Applications of Consensus Algorithms to Wireless Sensor Networks

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Luca Schenato (Univ. Padova)

Joint work w/

Alessandro Chiuso
University of Padova

Sandro Zampieri
University of Padova

Ruggero Carli
UC Santa Barbara

Ph.D. Students:
Damiano Varagnolo
Simone Del Favero
Saverio Bolognani
University of Padova

Outline

- Motivations and target applications
- Overview of consensus algorithms
- Application of consensus to WSN:
  - Sensor calibration
  - Least-square parameter identification
  - Time-synchronization
- Open problems
  - Identification
  - Estimation
  - Control

Outline

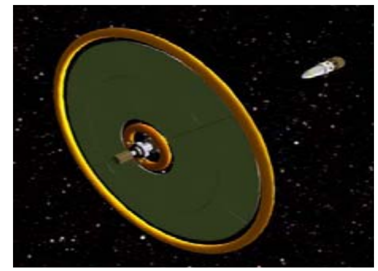
- Motivations and target applications
- Overview of consensus algorithms
- Application of consensus to WSN:
  - Sensor calibration
  - Least-square parameter identification
  - Time-synchronization
- Open problems
  - Identification
  - Estimation
  - Control

Networked Control Systems

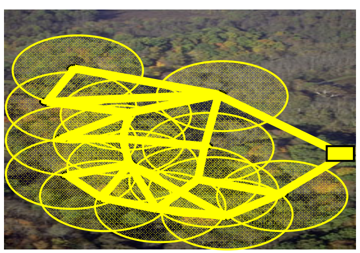
Drive-by-wire systems

Swarm robotics

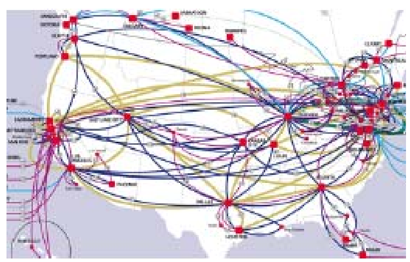
Smart structures:
space telescope & satellites mesh



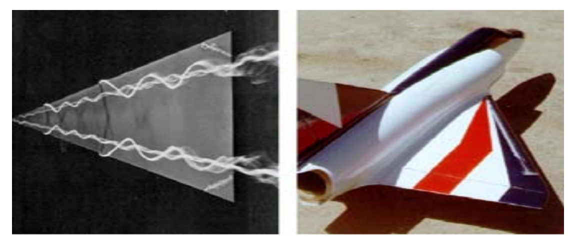
Wireless Sensor Networks



Traffic Control:
Internet and transportation



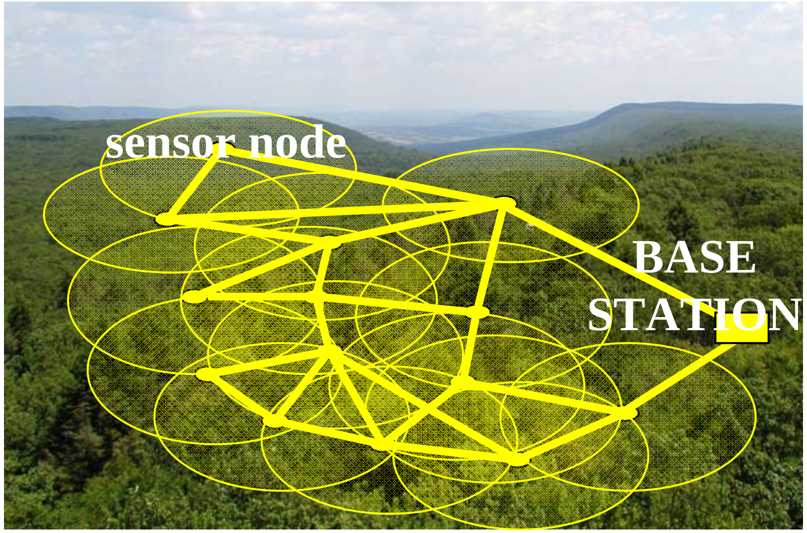
Smart materials & MEMS:
sheets of sensors and actuators



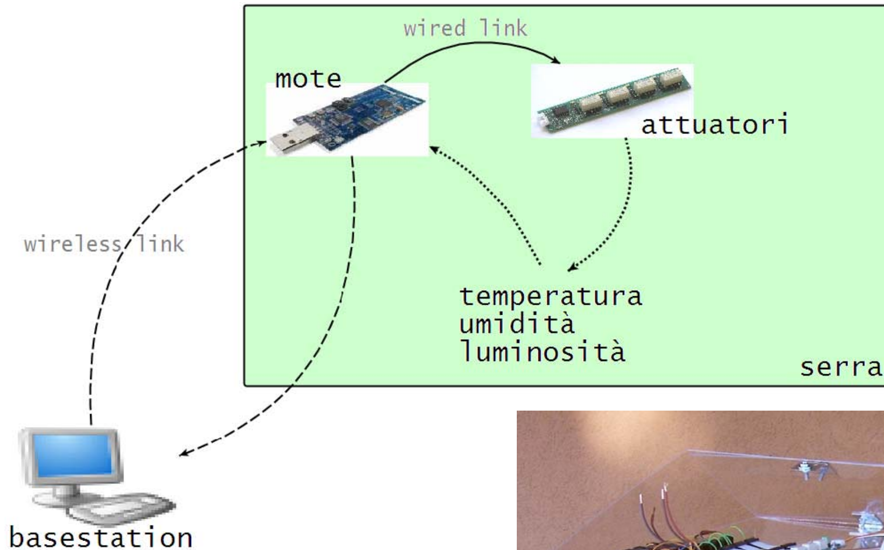
NCSs: physically distributed dynamical systems interconnected by a communication network

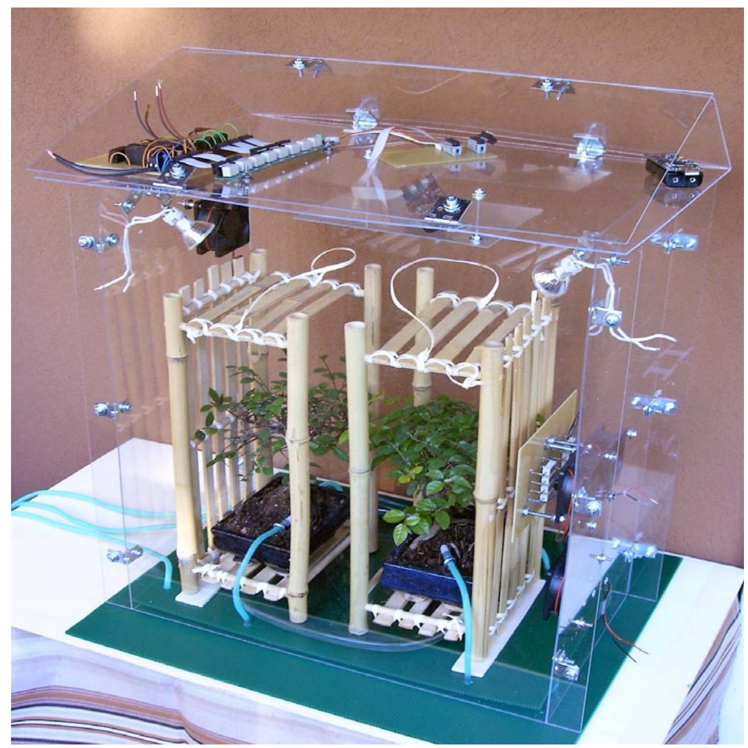
Wireless Sensor Actuator Networks (WSANs)

- Small devices
  - µController, Memory
  - Wireless radio
  - Sensors & Actuators
  - Batteries
- Inexpensive
- Multi-hop communication
- Programmable (micro-PC)



Applications: Smart Greenhouse

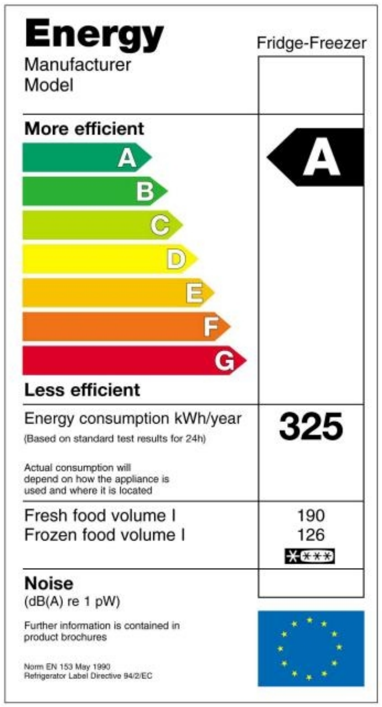




- Distributed estimation
- Distributed control
- Control under packet loss & random delay
- Sensor fusion
- Distributed time synchronization

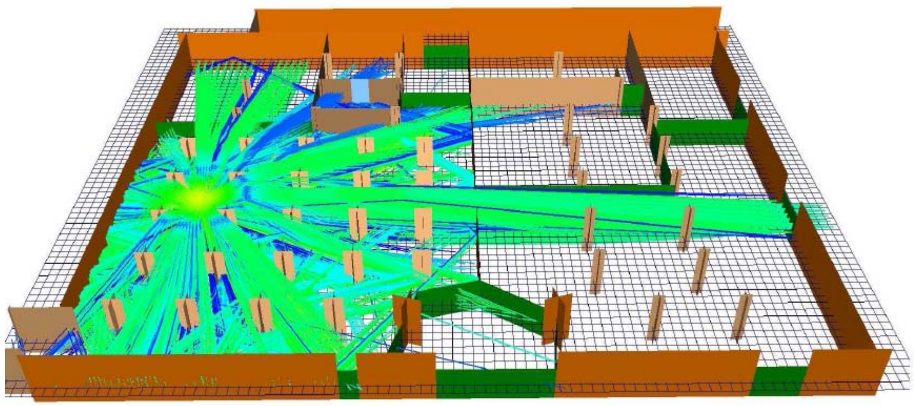
DEPARTMENT OF
INFORMATION
ENGINEERING
UNIVERSITY OF PADOVA

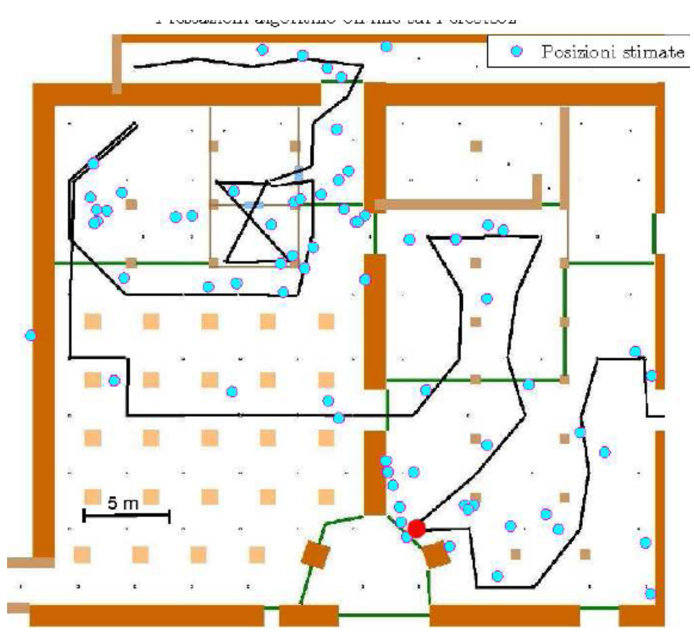
Applications: ThermoEfficiency Labeling



- Building thermodynamics model identification
- Sensor selection for identification
- Optimal sensor placement

Applications: Distributed Localization&Tracking

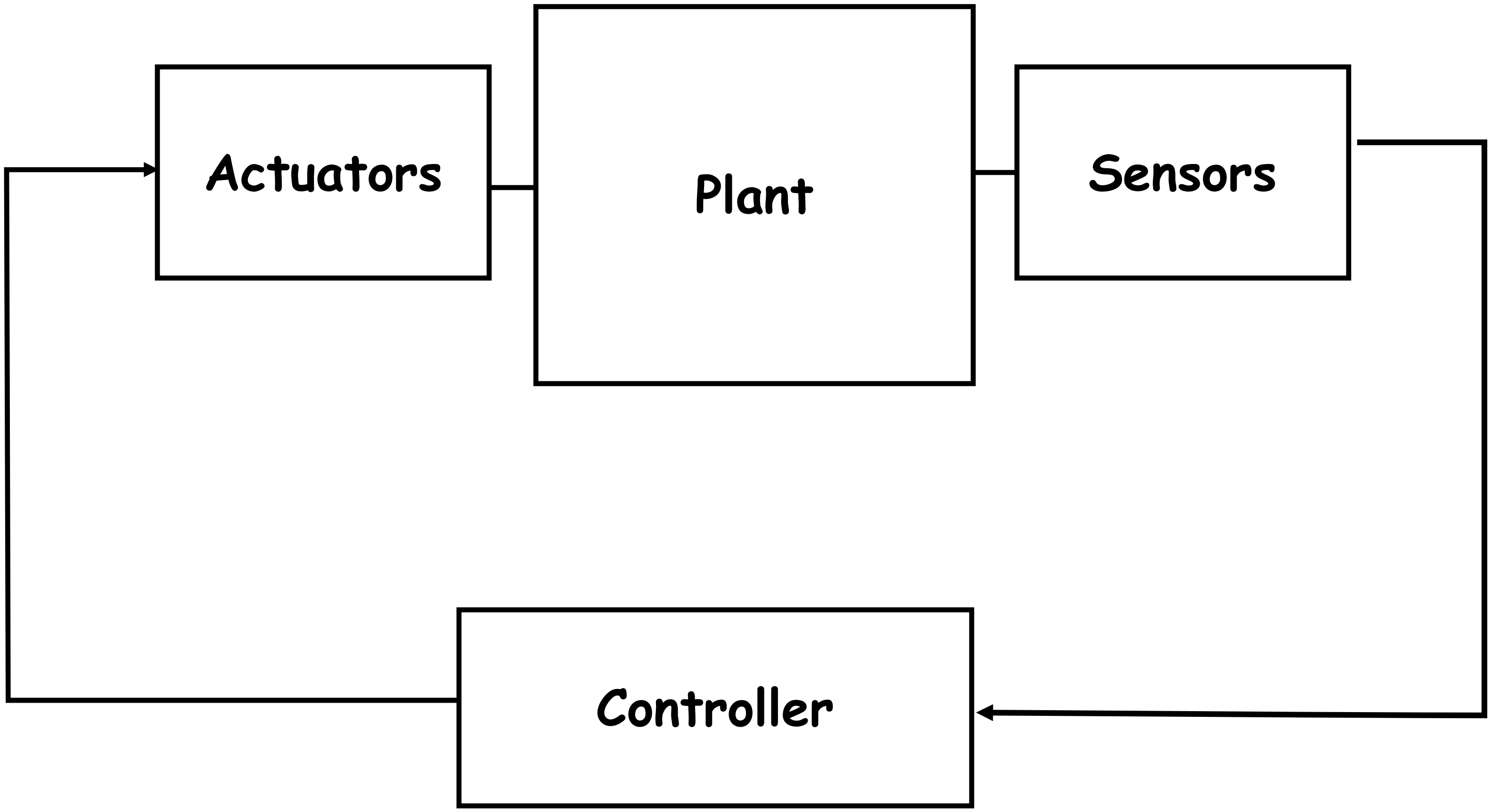




- Indoor radio signal modeling
- Real-time localization
- Distributed tracking
- Coordination

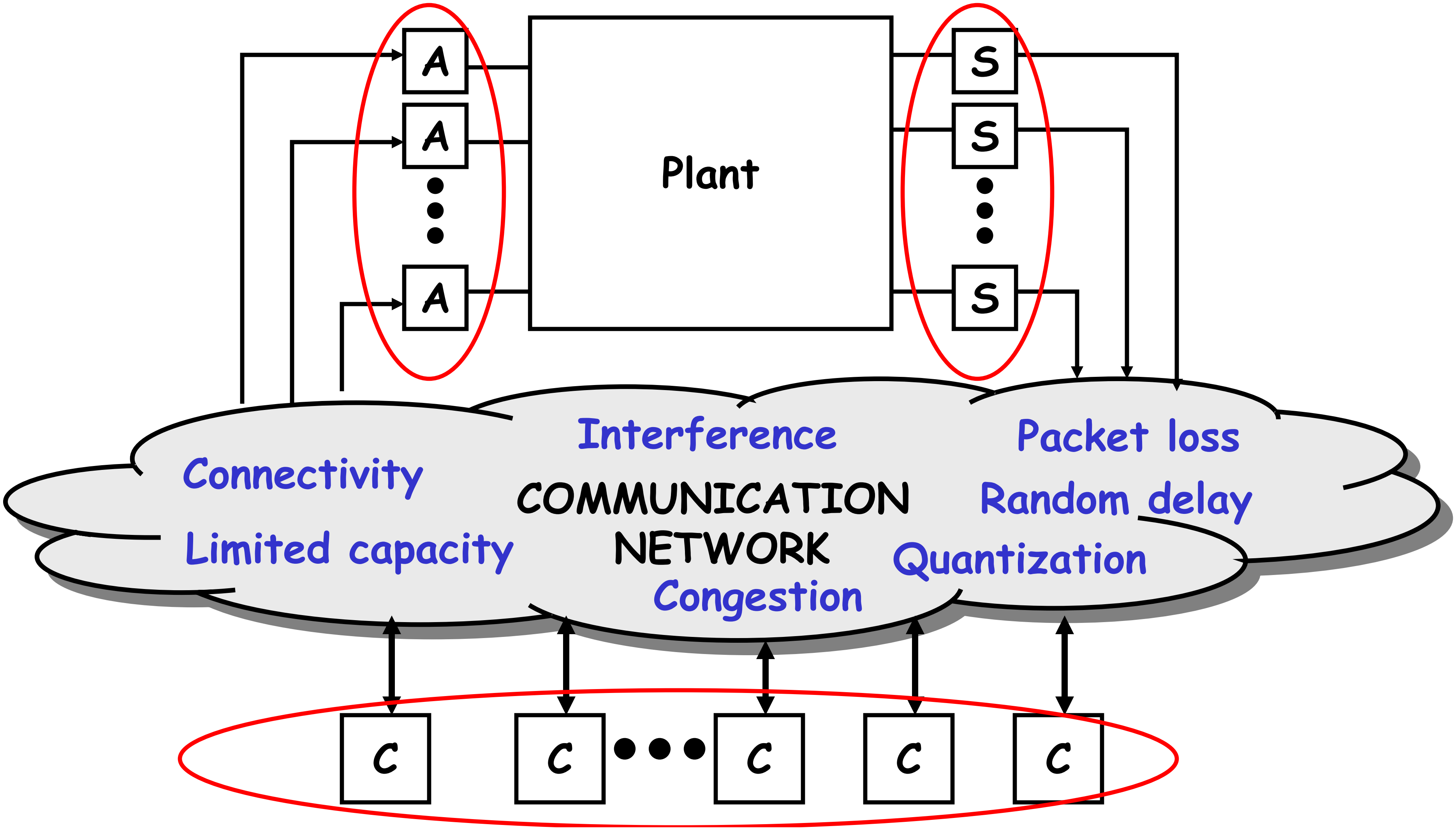
NCSs: what's new for control?

Classical architecture: Centralized structure



NCSs: what's new for control?

NCSs: Large scale distributed structure



DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Outline

- Motivations and target applications
- Overview of consensus algorithms
- Application of consensus to WSN:
  - Sensor calibration
  - Least-square parameter identification
  - Time-synchronization
- Open problems
  - Identification
  - Estimation
  - Control

The consensus problem

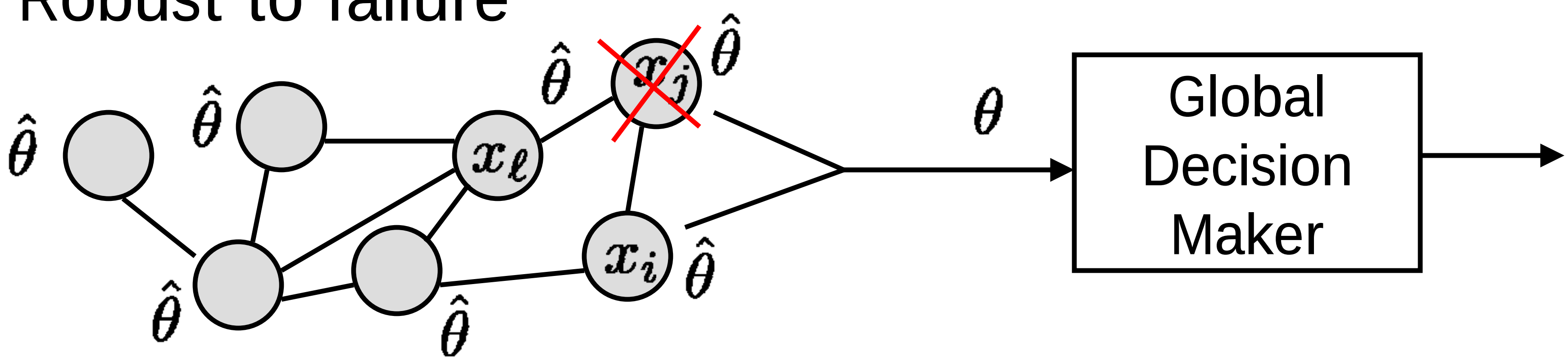
- Main idea
  - Having a set of agents to agree upon a certain value (usually global function) using only local information exchange (local interaction)
- Also known as:
  - Agreement problem (economics, signal processing, social networks)
  - Gossip algorithms (CS & communications)
  - Synchronization (statistical mechanics)
  - Rendezvous (robotics)
- Suitable for (noisy) sensor networks

Main features

- Distributed computation of general functions

$$\theta = f(x_1, \ldots, x_N) = f\left(\frac{1}{N}\sum_{i=1}^{N} g_i(x_i)\right) \qquad \text{(ex. } \theta = \frac{1}{N}\sum_{i=1}^{N} x_i \text{ for } f = g = ident \text{)}$$

- Computational efficient (linear & asynchronous)
- Independent of graph topology
- Incremental (i.e. anytime)
- Robust to failure



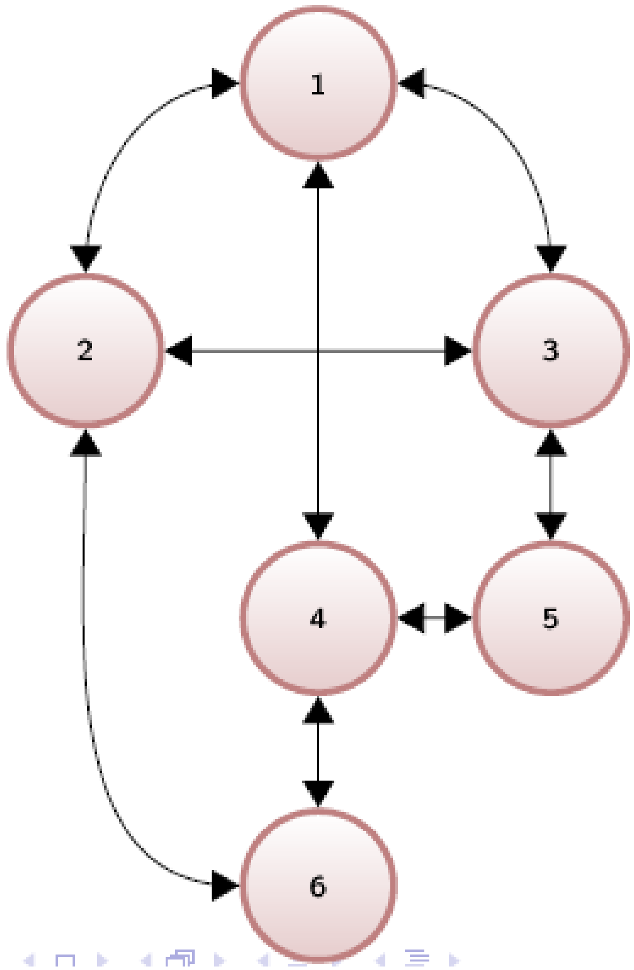
Some history (in control)

- Convergence of Markov Chains (60's) and Parallel Computation Alg.(70's)
- John Tsitsiklis "*Problems in Decentralized Decision Making and Computation* ", Ph.D thesis, MIT 1984
- A. Jadbabaie, J. Lin, and A. S. Morse "*Coordination of groups of mobile autonomous agents using nearest neighbor rules*", CDC' 02 (Axelby Best Paper Award TAC)
- Time-varying topologies (worst-case)
  - **L. Moreau**, "*Consensus seeking in multi-agent systems using dynamically changing interaction topologies*," IEEE, Transactions on Automatic Control, vol 50, No. 2, 2005
  - **M. Cao, A. S. Morse, and B. D. O. Anderson**. "*Reaching a Consensus in a Dynamically Changing Environment: A Graphical Approach*." SIAM Journal on Control and Optimization, Feb 2008
- Randomized topologies
  - **S. Boyd, A. Ghosh, B. Prabhakar, D. Shah** "*Randomized Gossip Algorithms*", TIT 2006
  - **F. Fagnani, S. Zampieri**, "*Randomized consensus algorithms over large scale networks*", JSAC 08
- Applications:
  - Vehicle coordination: Jadbabaie, Francis's group, Tanner, ...
  - Kalman Filtering: Olfati Saber-Murray, Alighanbari-How, Carli-Chiuso-Schenato-Zampieri
  - Generalized means: Giarre',Cortes
  - Time-synchronization: Solis-P.R. Kumar,Osvlado-Spagnolini, Carli-Chiuso-Schenato-Zampieri
  - WSN sensor calibration and parameter identification: Bolognani-DelFavero-Schenato-Varagnolo

Consensus formulation

Network of

- N agents
- Communication graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$
- i-th node neighbors: $\mathcal{N}(i)$
- Every node stores a variable: node i stores x_i.

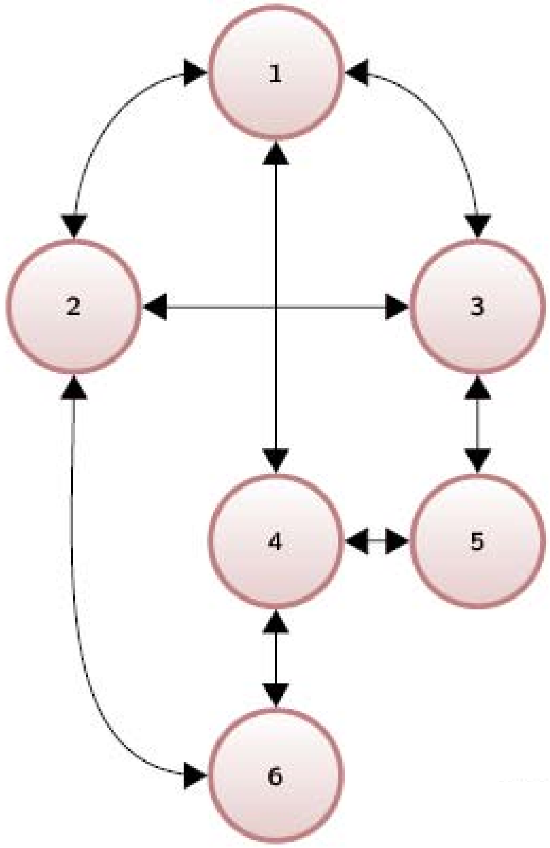


Consensus formulation (cont')

Definition (Recursive Distributed Algorithm adapted to the graph $\mathcal{G}$)

Any recursive algorithm where the i node's update law of depends only on the state of i and in its neighbors $j \in \mathcal{N}(i)$

$$x_i(t+1) = f(x_i(t), x_{j_1}(t), \ldots x_{j_{N_i}}(t)) \quad \text{with } j_1, \ldots, j_{N_i} \in \mathcal{N}(i)$$

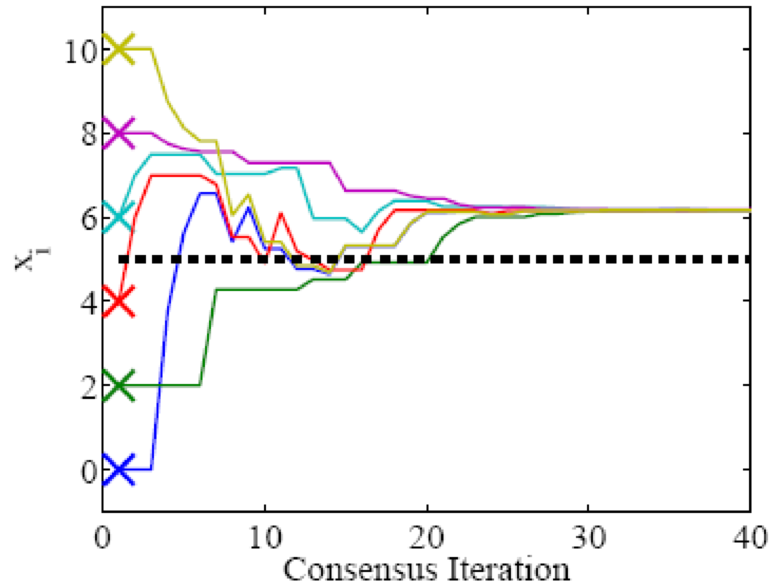


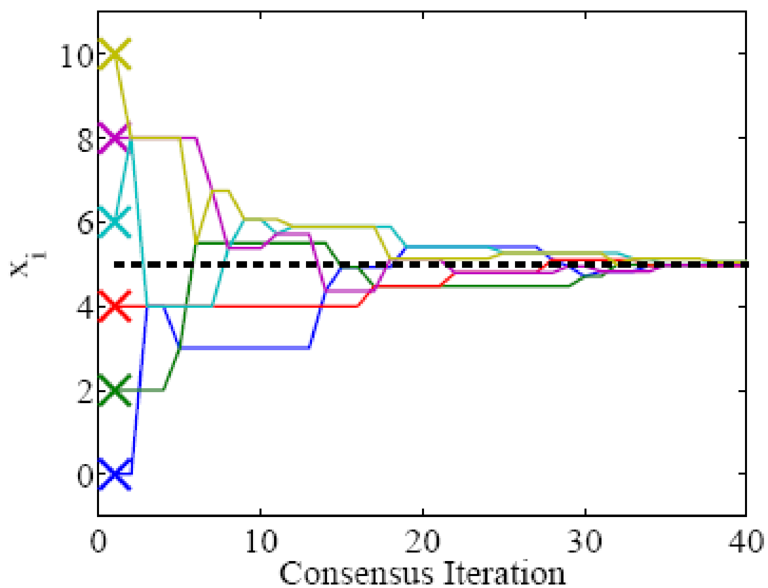
Consensus definition

Definition

A Recursive Distributed Algorithm adapted to the graph $\mathcal{G}$ is said to asymptotically achieve consensus if

$$x_i(t) \longrightarrow \alpha \qquad \forall i \in \mathcal{N}$$





Definition

A Recursive Distributed Algorithm adapted to the graph $\mathcal{G}$ is said to asymptotically achieve *average* consensus if

$$x_i(t) \longrightarrow \frac{1}{N} \sum_{i \in \mathcal{N}} x_i(0) \qquad \forall i \in \mathcal{N}$$

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Linear consensus

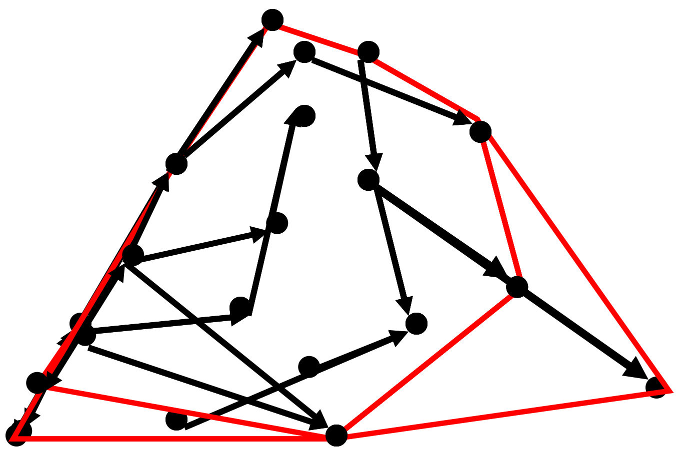
$$x_i(t+1) = p_{ii}x_i(t) + \sum_{j\in\mathcal{N}(i)} p_{ij}x_j(t)$$

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{bmatrix} \qquad x(t+1) = P(t)x(t)$$

Say $\mathcal{G}_P$ Graph associated to P, $P_{i,j} \neq 0 \iff (i,j) \in \mathcal{E}_P$,

$$\mathcal{G}_P \subseteq \mathcal{G} \qquad (\mathcal{N} \equiv \mathcal{N}_P, \quad \mathcal{E} \subseteq \mathcal{E}_P)$$

A robotics example: the rendezvous problem



$$x_i(t+1) = x_i(t) + u_i(t)$$
$$x_i(t+1) = p_{ii}x_i(t) + \sum_{j \in N(i)} p_{ij}x_j$$

Convex hull always shrinks.
If communication graph sufficiently connected, then shrinks to a point

Stochastic matrix

Definition (Stochastic Matrix)

If $P_{i,j} \geq 0$ and $\sum_j P_{i,j} = 1 \; \forall i$, than P is said to be stochastic

$$P\mathbb{1} = \mathbb{1} \qquad \mathbb{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Remark

If P is stochastic the linear algorithm can be written in both forms:

$$x_i(t+1) = p_{ii}x_i(t) + \sum_{j \in \mathcal{N}(i)} p_{ij}x_j(t)$$

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}(i)} p_{ij}\left(x_j(t) - x_i(t)\right)$$

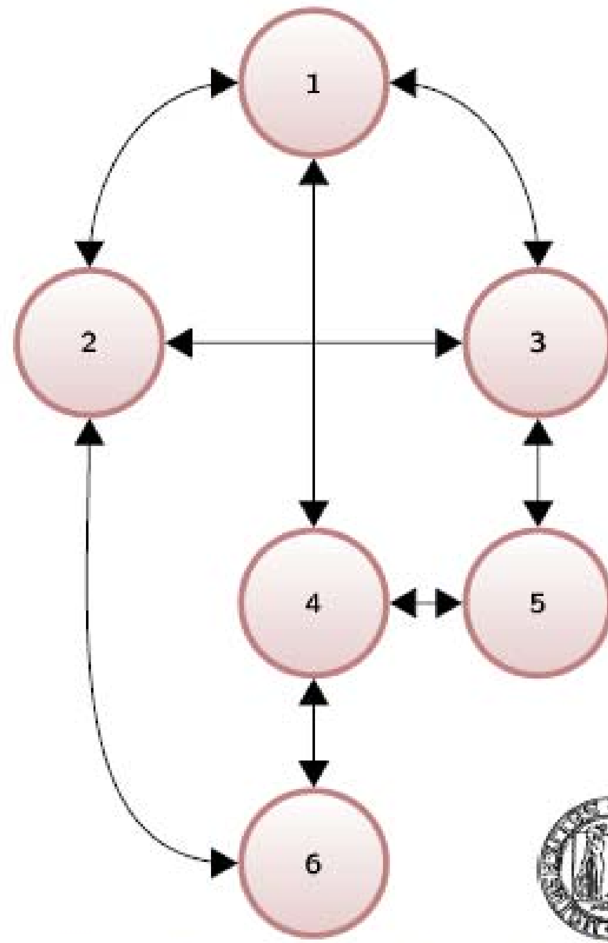
Constant matrix P

Synchronous Communication:

At each time all nodes communicate according to the communication graph

P(t)=P:

$$
P = \begin{bmatrix}
3/6 & 1/6 & 1/6 & 1/6 & 0 & 0 \\
1/6 & 3/6 & 1/6 & 0 & 0 & 1/6 \\
1/6 & 1/6 & 3/6 & 0 & 1/6 & 0 \\
1/6 & 0 & 0 & 3/6 & 1/6 & 1/6 \\
0 & 0 & 1/6 & 1/6 & 4/6 & 0 \\
0 & 1/6 & 0 & 1/6 & 0 & 4/6
\end{bmatrix}
$$



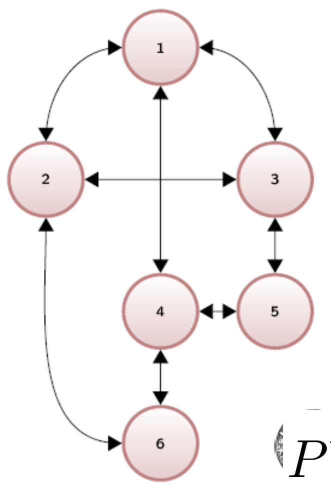
Convergence results

Theorem

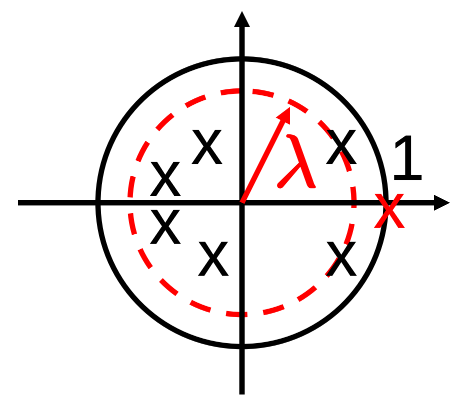
$P(t) = P$ stochastic.

- *If P such that $\mathcal{G}_P \subseteq \mathcal{G}$ is rooted then the algorithm achieves consensus*
- *If also P^T is stochastic (P doubly stochastic), then average consensus is achieved*

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} & 0 & 0 \\ p_{2,1} & p_{2,2} & p_{2,3} & 0 & 0 & p_{2,6} \\ p_{3,1} & p_{3,2} & p_{3,3} & 0 & p_{3,5} & 0 \\ p_{4,1} & 0 & 0 & p_{4,4} & p_{4,5} & p_{4,6} \\ 0 & 0 & p_{5,4} & p_{5,3} & p_{5,5} & 0 \\ 0 & p_{6,2} & 0 & p_{6,4} & 0 & p_{6,6} \end{bmatrix}$$



eigenvalues of P:



$P^t \xrightarrow{\lambda^t} \mathbb{1}\rho^T$, ρ is left eigenvector of 1

$\sum_i \rho_i = 1, \rho_i \geq 0, \quad (\rho_i > 0$ if strong. conn.)

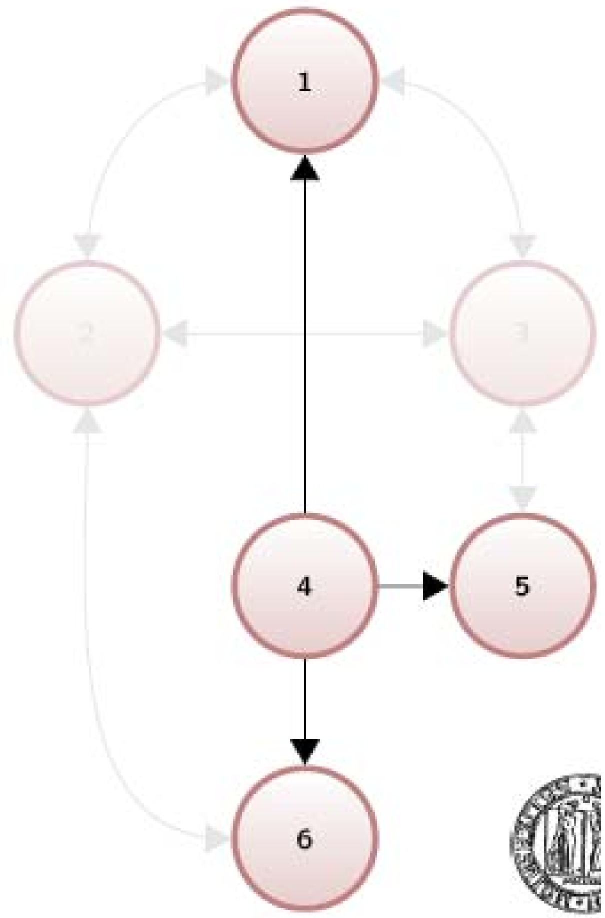
$\rho = \frac{1}{N}\mathbb{1}$ if P doubly stochastic

Time varying P(t): broadcast

Broadcast:

At each time one node randomly wakes up and broadcasts its information to all its neighbors.

$$P(t) = \begin{bmatrix} 3/4 & 0 & 0 & 1/4 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 3/4 & 0 \\ 0 & 0 & 0 & 1/4 & 0 & 3/4 \end{bmatrix}$$

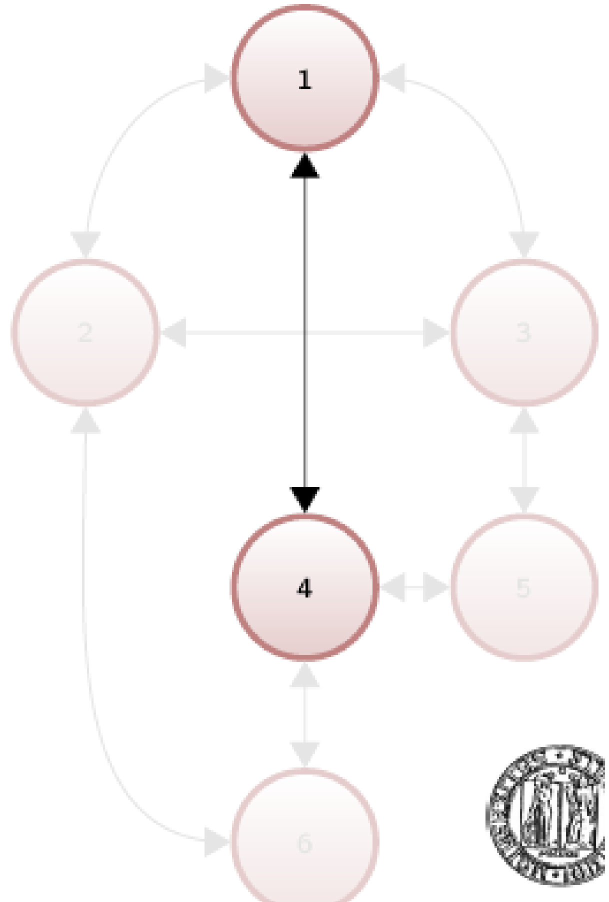


Time varying P(t): symmetric gossip

Symmetric Gossip:

At each time one node randomly wakes up and chose randomly a its neighbor. Those two nodes exchange information

$$P(t) = \begin{bmatrix} 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

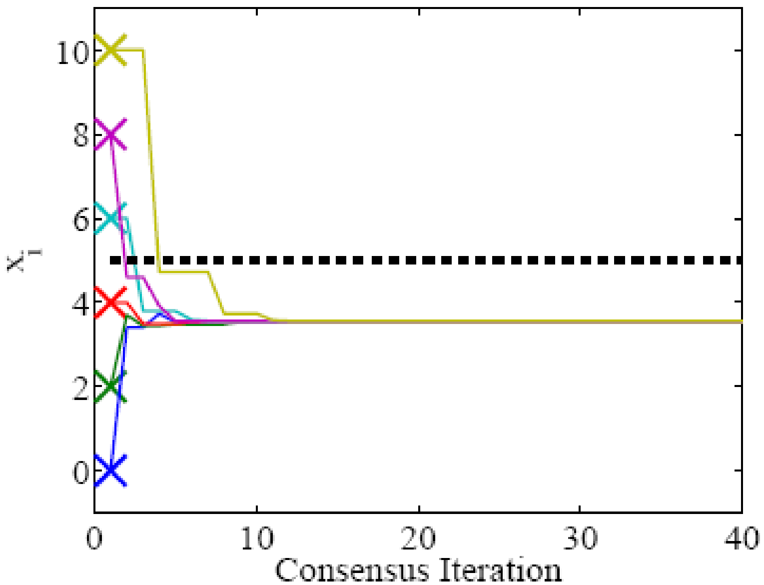


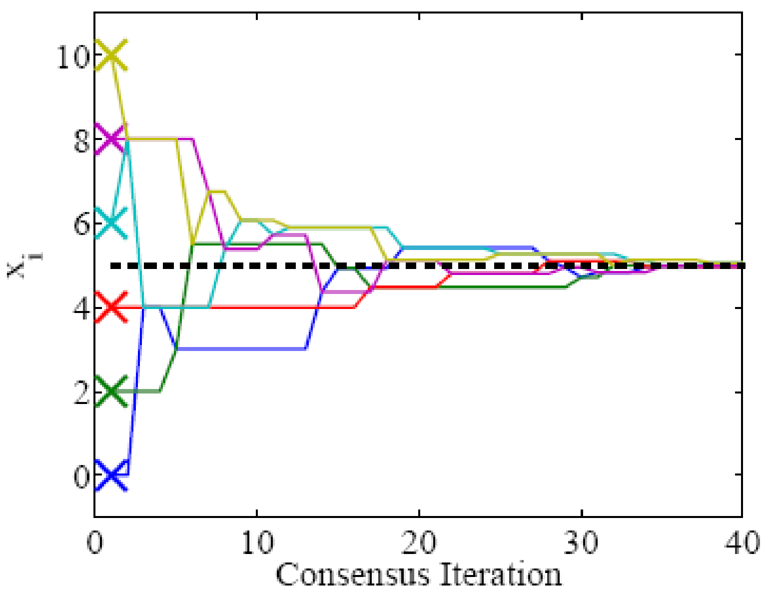
Consensus strategies for WSN

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Broadcast

- 1 message broadcasted, $|\mathcal{N}(i)|$ estimate updated
- Does not guarantee average consensus





Symmetric Gossip

- At least 3 messages exchanged, 2 estimate updated
- Guarantee average consensus

Convergence results: P=P(t) deterministic

Theorem

Suppose that $P_{ii}(t) > 0, \forall i, \forall t$ and that there exixts K such that $\mathcal{G}_\ell = \mathcal{G}_{P((\ell+1)K)} \cup \ldots \cup \mathcal{G}_{P(\ell K)}$ is rooted at some node j for all ℓ then

- the sequence $\{P(t)\}$ achieves consensus
- if also $P^T(t)$ are stochastic for all t, then the sequence $\{P(t)\}$ achieves *average* consensus

Remark:

Estimates of rate of convergence are very conservative (worst case)

L. Moreau, "*Consensus seeking in multi-agent systems using dynamically changing interaction topologies*," IEEE, Transactions on Automatic Control, vol 50, No. 2, 2005

M. Cao, A. S. Morse, and B. D. O. Anderson. "*Reaching a Consensus in a Dynamically Changing Environment: A Graphical Approach*." SIAM Journal on Control and Optimization, Feb 2008

Convergence results: P=P(t) randomized

Theorem

Suppose $\{P(t)\}$ is a sequence of i.i.d. stochastic random matrices. Suppose moreover $\mathcal{G}_{P(t)} \subseteq \mathcal{G}$ $\forall t$ and call $\bar{P} = \mathbb{E}[P]$.

- *If $\mathcal{G}_{\bar{P}}$ is rooted that consensus is achieved w.p.1*
- *If also $P(t)^T$ is stochastic for every t, then average consensus is achieved w.p.1*

Remark:

It is not sufficient $\bar{P}$ doubly stochastic to guarantee average consensus

$$x(t+1) = P(t)x(t) = P(t)P(t-1)\cdots P(0)x(0) = Q(t)x(0) \qquad (Q(t) = P^t \text{ if } P(t) = P)$$

$$Q(t) \to \mathbb{1}\rho^T, \ \mathbb{E}[\rho] = \frac{1}{N}\mathbb{1}, \ \boxed{Var(\rho) \sim \frac{1}{N}}$$

F. Fagnani, S. Zampieri, "*Randomized consensus algorithms over large scale networks*", IEEE Journal on Selected Areas in Communications, 2008

Generalized mean

$$x_i(t+1) = p_{ii}x_i(t) + \sum_{\mathcal{N}_i} p_{ij}x_j(t) \qquad x_i(t) \to \frac{1}{N}\sum_i x_i(0)$$

$$\theta = f(a_1, \dots, a_N) = f\left(\frac{1}{N}\sum_{i=1}^{N} g_i(a_i)\right)$$

$$x_i(0) = g_i(a)$$
$$\hat{\theta}_i(t) = f(x_i(t))$$

$$\text{Geometric mean: } \theta = \sqrt[n]{\prod_i a_i} = \exp\left(\frac{1}{N}\sum_i \log(a_i)\right)$$

$$x_i(0) = \log(a_i), \quad \hat{\theta}_i(t) = \exp(x(t)) \to \theta$$

$$\text{Armonic mean: } \theta = \left(\frac{1}{N}\sum_i \frac{1}{a_i}\right)^{-1} \qquad x_i(0) = \frac{1}{a_i}, \quad \hat{\theta}_i(t) = \frac{1}{x_i(t)} \to \theta$$

$$\text{Quadratic mean: } \theta = \sqrt{\frac{1}{N}\sum_i a_i^2} \qquad x_i(0) = a_i^2, \quad \hat{\theta}_i(t) = \sqrt{x_i(t)} \to \theta$$

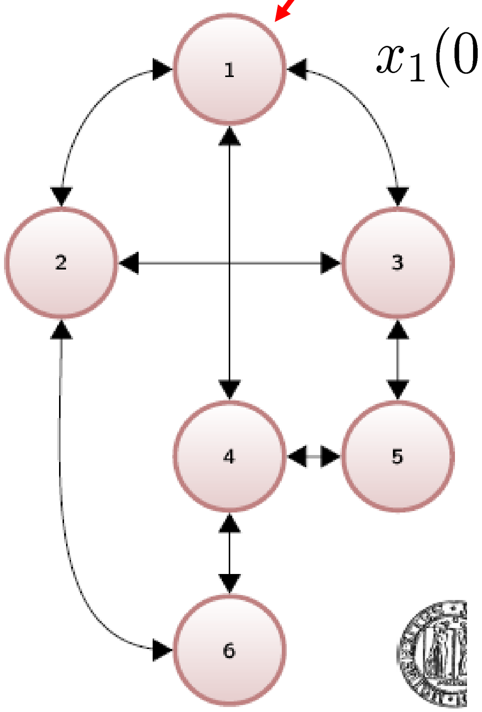
D. Bauso, L. Giarre' and R. Pesenti, "*Nonlinear protocols for Optimal Distributed Consensus in Networks of Dynamic Agents*", Systems and Control Letters, 2006

J. Cortés, *Distributed algorithms for reaching consensus on general functions,* Automatica 44 (3) (2008), 726-737

Node counting

$$x_i(t+1) = p_{ii}x_i(t) + \sum_{\mathcal{N}_i} p_{ij}x_j(t) \qquad x_i(t) \to \frac{1}{N}\sum_i x_i(0)$$

leader



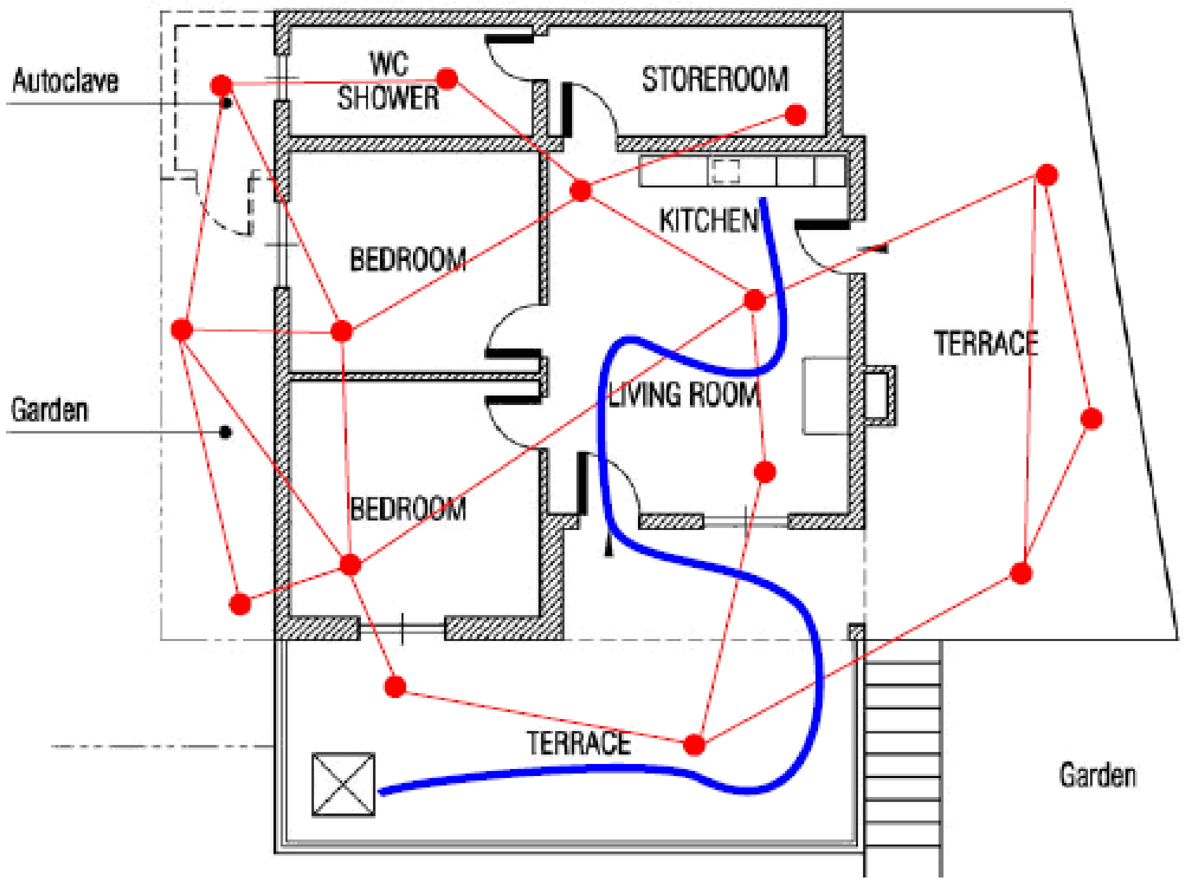
$$x_1(0) = 1, \quad x_i(0) = 0, i = 2, \ldots, N \implies \frac{1}{N}\sum_i x_i(0) = \frac{1}{N}$$

$$\hat{\theta}_i(t) = \frac{1}{x_i(t)} \to N$$

Outline

- Motivations and target applications
- Overview of consensus algorithms
- Application of consensus to WSN:
  - Sensor calibration
  - Least-square parameter identification
  - Time-synchronization
- Open problems
  - Identification
  - Estimation
  - Control

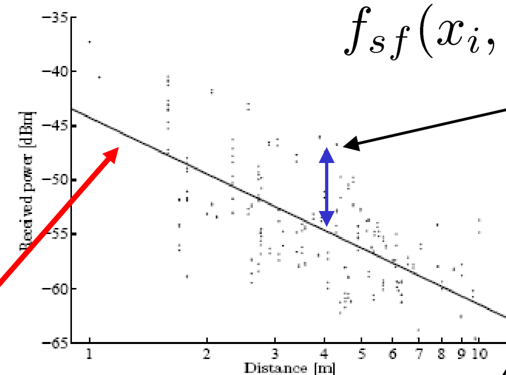
Localization with WSN



Localization with WSN

Each node

can measure the Radio Signal Strength Indicator, RSSI, i.e. the received signal power P_{rx} in *dBm*.



$$f_{sf}(x_i, x_j) = f_{sf}(x_j, x_i)$$

$$P_{rx}^{ij} = P_{tx}^{j} + \beta - 10\gamma \log_{10}(||x_i - x_j||) + f_{sf}(x_i, x_j) + v(t) + o_i$$

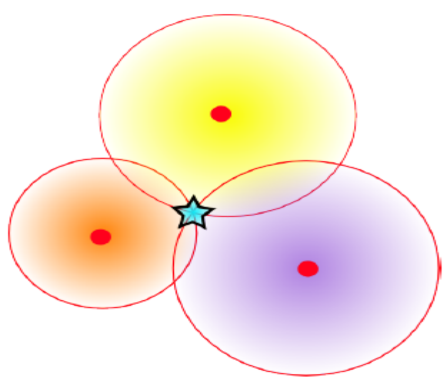
Transmi

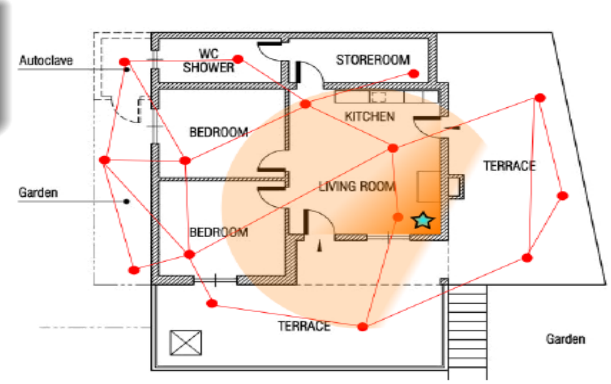
pow

Map Based

Most likely location that matches with pre-learned maps.

ıeter



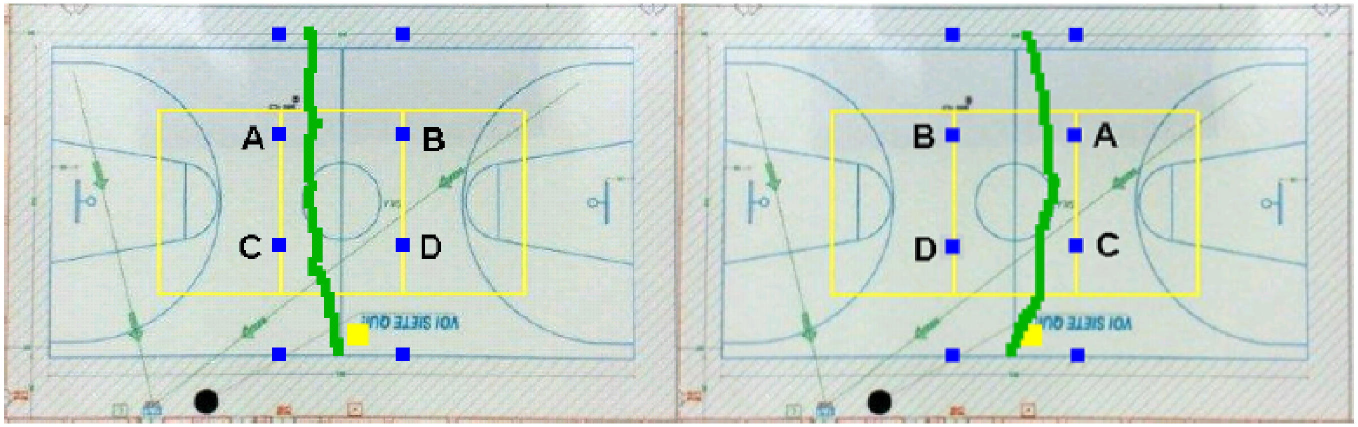


Range based

Triangulation (similarly to GPS)

Offset effect

Reception offset is particularly harmful for localization applications, Experiment inside a basketball court.[S 07][2]



[2][S 07] Courtesy of ST Microelectronics,
I. Solida, "Localization services for IEEE802.15.4/Zigbee devices. Mobile node tracking (in Italian)", Master Thesis, Department of information Engineering, University of Padua, 2007

WSN sensor calibration

Ideally:

- Estimate o_i: $\hat{o}_i$
- Use $\hat{o}_i$ to compensate the offset: $o_i - \hat{o}_i = 0$

Remember the previous example

What we propose is:

$o_i - \hat{o}_i = \alpha \qquad \alpha \cong 0$ equal for all nodes

All nodes overestimate or underestimate the distance similarly.
The errors, in the triangulation process, cancel out partially.

Calibration as consensus problem

Remark

If P is stochastic the linear algorithm can be written in both forms:

$$x_i(t+1) = p_{ii}x_i(t) + \sum_{j\in\mathcal{N}(i)} p_{ij}x_j(t)$$

$$x_i(t+1) = x_i(t) + \sum_{j\in\mathcal{N}(i)} p_{ij}\left(x_j(t) - x_i(t)\right) \qquad o_i - \hat{o}_i(t) = x(t)$$

$$o_i - \hat{o}_i(t+1) = o_i(t) - \hat{o}_i(t) + \sum_{j\in\mathcal{N}_i} p_{ij}\left((o_i - \hat{o}_i(t)) - (o_j - \hat{o}_j(t))\right)$$

$$\hat{o}_i(t+1) = \hat{o}_i(t) - \sum_{j\in\mathcal{N}_i} p_{ij}\left(P^{ij} - P^{ji} - \hat{o}_i(t) + \hat{o}_j(t)\right)$$

update equation

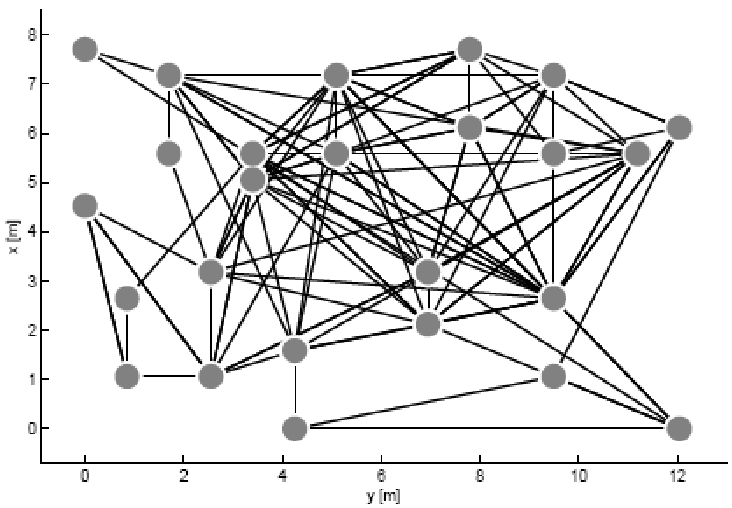
$$\hat{o}_i(t) \to o_i - \frac{1}{N}\sum_i o_i = o_i - \alpha \approx o_i$$

Steady state

Experimental Testbed

25 TMote-Sky nodes with Chipcon CC2420 RF Transceiver randomly placed inside a single conference room:

Network topology and nodes displacement:

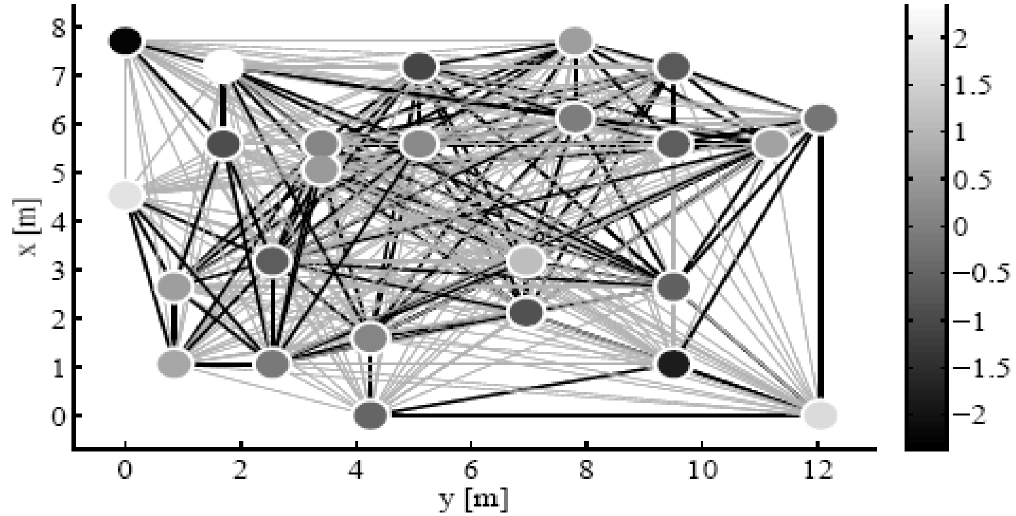


Kept just the links that safely carried the 75% of the sent messages over them

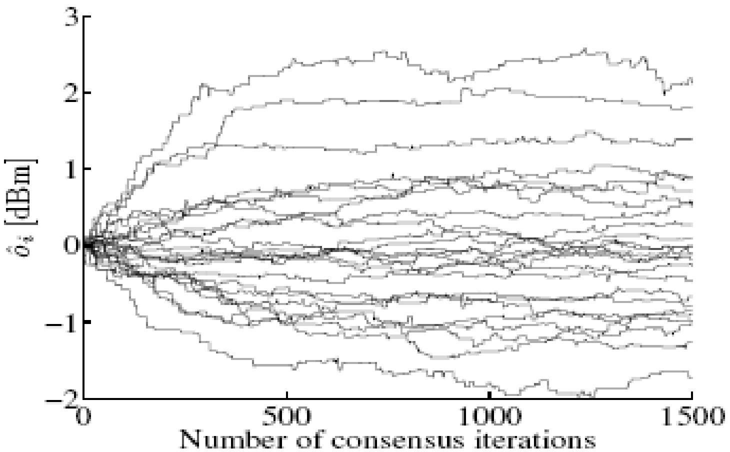
Experimental results

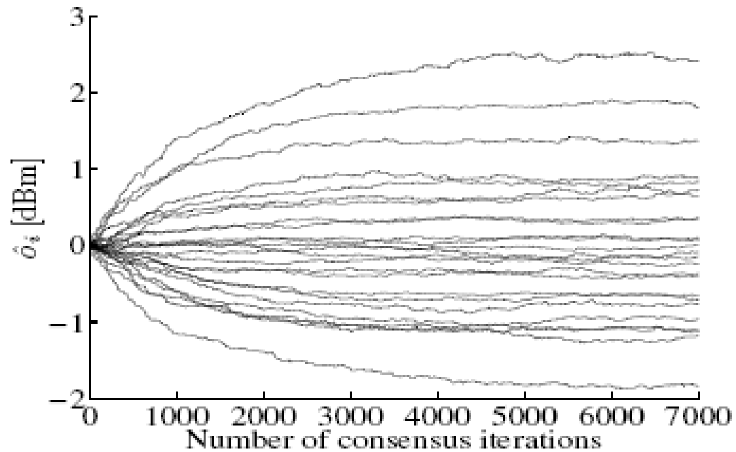
Links divided in 2 categories:

- Training links (black)
- Validation links (gray)



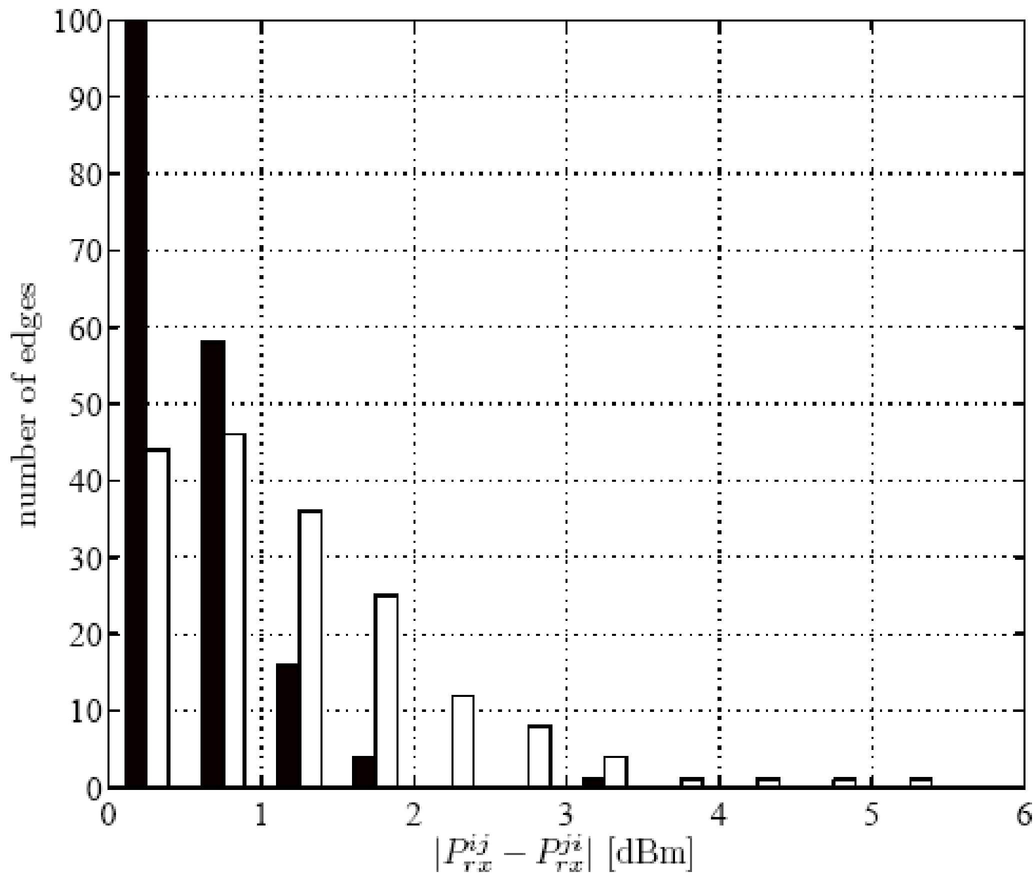
Estimate time evolution





DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

$$\Delta \bar{P}^{ij} = \bar{P}^{ij} - \bar{P}^{ji} = o_i - o_j$$



	Before	After
<0.5 dB	24%	56 %
<1	50%	88 %
>2dB	35%	0.6 %
Max	<6dB	<3.5dB

Effects of systematic errors when estimating distances

$1dB \longmapsto \cong 2m \pm 0.28m$.
$6dB \longmapsto$ uncertainty for $0.9m$ to $4.4m$ for an actual distance of $2m$.

$1dB \longmapsto \cong 10m \pm 1.4m$.

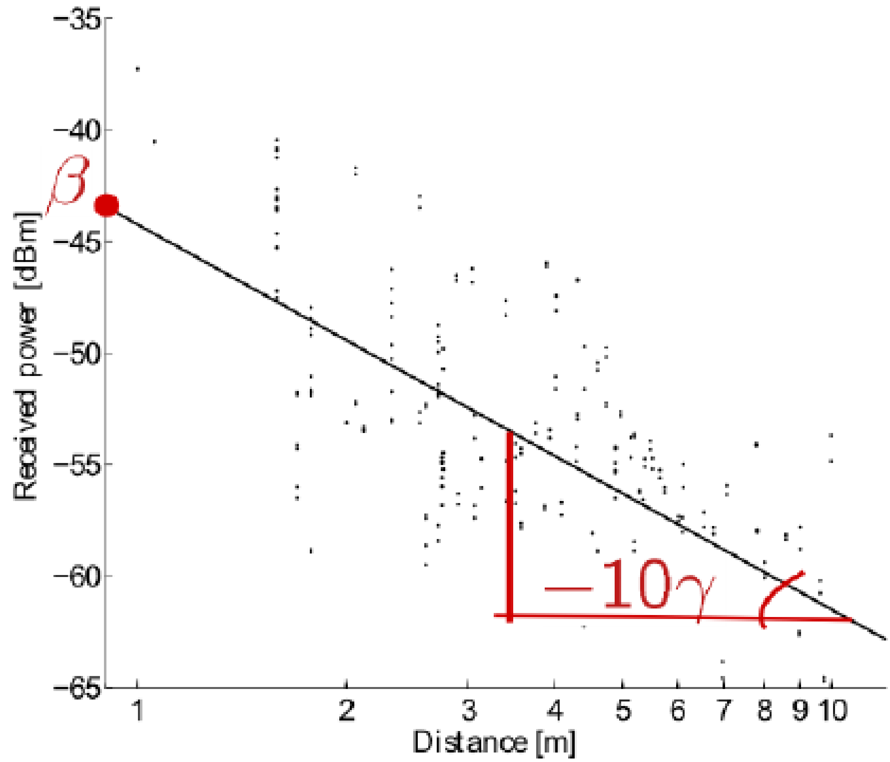
Parameter identification

Another important problem:

Accurately identify the wireless channel parameters β and γ.

In fact:

- Parameters extremely environment dependent
- $\gamma \in [1, 6]$
- Environment change hourly or daily



$$P_{rx}^{ij} = P_{tx}^{j} + \beta - 10\gamma \log_{10}(||x_i - x_j||) + f_{sf}(x_i, x_j) + v(t) + o_i$$

Modeling

Recall the Wireless Channel Model

$$\bar{P}_{rx}^{ij} + \hat{o}_i = P_{tx} - \beta - \gamma 10 \log_{10}(d_{ij}) + f_{sf}(\mathbf{x}_i, \mathbf{x}_j) + \underbrace{(o_i + \hat{o}_i)}_{\alpha} + w_i$$

$$\bar{P}_{rx}^{ij} + \hat{o}_i = \beta - \gamma 10 \log_{10}(d_{ij}) + w_i$$

For each link:

$$\underbrace{\bar{P}_{rx}^{ij} + \hat{o}_i}_{b_{ij}} = \underbrace{[1 \quad -10 \log_{10}(d_{ij})]}_{a_{ij}^T} \underbrace{\begin{bmatrix} \beta \\ \gamma \end{bmatrix}}_{\theta} + w_i$$

Modeling (cont'd)

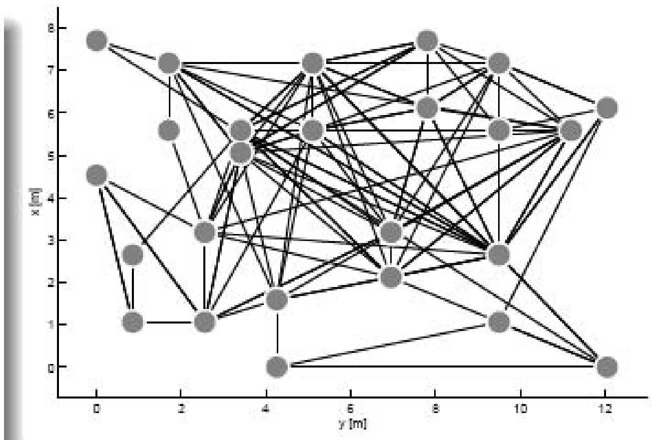
Each node

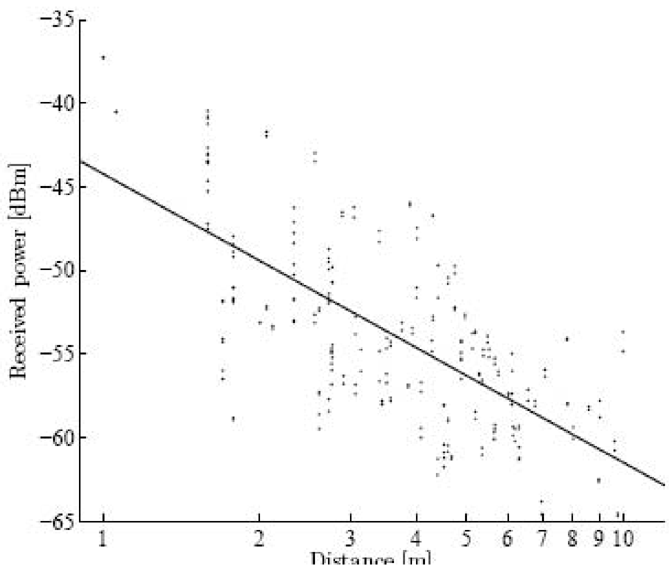
- knows its distance with its neighbor

$$d_{ij} \longrightarrow a_{ij}$$

- measures the strength of the message received form its neighbors

$$P_{ij} \longrightarrow b_{ij}$$





Globally the network collected

M couples measure-regressors: $(a_1, b_1), \ldots (a_M, b_M)$

For ease of notation, assume that

Each node stores one couple measure-regressor.

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Least-square Identification

Globally, the sensor network collected

M couples measure-regressors: $(a_1, b_1), \ldots (a_M, b_M)$.

Let us call

$$A = [a_1, \ldots, a_M]^T \text{ and } b = [b_1, \ldots, b_M].$$
$$b = A\theta + w$$

The least square estimate of θ,

given the measurements b is

$$\hat{\theta} = \arg\min_{\theta} ||A\theta - b|| = (A^T A)^{-1} A^T b$$

Consensus-based Identification

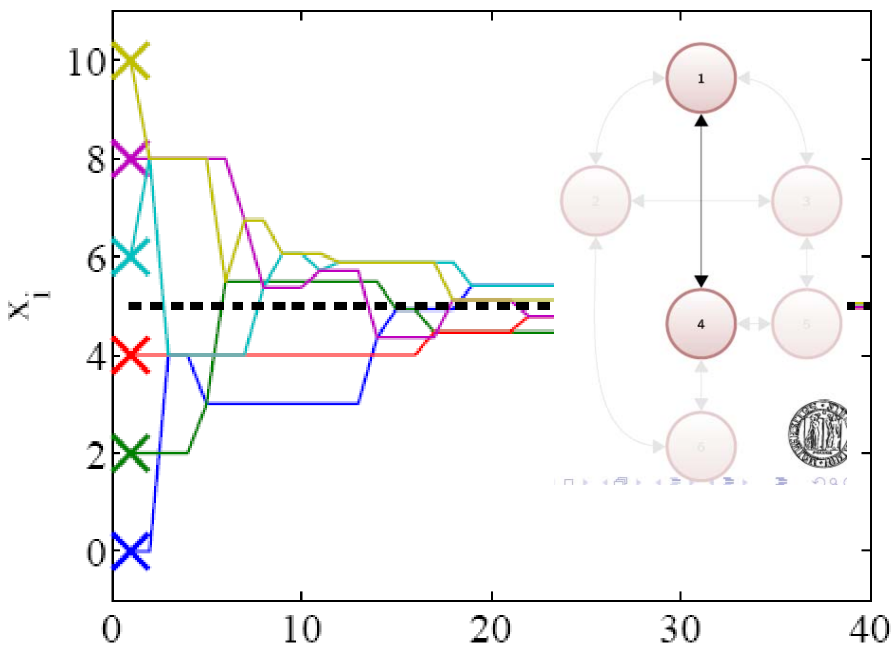
$$\hat{\theta} = \arg\min_{\theta} ||A\theta - b|| = (A^T A)^{-1} A^T b = (\frac{1}{N} \sum_{i \in \mathcal{N}} a_i a_i^T)^{-1} (\frac{1}{N} \sum_{i \in \mathcal{N}} a_i b_i)$$

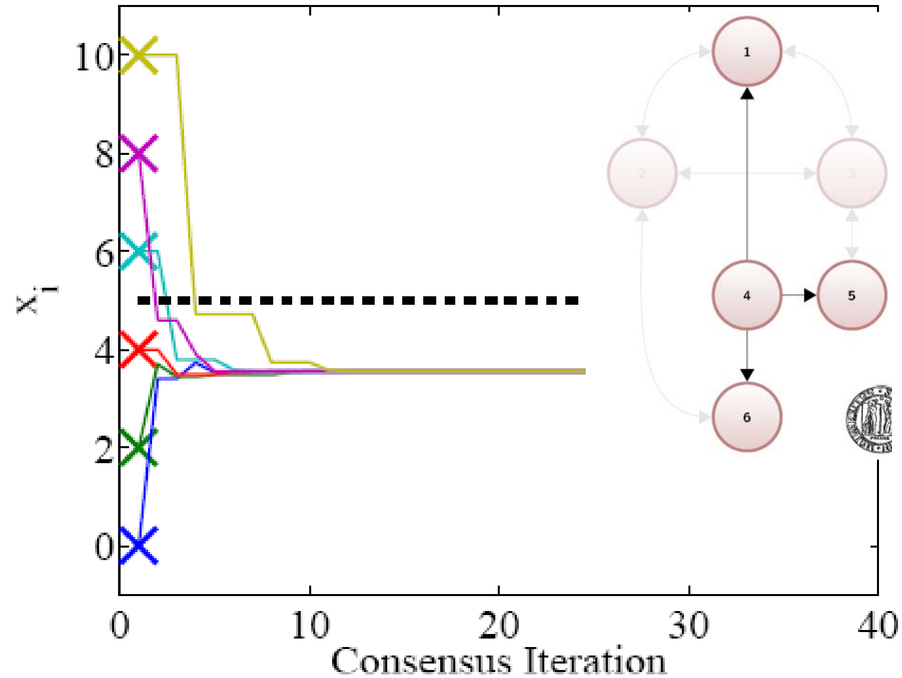
DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Communication schemes

Broadcast

- 1 message broadcasted, $|\mathcal{N}(i)|$ estimate updated
- Does not guarantee average consensus



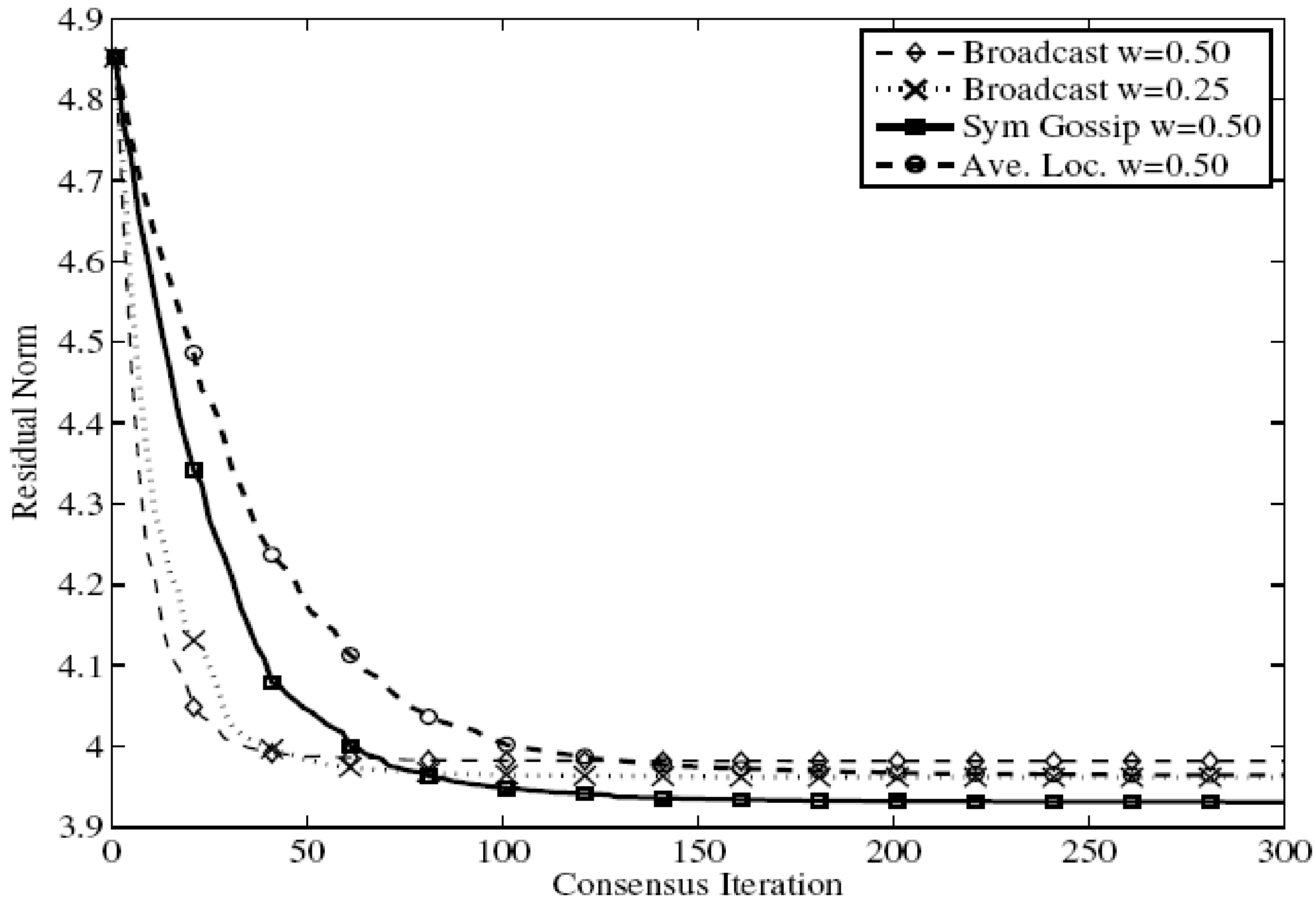


Symmetric Gossip

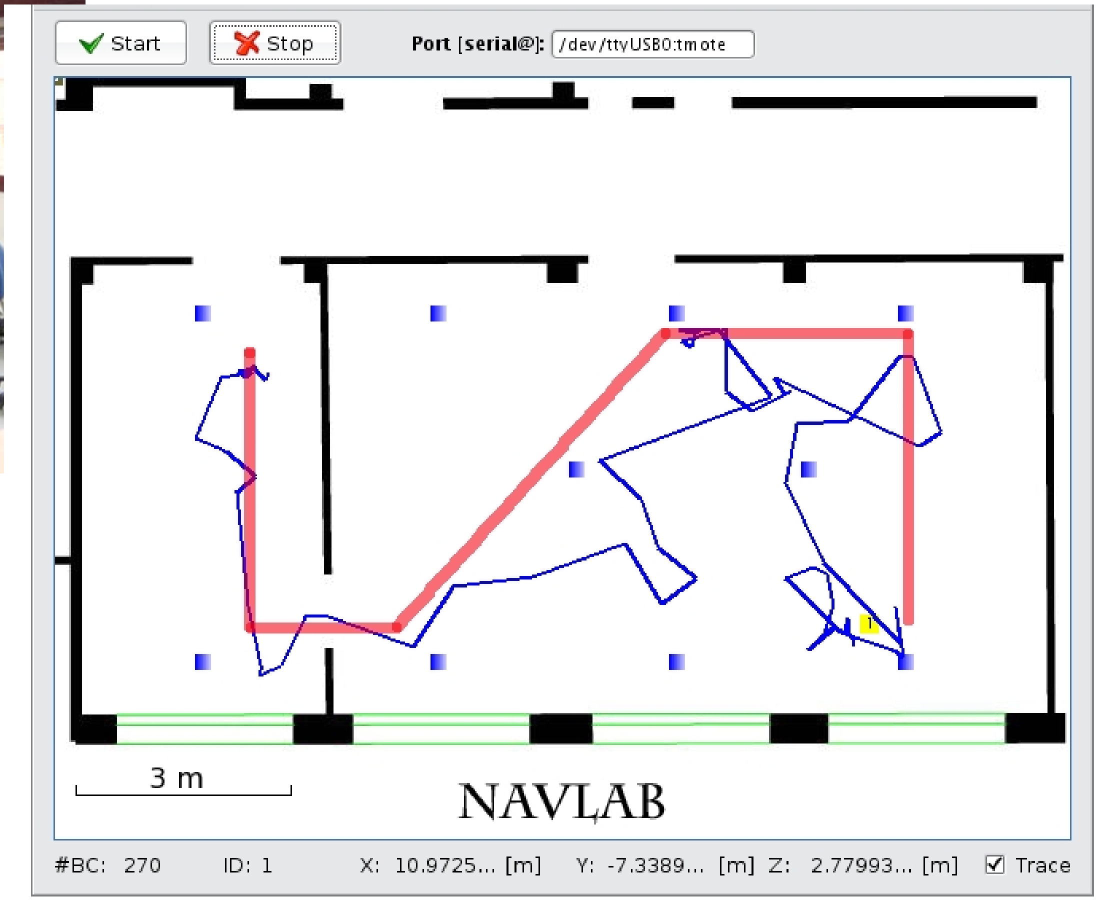
- At least 3 messages exchanged, 2 estimate updated
- Guarantee average consensus

Experimental results

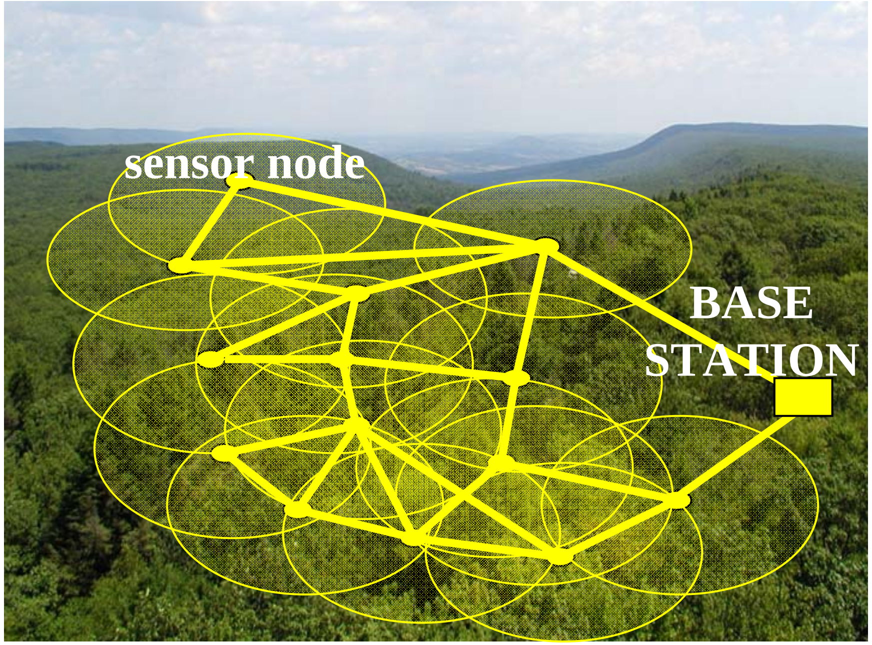
Residual: $\frac{1}{M}||A\hat{\theta} - b||^2$

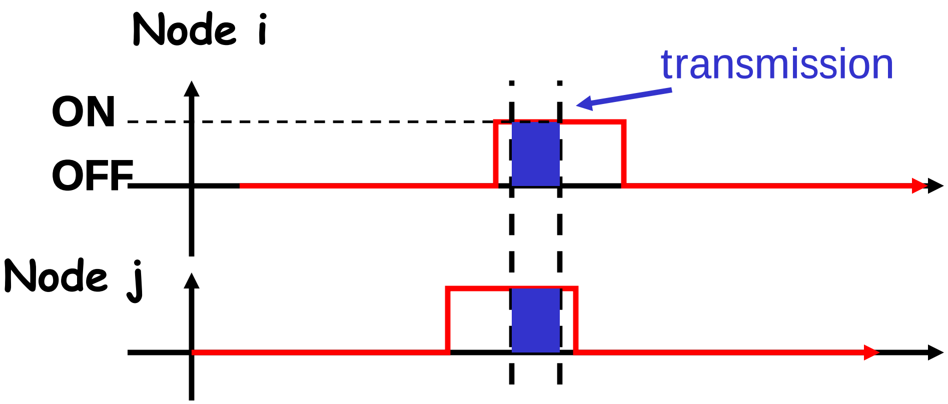


Tracking results

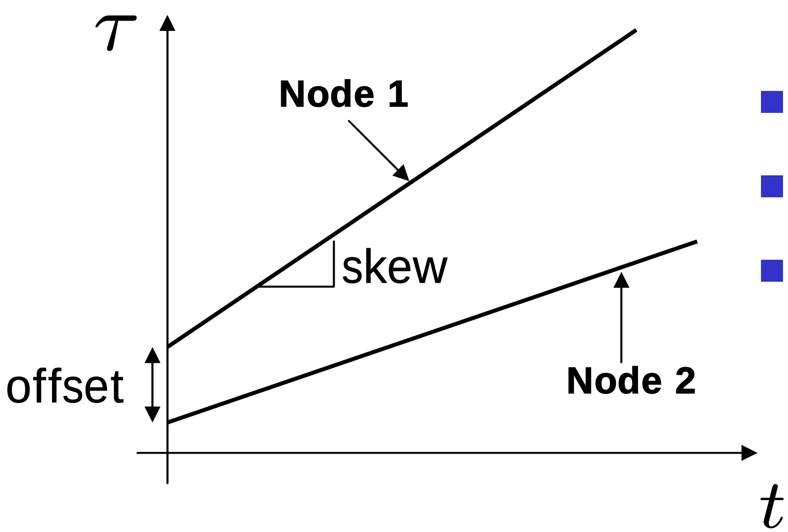


Time synchronization in sensor networks



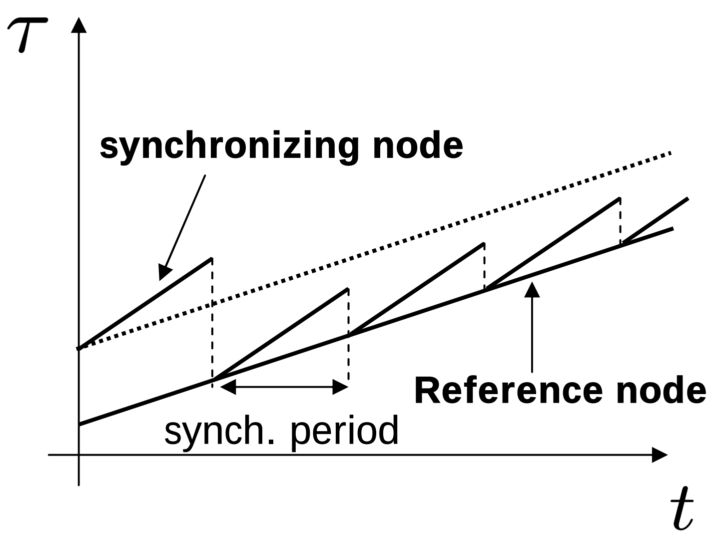


Clock characteristics & standard clock pair synch



- Offset: instantaneous time difference
- Skew: clock speed
- Drift: derivative of clock speed

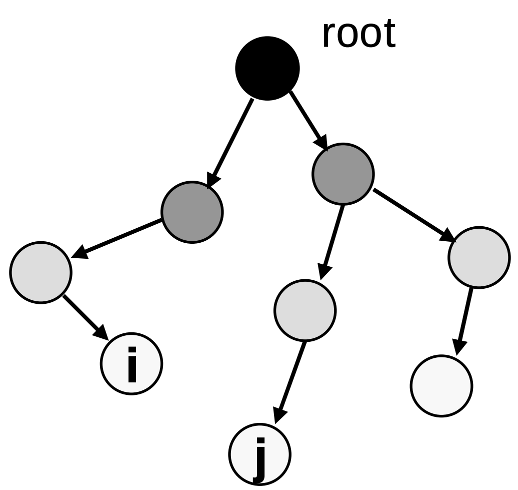
$$\tau_i = a_i t + b_i$$



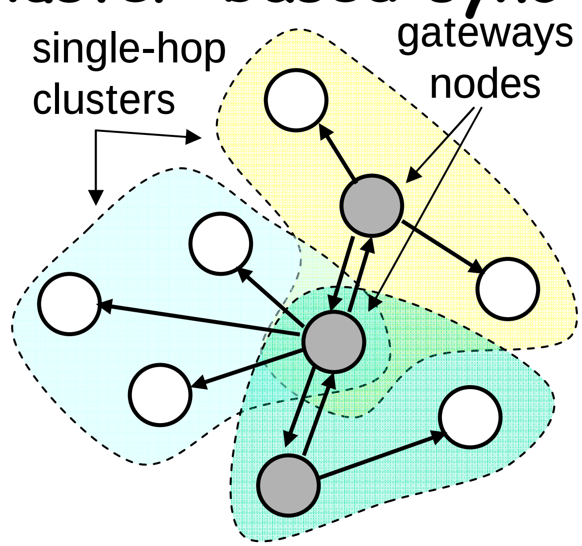
- Offset synch: periodically remove offset with respect to reference clock
- Skew compensation: estimate relative speed with respect to reference clock

State-of-the-art

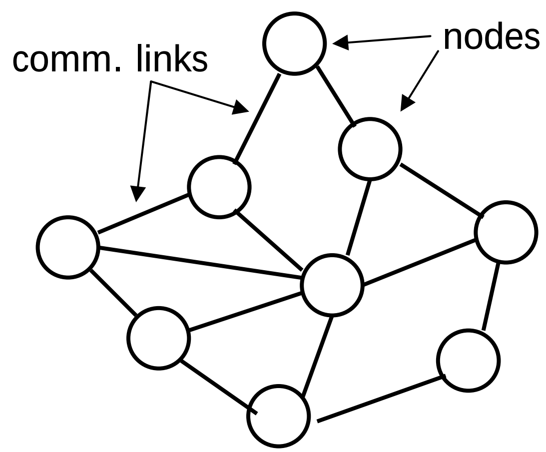
Tree-based sync



Cluster-based sync



Distributed



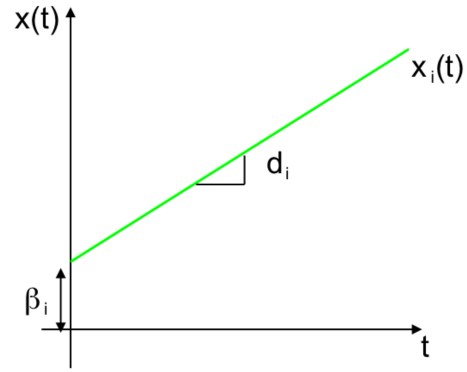
Modeling

MODEL: N clocks as discrete time integrators

$$x_i(t+1) = x_i(t) + d_i$$

d_i : *skew* (clock speed)

$x_i(0) = \beta_i$: *initial offset*



CONTROL: Assume that it is possible to control each clock by a local input $u_i(t)$:

$$x_i(t+1) = x_i(t) + d_i + u_i(t)$$

$$x(t+1) = x(t) + d + u(t)$$

GOAL: Clocks Synchronization

$$\lim_{t\to\infty} x_i(t) - x_j(t) = 0$$

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T \right) x(t) = 0$$

CONTROL: Proportional controller

$$u_i(t) = -\sum_{j\in\mathcal{N}(i)} k_{ij}(x_j(t) - x_i(t))$$

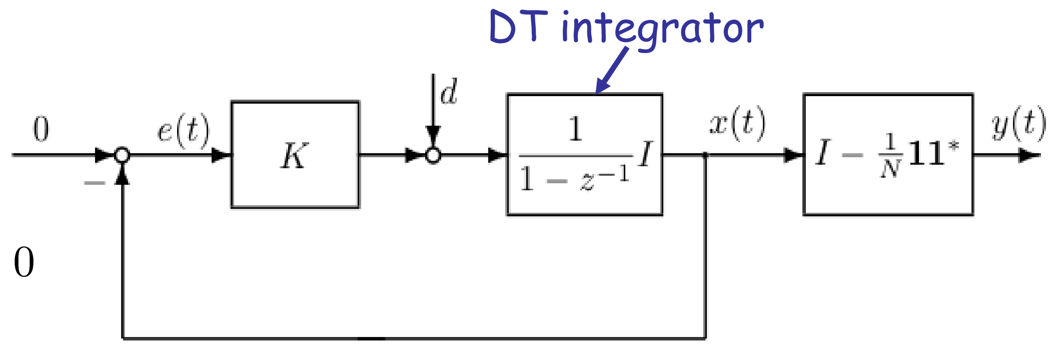
$$u(t) = -Kx(t)$$

P-control

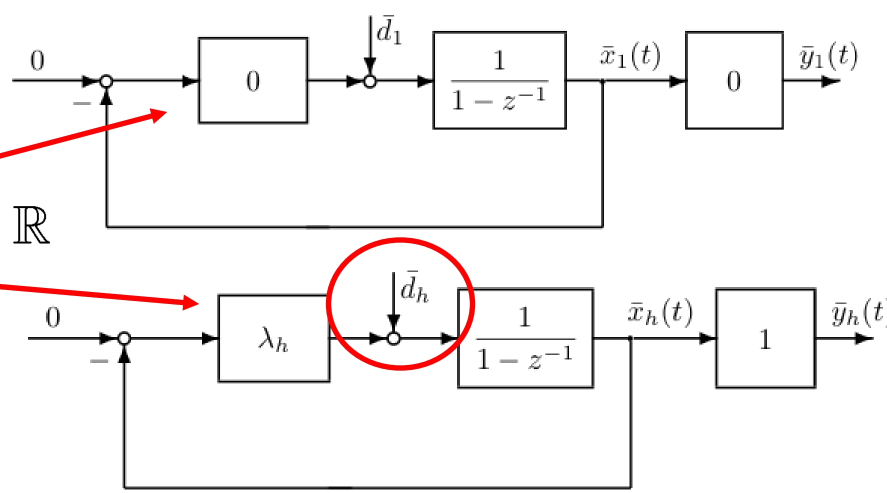
$$x(t+1) = x(t) + d + u(t)$$

$$u(t) = -Kx(t)$$

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right) x(t) = 0$$



If K symmetric:



eingenvalues of $K \in \mathbb{R}$

$\mathbf{h = 1}$

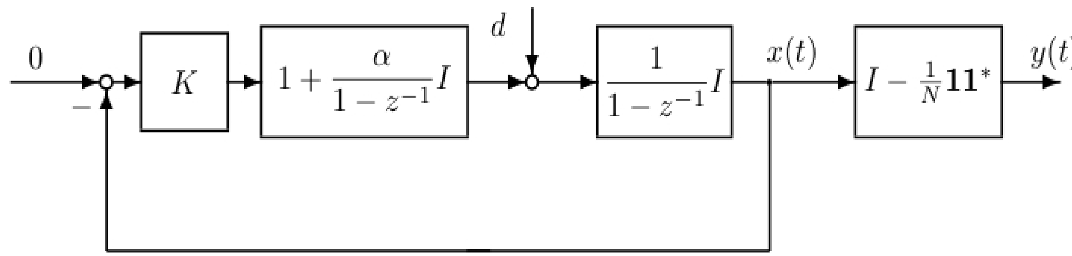
$\mathbf{h = 2, \ldots, N}$

PI-control

$$x(t+1) = x(t) + d + u(t)$$

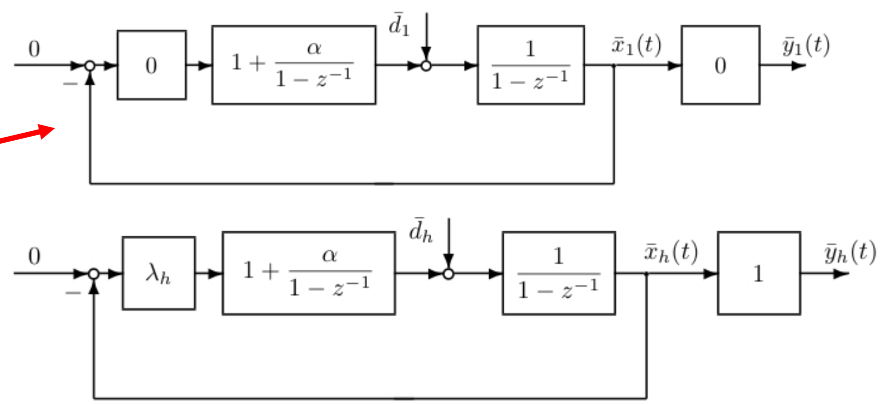
$$w(t+1) = w(t) - \alpha K x(t)$$
$$u(t) = w(t) - K x(t)$$

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right) x(t)$$



If K symmetric:

eingenvalues of $K \in \mathbb{R}$



$\mathbf{h = 1}$

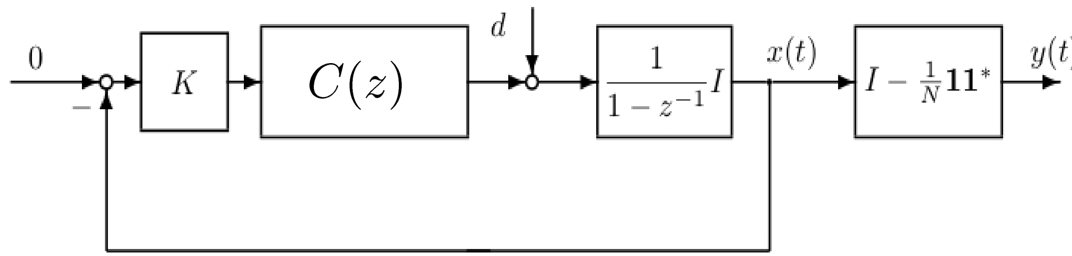
$\mathbf{h = 2, \ldots, N}$

C(z)-control

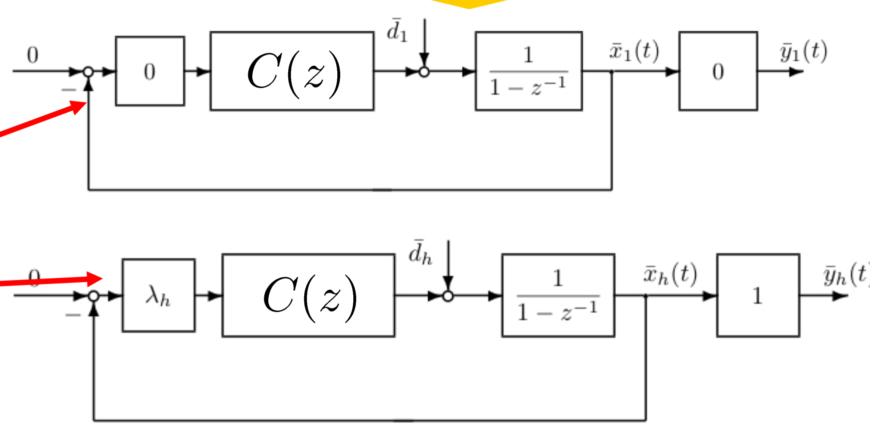
$$x(t+1) = x(t) + d + u(t)$$

$$u(t) \quad = \quad C(z)Kx(t)$$

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right) x(t)$$



If K symmetric:



$\mathbf{h = 1}$

eingenvalues of $K \in \mathbb{R}$

$\mathbf{h = 2, \ldots, N}$

Parameter design (undirected graphs)

GOAL: fastest rate of convergence

$$\begin{bmatrix} x(t+1) \\ w(t+1) \end{bmatrix} = \begin{bmatrix} I - K & I \\ -\alpha K & I \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix} + \begin{bmatrix} d \\ 0 \end{bmatrix}$$

- Suboptimal design (no topology needed):

$$k_{ij} = -\frac{1}{\max(d_i, d_j)+1} \; i \neq j, \quad \alpha = \frac{1}{2}, \quad \text{where } d_i \text{ is } \# \text{ of neighbors of node } i.$$

- Optimal design: almost convex problem (SDP + 1D non-convex search)

Model w/ noise

$$u_i(t) = -\sum_{j \in \mathcal{N}(i)} k_{ij}(x_j(t) - x_i(t))$$

white measurement noise   white process noise

$$\begin{bmatrix} x(t+1) \\ w(t+1) \end{bmatrix} = \begin{bmatrix} I-K & I \\ -\alpha K & I \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix} + \begin{bmatrix} -K \\ -\alpha K \end{bmatrix} v(t) + \begin{bmatrix} d \\ 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} n(t)$$

GOAL: smallest steady state mean square error: $J(K,\alpha) = \frac{1}{N} E[||y(\infty)||^2]$

- Suboptimal design still OK

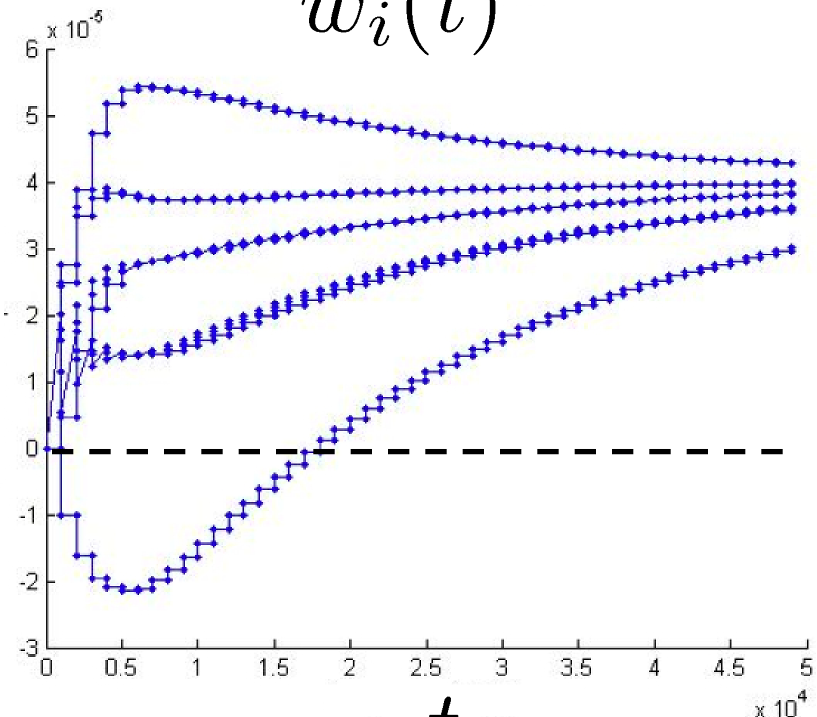
$$k_{ij} = \frac{1}{\max(d_i, d_j)+1}, \quad \alpha = \frac{1}{2}, \qquad \text{where } d_i \text{ is \# of neighbors of node } i.$$

- Optimal design: almost convex problem (Semidefinite programming in K+ 1D non-convex search in ff)

Simulations

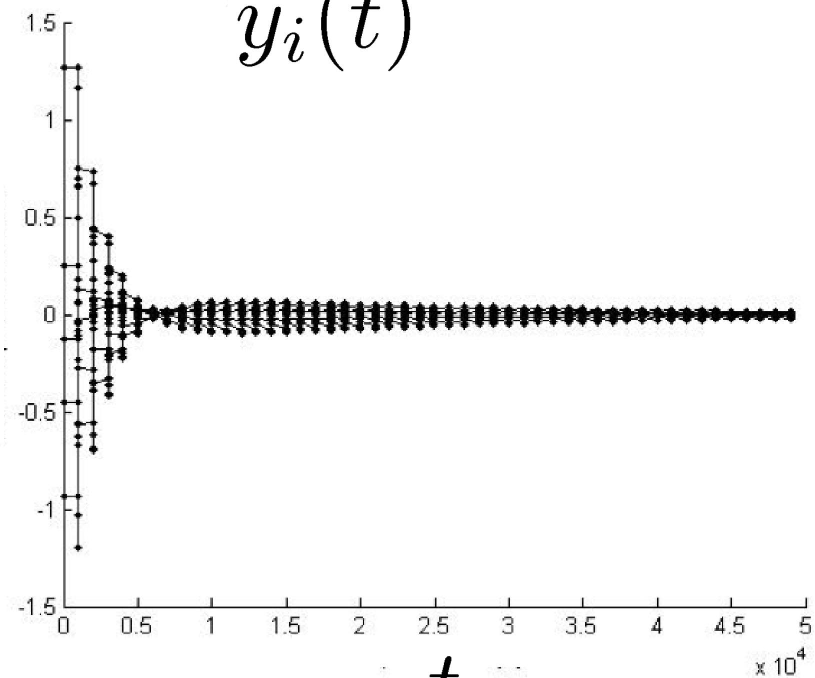
Model parameters based on experimental data from real WSN and pseudo-synchronous implementation

$w_i(t)$



t

$y_i(t)$



t

$$y_i(\infty) = 0$$

$$w_i(\infty) + d_i = w_j(\infty) + d_j$$

Outline

- Motivations and target applications
- Overview of consensus algorithms
- Application of consensus to WSN:
  - Sensor calibration
  - Least-square parameter identification
  - Time-synchronization
- Open problems
  - Identification
  - Estimation
  - Control

Consensus applications

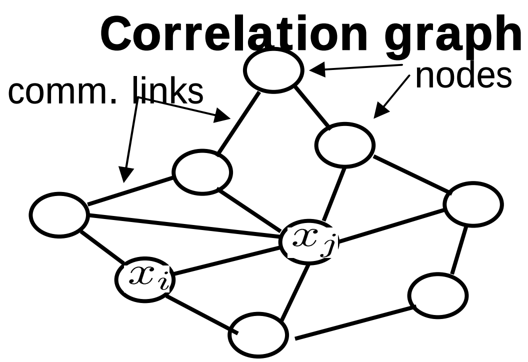
Which problem can be casted as a consensus problem ?

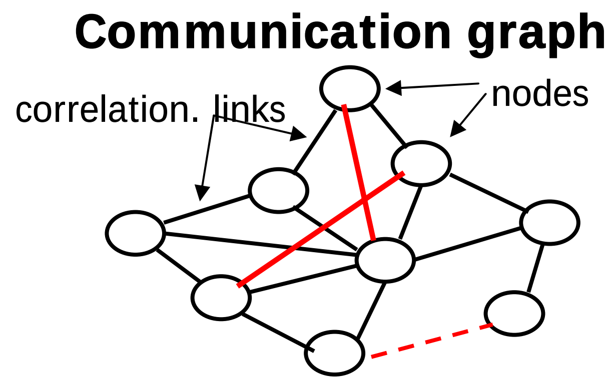
- Kalman filtering
- unbiased broadcast communication
- .....

Consensus algorithms == optimization tool

Identification: large scale structured systems

$$x \sim \mathcal{N}(0, \Sigma), \quad \Sigma^{-1} \text{ is sparse (graph model)}$$

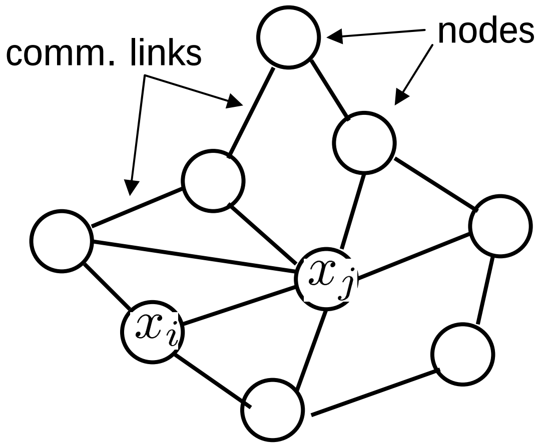


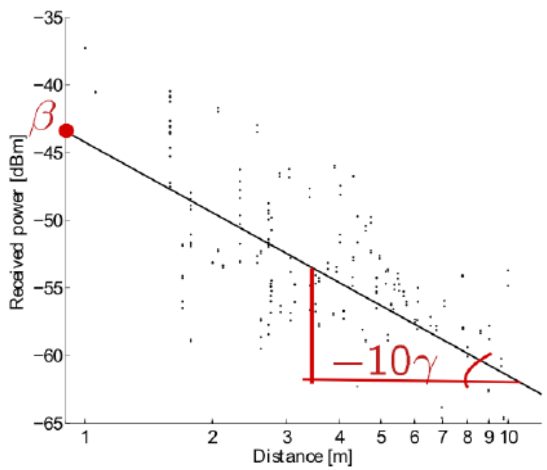


- Σ only partially known and noisy $\Rightarrow \Sigma^{-1}$ is full.
- communication graph $\neq$ correlation graph
- weak correlation, i.e. Σ^{-1} full w/ some small entries $\Rightarrow$ Graph identifiability
- what if dynamics also, i.e. $x_{t+1} = Ax_t + w_t$?
- if a node dies, i.e. remove row-column from Σ, how to compute Σ^{-1} ?
- how to do model reduction preserving graph structure ?
- is consensus relevant ?

Carlos Carvalho " *Structure and Sparsity in High-Dimensional Multivariate Analysis*", Ph.D. Theis, Duke Univ., 2007
A. P. Dempster, *Covariance Selection*, Biometrics, Vol. 28, No. 1, Special Multivariate Issue (Mar., 1972), pp. 157-175

Non-parametric estimation





Identification/Estimation of infinite dimensional space $f : R^n \rightarrow R$.

Centralized learning: $\hat{f}(\cdot) = \sum_{n=1}^{N} \alpha_n \Phi(x_n, \cdot)$

Totally decentralized learning: $\hat{f}_i(\cdot) = \sum_{n=1}^{N_i} \alpha_n^i \Phi(x_n, \cdot), \quad \mathbf{N_i << N}$

kernel

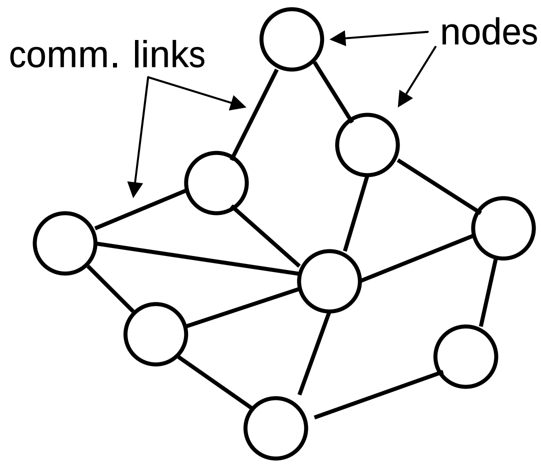
What to exchange ?

- all $(x_i, f(x_i))$ of neighbors ?
- most informative $(x_i, f(x_i))$ of neighbors ?
- smoothed observation of neighbors $(x_i, \hat{f}_i(x_i))$
- virtual observations $\hat{f}(\hat{x}_i, \hat{f}_i(\hat{x}_i))$

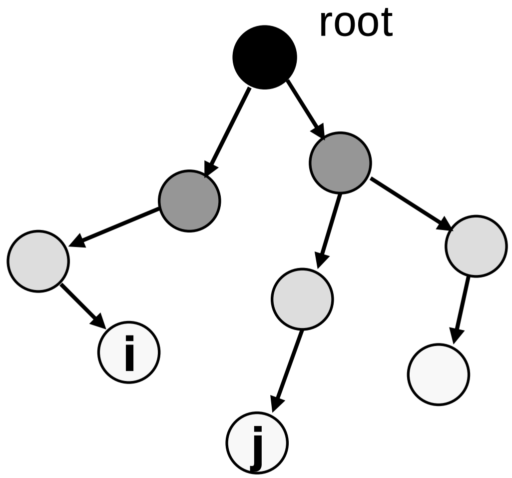
Predd, J.B.; Kulkarni, S.B.; Poor, H.V. "*Distributed learning in wireless sensor networks*" Signal Proce Magazine,2006

Soft Hierarchical Control

Time synchronization example:



P_{dist} symmetric:
slow convergence but robust



P_{hier} asymmetric:
fast convergence but fragile to node failure

$$P_{soft} = \alpha P_{dist} + (1-\alpha) P_{hier}, \qquad \text{optimal } \alpha \text{ depends on failure rate}$$

Q&A

THANK YOU