Secure Distributed State Estimation in Cyber-Physical Systems

Arash Mohammadi

Assistant Professor

UNIVERSITÉ Concordia UNIVERSITY

IEEE SPS Winter School on Secure and distributed Cyber-Physical Systems, Montreal, QC., Canada, November 4th, 2016

Email: arash.mohammadi@concordia.ca
Homepage: https://users.encs.concordia.ca/~arashmoh

Outline of the Talk

1. Behind the Scene!
2. Introduction to Cyber-Physical Systems (CPS).
3. State Estimation with a example of distributed Sensor Selection methodology.
4. Complex-valued Signals and Non-Circular Data Injection Attacks.

Thanks to Organizing Volunteers

Golnar Kalantar

Hamidreza Sadreazami

Somayeh Davar

Ali Al-Dulaimi

Soroosh Shahtalebi

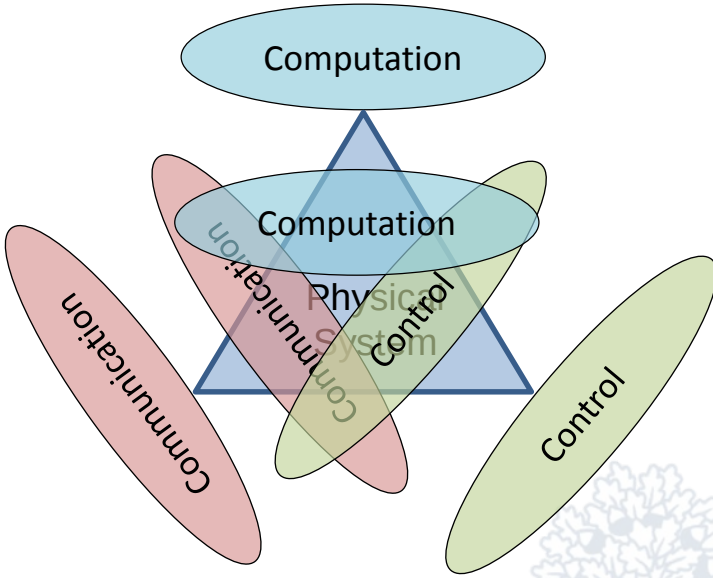
Muhammad Nasir Shafique

Amir Amini

Cyber-Physical Systems (CPS)

CPSs are

- Integration of computation, communication, and control with physical processes.
- The next generation embedded systems.
- CPSs are embedded computers and networks that monitor and control the physical processes, with feedback loops where physical processes affect computations and vice versa.



In CPSs interaction with the physical world is promoted to **"First Class Citizen"**

Industrial CPS
Courtesy of Kuka Robotics Corp.

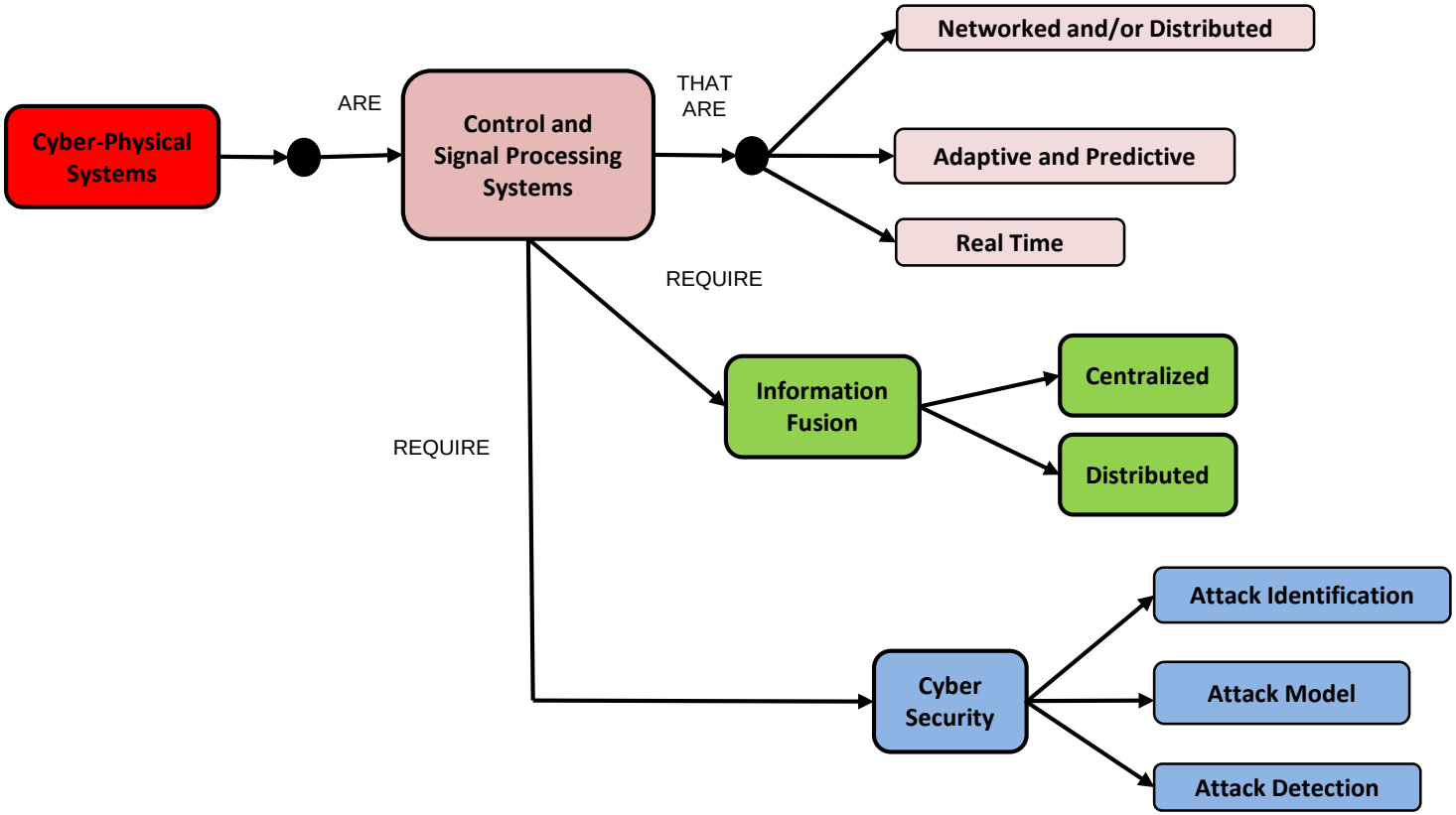
Main Goals

- Co-design the cyber and physical part of the system.
- Engineer a **"System of Systems"**

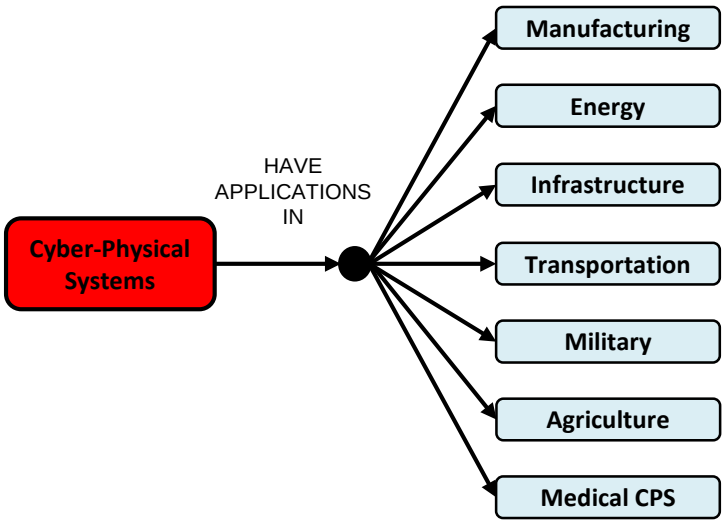
CPS- A Concept Map

Control Systems include algorithms that react to sensor data by issuing control signals via actuators to the physical components.

Signal Processing is an area of systems engineering that deals with measurements of time-varying physical quantities and operations on or analyses of signals.



Applications of CPSs



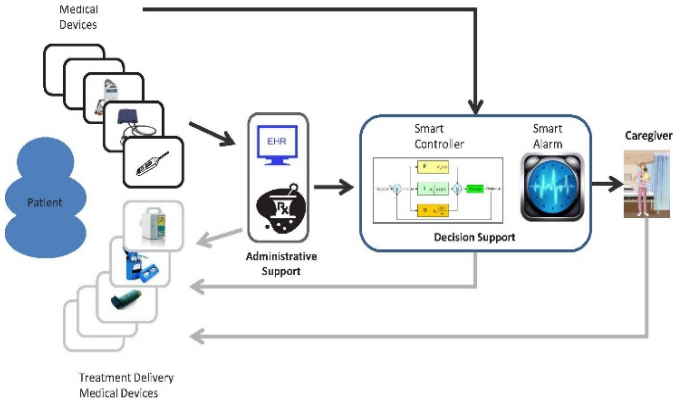
Distributed Robotic Garden

Courtesy of Distributed Robotic Laboratory, CSAIL, MIT

CPS Printing Press

Courtesy of Bosch-Rexroth

Smart Grids

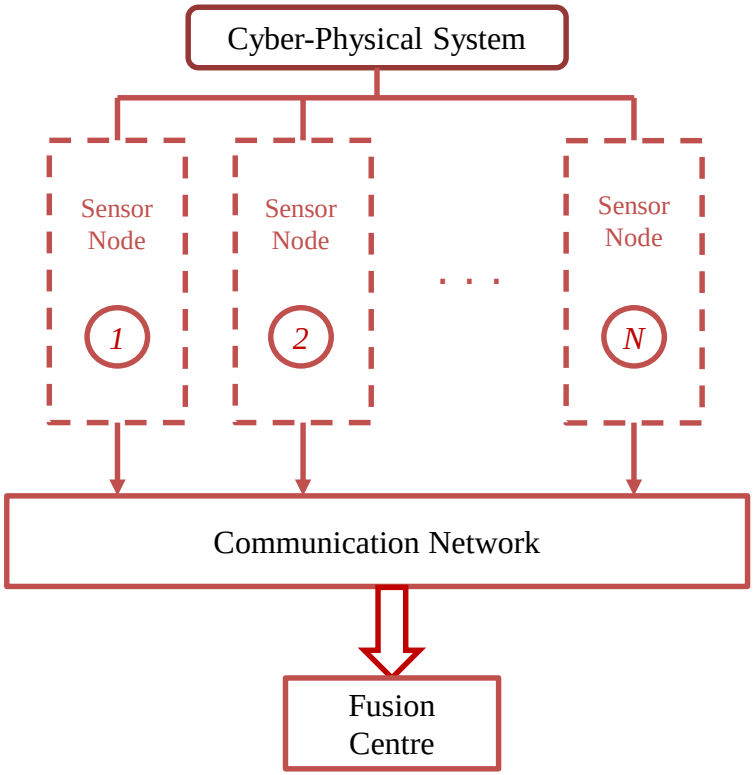


Medical CPS

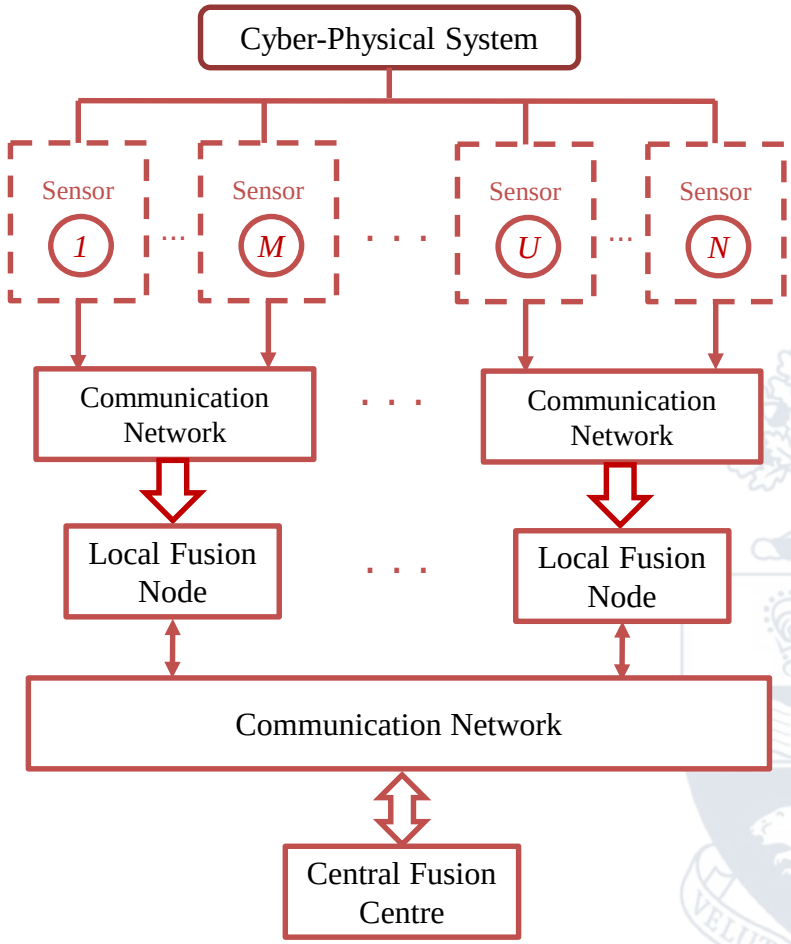
Transportation

Information Fusion

Information Fusion involves the combination of information into a new set of information towards reducing uncertainty.



Centralized Architecture



Hierarchical Architecture

Distributed Information Fusion

Problems of Centralized Information Fusion

- **Single point of failure**: Susceptible to failure with the shut down of the fusion centre.
- **Hot-Spot Problem**: Uneven energy consumption.
- **Routing Issues:** Establish and maintain routing tables.

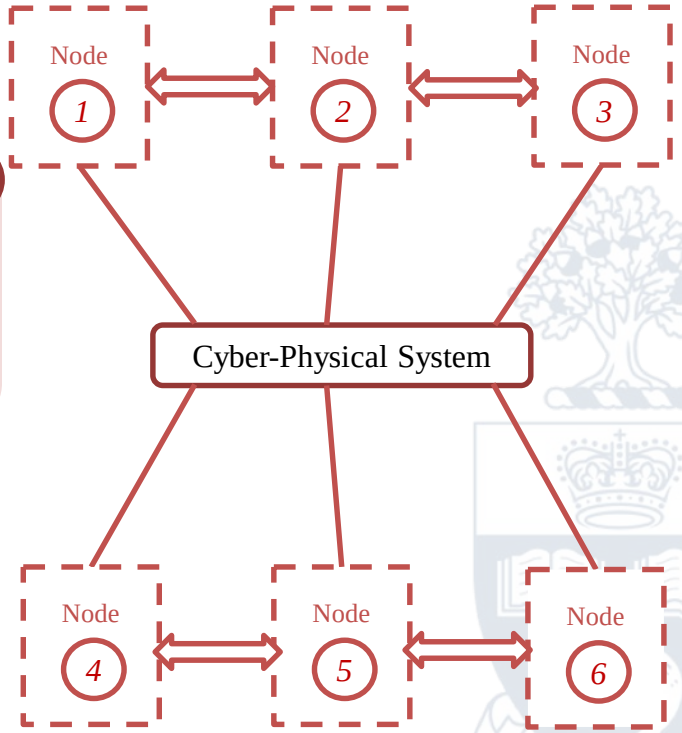
The process of discovering **global** information from **local** interactions among **dispersed** nodes.

Characteristics of Distributed Processing

- There is no fusion center.
- Global knowledge of network topology is not available locally.
- Communication occurs within local neighbourhoods.

Why Distributed Information Fusion?

- **Distributed Nature of Information**: Data available at dispersed locations (e.g., the cloud)
- **Big-Data**: Distributed processing of large data sets.
- **Robustness**
- **Latency**: Dependence on real-time processing.
- **Privacy and Security**



Distributed Architecture

Distributed Signal Processing

Issues of Distributed Processing

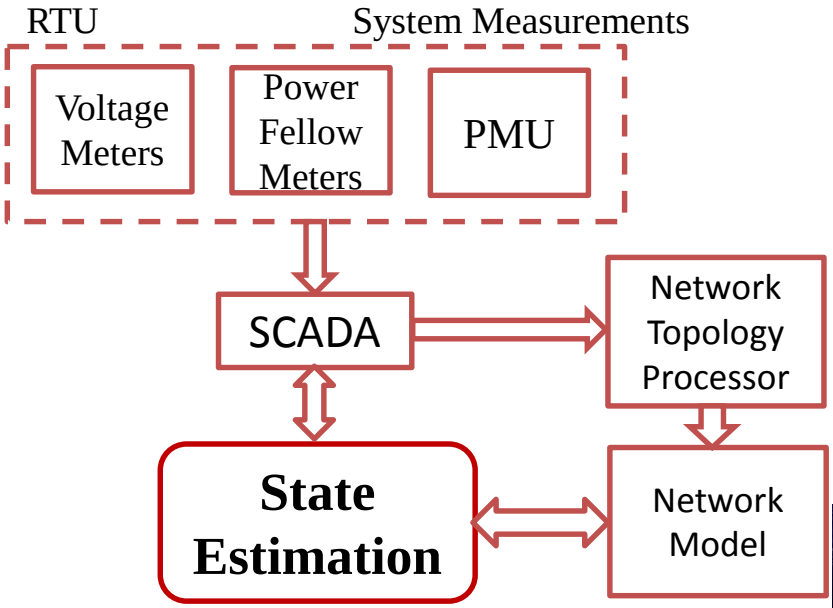
1. Local cooperation and dynamical and adjustment of network topology.
2. Restricted local resources such as power, bandwidth, and processing power.
3. Challenges of cyber protection due to scale and complexity of CPSs along with increased connectivity and automation.
4. CPSs require resilience, adaptability, scalability, and sustainability.
5. Need for communication and energy-efficient methods for distributed estimation.
6. Autonomous operation of mobile, multi-agent systems.

This Winter School: Solutions

1. Adaptation and Learning in Networked Agents (**Prof. Sayed**).
2. Sensor Management and Event-based Control/Estimation (**Prof. Varshney and Prof. Chen**).
3. Cyber-Physical Systems Security and the Smart Grid (**Prof. Kundur**).
4. Distributed and robust optimization algorithms (**Prof. Giannakis**).
5. Distributed Estimation and Tracking in CPSs (**Prof. Coates**).
6. Distributed intelligence in Mobile Multi-agent Network (**Prof. Khan**).

State Estimation in CPSs

The accuracy of state estimation plays a crucial role in computing control commands for efficient and safe operation of CPSs.

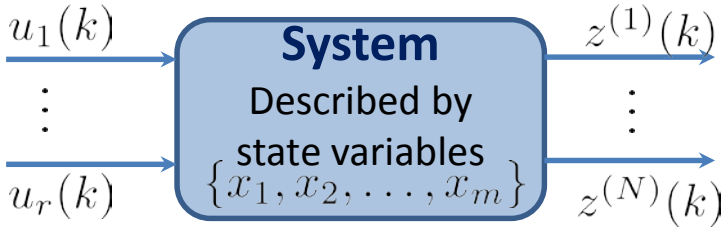


State Estimation Block diagram in Smart Grids

Northeast Blackout 2003

Autonomous vehicles

State-Space Model in CPSs



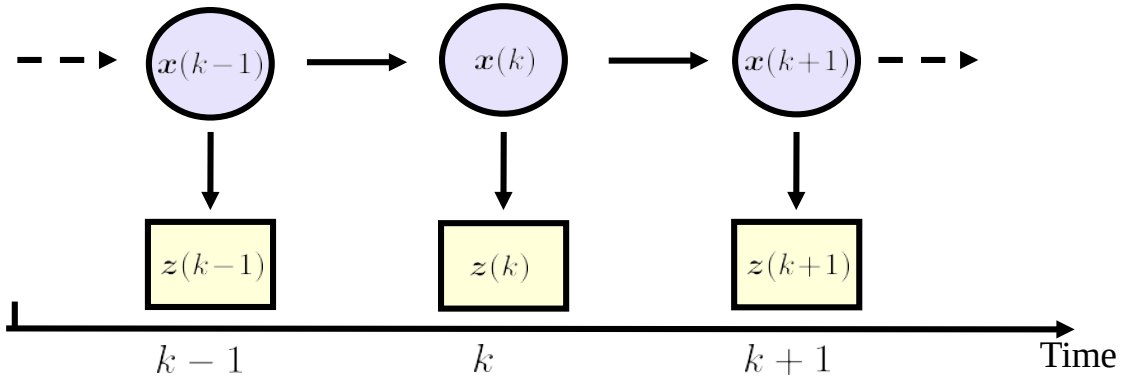
State Vector: A set of variables that describe the system and its response to a given set of inputs.

State Estimation: The process of inferring the state values from the observations.

System Model:
$$\boldsymbol{x}(k) = \boldsymbol{f}(\boldsymbol{x}(k-1)) + \boldsymbol{w}(k)$$

Observation Model:
$$\underbrace{\begin{bmatrix} \boldsymbol{z}^{(1)}(k) \\ \vdots \\ \boldsymbol{z}^{(N)}(k) \end{bmatrix}}_{\boldsymbol{z}(k)} = \underbrace{\begin{bmatrix} \boldsymbol{g}^{(1)}(\boldsymbol{x}(k)) \\ \vdots \\ \boldsymbol{g}^{(N)}(\boldsymbol{x}(k)) \end{bmatrix}}_{\boldsymbol{g}(\boldsymbol{x}(k))} + \underbrace{\begin{bmatrix} \boldsymbol{v}^{(1)}(k) \\ \vdots \\ \boldsymbol{v}^{(N)}(k) \end{bmatrix}}_{\boldsymbol{v}(k)},$$

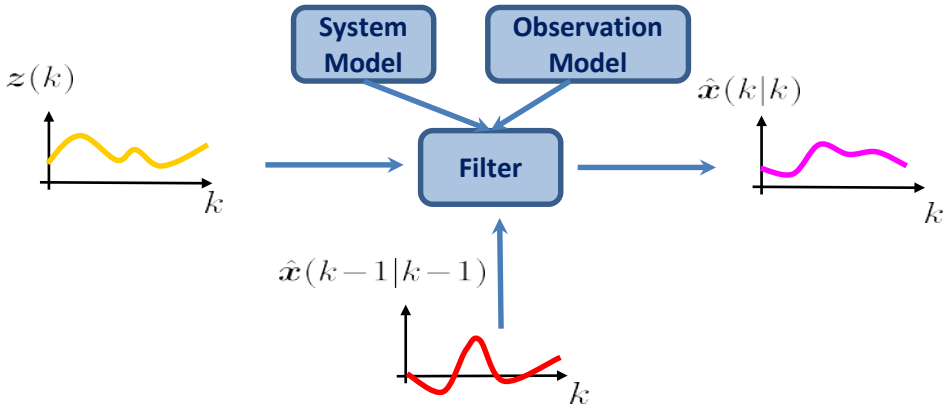
Dynamic System



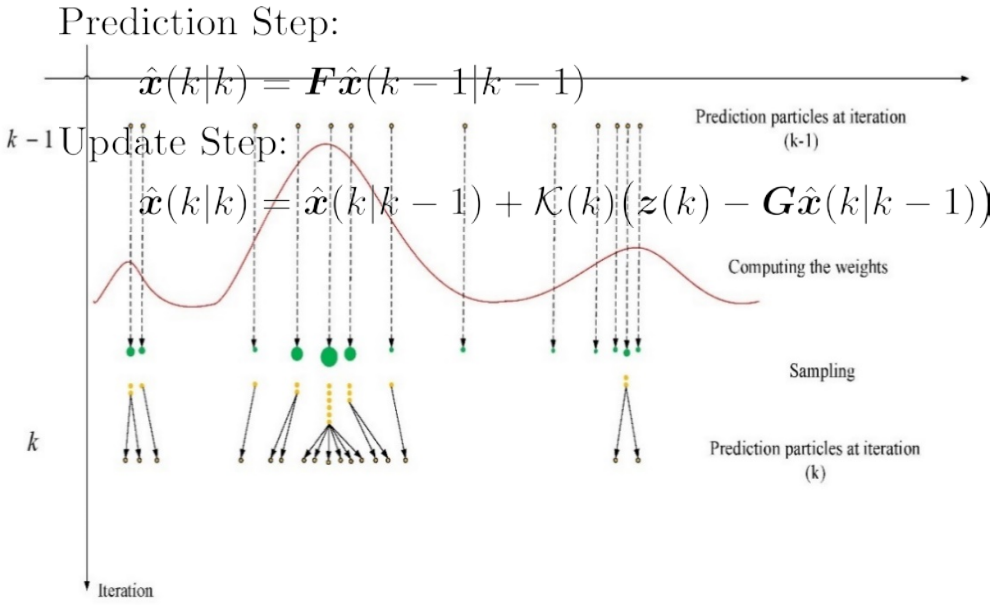
Filtering Solutions

Linear Systems

Kalman Filter: Systems with Linear dynamics and Gaussian uncertainties.



$$\text{System Model}: \quad \boldsymbol{x}(k) = \boldsymbol{F}\boldsymbol{x}(k-1) + \boldsymbol{w}(k)$$
$$\text{Observation Model}: \quad \boldsymbol{z}(k) = \boldsymbol{G}\boldsymbol{x}(k) + \boldsymbol{v}(k)$$



Nonlinear Systems

Particle Filter: General systems with Non-linear dynamics and Non-Gaussian uncertainties.

Prediction Step:

$$\mathbb{X}_i(k) \sim p\big(\boldsymbol{x}(k)|\boldsymbol{x}(k-1)\big)$$

Update Step:

$$W_i(k) \propto W_i(k-1)p\big(\boldsymbol{z}(k)|\mathbb{X}_i(k)\big)$$

State Estimate:

$$\hat{\boldsymbol{x}}(k|k) = \sum_{i=1}^{N_s} W_i(k)\mathbb{X}_i(k)$$

Distributed Sensor Selection

Sensor Selection

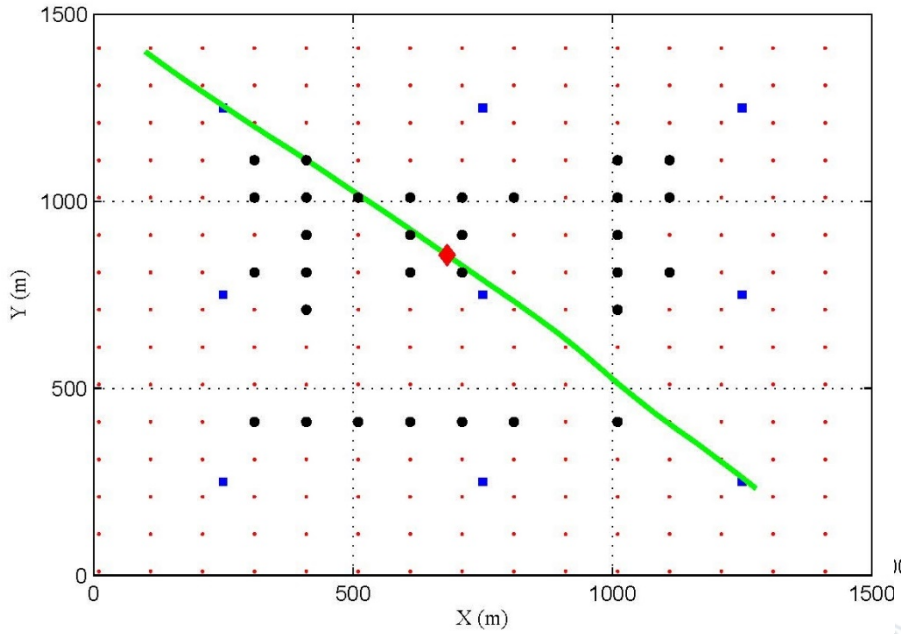
Dynamical activation of sensor nodes within a sensing task in CPSs.

- Random
- Nearest Neighbour

Adaptive Sensor Selection

A stochastic problem that involves optimization of a pre-defined cost function.

- The overall cost function is based on the PCRLB.
- The Posterior Cramer-Rao lower bound (PCRLB) is the performance that an optimal estimator achieves.
- **PCRLB** is an effective sensor selection criteria
  - It provides a measure of the achievable optimal performance.
  - It can be calculated predictively.
  - It is independent of the estimation methodology employed.



Random/Selected Sensors

Optimal Expression for dPCRLB

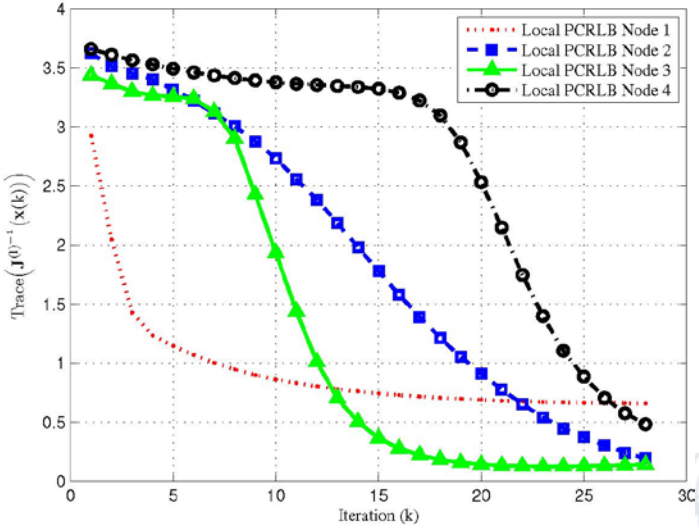
The mean square error (MSE) of the estimate of the state variables is lower bounded by

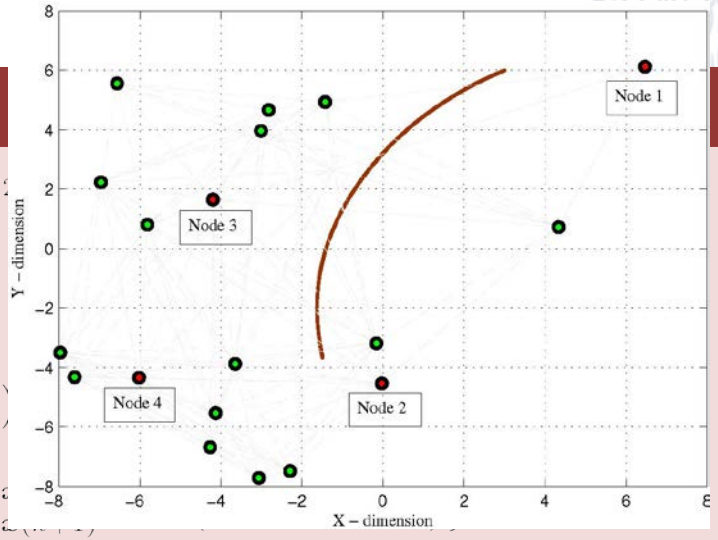
$$\mathbb{E}\Big\{\big(\hat{\boldsymbol{x}}(k)-\boldsymbol{x}(k)\big)\big(\hat{\boldsymbol{x}}(k)-\boldsymbol{x}(k)\big)^{T}\Big\}\geq\big[\boldsymbol{J}(\boldsymbol{x}(k))\big]^{-1}$$

A form of the FIM is defined as follows

$$\boldsymbol{J}\big(\boldsymbol{x}(k)\big)=\mathbb{E}\Big\{-\Delta^{\boldsymbol{x}(k)}_{\boldsymbol{x}(k)}\log p\big(\boldsymbol{x}(k)|\boldsymbol{z}(1:k)\big)\Big\}.$$

Without implementing a fusion centre, distributed algorithms are required to compute the PCRLB.





dPCRLB Theorem (Mohammadi & Asif, AES:2014)

$$\boldsymbol{J}\big(\boldsymbol{x}(k+1)\big)=\boldsymbol{C}^{22}(k)-\boldsymbol{C}^{21}(k)\big(\boldsymbol{J}\big(\boldsymbol{x}(k)\big)+\boldsymbol{C}^{11}(k)\big)^{-1}\boldsymbol{C}^{12}$$

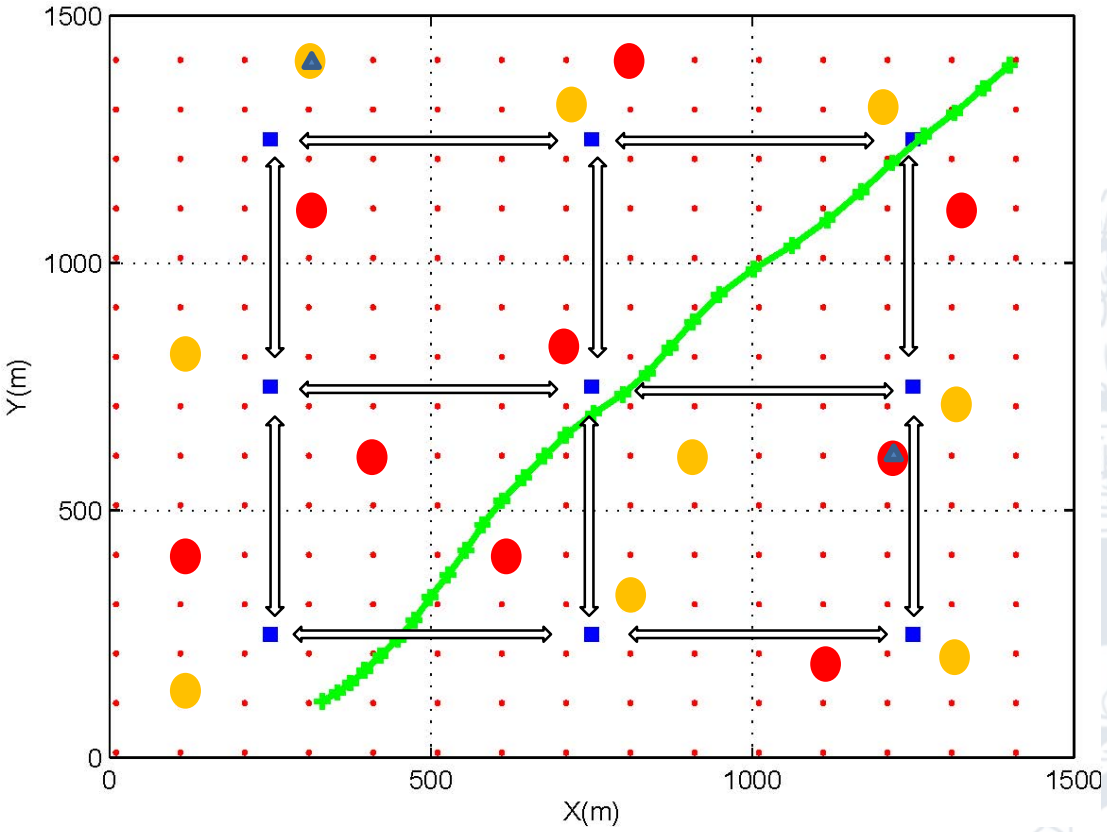
$$\begin{aligned}\boldsymbol{C}^{11}(k) &= \mathbb{E}\big\{-\Delta^{\boldsymbol{x}(k)}_{\boldsymbol{x}(k)}\log p\big(\boldsymbol{x}(k+1)|\boldsymbol{x}(k)\big)\big\},\\ \boldsymbol{C}^{12}(k) &= \big[\boldsymbol{C}^{21}(k)\big]^{T}=\mathbb{E}\big\{-\Delta^{\boldsymbol{x}(k+1)}_{\boldsymbol{x}(k)}\log p\big(\boldsymbol{x}(k+1)|\boldsymbol{x}(k)\big)\\ \boldsymbol{C}^{22}(k+1) &= \sum_{l=1}^{N}\boldsymbol{J}^{(l)}(\boldsymbol{x}(k))-\sum_{l=1}^{N}\boldsymbol{J}^{(l)}(\boldsymbol{x}(k+1|k))+\mathbb{E}\big\{-\Delta^{\boldsymbol{x}}\end{aligned}$$

Decentralized Sensor Selection

Sensor Selection (Mohammadi & Asif, SIGPRO:2015)

- Observation node selection is carried out in several iterations.
- During each iteration, first the best sensor for each fusion centre is picked. Then fusion nodes cooperate to select a sensor based on their previous selection decisions.

1. Initial Sensor Selection
2. Fusion to fusion collaboration
3. Select one sensor
4. Select the second sensor based on the selected sensors
5. Fusion to fusion collaboration
6. Select a sensor



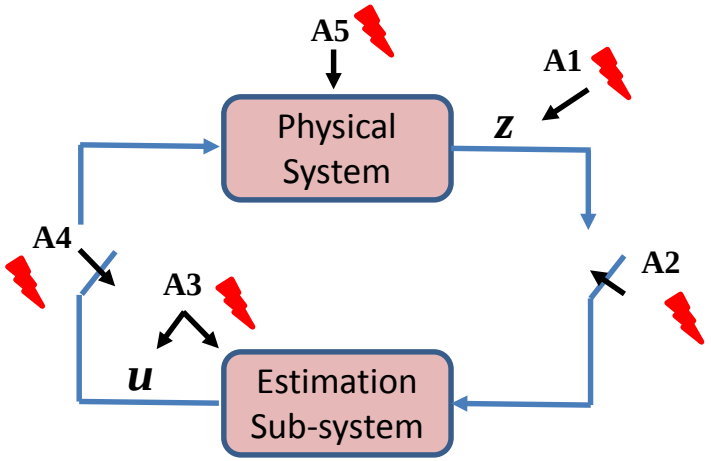
Integrity (A1 and A3)

1. Trustworthiness of sensor and control data packets.
2. Lack of integrity results in deception.

Availability (A2 and A4)

1. Ability of system components on being accessible.
2. Lack of availability results in denial of service (DoS) of sensor and control data.

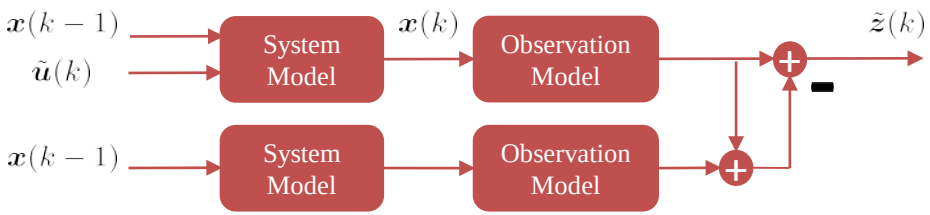
North Pole Toys



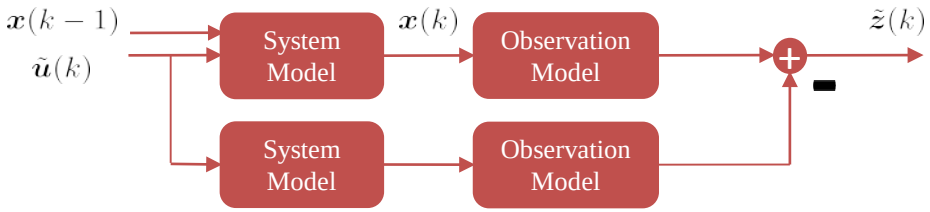
Power grid

Cyber Attacks on Estimation Sys

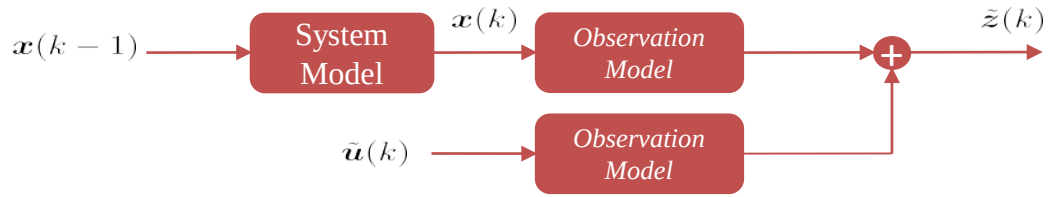
Replay Attack:



Covert Attack:

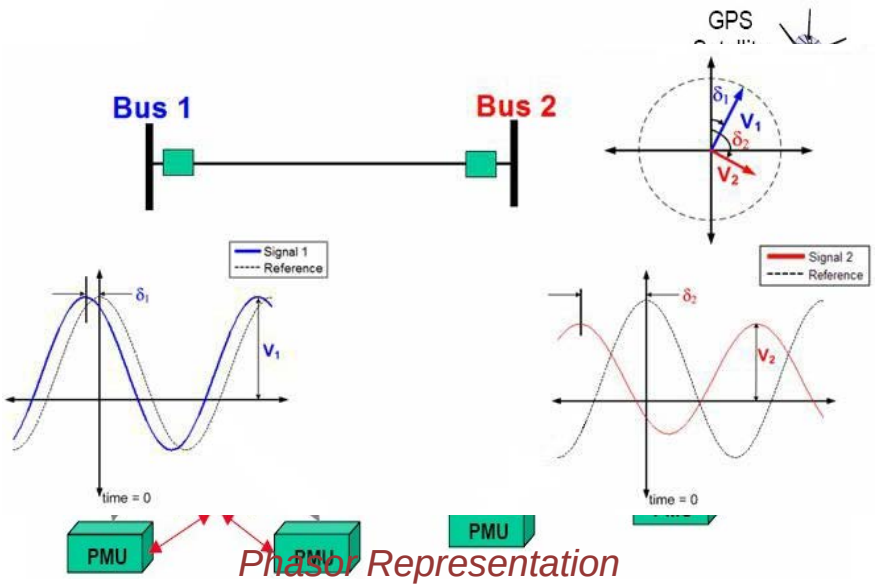


False Data Injection Attack:



False Data Injection Attacks in Smart Grid

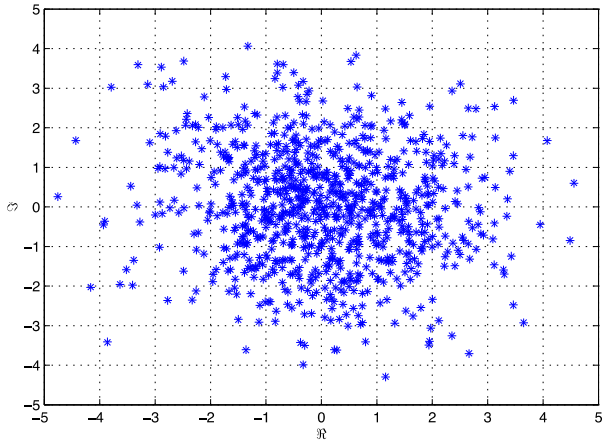
- Phasor Measurement Units (PMU) offer a new paradigm for state estimation in Smart Grid's monitoring system.
- PMU is considered to be the technology which makes the dream of smart grids, a reality.
- A PMU reports complex numbers (phasor representing)

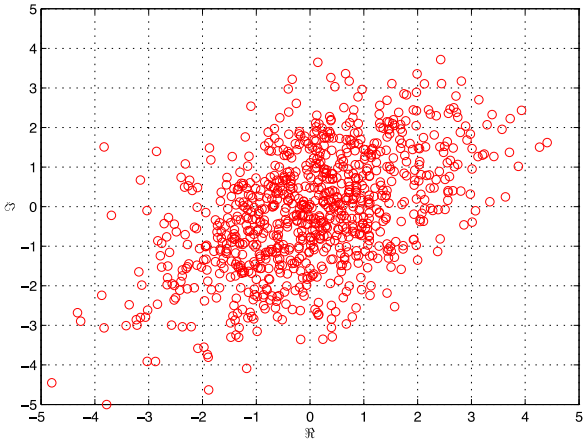


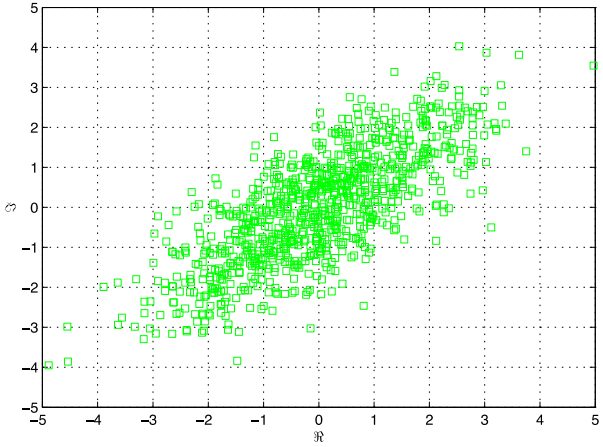
Phasor Representation

Smart Grid Monitoring Unis PMUs

Non-circular Signals







Complex-Valued Signals

- A complex-valued signal is non-circular provided that it is correlated with its complex conjugate, i.e., the pseudo-covariance matrix does not vanish.
- A non-circular signal has two specific properties:
  - ❑ The power of the signal is not equally distributed between its real and imaginary parts.
  - ❑ There exists a correlation between the real and imaginary parts of the signal.
- In this figure scatter plots of complex Gaussian random variables are illustrated where from left to right degree of non-circularity increases from 0 to 0.5.

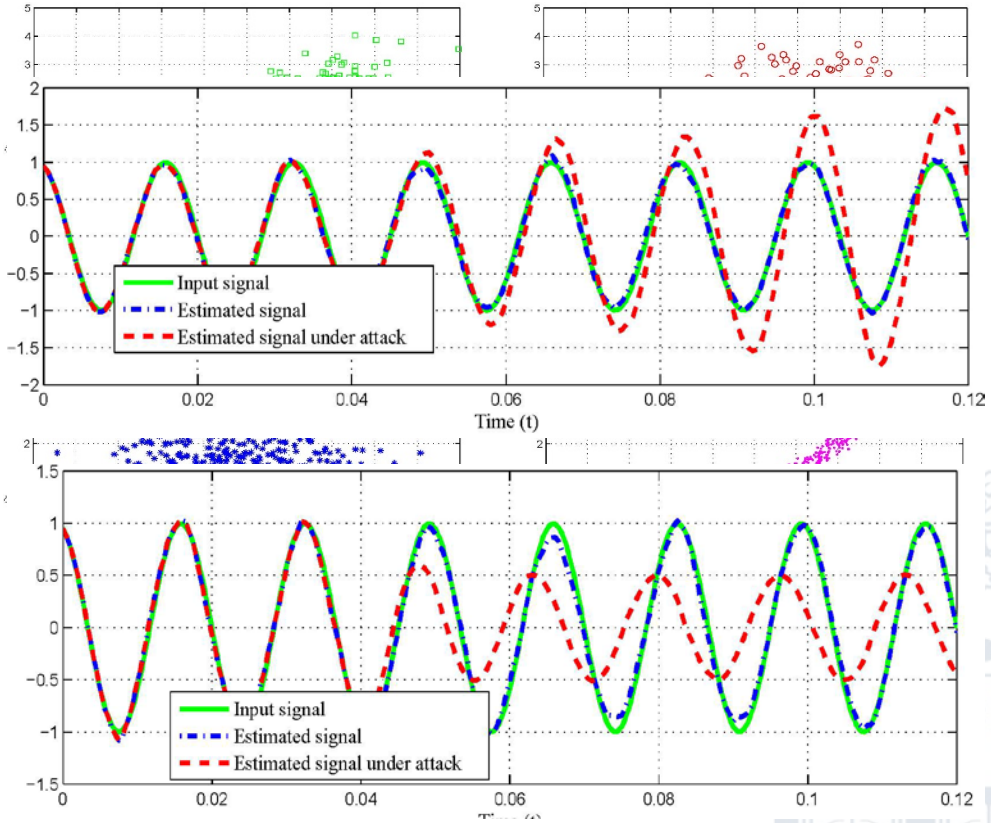
Complex-Valued Attacks

Non-Circular (Improper) Signal

- A complex-valued signal is correlated with its complex conjugate.
- Does not have statistical properties similar to real signals.
- Second-order statistical properties are represented by conventional covariance matrix and pseudo-covariance matrix.
- The processing model should depend on both the signal itself and its complex conjugate.

Non-Circular Attacks

- Introduce correlations between the real and imaginary parts of the signal.
- Need to design advanced models.
- Need to design advanced detectors.



Non-circular Attacks

Non-circular Attacks

Non-Circular Attack Model

- The proposed complex-valued attack adds non-circularity to the measurements via widely-linear transformation, i.e.,

$$z_k^a = D_k z_k + \tilde{D}_k z_k^*,$$

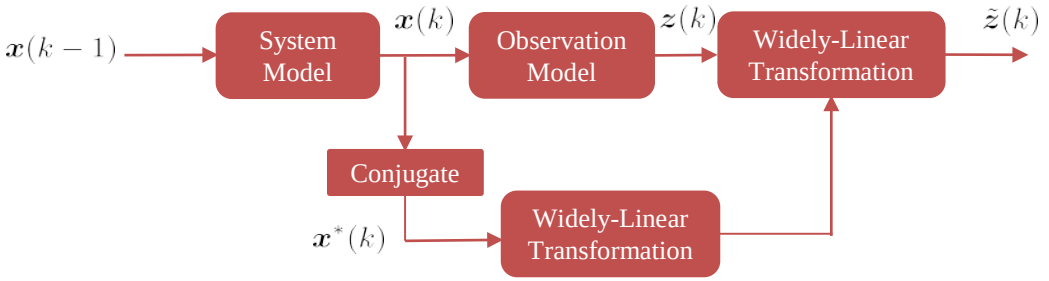
- Attacked observation is given by

$$z_k^a = \begin{bmatrix} D_k H_k & \tilde{D}_k H_k^* \end{bmatrix} \begin{bmatrix} x_k \\ x_k^* \end{bmatrix} + \begin{bmatrix} D_k & \tilde{D}_k \end{bmatrix} \begin{bmatrix} v_k \\ v_k^* \end{bmatrix}.$$

- This attack injects critical statistical information by modifying both the covariance and pseudo-covariance of the innovation sequence, i.e.,

$$\begin{aligned} S_k^a &= [H_k^a]^H P_{k|k-1} H_k^a + [\tilde{H}_k^a]^T P_{k|k-1} [\tilde{H}_k^a]^* + R_k^a \\ \tilde{S}_k^a &= [H_k^a]^H P_{k|k-1} \tilde{H}_k^a + [\tilde{H}_k^a]^T P_{k|k-1} [H_k^a]^* + \tilde{R}_k^a, \end{aligned}$$

Complex-valued Attack:



Non-circular Attack Detection

Conventional Approach

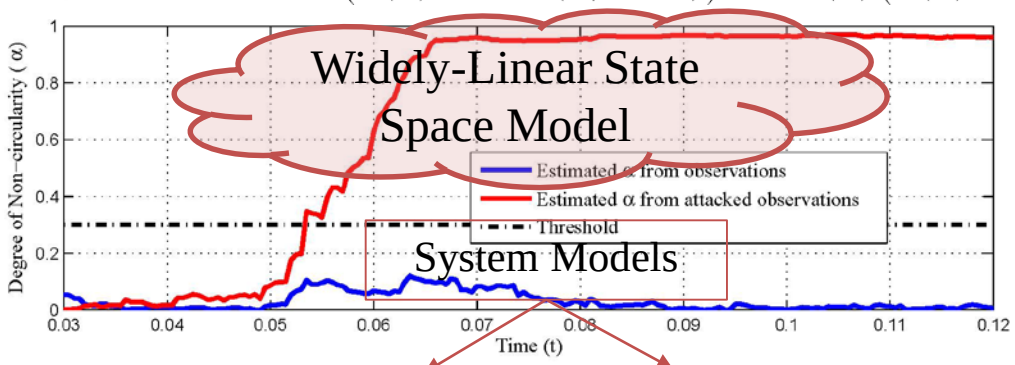
No to Linear State-Space Models

- Widely-Linear State-Space Model
- Augmented State-Space Model

No to Conventional Detection Algorithms

- Non-circularity coefficient.
- Transform detection problem to comparing distance between Gaussian distributions

$$\chi^2\text{- detection} \rightarrow \big(\boldsymbol{z}(k) - \boldsymbol{H}\boldsymbol{x}(k|k-1)\big)^T \boldsymbol{S}^{-2}(k)\big(\boldsymbol{z}(k) - \boldsymbol{H}\boldsymbol{x}(k|k-1)\big)$$



Widely-Linear State Space Model

System Models

$$\begin{aligned}\boldsymbol{x}(k) &= \boldsymbol{F}\boldsymbol{x}(k-1) + \tilde{\boldsymbol{F}}\boldsymbol{x}^*(k-1) + \boldsymbol{w}(k)\\ \boldsymbol{z}(k) &= \boldsymbol{G}\boldsymbol{x}(k) + \tilde{\boldsymbol{G}}\boldsymbol{x}^*(k) + \boldsymbol{v}(k)\end{aligned}$$

Non-circular Detection Algorithms

Observation Models

(Mohammadi & Plataniotis, TNNLS:2015)

Attack Detection

Detection via Statistical Distance Measure

- Compare implemented filter with the theoretical one.
- Compare the innovation sequences in distribution using Statistical Distance Measures.
- Detection is based on Bhattacharyya distance (BD) as:

$$d_{\mathrm{B}}(\underline{\boldsymbol{\nu}}^o, \underline{\boldsymbol{\nu}}_k^a) \underset{H_1}{\overset{H_0}{\lessgtr}} \tau.$$

- The BD between Complex-valued Innovations:

$$d_{\mathrm{B}}(\underline{\boldsymbol{\nu}}^o, \underline{\boldsymbol{\nu}}_k^a) = \frac{1}{4}\Big[\Re\Big\{(\boldsymbol{r}^a - \boldsymbol{r}_k^o)^H \boldsymbol{\Lambda}_k (\boldsymbol{r}^a - \boldsymbol{r}_k^o)\Big\}$$

$$+\Re\Big\{(\boldsymbol{r}^a - \boldsymbol{r}_k^o)^H \tilde{\boldsymbol{\Lambda}}_k ([\boldsymbol{r}^a]^* - [\boldsymbol{r}_k^o]^*)\Big\}\Big] + \underbrace{\frac{1}{2}\ln \frac{|\underline{\boldsymbol{\Gamma}}_k|}{\sqrt{|\underline{\boldsymbol{S}}^a||\underline{\boldsymbol{S}}_k^o|}}}_{\beta_k};$$

$$\text{where} \quad \underline{\boldsymbol{\Gamma}}_k = \frac{1}{2}\big(\underline{\boldsymbol{S}}^a + \underline{\boldsymbol{S}}_k^o\big) \triangleq \begin{bmatrix} \boldsymbol{\Gamma}_k & \tilde{\boldsymbol{\Gamma}}_k \\ \tilde{\boldsymbol{\Gamma}}_k^* & \boldsymbol{\Gamma}_k^* \end{bmatrix},$$

$$\text{and} \quad \underline{\boldsymbol{\Lambda}}_k \triangleq \quad \underline{\boldsymbol{\Gamma}}_k^{-1} = \begin{bmatrix} \boldsymbol{\Lambda}_k & \tilde{\boldsymbol{\Lambda}}_k \\ \tilde{\boldsymbol{\Lambda}}_k^* & \tilde{\boldsymbol{\Lambda}}_k^* \end{bmatrix}.$$

Summary Data

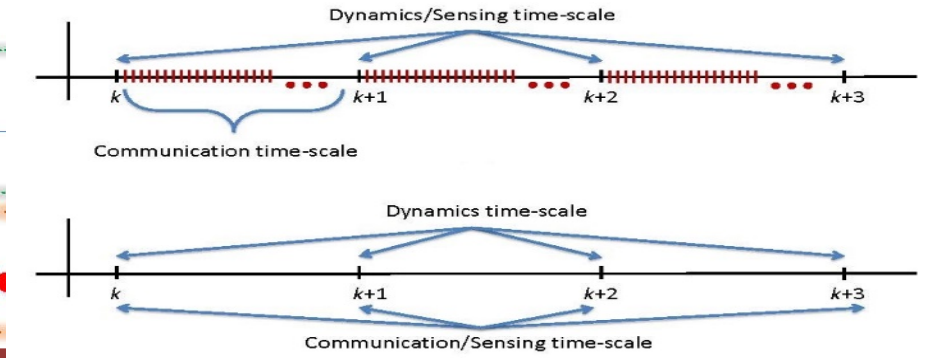
Availability of “big” monitoring data on CPSs brings significant challenges and opportunities to the traditional modeling approaches.

Reduced-order Distributed Estimation

- Reduced-order state-space model.
- Break the global system into several tightly coupled lower dimensional sub-systems.

Diffusive Strategies

- Sensing time-scale and dynamics time-scale are equal.



Social Information Fusion

Event driven Distributed Strategies

- The fusion of social media data with traditio
ement
urces.
ut
potentially fertile, territory.

Distributed Fusion with Structural Uncertainties.