浙江大学电气工程学院
College of Electrical Engineering Zhejiang University

网络安全导论

电力工控系统安全

1. 概述、基础知识
2. 加密与认证技术
3. 软件与通讯安全
4. 电力工控系统安全
5. 物联网终端安全
6. 智能无人系统安全

4.1 电力工控系统安全概述/CIA机密性、完整性、可用性

BACKGROUND & MOTIVATION

Goals of this lecture

- Introduction of risk management to control security （风险管理）
- Vulnerability assessment: a security index approach （脆弱性分析：安全因子）
- Applied robust attack detection through tractable optimizations （攻击检测）
- Other control security research frontier

US-Canada Blackout

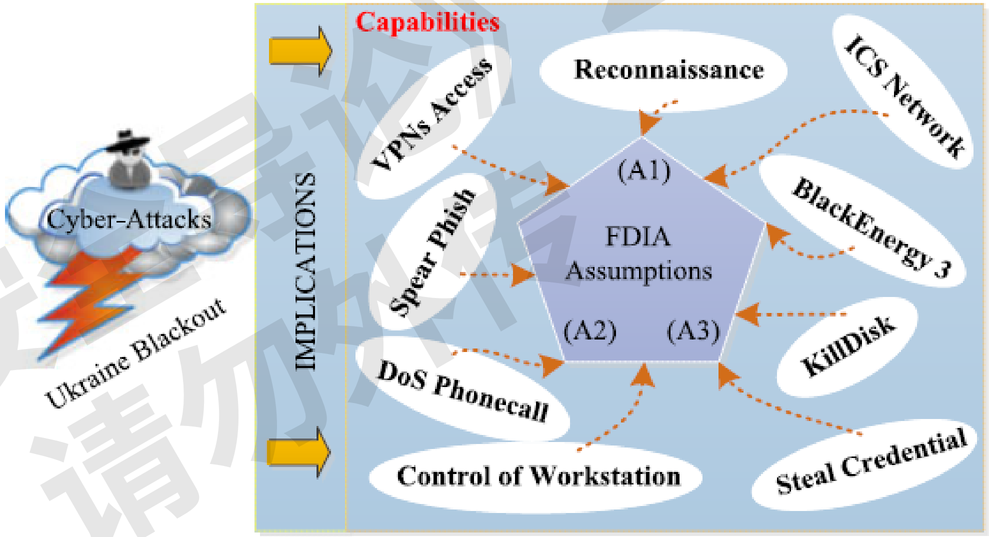
Normal failures have huge impact - US-Canada 2003 Blackout.

What about **intentional** cyber attacks?

Is this a "real" concern?

The 2015 Ukraine Blackout

- **Long-term reconnaissance**
- **BlackEnergy3:**
  - –delivered via spear phishing emails;
  - –initial access vector for the theft of authorized users' VPN credentials.
- **Telephonic denial-of-service attack:**
  - –frustrated reports of outages to call centers.
- **Modified "KillDisk" firmware attack:**
  - –erased master boot records **(主引导记录)** on workstations, thereby delaying restoration.



[Liang *et al.* IEEE Trans. Power Syst., 2017]

- **Primary Attack: Hijack of the Supervisory Control and Data Acquisition (SCADA) network**
  - –targeting of field devices;
  - –remote opening of substation breakers.

VPN: virtual private network ICS: industrial control system FDIA: false data injection attack

Cyber Security in Smart Grid

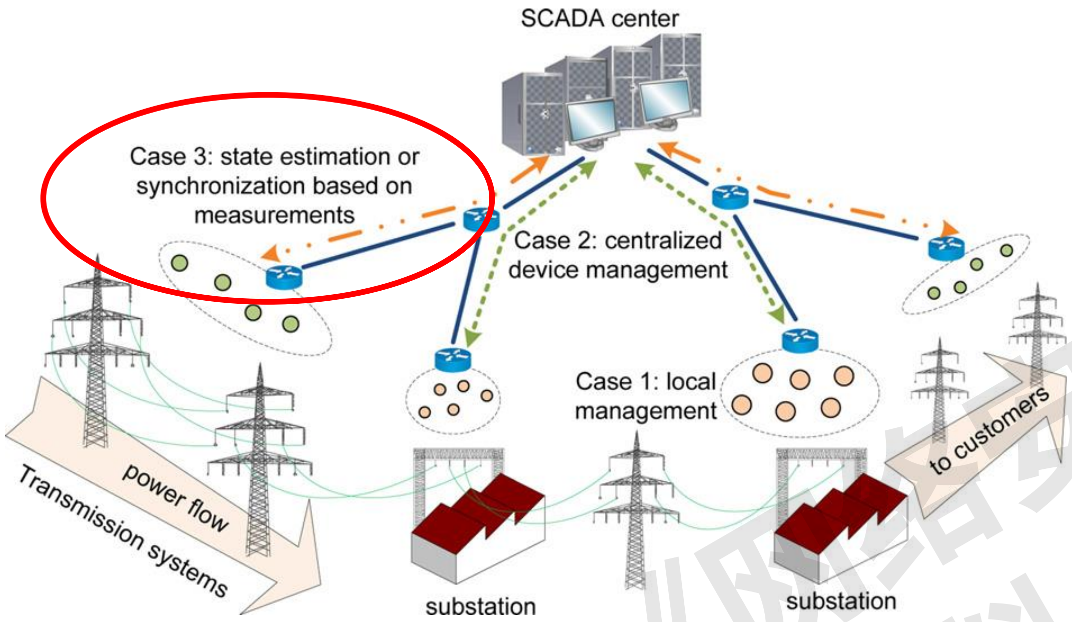
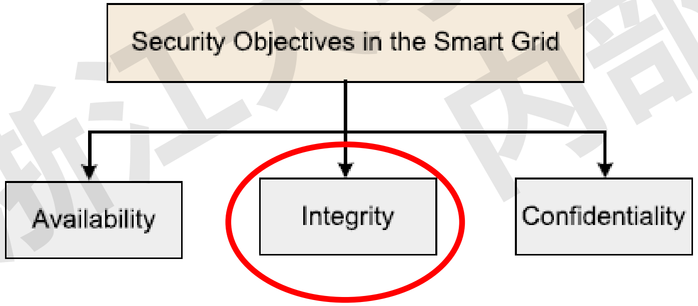
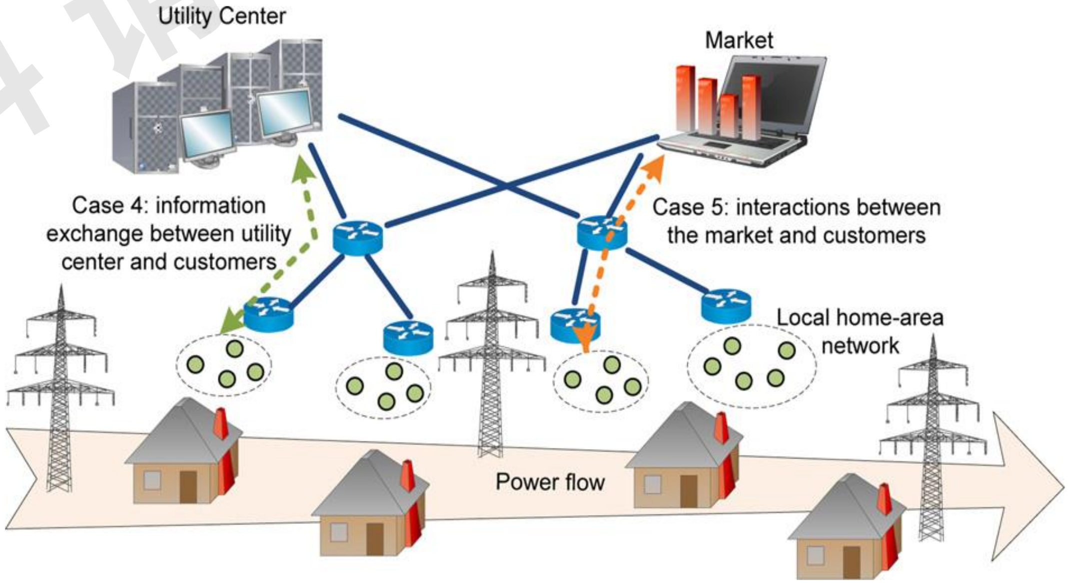


Table 1
Differences between the Internet and the Smart Grid communication network.

	The internet	Smart Grid communication network
Performance metric	Throughput and fairness	Message delay
Major traffic	Power-law	Periodic
Timing requirement	Delay-sensitive (100 ms) to best-effort	Time-critical (3 ms) to best-effort
Communication model	End-to-end	Two-way, limited peer-to-peer
Protocol stack	IPv4, IPv6	Proprietary, heterogeneous, IPv6

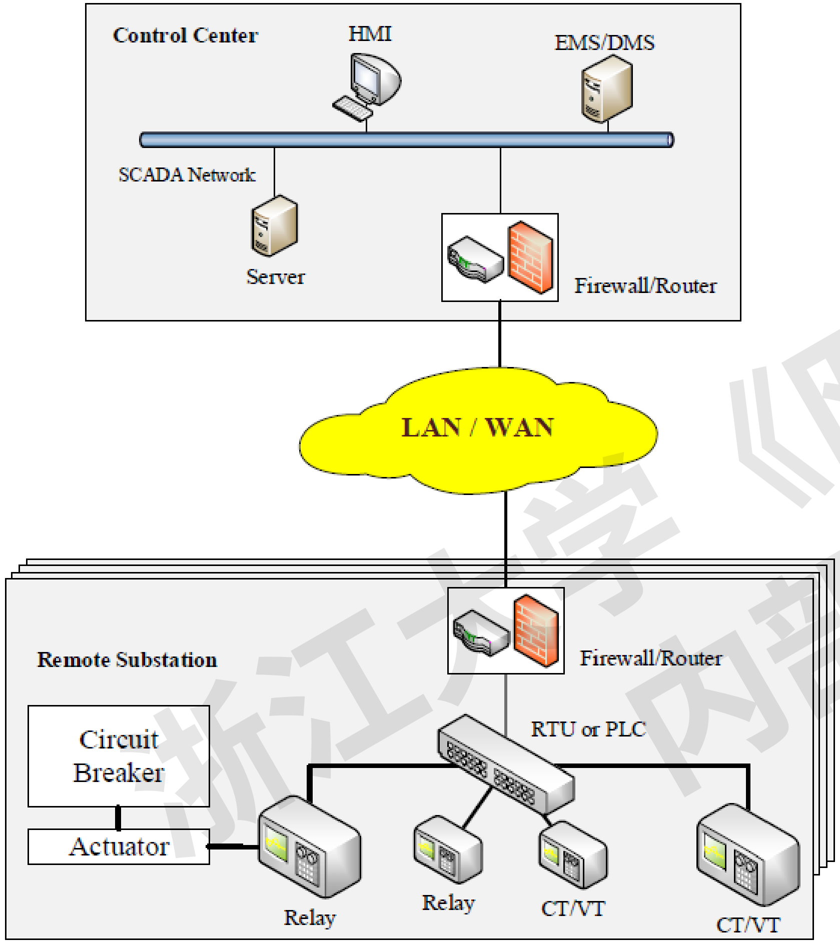




W. Wang, Z. Lu, "Cyber security in the Smart Grid: survey and challenges". Computer Network, 2013;57:1344–71.

SCADA Network

- **Architectures（通讯架构）**



- **Components**
  - **Sensors and control devices**, wired to Programmable Logic Controller (PLC), directly interfaced with the Remote Terminal Unit (RTU).
  - **Communication system**, connecting the SCADA master to the RTU/PLC in the remote field.
  - **Human machine interface (HMI)**
  - **Software**, e.g., Energy Management System (EMS), Demand Management System (DMS).

- **Protocols (通讯协议)**
  - **Traditional**: Distributed Network Protocol Ver. 3 (DNP3), IEC 60870-5-101& IEC 60870-5-104, etc.
  - **Modern:** IEC61850 (SV, MMS, GOOSE), standard suite for substation automation, etc.

CTs: current transformers   VTs: voltage transformers

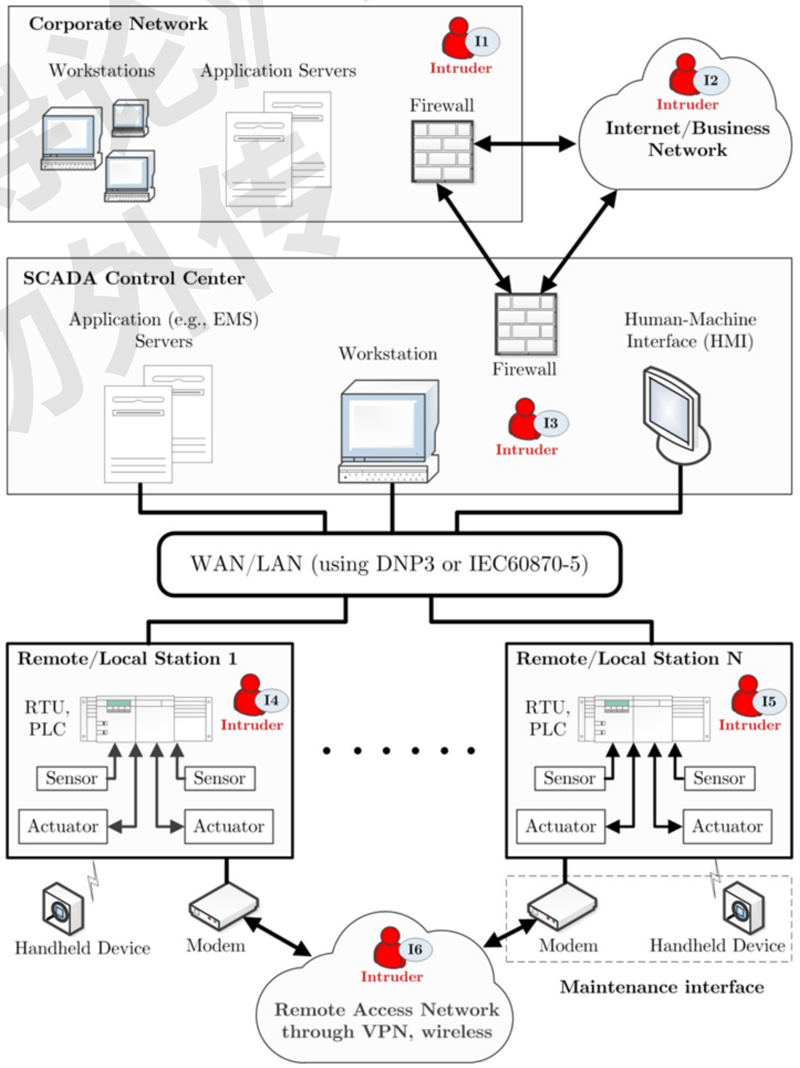
LAN: local-area network   WAN: wide-area network

SCADA Cyber Security Threats

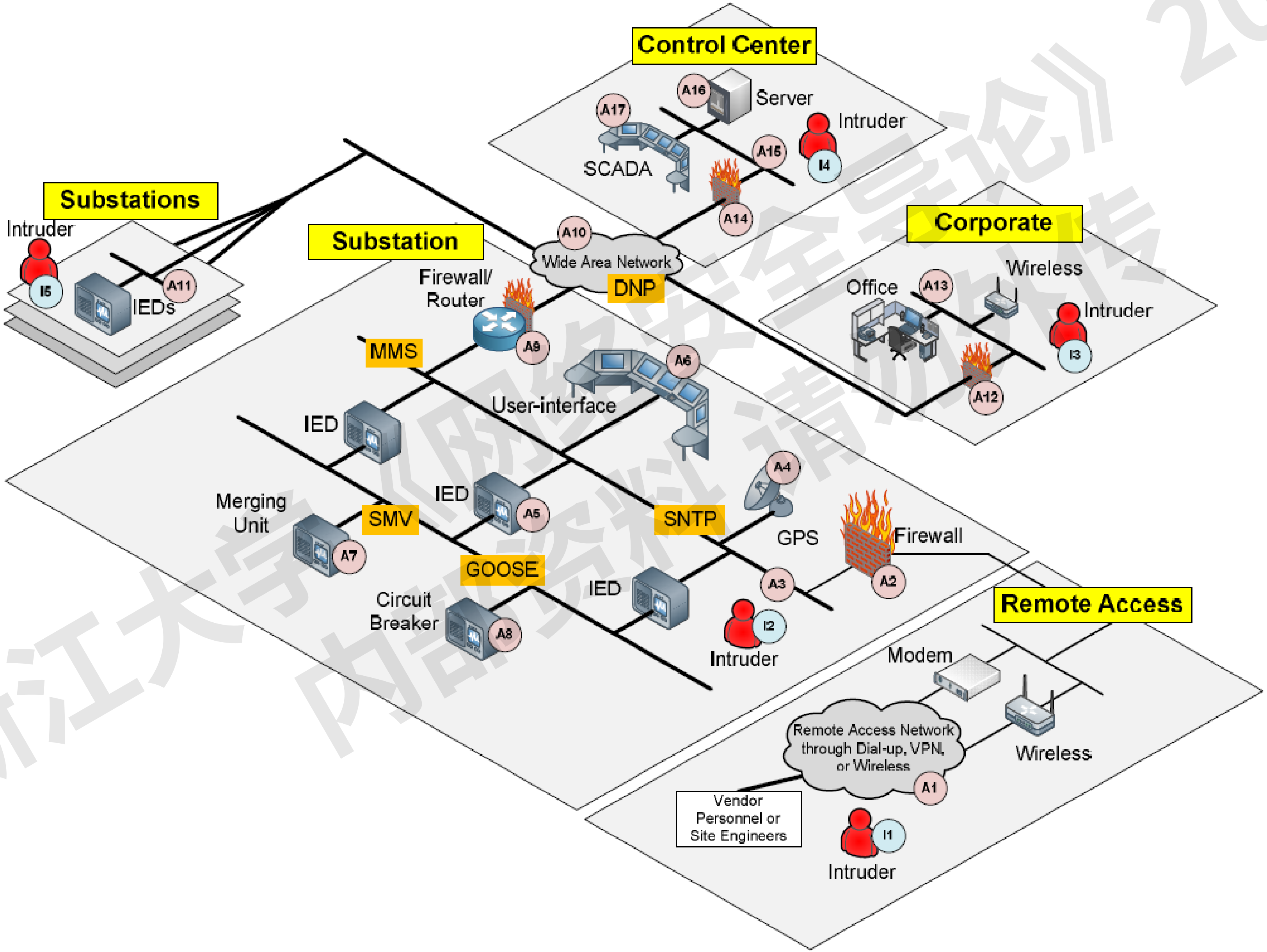
- “Most SCADA network protocols are not designed to provide robust security checks.”

  --*Vulnerability Analysis of Energy Delivery Control System*.

- “SCADA networks are more connected to Internet and corporate networks, leading to increased vulnerability to cyber threat.”



SCADA Cyber Security Threats



Cybersecurity Leads to Societal Costs

Attack

SCADA system

Power network

Societal cost

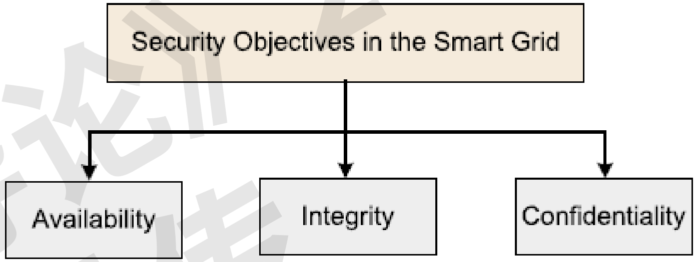
Security issues

Power system: susceptible to operational errors and external attacks

Smart grid technology makes the system even more vulnerable

Recall: C I A

Confidentiality, Integrity, Availability (CIA)

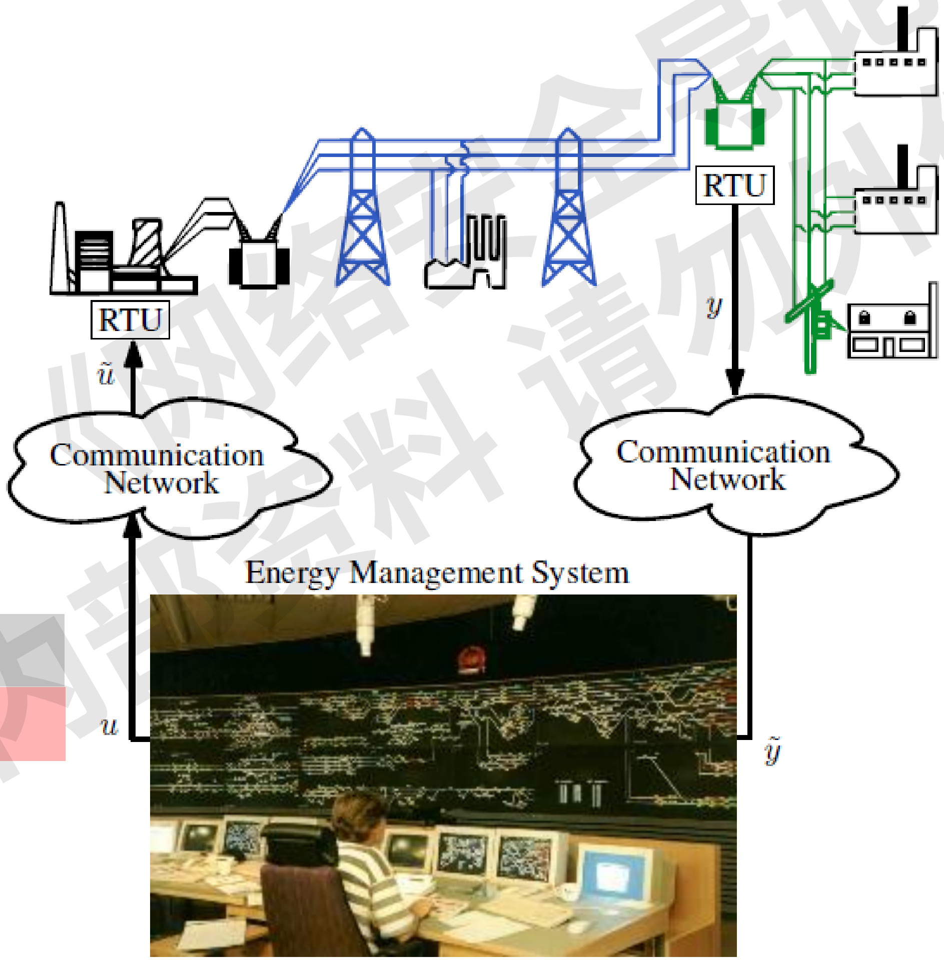


- **Confidentiality (机密性):**
  - –Confidentiality is roughly equivalent to *privacy*. Measures are designed to prevent sensitive information from reaching the wrong devices/people.
  - –**Confidentiality attacks:** eavesdropping attacks (窃听), packet sniffing attacks (挟持), etc.

- **Integrity (完整性):**
  - –Maintaining the consistency, accuracy, and trustworthiness of data over its entire life cycle.
  - –**Integrity attacks:** false data injection (FDI) attack, zero-dynamics attack, etc.

- **Availability (可用性): :**
  - –frustrated reports of outages to call centers.
  - –**Availability attacks:** denial of service (DoS) attacks, distributed DoS (DDoS) attacks, etc.

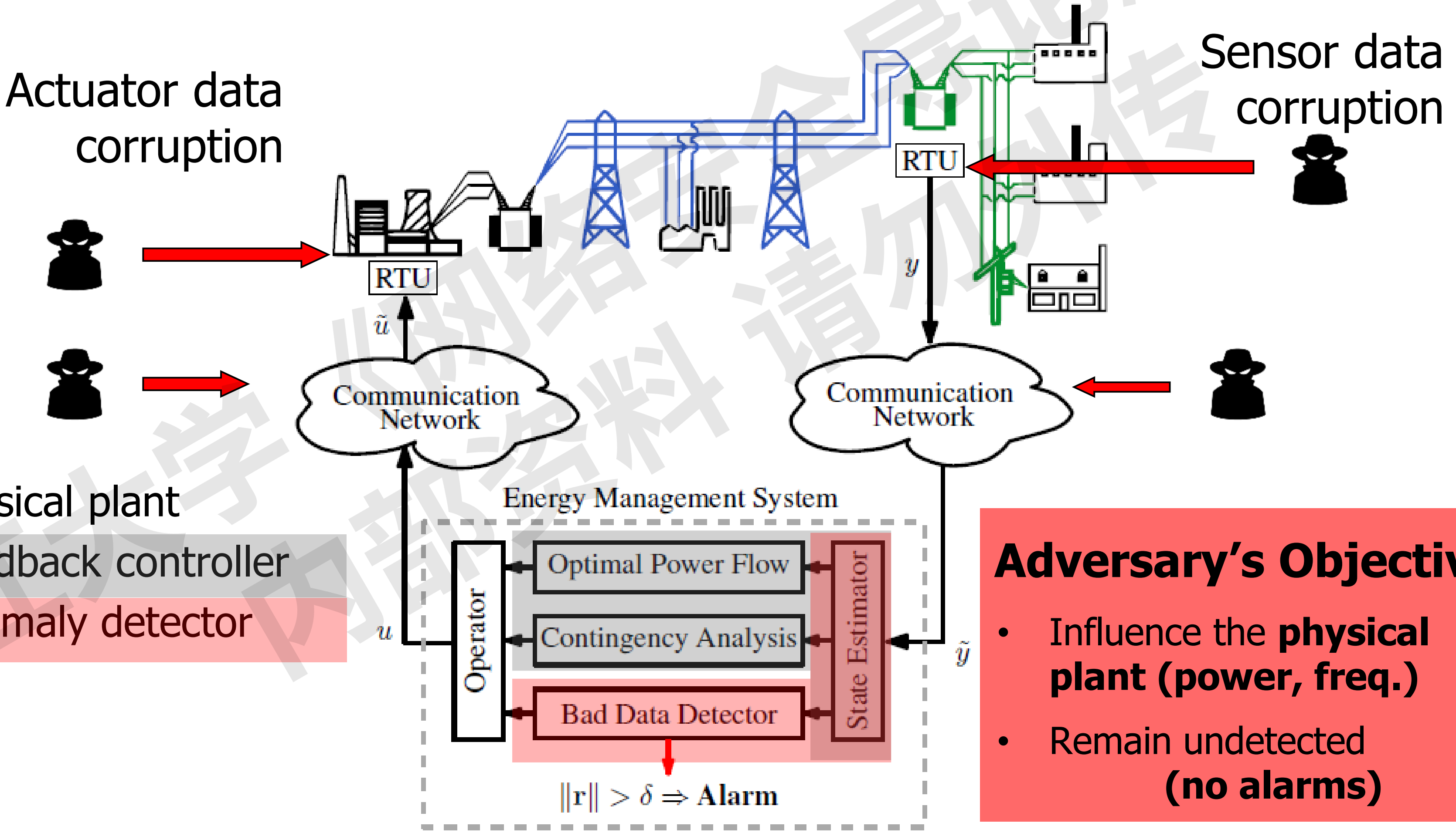
Power System Control



- Physical plant
- Feedback controller
- Anomaly detector

Power System Control

Closing the loop over corrupted data

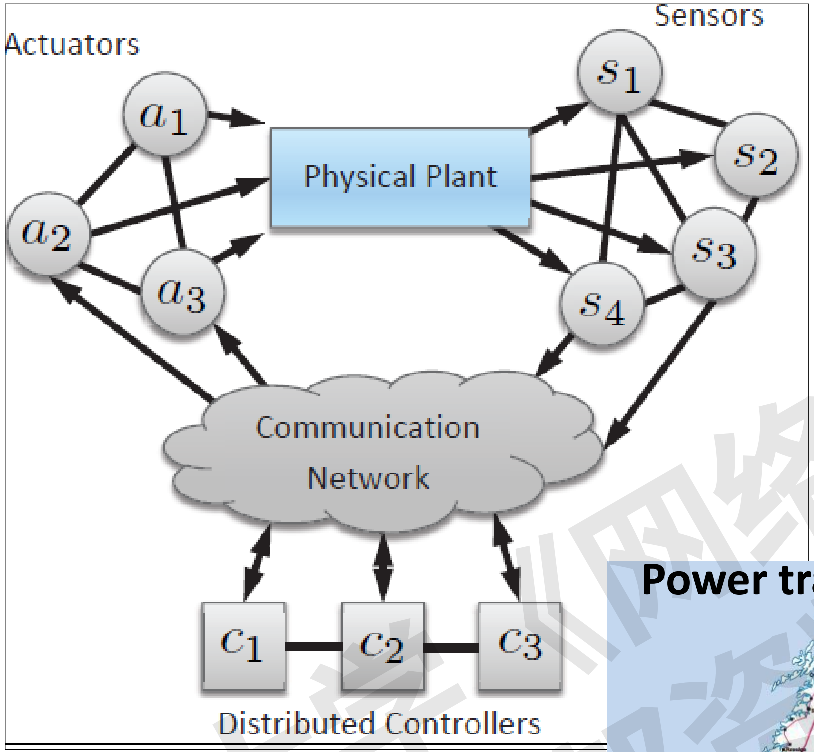


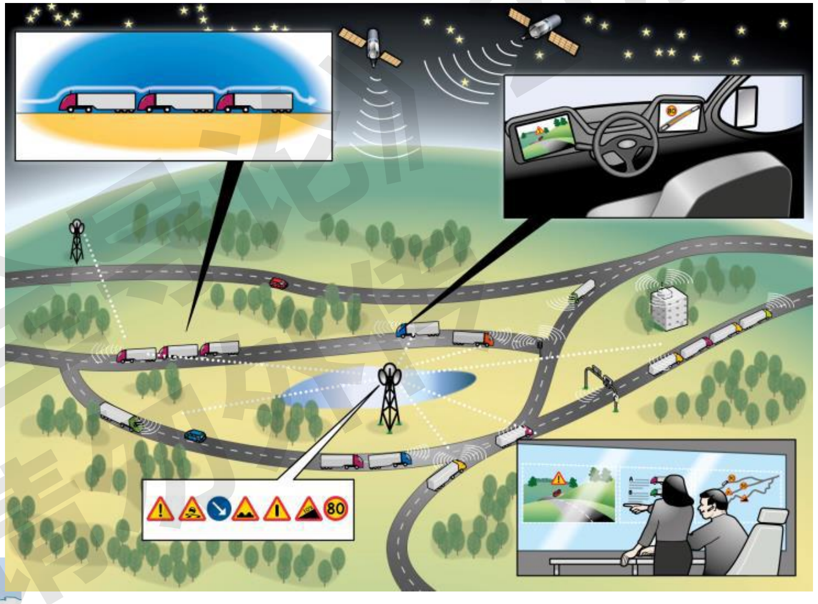
- Physical plant
- Feedback controller
- Anomaly detector

Adversary's Objective

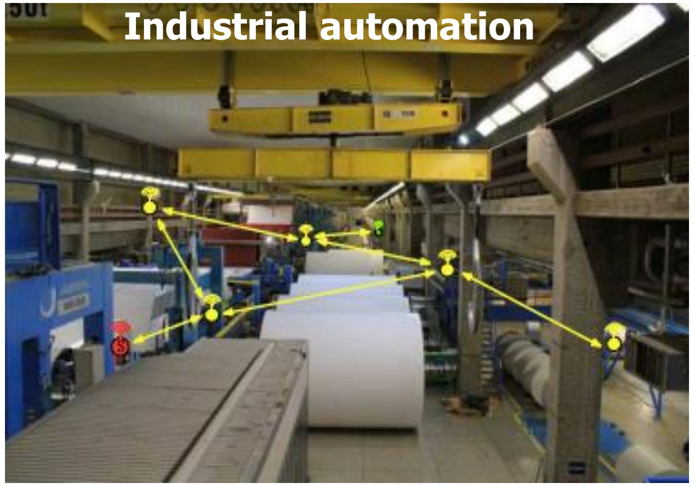
- Influence the **physical plant (power, freq.)**
- Remain undetected **(no alarms)**

Other Control Systems





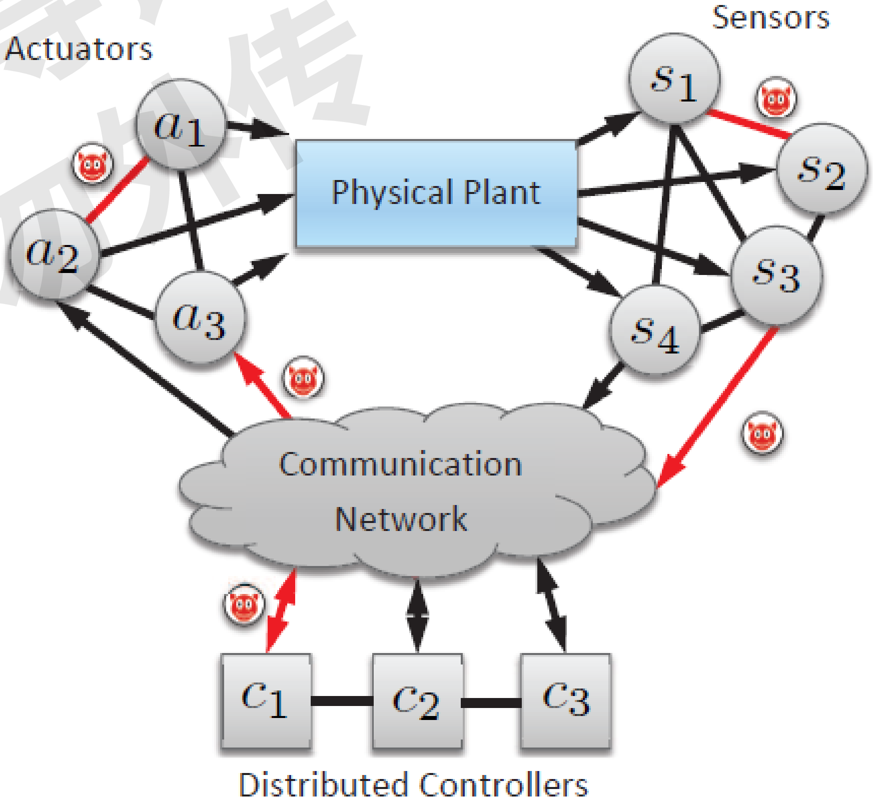
Power transmission



Control System Cyber Security

- Leads to **increased vulnerability** to cyber-threats with potential points of cyber attacks.
- **Cyber-attacks** can have dramatic physical impact.

- How to model adversaries and attacks?
- How to measure vulnerability? **（脆弱性）**
- How to compute consequences? **（攻击影响）**
- How to design protection and detection mechanisms?



- Related to work
  - **Modeling Frameworks （建模框架）**
  - **Cyber Risk Assessment（风险分析、风险管控）**
  - **Cyber Attack Detection（攻击检测）**

4.2 安全风险分析与管控/工控系统模型与隐式攻击/脆弱性分析与安全因子

CYBER RISK ANALYSIS & SECURITY METRICS

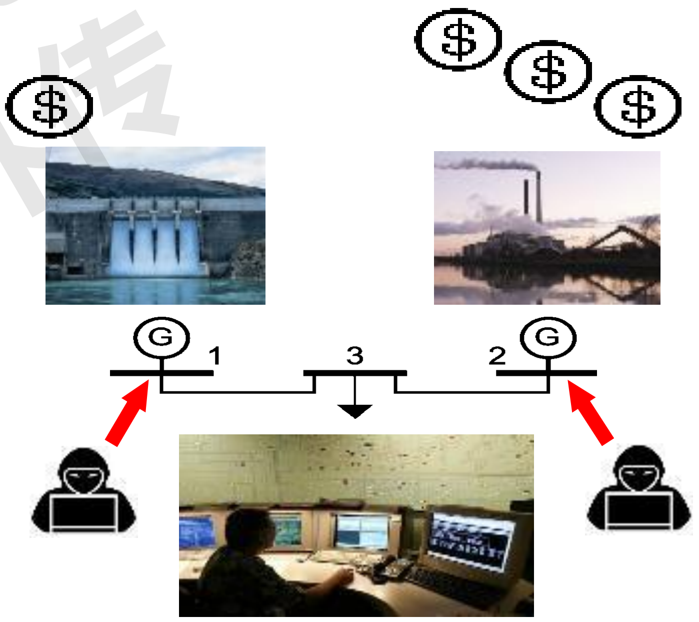
Why cyber risk management?

- Complex systems with numerous attack scenarios.

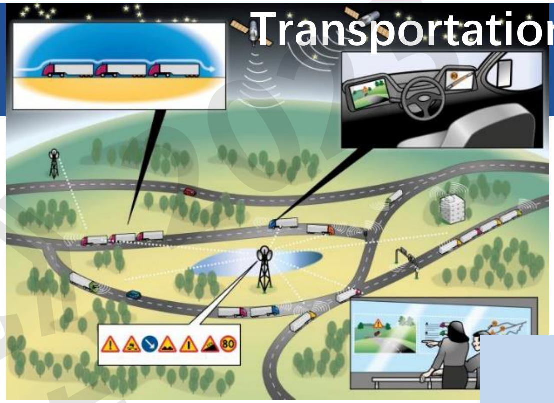
- Too costly to secure the entire system against all possible attack scenarios.

- What scenarios to **prioritize**? **(优先级)**

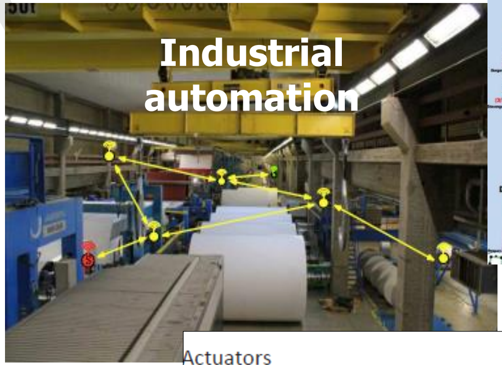
- What components to **protect/defend first**? **(防御策略)**



Why cyber risk management?

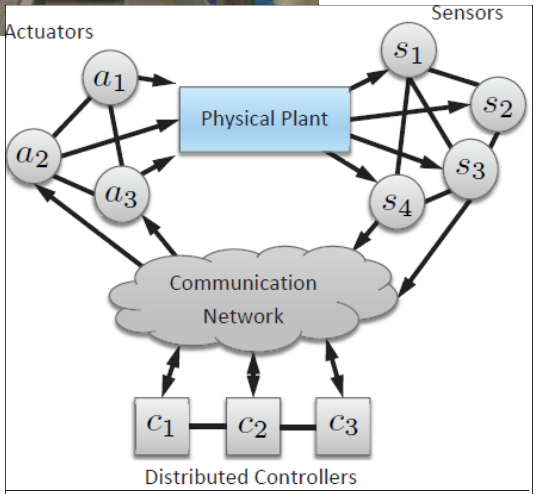


- Examples: **Critical infrastructures（关键基础设施）**
- power, transport, water, gas, oil are often with weak security guarantees



- What scenarios to **prioritize**?

- What components to **protect/defend first**?

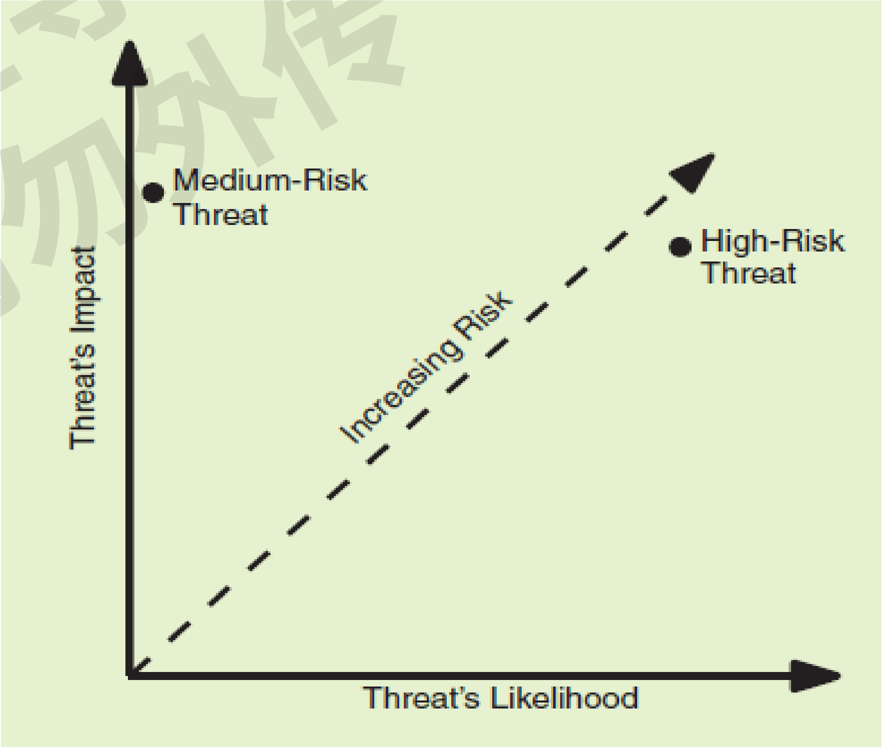


The Concept of Cyber Risk

- **Risk is a set of tuples**: [Kaplan & Garrick, 1981]

- **Attack scenario （攻击是什么？）**
- **Impact** of the attack **（攻击影响）**
- **Likelihood** of the attack **（攻击发生可能性）**

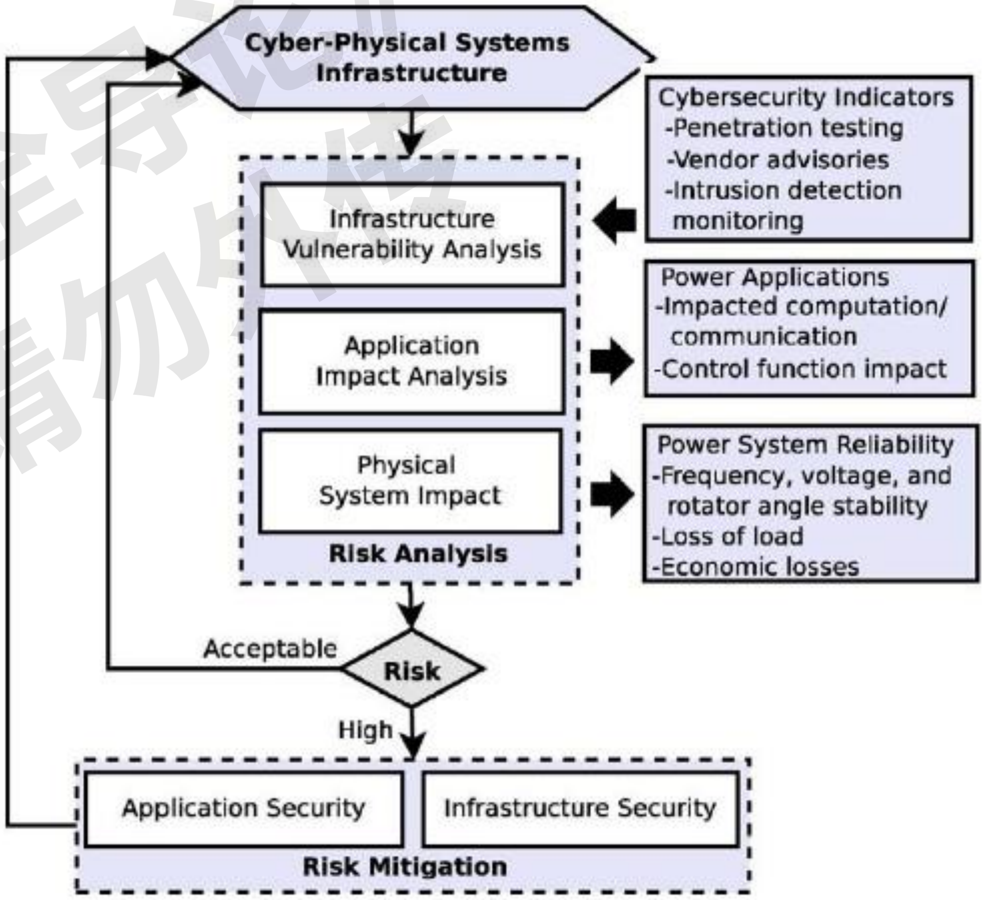
Risk = Likelihood * Impact



[Texeira *et al.* IEEE CSM, 2015]

Risk Management Cycle

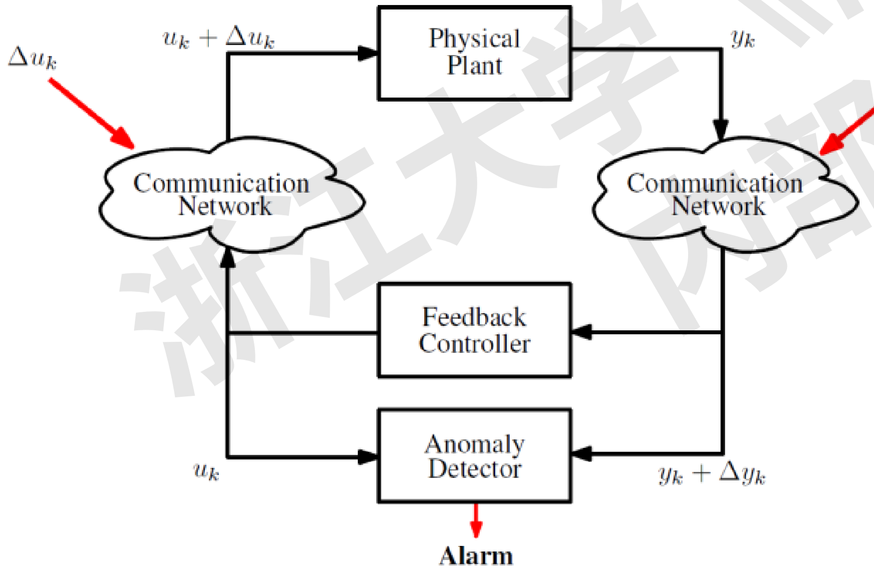
- Main steps in risk management
  - Scope definition （范围定义）
    - *Models, scenarios, objectives*
  - Risk Analysis （风险分析）
    - ***Attack** scenario*
    - ***Likelihood** Assessment*
    - ***Impact** Assessment*
  - Risk Treatment （风险处置）
    - *Prevention, detection, Mitigation*

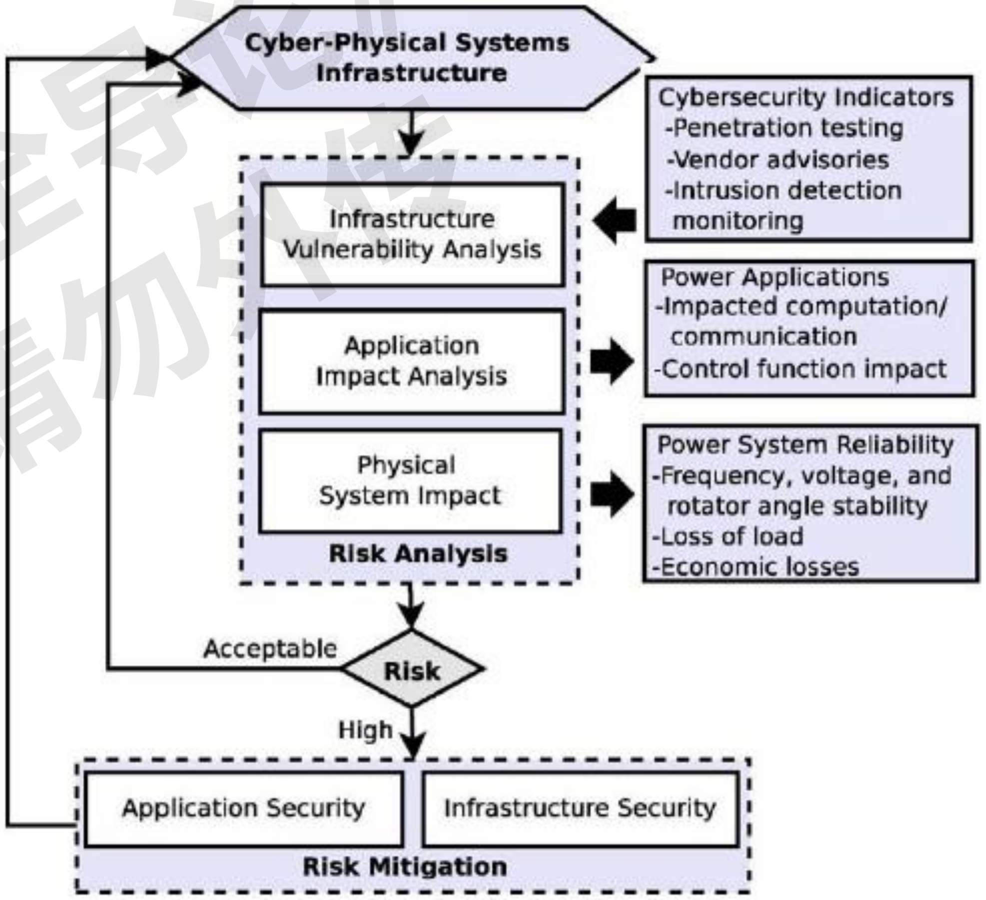


[Sridhar et al., Proc. IEEE, 2012]

Risk Management Cycle

- Risk is a set of tuples: [Kaplan & Garrick, 1981]
  - **Attack Scenario**
  - **Impact** of the attack
  - **Likelihood** of the attack
- How to model adversaries and attacks?
  - *Describe the system*
  - *Characterize the* ***attack scenario***

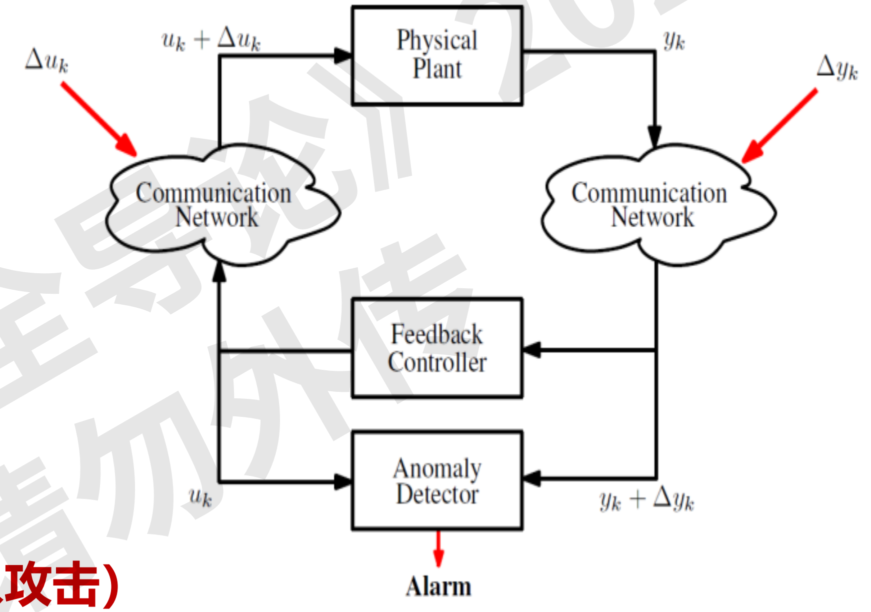


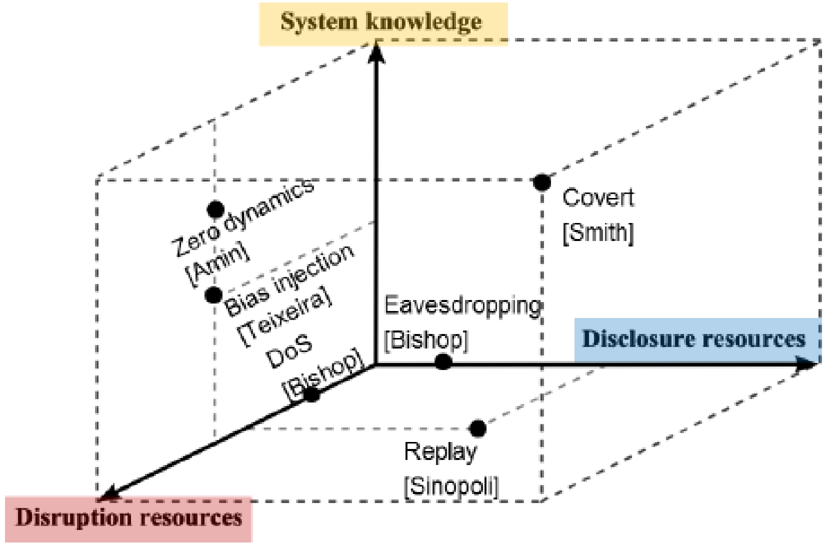


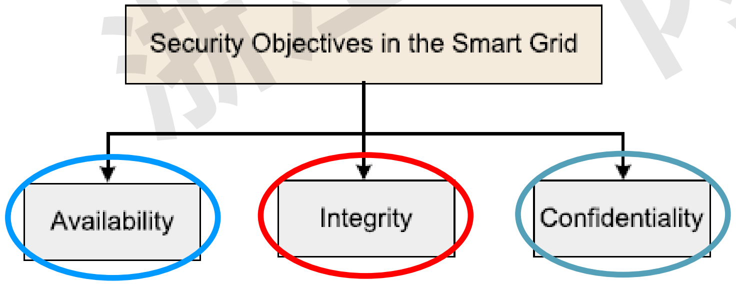
Risk Management Cycle

- How to model adversaries and attacks?
  - *Describe the system*
  - *Characterize the **attack scenario***

- **Attack scenarios**
  - ***Dos / DDoS attacks （拒绝服务攻击）***
  - ***False data injection (FDI) attacks （虚假数据注入攻击）***
  - ***Zero dynamics attacks （“零动态”攻击）***
  - ***Covert attacks***
  - ***Eavesdropping attacks （窃听攻击）***



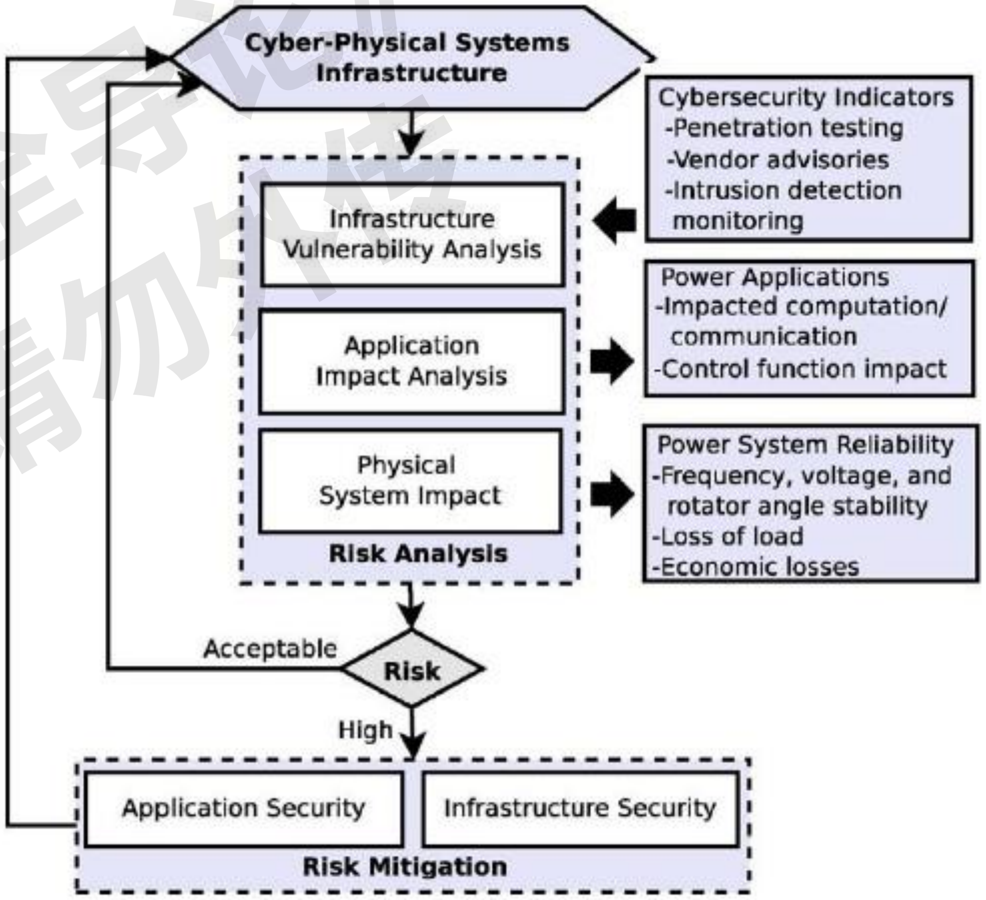




[Texeira *et al.* IEEE Control Syst. Mag. , 2015]

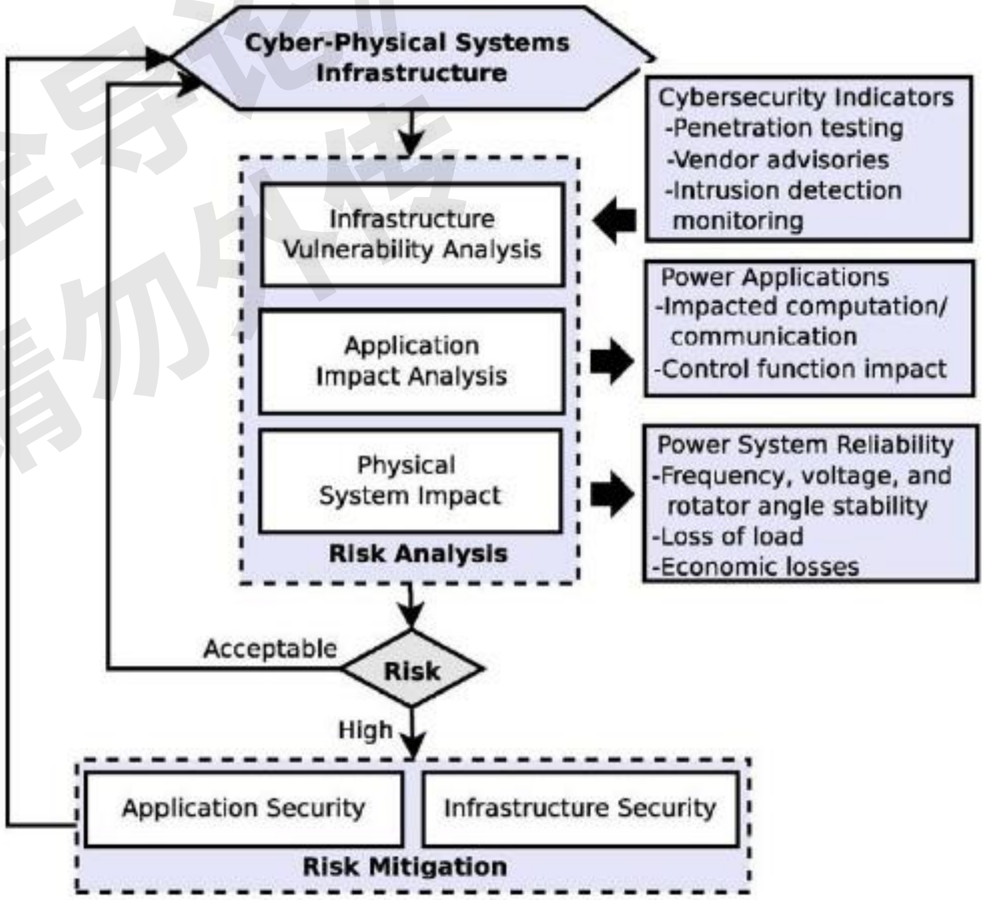
Risk Management Cycle

- Risk is a set of tuples: [Kaplan & Garrick, 1981]
  - **Attack Scenario**
  - **Impact** of the attack
  - **Likelihood** of the attack
- How to model adversaries and attacks?
  - *Describe the system*
  - *Characterize the attack* scenario
- How to measure vulnerability?
  - *Assess likelihood of attack*
    - ***Attack effort***
    - ***Amount of resources (knowledge, corrupted channels)***



Risk Management Cycle

- Risk is a set of tuples: [Kaplan & Garrick, 1981]
  - **Attack Scenario**
  - **Impact** of the attack
  - **Likelihood** of the attack
- How to model adversaries and attacks?
  - *Describe the system*
  - *Characterize the attack scenario*
- How to measure vulnerability?
  - *Assess likelihood of attack*
    - Attack effort
    - Amount of resources (knowledge, corrupted channels)
- How to compute consequences?
  - *Assess Impact on performance objectives*
    - ***Loss of performance***
    - ***Loss of stability***
    - ***Loss of desired properties***
    - ***Violation of safety constraints***

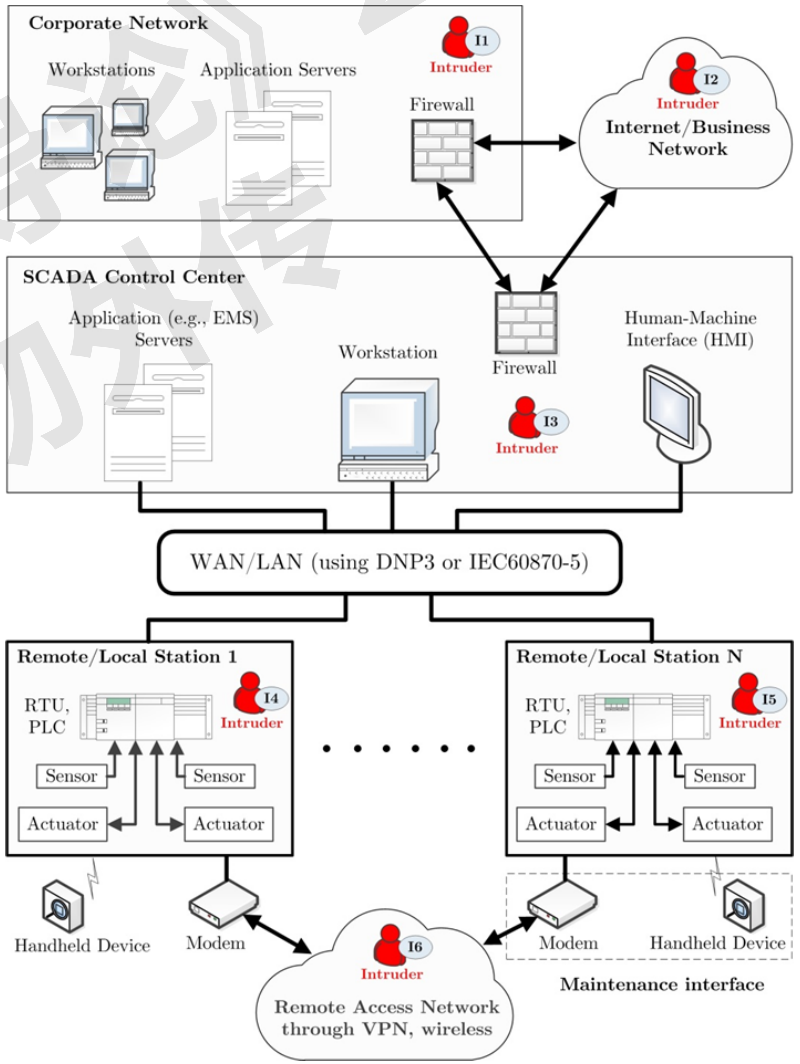


Likelihood Metrics (proxy，代理量值)

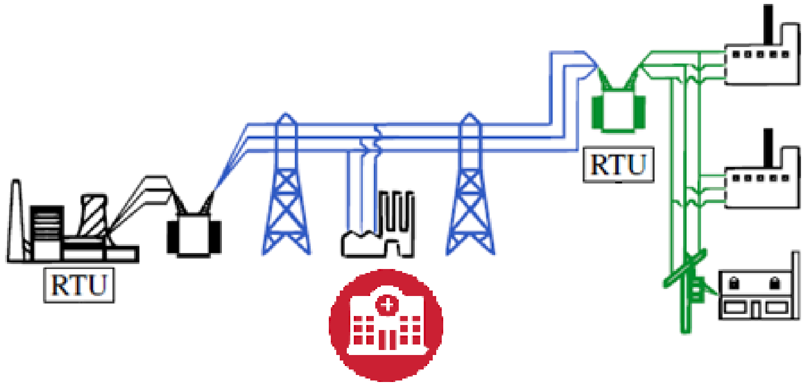
- Likelihood depends on **ICT infrastructure**

- **Successful attack**
  - successful initial infection
  - successful dissemination of malware
  - successful **infection of target device**
  - Successful **control of target device**

- **Likelihood metric:** probability of a successful attack
  - ***Hard to compute*** *– lack of historical data*
  - *Alternatively – **proxy metrics** that assess the **minimum attack effort***
  - ***E.g., number of infected target devices***



Impact Metrics (proxy)



- **Operation goals:**
  - No blackout
  - No blackout in **critical loads**
  - Efficiency
  - Quality of power supply
    - *Voltage @ 230v*
    - *Frequency @ 50Hz*

- **Impact Metrics:**
  - Loss of load
  - Loss of **critical loads**
  - Increase of costs
  - Reduced Quality of power supply
    - *(Maximum) voltage variation*
    - *(Maximum) Frequency variation*
  - Loss of stability/desired properties

Example of Likelihood Metrics

启发：拜占庭将军问题

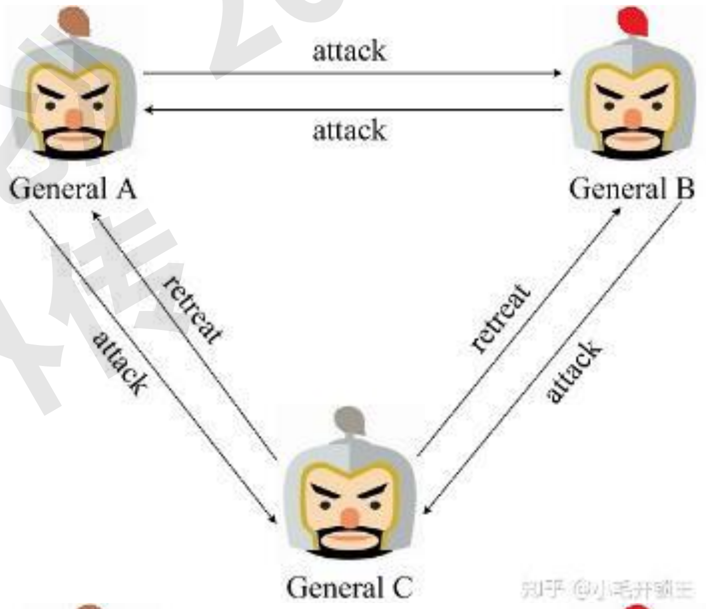
（The Byzantine Generals Problem）

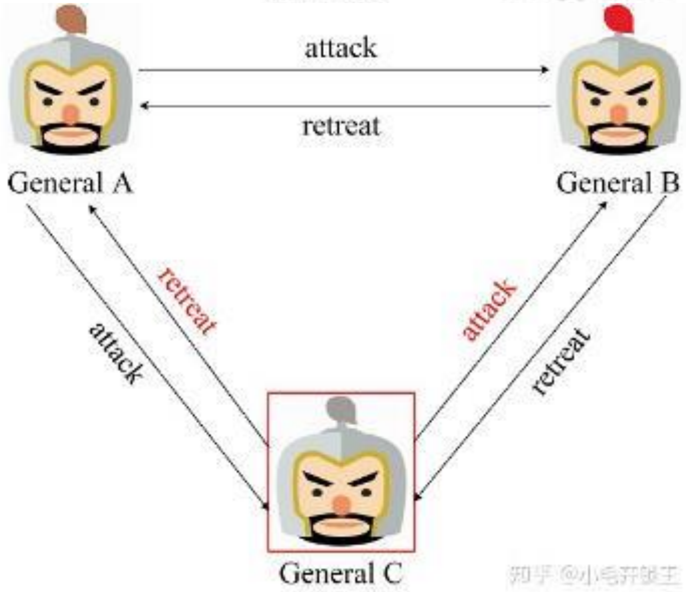
- 考虑以下情形：**n个将军**只能通过信使相互沟通，须制定共同行动计划，如**进攻(Attack)或者撤退(Retreat)**，且只有当**半数以上**的将军共同发起进攻时才能取得胜利，其中有**q个叛将**。问题是，**n-q个忠将**能够一直达成一致意见吗？

- 叛将可以做任何事：发不同意见给其他将军，不发表意见，篡改转发的意见...

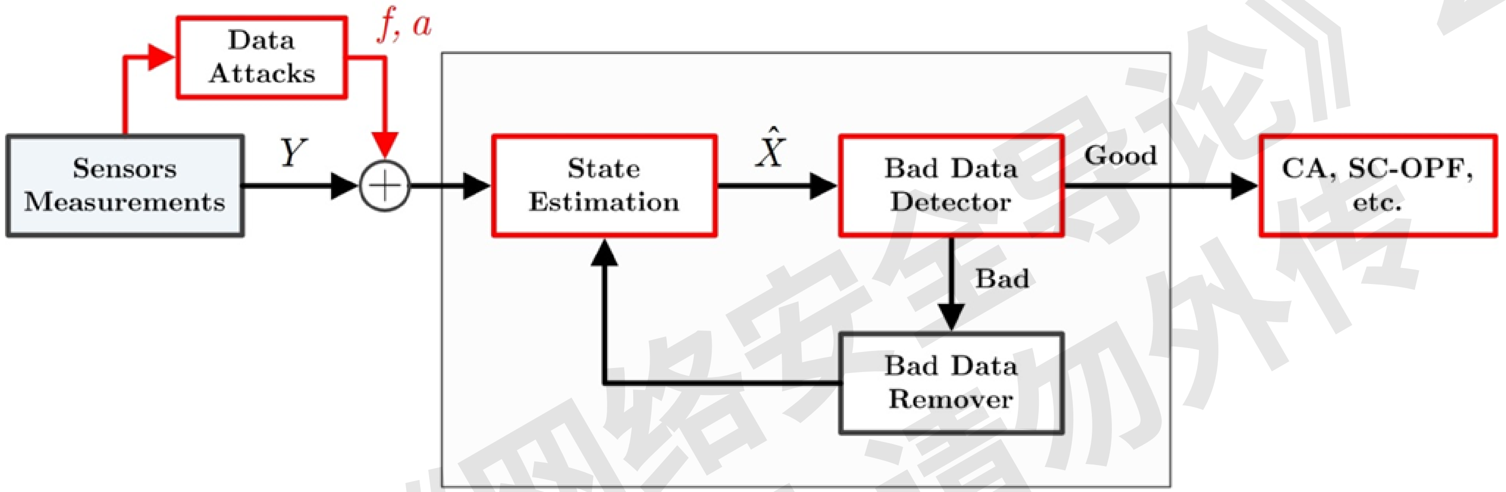
- **一致协议存在当且仅当 n >= 3q +1**

- 如果忠将可以对要发表的意见进行“**加密**”？





Example of Likelihood Metrics



- Cyber attacks on power system **State Estimation** process
- **Attacker's objectives**: 1) Attack is stealthy (undetectable); 2) Target measurements are corrupted.
- **Security Index**: *for attacks need **minimum effort***

$$\alpha^\star := \min_{f} \quad \|f\|_p \qquad \text{s.t.} \quad f \in \mathcal{S},\ f \in \mathcal{G}\,.$$

- $\mathcal{S}$: set of stealthy attacks
- $\mathcal{G}$: set of goals of attacks

$\mathcal{S}$ and $\mathcal{G}$ are scenario specific.

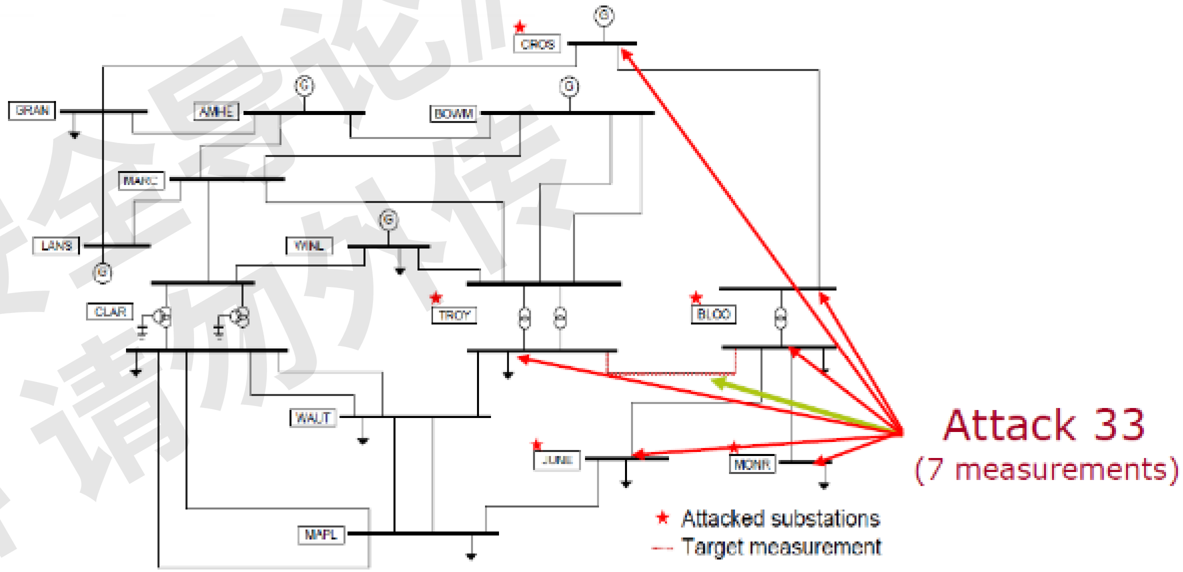
Example of Likelihood Metrics

– How many measurements must be corrupted to remain stealthy?

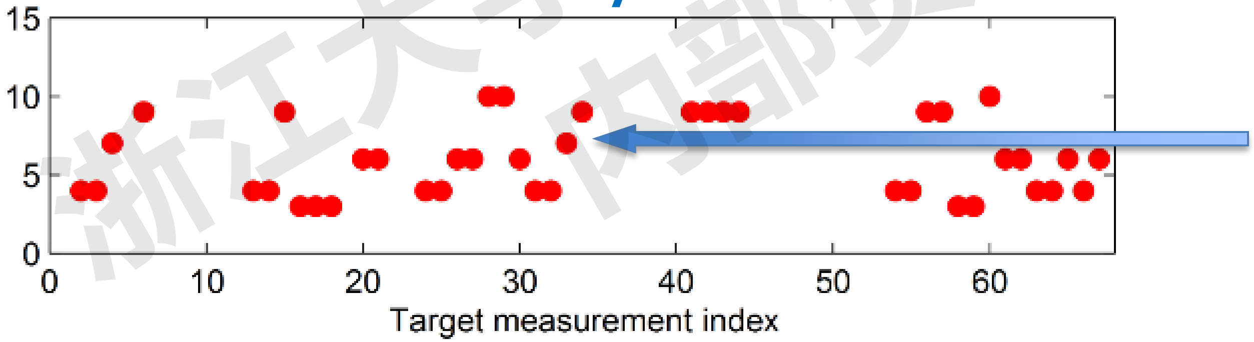
$$\alpha_i := \min_a \|a\|_0$$

$$\text{subject to} \quad a = H \triangle x,$$

$$a(i) \neq 0.$$



Security index



At least **7 measurements** involved in a stealth attack against measurement 33.

H. Sandberg, A. Teixeira, K. H. Johansson, "On security indices for state estimators in power networks," in First Workshop on Secure Control Systems (CPSWEEK 2010), Stockholm, Sweden, 2010.

Example of Likelihood Metrics

Security Metrics of 14-bus System

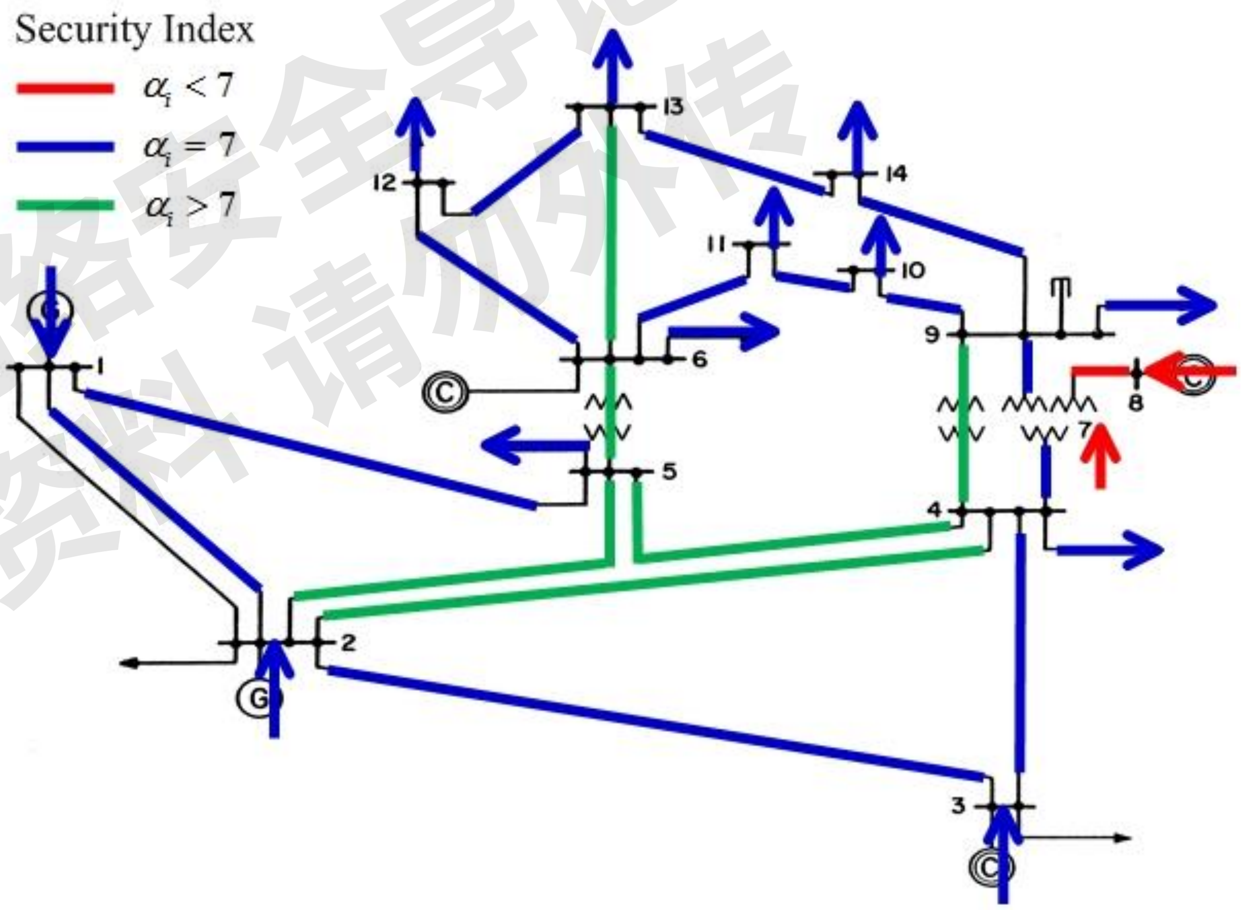
- IEEE 14-bus system

- State dimension

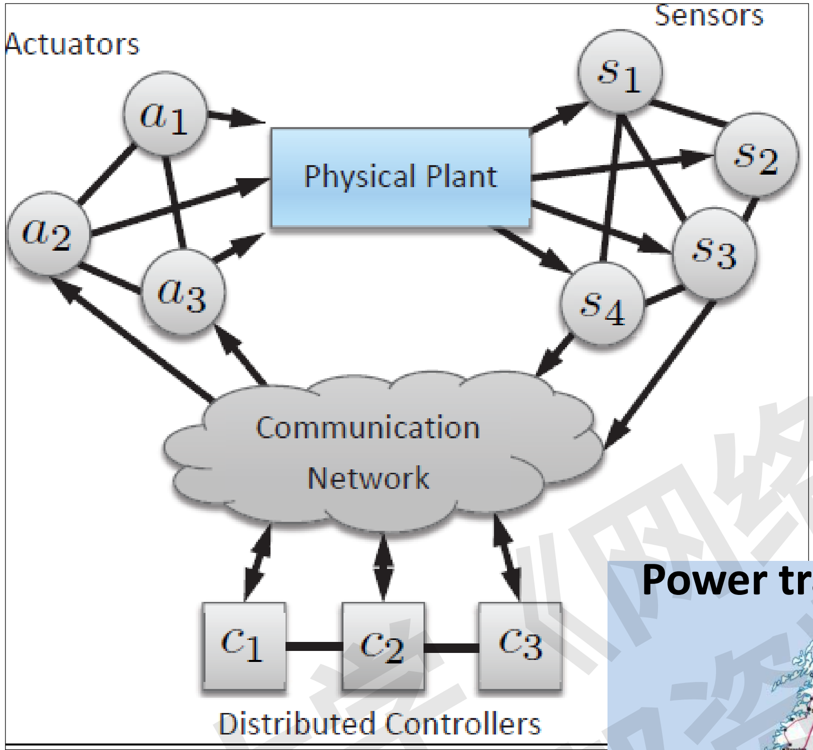
$$n_x = 14$$

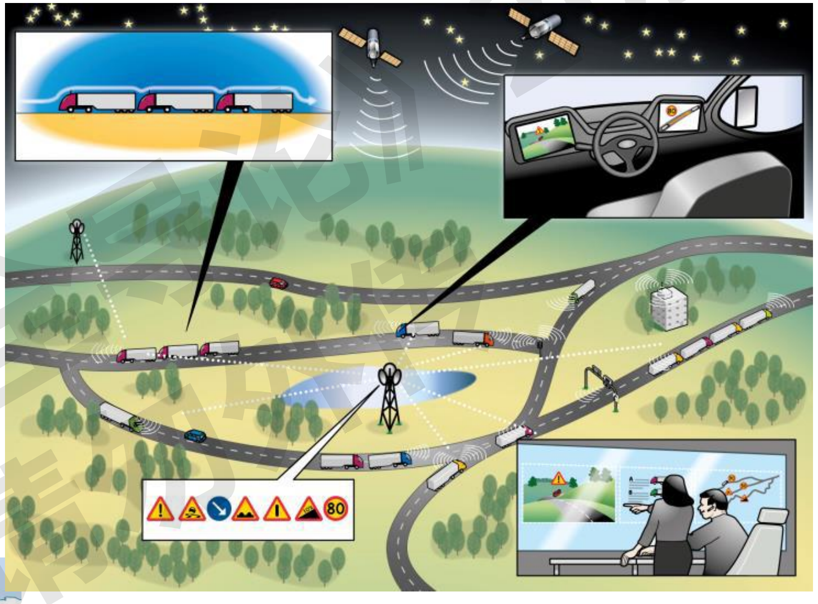
- Number of measurements

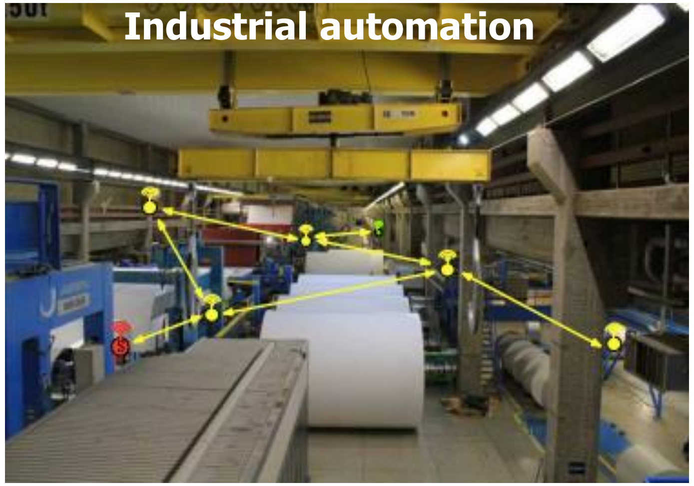
$$n_y = 54$$



Metrics Formulation for Control Systems







Basic Notations: Discrete Time-invariant System

$$x(k+1) = Ax(k) + Bu(k), \quad x(k) \in \mathbb{R}^n, \ u(k) \in \mathbb{R}^m$$
$$y(k) = Cx(k) + Du(k), \quad y(k) \in \mathbb{R}^p$$

- Unknown state: $x(k) \in \mathbb{R}^n$ ($x(0)$ in particular)
- Unknown input: $u(k) \in \mathbb{R}^m$ (e.g., natural disturbance)
- Known output (measurement): $y(k) \in \mathbb{R}^p$
- Known system model: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$
- Transfer function form: $y(z) = H(z)u(z)$
- **The Rosenbrock system matrix:**

$$P(z) = \begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} \in \mathbb{C}^{(n+p)\times(n+m)}$$

Basic Notations: Discrete Time-invariant System

$$x(k+1) = Ax(k) + Bu(k), \quad x(k) \in \mathbb{R}^n,\ u(k) \in \mathbb{R}^m$$
$$y(k) = Cx(k) + Du(k), \quad y(k) \in \mathbb{R}^p$$

- Unknown state: $x(k) \in \mathbb{R}^n$ ($x(0)$ in particular)
- Unknown input: $u(k) \in \mathbb{R}^m$ (e.g., natural disturbance)
- Known output (measurement): $y(k) \in \mathbb{R}^p$
- Known system model: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$

- **Definition:**

- 1. The input u is ***observable*** if $y(k) = 0$ for $k \geq 0$ implies $u(k) = 0$ for $k \geq 0$ ($x(0)$ unknown). **(可观性)**

- 2. The input u is ***detectable*** if $y(k) = 0$ for $k \geq 0$ implies $u(k) \to 0$ for $k \geq 0$ ($x(0)$ unknown). **(可测性)**

Input Observability and Detectability

$$x(k+1) = Ax(k) + Bu(k), \quad x(k) \in \mathbb{R}^n, u(k) \in \mathbb{R}^m$$
$$y(k) = Cx(k) + Du(k), \quad y(k) \in \mathbb{R}^p$$

- **The Rosenbrock system matrix:**

$$P(z) = \begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} \in \mathbb{C}^{(n+p)\times(n+m)}$$

- **Definition: suppose (A, B, C, D) is minimal realization （最小实现）,**
- 1. The input u is ***observable*** $\Leftrightarrow \forall\, z: rank\ P(z) = n + m$
- 2. The input u is ***detectable*** $\Leftrightarrow normalrank\ \ P(z) = n + m$, **and** $\sigma\big(P(z)\big) \subseteq \{z{:}\ |z| < 1\}$ **(正规秩：有理函数域上的秩)**

Input Observability and Detectability

$$x(k+1) = Ax(k) + Bu(k), \quad x(k) \in \mathbb{R}^n,\ u(k) \in \mathbb{R}^m$$
$$y(k) = Cx(k) + Du(k), \quad y(k) \in \mathbb{R}^p$$

- **The Rosenbrock system matrix:**

$$P(z) = \begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} \in \mathbb{C}^{(n+p)\times(n+m)}$$

- 2. The input u is ***detectable*** $\Leftrightarrow$ $normalrank\ \ P(z) = n + m$**, and** $\sigma\big(P(z)\big) \subseteq \{z\colon |z| < 1\}$

- $\sigma\big(P(z)\big)$ denotes the set of ***invariant zeros*** of the system. **(不变零点)**

Invariant Zeros （不变零点）

- **Definition:**
- A number $z_0 \in \mathbb{C}$ is an ***invariant zero*** of the system if and only if there exist vectors $x_0 \in \mathbb{C}^n$ (state-zero direction) and $u_0 \in \mathbb{C}^m$ (input-zero direction) such that the triple z_0, x_0, u_0 satisfies

$$P(z_0)\begin{bmatrix} x_0 \\ u_0 \end{bmatrix} = \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix}\begin{bmatrix} x_0 \\ u_0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

- Transmission zeros （**传递零点**） + uncontrollable/unobservable modes,
  Matlab command: `tzero`

- Minimal realization: any state-space model that is both controllable and observable; describe the system with the minimum number of states.

Example

$$x(k+1) = Ax(k) + Bu(k), \quad x(k) \in \mathbb{R}^n,\ u(k) \in \mathbb{R}^m$$
$$y(k) = Cx(k) + Du(k), \quad y(k) \in \mathbb{R}^p$$

$$A = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.9 \end{pmatrix} \quad B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.25 \end{pmatrix} \quad C = \begin{pmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0 & 0.4 \end{pmatrix}$$

- **Transfer function:** $y(z) = C(zI - A)^{-1}B + D = \begin{pmatrix} \frac{0.2}{z-0.9} & \frac{0.3}{z-0.8} \\ \frac{0.1}{z-0.9} & \frac{0.1}{z-0.9} \end{pmatrix}$
- **Invariant zeros:** $\sigma\big(P(z)\big) = \{1.1\}$
- **Normal rank:** $normalrank\ \ P(z) = 2$

- 1. The input u is ***observable:*** ***NO!***
- 2. The input u is ***detectable:*** ***NO!***

- With $x(0) = \begin{pmatrix} -0.705 \\ 0.470 \\ 0.352 \end{pmatrix}$ and $u(k) = 1.1^k \begin{pmatrix} -0.282 \\ 0.282 \end{pmatrix}$, then $y(k) = 0,\ k \geq 0$

Example

$$x(k+1) = Ax(k) + Bu(k), \quad x(k) \in \mathbb{R}^n,\ u(k) \in \mathbb{R}^m$$
$$y(k) = Cx(k) + Du(k), \quad y(k) \in \mathbb{R}^p$$

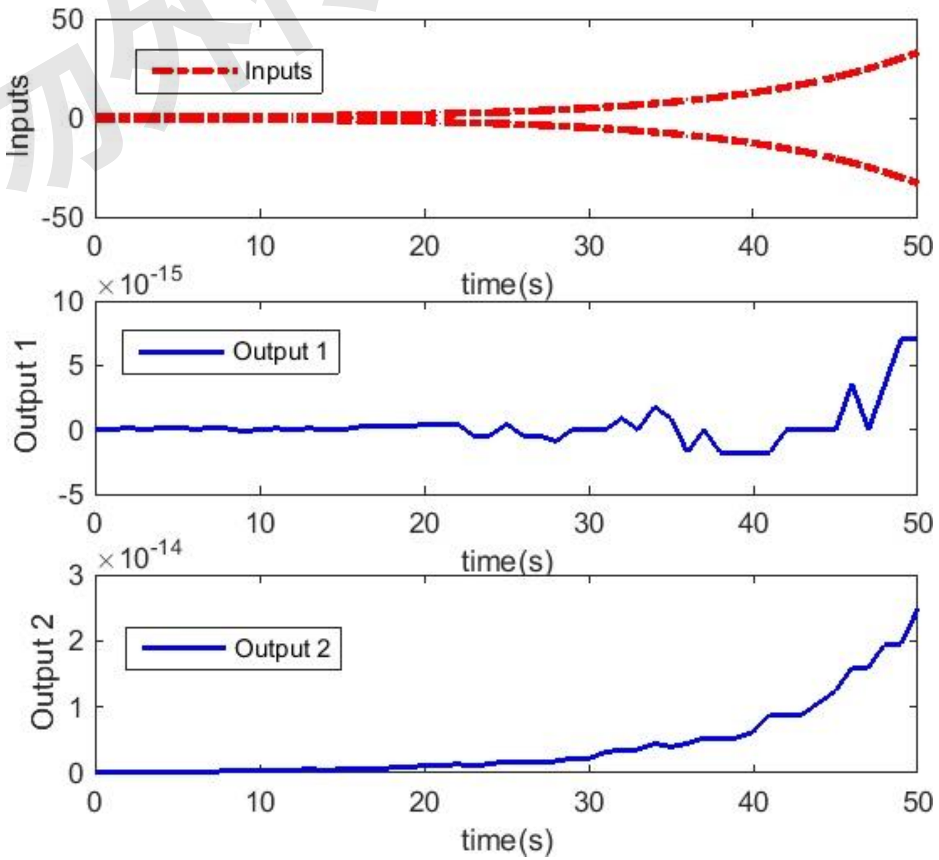
$$A = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.9 \end{pmatrix} \quad B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.25 \end{pmatrix} \quad C = \begin{pmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0 & 0.4 \end{pmatrix}$$

- 1. The input u is ***observable:*** **NO!**
- 2. The input u is ***detectable:*** **NO!**

- With $x(0) = \begin{pmatrix} -0.705 \\ 0.470 \\ 0.352 \end{pmatrix}$

  and $u(k) = 1.1^k \begin{pmatrix} -0.282 \\ 0.282 \end{pmatrix}$,

  then $y(k) = 0,\ k \geq 0$



Disturbance and Attack Model

$$x(k+1) = Ax(k) + B_d d(k) + B_a a(k)$$
$$y(k) = Cx(k) + D_d d(k) + D_a a(k)$$

- Unknown state: $x(k) \in \mathbb{R}^n$ ($x(0)$ in particular)
- Unknown (natural) disturbance: $d(k) \in \mathbb{R}^o$
- Unknown (malicious) false data injection (FDI) attack: $a(k) \in \mathbb{R}^m$
- Known output (measurement): $y(k) \in \mathbb{R}^p$
- Known system model: $A \in \mathbb{R}^{n\times n}$, $B_d \in \mathbb{R}^{n\times o}$, $B_a \in \mathbb{R}^{n\times m}$ $C \in \mathbb{R}^{p\times n}$,

  $D_d \in \mathbb{R}^{p\times o}$, $D_a \in \mathbb{R}^{p\times m}$

Disturbance and Attack Model

$$x(k+1) = Ax(k) + B_d d(k) + B_a a(k)$$
$$y(k) = Cx(k) + D_d d(k) + D_a a(k)$$

- **Definition:**
- 1. An attack signal a is ***persistent*** if $a(k) \nrightarrow 0$ as $k \to \infty$
- 2. A persistent attack signal a is ***undetectable*** if there exists a simultaneous (masking) disturbance signal d and initial state $x(0)$ such that $y(k) = 0, k \geq 0$.

  (**zero-dynamics attack，“零动态”攻击**)

[Sandberg, Teixeira]

Undetectable Attack

$$x(k+1) = Ax(k) + B_d d(k) + B_a a(k)$$
$$y(k) = Cx(k) + D_d d(k) + D_a a(k)$$

- Transfer function form: $y(z) = [G_d(z) \;\; G_a(z)] \begin{bmatrix} d(z) \\ a(z) \end{bmatrix}$
  - If $0 = G_d d + G_a a$, then clearly $y = G_a a = -G_d d$
  - It's impossible to distinguish between the undetectable attack and the masking disturbance, if they occur by themselves without the other.
- **How to build an attack signal a to be undetectable?**

Undetectable Attack

$$x(k+1) = Ax(k) + B_d d(k) + B_a a(k)$$
$$y(k) = Cx(k) + D_d d(k) + D_a a(k)$$

- **How to build an attack signal a to be undetectable?**
- **The Rosenbrock system matrix:**

$$P(z) = \begin{bmatrix} A - zI & B_d & B_a \\ C & D_d & D_a \end{bmatrix}$$

- **Proposition:**
- An attack signal $a(k) = {z_0}^k a_0$, $a_0 \in \mathbb{C}^m$, $z_0 \in \mathbb{C}$, is **undetectable** if and only if there exist $x_0 \in \mathbb{C}^n$ and $d_0 \in \mathbb{C}^o$ such that

$$P(z_0)\begin{bmatrix} x_0 \\ d_0 \\ a_0 \end{bmatrix} = 0$$

- The undetectable attack is also ***persistent*** if and only if $|z_0| \geq 1$.
- *Proof: recall the definition of invariant zeros.*

Example (Continued)

$$A=\begin{pmatrix}0.9 & 0 & 0\\ 0 & 0.8 & 0\\ 0 & 0 & 0.9\end{pmatrix}\quad B=\begin{pmatrix}0.5 & 0\\ 0 & 0.5\\ 0 & 0.25\end{pmatrix}\quad C=\begin{pmatrix}0.4 & 0.6 & 0\\ 0.2 & 0 & 0.4\end{pmatrix}$$

- **Transfer function:** $y(z)=C(zI-A)^{-1}B+D=(G_d(z)\quad G_a(z))$

$$G_d(z)=(),\qquad G_a(z)=\begin{pmatrix}\frac{0.2}{z-0.9} & \frac{0.3}{z-0.8}\\ \frac{0.1}{z-0.9} & \frac{0.1}{z-0.9}\end{pmatrix}$$

- **Invariant Zeros:** $\sigma\big(P(z)\big)=\{1.1\}$

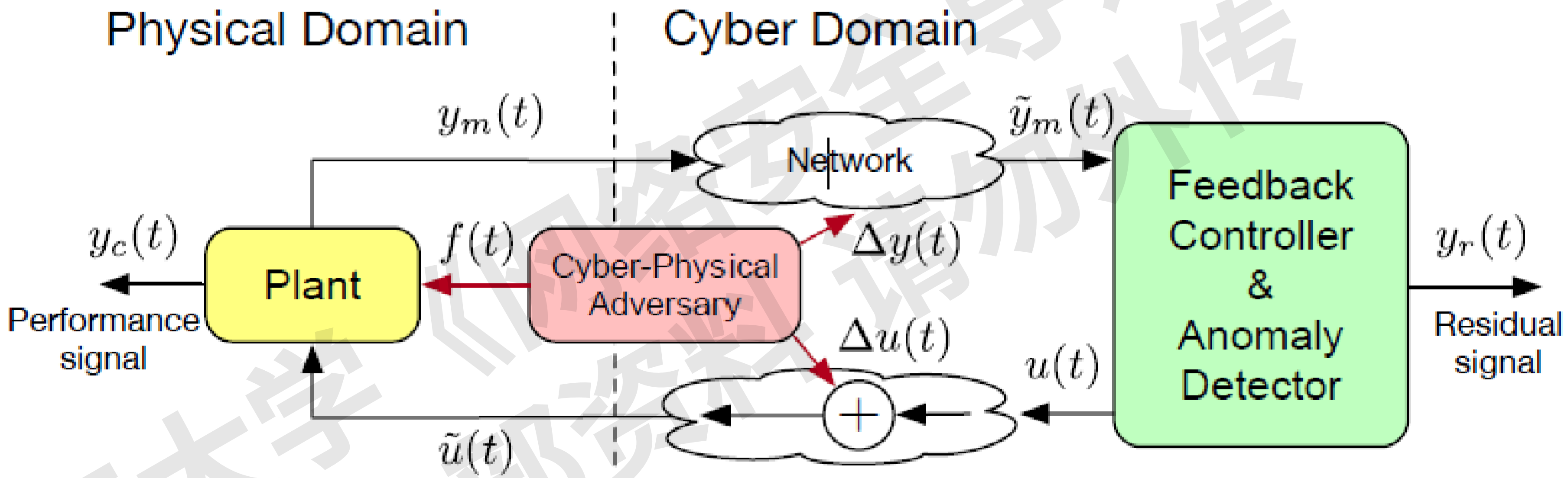
- **Undetectable attack:** $a(k)=1.1^k\begin{pmatrix}-0.282\\ 0.282\end{pmatrix}$

- **Masking initial state:** $x_0=\begin{pmatrix}-0.705\\ 0.470\\ 0.352\end{pmatrix}$

Undetectable Attack

- Suppose the operator observes the output $y(k)$, and ***does not know*** the true initial state $x(0)$ and true disturbance $d(k)$.
- Let $(x_0,\ d_0, a_0)$ be an undetectable attack, $0 = G_d d_0 + G_a a_0$ with initial state $x(0) = x_0$.
- Consider the cases:
  - **1. Un-attacked system: $y = G_d(-d_0)$, with initial state $x(0) = 0$**
  - **2. Attacked system: $y = G_a(a_0)$, with initial state $x(0) = x_0$**
  - If initial states $x(0) = 0$ and $x(0) = x_0$, disturbances $d = -d_0$ and $d = 0$ are equally likely, then it's impossible for the operator to decide which case is true! ⇒ ***Attack is undetectable!***

Undetectable Attack – Another Perspective



Undetectable Attack – Another Perspective

System dynamics:

$$x(k+1)=Ax(k)+Bu(k)+B_a a(k) \qquad a_k=\begin{bmatrix}\Delta u_k\\ \Delta y_k\end{bmatrix}$$

$$y(k)=Cx(k)+D_a a(k)$$

Output function: $y(k)=\Phi(x_0,a,k)\triangleq CA^k x_0 + C\sum_{i=1}^{k-1} A^{k-i-1}B_a a_i + D_a a_k$

Measurement trajectory under attack a: $y(k)=\Phi(x_0,a,k),\quad k\geq 0$

Definition: Attack a is *undetectable* if

$$\Phi(x_0,a,k)=\Phi(x_0^a,0,k)$$

for *some* initial state x_0^a for all $k\geq 0$

- **Interpretation:**

  Output under attack can be confused as an initial state without attack.

Undetectable Attack – Another Perspective

Undetectability requires (for all $k \geq 0$)

$$0 = \Phi(x_0, a, k) - \Phi(x_0^a, 0, k)$$

[linearity] $$0 = \Phi(x_0 - x_0^a, a, k)$$

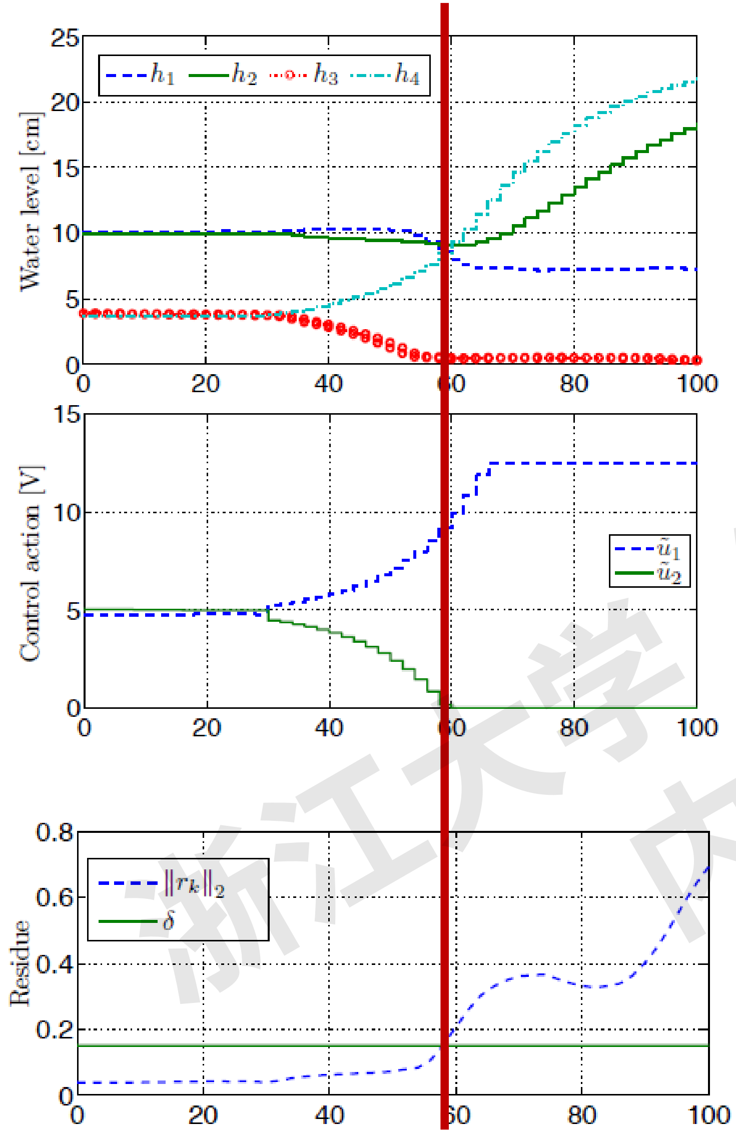
$$0 = CA^k \underbrace{(x_0 - x_0^a)}_{\triangleq \bar{x}_0^a} + C \sum_{i=1}^{k-1} A^{k-i-1} B_a a_i + D_a a_k$$

$$0 = \Phi(\bar{x}_0^a, a, k)$$

$\Leftrightarrow$ Must exist initial state $\bar{x}_0^a$ and input a_k yielding zero output.

This corresponds to the zero dynamics of the system

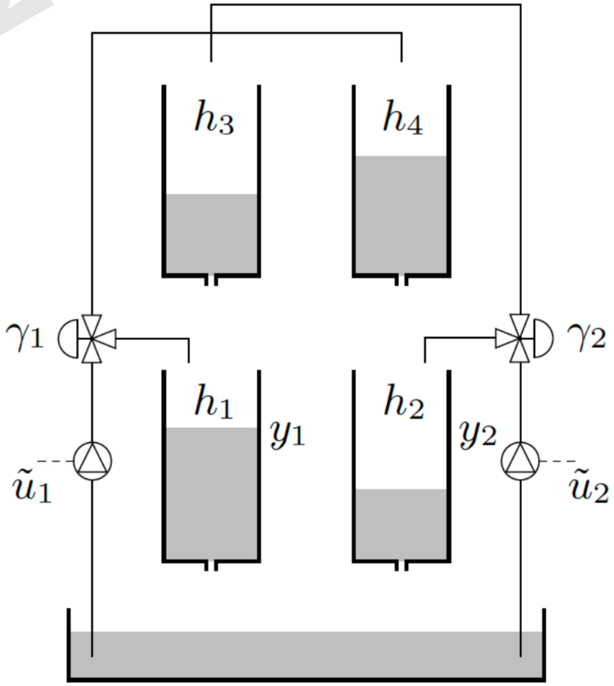
Example: Zero Dynamics Attack

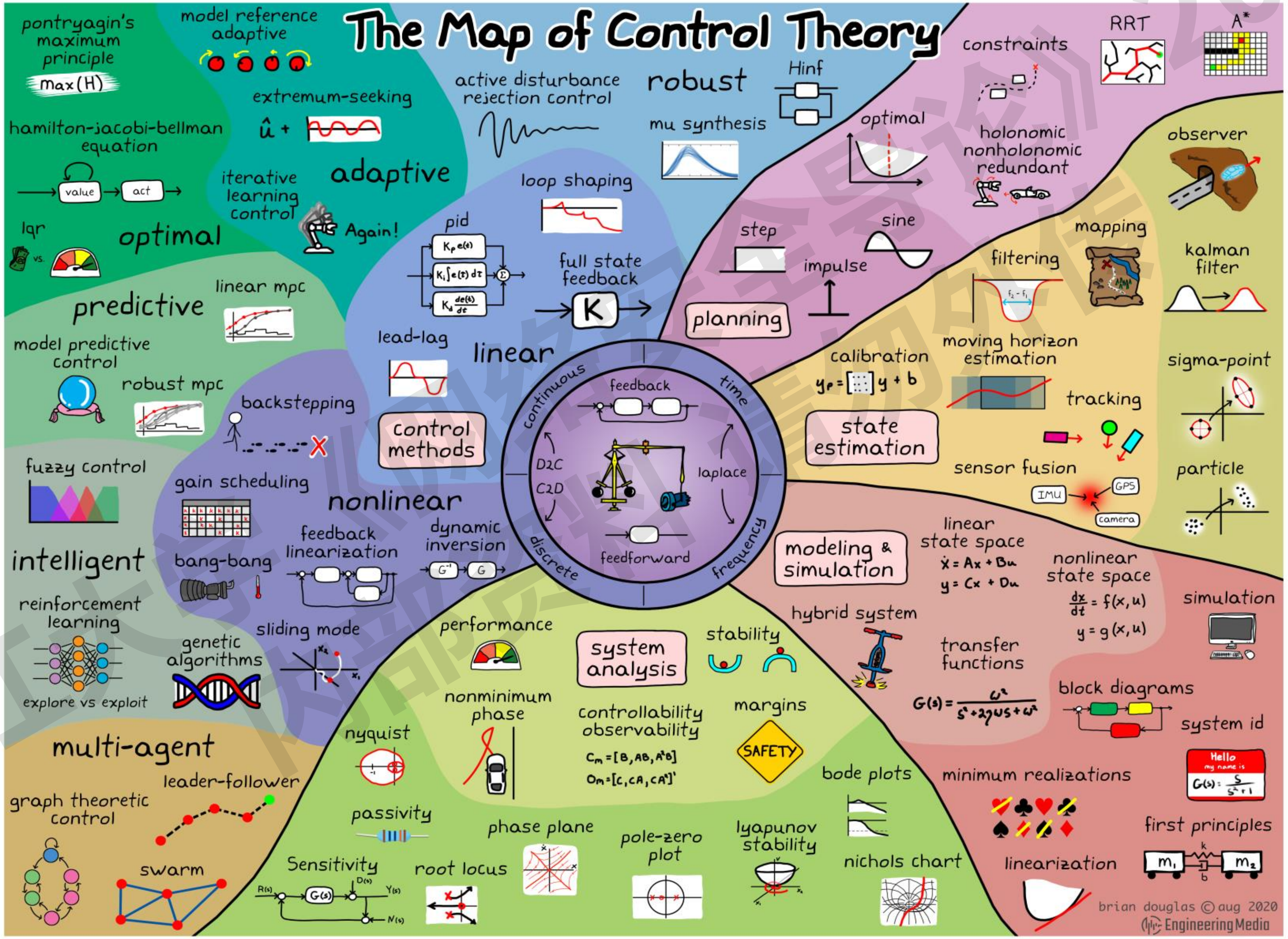


Attack Goal: Empty tank 3

Zero dynamics attack on both actuators - unstable zero

Tank 3 becomes empty





The Map of Control Theory
pontryagin's maximum principle
max(H)
model reference adaptive
extremum-seeking
hamilton-jacobi-bellman equation
value
act
iterative learning control
adaptive
Again!
lqr
vs.
optimal
predictive
linear mpc
model predictive control
robust mpc
active disturbance rejection control
robust
Hinf
mu synthesis
loop shaping
pid
full state feedback
K
lead-lag
linear
backstepping
control methods
gain scheduling
nonlinear
feedback linearization
dynamic inversion
bang-bang
sliding mode
fuzzy control
intelligent
reinforcement learning
explore vs exploit
genetic algorithms
multi-agent
leader-follower
graph theoretic control
swarm
continuous
time
feedback
D2C
C2D
laplace
feedforward
discrete
frequency
constraints
RRT
A*
optimal
holonomic
nonholonomic
redundant
step
sine
impulse
planning
observer
filtering
mapping
kalman filter
calibration
moving horizon estimation
sigma-point
state estimation
tracking
sensor fusion
IMU
GPS
camera
particle
modeling & simulation
linear state space
$\dot{x} = Ax + Bu$
$y = Cx + Du$
nonlinear state space
$\frac{dx}{dt} = f(x,u)$
$y = g(x,u)$
simulation
hybrid system
transfer functions
$G(s) = \frac{\omega^2}{s^2+2\zeta\omega s+\omega^2}$
block diagrams
system id
Hello my name is
$G(s) = \frac{s}{s^2+1}$
minimum realizations
first principles
m_1
m_2
linearization
performance
system analysis
stability
nonminimum phase
controllability observability
$C_m = [B, AB, A^2B]$
$O_m = [C, CA, CA^2]'$
margins
SAFETY
bode plots
nichols chart
nyquist
passivity
phase plane
pole-zero plot
lyapunov stability
Sensitivity
root locus
R(s)
G(s)
D(s)
Y(s)
N(s)
brian douglas © aug 2020
Engineering Media

Security Metrics for Vulnerability Assessment

- **Security Metric (安全因子) α_i:**

$$\alpha_i := \min_{|z_0|\geq 1, x_0, d_0, a_0^i} \|a_0^i\|_0$$

$$\text{subject to} \quad P(z_0)\begin{bmatrix} x_0 \\ d_0 \\ a_0^i \end{bmatrix} = 0$$

- **Notation:** $\|a\|_0 := |supp(a)|$, a^i denotes the vector a with i-th element non-zero.

- **Interpretation:**
  - Attacker persistently target signal component a_i (condition $|z_0| \geq 1$) .
  - α_i is the smallest number of signals that need to be attacked simultaneously to launch undetectable attack against a_i.

Security Metrics for Vulnerability Assessment

- **Security Metric (安全因子) α_i:**

$$\alpha_i := \min_{|z_0|\geq 1, x_0, d_0, a_0^i} \|a_0^i\|_0$$

$$\text{subject to} \quad P(z_0)\begin{bmatrix} x_0 \\ d_0 \\ a_0^i \end{bmatrix} = 0$$

- **Notation:** $\|a\|_0 := |supp(a)|$, a^i denotes the vector a with i-th element non-zero.

- **Remark 1:**
  - **NP-hard** in general, combinatorial optimization.
  - A generalized security metric form extends the static security metric in power system. [Sandberg, Teixeira]

Security Metrics

- **Security Metric (安全因子) α_i:**

$$\alpha_i := \min_{|z_0|\geq 1, x_0, d_0, a_0^i} \|a_0^i\|_0$$

$$\text{subject to} \quad P(z_0)\begin{bmatrix} x_0 \\ d_0 \\ a_0^i \end{bmatrix} = 0$$

- **Implication:**
  - **α_i is of interest to both the operator and the attacker.**
  - If the number α_i is large, it require significant coordinated resources by the attack to accomplish undetectable attacks. If α_i is small, these signals are critical!
  - **Quantitative risk assessment （量化风险评估）.**

Example (Continued)

$$A = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.9 \end{pmatrix} \quad B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.25 \end{pmatrix} \quad C = \begin{pmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0 & 0.4 \end{pmatrix}$$

- **Transfer function:** $y(z) = C(zI - A)^{-1}B + D = (G_d(z) \quad G_a(z))$

$$G_d(z) = (), \quad G_a(z) = \begin{pmatrix} \frac{0.2}{z-0.9} & \frac{0.3}{z-0.8} \\ \frac{0.1}{z-0.9} & \frac{0.1}{z-0.9} \end{pmatrix}$$

- **Invariant Zeros:** $\sigma\big(P(z)\big) = \{1.1\}$

- **Undetectable attack:** $a(k) = 1.1^k \begin{pmatrix} -0.282 \\ 0.282 \end{pmatrix} \Rightarrow a_0 = \begin{pmatrix} -0.282 \\ 0.282 \end{pmatrix}$

- **Masking initial state:** $x_0 = \begin{pmatrix} -0.705 \\ 0.470 \\ 0.352 \end{pmatrix}$

Example (Continued)

$$A = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.9 \end{pmatrix} \quad B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.25 \end{pmatrix} \quad C = \begin{pmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0 & 0.4 \end{pmatrix}$$

- **Transfer function:** $y(z) = C(zI - A)^{-1}B + D = (G_d(z) \quad G_a(z))$

$$G_d(z) = (), \qquad G_a(z) = \begin{pmatrix} \frac{0.2}{z-0.9} & \frac{0.3}{z-0.8} \\ \frac{0.1}{z-0.9} & \frac{0.1}{z-0.9} \end{pmatrix}$$

- **Invariant Zeros:** $\sigma\big(P(z)\big) = \{1.1\}$

- **Undetectable attack:** $a(k) = 1.1^k \begin{pmatrix} -0.282 \\ 0.282 \end{pmatrix} \Rightarrow a_0 = \begin{pmatrix} -0.282 \\ 0.282 \end{pmatrix}$

- Only one signal satisfies α_i constraints! Thus $\|a_0\|_0 = 2$, $\alpha_{1,2} = 2$!

Special Case: Sensor Attacks for Static System

$$P(z) = \begin{bmatrix} I - zI & 0 & 0 \\ C & 0 & D_a \end{bmatrix}$$

- $A = I$, $B_d = B_a = D_d = 0$ (only sensors attacked), this is the ***steady-state*** case.
- Space of eigenvectors x_0 is n-dimensional ⇒ Typically makes computation of α_i harder than in the dynamical case!
- Particular relevant case in ***Power Systems State Estimation*** under steady-state power flow model.

Special Case: Sensor Attacks for Static System

$$P(z) = \begin{bmatrix} I - zI & 0 & 0 \\ C & 0 & D_a \end{bmatrix}$$

- Particular relevant case in ***Power Systems State Estimation*** under steady-state power flow model.
  - $y = Cx$, the noise-less measurement model.
  - Computation of α_i is NP-hard, but power systems impose special structures in C (DC power flow matrix).
  - It recovers the state estimation security metric, where

$$\alpha_i := \min_{x_0, a} \|a\|_0$$

$$\text{subject to } Cx_0 + a = 0$$

$$a(i) \neq 0.$$

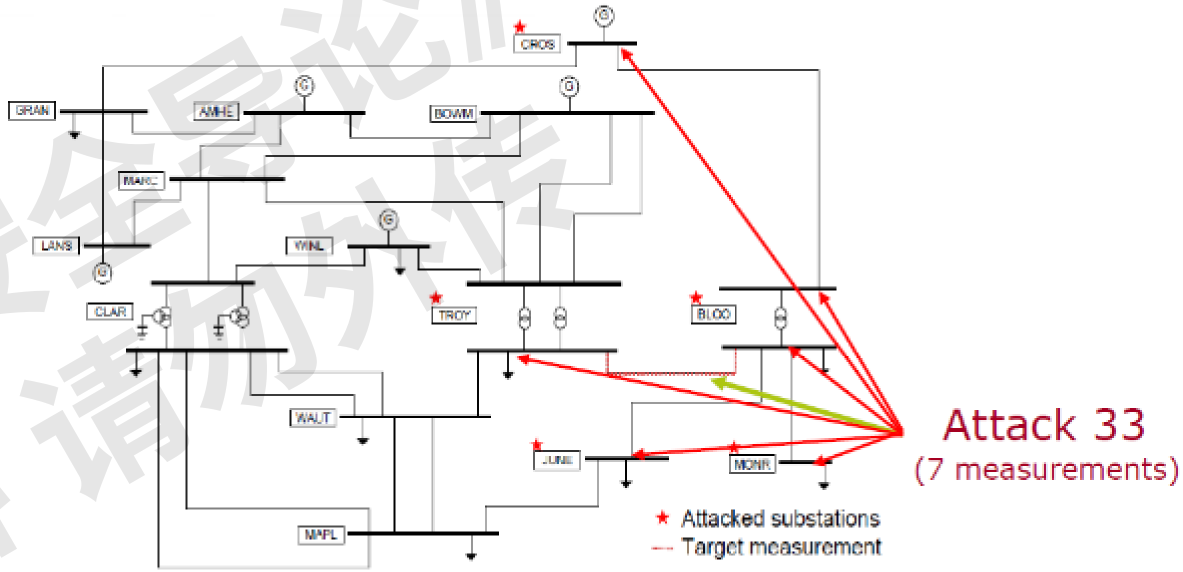
Example of Likelihood Metrics

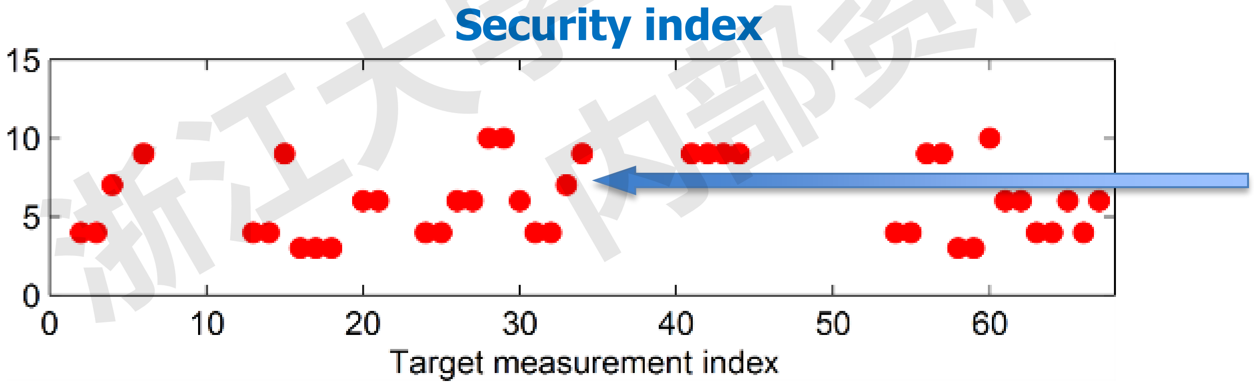
– How many measurements must be corrupted to remain stealthy?

$$\alpha_i := \min_a \|a\|_0$$

$$\text{subject to} \quad a = H \triangle x,$$

$$\text{a(i)} \neq 0.$$

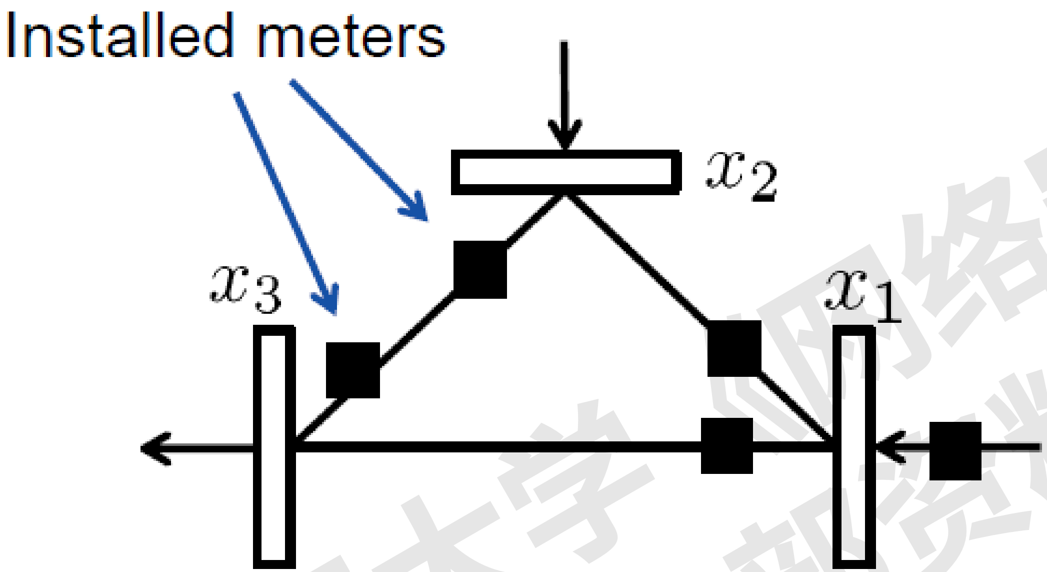




At least **7 measurements** involved in a stealth attack against measurement 33.

H. Sandberg, A. Teixeira, K. H. Johansson, "On security indices for state estimators in power networks," in First Workshop on Secure Control Systems (CPSWEEK 2010), Stockholm, Sweden, 2010.

Example: Power System State Estimation



$$x = [x_1, x_2, x_3]^T \in \mathbb{R}^3$$

$$y = [y_1, y_2, y_3, y_4, y_5]^T \in \mathbb{R}^5$$

$$C \in \mathbb{R}^{5\times 3}$$

$$C = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ 2 & -1 & -1 \end{pmatrix}$$

$$D_a = I_5$$

Example: Power System State Estimation

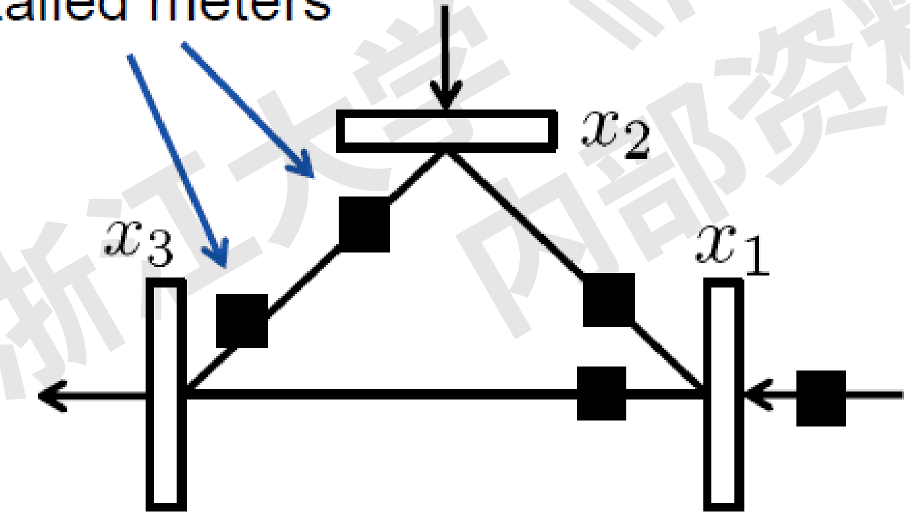
$$\alpha_i := \min_{x_0, a_0} \|a_0\|_0$$

subject to

$$0 = Cx_0 + D_a a_0$$

$$a_0(i) \neq 0,$$

Installed meters



For meter 1: $\alpha_1 = 3$

For meter 2: $\alpha_2 = 3$

For meter 3: $\alpha_3 = 4$

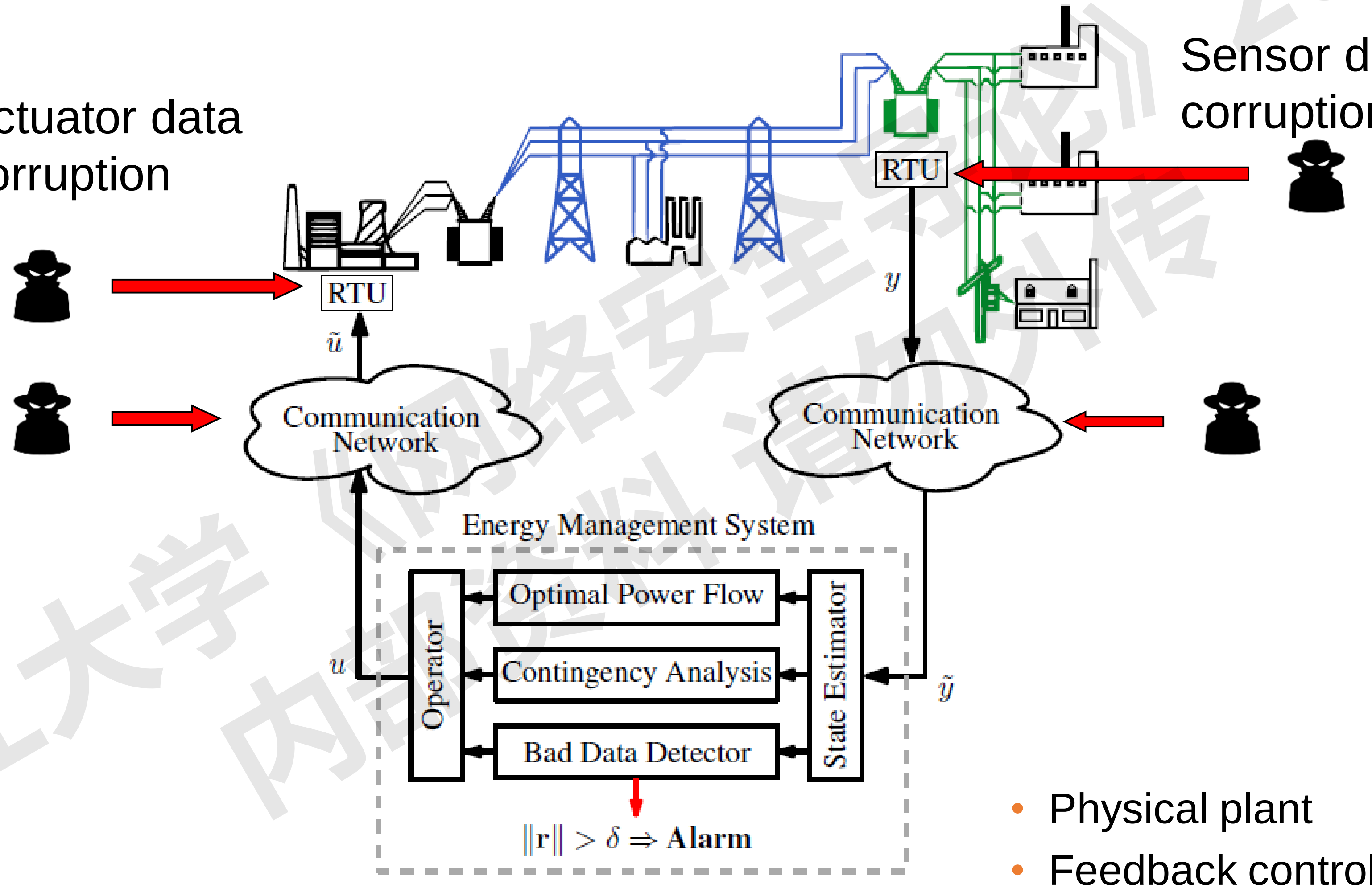
For meter 4: $\alpha_4 = 4$

For meter 5: $\alpha_5 = 3$

Example: Power System State Estimation

Actuator data corruption

Sensor data corruption



- Physical plant
- Feedback controller
- Anomaly detector

Example: Power System State Estimation

- **Simplifications:**
  - $\sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$
  - $V_i = 1pu$
  - No resistances or shunt elements

- **Only active power:**
  - $P_i = \sum B_{ij}(\theta_i - \theta_j)$
  - $P_{ij} = B_{ij}(\theta_i - \theta_j)$
  - Similar to a DC resistive network

- Noiseless Measurement model:

$$y = Cx$$

- Linear Least Squares Estimator:

$$\hat{x} = (C^T C)^{-1} C^T y$$

- Measurement residual:

$$r = y - C\hat{x}$$

- Bad Data Detector:

$$\|Wr(\hat{x})\|_p \gtrless \tau$$

Example: Power System State Estimation

$$C = \begin{bmatrix} P_1 D B^T \\ -P_2 D B^T \\ P_3 B D B^T \end{bmatrix} \begin{matrix} \text{(power flow measurements, “from” side)} \\ \text{(power flow measurements, “to” side)} \\ \text{(power injection measurements)} \end{matrix}$$

- B – directed incidence matrix of graph corresponding to power network topology.
- D – nonsingular diagonal matrix containing reciprocals of reactance of transmission lines.
- P_1, P_2, P_3 – meter selection matrices (rows of identity matrices).

More measurements than states, $p > n$, full redundancy!

Applies to all potential flow networks, e.g., water, gas…

Example: Power System State Estimation

$$C = \begin{bmatrix} P_1 D B^T \\ -P_2 D B^T \\ P_3 B D B^T \end{bmatrix}$$

(power flow measurements, "from" side)
(power flow measurements, "to" side)
(power injection measurements)

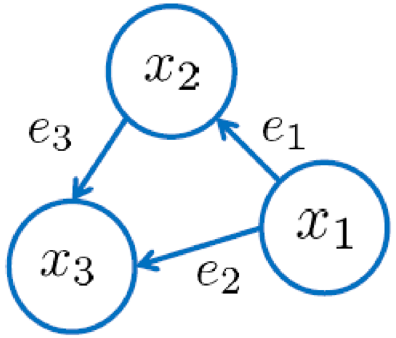
- B – directed incidence matrix of graph corresponding to power network topology.
- D – nonsingular diagonal matrix containing reciprocals of reactance of transmission lines.
- P_1, P_2, P_3 – meter selection matrices (rows of identity matrices).

More measurements than states, $p > n$, full redundancy!

Applies to all potential flow networks, e.g., water, gas…

Example: Power System State Estimation

- **DC power flow measurement matrix**



$$B^T = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

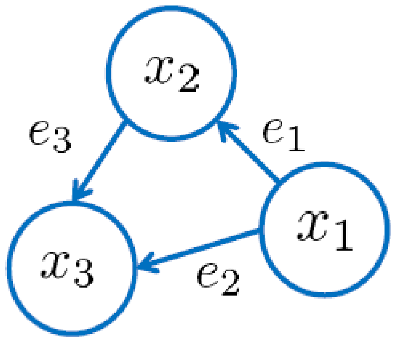
$$D = I$$

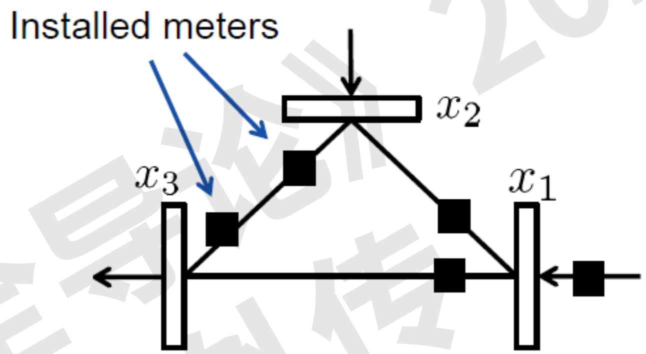
(all the reactances are 1)

Branch power flows: $$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = DB^T x = \begin{pmatrix} x_1 - x_2 \\ x_1 - x_3 \\ x_2 - x_3 \end{pmatrix}$$

Node Injections: $$BDB^T x = B \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} e_1 + e_2 \\ -e_1 + e_3 \\ -e_2 - e_3 \end{pmatrix}$$

Example: Power System State Estimation





$$C = \begin{bmatrix} P_1 D B^T \\ -P_2 D B^T \\ P_3 B D B^T \end{bmatrix}$$

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$B^T = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

$$P_2 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

$$D = \begin{pmatrix} 1/r_{12} & 0 & 0 \\ 0 & 1/r_{13} & 0 \\ 0 & 0 & 1/r_{23} \end{pmatrix}$$

$$P_3 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$$

r_{ij}: reactance of the branch

Example: Power System State Estimation

Security Index Computation

- **Method 1:** ℓ_1-norm relaxation for an approximate solution
  - The original security index problem: non-convex; ℓ_0-norm sparsity.
  - Minimization of the ℓ_1-norm always give rise to sparse solutions.
  - ℓ_1-norm relaxation is a good compromise between a sparse and a small attack vector.
  - $\|a\|_1 := \sum_i a(i)$

- A convex optimization problem:

$$\Rightarrow \boxed{\begin{aligned} \beta_i := \min_{a,\triangle x} \|a\|_1 \\ \text{subject to} \quad a = H \triangle x, \\ \mathrm{a(i)} = 1. \end{aligned}}$$

- It can be re-cast to a linear program (LP) （线性规划）.
- How is the relation between these two index, α_i and β_i?
- ℓ_1-norm relaxation provide an overestimate (upper bound) of the security index.

Security Index Computation

- **Method 2:** Big M method formulation

$$\alpha_i := \min_{a,\triangle x} \|a\|_0$$

subject to

$$a = H \triangle x,$$

$$\mathrm{a(i)} = 1.$$

$$\Rightarrow$$

$$\alpha_i := \min_{z_k,\triangle x,a} \sum_k z_k$$

subject to

$$a = H \triangle x,$$

$$\mathrm{a(i)} = 1,$$

$$-Mz_k \leq \mathrm{a} \leq Mz_k$$

$$z_k \in \{0,1\}$$

Elementwise

"∞"

- A mixed integer linear programming (MILP) problem. **(混合整数线性规划)**
- M is user-defined, greater than the maximum entry of $H\Delta x$ in absolute value.
- Not scale well. For large power system, the computation time explodes!

Optimization Solver

- Install the solver CPLEX and add the path in Matlab.
- Function *cplexmilp*

Detailed Description

Solve mixed integer linear programming problems.

```
x = cplexmilp(f,Aineq,bineq)
x = cplexmilp(f,Aineq,bineq,Aeq,beq)
x = cplexmilp(f,Aineq,bineq,Aeq,beq,sostype,sosind,soswt)
x = cplexmilp(f,Aineq,bineq,Aeq,beq,sostype,sosind,soswt,lb,ub)
x = cplexmilp(f,Aineq,bineq,Aeq,beq,sostype,sosind,soswt,lb,ub,ctype)
x = cplexmilp(f,Aineq,bineq,Aeq,beq,sostype,sosind,soswt,lb,ub,ctype,x0)
x = cplexmilp(f,Aineq,bineq,Aeq,beq,sostype,sosind,soswt,lb,ub,ctype,x0,options)
x = cplexmilp(problem)
[x,fval] = cplexmilp(...)
[x,fval,exitflag] = cplexmilp(...)
[x,fval,exitflag,output] = cplexmilp(...)
```

Finds the minimum of a problem specified by

```
min      f*x
st.      Aineq*x <= bineq
         Aeq*x    = beq
         lb <= x <= ub
```

Security Metric Computation

Build DC power flow matrix

- B – directed incidence matrix of graph corresponding to power network topology.
- D – nonsingular diagonal matrix containing reciprocals of reactance of transmission lines.

$$H = \begin{bmatrix} P_1 D B^T \\ -P_2 D B^T \\ P_3 B D B^T \end{bmatrix}$$

```
%% branch data
%   fbus   tbus    r      x       b      rateA  rateB  rateC  ratio  angle  status
mpc.branch = [
    1   4   0       0.0576  0       250 250 250 0   0   1   -360    360;
    4   5   0.017   0.092   0.158   250 250 250 0   0   1   -360    360;
    5   6   0.039   0.17    0.358   150 150 150 0   0   1   -360    360;
    3   6   0       0.0586  0       300 300 300 0   0   1   -360    360;
    6   7   0.0119  0.1008  0.209   150 150 150 0   0   1   -360    360;
    7   8   0.0085  0.072   0.149   250 250 250 0   0   1   -360    360;
    8   2   0       0.0625  0       250 250 250 0   0   1   -360    360;
    8   9   0.032   0.161   0.306   250 250 250 0   0   1   -360    360;
    9   4   0.01    0.085   0.176   250 250 250 0   0   1   -360    360;
];
```

Reactance values of each branch.

Compute Security Metric

- Big *M* method formulation

$$\alpha_i := \min_{z_k, \mathrm{V}x, a} \sum_k z_k$$

subject to

$$a = H\mathrm{V}x,$$

$$a(i) = 1,$$

$$-Mz_k \le a \le Mz_k$$

$$z_k \in \{0,1\}$$

Elementwise

$$"\infty"$$

$$\Rightarrow$$

Finds the minimum of a problem specified by

```
min     f*x
st.     Aineq*x <= bineq
        Aeq*x    = beq
        lb <= x <= ub
```

Parameters:

`f`	Double column vector for linear objective function
`Aineq`	Double matrix for linear inequality constraints
`bineq`	Double column vector for linear inequality constraints
`Aeq`	Double matrix for linear equality constraints
`beq`	Double column vector for linear equality constraints

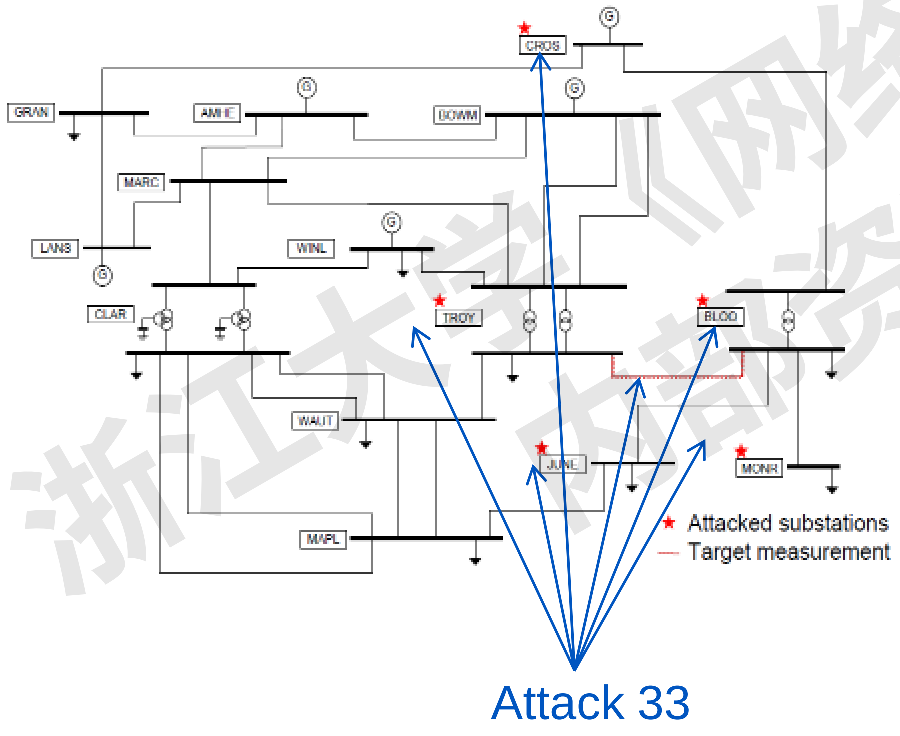
Source: *https://www.ibm.com/support/knowledgecenter/SSSA5P_12.6.2/ilog.odms.cplex.help/refmatlabcplex/html/cplexmilp-m.html*

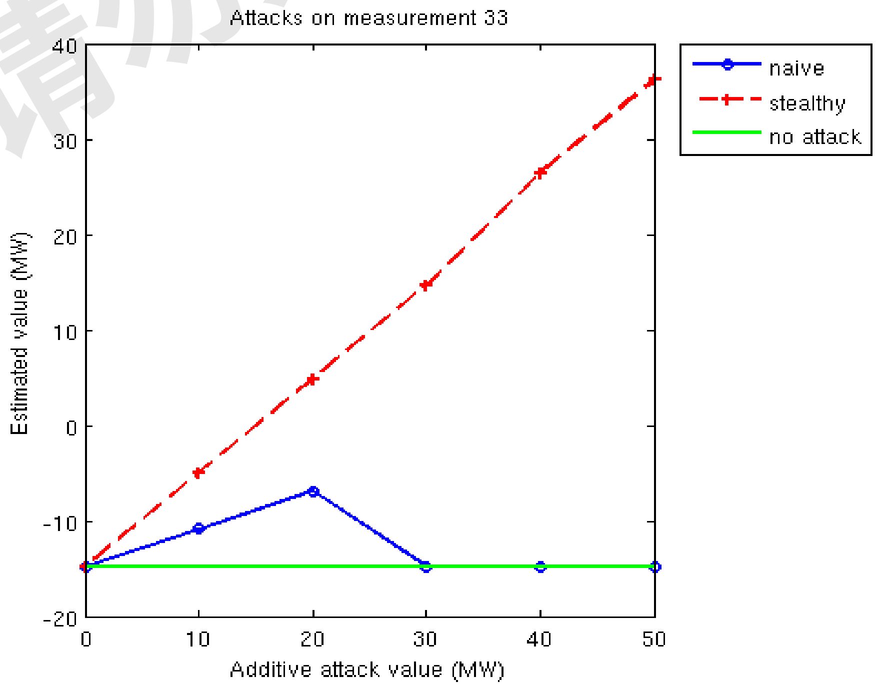
4.3 攻击检测与应用

ROBUST DYNAMIC ATTACK DETECTION

Motivation of Attack Detection

- **Undetectable Sensor Attacks vs. Static detection**
  - Attacks satisfy spatial correlations of the measurements. ✓
  - Static detector **（静态检测）**
    - detects corrupted measurements based on its statistical properties at each time step.

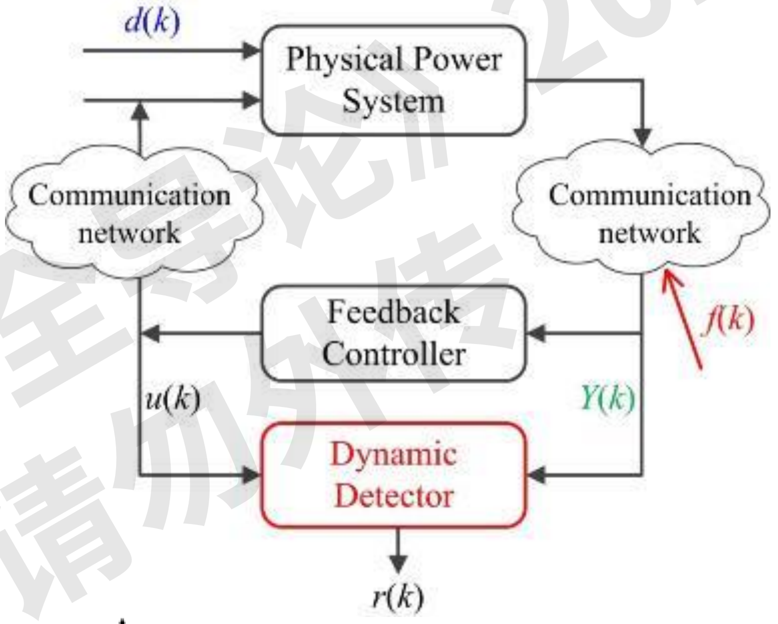


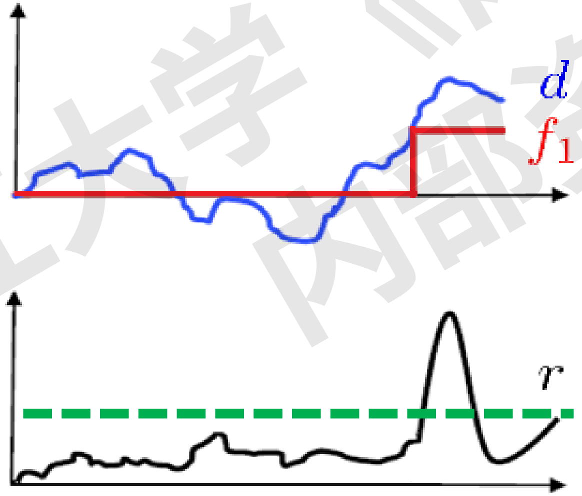


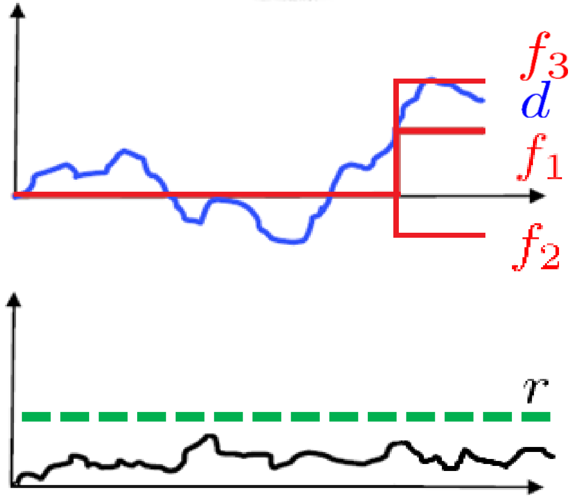
Challenges

Power system dynamics

- Large-scale, large number of dynamic states;
- Unknown disturbance, non-zero initial conditions;
- Multivariate attack signals.







State of the Art

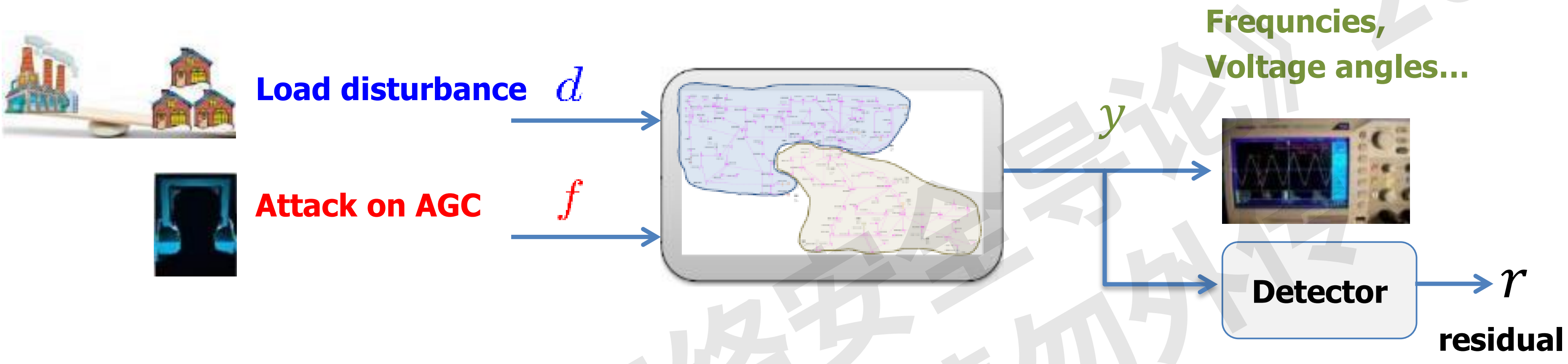
Existing methods

- Statistical method, e.g., cumulative sum – type [Li (TSG15)]
- Additional information from PMU, load forecasting, generation schedules, etc. [Ashok (TSG 16), Zhao (TPS 18)]
- Matrix separation (sparse optimization) [Liu (TSG 14)]
- Machine learning method, e.g., deep neural networks [James (TII 18)]

Related methods

- Linear model, observer based, [Nyberg (TAC 06)]
- Nonlinear model, scalable optimization based, [Peyman (TAC 16)]

Model Description



$$\begin{cases} X(k+1) = A_{cl}X(k) + B_d d(k) + B_f f(k) \\ Y(k) = CX(k) + D_f f(k) \end{cases}$$

$x(k) = \begin{bmatrix} X(k) \\ d(k) \end{bmatrix}$ → Unknown signals (disturbance & states)

$y(k) := Y(k)$ → Known signals (measurements, etc.)

$$L(q) := \begin{bmatrix} 0 \\ -I \end{bmatrix} \qquad H(q) := \begin{bmatrix} A_{cl} - qI & B_d \\ C & 0 \end{bmatrix} \qquad F(q) := \begin{bmatrix} B_f \\ D_f \end{bmatrix}$$

$$H(q)x(k) + L(q)y(k) + F(q)f(k) = 0$$

(DAE: Difference-Algebraic Equations)

Attack Detector Contruction

$$H(q)x(k) + L(q)y(k) + F(q)f(k) = 0$$

$$r(k) := \underbrace{N(q)L(q)y(k)}_{\text{Dynamic detector}}$$

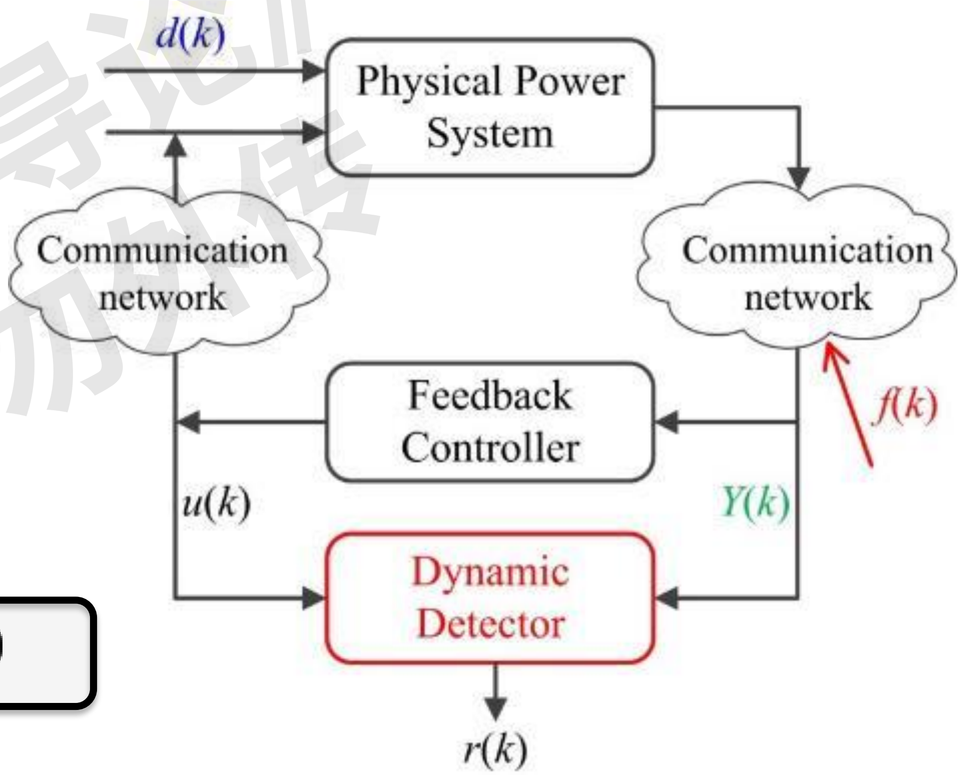
$$r(k) = -N(q)H(q)x(k) - N(q)F(q)F_f\alpha$$

$$\boxed{N(q)H(q) \equiv 0}$$

(I)

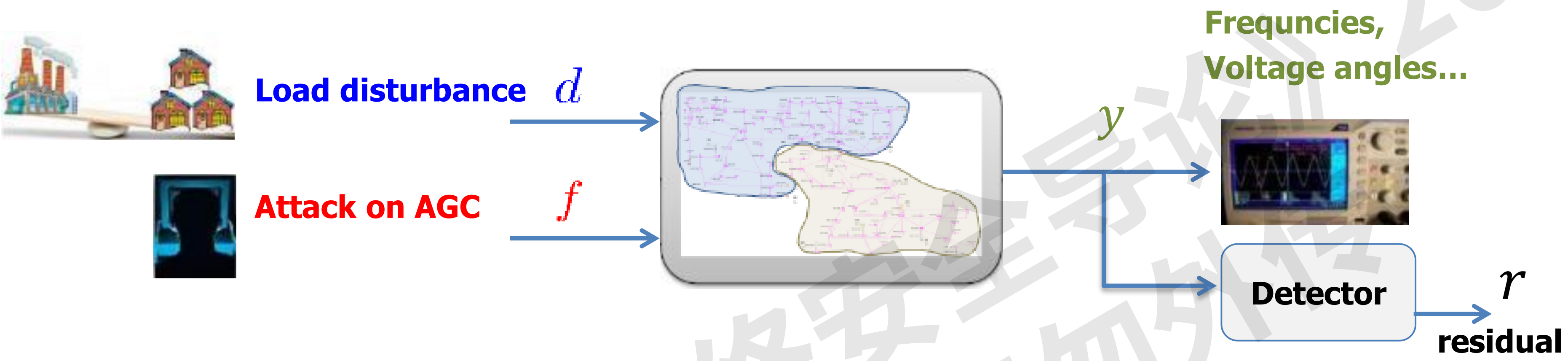
$$\boxed{\text{for all } \alpha,\ N(q)F(q)F_f\alpha \neq 0}$$

(II)



- (I): decoupled from internal system states and unknown disturbances;
- (II): robust detection of multivariate attack.

Robust Attack Detector



- (I): decoupled from internal system states and unknown disturbances;
- (II): robust detection of multivariate attack.

Maximin Optimization*:

$$\gamma^\star := \max_{N\in\mathcal{N}} \min_{\alpha\in\mathcal{A}} \left\{\mathcal{J}(N,\alpha)\right\}$$

Detector　　Attacker

Attack Detector – Transient Behavior

$$\mathcal{J}(N, \alpha) := \|\bar{N} V(\alpha)\|_\infty$$

Two Conclusive Messages:

- If the program (*) is feasible with $\gamma^\star > 0$, the filter catches the intruder **<u>even if the attacker knows everything about the detector!</u>**

- If the program (*) is infeasible or $\gamma^\star = 0$, there is no such diagnosis filter **<u>even if the attacker is blind to the detector!</u>**

Attack Detector –Steady-state Behavior

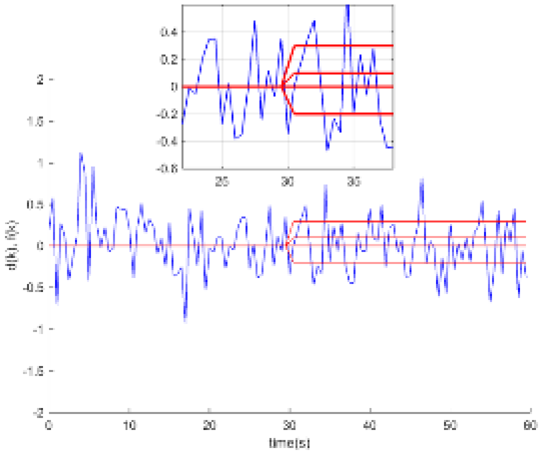
$$\mathcal{J}(N,\alpha) := |\bar{N}\bar{F}\alpha|$$

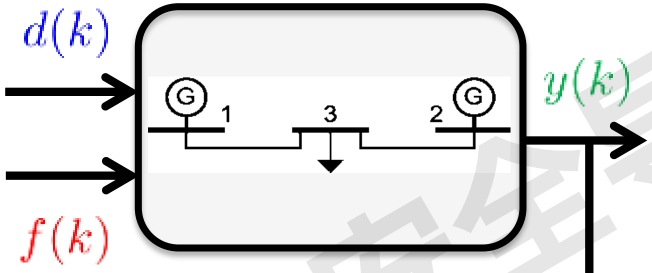
Nash equilibrium （纳什均衡，静态双人零和博弈）

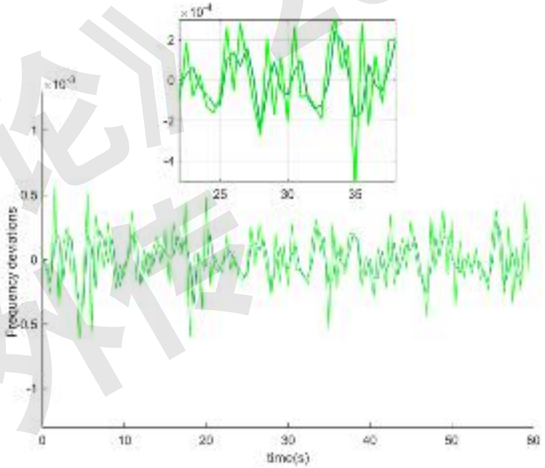
- If the **maximin** and **minimax** programs admit a positive optimal value, then detector can **detect all the admissible multivariate attacks along with a non-zero steady-state residual level;**

- If the optimal values coincide with zero, then there is **no linear detector being able to decouple the admissible attack from the natural disturbances in a long-term horizon.**

Numerical Results

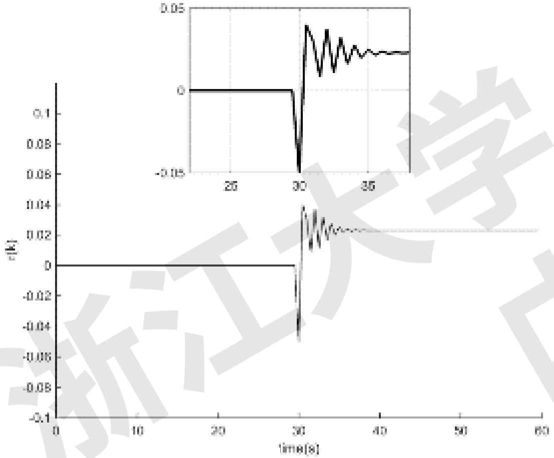


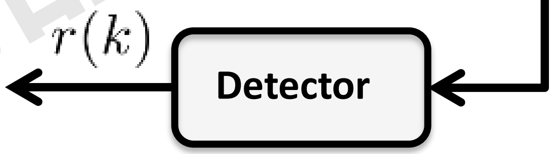


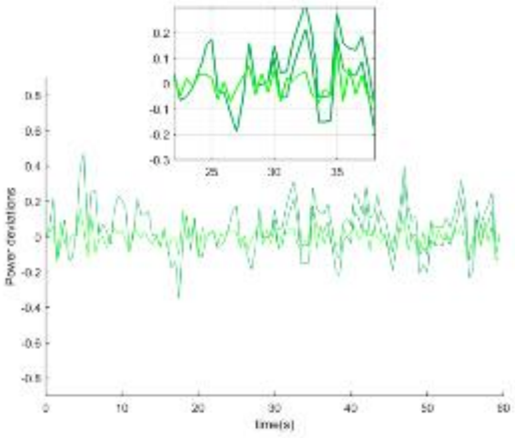


Case study:

- Load frequency control
- Three-area 39-bus system
- $y(k)$: freq. & mech. power
- $d(k)$: load disturbances
- $f(k)$: multivariate attack
- $r(k)$: diagnostic signal





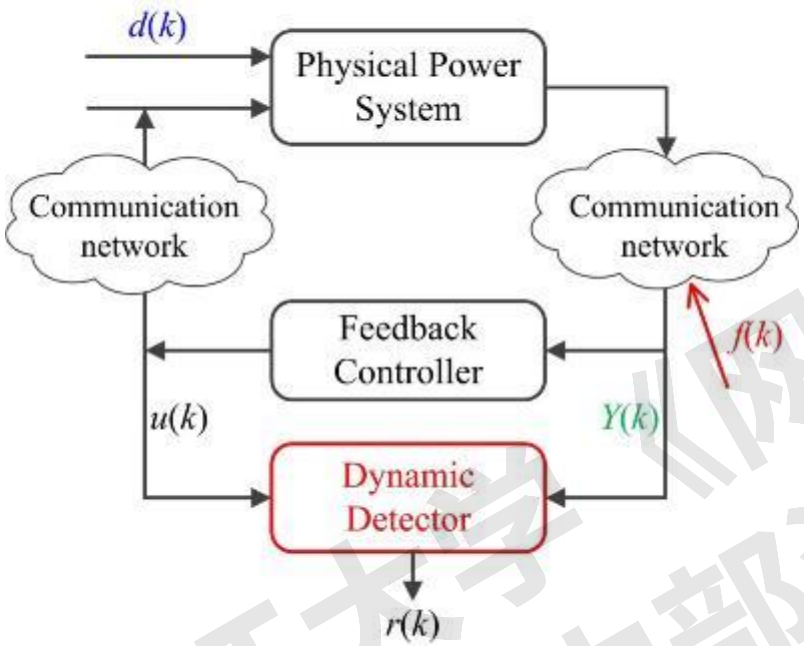


Alarm!

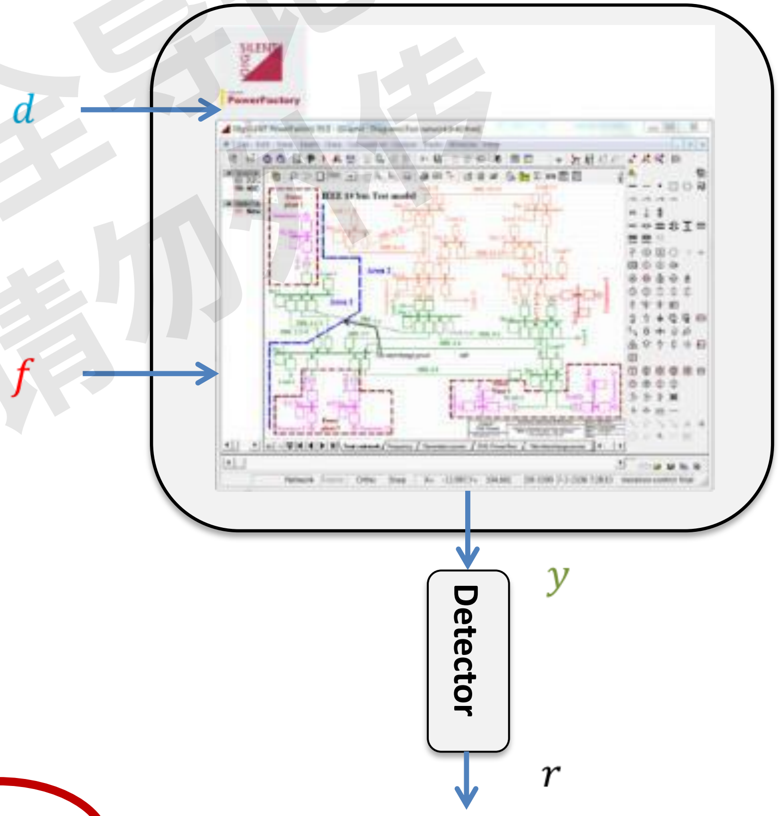
Blind!

Robustness?

- **What if model mismatch?**

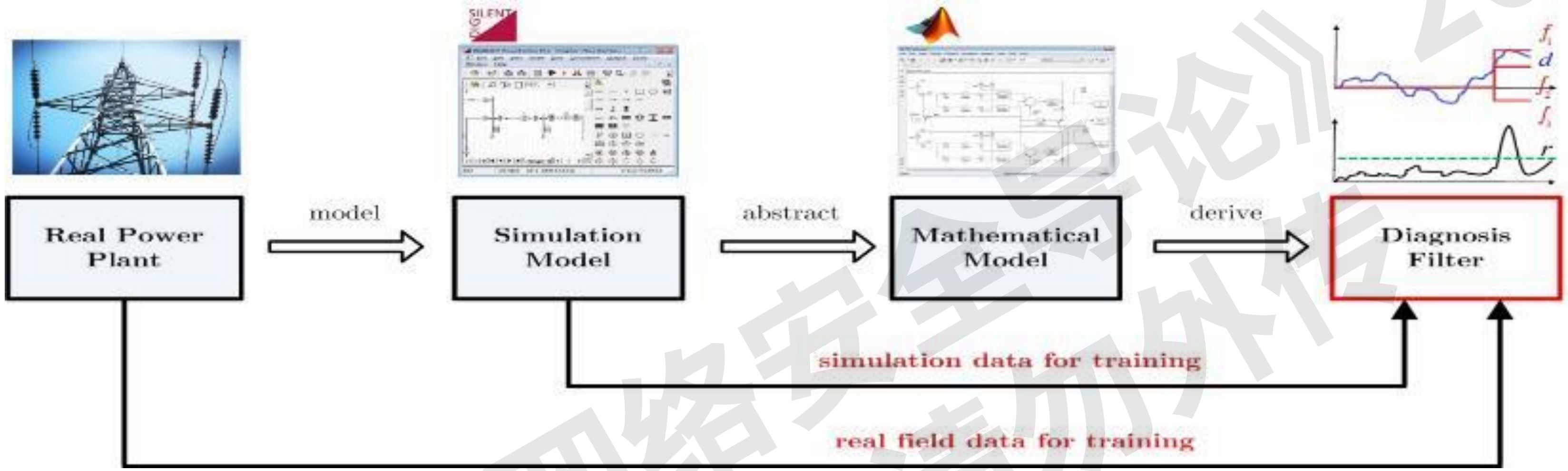


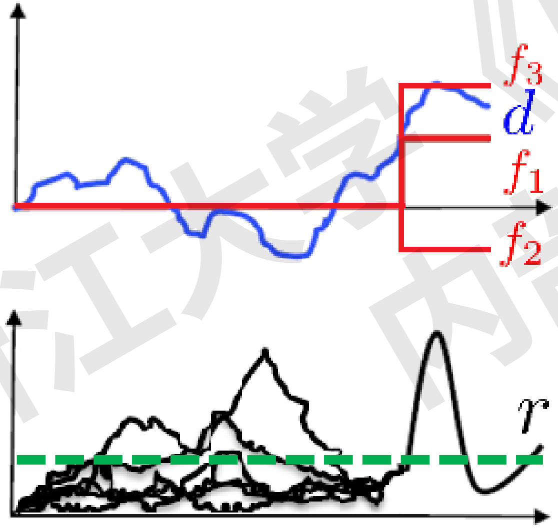
$$H(q)x(k) + L(q)y(k) + F(q)f(k) = 0$$



Model mismatch signature! $E(x(k)) + H(q)x(k) + L(q)y(k) + F(q)f(k) = 0$

Robustness?





Effects of model mismatch

$$E(x(k)) + H(q)x(k) + L(q)y(k) + F(q)f(k) = 0$$

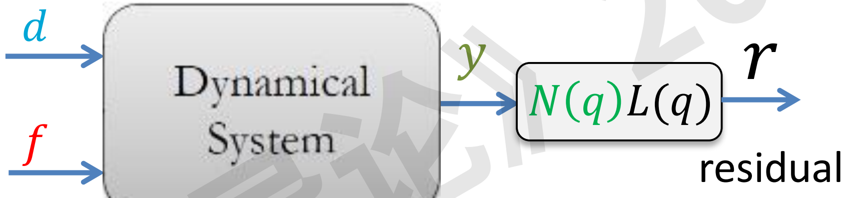
$$r(k) := N(q)L(q)y(k)$$

$$r(k) = -N(q)H(q)x(k) - N(q)F(q)F_f\alpha$$
$$- N(q)E(x(k)) \quad \text{(III)}$$

$$\boxed{\|(\text{III})\|^2_{\mathcal{L}_2} = N^T Q_d N}$$

Robustness? Convex Optimization Method

$$\underbrace{E(b)}_{\text{model mismatch}} + H(q)b + L(q)y + F(q)f = 0$$

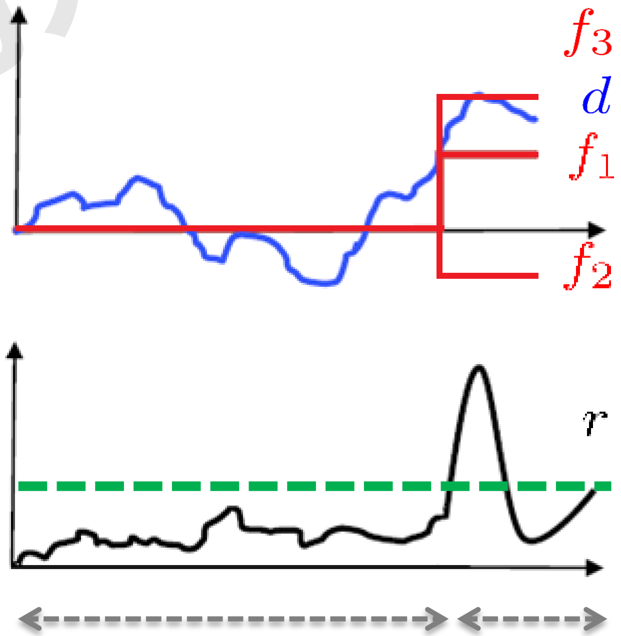


$$b := \begin{bmatrix} x \\ d \end{bmatrix}$$

$$\begin{cases} N(q)H(q) \equiv 0 \\ r = \underline{N(q)L(q)y} = \underbrace{-N(q)E(b)}_{\text{(I)}} \underbrace{-\,N(q)F(q)f}_{\text{(II)}} \end{cases}$$

$$\begin{cases} \min\limits_{N,\gamma} \quad \gamma \\ s.t. \quad N^T Q_{d_i} N \leq \gamma, \quad \forall\, i \leq m \\ \qquad \|NF\|_\infty \geq 1, \\ \qquad NH = 0. \end{cases}$$

$\Longrightarrow$ Optimizer N^*



K. Pan, P. Palensky and P. M. Esfahani, "Dynamic Anomaly Detection with High-fidelity Simulators: A Covex Optimization based Approach," ***IEEE Transactions on Smart Grid***, 2021.

Summary

- Smart grid/control system cyber security risk management.
- Undetectable attacks and masking initial states and disturbances.
- Security index α_i in control system and power system state estimation, and its computation.
- A robust detection approach for undetectable sensor attacks, utilizing the system dynamics information.