Secure Detection in the Presence of Integrity Attacks

Yilin Mo*, João Hespanha† and Bruno Sinopoli*

*: Department of Electrical and Computer Engineering
Carnegie Mellon University
†: Department of Electrical and Computer Engineering
University of California, Santa Barbara

Nov 3, 2011

Motivation

- Cyber Physical Systems (CPS) refer to the embedding of widespread sensing, computation, communication and control into physical spaces.
- Applications: aerospace, chemical processes, civil infrastructure, energy, manufacturing and transportation. Safety Critical
- The next generation CPS, such as smart grids, will make extensive use of information technology.
- Stuxnet raised significant concerns on CPS security.
- How to design secure CPS?

A Classical Detection Problem

- We want to decide whether the state θ is 1 or -1.

$$\theta = \begin{cases} -1 & \text{w.p. } 0.5 \\ +1 & \text{w.p. } 0.5 \end{cases}$$

- m sensors are measuring the system:

$$y_i = \theta + v_i.$$

- v_i are i.i.d. Gaussian noise with mean 0 and variance 1.
- A detector is a function $f : \mathbb{R}^m \to \{-1, 1\}$.
- The optimal detector with minimum detector error is a Naive Bayesian Detector:

$$\hat{\theta} = f(y) = \begin{cases} -1 & \text{if } \sum_i y_i < 0 \\ +1 & \text{if } \sum_i y_i \geq 0 \end{cases}$$

System Diagram

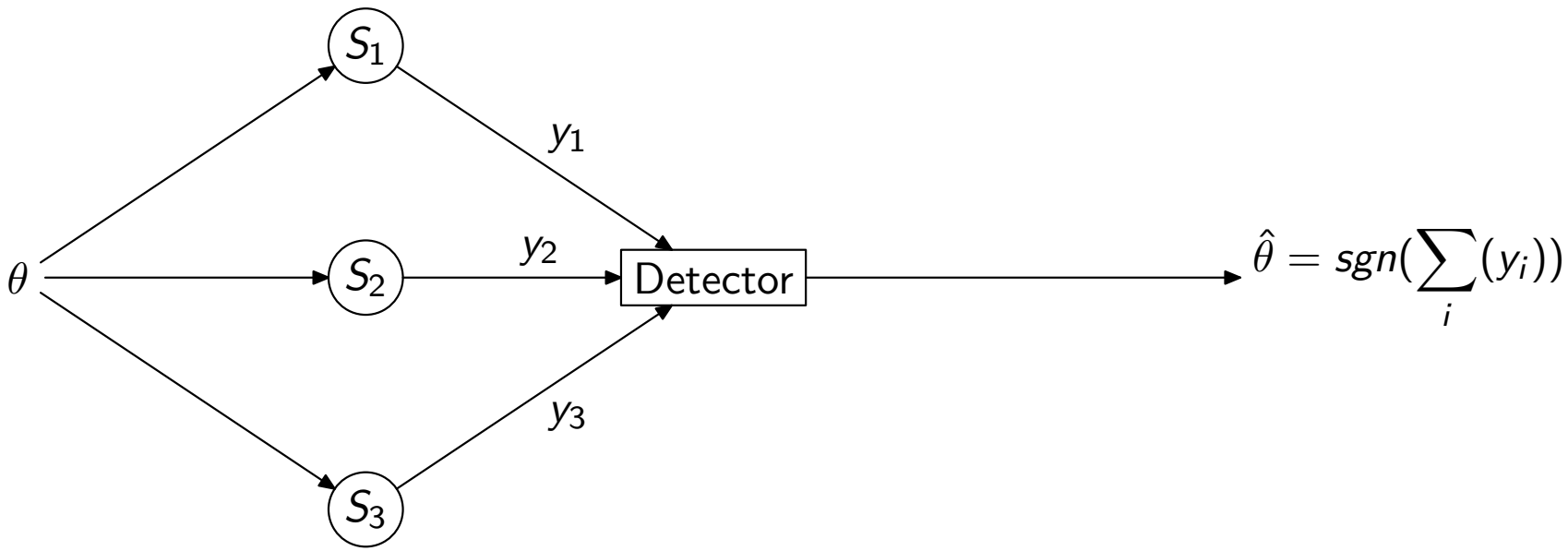


Figure: System Diagram

Suppose the attacker compromise one sensor. Thus, it can manipulate $\sum_i y_i$ arbitrarily. Hence, it has full control over $\hat{\theta}$.

Countermeasures

Information Security

- Tamper-resistant microprocessor, Software attestation, Secure Communication Protocol,...
- It is hard to guarantee security for every single sensor. (A single compromised sensor can totally ruin the Naive Bayes detector.)
- Physical attacks?

System Theory

- Bad data detection, Robust detection ($\varepsilon-$contamination, bounded total variation difference, ...)
- The uncertainty models of system theoretic approaches are usually not quite different from the cyber attacks.

Our goal: To design the optimal detector that can withstand Byzantine attacks from at most n sensors.

Attack Model

We assume the attacker knows the following:

- the detection algorithm f (Kerckhoffs' Principle);
- the true state θ;
- all measurements y.

The attacker can manipulate up to n measurements arbitrarily.

$$y' = y + u, \|u\|_0 \leq n.$$

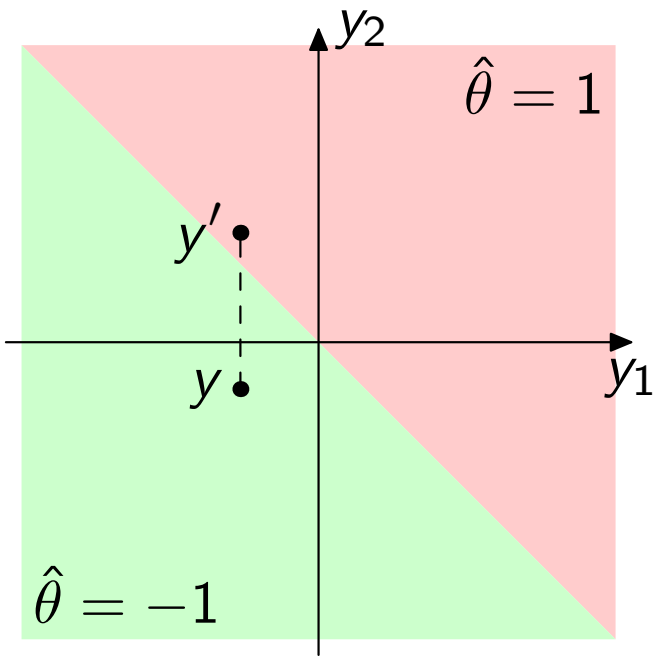
The attacker wants to minimize the probability of detection (or maximize probability of error):

$$P_d(f) = P(f(y') = \theta)$$

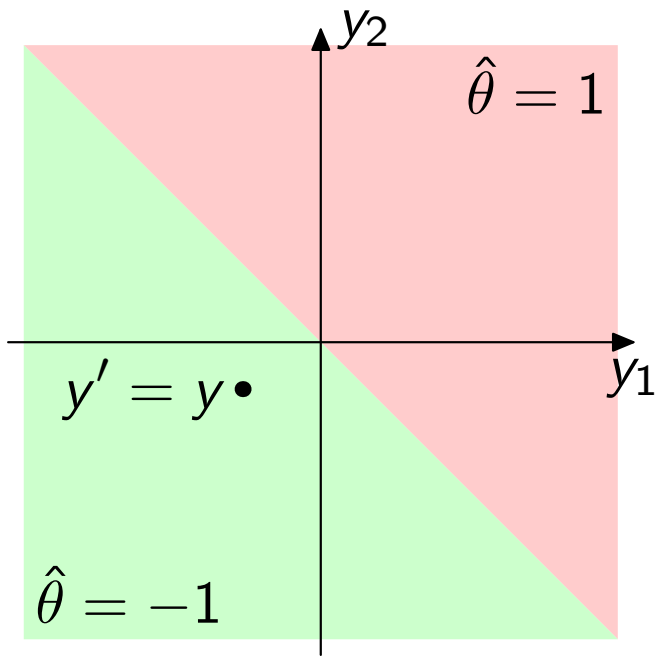
Attacker's Strategy

The optimal strategy for the attacker:

$$u = \begin{cases} argmin_{\|u\|_0 \le n} f(y+u) & (\theta = 1) \\ argmax_{\|u\|_0 \le n} f(y+u) & (\theta = -1) \end{cases}$$

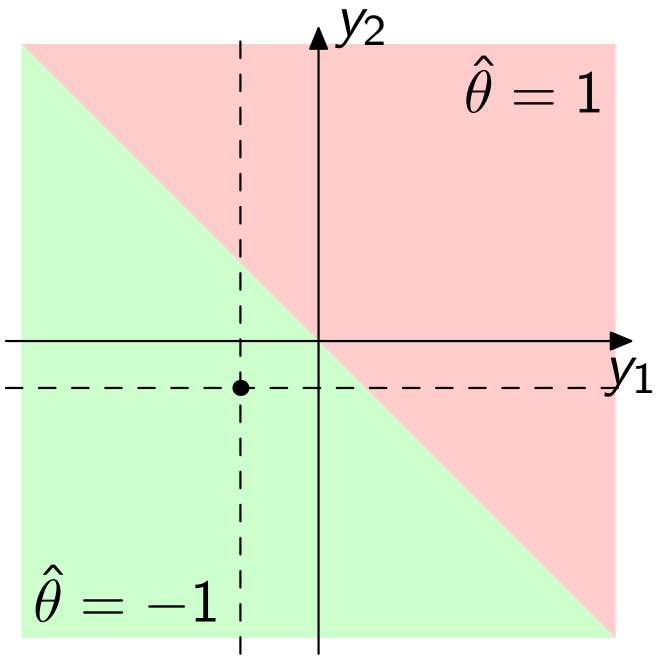


(a) $\theta = -1$

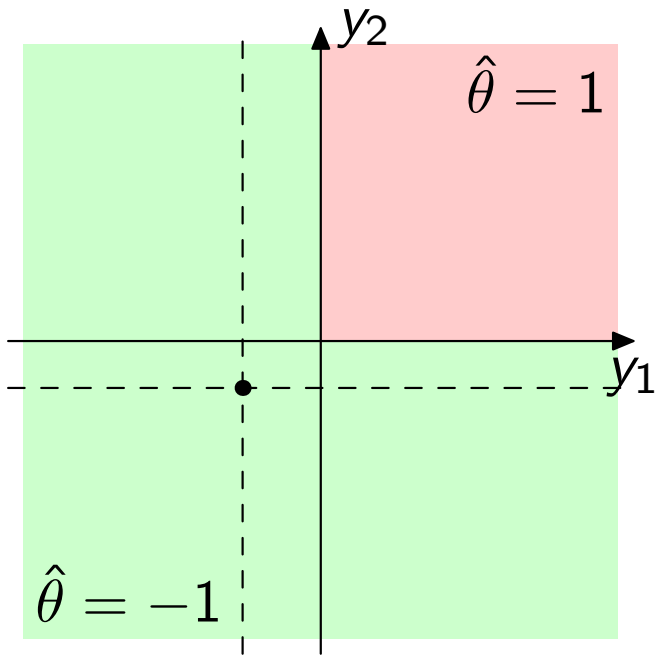


(b) $\theta = 1$

"Good" Point



(c) First Detector f_1



(d) Second Detector f_2

There are three categories of point in R^m:

1. for all $\|u\|_0 \leq n$, $f(y+u) = 1$. (Good)
2. for all $\|u\|_0 \leq n$, $f(y+u) = -1$. (Good)
3. for some $\|u\|_0 \leq n$, $f(y+u) = 1$ and for some $\|u\|_0 \leq n$, $f(y+u) = -1$. (Bad)

"Good" sets

Define

$$Y^+(f) \triangleq \{y : f(y+u) = 1,\ \forall \|u\|_0 \le n\},$$
$$Y^-(f) \triangleq \{y : f(y+u) = -1,\ \forall \|u\|_0 \le n\},$$

The detector can correctly detect θ if and only if

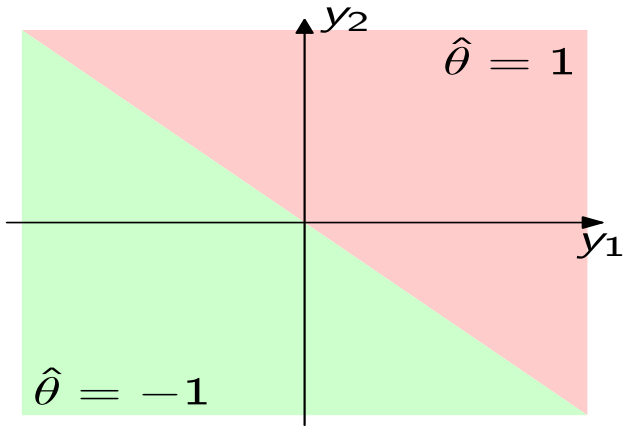
$$(\theta, y) \in \Big\{(-1, y) : y \in Y^-(f)\Big\} \cup \Big\{(+1, y) : y \in Y^+(f)\Big\}$$

As a result

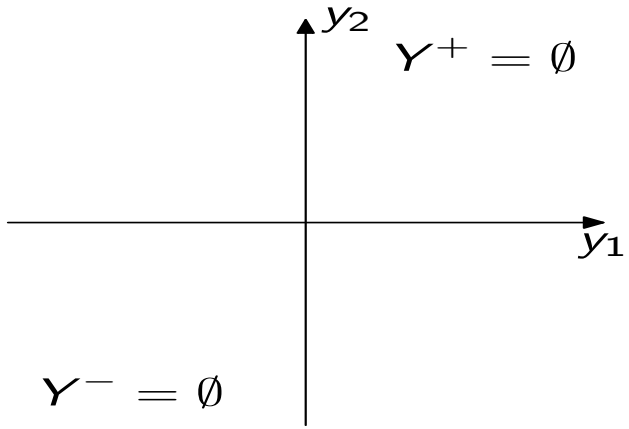
$$P_d(f) = p^- P(y \in Y^-(f) | \theta = -1) + p^+ P(y \in Y^+(f) | \theta = 1).$$

If $Y^-(f) \subseteq Y^-(g)$ and $Y^+(f) \subseteq Y^+(g)$, then $P_d(f) \le P_d(g)$.

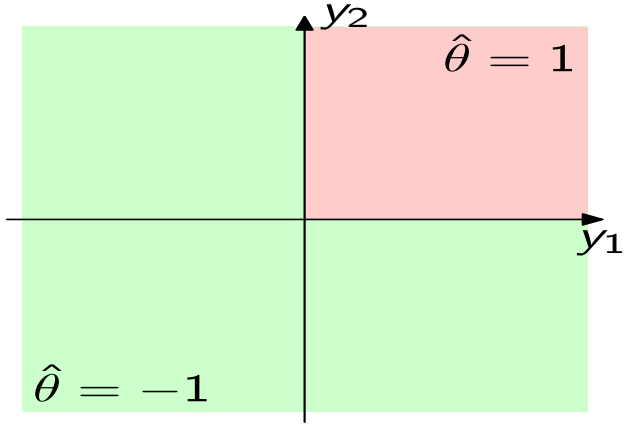
Y^- and Y^+ for different detectors



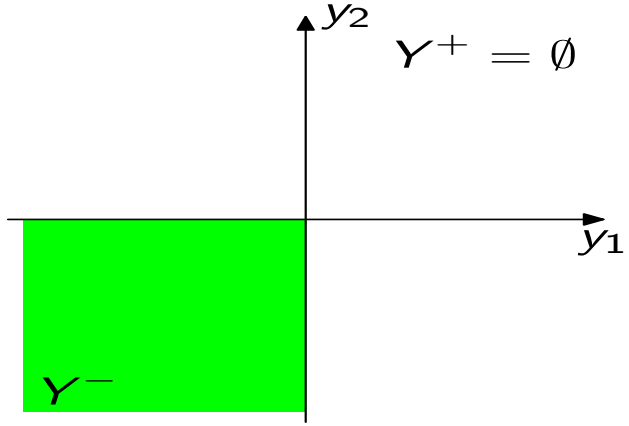
(e) First Detector f_1



(f) Y^- and Y^+ of f_1



(g) Second Detector f_2



(h) Y^- and Y^+ of f_2

A Hamming-like distance

Define metric $d : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{N}_0$ as

$$d(x, y) = \|x - y\|_0.$$

Let X, Y be two subsets of $\mathbb{R}^m$, define

$$d(X, Y) = min_{x \in X, y \in Y} d(x, y),\ d(x, Y) = d(\{x\}, Y).$$

Lemmas

Lemma

For any detector f, $d(Y^-(f), Y^+(f)) \geq 2n+1$.

Lemma

Given X^-, X^+ two subsets of $\mathbb{R}^m$, and $d(X^-, X^+) \geq 2n+1$, there exists a detector f, such that $X^- \subseteq Y^-(f)$ and $X^+ \subseteq Y^+(f)$.

Main Theorem

Theorem

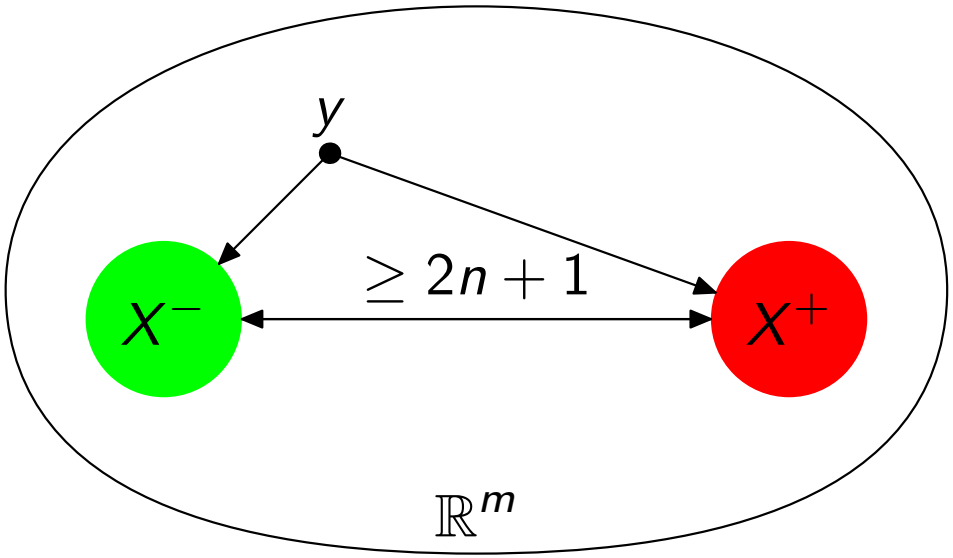
The optimal f^* is of the following form

$$f^*(y) = \begin{cases} 1 & d(y, X^-) \geq d(y, X^+) \\ -1 & d(y, X^-) < d(y, X^+), \end{cases}$$

where X^+ and X^- are the solutions of the following optimization problem:

$$\begin{aligned} \underset{X^+, X^-}{maximize} \quad & P(y \in X^- | \theta = -1)p^- + P(y \in X^+ | \theta = 1)p^+ \\ subject\ to \quad & d(X^-, X^+) \geq 2n + 1 \end{aligned}$$

The structure of the optimal detector



Main results

- If $n \geq m/2$, then the optimal detector is $f = 1$ or $f = -1$.
- The true optimal detector is difficult to compute when $n < m/2$.
- As a result, we propose a heuristic detector based on trimmed mean:
  1. The detector sorts all the y_is in descending order.
  2. The detector throws away n measurements with the largest y_is and n measurements with the least y_is.
  3. The detector sums the remaining $m - 2n$ y_is and compares it to 0. The detector chooses $\hat{\theta} = -1$ is the truncated sum is less than 0, otherwise the detector chooses $\hat{\theta} = 1$.

  It can be proved that the corresponding X^- and X^+ are

$$X^- = \{y \in \mathbb{R}^m : \sum_{i \in \mathcal{I}} y_i < 0, \forall |\mathcal{I}| = m - 2n, \mathcal{I} \subset \{1, \ldots, m\}\},$$

$$X^+ = \{y \in \mathbb{R}^m : \sum_{i \in \mathcal{I}} y_i \geq 0, \forall |\mathcal{I}| = m - 2n, \mathcal{I} \subset \{1, \ldots, m\}\}.$$

A Heuristic Detector

- Suppose we received the following measurements:

$$y_1 = -2,\ y_2 = 5,\ y_3 = 1,\ y_4 = -3,\ y_5 = 0.5.$$

- After sorting we have:

$$5,\ 1,\ 0.5,\ -2,\ -3.$$

- If $n = 0$, then we have $5 + 1 + 0.5 - 2 - 3 = 1.5 > 0$. Hence $\hat{\theta} = 1$.
- If $n = 1$, then we have $1 + 0.5 - 2 = -0.5 < 0$. Hence $\hat{\theta} = -1$.
- If $n = 2$, then we have $0.5 > 0$. Hence $\hat{\theta} = 1$.

Heuristic Detector for $n < m/2$

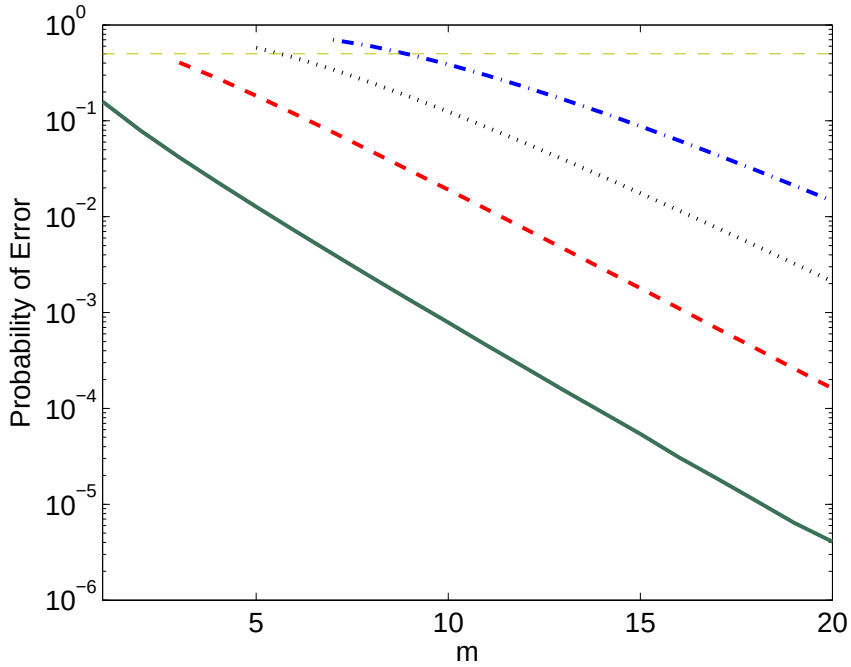


Figure: Probability of Error v.s. Number of Sensors(m). The green line: $n = 0$. The red line: $n = 1$. The black line: $n = 2$. The blue line: $n = 3$.

We can prove that the heuristic detector is asymptotically optimal.

Conclusion

- Information security alone is not sufficient for CPS security. We also need a more secure system theory.
- We present a robust detector design which can withstand integrity attacks on up to n sensors.
- In particular, if $n \geq m/2$, then the optimal detector is constant filter $f = 1$ or $f = -1$.
- For general $n < m/2$, we propose a heuristic detector design.