Secure Information Fusion in Cyber-Physical Systems

Yilin Mo

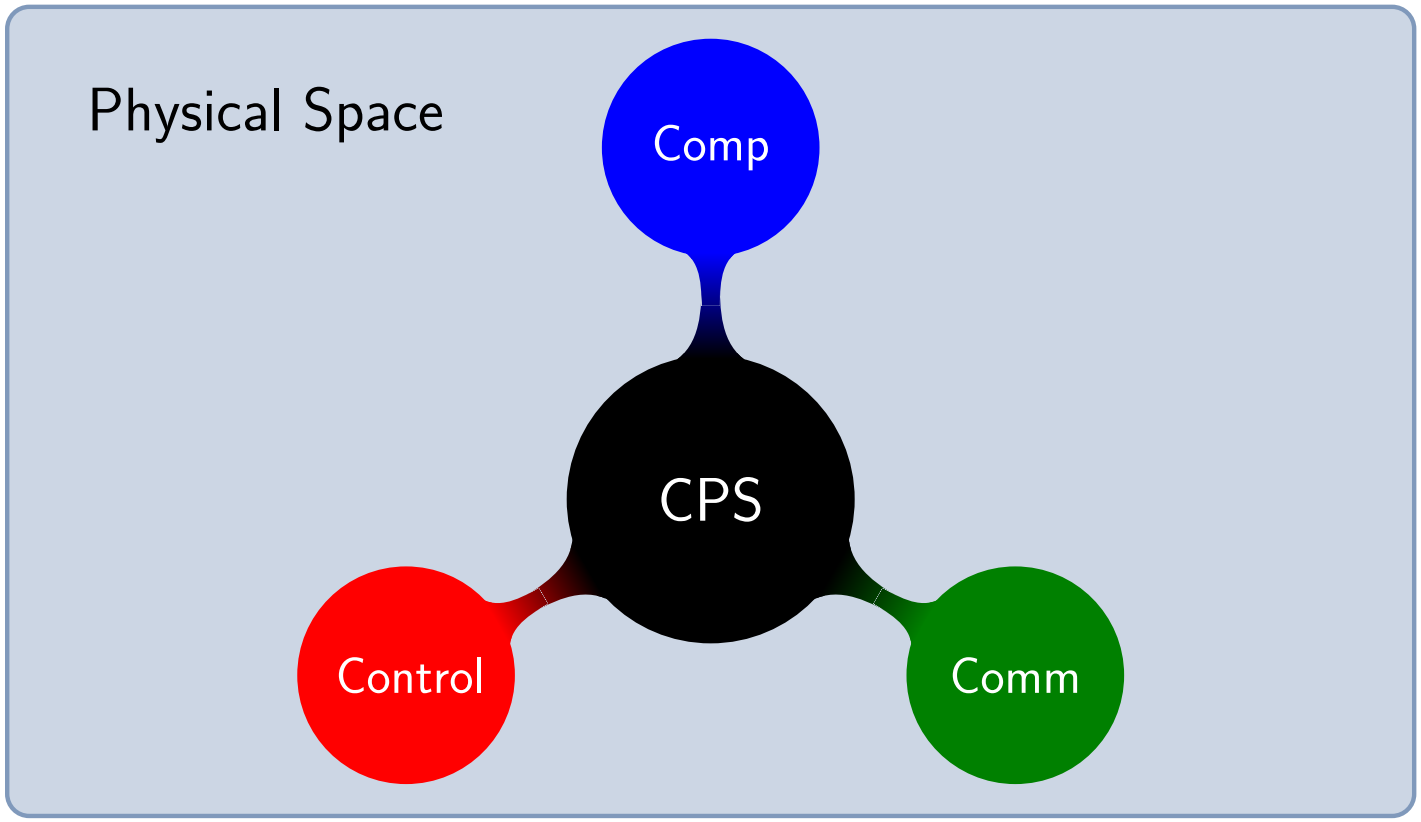
School of EEE
Nanyang Technological University

December 11, 2015 at Shanghai Jiao Tong University

Joint Work with Richard M. Murray, Duo Han, Lihua Xie

Cyber-Physical System

- Cyber-Physical Systems (CPSs) refer to the embedding of computation, communication and control into physical spaces.



- Applications: aerospace, chemical processes, civil infrastructure, energy, manufacturing and transportation.

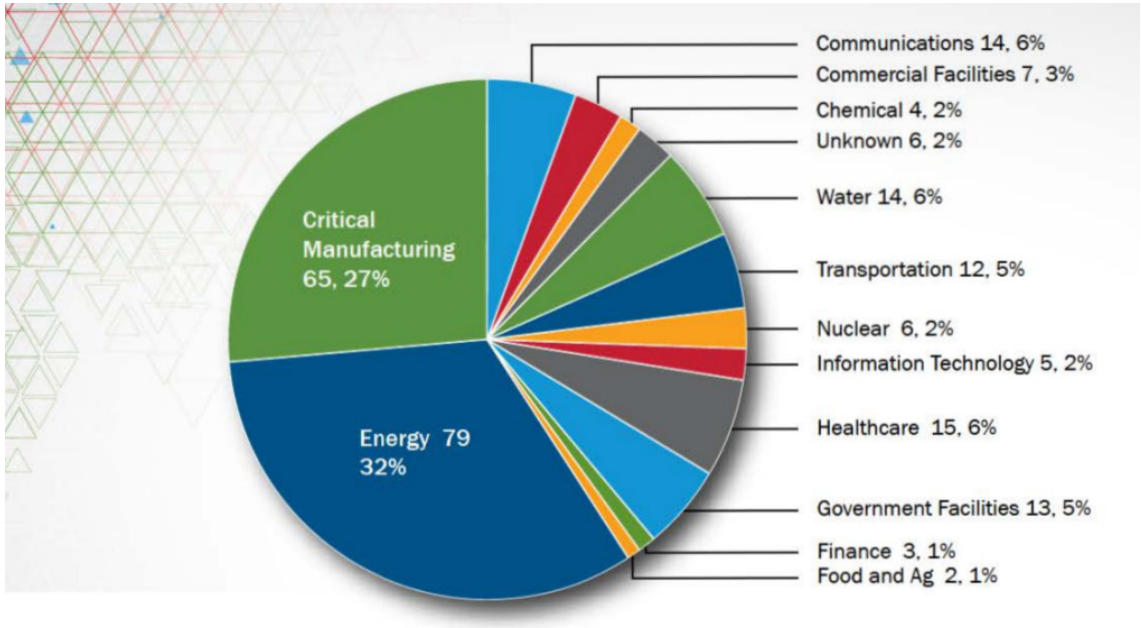
Security Threats for the CPS

- The next generation CPS: Smart Grids, Smart Buildings, Smart Home, Internet of Things, will make extensive use of widespread sensing and networking.
- As the CPSs become "smarter", they are also more vulnerable to malicious attacks.

Stuxnet

Stuxnet is the first discovered malware that spies on and subverts industrial control systems. It was discovered in June 2010.

Industrial Control Systems



FY 2014 incidents reported by sector (245 total).

In FY 2014, ICS-CERT (Industrial Control Systems Cyber Emergency Response Team) received and responded to 245 incidents as reported by asset owners and industry partners.

Industrial Control Systems

The scope of incidents encompassed a vast range of threats and observed methods for attempting to gain access to both business and control systems infrastructure, including but not limited to the following:

1. Unauthorized access and exploitation of Internet facing ICS/Supervisory Control and Data Acquisition (SCADA) devices,
2. Exploitation of zero-day vulnerabilities in control system devices and software,
3. Malware infections within air-gapped control system networks,
4. ...

Attack Through Compromised Supply Chain

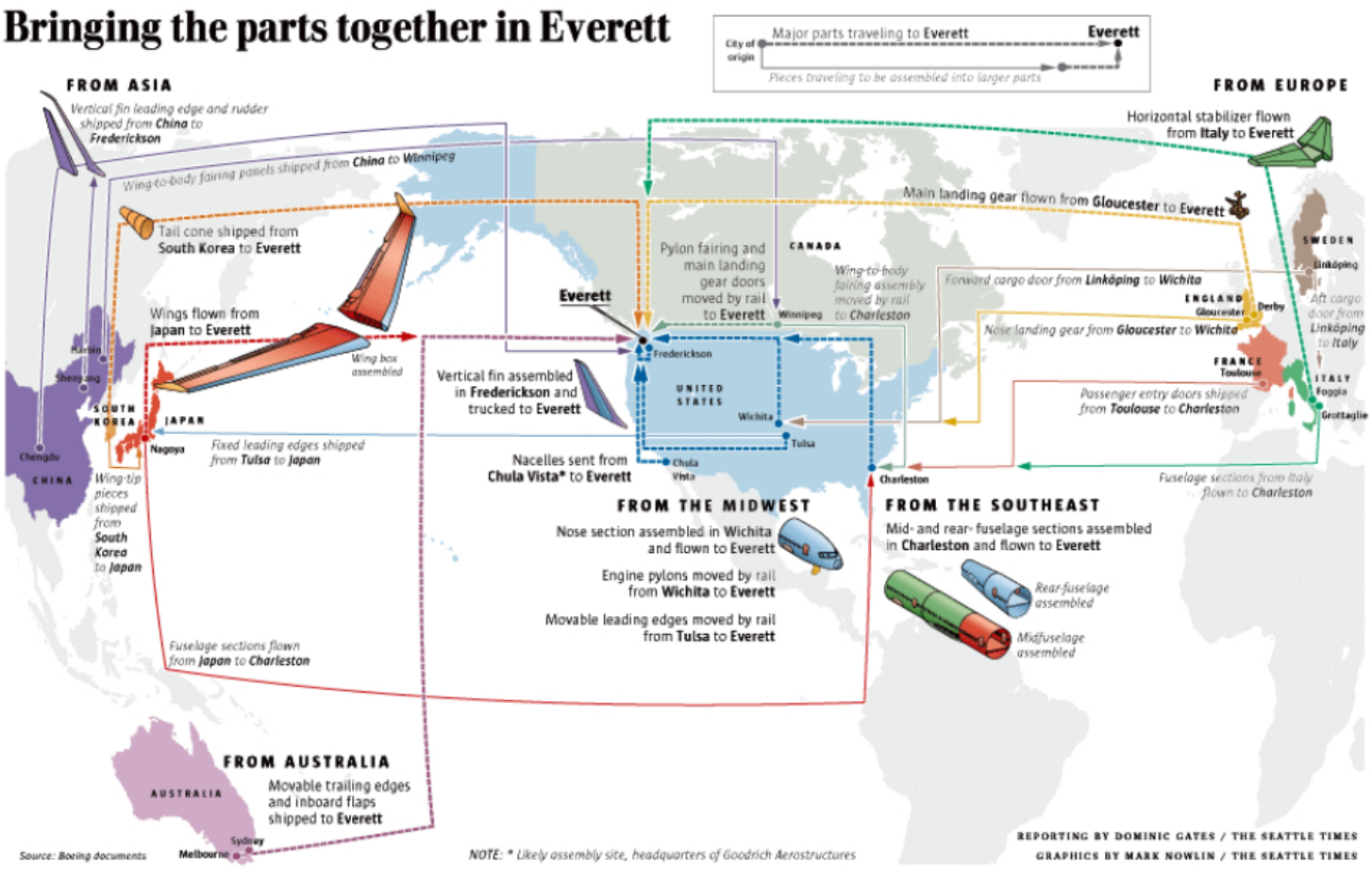


Figure : Boeing 787 outsourced 70% of its parts.

2003 Northeast Blackout

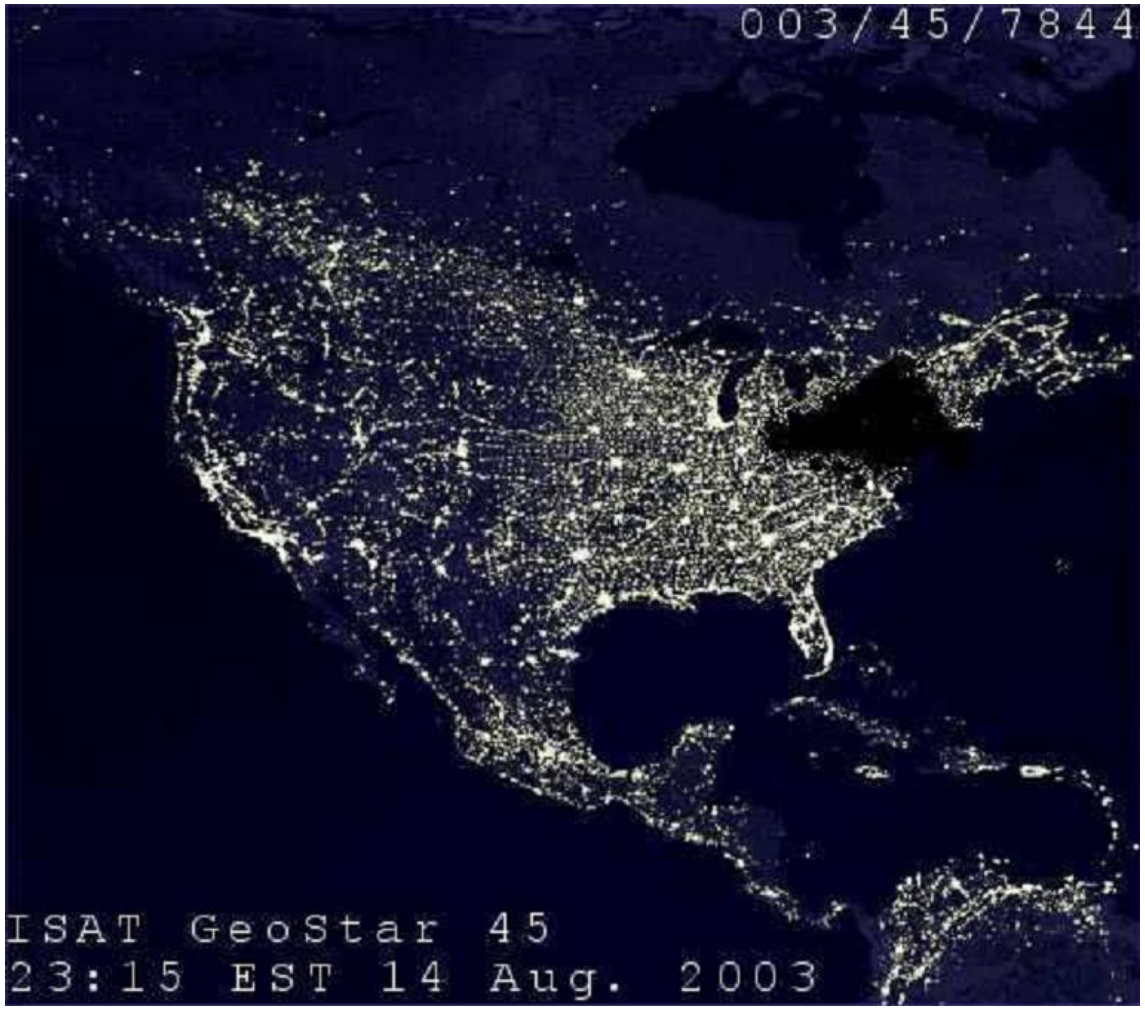


Figure : A successful attack on CPS can have devastating effects.

How to deal with CPS security threats

"知彼知己，百战不殆；不知彼而知己，一胜一负；不知彼，不知己，每战必殆。"

—《孙子·谋攻篇》

1. 知彼: Intrusion detection and isolation
2. 知己: Information Fusion

Overview of My Research in CPS

Networked Control System

Understanding the interaction between control and communication.

- Kalman filtering with intermittent observation
- Sensor/Actuator Scheduling: offline-schedules and event-based schedules.
- Energy efficiency of consensus algorithm

CPS Security

- Replay attack: we proposed countermeasures to Stuxnet one year before its discovery
- Deception attack
- CPS security and electricity market
- Information fusion

1 Introduction

2 Secure State Estimation

3 Privacy Preserving Consensus

4 Conclusion

Preliminary: State Estimation

1. We assume that $x \in \mathbb{R}^n$ is the state that we want to estimate.
2. m sensors are deployed to monitor the system. Denote $z_i \in \mathbb{R}$ as the measurement generated by sensor i.
3. Denote $z = \begin{bmatrix} z_1 & \dots & z_m \end{bmatrix}^T$ as the collection of all sensory data.
4. We assume the following sensor model:

$$z = Hx + w,$$

   where H is a matrix of proper dimension and w represent random noise.
5. The optimal state estimator is of the form

$$\hat{x} = Kz, \text{ where } K = (H^T H)^{-1} H^T.$$

Preliminary: Secure State Estimation

Suppose the following sensory model:

$$y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} x + w.$$

The optimal estimator is given by

$$\hat{x} = \frac{1}{3}\left(y_1 + y_2 + y_3\right).$$

This estimator is not resilient to a single malicious sensor.

Problem Formulation

1. Assume that at most p sensors are compromised.
2. The sensory model is given by

$$y = z + a = Hx + w + a,$$

   where a is a p-sparse vector indicating the attacker's action.
3. We will call an estimator $\hat{x} = g(y)$ to be resilient if

$$\|g(y) - g(z)\| \text{ is bounded for all } p\text{-sparse } a.$$

4. The total state estimation error can be written as

$$x - g(y) = x - g(z) + g(z) - g(y).$$

   Resiliency means that the state estimation error caused by the adversary is bounded!

Why do not use 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 linear estimator.)
- Physical attacks?
- A resilient estimator can provide an additional layer of protection.

Why not use bad data detection?

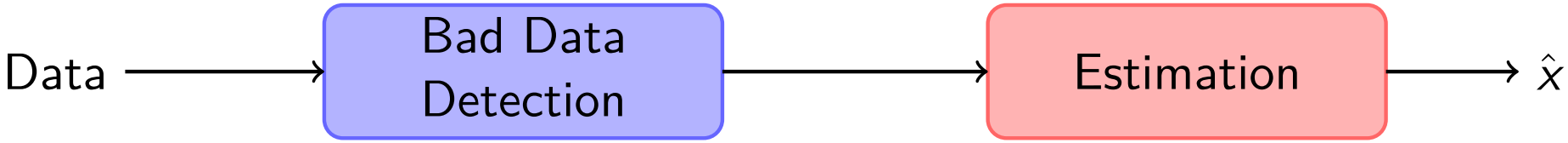


Figure : Bad Data Detection Scheme

1. A typical bad data detector will first compute $\hat{x} = Ky$ and then the residue $r = y - H\hat{x}$.
2. The bad data detector will then remove the sensors with large residue.
3. Drawbacks: It is designed for *random failures*, not necessarily attacks.
4. One can view the Bad data detector + estimator as a big estimator.

Why not use bad data detection?

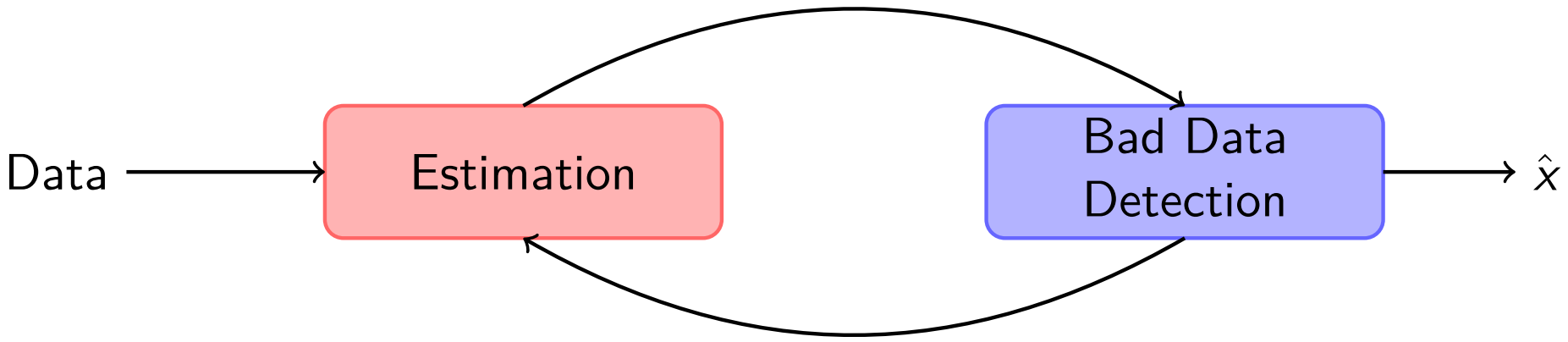


Figure : Bad Data Detection Scheme

1. A typical bad data detector will first compute $\hat{x} = Ky$ and then the residue $r = y - H\hat{x}$.
2. The bad data detector will then remove the sensors with large residue.
3. Drawbacks: It is designed for *random failures*, not necessarily attacks.
4. One can view the Bad data detector + estimator as a big estimator.

Why not use bad data detection?

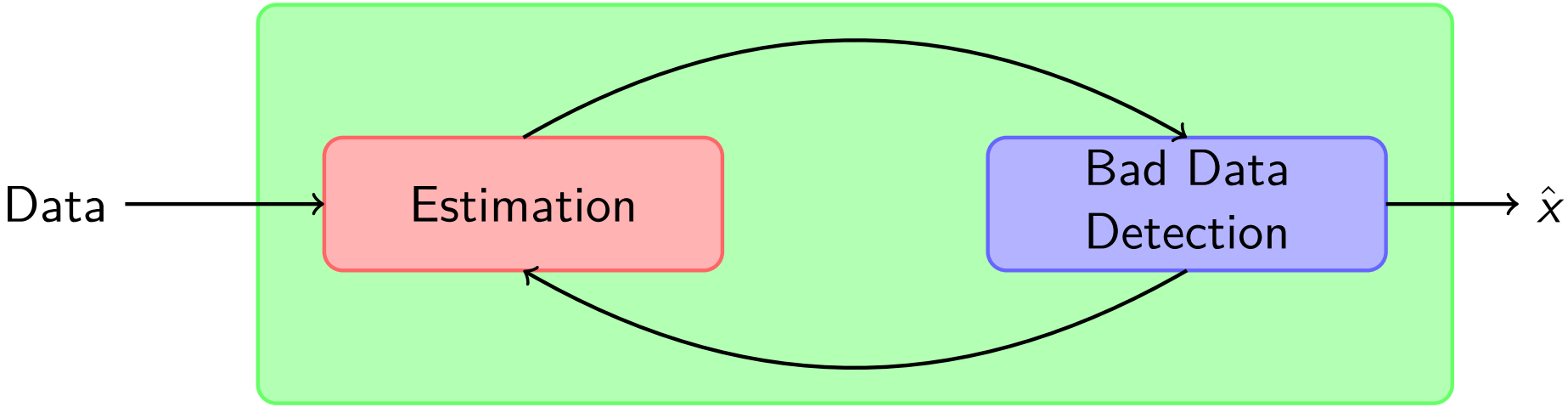


Figure : Bad Data Detection Scheme

1. A typical bad data detector will first compute $\hat{x} = Ky$ and then the residue $r = y - H\hat{x}$.
2. The bad data detector will then remove the sensors with large residue.
3. Drawbacks: It is designed for *random failures*, not necessarily attacks.
4. One can view the Bad data detector + estimator as a big estimator.

Some Related Research

1. The following estimator is resilient:

$$\begin{array}{ll} \underset{\hat{x},a,w}{minimize} & \|w\|^2 \\ \text{subject to} & y = H\hat{x} + w + a,\ \|a\|_0 \leq p. \end{array}$$

   The drawback: The optimization problem is not convex and solving it is computationally difficult.

2. We can use convex optimization based estimator, e.g.,

$$\begin{array}{ll} \underset{\hat{x},a,w}{minimize} & \|w\|^2 + \|a\|_1 \\ \text{subject to} & y = H\hat{x} + w + a. \end{array}$$

   However, there is no guarantee on the resiliency.

A General Convex Optimization Based Estimator

We consider the following estimator

$$\hat{x} = g(y) \triangleq \arg\min_{\hat{x}} \sum_{i=1}^{m} f_i(y_i - H_i\hat{x}),$$

where the following properties of function $f_i : \mathbb{R} \mapsto \mathbb{R}$ are assumed:

1. f_i is convex.
2. f_i is symmetric, i.e., $f_i(u) = f_i(-u)$.
3. f_i is non-negative and $f_i(0) = 0$.

If we define the residue $r_i \triangleq y_i - H_i\hat{x}$, then we are trying to minimize the sum of some cost function related to the residue.
The estimator can be computed efficiently via convex optimization. (It is also easy be computed in a parallel fashion.)

A General Convex Optimization Based Estimator

Our proposed estimator is very general since we can choose the right f_i to get the following estimator:

1. Linear Estimator:

$$g(y) = \arg\min_{\hat{x}} \|y - H\hat{x}\|_2^2 = \arg\min_{\hat{x}} \sum_{i=1}^{m} (y_i - H_i\hat{x})^2.$$

2. L_1 Estimator:

$$g(y) = \arg\min_{\hat{x}} \|y - H\hat{x}\|_1 = \arg\min_{\hat{x}} \sum_{i=1}^{m} |y_i - H_i\hat{x}|.$$

A General Convex Optimization Based Estimator

3. LASSO:

$$g(y) = \arg\min_{\hat{x}} \|w\|^2 + \lambda\|a\|_1, \text{ s.t. } y = H\hat{x} + w + a.$$

We can define $r_i = y_i - H_i\hat{x} = w_i + a_i$. Then the problem can be rewritten as

$$g(y) = \arg\min_{\hat{x}} \sum_{i=1}^{m} \min_{w_i} \left(\|w_i\|^2 + \lambda\|r_i - w_i\|_1\right), \text{ s.t. } r_i = y_i - H\hat{x}.$$

As a result, we can rewrite g as

$$g(y) = \arg\min_{\hat{x}} \sum_{i=1}^{m} f(y_i - H_i\hat{x})$$

where $f(r) = \min_w w^2 + \lambda|r - w|$.

Example

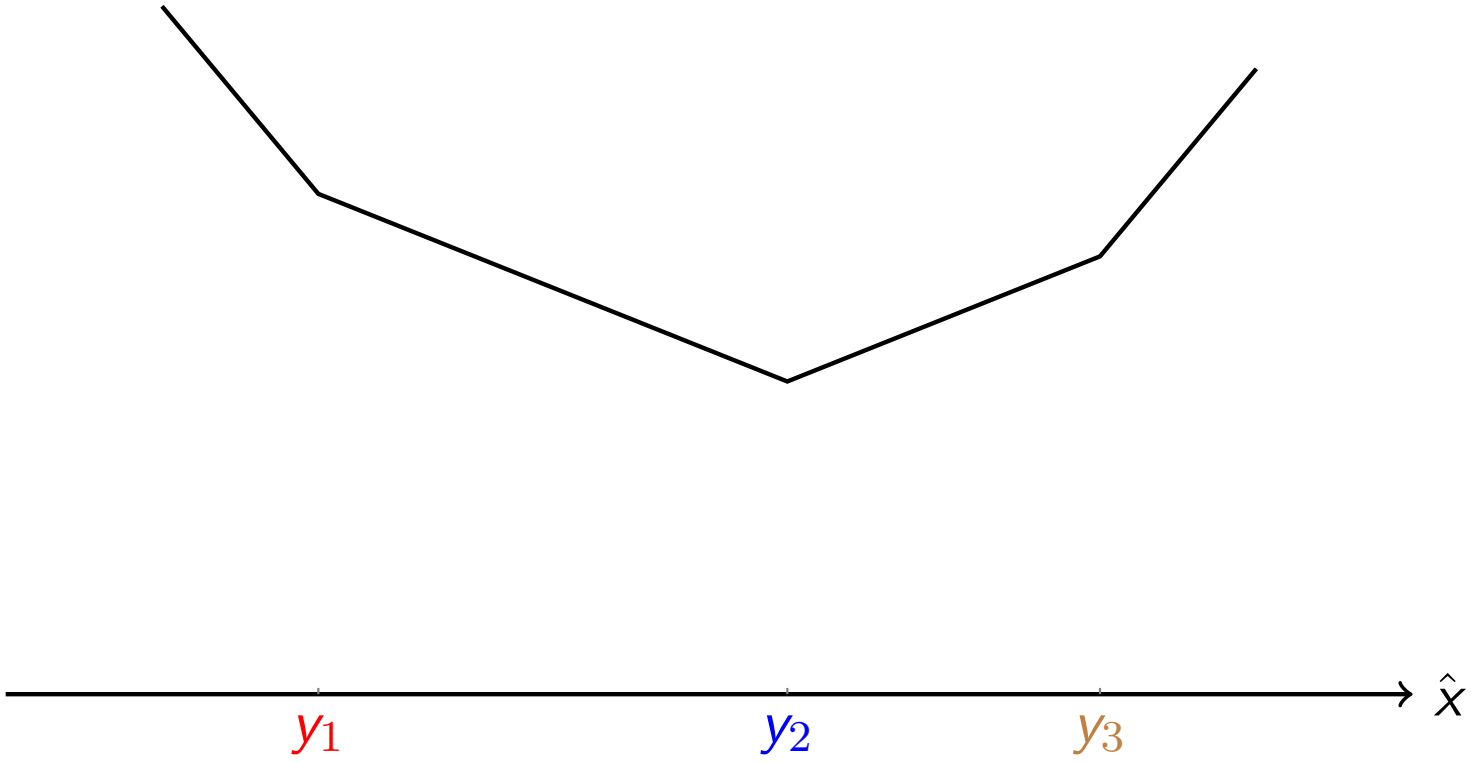
Suppose the following sensory model:

$$y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} x + w + a, \ \|a\|_0 \leq 1.$$

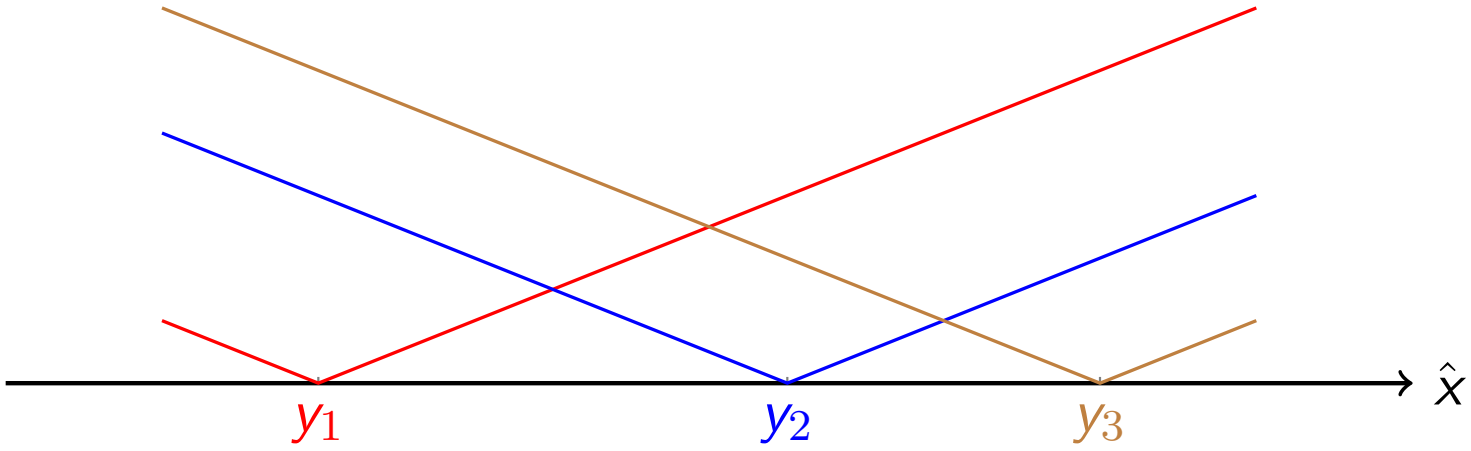
Then the following estimator is resilient (in fact, it is the median of y_i):

$$g(y) = \arg\min_{\hat{x}} |y_1 - \hat{x}| + |y_2 - \hat{x}| + |y_3 - \hat{x}|.$$

Example



Interpretation



If we interpret the function $|y_i - \hat{x}|$ as a potential function generate by sensor i, then we can see that sensor i is dragging $\hat{x}$ towards y_i with 1 unit of force. The equilibrium point will be at the middle y_i.

Another Example

Suppose the following sensory model:

$$y = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} x + w + a, \ \|a\|_0 \leq 1.$$

Then the following estimator is not resilient:

$$g(y) = \arg\min_{\hat{x}} |y_1 - \hat{x}| + |y_2 - \hat{x}| + |y_3 - 3\hat{x}|.$$

In fact, we can rewrite it as

$$g(y) = \arg\min_{\hat{x}} |y_1 - \hat{x}| + |y_2 - \hat{x}| + 3\left|\frac{y_3}{3} - \hat{x}\right| = \frac{y_3}{3}$$

Sensor 3 generates 3 unit of force comparing to 1 unit of force from sensor 1 and 2.

Sufficient Condition For Resiliency

Theorem

If the following conditions hold, then the estimation is resilient:

1. *For all i, the following limit is well-defined:*

$$\lim_{t\to\infty} \frac{f_i(t)}{t} = \alpha_i < \infty.$$

2. *For any $u \neq 0$ and any index set $\mathcal{I}$ of cardinality p, the following inequality hold:*

$$\sum_{i\in\mathcal{I}} |\alpha_i H_i u| < \sum_{i\in\mathcal{I}^c} |\alpha_i H_i u|.$$

Roughly speaking $|\alpha_i H_i u|$ is the force from sensor i on the direction u. The condition can be interpreted as the force from any p sensors should be smaller than the remaining $m-p$ sensors on all direction.

Necessary Condition For Resiliency

Theorem

If the one of the following conditions is violated, then the estimation is not resilient:

1. *There exists an i, such that*

$$\lim_{t\to\infty} \frac{f_i(t)}{t} = \infty.$$

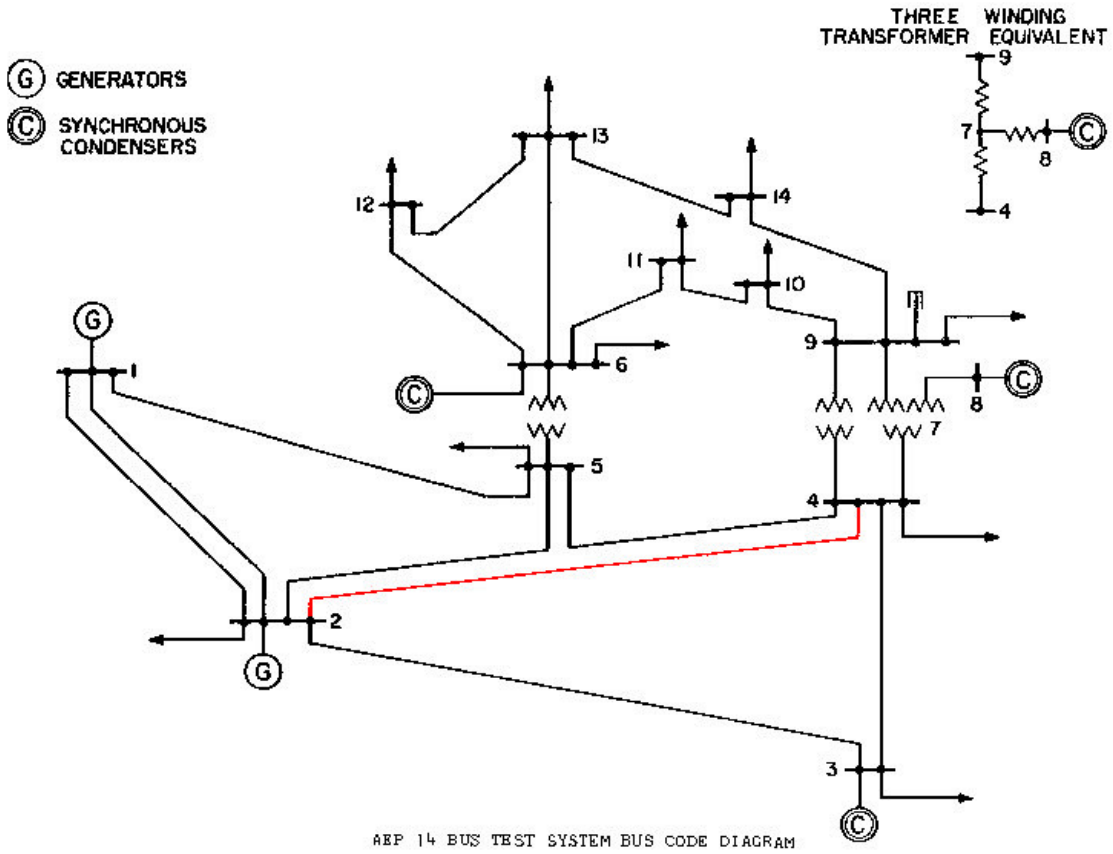
2. *There exists a $u \neq 0$ and an index set $\mathcal{I}$ of cardinality p, such that*

$$\sum_{i\in\mathcal{I}} |\alpha_i H_i u| > \sum_{i=\mathcal{I}^c} |\alpha_i H_i u|.$$

Notice that we only have a trivial gap for the case:

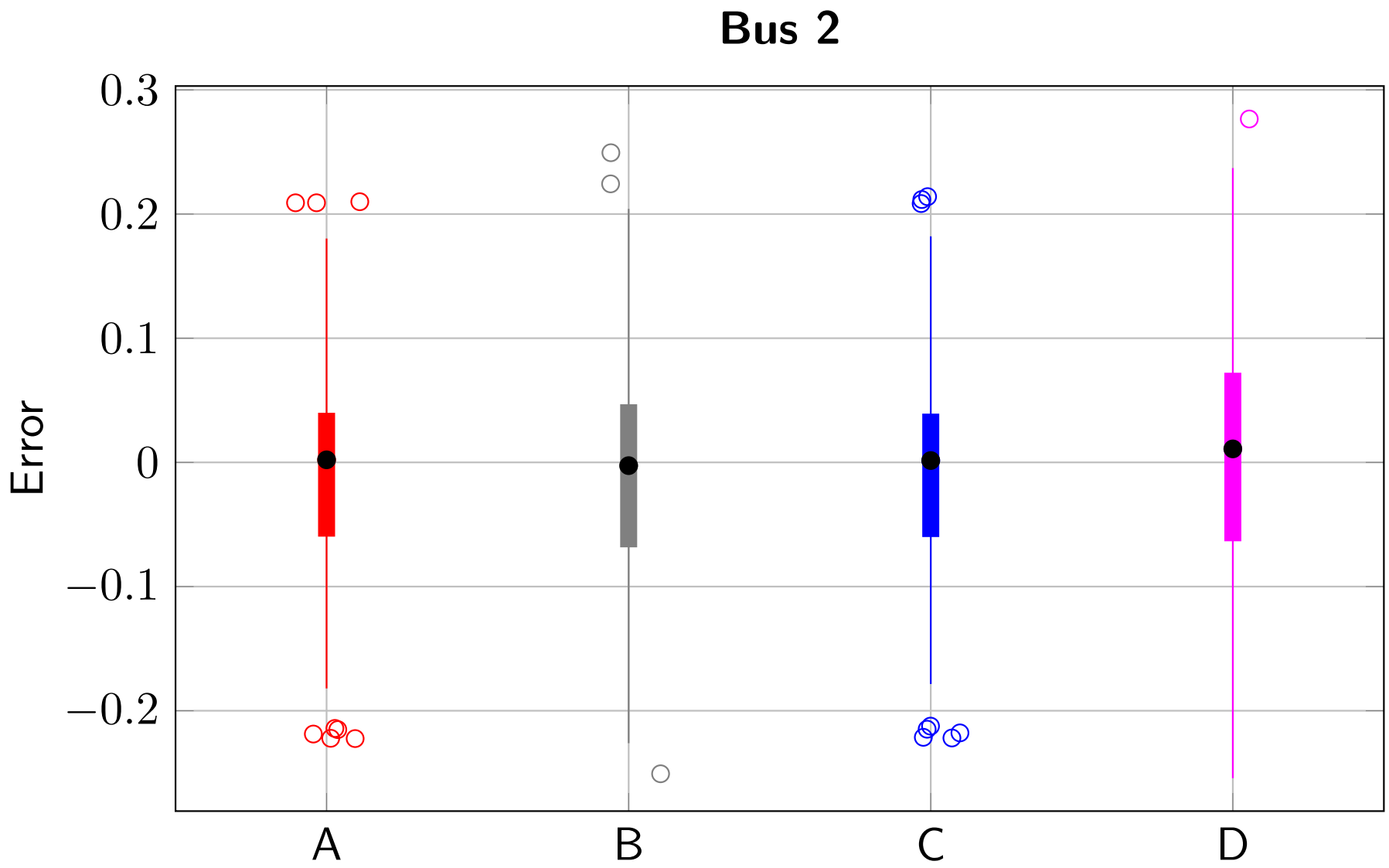
$$\sum_{i\in\mathcal{I}} |\alpha_i H_i u| = \sum_{i=\mathcal{I}^c} |\alpha_i H_i u|.$$

Simulation IEEE 14-bus System

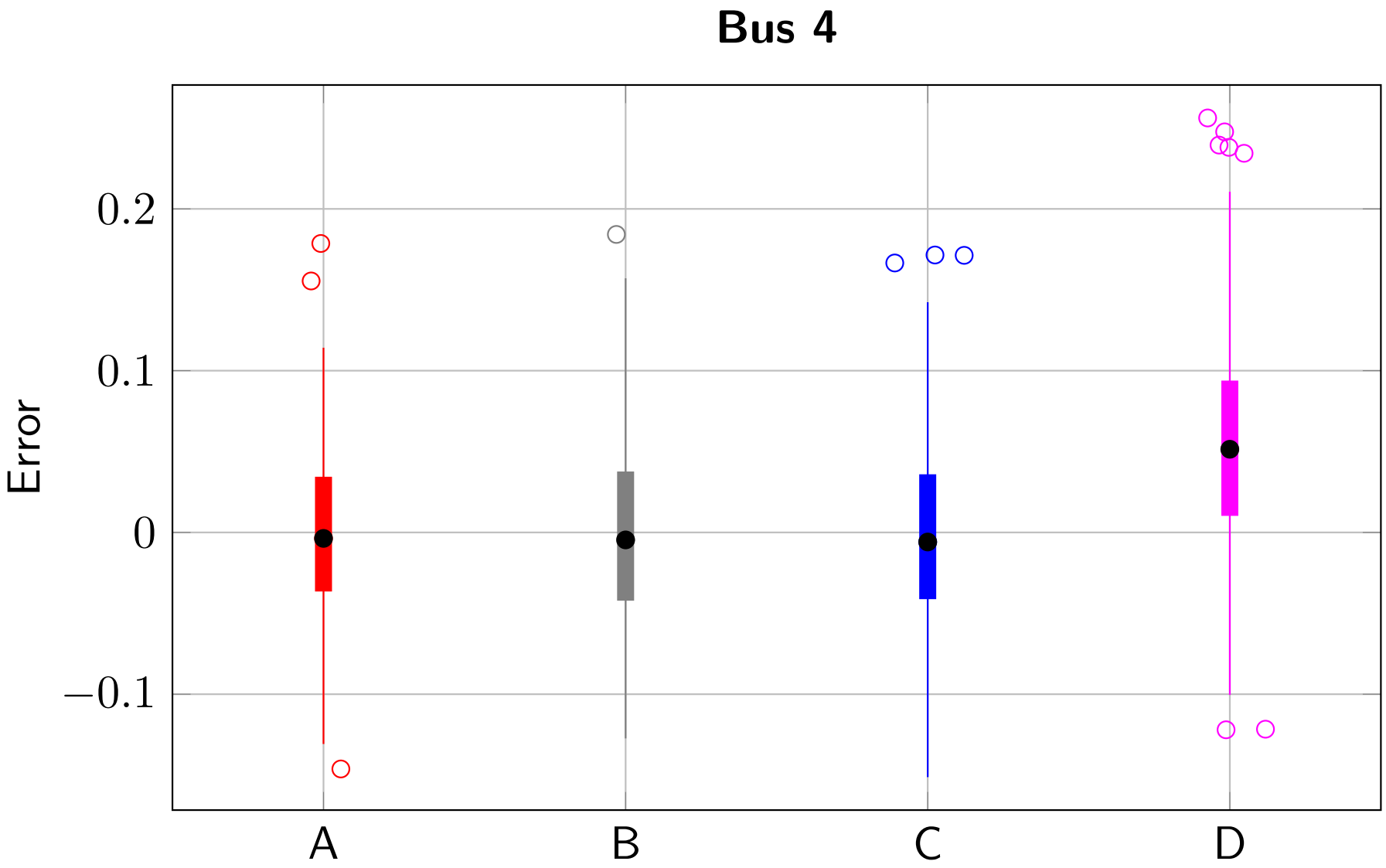


We assume $p = 1$ and the flow sensor on the red line is being attacked.

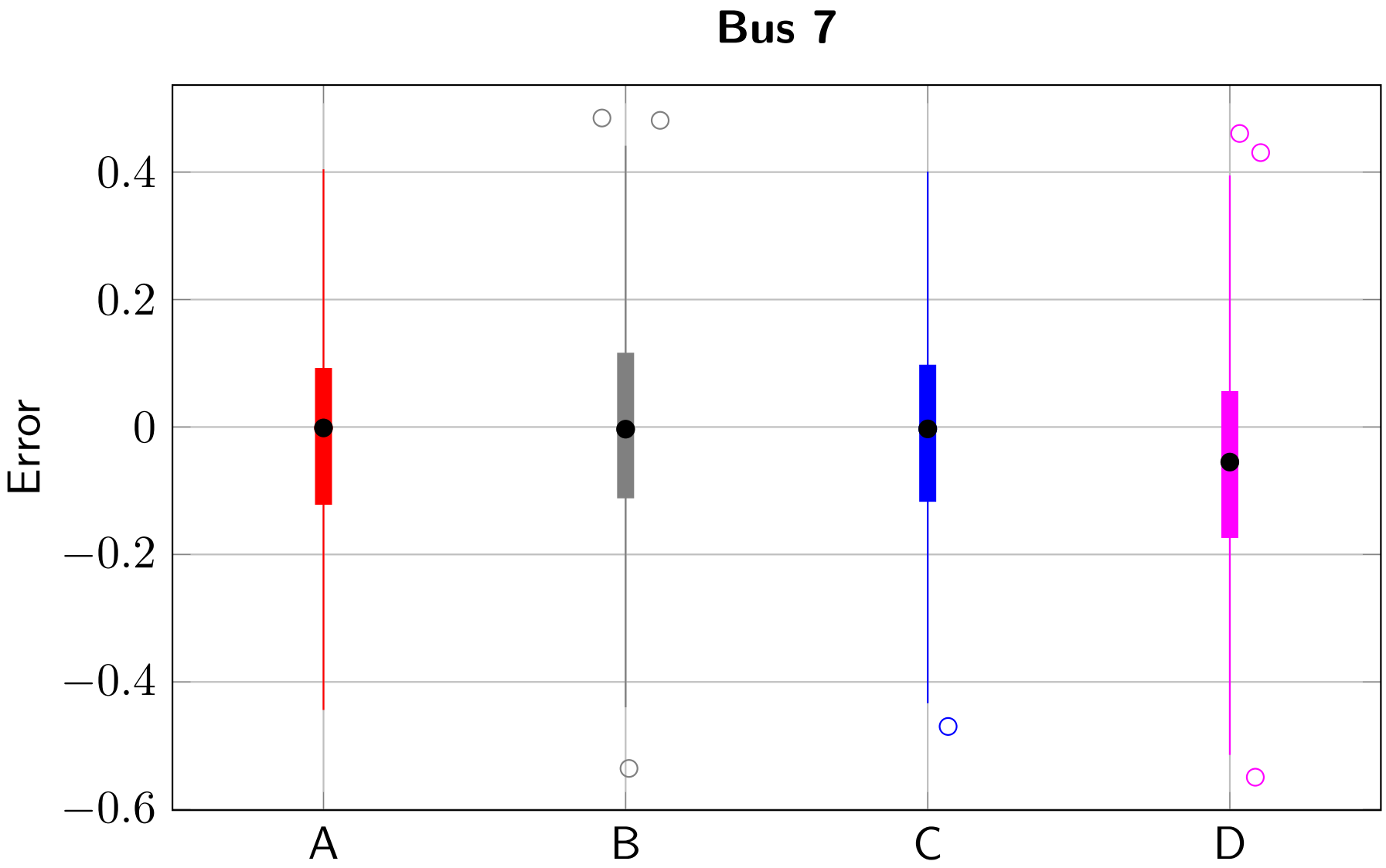
Simulation IEEE 14-bus System



Simulation IEEE 14-bus System



Simulation IEEE 14-bus System



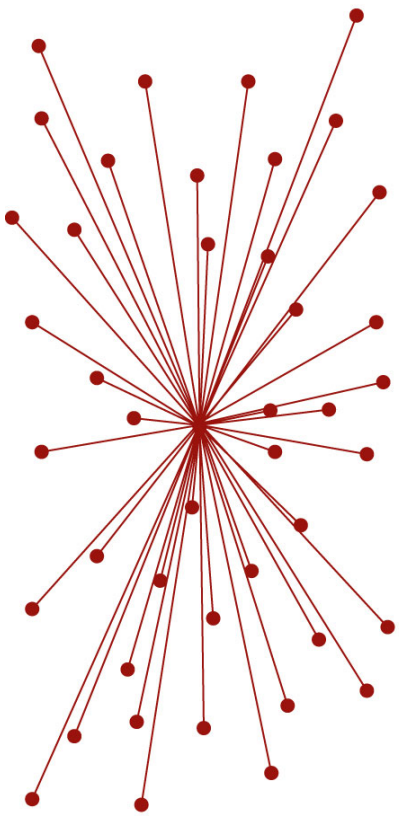
1 Introduction

2 Secure State Estimation

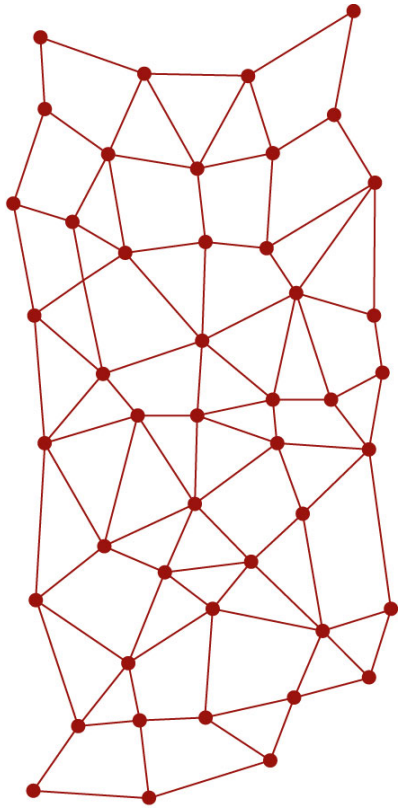
3 Privacy Preserving Consensus

4 Conclusion

Centralized v.s. Distributed Algorithm



Centralized Network



Distributed Network

Benefit of distributed algorithm: Reliability, Scalability, Transparency, Incremental Growth, Privacy

Preliminary: Average Consensus

- We model the network composed of n agents as an *undirected* and *connected* graph $G = \{V, E\}$.
- Define the neighborhood of sensor i as $\mathcal{N}(i)$.
- Each agent has an initial scalar state $x_i(0)$.
- Update equation:

$$x_i(k+1) = a_{ii}x_i(k) + \sum_{j \in \mathcal{N}(i)} a_{ij}x_j(k).$$

- Update equation in matrix form:

$$x(k+1) = Ax(k),$$

where we assume that A is *symmetric*.

Preliminary: Average Consensus

Define the average vector as

$$\bar{x} \triangleq \frac{\mathbf{1}' x(0)}{n} \mathbf{1}.$$

The following conditions are necessary and sufficient for all the agents to reach average consensus:

(A1) $\lambda_1 = 1$ and $|\lambda_i| < 1$ for all $i = 2, \ldots, n$.

(A2) $A\mathbf{1} = \mathbf{1}$, i.e., $\mathbf{1}$ is an eigenvector of A.

Applications of Consensus

- Dynamic load balancing;
- Vehicle formation control;
- Rendezvous problem;
- Social network;
- Clock Synchronization;
- Localization;
- Distributed Estimation;
- Distributed Optimization;
- ...

Challenges

- Malicious Attacker: What if some agent do not follow the update rule:

$$x_i(k+1) = a_{ii}x_i(k) + \sum_{j \in \mathcal{N}(i)} a_{ij}x_j(k).$$

  (Pasqualetti, Bicchi and Bullo, 2012): If there are k malicious agents, then
  - a benign node can detect the presence of malicious agents if the size of the minimum cut of the graph is at least $k+1$.
  - a benign node can identify the set of the malicious agents if the size of the minimum cut of the graph is at least $2k+1$.
- Curious Attacker: What if an agent wants to infer the initial state of other agents?
- Malicious & Curious Attacker?: Separation principle.

Privacy Concerns

- Without loss of generality, we consider agent n wants to estimate the initial states of other agents.
- Denote the neighborhood of agent n as

$$\mathcal{N}(n) = \{j_1, \ldots, j_m\}.$$

- Define

$$C \triangleq \begin{bmatrix} e_{j_1} & \ldots & e_{j_m} & e_n \end{bmatrix}' \in \mathbb{R}^{(m+1)\times n},$$
$$y(k) \triangleq \begin{bmatrix} x_{j_1}(k) & \ldots & x_{j_m}(k) & x_n(k) \end{bmatrix} = Cx(k).$$

- The information set for agent n at time k is

$$\mathcal{I}(k) = \{y(0), \ldots, y(k)\}.$$

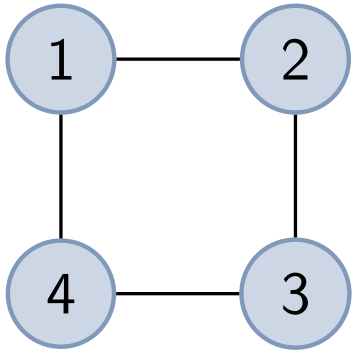
Privacy Concerns

For agent n, the problem becomes a standard estimation problem of the following linear system:

$$x(k+1) = Ax(k),\ y(k) = Cx(k).$$

Since the system is noiseless, if x_i is in the observable space of (A, C), then agent n can perfectly recover the initial state $x_i(0)$ of agent i.
The privacy of agent i is breached.

Example



1. At time step 0, agent 4 knows $x_1(0)$ and $x_3(0)$.
2. At time step 1, agent 4 knows $x_1(1)$, which is

$$x_1(1) = a_{11}x_1(0) + a_{14}x_4(0) + a_{12}x_2(0).$$

For agent 4, it can infer all the initial state $x(0)$ in two steps.

Privacy Preserving Average Consensus

1. At time k, each agent generates a standard normal distributed random variable $v_i(k)$ with mean 0 and variance 1. We assume that all the random variables $\{v_i(k)\}_{i=1,\ldots,n,\, k=0,1,\ldots}$ are jointly independent.
2. Each agent then adds a random noise $w_i(k)$ to its state $x_i(k)$, where

$$w_i(k) = \begin{cases} v_i(0) & \text{, if } k = 0 \\ \varphi^k v_i(k) - \varphi^{k-1} v_i(k-1) & \text{, otherwise} \end{cases},$$

   where $0 < |\varphi| < 1$ is a constant for all agents. Define the new state to be $x_i^+(k)$, i.e.,

$$x_i^+(k) = x_i(k) + w_i(k).$$

3. Each agent then communicates with its neighbors and update its state to the average value, i.e.,

$$x_i(k+1) = a_{ii} x_i^+(k) + \sum_{j \in \mathcal{N}(i)} a_{ij} x_j^+(k).$$

4. There is no requirement for additional communication.

Privacy Preserving Average Consensus

$$
\begin{array}{lccccccc}
k: & 0 & & 1 & & 2 & & 3 \\
w_i(k): & v_i(0) & \longrightarrow & \begin{array}{c}\varphi v_i(1)\\ + \\ -v_i(0)\end{array} & \longrightarrow & \begin{array}{c}\varphi^2 v_i(2)\\ + \\ -\varphi v_i(1)\end{array} & \longrightarrow & \begin{array}{c}\varphi^3 v_i(3)\\ + \\ -\varphi^2 v_i(2)\end{array}
\end{array}
$$

Privacy Preserving Average Consensus

- Update equation in matrix form:

$$x(k+1) = Ax^{+}(k) = A(x(k) + w(k)).$$

- Agent n only receives a noisy version of the state:

$$y(k) = Cx^{+}(k) = C(x(k) + w(k)).$$

Performance Metric

Define the error vector as

$$e(k) \triangleq x(k) - \bar{x}.$$

Define the mean square convergence rate as

$$\rho \triangleq \limsup_{k\to\infty} \left(\sup_{z(0)\neq 0} \frac{\mathbb{E}_v e(k)' e(k)}{e(0)' e(0)} \right)^{1/k},$$

Performance Metric

- The new information set of agent n at time k is

$$\mathcal{I}(k) \triangleq \{x_n(0), y(0), \ldots, y(k)\}.$$

- Denote the maximum likelihood estimate of $x(0)$ given $\mathcal{I}(k)$ as $\hat{x}(0|k)$, the covariance of which is defined as $P(k)$.
- Clearly, $\mathcal{I}(k) \subset \mathcal{I}(k+1)$, which implies $P(k) \geq P(k+1)$. Hence, we can define

$$P = \lim_{k \to \infty} P(k).$$

  P is best estimation performance that can be achieved by agent n (with infinite observations).
- If $P_{ii} = 0$, then agent n can infer the initial state $x_i(0)$ of agent i without any error (asymptotically).

Our goal is to make ρ *small*, while making P *large*.

Convergence Result

Theorem

For any initial condition $x(0)$, $x(k)$ converges to $\bar{x}$ in the mean square sense. Furthermore, the mean square convergence rate ρ equals

$$\rho = \max(|\varphi|^2, |\lambda_2|^2, |\lambda_n|^2).$$

Estimation Performance

Theorem

Suppose that $1 > \varphi > \|\tilde{A}\|$. *P is given by the following equality:*

$$P = \begin{bmatrix} \mathcal{Q}_2 \left[\mathcal{Q}_2' (I - \tilde{A})^{-1} Y (I - \tilde{A})^{-1} \mathcal{Q}_2 \right]^{-1} \mathcal{Q}_2' & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix}$$

where $Y = \lim_{k \to \infty} Y(k)$ *is the limit of the following recursive Riccati equations:*

$$Y(0) = \tilde{A} \mathcal{U} \tilde{A},$$

$$Y(k+1) = \tilde{A} \mathcal{U} \tilde{A} + \varphi^{-2} \tilde{A} \left[Y^{+}(k) - Y^{+}(k) \left(\varphi^2 I + Y^{+}(k) \right)^{-1} Y^{+}(k) \right] \tilde{A},$$

where

$$Y^{+}(k) = \mathcal{V} Y(k) \mathcal{V}.$$

Estimation Performance

Define the essential neighborhood of sensor i as

$$\mathcal{N}_e(i) \triangleq \{j \in V : j \neq i,\ a_{ij} \neq 0\}.$$

Node j is called a super neighbor of i if

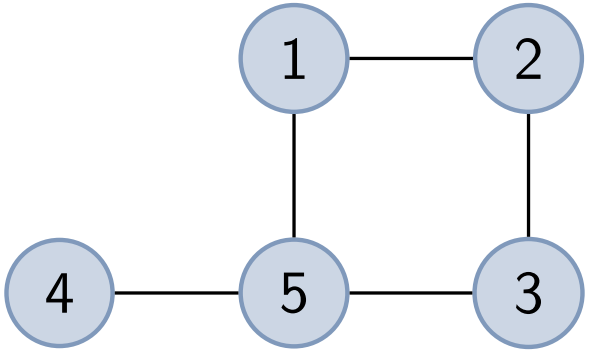
1. j is a neighbor of i.
2. j is also a neighbor of all i's essential neighbors.

Theorem

The ith node's privacy is breached to node j if and only if j is a super neighbor of i.

The condition is can be checked locally. In other words, your initial state can only be leaked to your neighbors.

Estimation Performance



$$\mathcal{N}(5)\bigcup\{5\} = \{1,3,4,5\}.$$

$\mathcal{N}(4)\bigcup\{4\} = \{4,5\}$. Agent 5 can perfectly infer $x_4(0)$.

$\mathcal{N}(1)\bigcup\{1\} = \{1,2,5\}$. Agent 5 cannot perfectly infer $x_1(0)$ if $a_{12} \neq 0$.

Fundamental Limitation

- If $P_{ii} = 0$ using our average consensus algorithm, it is possible to design another noise sequence $\{w_i(k)\}$ to make $P_{ii} \neq 0$?
- Consider a more general noise model:
  1. $\mathbb{E}w_i(k) = 0$.
  2. $\mathbb{E}w_i(k_1)w_j(k_2) = 0$, when $i \neq j$.
- The second condition implies that the sensors are not collaborating.

Fundamental Limitation

Theorem

$x(k)$ converges to $\bar{x}$ in the mean squared sense, i.e.,

$$\lim_{k\to\infty} \mathbb{E}\,\|x(k) - \bar{x}\|^2 = 0,$$

implies that for each agent i:

$$\lim_{k\to\infty} \mathbb{E}\left(\sum_{t=0}^{k} w_i(t)\right)^2 = 0,\ \forall i = 1, \ldots, n.$$

Since the sensors are not collaborating, average consensus is achieved if and only if the noise from each sensor sums to 0 asymptotically (in the mean squared sense).

Fundamental Limitation

Theorem

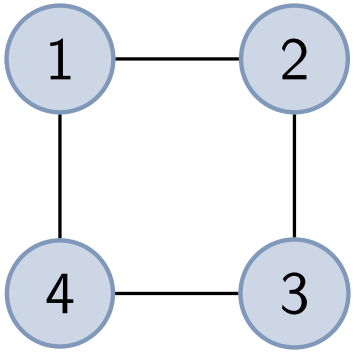
Suppose that

$$\lim_{k\to\infty} \mathbb{E}\left(\sum_{t=0}^{k} w_i(t)\right)^2 = 0, \forall i = 1, \ldots, n.$$

Then ith node's privacy is breached to node j if *j is a super neighbor of i.*

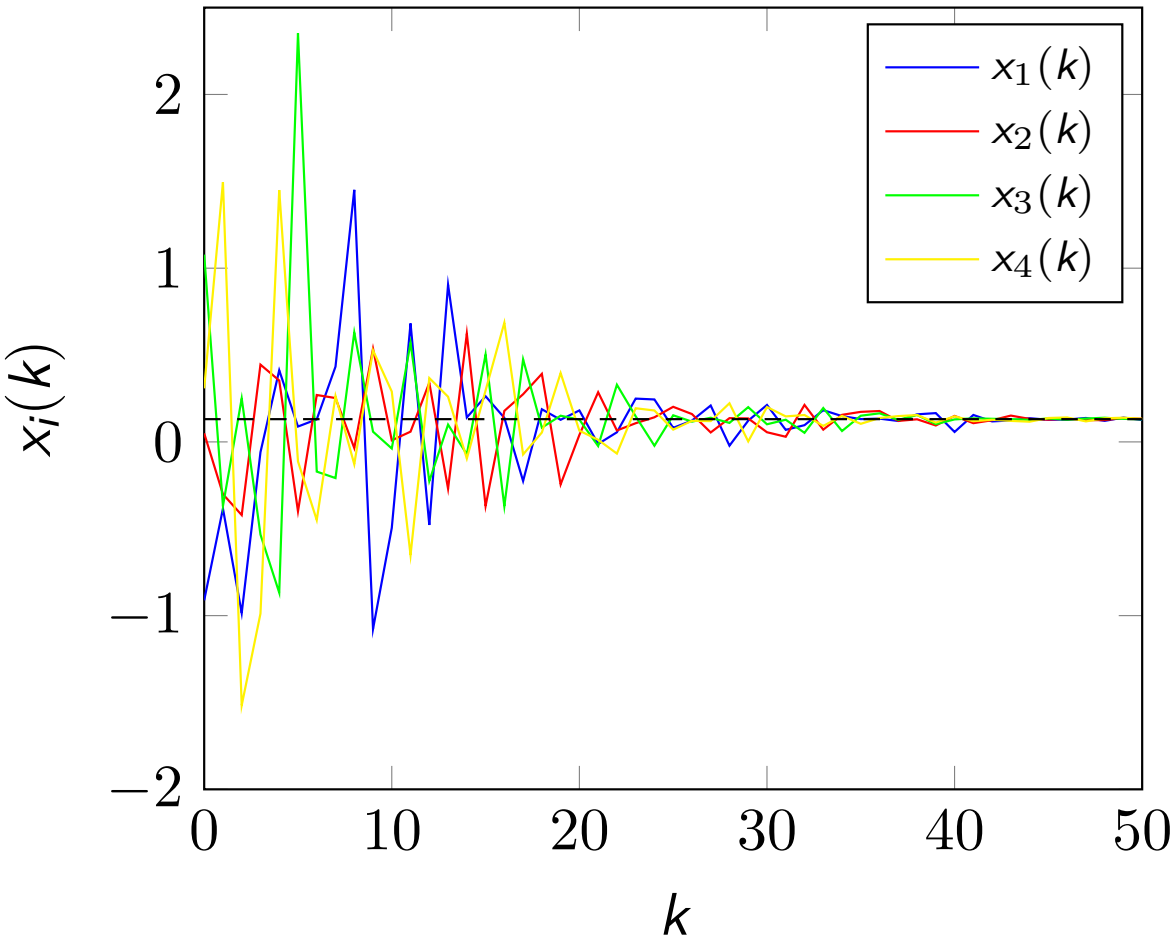
Comparing it to the previous theorem, we know that our proposed strategy achieves "minimum" privacy breach.

Numerical Example

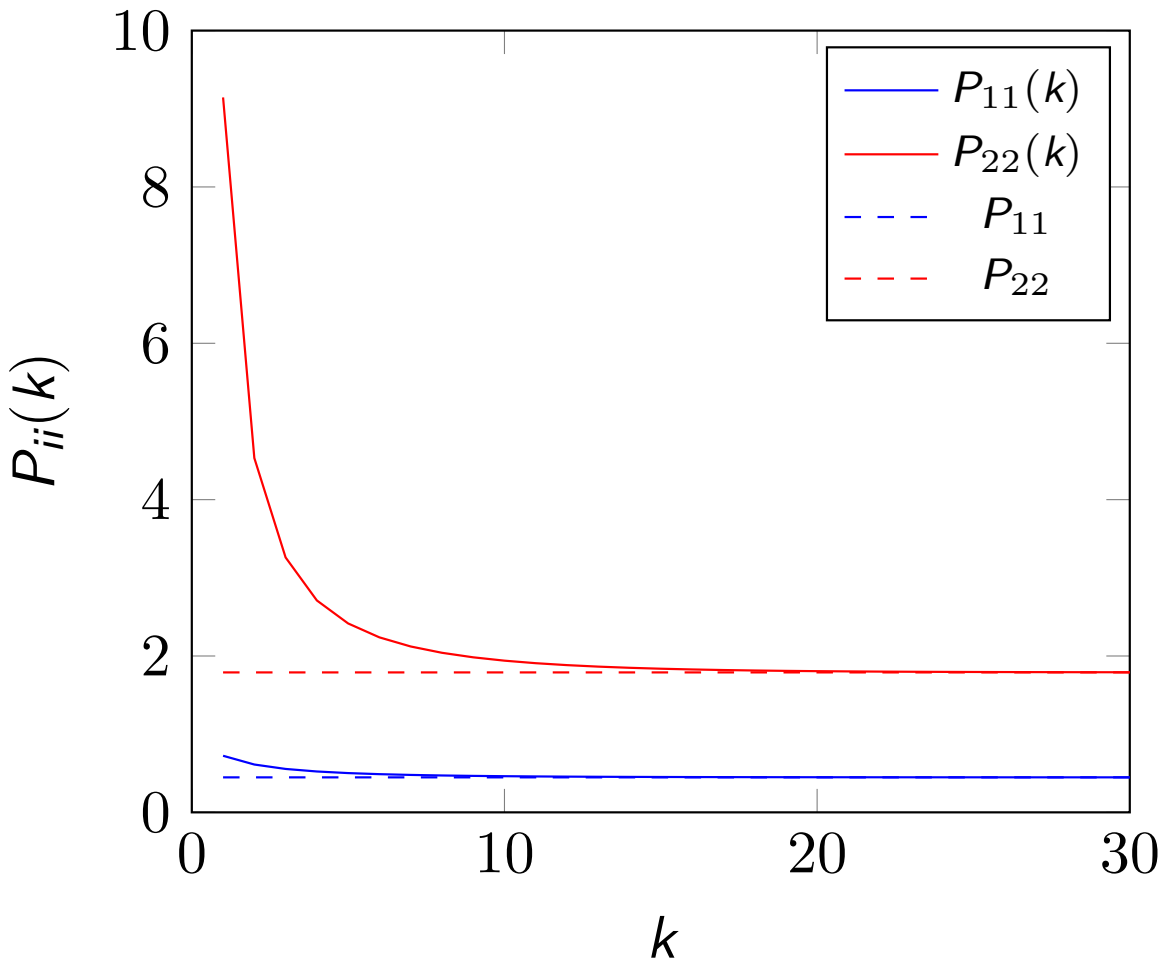


- $a_{ii} = a_{ij} = 1/3$, for $j \in \mathcal{N}(i)$.
- $\|\tilde{A}\| = 0.805$.
- We choose $\varphi = 0.9$.

Numerical Example: Convergence Result



Numerical Example:Estimation Performance



Conclusion

We consider two different information fusion schemes: state estimation and average consensus and design low weight algorithms to address the security and privacy challenges in these settings.

Advertisement

- If you want to learn more about my research, you can check my homepage: `https://yilinmo.github.io`
- I am currently looking for prospective Ph.D. student and Postdoc. If you are interested, please let me know. My email address is ylmo@ntu.edu.sg.
  1. PhD: Tuition will be covered, Stipend is 2000/2500SGD (before/after Qual)
  2. PhD: Need TOEFL or IELTS. NTU will organize an English test if you do not have GRE.
  3. Postdoc: Salary $\sim$ 4500SGD

Thank You!