Resilient Algorithms for Distributed Coordination and Decision-Making in Large-Scale Networks

Shreyas Sundaram

School of Electrical and Computer Engineering

Purdue University

https://engineering.purdue.edu/~sundara2/

PURDUE ENGINEERING

Funding sources:

NSF

Sandia National Laboratories

Large-Scale Systems Monitored by a Network of Agents

Setting:

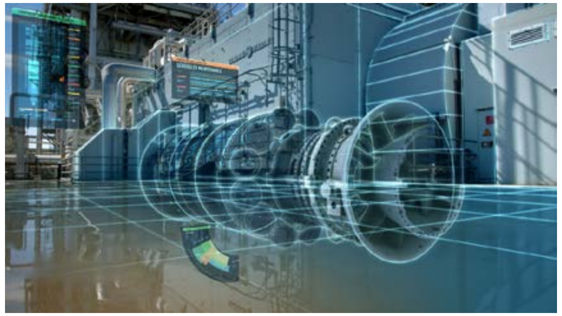
- Large-scale system monitored by a network of agents
- Each agent periodically receives signals about the state of the system
    - Each agent's signals are only partially informative
- Network can be mobile, time-varying, **contain adversaries**

Objective:
Formulate algorithms that allow all regular nodes in the network to cooperatively estimate the state of the entire system

Monitoring / surveillance with autonomous teams

EPFL

Smart factories

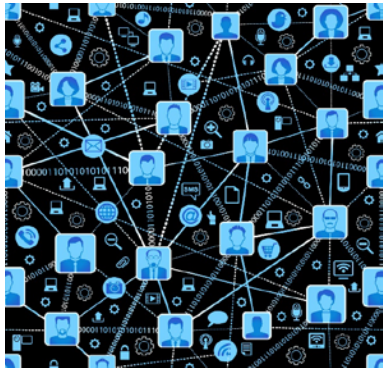


GE

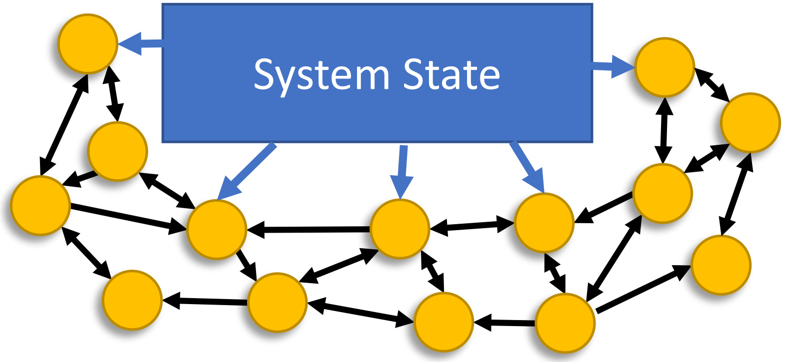
Smart cities

TechRepublic

Social Networks



MIT IDSS



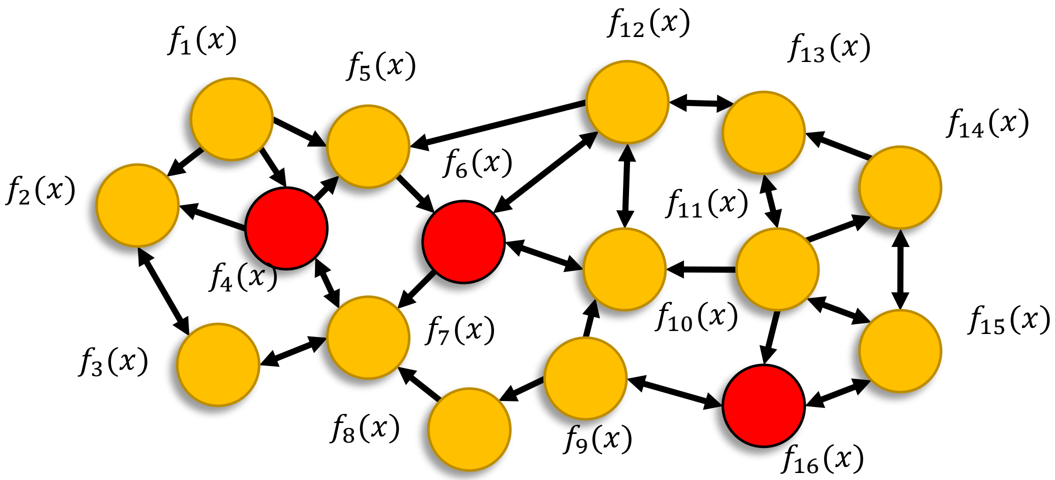
Specific instances

- **Distributed consensus:** each node v_i has a local (static) measurement, and all nodes must converge to the same function of their local measurements
- **Distributed optimization:** each node has a local function $f_i(x)$ and the nodes must cooperatively calculate the minimizer of the sum of their local functions
- **Distributed state estimation:** the nodes are each measuring different parts of a dynamical system, and must cooperate to estimate the global system state
- **Distributed hypothesis testing:** the nodes must cooperate to identify the true state of the world from a set of possible hypothesis, based on stochastic measurements

There exist various distributed algorithms to solve versions of these problems

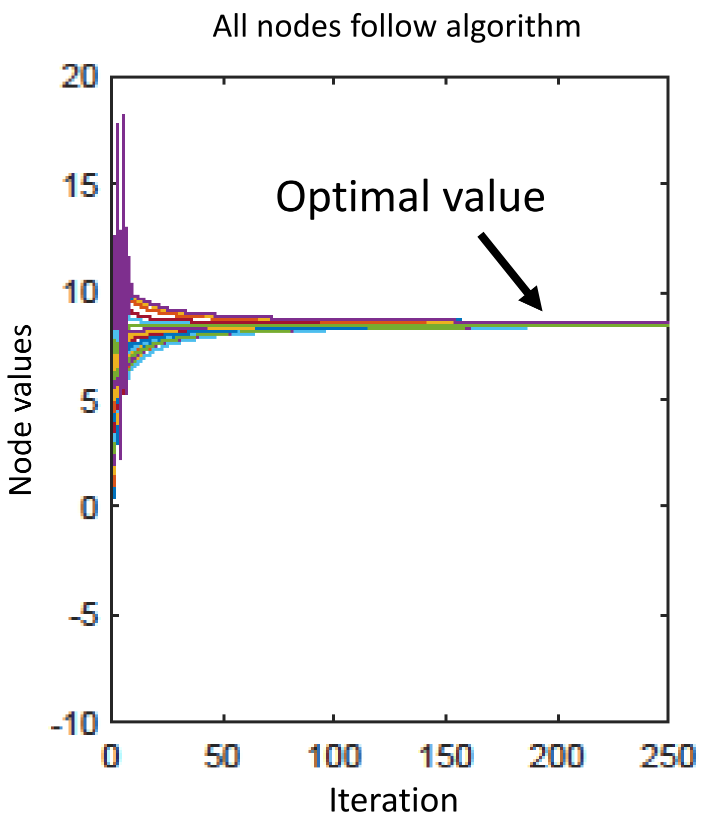
The Need for Resilience

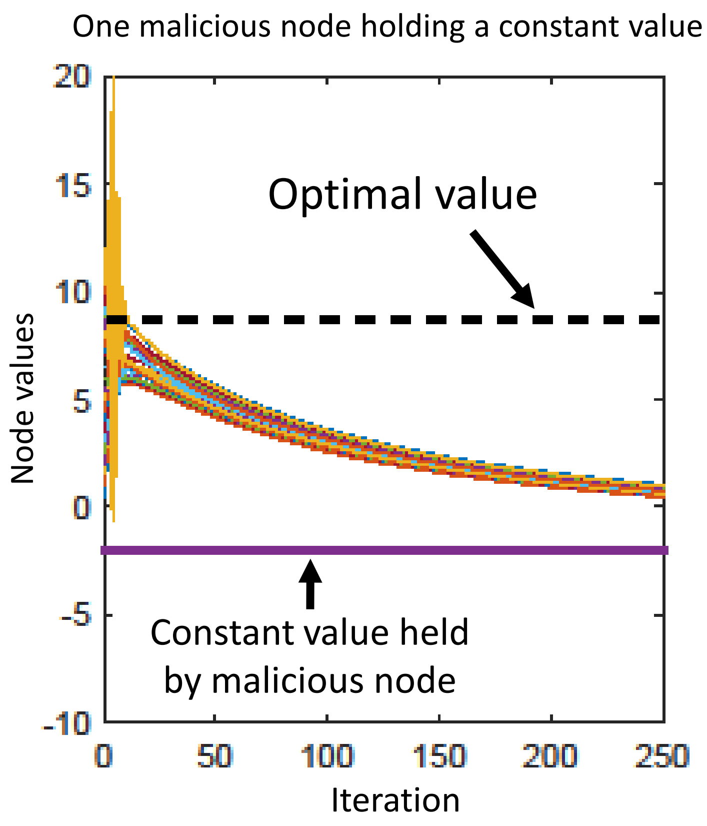
- What happens if certain nodes **fail** or are **compromised by an attacker**?



- Attacks can be coordinated, based on “insider knowledge”, targeted against vulnerable nodes, etc.

Illustration of vulnerabilities in distributed consensus/optimization algorithms





Considerations for Resilient Algorithms

- What do the “normal” nodes know?
  - Entire network topology versus only their local neighborhoods
  - Nominal behavior of all nodes versus only local dynamics

- How much computation/storage do the normal nodes have?
  - Extensive computations with lots of stored data versus simple computations on limited data

- What are the objectives for the normal nodes?
  - Calculate the desired value exactly versus calculate an approximate value

Considerations for Resilient Algorithms (cont'd)

- What kinds of misbehavior need to be overcome?
  - Node drops out of the network ("crashes")
  - Node updates its state according to a known model

  } **"Faulty"**

  - Node updates its state in an arbitrary (unknown) manner (**"Malicious"**)
  - Node can send conflicting values to different neighbors (**"Byzantine"**)

- How many misbehaving nodes can there be?
  - **F-total**: Up to F misbehaving nodes in the entire network
  - **F-local**: Up to F misbehaving nodes in the neighborhood of each normal node

Answers to the above questions will dictate the conditions on the network topology required to design resilient algorithms

The Role of Network Connectivity

- **Classical result:** If there are up to F malicious nodes, all nodes can reliably exchange information if and only if network is **(2F+1)-connected**
  - [Dolev et al., ‘93], [Lynch, ‘96], [Sundaram & Hadjicostis, ‘11], [Pasqualetti et. al, ’12], …

- Typical assumptions:
  - All nodes know the entire network topology and nominal dynamics of the other nodes
  - Each node can store data and perform extensive computations

- **Need *scalable* algorithms** and mechanisms to overcome adversarial behavior in large-scale networks
  - Shouldn’t require regular nodes to know network topology
  - Tradeoff between knowledge and achievable objectives

Local-Filtering Dynamics for Resilient Consensus

Local Filtering Dynamics for Consensus

- Suppose each node v_i starts with an initial value $x_i(0)$
- Mechanism:
  - At each time-step t, each node v_i receives values from its neighbors
  - v_i **removes the F highest and lowest values** in its neighborhood, updates its state as a weighted average of remaining values

$$x_i(t+1) = w_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{J}_i(t)} w_{ij}(t)x_j(t)$$

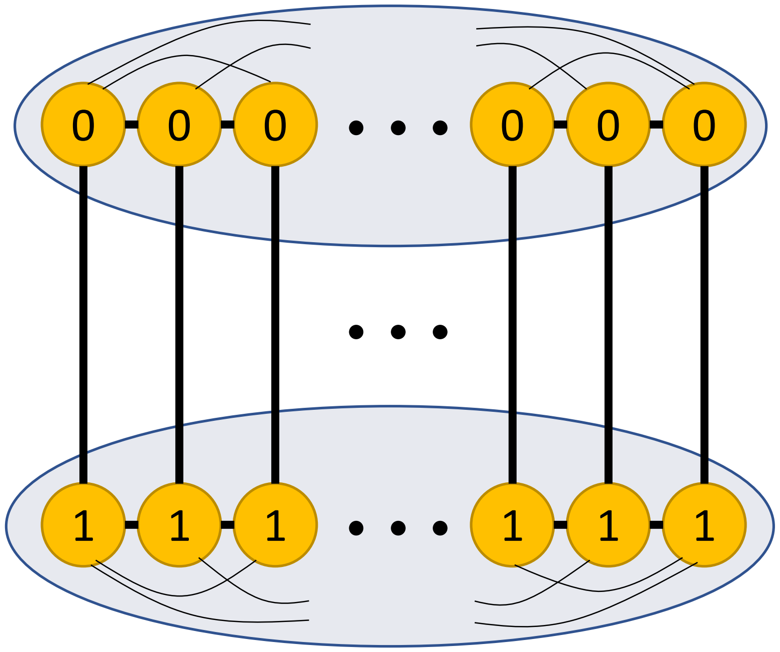
Neighbors after removing extreme values

- Weights $w_{ii}(t)$ and $w_{ij}(t)$ specify a convex combination at each time-step

Dolev, et al., '86; Azadmanesh et al., '90s; Vaidya et al., '12; LeBlanc, Zhang, Koutsoukos and Sundaram '12, '13

Failure of Convergence

- Network where convergence does not occur:



Fully-connected graph with $n/2$ nodes
Initial value 0

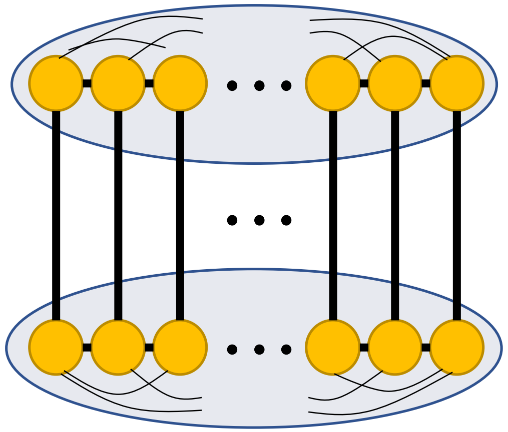
One-to-one edges between sets

Fully-connected graph with $n/2$ nodes
Initial value 1

- Connectivity of graph is $n/2$, but **no node ever uses a value from opposite set**

Insufficiency of Connectivity as a Metric

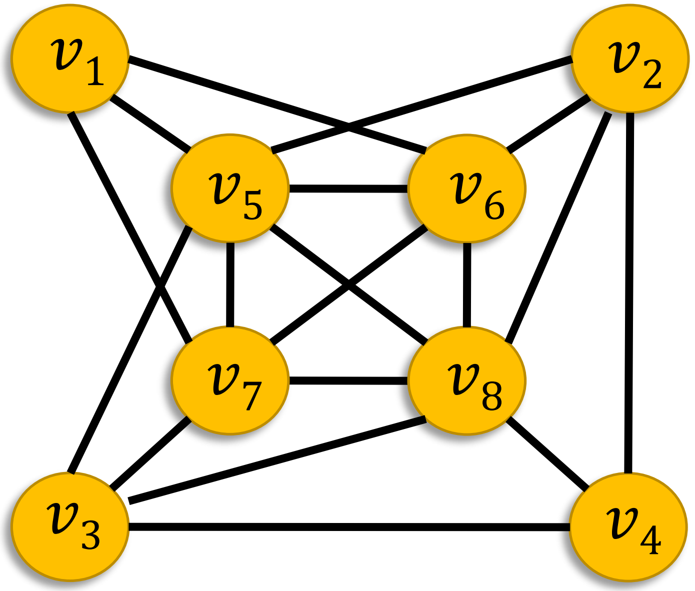
- Graph contains sets where no node in any set has **enough neighbors** outside the set
  - i.e., all outside information is filtered away by each node



- **Need a new topological property** to characterize conditions under which local filtering will succeed

Robust Graphs

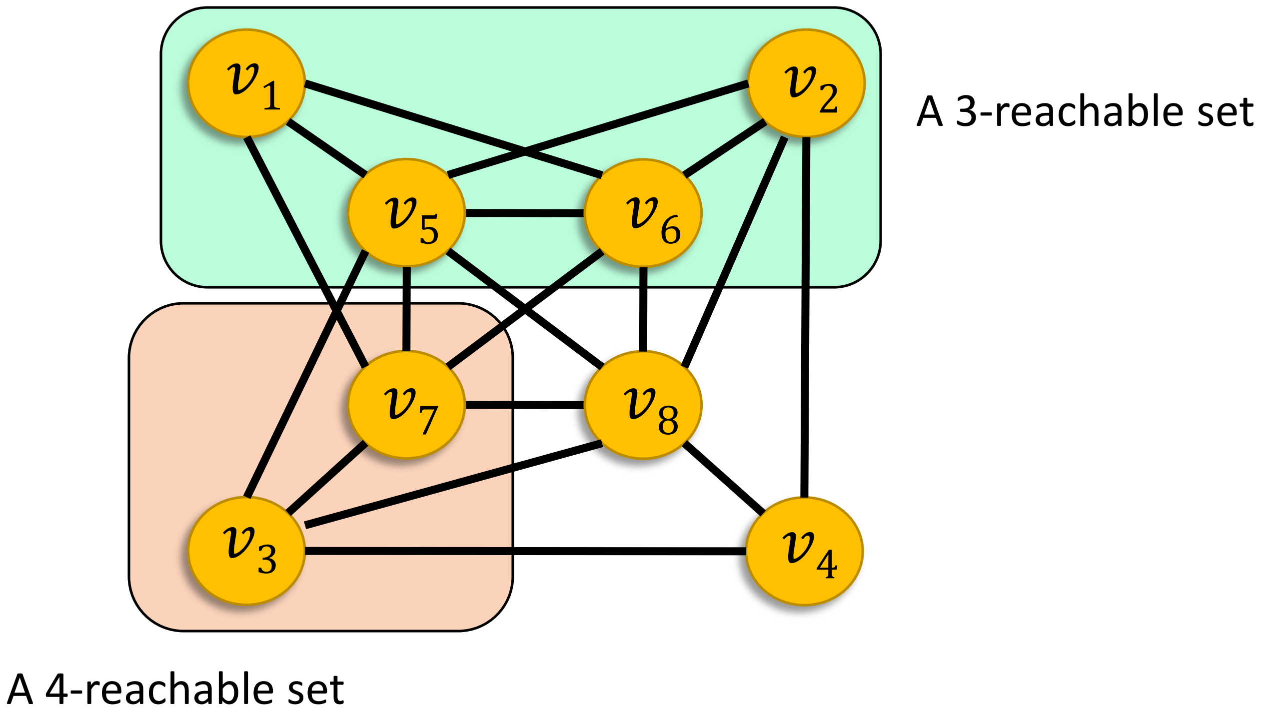
- We introduce the following definitions
  - A set S is ***r*-reachable** if it has a node that has at least r neighbors outside the set



Zhang & Sundaram, ACC 2012; LeBlanc, Zhang, Koutsoukos and Sundaram, IEEE JSAC 2013

Robust Graphs

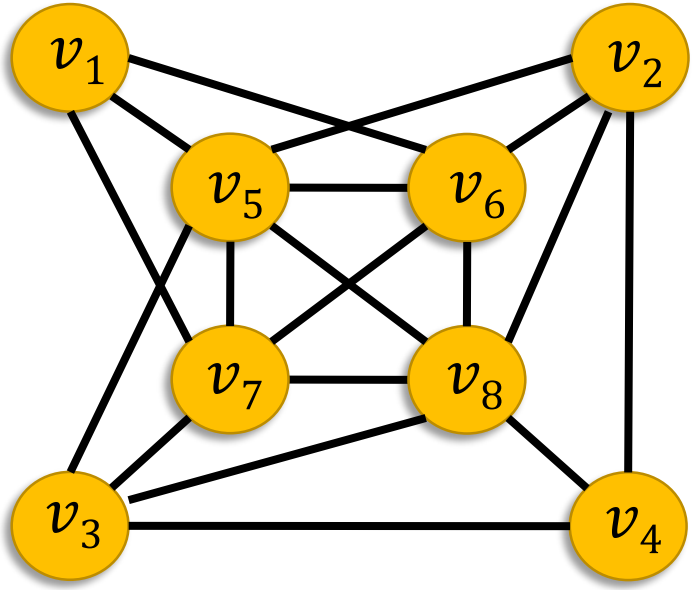
- We introduce the following definitions
  - A set S is ***r*-reachable** if it has a node that has at least r neighbors outside the set



Zhang & Sundaram, ACC 2012; LeBlanc, Zhang, Koutsoukos and Sundaram, IEEE JSAC 2013

Robust Graphs

- A graph is ***r*-robust** if for every pair of disjoint subsets, at least one of the sets is r-reachable

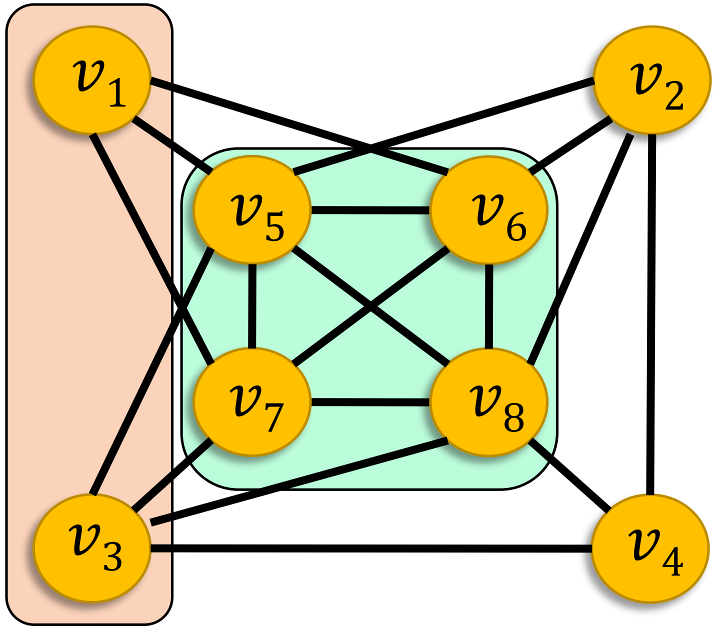


3-robust graph:
For every pair of disjoint subsets of nodes, at least one subset is 3-reachable

Zhang & Sundaram, ACC 2012; LeBlanc, Zhang, Koutsoukos and Sundaram, IEEE JSAC 2013

Robust Graphs

- A graph is ***r*-robust** if for every pair of disjoint subsets, at least one of the sets is r-reachable



3-robust graph:
For every pair of disjoint subsets of nodes, at least one subset is 3-reachable

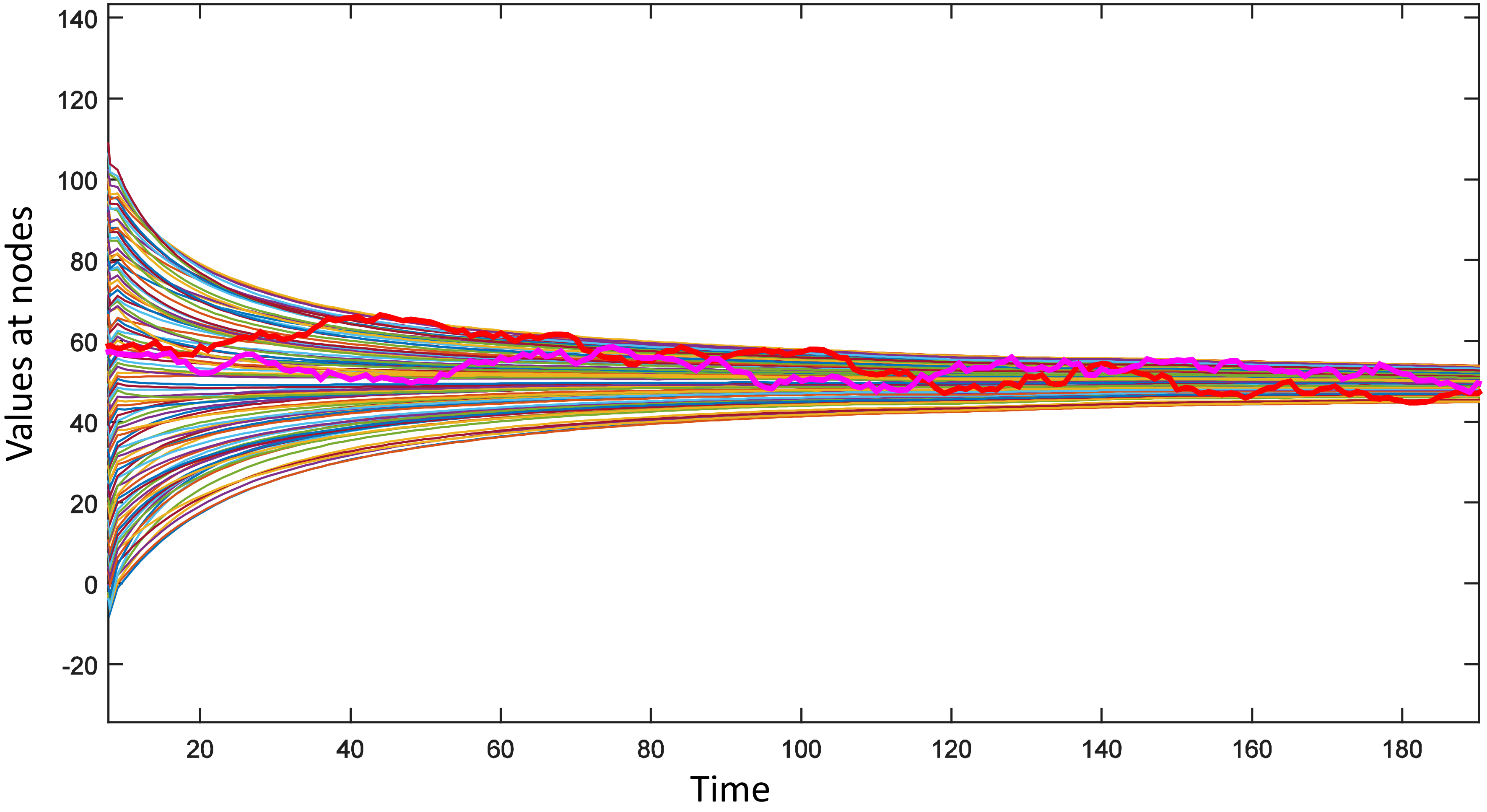
Zhang & Sundaram, ACC 2012; LeBlanc, Zhang, Koutsoukos and Sundaram, IEEE JSAC 2013

Condition for Resilient Consensus under Local-Filtering

Theorem:
If network is $(2F + 1)$-robust, normal nodes will reach consensus in the convex hull of their initial values despite actions of any F-local set of Byzantine nodes

- **F-local set:** up to F adversaries in neighborhood of *every* node
- **Takeaway point:** If the graph satisfies the required "robustness" property, local-filtering algorithm provides strong resilience guarantees against a potentially large number of worst-case adversaries

LeBlanc, Zhang, Koutsoukos and Sundaram, IEEE JSAC 2013; Zhang, Fata and Sundaram, IEEE TCNS 2015



140
120
100
80
60
40
20
0
-20
Values at nodes
20
40
60
80
100
120
140
160
180
Time

Robustness of Complex Networks

- r-robustness and r-connectivity coincide in various models for complex networks:
  - Erdos-Renyi random graphs (Zhang, Fata & Sundaram, TCNS 2015)
  - 1-D geometric random graphs (Zhang, Fata & Sundaram, TCNS 2015)
  - Preferential attachment graphs (Zhang, Fata & Sundaram, TCNS 2015)
  - Random intersection graphs (Zhao, Yagan & Gligor, CDC 2014)
  - Random k-partite graphs (Shahrivar, Pirani & Sundaram, Automatica 2017)
  - Circulant graphs (Usevitch & Panagou, CDC 2017)

Takeaway points:

- Although r-robustness is stronger than r-connectivity, the properties occur simultaneously in many large-scale networks
- Such networks will be conducive to applying local-filtering dynamics for resilient coordination

"Local-Filtering" is a promising scalable mechanism for resilient distributed coordination in large-scale networks

Applications of Local-Filtering in Distributed Optimization and State Estimation

Distributed Optimization

- Each node v_i in the network has a local convex function $f_i: \mathbb{R} \rightarrow \mathbb{R}$
- Nodes wish to calculate (in a distributed manner) $\arg\min_{x \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} f_i(x)$
- Common approach: **consensus-based distributed optimization**
  - Each node updates its estimate of the optimal parameter as

$$x_i(t+1) = \underbrace{w_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{N}_i} w_{ij}(t)x_j(t)}_{\text{Consensus Step}} \underbrace{- \alpha_t d_i(t)}_{\text{Gradient Step}}$$

  - $d_i(t)$ is a subgradient of $f_i(x)$ evaluated at $w_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{N}_i} w_{ij}(t)x_j(t)$
  - $\alpha_t \in \mathbb{R}_{\geq 0}$ is a stepsize

Resilient Distributed Optimization via Local-Filtering Dynamics

- To obtain resilience, apply local-filtering

$$x_i(t+1) = w_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{J}_i(t)} w_{ij}(t)x_j(t) - \alpha_t d_i(t)$$

Neighbors after removing extreme values

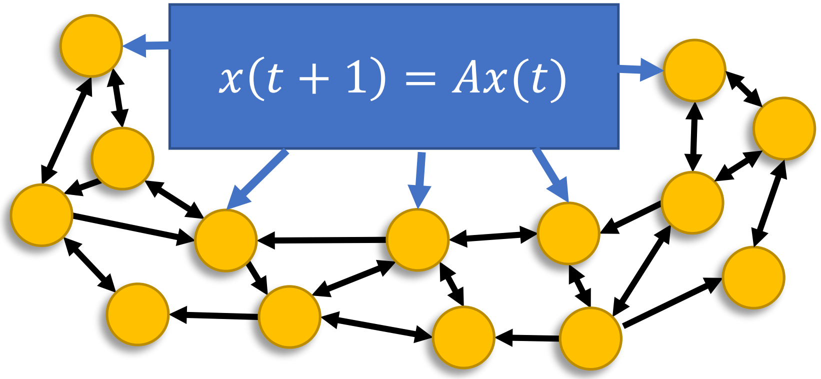
Theorem:
Suppose network is $(2F+1)$-robust and that $\alpha_t \to 0$ and $\sum \alpha_t = \infty$ in the Local-Filtering distributed optimization dynamics.

Then, all regular nodes asymptotically reach consensus and converge to the convex hull of the local minimizers of the regular nodes, regardless of actions of any F-local set of Byzantine adversaries.

Sundaram & Gharesifard, Allerton 2015, TAC 2019; Su & Vaidya, 2015

Distributed State Estimation

- Consider a dynamical system, monitored by a network of nodes:



Aritra Mitra

- Each node v_i obtains the state measurement

$$y_i(t) = C_i x(t)$$

- Nodes seek to cooperatively estimate the full state $x(t)$

Contribution: A fully distributed state estimator that allows all normal nodes to asymptotically recover the state despite F-local Byzantine adversaries.

Problem and Challenges

- For simplicity, consider a scalar dynamical system of the form:

$$x(t+1) = ax(t), \qquad a \in \mathbb{R}, |a| \geq 1$$

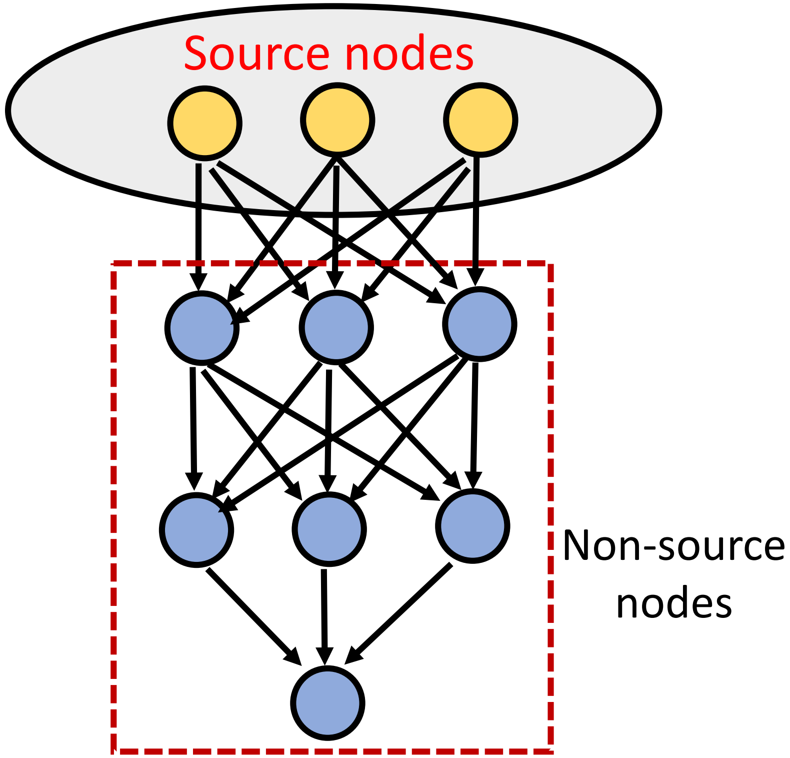
- For this system, nodes with non-zero measurements can estimate the state themselves **without** communicating with neighbors
- We call such nodes the "**source nodes**", denoted by set S
- A non-source node must communicate with (potentially adversarial) neighbors

- **Key Question:** How does a non-source node process the information received from neighbors to asymptotically estimate $x(t)$?
  - Require **redundancy** in both **measurements** (source nodes) and **network structure** (for information diffusion)

Mode Estimation Directed Acyclic Graph (MEDAG)

- For a given $F \in \mathbb{N}$, define a **Mode Estimation Directed Acyclic Graph** (MEDAG) to be a DAG where:
  - The root nodes are the source nodes S
  - Each non-root node has at least $(2F + 1)$ parents
- Such graphs capture the required redundancy in both measurements and topology
- When does a given graph contain a MEDAG with respect to a given source set?
  - We show a graph-theoretic notion similar to "(2F+1)-robustness" is required for MEDAG to exist: "**strong (2F+1)-robustness with respect to S**"
  - If graph contains MEDAG, it can be found in polynomial time via a distributed algorithm

MEDAG for $F = 1$



Mitra & Sundaram, CDC 2016, Automatica 2019, Autonomous Robots 2019

Local Filtering Dynamics for Resilient Distributed State Estimation

- Suppose the network contains a MEDAG
- Each non-source node v_i applies a two-stage filtering strategy:
  - At each time-step, it only listens to its parents in the MEDAG, denoted P_i.
  - It sorts the estimates received from P_i from highest to lowest. removes the F highest and F lowest values, and takes a convex combination of the rest to update its state estimate:

$$\hat{x}_i(t+1) = a \sum_{v_j \in \mathcal{J}_i(t)} w_{ij}(t)\hat{x}_j(t)$$

Estimate of state $x(t+1)$ at node v_i

Set of parents whose estimates are used at time t

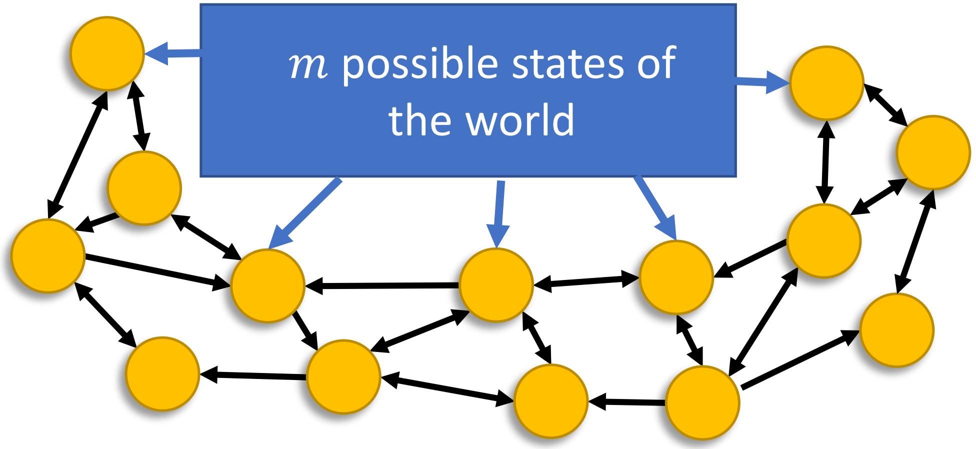
Main Result for Resilient Distributed State Estimation

Theorem:
Suppose the network is strongly (2F+1)-robust with respect to $\mathcal{S}$. Then by applying local-filtering, all regular nodes can asymptotically estimate the state despite the actions of any F-local set of Byzantine nodes.

- Key benefit of approach: Each step of our algorithm can be implemented in a fully distributed and secure manner
- Can be extended directly to more general (non-scalar) sytems

Mitra & Sundaram, CDC 2016, Automatica 2019, Autonomous Robots 2019

Resilient Distributed Hypothesis Testing



Aritra Mitra

- **Problem:** nodes have to cooperatively identify the true state of the world (out of m possible hypotheses) based on stochastic signals
- **Contribution:** a new distributed hypothesis testing algorithm that is provably resilient to F-local Byzantine adversaries
- See poster by Aritra Mitra (at this workshop), and talk tomorrow at 10:00am!

Mitra, Richards & Sundaram, ACC 2019

Summary

- Resilient algorithms require appropriate notions of network "**redundancy**" in order to overcome adversaries
  - Specific notion of redundancy depends on the nature of the algorithm, assumptions about adversaries, etc.
- Traditional graph property for resilience to F-total adversaries: **2F+1 connectivity**
  - Corresponding algorithms require strong assumptions about network topology and capabilities of normal nodes
- Formulated a class of scalable algorithms for resilience against F-local adversaries
  - Based on "**Local Filtering**" dynamics, where normal nodes ignore extreme values from neighbors
  - Requires a new graph property: **(2F+1)-robustness**
  - Local filtering can be used as a building block for resilience in a variety of applications

References

- S. Sundaram and C. N. Hadjicostis, "*Distributed Function Calculation via Linear Iterative Strategies in the Presence of Malicious Agents*." IEEE Transactions on Automatic Control, vol. 56, no. 7, pp. 1495 - 1508, July 2011
- H. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "*Resilient Asymptotic Consensus in Robust Networks*." IEEE Journal on Selected Areas in Communications: Special Issue on In-Network Computation, vol. 31, no. 4, pp. 766 - 781, Apr 2013.
- H. Zhang, E. Fata, and S. Sundaram, "*A Notion of Robustness in Complex Networks*." IEEE Transactions on Control of Network Systems, vol. 2, no. 3, pp. 310 - 320, Sept 2015.
- S. Sundaram and B. Gharesifard, "*Distributed Optimization Under Adversarial Nodes*." IEEE Transactions on Automatic Control, 2019.
- A. Mitra, J. A. Richards, S. Bagchi, and S. Sundaram, "*Resilient distributed state estimation with mobile agents: overcoming Byzantine adversaries, communication losses, and intermittent measurements*", Autonomous Robots, March 2019.
- A. Mitra and S. Sundaram, "*Byzantine-Resilient Distributed Observers for LTI Systems*" Automatica, 2019.
- A. Mitra, J. A. Richards, and S. Sundaram, "*A New Approach for Distributed Hypothesis Testing with Extensions to Byzantine-Resilience*", American Control Conference, July 2019.

Thank you!