Lecture: Distributed Algorithms for Gossiping

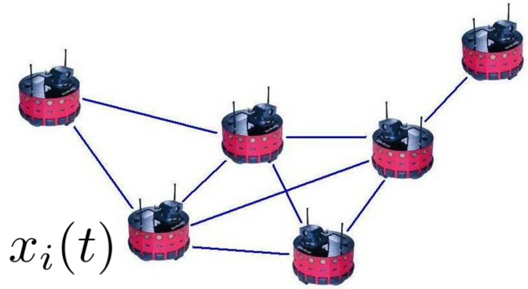
Shaoshuai Mou

PURDUE UNIVERSITY

Review

agent's dynamics: $x_i(t+1) = u_i$

control input: $u_i = f_i(x_j(t), j \in \mathcal{N}_i)$



$x_i(t)$

distributed — Each agent's input can only be local information/measurements from its neighbors.

❑ Distributed Consensus

- **Objective:** $x_1(t) = x_2(t) = \cdots = x_m(t) = x^*$ ***just a constant***

- **Update:**

$$x_i(t+1) = \sum_{j \in \mathcal{N}_i} w_{ij} x_j(t)$$

$$w_{ij} = \begin{cases} > 0, & j \in \mathcal{N}_i \\ 0, & \text{otherwise.} \end{cases} \qquad \sum_{j=1}^{m} w_{ij} = 1$$

Local Weighted Average (convex combination)

w_{ij} : the weight assigned by agent *i* to agent *j*

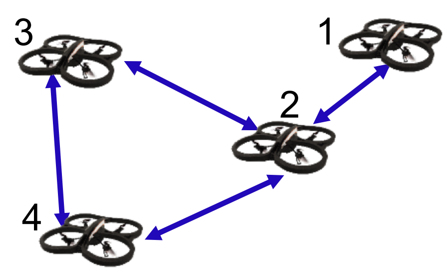
- **Example:**

$$x_1(t+1) = \frac{1}{2}x_1(t) + \frac{1}{2}x_2(t)$$

$$x_2(t+1) = \frac{1}{4}x_1(t) + \frac{1}{4}x_2(t) + \frac{1}{4}x_3(t) + \frac{1}{4}x_4(t)$$

$$x_3(t+1) = \frac{1}{3}x_2(t) + \frac{1}{3}x_3(t) + \frac{1}{3}x_4(t)$$

$$x_4(t+1) = \frac{1}{3}x_2(t) + \frac{1}{3}x_3(t) + \frac{1}{3}x_4(t)$$



❑ Consensus to the Global Average

- **Objective:** $x_1(t) = x_2(t) = \cdots = x_m(t) = x^*$ $\quad x^* = \frac{1}{m}\sum_{i=1}^{m} x_i(0) = \frac{1}{m}\mathbf{1}'x(0)$ $\quad \mathbf{1} = \begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$

  consensus — to a specific value (the global average)

- **Update:** *Local weighted average* — Metropolis Weights

$$w_{ij} = \begin{cases} > 0, & j \in \mathcal{N}_i \\ 0, & \text{otherwise.} \end{cases}$$

$$\sum_{j=1}^{m} w_{ij} = 1$$

$$x_i(t+1) = \sum_{j \in \mathcal{N}_i} w_{ij} x_j(t) \qquad w_{ij} = \begin{cases} \min\{\frac{1}{d_i}, \frac{1}{d_j}\} & j \in \mathcal{N}_i,\ j \neq i; \\ 1 - \sum_{j \in \mathcal{N}_i,\ j \neq i} w_{ij} & j = i \end{cases}$$

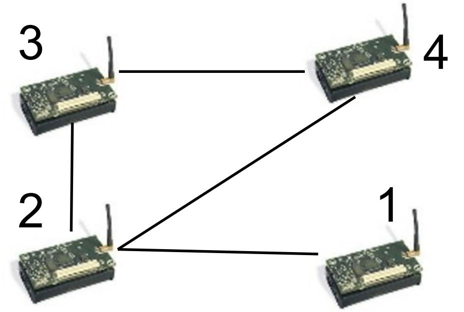
- **Example:**

$$x_1(t+1) = \frac{3}{4}x_1(t) + \frac{1}{4}x_2(t)$$

$$x_2(t+1) = \frac{1}{4}x_1(t) + \frac{1}{4}x_2(t) + \frac{1}{4}x_3(t) + \frac{1}{4}x_4(t)$$

$$x_3(t) = \frac{1}{4}x_2(t) + \frac{5}{12}x_3(t) + \frac{1}{3}x_4(t)$$

$$x_4(t) = \frac{1}{4}x_2(t) + \frac{1}{3}x_3(t) + \frac{5}{12}x_4(t)$$



Any other way for distributed averaging?

Gossiping

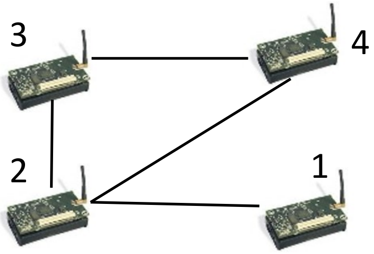
bi-directional communications & coordination

Definition: One pair of agents *i* and *j* ***gossip*** at time *t by simultaneously updating their states to be the average, namely*

$$x_i(t+1) = x_j(t+1) = \frac{1}{2}(x_i(t) + x_j(t))$$

All the other nodes keep unchanged:

$$x_k(t+1) = x_k(t), k \neq i, k \neq j$$



- *Example:* In the four-agent network, write out all agents' state update when 2 and 4 gossip.

$$\left.\begin{aligned}
x_1(t+1) &= x_1(t) \\
x_2(t+1) &= \frac{1}{2}x_2(t) + \frac{1}{2}x_4(t) \\
x_3(t+1) &= x_3(t) \\
x_4(t+1) &= \frac{1}{2}x_2(t) + \frac{1}{2}x_4(t)
\end{aligned}\right\} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

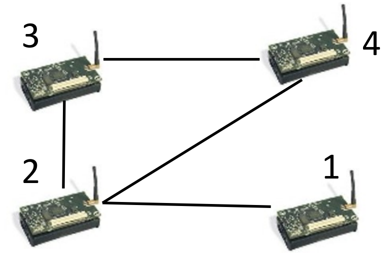
compact form

Try by yourself.

Gossiping Matrix

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

- ***Gossiping Matrix***: Agent i and j gossip in an m-agent network, its corresponding state update matrix $M \in \mathbb{R}^{m \times m}$ is such that
  - $M_{ii} = M_{ij} = M_{ji} = M_{jj} = \frac{1}{2}$
  - all the other diagonal elements are equal to 1.
  - all the other off-diagonal elements are equal to 0.



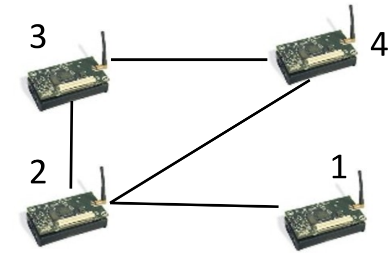
2 and 4 gossip (2,4)

Try by yourself.

1 and 2 gossip (1,2)

$$M_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix} \qquad M_{12} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Periodic Gossiping



- Periodic Gossiping Sequence: (1,2), (3,4), (2,3), (2,4) (1,2), (3,4), (2,3), (2,4), ...

In one period, all edges are covered, and the network is connected

	1	2	3	4	
$t=0$	$x_1(0)$	$x_2(0)$	$x_3(0)$	$x_4(0)$	
$t=1$	$x_1(1) = \overset{\text{average}}{\frac{x_1(0)+x_2(0)}{2}}$	$= x_2(1)$	$x_3(1) = x_3(0)$	$x_4(1) = x_4(0)$	1 and 2 gossip
$t=2$	1 unchanged	1 unchanged.	3 and 4 take average		3 and 4 gossip
$t=3$	1 unchanged.	2 and 3 take average		4 unchanged.	2 and 3 gossip
$t=4$	1 unchanged	2 and 4 take average	3 unchanged.		2 and 4 gossip

$$x(2) = M_{34}x(1)$$

$$M_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

$$x(1) = M_{12}x(0)$$

$$M_{12} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$x(4) = M_{24}x(3)$$

$$M_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix}$$

$$x(3) = M_{23}x(2)$$

$$M_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- The update within one gossip period: (1,2), (3,4), (2,3), (2,4)

$$x(4) = M_{24}x(3) = M_{24}M_{23}x(2) = M_{24}M_{23}M_{34}x(1) = M_{24}M_{23}M_{34}M_{12}x(0)$$

$$A$$

$$x(4(t+1)) = Ax(4t)$$

$$= \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/8 & 1/8 & 3/8 & 3/8 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/8 & 1/8 & 3/8 & 3/8 \end{bmatrix}$$

Do all $x_i(t)$ converge to reach a consensus, which is the global average?

*Does $x(t)$ **converge** to be constant?*

*At the convergence point, is a **consensus** reached?*

Is the consensus value the global average?

$$x_1 = x_2 = x_3 = x_4 = x^*$$

$$x^* = \frac{1}{m}\sum_{i=1}^{m} x_i(0)$$

$$x(4t) = A^t x(0)$$

$$x(t) \to x^* \mathbf{1}$$

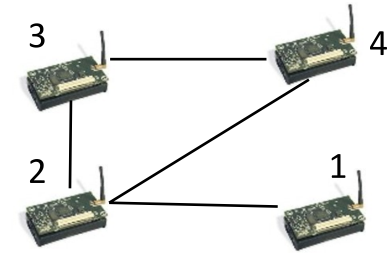
$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$x^*$$

$$\lim_{t\to\infty} x(4t) = \lim_{t\to\infty} A^t x(0) = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix} x(0) = \mathbf{1} \cdot \frac{1}{4} \cdot \mathbf{1}' x(0)$$

$$\underbrace{A^t \to A^*}_{\text{converge}} = \underbrace{\frac{1}{m}\mathbf{1}\mathbf{1}'}_{\text{Consensus+ Global Average}}$$

Some Concepts Related to Gossiping



- **Periodic Gossiping:**

  Is this necessary in practice?

  - At each time, there is **only one pair** of agents to perform gossip.
  - *In one period, **all edges are covered**, and the network is connected*

  (1,2), (3,4), (2,3), (2,4), (1,2), (3,4), (2,3), (2,4), (1,2), (3,4), (2,3), (2,4), ...

  (+): Convergence rate can be tuned and controlled.

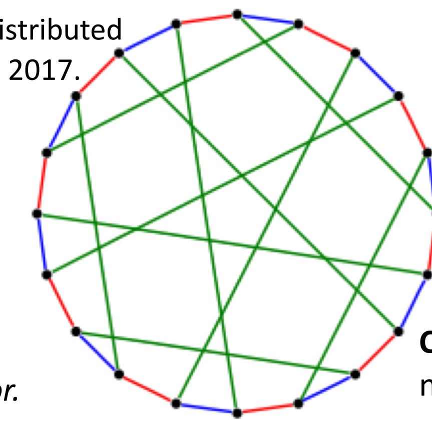
❖ **Related Paper:** J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, C. Yu. Deterministic Gossiping. Proceedings of IEEE, 2011.

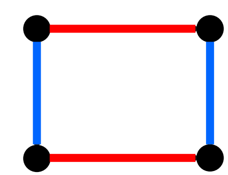
- **Multi-Gossiping:** (1,2), (3,4), (2,3), (2,4) *One period requires four time-steps.* 3

  Disjoint gossiping pairs could be implemented simultaneously. (+): Faster convergence.

❖ **Related Paper:** C. Yu, B. D. O. Anderson, S. Mou, J. Liu, F. He, A. S. Morse. Distributed Averaging using periodic gossiping. IEEE Transactions on Automatic Control, 2017.

In order to achieve **faster convergence**, one may enable **disjoint pairs** of agents to gossip at one-time step as many as possible.





How to find these disjoint gossip pairs?

❖ **Edge Coloring** Problem in Graph Theory: *Assign colors to edges of a graph such that no two incident edges have the same color.*

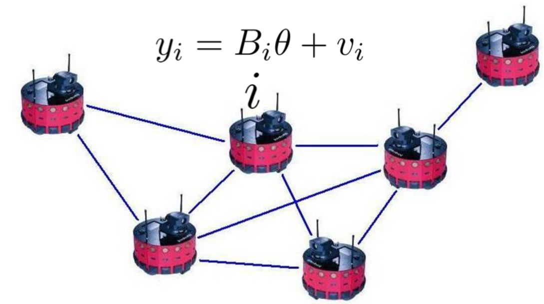
Chromatic index: minimum number of colors needed.

(-): Periodic gossiping requires all agents to be synchronized to determine the gossip order in one period.

- **Randomized Gossiping: Agents have their own clocks, wake up randomly and choose one neighbor to gossip.**

S. Boyd, A. Ghosh, B. Prabhakar, D. Shah. Randomized Gossip Algorithms. IEEE Transactions on Information Theory, 2006.

Application of Distributed Averaging in **Distributed Estimations**



- Utilize a multi-agent network to achieve an important parameter vector θ , which is not directly observable/available.

- **Local Measurement**: Each agent *i* observes/measures a linear combination

$$y_i = \underbrace{B_i}_{\text{known to } i}\theta + \underbrace{v_i}_{\text{noise}}$$

$v_1, v_2, \ldots, v_m$ are independent jointly-Gaussian.

$E[v_i] = 0 \qquad E[v_i v_i'] = \Lambda_i = \Lambda_i'$

- **Global Goal**: Achieve a *nice estimate* to θ

$\hat{\theta}^*$ minimizes the following objective function

convex

$$F(\hat{\theta}) = \sum_{i=1}^{n} \underbrace{(y_i - B_i\hat{\theta})'}_{\text{estimation error}}\Lambda_i^{-1}(y_i - B_i\hat{\theta})$$

quadratic form $x'Ax$

each agent's estimation error is weighted by Λ_i^{-1}

Accurate (inaccurate) measurements are with high (low) weights.

$$\frac{\partial F}{\partial \theta}\Big|_{\hat{\theta}=\hat{\theta}^*} = 0$$

$$\sum_{i=1}^{n} B_i'\Lambda_i^{-1}(y_i - B_i\hat{\theta}^*) = 0$$

$$\sum_{i=1}^{n} \underbrace{B_i'\Lambda_i^{-1}B_i}_{A_i}\hat{\theta}^* = \sum_{i=1}^{n} \underbrace{B_i'\Lambda_i^{-1}y_i}_{b_i}$$

is **available** to each agent *i* since it knows B_i, Λ_i, y_i

A_i is positive definite

Problem 1: Find $\hat{\theta}^*$ such that

$$(A_1 + A_2 + \cdots + A_m)\hat{\theta}^* = (b_1 + b_2 + \cdots + b_m)$$

in a multi-agent system, where each agent *i* knows A_i, b_i

A_i is positive definite

- *Method 1:* Distributed Algorithm for Solving Linear Equation.

J. Lu, C. Y. Tang. A Distributed Algorithm for Solving Positive Definite Linear Equations Over Networks With Membership Dynamics, IEEE Transactions on Control of Network Systems, 2018.

$$\hat{\theta}^* = \Big(\overbrace{\frac{1}{n}\sum_{i=1}^{n} A_i}^{X^*}\Big)^{-1} \Big(\overbrace{\frac{1}{n}\sum_{i=1}^{n} b_i}^{Z^*}\Big)$$

global average

Problem 2: Let each agent *i* controls two state variables $X_i(t), Z_i(t)$

with initialization $X_i(0) = A_i, Z_i(0) = b_i$

Develop a distributed update for each agent to separately update $X_i(t), Z_i(t)$ such that $X_i(t) \to X^*, \quad Z_i(t) \to Z^*$ Then one has $\hat{\theta}^* = (X^*)^{-1} Z^*$

- *Method 2:* Distributed Consensus Averaging Algorithms.

Summary

- ***Gossiping:*** One pair of agents *i* and *j* ***gossip*** at time *t by simultaneously updating their states to be the average, and* all the other nodes keep unchanged.

$$x_i(t+1) = x_j(t+1) = \frac{1}{2}(x_i(t) + x_j(t))$$

$$x_k(t+1) = x_k(t), k \neq i, k \neq j$$

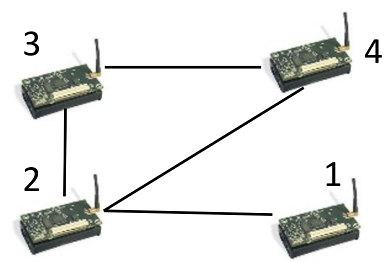
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix}$$

Its corresponding *gossiping matrix* is

- $M_{ii} = M_{ij} = M_{ji} = M_{jj} = \frac{1}{2}$
- all the other diagonal elements are equal to 1.
- all the other off-diagonal elements are equal to 0.

- **Periodic Gossiping Sequence:**

(1,2), (2,3), (3,4), (2,4), (1,2), (2,3), (3,4), (2,4), ...



In one period, all edges are covered, and the network is connected.

$$x(4(k+1)) = Ax(4k) \qquad A = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/8 & 1/8 & 1/4 & 1/2 \\ 3/16 & 3/16 & 3/8 & 1/4 \\ 3/16 & 3/16 & 3/8 & 1/4 \end{bmatrix}$$