Chapter 8

Sensor Manipulation Games in Cyber Security

João P. Hespanha*

Center for Control Dynamical-Systems, and Computation, University of California, Santa Barbara, USA

*hespanha@ece.ucsb.edu

Abstract: This chapter addresses the problem of making decisions based on sensor measurements that may have been manipulated by an adversary. For concreteness, we focus our attention on making binary decisions that, in the context of cyber security, could correspond to denying access to a sensitive resource, flagging a computer as compromised, deauthorizing a user, closing a firewall, etc. The chapter presents a game theoretical treatment of sensor manipulation and considers two types of sensor manipulation: In *measurement manipulation games*, the attacker is able to manipulate the measurements of M out of N sensors available to the defender, but the latter does not know which sensors have been manipulated. In *sensor reveal games*, the attacker exposes to the defender the measurement of a single sensor out of N sensor possibilities, with the caveat that revealing data from non-informative sensors may be costly and interpreted by the defender as strong indication that an attack is afoot. These games cover different aspects of sensor manipulation: Measurement manipulation games

mostly explore the challenges that arise from not knowing which sensors have been manipulated and how this impairs the defender's ability to make correct decisions. In sensor reveal games the defender actually knows which sensor has been revealed, but the challenge is to explore the information that the attacker implicitly provides by selecting a particular sensor.

Keywords: Cyber Security, Noncooperative Games, Partial Information, Sensor Manipulation

8.1. Introduction

This chapter addresses the problem of making decisions based on sensor measurements that may have been manipulated by an adversary. For concreteness, we focus our attention on making a binary decision that, in the context of cyber security, could correspond to denying access to a sensitive resource, flagging a computer as compromised, deauthorizing a user, closing a firewall, etc. Such decisions are typically based on measurements collected by cyber security sensors that analyze records of events (logs) and provide recommendations on what the binary decision should be. Sophisticated cyber security systems typically rely on multiple such sensors to maximize their ability to catch attacks, while maintaining a small probability of false alarms. The use of multiple cyber sensors also provides protection against adversaries that may have compromised some of these sensors, either by disabling a sensor or actively manipulating its output.

Noncooperative game theory provides a mathematical framework to reason about decision making by a group of agents (players) in which the decision of one agent affects the costs/rewards incurred by the other agents. To apply ideas from game theory to cyber security, we regard the cyber defense system

as one player and the attacker as the other player. In the simplest form of the game, the defender wants to minimize the probability of making the wrong decision, while the attacker wants to maximize this probability. However, we shall see below that the goals of the players may be more complex. While the defender's decision is assumed to be binary, we consider a much richer set of choices for the attacker, as its decision space involves selecting if and how to manipulate sensor measurements.

We shall consider two types of sensor manipulation games that differ on the type of sensor manipulation available to the attacker. In *measurement manipulation games*, the attacker is able to manipulate the measurements of M out of N sensors available to the defender, but the latter does not know which sensors have been manipulated. In *sensor reveal games*, the attacker exposes to the defender the measurement of a single sensor out of N sensor possibilities, with the caveat that revealing data from non-informative sensors may be costly and interpreted by the defender as strong indication that an attack is afoot. These games cover different aspects of sensor manipulation: Measurement manipulation games mostly explore the challenges that arise from not knowing which sensors have been manipulated and how this impairs the defender's ability to make correct decisions. In sensor reveal games the defender actually knows which sensor has been revealed, but the challenge is to explore the information that the attacker implicitly provides by selecting a particular sensor. Several combinations of these games are possible (and relevant).

The remainder of this chapter discusses these two types of games and is mostly based on results from [Vamvoudakis et al., 2014] for measurement manipulation games and from [Hespanha and Garagic, 2019] for sensor reveal games. The goal of this chapter is to provide a tutorial view on the use of

game theory to model and solve sensor manipulations problems that arise in the cyber security domain. We refer the reader to the following two recent surveys for excellent literature reviews on the application of game theory to the cyber security domain and to the detection of cyber attacks [Do et al., 2017, Giraldo et al., 2018]. Chapter 7 of this book also considers the detection of cyber attacks [Kamhoua et al., 2020, Chapter 7], but the focus there in the detection of multi-step dynamic attacks, rather than sensor manipulation.

8.2. Measurement Manipulation Games

One can model making a binary decision as estimating the value of a Bernoulli random variable $\theta \in \{0, 1\}$, with the understanding that if the defender's estimate $\hat{\theta}$ is equal to θ then the "right" decision has been made. The Bernoulli parameter

$$\mathrm{P}(\theta = 1) = 1 - \mathrm{P}(\theta = 0) = p \tag{8.1}$$

expresses any a-priori information that we may have on θ. To make its decision, the defender has available a vector $Y \coloneqq (Y_1, Y_2, \ldots, Y_N)$ of N binary "noisy" sensor measurements that provide recommendations on how $\hat{\theta}$ should be selected. We assume that the measurements $Y_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, N\}$ are conditionally independent, given θ. Specifically, $Y_i = 1$ means that sensor i recommends selecting $\hat{\theta} = 1$, with the understanding that

$$\mathrm{P}(Y_i = 1 | \theta; Y_{j \neq i}) = \begin{cases} p_{\mathrm{err}} & \theta = 0 \\ 1 - p_{\mathrm{err}} & \theta = 1, \end{cases} \qquad \forall i \in \{1, 2, \ldots, N\}, \tag{8.2}$$

where p_{err} denotes the sensor's error probability that, for simplicity, is assumed the same for every sensor.

Measurement manipulation arises because the defender must build its estimate $\hat{\theta}$ of θ based on a version $Z := (Z_1, Z_2, \ldots, Z_N)$ of the measurement vector Y that may have been "corrupted" by an attacker. Specifically, with a probability $p_{\text{att}} \in [0, 1]$, the attacker manipulated the values of $M \leq N$ entries of Y and therefore only $N - M$ of the entries of Z are guaranteed to match those of Y, but the defender does not know which. The probability p_{att} should be viewed as a design parameter that reflects how certain the estimator is that the measurements have been manipulated. For $p_{\text{att}} = 0$, there is no attack and the estimation of θ is a standard Bayesian estimation problem subject to stochastic measurement errors. However, for $p_{\text{att}} > 0$ the solution to this problem requires game theoretical tools.

Our problem can be viewed as a two-player partial information game: The *defender* must select an estimate $\hat{\theta}$ for θ based on the vector Z of possibly corrupted sensor measurements. Since the defender does not know which measurements have been manipulated and all sensors are assumed equally reliable, the defender's decision must be based solely on the total number of zeros and ones that appear in the N-vector Z. This means that the defender's policy is a function μ that maps integers in $\{0, 1, \ldots, N\}$ into the binary estimate:

$$\hat{\theta} = \mu\Big(\sum_{i=1}^{N} Z_i\Big). \tag{8.3}$$

Since the domain of μ has $N + 1$ elements and its codomain has 2 elements, the set $\mathcal{U}$ of all possible estimation policies contains 2^{N+1} policies.

The *attacker* is able to manipulate M out of the N entries of Z and must

thus decide how many of these M measurements should be set to report a zero. We assume that the attacker knows the value of θ but not the measurements Y_i that it cannot manipulate. The attack policy δ is thus a function that maps the binary variable $\theta \in \{0, 1\}$ to the integer set $\{0, 1, \ldots, M\}$. Since the domain of δ has 2 elements and its codomain has $M+1$ elements, the set $\mathcal{D}$ of all possible attack policies contains $(M + 1)^2$ policies.

In measurement manipulation games, we consider a simple zero-sum formulation in which the defender wants to minimize the probability of making the wrong decision

$$J_{\text{def}} \coloneqq \mathrm{P}(\hat{\theta} \neq \theta), \tag{8.4}$$

whereas the attacker wants to maximize it.

8.2.1. Saddle-Point Equilibria

The model described above for measurement manipulation defines a zero-sum noncooperative game in which the detector selects a policy $\mu \in \mathcal{U}$ and the attacker a policy $\delta \in \mathcal{D}$ to minimize and maximize, respectively, the probability of an estimation error (8.4) [Kamhoua et al., 2020, Chapter 2]. Since the sets of policies are finite, we have a (finite) matrix game defined by a 2^{n+1} by $(M + 1)^2$ matrix A, with each row corresponding to an estimation policy μ for the defender and each column to an attack policy δ. Straightforward computations show that the entry of A corresponding to a defender policy μ and an attack policy δ that sets to 0 and to 1 a number of sensors equal to $\delta(\theta)$ and $M - \delta(\theta)$,

respectively, is given by

$$
\begin{aligned}
&\mathrm{P}(\hat{\theta} \neq \theta) = \\
&(1-p)\Bigg(p_{\mathrm{att}} \sum_{k=M-\delta(0)}^{N-\delta(0)} \mu(k) \binom{N-M}{k-M+\delta(0)} p_{\mathrm{err}}^{k-M+\delta(0)} (1-p_{\mathrm{err}})^{N-k-\delta(0)} \\
&\qquad\qquad + (1-p_{\mathrm{att}}) \sum_{k=0}^{N} \mu(k) \binom{n}{k} p_{\mathrm{err}}^{k} (1-p_{\mathrm{err}})^{N-k} \Bigg) \\
&\quad + p\Bigg(p_{\mathrm{att}} \sum_{k=M-\delta(1)}^{N-\delta(1)} \big(1-\mu(k)\big) \binom{N-M}{k-M+\delta(1)} (1-p_{\mathrm{err}})^{k-M+\delta(1)} p_{\mathrm{err}}^{N-k-\delta(1)} \\
&\qquad\qquad + (1-p_{\mathrm{att}}) \sum_{k=0}^{N} \big(1-\mu(k)\big) \binom{n}{k} (1-p_{\mathrm{err}})^{k} p_{\mathrm{err}}^{N-k} \Bigg). \qquad (8.5)
\end{aligned}
$$

Often, this game does not have pure saddle-point equilibria and the players will need to seek for mixed policies, which correspond to selecting probability distributions over the sets of actions $\mathcal{U}$ and $\mathcal{D}$ [Başar and Olsder, 1995, Hespanha, 2017]. While the number of rows of the matrix A defined above grows exponentially with the number of sensors, it turns out that a large number of rows can be ignored using policy domination and this game has mixed saddle-point policies that involve randomization only over a small number of pure policies that we define below:

1. The attacker's *deception rule* is the pure policy

$$
\delta(\theta) = \delta_{\mathrm{dec}}(\theta) \coloneqq \begin{cases} 0 & \theta = 0, \\ M & \theta = 1, \end{cases}
$$

   which sets all M manipulated sensors equal to 1 when $\theta = 0$ and all M sensors equal to 0 when $\theta = 1$.

2. The attacker's *no-deception rule* is the pure policy

$$\delta(\theta) = \delta_{\mathrm{no-dec}}(\theta) \coloneqq \begin{cases} M & \theta = 0, \\ 0 & \theta = 1, \end{cases}$$

which sets all M manipulated sensors equal to θ.

3. The detector's *majority rule* is the pure policy

$$\mu\Big(\sum_{i=1}^{N} Z_i\Big) = \mu_{\mathrm{maj}}\Big(\sum_{i=1}^{N} Z_i\Big) \coloneqq \begin{cases} 0 & \sum_{i=1}^{N} Z_i \leq \frac{N-1}{2}, \\ 1 & \sum_{i=1}^{N} Z_i \geq \frac{N+1}{2}, \end{cases}$$

which corresponds to setting $\hat{\theta} = 0$ if more than half the sensors reported the value 0.

4. The detector's *no-consensus rule* is the pure policy

$$\begin{aligned} \mu\Big(\sum_{i=1}^{N} Z_i\Big) &= \mu_{\mathrm{no-cons}}\Big(\sum_{i=1}^{N} Z_i\Big) \\ &\coloneqq \begin{cases} 0 & 0 < \sum_{i=1}^{N} Z_i \leq \frac{N-1}{2} \text{ or } \sum_{i=1}^{N} Z_i = N, \\ 1 & n > \sum_{i=1}^{N} Z_i \geq \frac{N+1}{2} \text{ or } \sum_{i=1}^{N} Z_i = 0. \end{cases} \end{aligned}$$

This somewhat unexpected policy is like the majority rule, except that if all sensors agree on a particular value (i.e., $Z_i = 1$, $\forall i$ or $Z_i = 0$, $\forall i$), the estimate $\hat{\theta}$ should take the opposite value.

The deception rule seems intuitive, since the attacker should try to deceive the defender by reporting the opposite of θ. Similarly, the majority rule also seems reasonable, at least when the attacker can manipulate less than half of the sensors. To gain intuition on why the remaining policies need to be

considered, suppose that the estimator observes that all sensors report the same value: Three options are possible: 1) there was no attack and all sensors are correct, 2) the attacker did not flip any bit and all sensors are correct, or 3) the attacker flipped M bits and the remaining $N - M$ sensors reported incorrect values. In cases 1) and 2) the defender should use the majority rule, but in case 3) the estimator should choose precisely what we called the non-consensus rule. At a saddle point, the attacker will choose the probabilities of using the deception versus no-deception rules to keep the estimator guessing. Because of this, it is not surprising to discover that the optimal estimator's policy will sometimes select the majority rule (typically with high probability) and sometimes the no-consensus rule (typically with low probability).

The following result provides an explicit formula for a mixed saddle-point equilibrium for this game.

Theorem 8.1: *Consider* $p = 1/2$ *in* (8.5) *and an odd number of sensors* $N \geq 3$*, for which*

$$M \leq \min\Big\{\frac{N-1}{2}, \frac{N+1}{2} - \frac{p_{\rm err}}{1-p_{\rm err}}\frac{N-1}{2}\Big\}, \tag{8.6}$$

$$p_{\rm err} \leq \frac{2}{N+1}, \qquad p_{\rm err} \leq \Big(1 + \frac{\frac{N-1}{2}!(N-M)!}{\frac{N-2M+1}{2}!}\Big)^{-1}, \tag{8.7}$$

and, for the case $M \geq 2$*, further assume that*

$$p_{\rm att} \leq \frac{1}{1 + \frac{1}{n}\binom{N-M}{M-1}\frac{p_{\rm err}^{N-2M+1}(1-p_{\rm err})^{M-1}}{p_{\rm err}(1-p_{\rm err})^{N-1} - p_{\rm err}^{N-1}(1-p_{\rm err})}}. \tag{8.8}$$

In this case, the value of the game is given by

$$v^* = \alpha + p_{\text{att}} \min\left\{\gamma, \gamma + \frac{\eta(\beta - p_{\text{err}}^{N-M})}{(1-p_{\text{err}})^{N-M} + p_{\text{err}}^{N-M}}\right\} \tag{8.9}$$

and a mixed saddle-point policy corresponds to selecting

$$\begin{cases} \mu_{\text{maj}} & w.p.\ 1 - p_{\text{no-cons}}, \\ \mu_{\text{no-cons}} & w.p.\ p_{\text{no-cons}}, \end{cases} \qquad \begin{cases} \delta_{\text{dec}} & w.p.\ 1 - p_{\text{no-dec}}, \\ \delta_{\text{no-dec}} & w.p.\ p_{\text{no-dec}}, \end{cases} \tag{8.10}$$

where

$$p_{\text{no-cons}} \coloneqq \begin{cases} \Pi\Big(\frac{\eta}{(1-p_{\text{err}})^{N-M} + p_{\text{err}}^{N-M}}\Big) & p_{\text{att}} \geq \frac{(1-p_{\text{err}})^N - p_{\text{err}}^N}{p_{\text{err}}^{N-M} + (1-p_{\text{err}})^N - p_{\text{err}}^N}, \\ 0 & p_{\text{att}} < \frac{(1-p_{\text{err}})^N - p_{\text{err}}^N}{p_{\text{err}}^{N-M} + (1-p_{\text{err}})^N - p_{\text{err}}^N}, \end{cases}$$

$$p_{\text{no-dec}} \coloneqq \Pi\left(\frac{p_{\text{err}}^{N-M} - \beta}{(1-p_{\text{err}})^{N-M} + p_{\text{err}}^{N-M}}\right),$$

$$\alpha \coloneqq (1 - p_{\text{att}}) \sum_{k=0}^{\frac{N-1}{2}} \binom{n}{k} p_{\text{err}}^{N-k} (1-p_{\text{err}})^k,$$

$$\beta \coloneqq \frac{1 - p_{\text{att}}}{p_{\text{att}}} \big((1-p_{\text{err}})^N - p_{\text{err}}^N\big),$$

$$\gamma \coloneqq \sum_{k=0}^{\frac{N-1}{2}} \binom{N-M}{k} p_{\text{err}}^{N-M-k} (1-p_{\text{err}})^k,$$

$$\eta \coloneqq \sum_{k=\frac{N+1}{2}-M}^{\frac{N-1}{2}} \binom{N-M}{k} p_{\text{err}}^{N-M-k} (1-p_{\text{err}})^k,$$

and $\Pi : \mathbb{R} \to \mathbb{R}$ *denotes the projection function into the interval* $[0, 1]$*:*

$$\Pi(x) = \begin{cases} 0 & x < 0, \\ x & x \in [0, 1], \\ 1 & x > 1. \end{cases} \qquad \square$$

An important (and convenient) feature of the mixed saddle-point equilibrium in (8.10) is that the defender does not need to know the precise value of the attack probability p_{att} to implement its policy. This is because, to compute $p_{\text{no-cons}}$, the defender only needs to know whether or not

$$p_{\text{att}} \geq \frac{(1 - p_{\text{err}})^N - p_{\text{err}}^N}{p_{\text{err}}^{N-M} + (1 - p_{\text{err}})^N - p_{\text{err}}^N}, \tag{8.11}$$

which is convenient because, in real applications, the value of p_{att} may be very hard to know. It is also worth noting that, when the error probability p_{err} is small, the right-hand side of (8.11) is close to 1 and therefore $p_{\text{no-cons}}$ is only nonzero for values of p_{att} very close to one. Otherwise, $p_{\text{no-cons}} = 0$ and the majority rule is always used.

While the detector's policy may depend little on p_{att}, that is not the case for the saddle-point probability of an estimation error v^*. In fact, one can show that this probability scales with the number of sensors as

$$p_{\text{err}}^{\frac{N+1-2M}{2}},$$

which shows that each one of the M potentially compromised sensors effectively decreases the total number of useful sensors from N to $N - 2M$.

Theorem 8.1 is only valid for $p = 1/2$ and parameters M and N that satisfy

(8.6)–(8.8). The restriction to the case $p = 1/2$ mostly simplifies the formulas and could be easily lifted. The assumption in (8.6) poses an upper bound on the maximum number of sensors that can be attacked and is not surprising since, even without sensor errors (i.e., $p_{\text{err}} = 0$), the detector can only make use of the measurements if less than half of the sensors have been attacked. The inequality (8.8) is also not restrictive since the left-hand side is typically very close to one. However, the inequalities in (8.7) are restrictive when N is large, because the right-hand side of both inequalities converges to zero as N goes to infinity. This limitation of Theorem 8.1 is the main motivation for the results in the next section.

Before proceedings it is worth noting that, in general, the mixed saddle-point in (8.10) is not unique and there will exist other mixed saddle-point involving different sets of pure policies. However, all other mixed saddle-point will have the value in (8.9), as all mixed saddle-point of a zero-sum game must lead to the same value. Moreover, because of the order interchangeability property of zero-sum games [Hespanha, 2017], the mixed defender's policy in (8.10) remains a security policy regardless of whether or not the attackers restricts theirs attention to the deception deception/no-deception policies considered above.

8.2.2. Approximate Saddle-Point Equilibrium

In view of (8.7), the estimation policy provided by Theorem 8.1 is only optimal for a number of sensors N roughly below $2/p_{\text{err}}$. We shall see in this section that when the number of sensors is large and the probability of sensor error is not very small, a threshold estimation policy like the majority rule is al-

most optimal. However, the proof of this result requires a completely different approach, which is described next.

The random variable

$$\bar{Z} \coloneqq \sum_{i=1}^{N} Z_i,$$

that the detector uses in the estimation policy (8.3) may have different distributions depending on the value of θ and whether or not there is an attack. Specifically,

$$\bar{Z} = \begin{cases} R & \text{w.p. } 1 - p_{\text{att}}, \\ S + W & \text{w.p. } p_{\text{att}}, \end{cases} \tag{8.12}$$

where the random variable R equals the sum of all the Y_i and therefore its distribution is

$$R \sim \begin{cases} \text{Binom}(N, p_{\text{err}}) & \theta = 0, \\ \text{Binom}(N, 1 - p_{\text{err}}) & \theta = 1; \end{cases} \tag{8.13}$$

the random variable S equals the sum of the $N - M$ sensors that have not been compromised by the attacker and therefore has distribution

$$S \sim \begin{cases} \text{Binom}(N - M, p_{\text{err}}) & \theta = 0, \\ \text{Binom}(N - M, 1 - p_{\text{err}}) & \theta = 1; \end{cases} \tag{8.14}$$

and the random variable W equals the sum of the readings of the M sensors compromised by the attacker. The distribution of W is selected by the attacker and may depend on the value of θ, with the constraint that its support must

lie in the set $\{0, 1, \ldots, M\}$.

This perspective motivates the following general problem: Suppose that one wants to estimate the random variable θ with Bernoulli distribution (8.1) based on a measurement $\bar{Z}$ of the form (8.12) where the conditional distributions of R and S given θ are known and the conditional distribution of W given θ is selected by an adversary, but its support is limited to a given subset $I \subset \mathbb{R}$. As before, we formulate this as a zero-sum game where the estimator wants to minimize the probability of an estimation error, whereas the attacker wants to maximize this probability.

We allow the *estimation policy* to be stochastic and represent it by a function $p_{\hat{\theta}=1} : \mathbb{R} \to [0, 1]$, with the understanding that, when the estimator observes a value $\bar{z} \in \mathbb{R}$ for (8.12), the estimate of θ is given by

$$\hat{\theta} = \begin{cases} 1 & \text{w.p. } p_{\hat{\theta}=1}(\bar{z}), \\ 0 & \text{w.p. } 1 - p_{\hat{\theta}=1}(\bar{z}). \end{cases}$$

Denoting by ρ_x, σ_x, and ω_x, respectively, the conditional distributions of R, S, and W given that $\theta = x \in \{0, 1\}$, we can use the law of total probability, to express the probability of an estimation error as follows:

$$\begin{aligned} J(p_{\hat{\theta}=1}, \omega_0, \omega_1) \coloneqq \mathrm{P}(\hat{\theta} \neq \theta) = p + p_{\mathrm{att}}(1-p) \int_I \int_{\mathbb{R}} p_{\hat{\theta}=1}(\bar{y} + \bar{w}) \sigma_0(d\bar{y}) \omega_0(d\bar{w}) \\ - p_{\mathrm{att}}\, p \int_I \int_{\mathbb{R}} p_{\hat{\theta}=1}(\bar{y} + \bar{w}) \sigma_1(d\bar{y}) \omega_1(d\bar{w}) \\ + (1 - p_{\mathrm{att}})(1-p) \int_{\mathbb{R}} p_{\hat{\theta}=1}(\bar{y}) \rho_0(d\bar{y}) - (1 - p_{\mathrm{att}}) p \int_{\mathbb{R}} p_{\hat{\theta}=1}(\bar{y}) \rho_1(d\bar{y}). \quad (8.15) \end{aligned}$$

For the above formula to be well defined, we assume that the attacker is only allowed to select distributions (ω_0, ω_1) in a set $\mathcal{A}$ containing all pairs of distri-

butions (ω_0, ω_1) for which the integrals in (8.15) exist.

The result that follows is based on first considering an auxiliary game that replaces the binomial distributions defined by (8.13)–(8.14) by Gaussian distributions and then using the mixed saddle-point for this "relaxed" version of the game to compute ϵ-saddle-point of our original game. We recall that a pair $(u^*, d^*) \in \mathcal{U} \times \mathcal{D}$ is an ϵ-saddle-point with respect to a criterion $J : \mathcal{U} \times \mathcal{D} \to \mathbb{R}$ when

$$J(u^*, d^*) - \epsilon \leq J(u^*, d^*) \leq J(u, d^*) + \epsilon, \quad \forall u \in \mathcal{U}, d \in \mathcal{D}.$$

For $\epsilon = 0$, an ϵ-saddle-point is just a regular saddle-point.

Theorem 8.2: *Suppose that*

$$M < N(1 - 2p_{\text{err}}), \quad p_{\text{err}} \in (0, 1/2),$$

and that the function

$$g(\bar{z}) \coloneqq \frac{1 - p_{\text{att}}}{\sqrt{np_{\text{err}}(1 - p_{\text{err}})2\pi}} \Big((1-p) e^{-\frac{(\bar{z} - np_{\text{err}})^2}{2np_{\text{err}}(1-p_{\text{err}})}} - p\, e^{-\frac{(\bar{z} - n(1-p_{\text{err}}))^2}{2np_{\text{err}}(1-p_{\text{err}})}} \Big)$$
$$+ \frac{p_{\text{att}}}{\sqrt{(N-M)p_{\text{err}}(1 - p_{\text{err}})2\pi}} \Big((1-p) e^{-\frac{(\bar{z} - M - (N-M)p_{\text{err}})^2}{2(N-M)p_{\text{err}}(1-p_{\text{err}})}} - p\, e^{-\frac{(\bar{z} - (N-M)(1-p_{\text{err}}))^2}{2(N-M)p_{\text{err}}(1-p_{\text{err}})}} \Big),$$

has a unique zero $\bar{z} = z^*$. Then the function

$$p^*_{\hat{\theta}=1}(\bar{z}) \coloneqq \begin{cases} 0 & \bar{z} < \bar{z}^*, \\ 1 & \bar{z} \geq \bar{z}^*, \end{cases} \tag{8.16}$$

for the defender and the Dirac distributions

$$\omega_0^*(\bar{w}) = \boldsymbol{\delta}(\bar{w} - M), \qquad \omega_1^*(\bar{w}) = \boldsymbol{\delta}(\bar{w}), \tag{8.17}$$

for the attacker form a 2ϵ*-saddle-point with value* $J(f^*, \omega_0^*, \omega_1^*)$ *satisfying*

$$\Big| J(f^*, \omega_0^*, \omega_1^*) - p - \int_{z^*}^{\infty} g(\bar{z}) d\bar{z} \Big| \leq \epsilon = O\Big(\frac{1}{\sqrt{N-M}}\Big). \qquad \square$$

While the problem formulation enables both players to select mixed policies, the 2ϵ-saddle-point defined by (8.16)–(8.17) actually corresponds to pure policies of the form

$$\hat{\theta} = \begin{cases} 0 & \bar{Z} := \sum_{i=1}^{N} Z_i < z^*, \\ 1 & \bar{Z} := \sum_{i=1}^{N} Z_i \geq z^*, \end{cases} \qquad \delta(\theta) = \delta_{\text{dec}}(\theta) := \begin{cases} 0 & \theta = 0, \\ M & \theta = 1 \end{cases}$$

and when $p = 1/2$ and there are no sensor errors ($p_{\text{err}} = 0$), the threshold value z^* in Theorem 8.2 is precisely equal to $N/2$ for every attack probability and we obtain the majority rule μ_{maj}. As the a-priori probability p deviates from $1/2$ and p_{err} increases, the threshold z^* changes to reflect the a-priori information and the probability of sensor error.

8.3. Sensor Reveal Games

In sensor reveal games, we continue to consider a scenario where the defender wants to estimate an unknown variable θ based on a vector $Y := (Y_1, Y_2, \ldots, Y_N)$ of N binary "noisy" sensor measurements that provide recommendations for $\hat{\theta}$. However, now the N sensors are heterogeneous with the

model (8.2) replaced by

$$\mathrm{P}(Y_i = 1|\theta = 0; Y_{j\neq i}) = p^i_{\mathrm{fp}}, \qquad \mathrm{P}(Y_i = 0|\theta = 1; Y_{j\neq i}) = p^i_{\mathrm{fn}},$$

$\forall i \in \{1, 2, \ldots, N\}$, where p^i_{fp} and p^i_{fn} can be regarded as the probabilities of a false positive and a false negative associated with the ith sensor. This model is more general than the one in (8.2) because we now allow the two types of errors (false positives and false negatives) to have different probabilities and also allow different sensors to have different error probabilities (thus the superscript ^{i} in p^i_{fp} and p^i_{fn}).

A more fundamental difference with respect to Section 8.2 is that the defender's decision $\hat{\theta}$ is now based on a single sensor measurements Y_σ revealed by the attacker, where $\sigma \in \{1, 2, \ldots, N\}$ denotes the index of the sensor that the attacker selects to reveal, which the defender can observe and use to inform its decision. Another difference is that we now enable the attacker to select the actual value of the unknown variable θ. This is reasonable when $\theta = 1$ corresponds to some form of cyber attack launched by the same agent that manipulates the sensors.

A sensor reveal game can also be viewed as a two-player partial information game: The *defender* selects an estimate $\hat{\theta}$ for θ based on the pair (σ, Y_σ) that includes the index σ of the sensor revealed by the attacker and the corresponding sensor recommendation Y_σ. This means that deterministic (pure) policies for the defender are functions μ that map the pair $(\sigma, Y_\sigma) \in \{0, 1\} \times \{1, 2, \ldots, N\}$ into the binary estimate:

$$\hat{\theta} = \mu(\sigma, Y_\sigma). \tag{8.18}$$

The *attacker* makes two decisions: it decides whether or not to launch an attack (i.e., selects the value of $\theta \in \{0, 1\}$) and also the index σ of which sensor to reveal. In the model considered here, the attacker can decide which sensor to reveal (i.e., can select the sensor index σ), but does not know which measurements Y_σ will be reported to the defender. In particular, the attacker does not know a-priori whether or not the sensor recommendation Y_σ is correct in the sense that it is equal to θ.

We also consider a more sophisticated cost structure for the players: The defender seeks to minimize a cost of the form

$$J_{\text{def}} \coloneqq A\,\mathrm{P}(\hat{\theta} = 1, \theta = 0) + B\,\mathrm{P}(\hat{\theta} = 0, \theta = 1), \tag{8.19}$$

where $A > 0$ and $B > 0$ are parameters that establish the cost of a false detection and of a missed detection, respectively. When $A = B$, the defender's cost (8.19) matches the measurement manipulation game cost in (8.4), but in sensor manipulation games we go beyond zero-sum games and consider a cost for the attacker that includes terms related to the benefit reaped from launching an attack and penalties for revealing different sensors. Specifically, we assume that the attacker minimizes a cost of the form

$$J_{\text{att}} \coloneqq -R\theta + C\,\mathrm{P}(\hat{\theta} = 1, \theta = 1) - F\,\mathrm{P}(\hat{\theta} = 1, \theta = 0) + S_\sigma, \tag{8.20}$$

where $R \geq 0$ is a reward associated with engaging in the cyber attack (i.e., choosing $\theta = 1$), $C \geq 0$ is the cost of being caught, S_σ the cost of revealing sensor σ, and $F \geq 0$ is a rewards to the attacker for generating a false alarm. In what follows, we assume that 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.,

$$R - C < F, \tag{8.21}$$

which excludes the trivial solution for the attacker to always select $\theta = 1$.

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., sensors with large values for p^i_{fp} and p^i_{fn}). Here, "costly" may mean that S_i is large but also that revealing the recommendation of a sensor that is not very informative can be taken as an indication that $\theta = 1$.

Consistent with standard terminology, a *pure Nash equilibrium* for this game is thus a (pure) defender's policy μ^* and a (deterministic) choice (θ^*, σ^*) for the attacker such that

1. when the defender uses the decision rule (8.18) with $\mu = \mu^*$, the attacker's cost J_{att} in (8.20) 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 (8.19) is minimized by using $\mu = \mu^*$ in (8.18).

A *mixed Nash equilibrium* follows a similar definition, but with the deterministic choices replaced by distributions over the sets of all deterministic policies. Specifically, a mix Nash equilibrium consists of a probability distribution over all possible policies $\mu : \{0, 1\} \times \{1, 2, \ldots, N\} \to \{0, 1\}$ 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 is no need to randomize over the

sensor selection σ, just over μ and θ.

8.3.1. Nash Equilibria

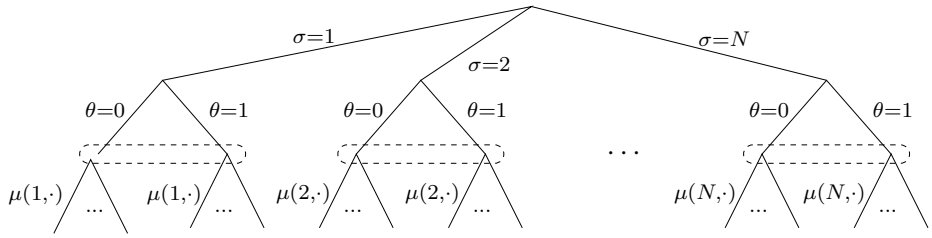

Figure 8.1: Extensive form representation of the sensor reveal game, with the top two branches corresponding to attacker decisions (which sensor σ to select and whether or not to set $\theta = 1$) and the bottom branch to the defender's selection of the maps $Y_i \mapsto \mu(i, Y_i)$, $i \in \{1, 2, \ldots, N\}$ that define the estimate in (8.18).

For the purpose of computing Nash equilibria for this game, it is convenient to consider the extensive form decision tree depicted in Figure 8.1, where the branches represent the players' 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 [Kamhoua et al., 2020, Chapter 2], [Hespanha, 2017]. 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 map $Y_i \mapsto \mu(i, Y_i)$ that defines the estimate

$\hat{\theta}$ in (8.18) as a function of Y_i. Since each Y_i can only take two values $\{0, 1\}$ and $\mu(i, Y_i)$ also only takes two values, there are three possible choices for the map $Y_i \mapsto \mu(i, Y_i)$:

$$\hat{\theta} = \mu(i, Y_i) \coloneqq 0, \quad \forall Y_i \in \{0, 1\}, \text{ or}$$
$$\hat{\theta} = \mu(i, Y_i) \coloneqq 1, \quad \forall Y_i \in \{0, 1\}, \text{ or}$$
$$\hat{\theta} = \mu(i, Y_i) \coloneqq Y_i, \quad \forall Y_i \in \{0, 1\}.$$

The first two options correspond to ignoring the sensor recommendation Y_i and always setting $\hat{\theta} = 0$ or $\hat{\theta} = 1$, respectively, whereas the third option corresponds to always agreeing with the sensor recommendation and setting $\hat{\theta} = Y_i$. Technically, there is a fourth option of always disagreeing with the sensor recommendation and setting $\hat{\theta} = 1 - Y_i$, but this policy would always be dominated by one of the other three, as long as the sensor is somewhat "informative" in the sense that its error probabilities satisfy

$$p^i_{\text{fp}} + p^i_{\text{fn}} \leq 1. \tag{8.22}$$

Straightforward computations can be used to show that the problem corresponding to the subtree $\sigma = i$ in Figure 8.1 can be represented by the following

2×3 bi-matrix game

$$A^i_{\text{att}} \coloneqq \begin{bmatrix} & \hat{\theta}{=}0 & \hat{\theta}{=}1 & \hat{\theta}{=}Y_i \\ \theta{=}0 & S_i & S_i - F & S_i - Fp^i_{\text{fp}} \\ \theta{=}1 & S_i - R & S_i - R + C & S_i - R + C\big(1 - p^i_{\text{fn}}\big) \end{bmatrix}, \tag{8.23a}$$

$$B^i_{\text{def}} \coloneqq \begin{bmatrix} & \hat{\theta}{=}0 & \hat{\theta}{=}1 & \hat{\theta}{=}Y_i \\ \theta{=}0 & 0 & A & Ap^i_{\text{fp}} \\ \theta{=}1 & B & 0 & Bp^i_{\text{fn}} \end{bmatrix}, \tag{8.23b}$$

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 is labeled with the corresponding policies for the attacker and defender, respectively (as enumerated above).

The following result provides explicit formulas for a mixed Nash equilibrium for the bimatrix game (8.23a) associated with the attacker's decision to reveal sensor $\sigma = i$.

Theorem 8.3: *Assuming that* (8.21) *holds and that the ith sensor is informative in the sense of* (8.22)*, the bimatrix game* (8.23a) *has a mixed Nash equilibrium of the form*

$${y^i_{\text{att}}}^* = \begin{cases} \begin{bmatrix} \frac{B(1-p^i_{\text{fn}})}{Ap^i_{\text{fp}}+B(1-p^i_{\text{fn}})} & \frac{Ap^i_{\text{fp}}}{Ap^i_{\text{fp}}+B(1-p^i_{\text{fn}})} \end{bmatrix}' & \bar{C}^i \geq R, \\ \begin{bmatrix} \frac{Bp^i_{\text{fn}}}{A(1-p^i_{\text{fp}})+Bp^i_{\text{fn}}} & \frac{A(1-p^i_{\text{fp}})}{A(1-p^i_{\text{fp}})+Bp^i_{\text{fn}}} \end{bmatrix}' & \bar{C}^i < R, \end{cases} \tag{8.24a}$$

$${z^i_{\text{def}}}^* = \begin{cases} \begin{bmatrix} \frac{\bar{C}^i - R}{\bar{C}^i} & 0 & \frac{R}{\bar{C}^i} \end{bmatrix}' & \bar{C}^i \geq R, \\ \begin{bmatrix} 0 & \frac{R-\bar{C}^i}{C+F-\bar{C}^i} & \frac{C+F-R}{C+F-\bar{C}^i} \end{bmatrix}' & \bar{C}^i < R, \end{cases} \tag{8.24b}$$

with values

$$J_{\text{att}}^{i\ *} = S_i - F \begin{cases} \frac{Rp_{\text{fp}}^i}{\bar{C}^i}, & \bar{C}^i \geq R, \\ \frac{(R-C)(1-p_{\text{fp}}^i)+Cp_{\text{fn}}^i}{C+F-\bar{C}^i}, & \bar{C}^i < R, \end{cases} \tag{8.25a}$$

$$J_{\text{def}}^{i\ *} = \begin{cases} \frac{ABp_{\text{fp}}^i}{Ap_{\text{fp}}^i+B(1-p_{\text{fn}}^i)}, & \bar{C}^i \geq R, \\ \frac{ABp_{\text{fn}}^i}{A(1-p_{\text{fp}}^i)+Bp_{\text{fn}}^i}, & \bar{C}^i < R, \end{cases} \tag{8.25b}$$

where $\bar{C}^i \coloneqq C(1-p_{\text{fn}}^i) + Fp_{\text{fp}}^i$. □

We conclude from Theorem 8.3 that the attacker minimizes its cost by revealing the sensor i that leads to the smallest value of its cost $J_{\text{att}}^{i\ *}$ in (8.25a). This provides the last piece of the policy for the attacker:

$$\sigma = \arg\min_i S_i - F \begin{cases} \frac{Rp_{\text{fp}}^i}{\bar{C}^i} & \bar{C}^i \geq R, \\ \frac{(R-C)(1-p_{\text{fp}}^i)+Cp_{\text{fn}}^i}{C+F-\bar{C}^i} & \bar{C}^i < R, \end{cases}$$

which corresponds to the top branch of the decision tree depicted in Figure 8.1. 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 (8.24a) for $i = \sigma$. The defender's selection of the maps $Y_i \mapsto \mu(i, Y_i)$ will typically also be mixed and given by the distributions in (8.24b), which typically depend on the sensor $\sigma = i$ that has been revealed.

8.4. Conclusions and Future Work

We considered sensor manipulation problems in which an attacker tries to induce a cyber-defense system to make the wrong decision. In measurement manipulation games the cyber-defense system is forced to make a decision based on a number of sensor recommendations that may have been altered by an attacker, whereas in sensor-reveal games the attacker selects which sensors the defender is allowed to use in making a decision.

The results for measurement manipulation games in Section 8.2 can be summarized as follows: For scenarios with a relatively small number of sensors (roughly $N \leq 2/p_{\text{err}}$) the game has a saddle-point equilibrium in mixed policies and the corresponding policy for the defender involves randomizing between a majority policy μ_{maj} that agrees with the most sensors and a somewhat unexpected policy $\mu_{\text{no-cons}}$ that most of the times follows a majority-like rule, but decides in opposition to all the sensor recommendations when these are in consensus. For the general case (which includes $N > 2/p_{\text{err}}$), it is shown that there exists an approximate Nash equilibrium corresponding to a deterministic voting-like decision rule that follows the sensor recommendations when more than a certain number of them agrees. While this rule is not optimal, the level of suboptimality decreases to zero as the number of sensors grows.

The results for sensor reveal games in Section 8.3 provide a closed form solution for mixed Nash equilibrium policies for both players. Under mild assumptions, the defender's policy involves randomizing between three deterministic policies: a policy that ignores Y_σ and always selects $\hat{\theta} = 0$, another one that also ignores Y_σ and always selects $\hat{\theta} = 1$, and a third policy that accepts the sensor revealed and selects $\hat{\theta} = Y_\sigma$. The probabilities associated with each

of these policies depend on the index σ of the sensor that has been revealed.

In the interest of clarity, this paper is focused on stylized versions of decision making processes based on data that has been manipulated. For practical application, complexity may need to be added to these models, especially in what regards the nature of the sensor measurements. This is particularly true for the sensor-reveal games, for which we restricted our attention to decisions based on a single binary sensor $Y_\sigma \in \{0, 1\}$. The paper [Hespanha and Garagic, 2019] on which Section 8.3 is based considers more general (non-binary) sensors, but further research is needed to extend this work to sensor models for which the Y_i are not necessarily conditionally independent, given θ. Similarly, for many cyber-defense problems a binary decision $\hat{\theta} \in \{0, 1\}$ is too restrictive, as one may want to decide among one of several alternative defense actions, some of which could include deploying additional cyber defense sensors. This chapter touches upon a rich area of research with numerous open problems that are crucial for the design of cyber-defense mechanisms robust to information manipulation. Among these it is worth highlighting that cyber-defense mechanisms interact with attackers in a repeated fashion over time, which provides opportunities for mutual learning [Kamhoua et al., 2020, Part 3].

Bibliography

T. Başar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press, London, 1995.

Cuong T Do, Nguyen H Tran, Choongseon Hong, Charles A Kamhoua, Kevin A Kwiat, Erik Blasch, Shaolei Ren, Niki Pissinou, and Sundaraja Sitharama

Iyengar. Game theory for cyber security and privacy. *ACM Computing Surveys (CSUR)*, 50(2):1–37, 2017.

Jairo Giraldo, David Urbina, Alvaro Cardenas, Junia Valente, Mustafa Faisal, Justin Ruths, Nils Ole Tippenhauer, Henrik Sandberg, and Richard Candell. A survey of physics-based attack detection in cyber-physical systems. *ACM Computing Surveys (CSUR)*, 51(4):1–36, 2018.

João Pedro Hespanha. *Noncooperative Game Theory: An Introduction for Engineers and Computer Scientists*. Princeton Press, Princeton, New Jersey, June 2017.

João Pedro Hespanha and Denis Garagic. Sensor-reveal games. In *Proc. of the 58th Conf. on Decision and Contr.*, December 2019.

Charles A. Kamhoua, Christopher D. Kiekintveld, Fei Fang, and Quanyan Zhu, editors. *Game Theory and Machine Learning for Cyber Security*. Wiley-IEEE Press, 2020.

Kyriakos G. Vamvoudakis, João Pedro Hespanha, Bruno Sinopoli, and Yilin Mo. Detection in adversarial environments. *IEEE Trans. on Automatic Control,* Special Issue on the Control of Cyber-Physical Systems, 59(12): 3209–3223, December 2014.