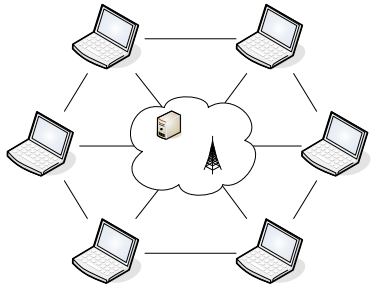
Identifying Cyber Attacks via Local Model Information

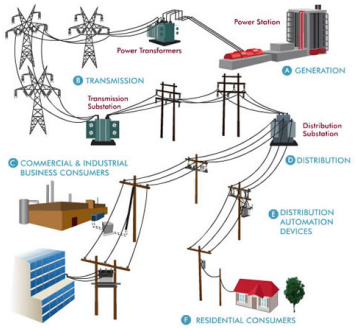
Fabio Pasqualetti
Ruggero Carli Antonio Bicchi Francesco Bullo

Center for Control, Dynamical systems and Computation
University of California, Santa Barbara

Conference on Decision and Control, Atlanta, GA, USA
December 15 - 17, 2010

Network systems and cyber-physical attacks





- *Physical attack:* compromise networks dynamics
- *Cyber attack:* compromise integrity of data

How do we enforce security and reliability?

Linear networks with misbehaving components

Network dynamics

$$x(t+1) = Ax(t) + Bu(t)$$

- if A is row stochastic and primitive, then *consensus* algorithm
- physical and cyber attacks are modeled by an exogenous input
- the nonzero entries of B determine the *misbehaving* components

Monitor measurements model

$$y_j(t) = C_j x(t)$$

- the output matrix is determined by the interaction topology

Detection and identification: known results

By knowing the system matrix A, observer j

- detects the attack B if and only if (A, B, C_j) has no zero dynamics
- identifies the attack B if and only if $(A, [B\ \bar{B}], C_j)$ has no zero dynamics, where $\bar{B}$ is any possible attack matrix
- implements effective combinatorial security controls

For a network A with connectivity k, there exists

- an undetectable attack if $\text{rank}(B) \geq k$
- an unidentifiable attack if $\text{rank}(B) \geq \lceil k/2 \rceil$

S. Sundaram and C. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Transactions on Automatic Control*

F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control* (submitted)

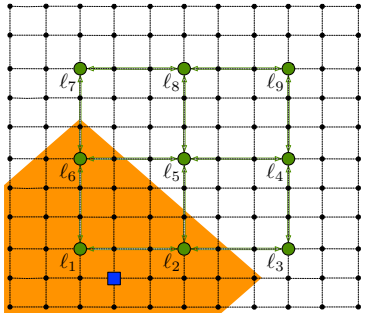
Detection and identification via local model information

Limitation of existing procedures

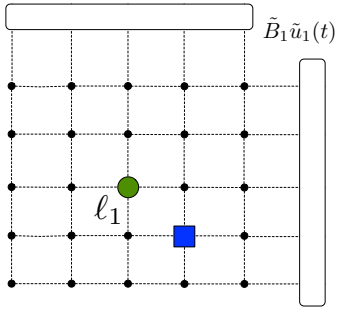
- the whole network topology has to be known by each observer
- high computational complexity
- numerically unreliable

Our approach

- geographically deploy leaders
- cooperation through a leader graph
- implement local security controls



Difficulty of a local approach

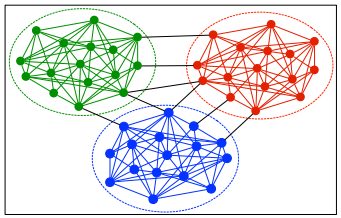


- the unmodeled dynamics act as an unknown and unmeasurable input

$$x_1(t+1) = A_1 x_1(t) + B_1 u_1(t) + \tilde{B}_1 \tilde{u}_1(t)$$

- Detection via network decomposition: $\tilde{B}_1\tilde{u}_1(t)$ is "small"
- Detection via leaders cooperation: $\tilde{B}_1\tilde{u}_1(t)$ is "large"

Weakly interconnected subnetworks



- Identify groups of strongly interacting components
- assign one leader to each group

$$A = \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_m \end{bmatrix} + \varepsilon \begin{bmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1m} \\ \Delta_{21} & \ddots & \ddots & \Delta_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ \Delta_{m1} & \Delta_{m2} & \cdots & \Delta_{mm} \end{bmatrix}$$

- if ε is "small", then different groups are weakly coupled
- determine the network partitioning that minimizes ε
- if A is a consensus matrix, then $\|\Delta\|_\infty = 2$

Local detection and identification

Each leader performs detection of misbehaving parts inside block

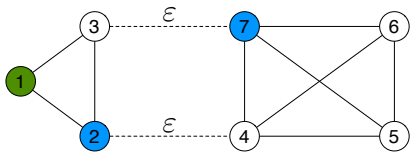
- neglect interaction with other blocks

Theorem (Local detectability)

Given $u_{\max}$, there exists α such that, if each input signal takes value in $\{u : \varepsilon\alpha u_{\max} \leq \|u\|_\infty \leq u_{\max}\}$, then local detection is successful

- there exists a critical value ε^* below which the local effect of misbehaving parts is larger than the effect of unmodeled dynamics

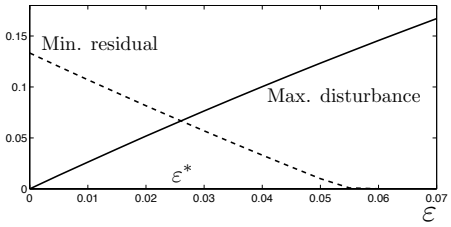
An example of detection via network decomposition



Residual generator

$$w(t+1) = Fw(t) + Ey_{\ell_1}(t)$$
$$r(t) = Mw(t) + Hy_{\ell_1}(t)$$

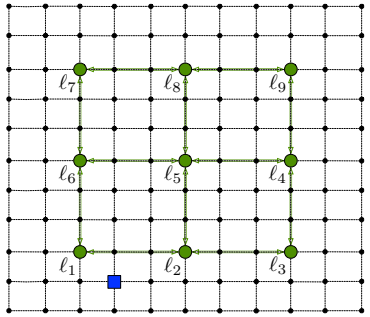
$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3}-\varepsilon & \frac{1}{3} & \varepsilon & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}-\varepsilon & 0 & 0 & 0 & \varepsilon \\ 0 & 0 & \varepsilon & \frac{1}{4} & \frac{1}{4}-\varepsilon & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \varepsilon & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}-\varepsilon \end{bmatrix}$$



- if $\varepsilon < \varepsilon^*$, then observer 1 detects and identifies the misbehaving component 2 by means of a local residual generator

Hierarchical framework

For networks without clusters ($\varepsilon > \varepsilon^*$), leaders cooperation is needed



$$C = \begin{bmatrix} C_{\ell_1} \\ C_{\ell_2} \\ \vdots \\ C_{\ell_m} \end{bmatrix} \quad O^s = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{s-1} \end{bmatrix}$$

$$O_i^s = \begin{bmatrix} C_{\ell_i} \\ C_{\ell_i}A \\ \vdots \\ C_{\ell_i}A^{s-1} \end{bmatrix} \quad O^s = \begin{bmatrix} O_1^s \\ O_2^s \\ \vdots \\ O_m^s \end{bmatrix}$$

- extension to process and measurement noise

"Distributed estimation and detection under local information," in *IFAC Workshop on Distributed Estimation and Control in Networked Systems*,

State estimation

Measurements

$$Y_i^s = \begin{bmatrix} y_{\ell_i}(0) \\ y_{\ell_i}(1) \\ \vdots \\ y_{\ell_i}(s-1) \end{bmatrix} \quad Y^s = \begin{bmatrix} Y_1^s \\ Y_2^s \\ \vdots \\ Y_m^s \end{bmatrix}$$

In the absence of misbehaving components

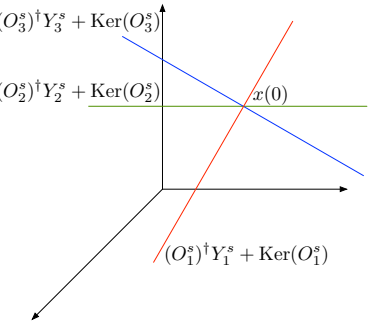
$$Y^s = O^s x(0)$$

- if $\mathrm{Ker}(O^s) = \{0\}$, then the state $x(0)$ can be recovered from the measurements Y^s

A geometric interpretation

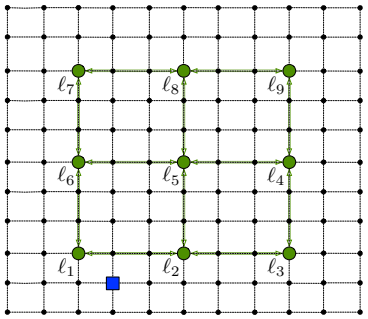
If $\mathrm{Ker}(O^s) = \{0\}$, the unique solution to $Y^s = O^s x(0)$ is the intersection of the affine subspaces defined by the block-rows of $Y^s = O^s x(0)$

$$\begin{bmatrix} Y_1^s \\ Y_2^s \\ \vdots \\ Y_m^s \end{bmatrix} = \begin{bmatrix} O_1^s \\ O_2^s \\ \vdots \\ O_m^s \end{bmatrix} x(0)$$



- similarities with iterative Kaczmarz method

An algorithmic solution



- the leader graph is connected

Leaders iterate the following operations

- transmit local estimate and uncertainty subspace
- receive estimates from neighbors
- update local estimate and uncertainty subspace

(Diffusive) Estimation algorithm

Input : Local observability matrix O_i^s, Local measurements Y_i^s;
Require : Global observability, i.e., $\mathrm{Ker}(O^s) = 0$;

transmit $\hat{x}_i = (O_i^s)^\dagger Y_i^s$, $\quad \mathcal{V}_i = \mathrm{Ker}(O_i^s)$;

while $\mathcal{V}_i \neq \{0\}$ **do**
  for $j \in N_i$ **do**
    receive $\hat{x}_j$ and $\mathcal{V}_j$;
    set $\hat{x}_i = \hat{x}_i \perp \left((\hat{x}_i + \mathcal{V}_i) \cap (\hat{x}_j + \mathcal{V}_j)\right)$;
    set $\mathcal{V}_i = \mathcal{V}_i \cap \mathcal{V}_j$;
  transmit $\hat{x}_i$ and $\mathcal{V}_i$;
return $\hat{x}_i$;

Detecting misbehaving components

If $Bu(t) \neq 0$, the measurements equation becomes

$$Y^s = O^s x(0) + F^s U^s$$

$$F_i^s = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ C_{\ell_i} B & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ C_{\ell_i} A^{s-2} B & \cdots & C_{\ell_i} B & 0 \end{bmatrix} \quad F^s = \begin{bmatrix} F_1^s \\ F_2^s \\ \vdots \\ F_m^s \end{bmatrix} \quad U^s = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(s-1) \end{bmatrix}$$

Assume that the attack is detectable from the measurements vector

- $\mathrm{Im}(O^s) \cap \mathrm{Im}(F^s) = \emptyset$

(Diffusive) Detection algorithm

If $F^s U^s \neq 0$, the system $Y^s = O^s x$ has no solution

- the affine subspaces have empty intersection

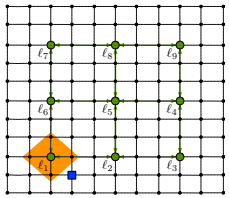
Input : O_i^s, Y_i^s;
Require : Global observability, i.e., $\mathrm{Ker}(O^s) = 0$,
Attack detectability, i.e., $\mathrm{Im}(O^s) \cap \mathrm{Im}(F^s) = \{0\}$;

transmit $\mathcal{S}_i = (O_i^s)^\dagger Y_i^s + \mathrm{Ker}(O_i^s)$;

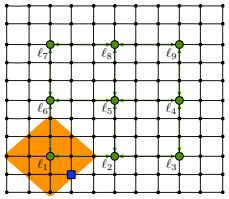
for *diam*$(G^{(\ell)})$ *iterations* **do**
  for $j \in N_i$ **do**
    set $\mathcal{S}_i = \mathcal{S}_i \cap \mathcal{S}_j$;
  transmit $\mathcal{S}_i$;
if $\mathcal{S}_i = \emptyset$ **then return** *Alarm*

Why local model information

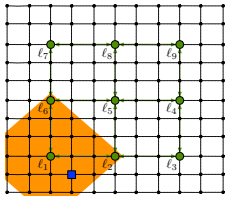
The computation of the matrix O_i^s only requires local model information



C_{ℓ_1}



$C_{\ell_1} A$



$C_{\ell_1} A^2$

- for the matrix $C_{\ell_1} A^s$, only the colored regions are needed
- place leaders such that $\mathrm{Ker}(O^s) = \{0\}$ with $s < \mathrm{diam}(G)$
- scalability properties against network dimension

Summary and extensions

Attack detection via local model information

- if the network is sufficiently weakly interconnected, then each leader performs detection of a certain class of attacks
  - local model information
  - finite time detection of a class of attacks
  - no communication overhead
- if a hierarchical structure of leaders is present, then state estimation and attack detection are possible
  - local model information
  - finite time estimation and detection
  - communication overhead

Identifying Cyber Attacks via Local Model Information

Fabio Pasqualetti
Ruggero Carli Antonio Bicchi Francesco Bullo

Center for Control, Dynamical systems and Computation
University of California, Santa Barbara

Conference on Decision and Control, Atlanta, GA, USA
December 15 - 17, 2010