_i(t) + \gamma y(t), \\ \bar{z}(t+1) &= (1-\gamma)\bar{z}(t) + \gamma z(t), \gamma \in (0,1), \end{aligned}$$

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. Current and Future Work

Theorem 1 is restricted to the case of $M_i = 1$. However, it is possible to formulate a version of Theorem 1 for the general case $M_i \geqslant 1$, including the case of infinitely many sets. Preliminary results show that even for $M_i > 1$, the mixed Nash policy used to select the set Y_i will still only randomize among two sets in (5).

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 a 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^i_{att} and B^i_{def} in (6). 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

[1] URL https://globalfishingwatch.org.

[2] J. Baker. Transparent transhipping: detecting illegal fishing with satellite data. *Ship Technology*, 2018. URL https://www.ship-technology.com/features/global-fishing-watch/.

[3] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[4] U. Berger. Fictitious play in $2 \times n$ games. *J. of Economic Theory*, 120:139–154, 2005.

[5] P. Bernhard, A.-L. Colomb, and G. P. Papavassilopoulos. Rabbit and hunter game: Two discrete stochastic formulations. *Comput. Math. Applic.*, 13(1–3):205–225, 1987.

[6] S. Emmert. Will fishing vessels find ways to avoid detection via ais now that it's possible to track them so accurately? *Global Fishing Watch*, 2018. URL https://globalfishingwatch.org/faq-items/will-fishing-vessels-find-ways-to-avoid-detection-via-ais-now-that-its-possible-to-track-them-so-accurately/.

[7] J. P. Hespanha. *Noncooperative Game Theory: An Introduction for Engineers and Computer Scientists*. Princeton Press, Princeton, New Jersey, June 2017.

[8] Y. Ji, S. Lee, E. Downing, W. Wang, M. Fazzini, T. Kim, A. Orso, and W. Lee. Rain: Refinable attack investigation with on-demand inter-process information flow tracking. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 377–390. ACM, 2017.

[9] G. Kimeldorf. Duels: An overview. In M. Shubik, editor, *Mathematics of Conflict*, pages 55–72. North-Holland, Amsterdam, 1983.

[10] D. Koller, N. Megiddo, and B. von Stengel. Efficient computation of equilibria for extensive two-person

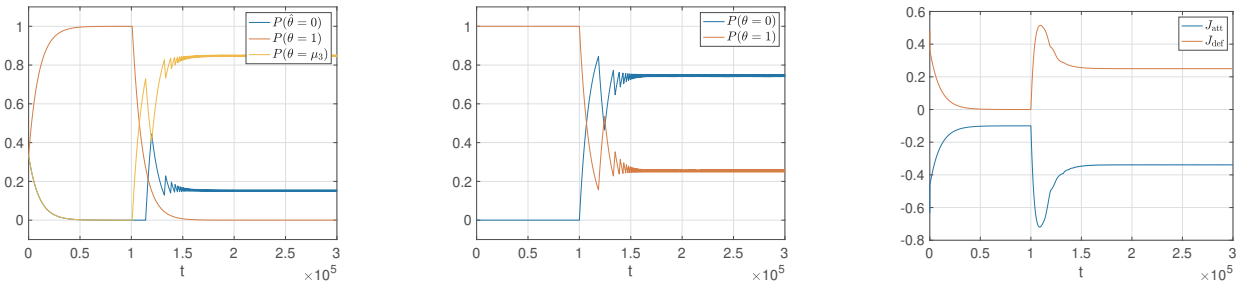

Fig. 2. Evolution of the defender's mixed policy (left), the attacker's mixed policy (middle), and both players costs (right) under fictitious play for the defender. Until time 10^5 the attacker uses a fixed policy $y^\dagger = [0\ 1]'$ and after that it also uses fictitious play. Throughout the simulation the attacker did not switch sensors and the following parameters values were used: $A = 1.5$, $B = 1$, $R = 1.1$, $C = 1$, $F = 2$, $S_i = 0$, $M_i = 1$, $p^i_{\text{fp}}(\mathcal{Y}_i^{(1)}) = 0.2$, $p^i_{\text{fn}}(\mathcal{Y}_i^{(1)}) = 0.1$.

games. *Games and Economic Behavior*, 14(2):247–259, 1996.

[11] C. Melolidakis. Stochastic games with lack of information on one side and positive stop probabilities. In Raghavan et al. [15], pages 113–126.

[12] A. Nayyar, A. Gupta, C. Langbort, and T. Başar. Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games. *IEEE Trans. on Automat. Contr.*, 59(3):555–570, Mar. 2013.

[13] G. J. Olsder and G. P. Papavassilopoulos. About when to use a searchlight. *J. Mathematical Anal. and Applications*, 136:466–478, 1988.

[14] G. J. Olsder and G. P. Papavassilopoulos. A Markov chain game with dynamic information. *J. Opt. Theory and Applications*, 59(3):467–486, Dec. 1988.

[15] T. E. S. Raghavan, T. S. Ferguson, and T. Parthasarathy, editors. *Stochastic Games and Related Topics: In Honor of Professor L. S. Shapley*, volume 7 of *Theory and Decision Library, Series C, Game Theory, Mathematical Programming and Operations Research*. Kluwer Academic, Dordrecht, 1991.

[16] L. S. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, pages 1–29. Princeton University Press, 1964.

[17] S. Sorin and S. Zamir. "Big Match" with lack of information on one side (III). In Raghavan et al. [15], pages 101–112.

[18] K. G. Vamvoudakis, J. P. Hespanha, B. Sinopoli, and Y. Mo. Detection in adversarial environments. *IEEE Trans. on Automatic Control,* Special Issue on the Control of Cyber-Physical Systems, 59(12):3209–3223, Dec. 2014.

[19] M. Wakaiki, P. Tabuada, and J. P. Hespanha. Supervisory control of discrete-event systems under attacks. *Dynamic Games and Applications,* Special Issue on Dynamic Games in Cyber Security, 9(4):965–983, Dec. 2019. To appear.