Information Theory for Mobile Ad-Hoc Networks (ITMANET): The FLoWS Project

Distributed Asynchronous Optimization Methods for General Performance Metrics

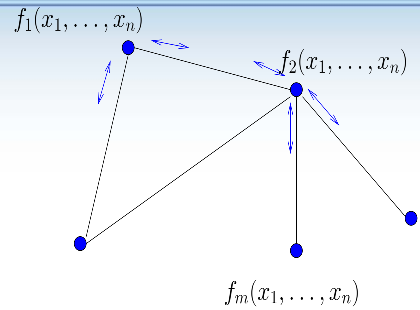
Asu Ozdaglar

Joint work with Angelia Nedic

Distributed Asynchronous Optimization Methods for General Performance Metrics

FLoWS

STATUS QUO



Existing methodology based on Lagrangian relaxation and duality does not lend itself to distributed algorithms for general non-separable (coupled) user performance metrics in wireless networks with time-varying connectivity

NEW INSIGHTS

Subgradient methods with simple consensus (averaging) policies lead to decentralized algorithms that
•optimize general performance metrics,
•are robust against changes in network topology

ACHIEVEMENT DESCRIPTION

MAIN RESULT:

- **Development of a distributed computational method for optimizing the sum of performance measures of users**
- **The method operates asynchronously under time-varying connectivity**
- **We provide convergence rate results that explicitly characterize the impact of the system and algorithm parameters on the quality of generated solutions.**

HOW IT WORKS:

- **Each user maintains an information state, which is an estimate of the optimal solution.**
- **The update rule for each user involves combining his information state with that of his current neighbors and performing a subgradient step using his local performance measure.**

ASSUMPTIONS AND LIMITATIONS:

- **The model is unconstrained.**
- **The communication bandwidth constraints have not been taken into account.**

END-OF-PHASE GOAL

•We will extend the model to include local (potentially time-varying) constraints for each user.

•We will explore the effect of bandwidth constraints (i.e., quantized information exchange) on the performance of the algorithms.

COMMUNITY CHALLENGE

Design of optimization algorithms that address the challenges and constraints associated with large-scale time-varying networks

Distributed optimization algorithms for general performance metrics and time-varying connectivity

Motivation

- Increasing interest in distributed optimization and control of ad hoc wireless networks, which are characterized by:
  – Lack of centralized control and access to information
  – Time-varying connectivity
- Control-optimization algorithms deployed in such networks should be:
  – Distributed relying on local information
  – Robust against changes in the network topology
  – Easily implementable
- Existing theory does not lend itself to distributed algorithms for general non-separable user performance metrics in wireless networks with time-varying connectivity

Multi-Agent Optimization Model

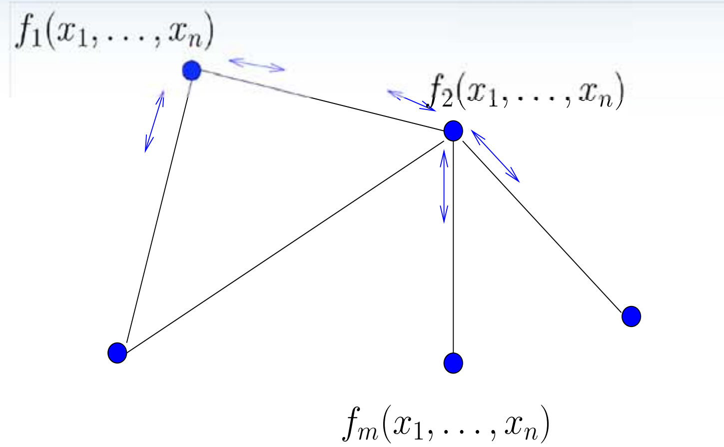
- Consider a network with node set

  $V = \{1, \ldots, m\}$

- Agents want to cooperatively solve the problem

  $$\min_{x \in X} \sum_{i=1}^{m} f_i(x)$$

- Function $f_i(x) : R^n \rightarrow R$ is a performance measure known only by node *i*



Examples:

- Performance measure: Utility-latency
- Parameter estimation from local sensor measurements

Agent Update Rule

FLoWS

- Agents update and exchange information at discrete times $t_0, t_1, \ldots$
- Agent *i* information state is denoted by $x^i(k) \in R^n$ at time *tk*
- Agent *i* updates his information state according to:($X = R^n$)

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

$a_j^i(k)$ are weights, $\alpha^i(k)$ is stepsize, $d^i(k)$ is subgradient of f_i at $x^i(k)$

- Time-varying communication is modeled by matrix $A(k)$ [columns $a^i(k)$]
- **Main Novelty:**

$\boxed{\text{Approximate subgradient step and consensus policy}}$

Weights

Assumption (Weights Rule):

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

(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.

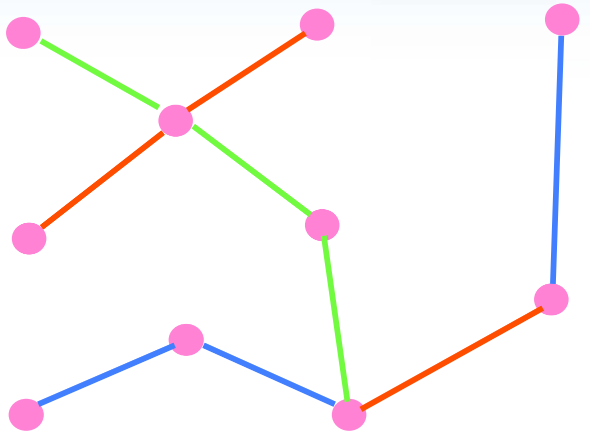
Example: Equal neighbor weights

$$a_j^i(k) = \frac{1}{n_i(k)+1}$$

where $n_i(k)$ is the number of agents communicating with *i* (his neighbors) at slot *k*

Information Exchange

- Information state of agent *i* influences information state of any other agent infinitely often – connectivity
- Agent *j* send his information to neighboring agent *i* within a bounded time interval – bounded intercommunications



•At slot k, information exchange may be represented by a directed graph (V, E_k) with $E_k = \{(j,i) \mid a_j^i(k) > 0\}$

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

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

•Assumption (Bounded Intercommunication Interval) There is $B \geq 1$ s.t.

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

Linear Dynamics and Transition Matrices

- Compact representation of agent local-update relation: for $k \geq s$

$$
\begin{aligned}
x^i(k+1) &= \sum_{j=1}^{m} [\Phi(k,s)]_j^i x^j(s) \\
&- \sum_{r=s}^{k-1} \left(\sum_{j=1}^{m} [\Phi(k,r+1)]_j^i \alpha^j(r) d_j(r) \right) - \alpha^i(k) d_i(k).
\end{aligned}
$$

where $\Phi(k,s)$ are transition matrices from time s to k:

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

- We analyze convergence properties of the distributed method by establishing:
  - Convergence of transition matrices
  - Convergence of stopped "subgradient updates"

Convergence of Transition Matrices

Proposition: Let weights rule, connectivity, and bounded intercommunication interval assumptions hold:

- The limit $\bar{\Phi}(s) = \lim_{k\to\infty} \Phi(k,s)$ exists for each s.
- The limit matrix $\bar{\Phi}(s)$ has identical columns and the columns are stochastic, i.e.,

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

  where $\phi(s) \in R^m$ is a stochastic vector for each s.
- 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 $B_0 = (m-1)B$.

"Stopped Model"

- Consider the local-update relation with $\alpha^i(k) = \alpha$ for $k \geq s$

$$
\begin{aligned}
x^i(k+1) &= \sum_{j=1}^{m} [\Phi(k,s)]_j^i x^j(s) \\
&- \alpha \sum_{r=s}^{k-1} \left(\sum_{j=1}^{m} [\Phi(k,r+1)]_j^i \; d_j(r) \right) - \alpha d_i(k).
\end{aligned}
$$

- Suppose agents cease computing subgradients but keep exchanging their estimates: for a $\bar{k} \geq 0$, $d_j(k) = 0$ for all j and $k \geq \bar{k}$
- It can be seen that the stopped process takes the form:

$$
\bar{x}^i(k+1) = \sum_{j=1}^{m} \sum_{j=1}^{m} [\Phi(k,0)]_j^i x^j(0) \; - \alpha \sum_{r=1}^{\bar{k}} \left(\sum_{j=1}^{m} [\Phi(k,r)]_j^i d_j(r-1) \right).
$$

Using $\lim_{k\to\infty} [\Phi(k,s)]_j^i \phi_j(s)$ for all i, we see that the limit vector $\lim_{k\to\infty} \bar{x}_i(k)$ exists and is independent of i, but dependent on $\bar{k}$

Behavior of Stopped Process

- Stopped process in the limit is described by:

$$y(k+1) = y(k) - \alpha \sum_{j=1}^{m} \phi(k)_j d^j(k)$$

- These iterations would correspond to an "approximate" subgradient method for minimizing $\sum_j f_j(x)$ provided that the values $\phi(k)_j$ are the same for all *j.*
- This is true for example when the following holds:
  - Assumption (Double Stochasticity)

    The matrices $A(k)$ are doubly stochastic
- Can be ensured when the agents exchange their information simultaneously and coordinate the selection of the weights $a^i_j(k)$

Main Convergence Result

Proposition: Let the subgradients of *fi* be uniformly bounded by a constant *L*. Then for every *i*, the averages $\hat{x}^i(k)$ of estimates $x^i(0), \ldots, x^i(k-1)$ are such that

$$f(\hat{x}^i(k)) \leq f^* + \frac{m\ dist^2(y(0), X^*)}{2\alpha k} + \frac{\alpha L^2 C}{2} + 2\alpha m \hat{L}_1 L C_1,$$

where $f = \sum_i f_i$, f^* is the optimal value, and X^* is the optimal set of the problem, $y(0) = \frac{1}{m} \sum_i x^i(0)$, $C = 1 + 8mC_1$ and

$$C_1 = 1 + \frac{m}{1 - (1 - \eta^{B_0})^{\frac{1}{B_0}}} \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}},$$

with η minimal weight, $B_0 = (m-1)B$, B intercommunication bound.

- Estimates are per iteration
- Capture tradeoffs between accuracy and computational complexity

Conclusions

- We presented a general distributed computational model for general performance measures
- We provided convergence analysis and convergence rate estimates
  - For unconstrained problem
  - Without communication delays in the system
- Future Work:
  - Extension of the method to constrained optimization
  - Effects of delay and quantization of information states
- Paper:
  - Nedic and Ozdaglar "Distributed Asynchronous Subgradient methods for Multi-Agent Optimization," MIT LIDS Technical Report 2575, submitted for publication May 2007