Consensus Networks with Misbehaving Agents

Fabio Pasqualetti, Antonio Bicchi, Francesco Bullo

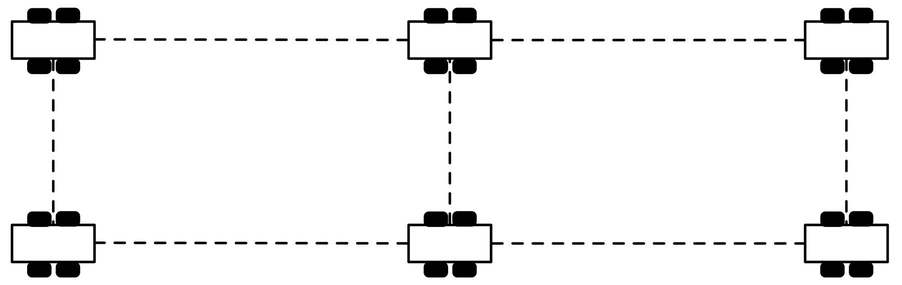
Center for Control, Dynamical systems and Computation
Department of Mechanical Engineering
University of California at Santa Barbara

Multi-agent system

Network of autonomous agents able to sense, communicate and process information

- rendezvous
- formation
- clock synchronization
- load balancing

Linear consensus network



Structure:

- each agent is represented by a vertex of a graph
- exchange data with neighboring agents

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

Convergence:

- A is row-stochastic and primitive

Misbehaving agent

A misbehaving agent updates its state differently than specified by the nominal protocol A

- modeled by an exogenous input

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

- each column of B has one nonzero entry
- the input function u is arbitrary

Misbehaving agent: $\exists t$ such that $u_i(t) \neq 0$

Malicious agent: the input u_i is arbitrary

Faulty agent: $\nexists F$ such that $u_i(t) = Fx(t), \forall t$

Local observation

Each agent j observes directly the state of its neighbors

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

- each row of C_j has one nonzero entry

$$B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Problem definition

Each agent knows A, and relies only on its output:

(Detection) Detect the presence of misbehaving agents in the network

(Identification) Identify the misbehaving agents in the network

Detection filter

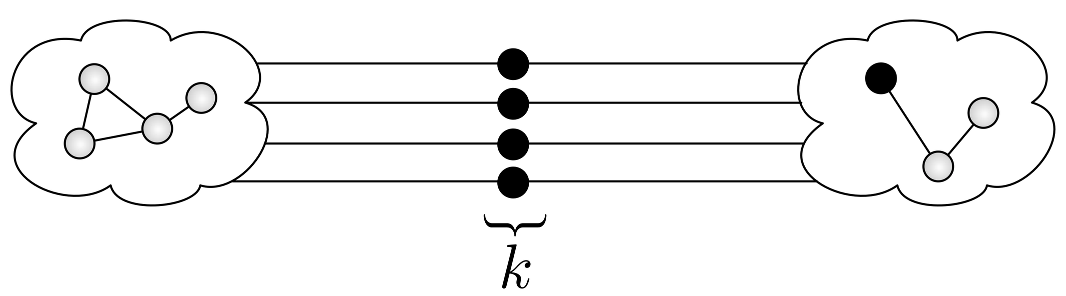
Let the zero dynamics be stable, then the filter

$$z(t+1) = (A + GC_j)z(t) - Gy_j(t)$$
$$\tilde{x}(t) = Lz(t) + Hy_j(t)$$

$$G = -A_{N_j} \quad H = C_j^T \quad L = I - HC_j$$

allows the detection of the misbehaving agents

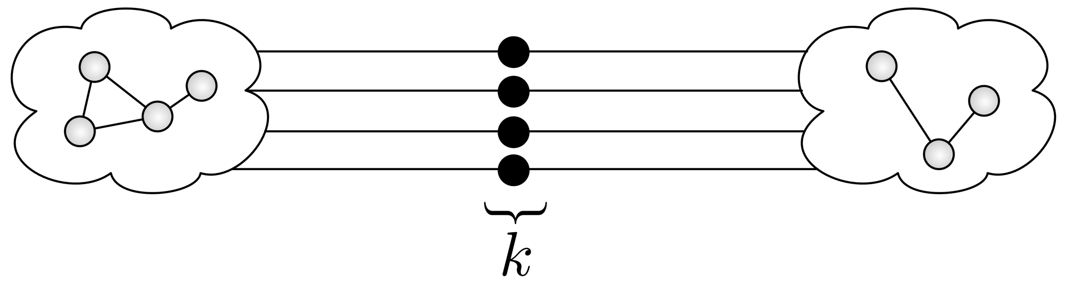
Zero dynamics and connectivity



Let K denote the set of misbehaving agents in a k-connected network

- $\exists$ K, j, with $|K| > k$, such that (A, B, C_j) is not left-invertible
- $\exists$ K, j, $|K| = k$, such that (A, B, C_j) has nontrivial zero dynamics

Detection of misbehaving agents



At most $k-1$ misbehaving agents can be detected in a k-connected network

Identification of misbehaving agents

The set K_1 and K_2 are distinguishable if and only if $(A, [B_1 \; B_2], C_j)$ has no zero dynamics

At most $\lfloor \frac{k-1}{2} \rfloor$ misbehaving agents can identified in a k-connected network

Identification of faulty agents

A zero input satisfies

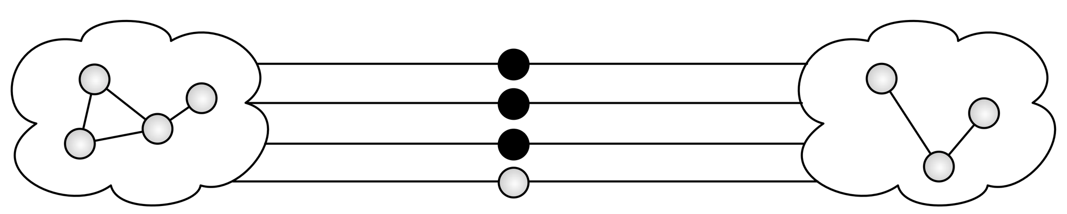
$$(C_j A^\nu B)u(t) = CA^{\nu+1}x(t)$$

- faulty agents do not inject zero inputs
- the faulty set K_1 is indistinguishable from the faulty set K_2 if $\mathcal{Y}_{K_1} \subseteq \mathcal{Y}_{K_2}$

At most $k-1$ faulty agents can be identified in a k-connected network

Generic detection and identification

A linear system is generic if its entries are either zeros or free independent parameters



A linear system has generically no zero dynamics if the number of inputs is less than the connectivity of its associated graph

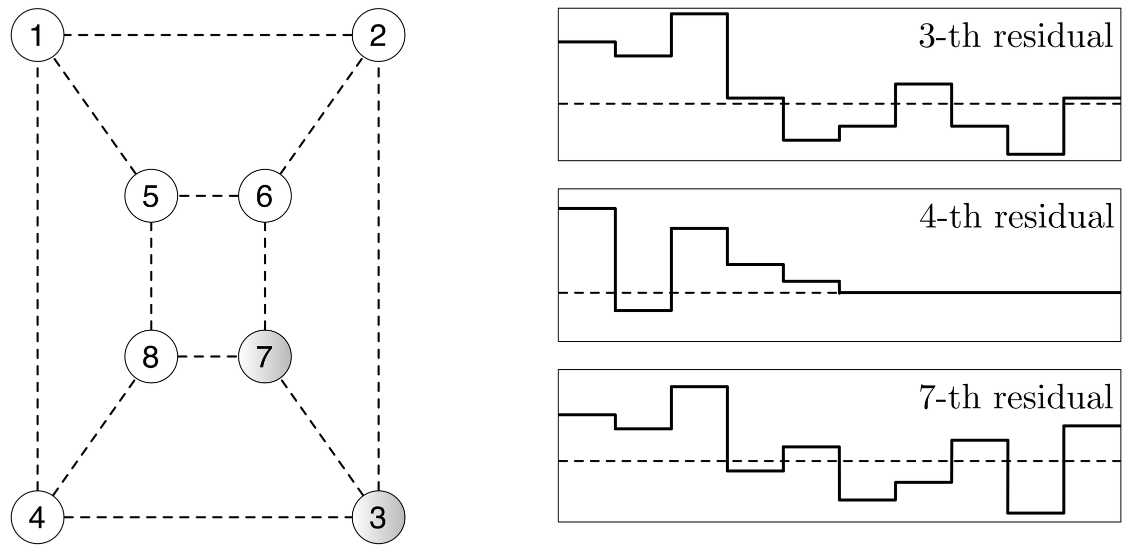
Identification algorithm

Input : consensus matrix, number of misbehaving agents k

Require: $k+1$ (resp. $2k+1$) connectivity

while *the misbehaving agents are unidentified* **do**
  exchange data with neighbors
  update state
  evaluate residual functions
  if *every i_{th} residual is nonzero* **then**
    agent i is recognized as misbehaving

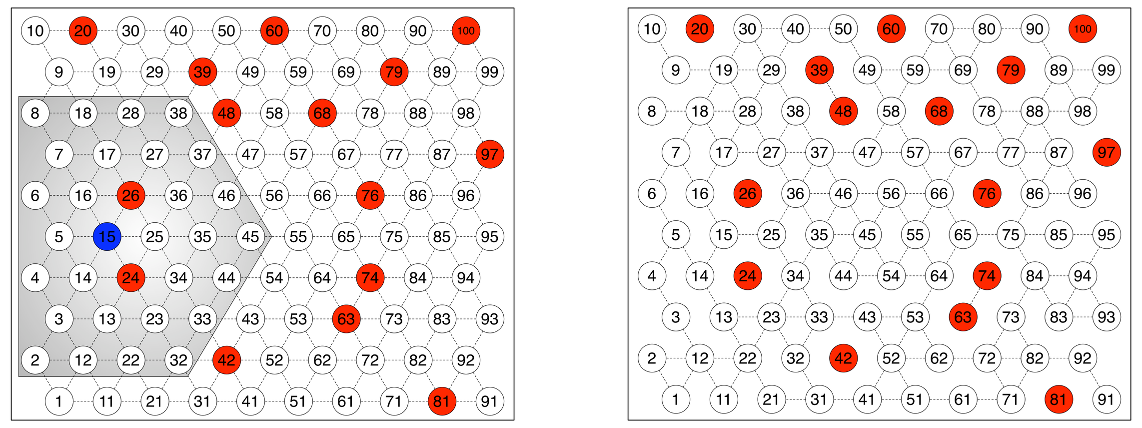
An example



The network is **3**-connected

- **2** faulty agents are *generically* identifiable
- **1** malicious agent is *generically* identifiable

Ongoing research



Each agent only knows the structure of its d-neighborhood

$$x_j^d(t+1) = A_j^d x_j^d(t) + B_j^d u_K(t) + B_D u_D(t)$$
$$y_j(t) = C_j^d x_j^d(t)$$

- the term $B_D u_D(t)$ appears in the residuals
- robust residual generation

Center for Control, Dynamical Systems and Computation - CCDC - Santa Barbara, California
fabiopas@engineering.ucsb.edu
http://www.fabiopas.it