Randomized consensus algorithms over large scale networks

ITA Workshop 2007

February 1, 2007

Fabio Fagnani

Dipartimento di Matematica, Politecnico di Torino, ITALY

joint work with

Sandro Zampieri

Department of Information Engineering, Università di Padova, ITALY

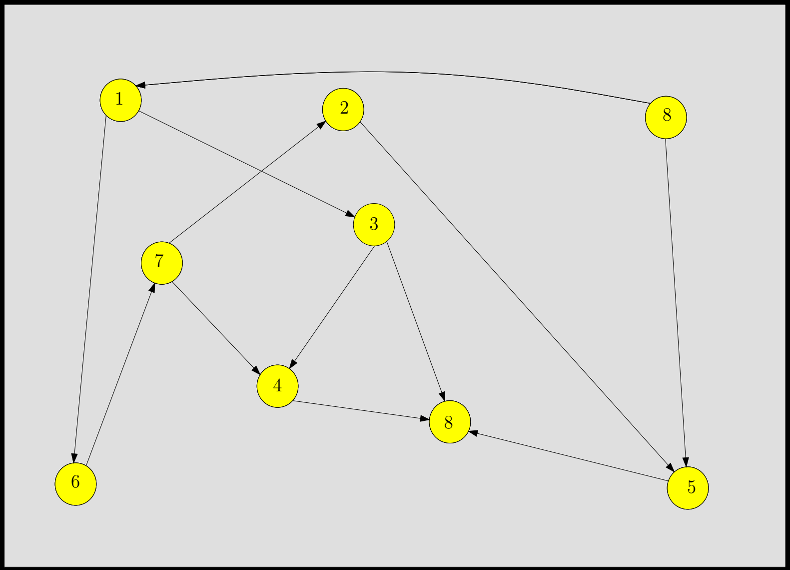
Outline

- Problem formulation and motivation
- Deterministic consensus algorithms
- Random consensus algorithms
- Mean square analysis, examples
- Concentration results
- Conclusions

Problem formulation

$\mathcal{G} = (V, E)$ directed graph.
$V = \{1, \ldots, N\}$

$\forall i \in V$, a measure $x_i \in \mathbb{R}$



GOAL:
compute $X_A := N^{-1} \sum x_i$,
iteratively, exchanging information along available edges.
in a decentralized fashion

APPLICATIONS:

- Load balancing in computer networks
- Data fusion in sensor networks
- Coordination of multi-agent systems

Consensus algorithms I

$P(t)$ $N \times N$ stochastic matrix. $(P(t)_{ij} \geq 0, P(t)\mathbb{1} = \mathbb{1})$

$x(t)_i$ estimation of x_A by agent i at time t. $x(t) \in \mathbb{R}^N$

$$x(t+1) = P(t)x(t)\,, \qquad x(0)_i = x_i$$

$x(t) = Q(t)x(0)$, $Q(t) = \prod_{s=0}^{t-1} P(s)$,

- CONSENSUS: $Q(t) \to \mathbb{1}\rho^T \Rightarrow x(t) \to \mathbb{1}\rho^T x(0)$
- AVERAGE CONSENSUS: $\rho = N^{-1}\mathbb{1}$
- $P(t)$ ADAPTED to $\mathcal{G}$: $P(t)_{ij} > 0 \Rightarrow (j, i) \in E$.

Consensus algorithms II

$\mathcal{G} = (V, E)$ strongly connected

P stochastic $P_{ij} > 0 \;\Leftrightarrow\; (j, i) \in E$, $P_{ii} > 0 \;\forall i$

- $P(t) = P$ achieves consensus
- If $\mathbb{1}^T P = \mathbb{1}^T$ then achieves av. consensus

A possible construction:

$A_{\mathcal{G}} \in \{0, 1\}^{N \times N}$: $(A_{\mathcal{G}})_{ij} = 1 \Leftrightarrow (j, i) \in E$

$D_{\mathcal{G}} = \mathrm{diag}(\nu_1, \dots \nu_N), \quad \nu_j = |\{i | (i, j) \in E\}|$

- $P = kI + (1 - k) D_{\mathcal{G}}^{-1} A_{\mathcal{G}}$ achieves consensus
- $\mathcal{G}$ undirected $\Rightarrow$ P symmetric $\Rightarrow$ av. consensus

Tsitsiklis, Cybenko, Morse, Olfati-Saber, Murray, Francis,...

Randomized algorithms

$P(t)$ chosen randomly at every time step

- PROBABILISTIC CONSENSUS:
  $Q(t) \longrightarrow \mathbb{1}\rho^T$ a.s. $\quad \rho$ random stochastic vector

Mean evolution: $\overline{P} = \overline{P(t)}, \ \overline{x(t)} = \overline{P}^t x(0)$.

THEOREM: (Cogburn 1987)
$P(t)$ achieves prob. cons. $\Leftrightarrow$ $\overline{P}$ achieves cons.

Example I

SYMMETRIC GOSSIP (Boyd et al. AC-2006)

$\mathcal{G}$ undirected, strongly connected

At every time t:

- Choose $i \in V$ randomly, and a neighbor j of i randomly
- i and j exchange their current estimations
- $x(t+1)_i = kx(t)_i + (1-k)x(t)_j$
- $x(t+1)_j = kx(t)_j + (1-k)x(t)_i$

$$\overline{P}_{ij} = \frac{1}{N}\left[\frac{1}{\nu_i} + \frac{1}{\nu_j}\right]k, \quad (i,j) \in E$$

$$\Downarrow$$

Average probabilistic consensus

Example II

IN-GOSSIP $\mathcal{G}$ strongly connected

At every time t:

- Every $i \in V$ chooses an in-neighbor j randomly
- i receives the estimation of j
- $x(t+1)_i = kx(t)_i + (1-k)x(t)_j$

$$\overline{P} = (1-k)I + kD_{\mathcal{G}}^{-1}A_{\mathcal{G}}$$

$$\Downarrow$$

Probabilistic consensus

Average is not preserved in general!

Examples III

BROADCASTING $\mathcal{G}$ strongly connected

At every time t:

- A node $i \in V$ is chosen randomly
- i sends its estimation to all its out-neighbor
- $x(t+1)_j = kx(t)_j + (1-k)x(t)_i$, if $(i,j) \in E$

$$\overline{P} = I + \frac{k}{N}[A_{\mathcal{G}} - D_{\mathcal{G}}]$$

$$\Downarrow$$

Probabilistic consensus

Average is not preserved in general!

Complexity and performance

WHY RANDOMNESS?

- Computation and transmission consume energy. Keep them low!
- In many applications, an agent can only receive data from just one neighbor.
- Under these limitations, random algorithms perform better

Complexity and performance

WHY RANDOMNESS?

- Computation and transmission consume energy. Keep them low!
- In many applications, an agent can only receive data from just one neighbor.
- Under these limitations, random algorithms perform better

WHY ASYMMETRIC SCHEMES?

- Transmission links can be quite asymmetric (sensor networks)
- In many applications it is sufficient to get close to the average

Mean square analysis

Two performance figures:

- $R = \limsup_{t \to +\infty} (\mathbb{E}||x(t) - \mathbb{1}x_A(t)||^2)^{1/t}$
- $\delta = \mathbb{E}[|x_A(\infty) - x_A(0)|^2]$

We can write

$$\bar{d}(t) := \mathbb{E}||x(t) - \mathbb{1}x_A(t)||^2 = x^*(0)\Delta(t)x(0)$$

where

$$\Delta(t+1) = \mathcal{L}(\Delta(t)) := \mathbb{E}[P(0)^*\Delta(t)P(0)],$$
$$\Delta(0) = I - N^{-1}\mathbb{1}\mathbb{1}^*$$

Mean square analysis

Assume that $Q(t) \to \mathbb{1}\rho^*$

Then,

$$\delta = x(0)^* B x(0)$$

$$B = \mathbb{E}[\rho\rho^*] - 2N^{-1}\mathbb{E}[\rho]\mathbb{1}^* + N^{-2}\mathbb{1}\mathbb{1}^*$$

REMARK:

$$\mathbb{E}[\rho]\bar{P} = \mathbb{E}[\rho] \;\text{ and }\; \mathcal{L}(\mathbb{E}[\rho\rho^*]) = \mathbb{E}[\rho\rho^*]$$

Spectral structure of $\overline{P}$ and $\mathcal{L}$ $\Rightarrow$ R and δ!

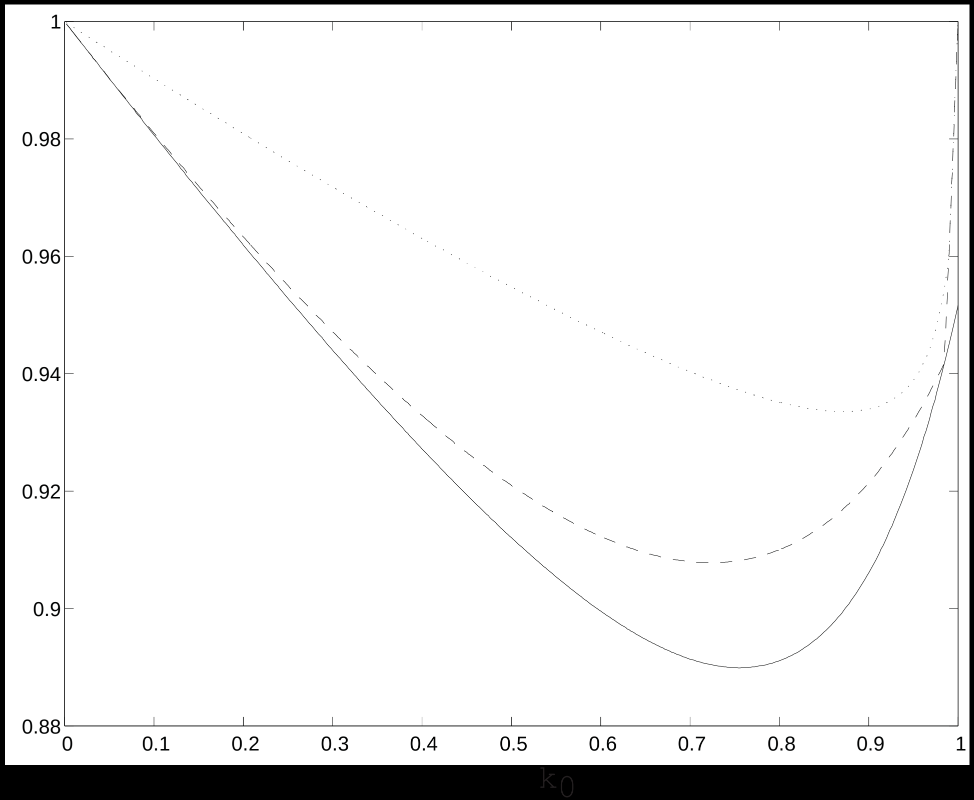
The case when $\mathcal{G}$ is complete

- SYMMETRIC GOSSIP $k = 1/2$
  $R_{sym} = 1 - \frac{1}{(N-1)}, \quad \delta_{sym} = 0$
- IN-GOSSIP $k = 1/2$
  $R_{in} = \frac{1}{2} + O(N^{-1}), \quad \delta_{in} = \left(\frac{1}{2N} + o(\frac{1}{N})\right) \overline{d}(0)$
- BROADCASTING
  $R_{br} = (1-k)^2$, $\delta_{br} = \frac{k}{2-k}\overline{d}(0)$

REMARKS:

- $\delta_{in} \to 0$ for $N \to +\infty$!
- $R_{sym}^N \to e^{-1} < R_{in}$
- Some results for Cayley graphs. $\delta_{in} \to 0$?

The cycle graph



R as a function of k: symmetric-gossip (dashed line), in-gossip (dotted line) broadcasting (solid line) for $N = 20$.

For symmetric and broadcasting, rates are powered to N.

Concentration results

Assume $\overline{P}$ achieves average consensus.

$d(t) = ||x(t) - \mathbb{1}x_A(t)||^2,$

$$\mathbb{P}[|d(t) - \overline{d}(t)| \geq \delta] \leq \exp\left(-\frac{\delta^2 \alpha_N}{K||x(0)||_\infty^4 t}\right)$$

(Symmetric: $\alpha_N = N^2$, In, Broad.: $\alpha_N = N$)

Similar results for $\beta(t) = [x_A(t) - x_A(0)]$.

Mean square analysis meaningful for $N >> t$!

Conclusions

SUMMARY:

- Random consensus algorithms which do not need symmetric communication but yield consensus near the average.
- In some cases precise analytical results
- Concentration results

FUTURE RESEARCH:

- Complete analysis for Cayley graphs
- Extend analysis to other examples (random graphs)
- Robustness to link failures, packet drops.