CONSENSUS PROBLEM IN MULTI-AGENT SYSTEMS

Asu Ozdaglar

Joint work with Angelia Nedić UIUC, USA

Joint EUROPT-OMS Conference
July 2007

Department of Electrical Engineering & Computer Science
MASSACHUSETTS INSTITUTE OF TECHNOLOGY, USA

Motivation

- Increasing interest in distributed control and coordination of networks consisting of multiple autonomous (potentially mobile) agents
- Motivated by many emerging networking applications, such as ad hoc wireless communication networks and sensor networks, characterized by:
  - Lack of centralized control and access to information
  - Time-varying connectivity
- Control algorithms deployed in such networks should be:
  - Completely distributed relying on local information
  - Robust against changes in the network topology
  - Easily implementable

Consensus Problem

- Canonical problem that appears in the coordination of multi-agent systems is the consensus problem
- Goal: Given initial values (scalar or vector) of agents, establish conditions under which through local interactions and computations, agents asymptotically agree upon a common value, i.e., reach a consensus
- Examples:
  - **Control of moving vehicles (UAVs)**: alignment of the heading angles
  - **Information processing in sensor networks**: computing averages of initial local observations (i.e., consensus on a particular value)
  - **Design of distributed optimization algorithms**: need a mechanism to align estimates of decision variables maintained by different agents/processors

Our Work

- We study consensus problem for multi-agent systems
- Main Results:
  - General distributed asynchronous computational model for reaching consensus
  - Convergence analysis and convergence rate estimates for time-varying topologies under general connectivity assumptions
- This talk has two parts:
  - Analysis with no communication delay (i.e., there is no delay in locally delivering information from one agent to another)
  - Analysis with communication delay
- Part of the work not discussed here:
  - Simultaneous optimization and consensus:
    - See talk by Nedić in EURO XXII on Monday

Related Literature

- Parallel and Distributed Algorithms:
  - General computational model for dist asynchronous optimization
    - Tsitsiklis 84, Bertsekas and Tsitsiklis 95
- Consensus and Cooperative Control:
  - Analysis of group behavior (flocking) in dynamical-biological systems
    - Vicsek 95, Reynolds 87, Toner and Tu 98
  - Mathematical models of consensus and averaging
    - Jadbabaie *et al.* 03, Olfati-Saber Murray 04, Boyd *et al.* 05

- Previous literature:
  - Focus on convergence to consensus
  - No explicit convergence rate estimates (except for specific cases; Olshevsky and Tsitsiklis 06, Cao *et al.* 06)
  - Limited focus on communication delay case

Model

- A network with m agents with node set $V = \{1, \ldots, m\}$
- Agents update and (potentially) send their information at discrete times $t_0, t_1, t_2, \ldots$
- We use $x^i(k) \in \mathbb{R}^n$ to denote information state of agent i at time t_k

Agent Update Rule:

- Agent i updates his information state by

$$x^i(k+1) = \sum_{j=1}^{m} a_j^i(k) x^j(k),$$

  where $a^i(k) = (a_1^i(k), \ldots, a_m^i(k))'$ is a vector of weights
- The vector $a^i(k)$ represents agent i's neighbor relations at slot k
- Dynamics governed by a *switched linear system*

Weights

Assumption (*Weights Rule*) For all k, we have

(a) There exists a scalar $\eta \in (0,1)$ s.t. for all $i \in \{1,\dots,m\}$,

 (i) $a_i^i(k) \geq \eta$

 (ii) $a_j^i(k) \geq \eta$ for all j communicating directly with i in (t_k, t_{k+1}).

 (iii) $a_j^i(k) = 0$ for all j otherwise.

(b) The vectors $a^i(k)$ are stochastic, i.e., $\sum_{j=1}^m a_j^i(k) = 1$ for all i.

Example: Equal neighbor weights $a_j^i(k) = \frac{1}{n_i(k)+1}$, where $n_i(k)$ is the number of agents communicating with i (his neighbors) at slot k

Information Exchange

At slot k, information exchange may be represented by a directed graph (V, E_k) where

$$E_k = \{(j,i) \mid a^i_j(k) > 0\}$$

Assumption (*Connectivity*) The graph (V, E_∞) is connected, where

$$E_\infty = \{(j,i) \mid (j,i) \in E_k \text{ for infinitely many indices } k\}.$$

- Information state of agent i influences information state of any other agent infinitely often

Assumption (*Bounded Intercomm Interval*) There is some $B \geq 1$ s.t.

$$(j,i) \in E_k \cup E_{k+1} \cup \cdots \cup E_{k+B-1} \qquad \text{for all } (j,i) \in E_\infty \;\text{ and }\; k \geq 0.$$

- Agent j send his information to neighboring agent i at least once every B consecutive time slots.

Evolution of Information States

Notation: For a matrix A, we write

$$[A]_i^j : (i,j)^{th} \text{ entry}, \quad [A]_i : i^{th} \text{ row}, \quad [A]^j : j^{th} \text{ column}$$

- Let $A(s)$ denote the matrix whose i^{th} column is the vector $a^i(k)$
  - By Weights Rule(b), $A'(s)$ is a stochastic matrix
- By the linearity of the dynamics, the iterates satisfy

$$x^i(k+1) = \sum_{j=1}^{m} [A(s)A(s+1)\cdots A(k-1)a^i(k)]_j x^j(s)$$

- We introduce the transition matrices

$$\Phi(k,s) = A(s)A(s+1)\cdots A(k-1)A(k) \qquad \text{for all } k \geq s$$

- Then: $\quad x^i(k+1) = \sum_{j=1}^{m} [\Phi(k,s)]_j^i x^j(s)$

Properties of Transition Matrices

Lemma: Let Weights Rule (a), Connectivity, and Bounded Intercommunication Interval assumptions hold. We then have

$$[\Phi(s+(m-1)B-1,s)]_j^i \geq \eta^{(m-1)B} \qquad \text{for all } s,\ i, \text{ and } j,$$

where η is the lower bound on weights and B is the intercommunication interval bound.

- We introduce the matrices $D_k(s)$ as follows: for a fixed $s \geq 0$,

  $$D_k(s) = \Phi'\left(s+kB_0-1, s+(k-1)B_0\right) \qquad \text{for } k = 1, 2, \ldots,$$

  where $B_0 = (m-1)B$.

- By the previous lemma, all entries of $D_k(s)$ are positive.

Convergence of Transition Matrices

Lemma: Let Weights Rule, Connectivity, and Bounded Intercommunication Interval assumptions hold. For each $s \geq 0$, we have:

(a) The limit $\bar{D}(s) = \lim_{k\to\infty} D_k(s) \cdots D_1(s)$ exists.

(b) The limit $\bar{D}(s)$ is a stochastic matrix with identical rows.

(c) The convergence of $D_k(s) \cdots D_1(s)$ to $\bar{D}(s)$ is geometric: $\forall\, x \in \mathbb{R}^m$,

$$\left\| \left(D_k(s) \cdots D_1(s)\right) x - \bar{D}(s)x \right\|_\infty \leq 2\left(1 + \eta^{-B_0}\right)\left(1 - \eta^{B_0}\right)^k \|x\|_\infty$$

In particular, for every j, the entries $[D_k(s) \cdots D_1(s)]_i^j$, $i = 1, \ldots, m$, converge to the same limit $\phi_j(s)$ as $k \to \infty$ with a geometric rate:

$$\left| [D_k(s) \cdots D_1(s)]_i^j - \phi_j(s) \right| \leq 2\left(1 + \eta^{-B_0}\right)\left(1 - \eta^{B_0}\right)^k$$

where η is the lower bound on weights, B is the intercommunication interval bound, and $B_0 = (m-1)B$.

Proof Idea

- We show that the sequence $\{(D_k \cdots D_1)x\}$ converges for every $x \in \mathbb{R}^m$
- Consider the sequence $\{x_k\}$ with $x_k = D_k \cdots D_1 x$ and write x_k as

$$x_k = z_k + c_k e, \qquad \text{where } c_k = \min_{1 \le i \le m} [x_k]_i$$

- Using the property that each entry of the matrix D_k is positive, we show

$$\|z_k\|_\infty \le \left(1 - \eta^{B_0}\right)^k \|z_0\|_\infty \qquad \text{for all } k.$$

  Hence $z_k \to 0$ with a geometric rate.
- We then show that the sequence $\{c_k\}$ converges to some $\bar{c} \in \mathbb{R}$ and use the contraction constant to establish the rate estimate
- The final relation follows by picking $x = e_j$, the j^{th} unit vector

Convergence of Transition Matrices

Proposition: Let Weights Rule, Connectivity, and Bounded Intercommunication Interval assumptions hold.

(a) The limit $\bar{\Phi}(s) = \lim_{k\to\infty} \Phi(k,s)$ exists for each s.

(b) The limit matrix $\bar{\Phi}(s)$ has identical columns and the columns are stochastic, i.e.,

$$\bar{\Phi}(s) = \phi(s)e',$$

where $\phi(s) \in \mathbb{R}^m$ is a stochastic vector for each s.

(c) For every i, $[\Phi(k,s)]_i^j$, $j = 1, ..., m$, converge to the same limit $\phi_i(s)$ as $k \to \infty$ with a geometric rate, i.e., for all i, j and all $k \geq s$,

$$\left| [\Phi(k,s)]_i^j - \phi_i(s) \right| \leq 2\, \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} \, \left(1 - \eta^{B_0}\right)^{\frac{k-s}{B_0}}$$

where η is the lower bound on weights, B is the intercommunication interval bound, and $B_0 = (m-1)B$.

Model with Delays

- Assume now that there is delay in delivering information of j to i
  - Models communication delay over wireless links
- In the presence of delay, agent i updates his information state by

$$x^i(k+1) = \sum_{j=1}^{m} a_j^i(k) x^j(k - t_j^i(k)),$$

where $t_j^i(k)$ is the delay in passing information from j to i

Assumption (*Bounded Delays*)

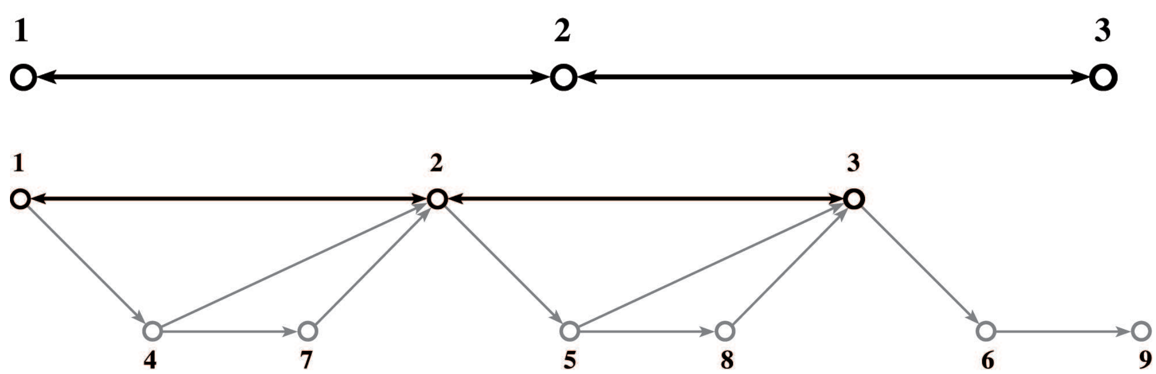
(a) $t_i^i(k) = 0$ for all agents i and all $k \geq 0$.

(b) $t_j^i(k) = 0$ for all agents i and j such that $a_j^i(k) = 0$.

(c) There is an integer B_1 such that $0 \leq t_j^i(k) \leq B_1 - 1$ for all agents i, j, and all k.

Enlarged Linear System

- Under Bounded Delays assumption, the system with delays can be reduced to a system without delays by state augmentation
- For each agent i, we associate a new agent for every possible delay value
  - Sufficient to add $m(B_1 - 1)$ new agents handling delays
- We refer to the original agents as **computing agents** (indexed by $1, \ldots, m$) and the new agents as **non-computing agents** (indexed by $m + 1, \ldots, (B_1 - 1)m$)
- An example with 3 agents and delay bound $B_1 = 3$



Information State Evolution

- Let $\tilde{x}^i(k)$ denote information state of agent i of the enlarged system
- The evolution of the states $\tilde{x}^i(k)$ for all agents i in the enlarged system:

$$\tilde{x}^i(k+1) = \sum_{h=1}^{mB_1} \tilde{a}_h^i(k)\tilde{x}^h(k),$$

where weights $\tilde{a}_h^i(k)$ for **computing agents** $i \in \{1, \ldots, m\}$ are

$$\tilde{a}_h^i(k) = \begin{cases} a_j^i(k) & \text{if } h = j + tm, \;\; t = t_j^i(k) \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } k \geq 0,$$

while weights $\tilde{a}_h^i(k)$ for **noncomputing agents** $i \in \{m+1, \ldots, mB_1\}$ are

$$\tilde{a}_h^i(k) = \begin{cases} 1 & \text{for } h = i - m \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } k \geq 0.$$

Convergence Analysis

- Let $\tilde{A}(k)$ denote the matrix whose columns are given by $\tilde{a}^i(k)$.
  - Under Weights Rule, $\tilde{A}'(k)$ is a stochastic matrix for all k.
- Similar to the previous analysis, we define the transition matrices

$$\tilde{\Phi}(k,s) = \tilde{A}(s)\tilde{A}(s+1)\cdots\tilde{A}(k-1)\tilde{A}(k) \qquad \text{for all } k \geq s.$$

Lemma:

(a) For any computing nodes $i, j \in \{1, \ldots, m\}$,

$$[\tilde{\Phi}(k,s)]_j^i \geq \eta^{k-s+1} \qquad \text{for all } k \geq s + (m-1)(B + B_1).$$

(b) For any computing node $j \in \{1, \ldots, m\}$, we have

$$[\tilde{\Phi}(s + (m-1)B + mB_1 - 1, s)]_j^i \geq \eta^{(m-1)B + mB_1} \qquad \text{for all } i.$$

Convergence Rate

Proposition:

(a) The limit $\tilde{\Phi}(s) = \lim_{k\to\infty} \tilde{\Phi}(k,s)$ exists for each s.

(b) The limit matrix $\tilde{\Phi}(s)$ has identical columns and the columns are stochastic, i.e.,

$$\tilde{\Phi}(s) = \tilde{\phi}(s)e',$$

where $\tilde{\phi}(s) \in \mathbb{R}^{mB_1}$ is a stochastic vector for each s.

(c) For every $i \in \{1,\ldots,mB_1\}$, the entries $[\tilde{\Phi}(k,s)]_i^j$, $j = 1, ..., mB_1$, converge to the same limit $\tilde{\phi}_i(s)$ as $k \to \infty$ with a geometric rate, i.e.,

$$\left|[\tilde{\Phi}(k,s)]_i^j - \tilde{\phi}_i(s)\right| \leq 2\,\frac{1+\eta^{-B_2}}{1-\eta^{B_2}}\,\left(1-\eta^{B_2}\right)^{\frac{k-s}{B_2}},$$

where B is the intercommunication interval bound, B_1 is the delay bound, and $B_2 = (m-1)B + mB_1$.

Conclusions

- We presented a general distributed computational model for the consensus problem
- We provided convergence analysis and convergence rate estimates with and without communication delay
- Our estimates highlight the dependence of convergence rate on key system parameters
- Ongoing work:
  - Distributed asynchronous subgradient methods for constrained multi-agent optimization
  - Effects of quantization of information states on consensus