# Distributed Averaging in Dynamic Networks

Bharath B. N.

Signal Processing for Communications Lab
ECE Dept., Indian Institute of Science
Bangalore-560012

September 22, 2012

# Outline

- Static Networks
  - Algorithm to find Average
  - Algorithm to find the Minimum of values
- Dynamic Networks
  - “Dynamic Aware" Algorithm to find the Minimum of values
  - “Dynamic Aware" Algorithm to find Average
- Conclusions

- Communication network is modeled as $G := (V, E)$, $|V| = n$
- Error free links with infinite capacity
- Connected, i.e., for all $u, v \in V$, $d(u, v) < \infty$
- Each node observes a *fixed value*, i.e., $X_v$, $v \in V$
- Time is discrete $t = 1, 2, \ldots$
- Communication is instantaneous (See the white board)

## Goal

- Find $f(X_1, X_2, \ldots, X_n)$ at each node
- Examples
  - Average: $f(X_1, X_2, \ldots, X_n) := \frac{1}{n}\sum_{i=1}^{n} X_i$
  - Maximum: $f(X_1, X_2, \ldots, X_n) := \max\{X_1, X_2, \ldots, X_n\}$
  - Minimum: $f(X_1, X_2, \ldots, X_n) := \min\{X_1, X_2, \ldots, X_n\}$
  - Majority: $f(X_1, X_2, \ldots, X_n) := \text{Majority}\{X_1, X_2, \ldots, X_n\}$, $X_i \in \{0, 1\}$

- Find an ALG. that enables $v \in V$ to compute $\hat{f}(X_1, X_2, \ldots, X_n)$ as an estimate of $f(X_1, X_2, \ldots, X_n)$
- ALG should have the following properties:
  - Distributed
  - Should not scale with $|V| = n$
  - (Accurate) For small $\epsilon > 0$, $\left|\hat{f}(X_1, X_2, \ldots, X_n) - f(X_1, X_2, \ldots, X_n)\right| < \epsilon$, for all $v \in V$
  - Low complexity
- What is the best ALG. in terms of the rate of convergence?

- Observation: $X_v, v \in V$
- Find $f(X_1, X_2, \ldots, X_n) := \frac{1}{n}\sum_{i=1}^{n} X_i$

**ALG. for finding the average**

- $\hat{X}_v(t+1) := \sum_{u \in V} A_{uv}\hat{X}_u(t)$, for all $v \in V$, and $t = 1, 2, \ldots$
- In matrix form, $\hat{X}(t+1) := A\hat{X}(t)$

- Goal: Find the matrix $A$ that results in a "good" estimate of the average
- Observation: Distributed ALG. is possible if $A_{uv} = 0$ for all $(u, v) \notin E$

- Consider the matrix with the following properties:

**Properties of the matrix $A$**

- $A_{uv} = 0$ for all $(u, v) \notin E$
- $\sum_{u \in V} A_{uv} = 1$ for all $v \in V$
- $\sum_{v \in V} A_{uv} = 1$ for all $u \in V$
- The matrix $A$ is symmetric
- Directed graph $G(A) = (V, E(A))$ is connected where a directed edge $(u, v) \in E(A)$ iff $A_{uv} > 0$

- Fancy name: Doubly stochastic, graph $G$ conformant and irreducible matrix

**Theorem**

*Consider the update rule*

$$\hat{X}(t+1) := A\hat{X}(t),$$

*where A is a doubly stochastic, graph G conformant and irreducible matrix. The above eventually reaches the true average, i.e.,* $\hat{X}(t) \to X_{\text{avg}}\mathbf{1}$ *as* $t \to \infty$*, where* $X_{\text{avg}} := \frac{1}{n}\sum_{v \in V} X_v$

*Proof*: First, we make the following observations:

- $X(0) := [X_1, X_2, \ldots, X_n]$
- $A$ is a symmetric matrix
- $\mathbf{1}$ is an eigenvalue of $A$, i.e., $A\frac{1}{n}\mathbf{1} = \frac{1}{n}\mathbf{1}$
- $A$ has eigenvalues $1 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq -1$

Note that $\hat{X}(t+1) := A\hat{X}(t) \Rightarrow \hat{X}(t+1) = A^{t+1}X(0)$

- Let $v_1, v_2, \ldots, v_n$ be the eigenvectors of $A$
- Need to prove $\hat{X}(t+1) = A^{t+1}\hat{X}(0) \rightarrow X_{\text{avg}}\mathbf{1}$
- $X(0) = \sum_{i=1}^{n} \langle X(0), v_i \rangle v_i$ implies

$$
\begin{aligned}
A^{t+1}\sum_{i=1}^{n} \langle X(0), v_i \rangle v_i &= \sum_{i=1}^{n} \langle X(0), v_i \rangle \lambda_i A^t v_i \\
&= \sum_{i=1}^{n} \langle X(0), v_i \rangle \lambda_i^2 A^{t-1} v_i \\
&\star \\
&\star \\
&\star \\
&= \sum_{i=1}^{n} \langle X(0), v_i \rangle \lambda_i^{t+1} v_i
\end{aligned}
$$

## Main Result: Proof Cont.

- As $t \to \infty$

$$\sum_{i=1}^{n} \langle X(0), v_i \rangle \lambda_i^{t+1} v_i \to \langle X(0), v_1 \rangle v_1 \qquad (1)$$

  since $\lambda_i < 1$ for $i > 1$
- But $v_1 = \frac{1}{n}\mathbf{1}$, and $\langle X(0), v_1 \rangle = X_{\mathrm{avg}}$ $\square$

- How fast does the ALG. converge to the average?
- Depends on how fast $A^t(u,:)$ goes to **1** for all $u \in V$

## Mixing Time

- The following quantifies the speed of convergence at time $t$:

$$\tau_\epsilon(u) = \inf\left\{n \geq t : \frac{1}{2}\sum_{v \in V} |A(u,v)^n - 1| < \epsilon\right\} \tag{2}$$

- Bound $\tau_\epsilon(t)$ to get an estimate of "the speed of convergence" (Not now!)

- How do we compute the minimum of all the values in a network in a distributed way?

**ALG. to find the minimum**

ALGORITHM updates the value at each node as

$$\hat{X}_v(t) = \min\{X_v(t), \min_{u \in N(v)} \hat{X}_u(t-1)\}, \tag{3}$$

where the neighbor is defined as $N(v) := \{u \in V : (u, v) \in E\}$.

- Requires at least $D := \max_{u,v \in V} d(u, v)$ units of time to compute the minimum in a network

## Different types of dynamics

- Dynamically varying topology
  - The link changes dynamically
  - The nodes change dynamically (sleep and awake)
- The observations change (focus of this talk)

## In this talk

- The observations change according to the following rule:
  - Communication network is modeled as $G := (V, E)$, $|V| = n$ with Error free links and connected, i.e., for all $u, v \in V$, $d(u, v) < \infty$
  - Each node observes: $X_v(t)$, $v \in V$, and $t = 1, 2, \ldots$
- Note that the observation is changing!

**Bounded Additive Change (BAC)**

Given a fixed $\delta > 0$, and for any $v \in V$, $|X_v(t+1) - X_v(t)| \leq \delta$ for any $t \geq 0$

**Bounded Multiplicative Change (BMC)**

Given a fixed $\delta > 0$, and for any $v \in V$, $e^{-\delta} \leq \frac{X_v(t+1)}{X_v(t)} \leq e^{\delta}$ for any $t \geq 0$

- The following update rule fails

$$\hat{X}_v(t) = \min\left\{X_v(t), \min_{u \in N(v)} \hat{X}_u(t-1)\right\}, \quad (4)$$

when $\hat{X}_v(t) = 0$ for some $t \geq 0$!

- Remedy: Forget the past

$$\hat{X}_v(t) = \min\left\{X_v(t), \min_{u \in N(v)} \hat{X}_u(t-1)e^{\delta}\right\}, \quad (5)$$

- The goal is to show that the above update rule is “good"

**Theorem**

*Consider the update rule*

$$\hat{X}_v(t) = \min\left\{X_v(t), \min_{u \in N(v)} \hat{X}_u(t-1)e^{\delta}\right\},$$

*then, for any $t \geq D$,*

$$\left(\min_{u \in V} X_u(t)\right) \leq \hat{X}_v(t) \leq e^{2D\delta}\left(\min_{u \in V} X_u(t)\right)$$

## Simplification

$$\hat{X}_v(t) = \min\left\{X_v(t), \min_{u \in N(v)} \hat{X}_u(t-1)e^{\delta}\right\}$$

$$= \min\left\{X_v(t), \min_{u \in N(v)} X_u(t-1)e^{\delta}, \min_{w \in N_2(v)} \hat{X}_w(t-2)e^{2\delta}\right\}$$

$\star$

$\star$

$\star$

$$= \min\left\{\min_{k=0}^{m} \min_{u \in N_k(v)} \{X_u(t-k)e^{k\delta}\}, \min_{w \in N_{m+1}(v)} \hat{X}_w(t-m-1)e^{(m+1)\delta}\right\}$$

believe me!

## Simplification

- As $m \to \infty$, we have

$$\hat{X}_v(t) = \min_{k \geq 0} \min_{u \in N_k(v)} \{X_u(t-k)e^{k\delta}\}$$
$$= \min_{u \in V} \min_{k \geq d(u,v)} \{X_u(t-k)e^{k\delta}\}$$
$$= \min_{u \in V} \left\{X_u(t-d(u,v))e^{d(u,v)\delta}\right\}$$

since $\hat{X}_v(s) := \infty$ for all $v \in V$, and $s < 0$

- BMC results in $X_u(t) \leq X_u(t-d(u,v))e^{d(u,v)\delta}$, which implies

$$\left(\min_{u \in V} X_u(t)\right) \leq \hat{X}_v(t)$$

## Proof Cont.

### Simplification

- Now, we prove $\hat{X}_v(t) \leq e^{2D\delta}\left(\min_{u\in V} X_u(t)\right)$

$$\begin{aligned}\hat{X}_v(t) &= \min_{u\in V}\left\{X_u(t-d(u,v))e^{d(u,v)\delta}\right\}\\ &\leq \min_{u\in V}\left\{X_u(t)e^{2d(u,v)\delta}\right\} \quad \text{for } t \geq d(u,v)\\ &\leq e^{2D\delta}\min_{u\in V} X_u(t) \quad \square\end{aligned}$$

## Why Minimum?

- It turns out that the average can be computed using the minimum
- $Y_i \sim \exp\{X_i\}$, independent rv's, then

$$Y := \min\{Y_1, Y_2, \ldots, Y_n\} \sim \exp\left\{\sum_{j=1}^{n} X_i\right\}$$

- This gives us the clue to find the average

## ALG. for computing the average under BMC

Given $p \in (0,1)$ and $\epsilon \in (0, 0.35)$, let $m = \lceil \frac{3\ln(2/p)}{\epsilon^2} \rceil$. For each $v \in V$ and $t \geq 0$, define $Y_{v,i}$ with $1 \leq i \leq m$ as follows:

- for $t = 0$, generate $Y_{v,i}(0) \sim \exp\{X_v(0)\}$ independent of all other values
- for $t \geq 1$,

$$Y_{v,i}(t+1) := \frac{X_v(t)}{X_v(t+1)} Y_{v,i}(t)$$

Update rule:

$$\hat{Y}_{v,i}(t) := \min\left\{ Y_{v,i}(t), \min_{u \in N(v)} \hat{Y}_{u,i}(t) e^{m\delta} \right\},$$

and the estimate at node $v$ is given by $\hat{X}_v(t) = \frac{e^{m(D+1)\delta}}{\hat{Y}_v(t)}$, where $\hat{Y}_v(t) := \frac{1}{m} \sum_{i=1}^{m} \hat{Y}_{v,i}(t)$

**Theorem**

*For the algorithm described above, we have*

$$(1-\epsilon)e^{-m(D+1)\delta} \leq \frac{\hat{X}_v(t)}{\sum_{v\in V} X_v(t)} \leq (1+\epsilon)e^{m(D+1)\delta}$$

*for all* $t \geq mD$

## Beyond this Talk

- Proof of the sufficient conditions under BMC
- Lower bounds (Necessary conditions) under BMC
- Some analysis for the BAC
- Example networks

## Summary

- “Dynamic aware" distributed ALG. for estimating the average of values which is nearly optimal
- The results captured tradeoff between the dynamic range and the accuracy

## Criticism

- Communication links are assumed to be perfect
- Bounded variations
- Links are assumed to be fixed
- Node dynamics are not accounted for

## References

S. Rajagopalan and D. Shah. “Distributed Averaging in Dynamic Networks",IEEE Journal of Selected Topics in Signal Processing, 2011.