



Distributed algorithms for optimization and control over networks

DEIB PhD Course, Politecnico di Milano February 10-14, 2020







Kostas Margellos

University of Oxford, UK

Alessandro Falsone, Simone Garatti, Maria Prandini Politecnico di Milano, Italy





Maria Prandini DEIB – Politecnico di Milano, Italy maria.prandini@polimi.it tel. 02-2399 3441 e-mail: maria.prandini@polimi.it http://home.deib.polimi.it/prandini/



Those of you who registered to the course as "single course attendees" should contact the PhD secretariat to officially appear in the list of attendees.



Introduce to the analysis and design of distributed decision making schemes for multi-agent systems seeking convergence to an optimal cooperative solution.



Introduce to the analysis and design of distributed decision making schemes for multi-agent systems seeking convergence to an optimal cooperative solution.

Motivating applications: energy and transportation systems





The course is structured in 4 parts



Part 1: Motivation and illustrative applications

Introduction to decision making problems arising in smart grid control and optimization, and in coordination and control for electric vehicle fleets.



Part 1: Motivation and illustrative applications

Introduction to decision making problems arising in smart grid control and optimization, and in coordination and control for electric vehicle fleets.

Part 2: Mathematical tools

Introduction to the mathematical tools of graph theory, convex analysis, optimization, duality theory that constitute the theoretical backbone for the analysis and design of cooperative algorithms



Part 3: Distributed cooperative algorithms

Primal-based and dual-based algorithms for distributed cooperative decision making will be illustrated, resting on the math tools of Part 2

Part 4: Distributed optimization in uncertain networks

The algorithmic solutions described in Part 3 will be extended to the case when the multi-agent optimization problem is affected by uncertainty.





Students will be evaluated based on a small project or on the study of an advanced topic related to the course.

Project and topic should be agreed upon with the organizer of the course.





Monday, February 10

11:30 – 13:00 Motivation [MP] 14:30 - 17:00 Math tools [MP]

Tuesday, February 11

09:30 - 11:00 & 11:15 - 12:00 Math tools [MP]

14:30 - 16:00 Math tools [KM]

Wednesday, February 12

09:30 – 11:00 & 11:15 – 12:00 Primal-based algo [KM] 14:30 - 16:00 Primal-based algo [KM]

Thursday, February 13

- 09:30 11:00 & 11:15 12:00 Duality-based algo [AF]
- 14:30 16:00 Duality-based algo [AF]

Friday, February 14

09:30 - 11:00 & 11:15 - 12:00 Distributed uncertain opt [SG]

14:30 - 16:00 Distributed uncertain opt [SG]





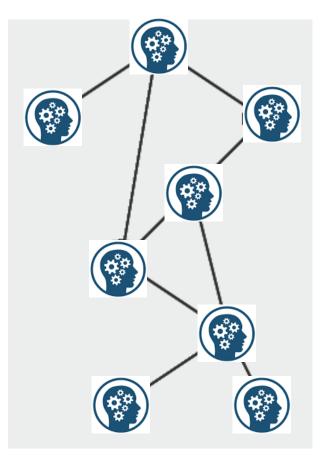


Large-scale system composed of multiple sub-systems (agents) with

- computational power
- communication capabilities







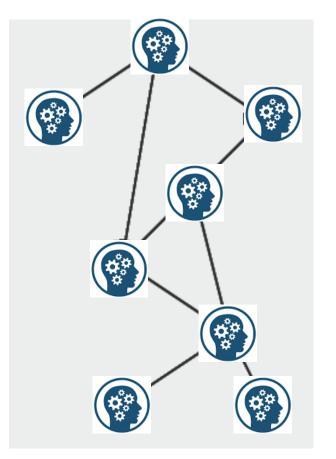
Large-scale system composed of multiple sub-systems (agents) with

- computational power
- communication capabilities

Agents can communicate according to a given communication graph to possibly agree on a solution to a coupled decision problem.







Large-scale system composed of multiple sub-systems (agents) with

- computational power
- communication capabilities

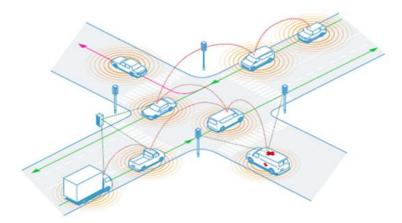
Agents can communicate according to a given communication graph to possibly agree on a solution to a coupled decision problem.

→ network system





Transportation systems



Robotic network



Energy systems



Social network





- Large scale systems
- Multi-agent multiple interacting entities/users
- Heterogeneous different physical or technological constraints per agent; different objectives per agent

- Large scale systems
- Multi-agent multiple interacting entities/users
- Heterogeneous different physical or technological constraints per agent; different objectives per agent

Challenges

Computation: problem size too big

- Large scale systems
- Multi-agent multiple interacting entities/users
- Heterogeneous different physical or technological constraints per agent; different objectives per agent

Challenges

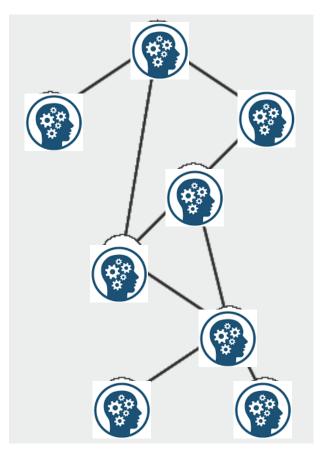
- Computation: problem size too big
- Communication: not all communication links at place; link failure; communication graph given, not to be designed

- Large scale systems
- Multi-agent multiple interacting entities/users
- Heterogeneous different physical or technological constraints per agent; different objectives per agent

Challenges

- Computation: problem size too big
- Communication: not all communication links at place; link failure; communication graph given, not to be designed
- Information privacy: agents may not want to share information with everyone

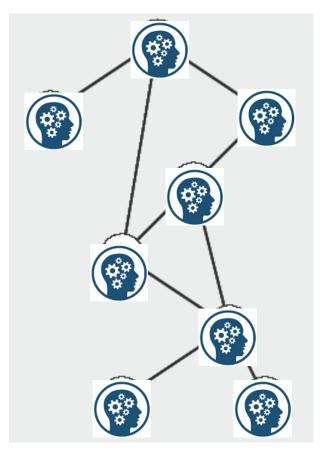




Large-scale system composed of multiple **cooperative agents** with

- computational power
- communication capabilities





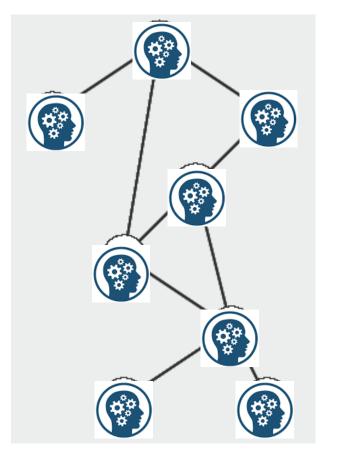
Large-scale system composed of multiple **cooperative agents** with

- computational power
- communication capabilities

Goal

Solving optimally a coupled decision making problem involving the overall network, so as to get the social welfare solution



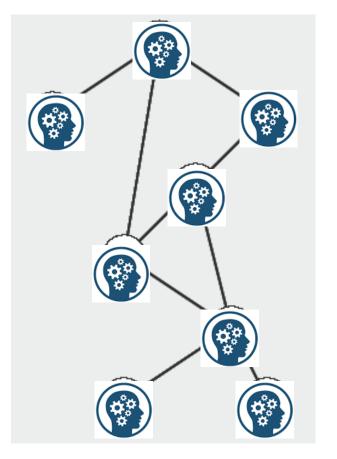


The decision making problems can typically be formulated as an optimization program for the overall system

where a cost function, typically the sum of local costs, is minimized subject to

- coupled decisions
- local information
- communication constraints





The decision making problems can typically be formulated as an optimization program for the overall system

where a cost function, typically the sum of local costs, is minimized subject to

- coupled decisions
- local information
- communication constraints

If the optimization problem has a separable structure, then, a distributed solution can be adopted



The large scale optimization problem is addressed by distributing computations locally to the agents.

Each agent iteratively solves a smaller optimization problem based on its local information (local objective and constraints) and the information received from its neighbors, till convergence, ideally to the global optimum



Scalable methodology

- Communication: Only between neighbors
- Computation: Only local; in parallel for all agents







- Communication: Only between neighbors
- Computation: Only local; in parallel for all agents
- Information privacy
 - Agents do not reveal information about their preferences (encoded by objective and constraint functions) to each other





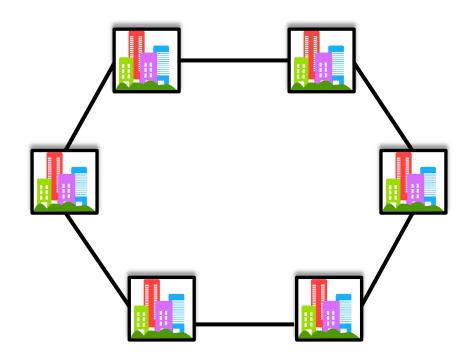
- Scalable methodology
 - Communication: Only between neighbors
 - Computation: Only local; in parallel for all agents
- Information privacy
 - Agents do not reveal information about their preferences (encoded by objective and constraint functions) to each other
- Resilience to communication failures



- Scalable methodology
 - Communication: Only between neighbors
 - Computation: Only local; in parallel for all agents
- Information privacy
 - Agents do not reveal information about their preferences (encoded by objective and constraint functions) to each other
- Resilience to communication failures
- Numerous applications
 - Wireless networks
 - Optimal power flow
 - Electric vehicle charging control
 - Energy management in building networks

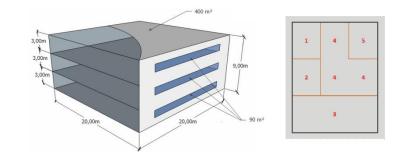
Energy management in a building network



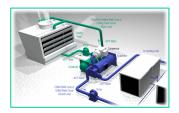


Building network

m buildings, divided into zones

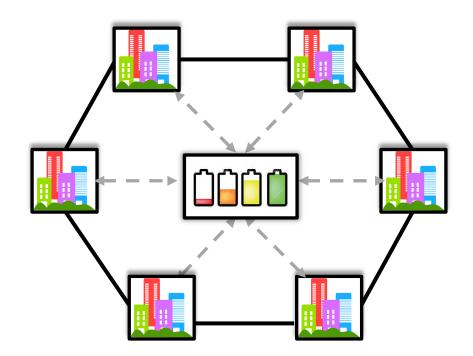


a chiller unit for each building



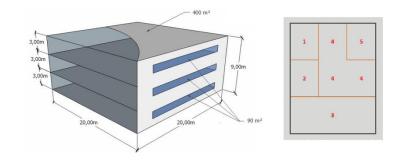
Energy management in a building network



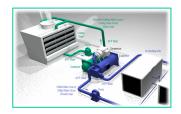


Building network

m buildings, divided into zones



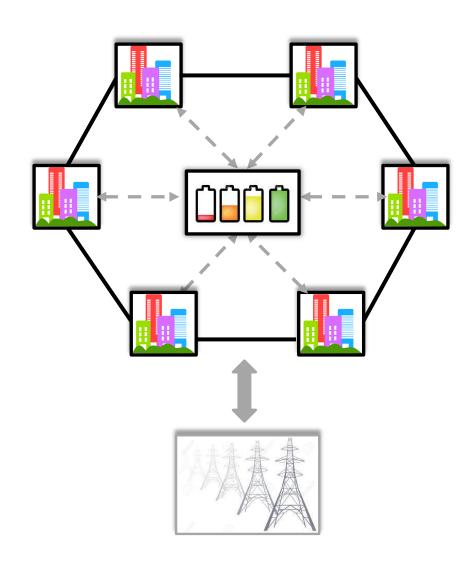
a chiller unit for each building



a shared cooling network

Energy management in a building network





Building network

 connected to the main grid for electrical energy provision



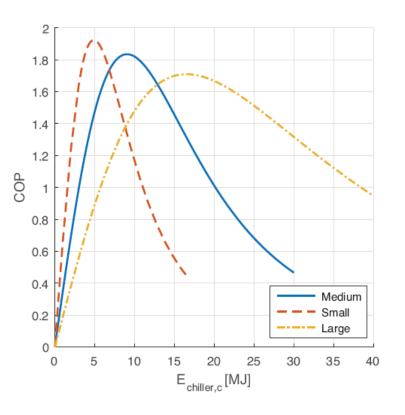
minimize the electrical energy cost for cooling a building network by operating the chillers at their maximum efficiency, possibly shifting the

cooling energy request in time by exploiting the shared cooling network.



minimize the electrical energy cost for cooling a building network by operating the chillers at their maximum efficiency, possibly shifting the cooling energy request in time by exploiting the shared cooling network.

Coefficient Of Performance: $COP = \frac{E_{ch}}{E_l}$





minimize the electrical energy cost for cooling a building network by operating the chillers at their maximum efficiency, possibly shifting the cooling energy request in time by exploiting the shared cooling network.

Set the energy exchanges with the cooling network of all buildings so as to

- 1. minimize the electrical energy cost
- 2. guarantee comfort conditions

over a 1-day time horizon



minimize the electrical energy cost for cooling a building network by operating the chillers at their maximum efficiency, possibly shifting the cooling energy request in time by exploiting the shared cooling network.

Set the energy exchanges with the cooling network of all buildings so as to

- 1. minimize the electrical energy cost
- 2. guarantee comfort conditions

over a 1-day time horizon

Decision variables: energy exchanges with the cooling network (the storage)



minimize cost of all chillers electrical energy consumption

- with respect to decision variables
- subject to
- 1. Chiller energy request = building energy request storage energy
- 2. Storage dynamics
- 3. Storage limits, chiller limits, comfort limits

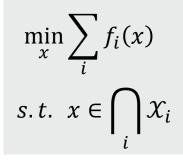


minimize cost of all chillers electrical energy consumption

with respect to decision variables

subject to

- 1. Chiller energy request = building energy request storage energy
- 2. Storage dynamics
- 3. Storage limits, chiller limits, comfort limits



Constrained optimization problem

x : decision vector (decision variables along 1-day horizon) $f_i(x)$: energy cost due to the chiller of building *i* \mathcal{X}_i : comfort and actuation constraints





$$\min_{x} \sum_{i} f_{i}(x)$$
s.t. $x \in \bigcap_{i} X_{i}$

x: decision vector (control input along some finite horizon) $f_i(x)$: energy cost due to the chiller of building *i* X_i : comfort and actuation constraints

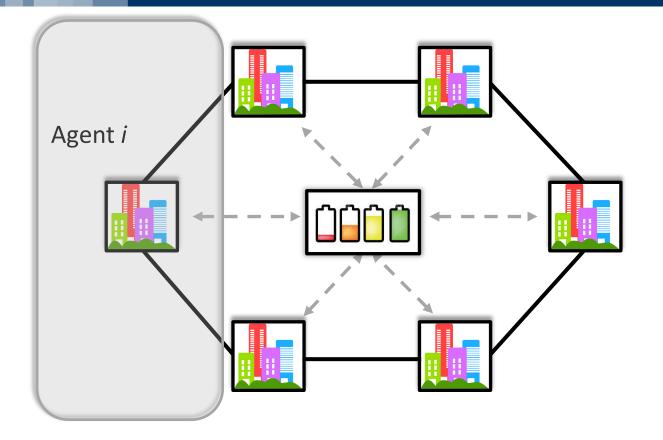
Issues

- Computation problem size may be too big
- Communication large amount of local info should be transmitted
- Information privacy

buildings may not want to share their consumption profiles



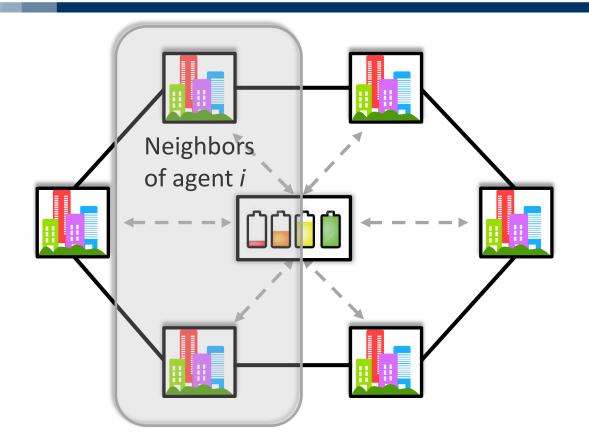




Step 1: agent *i* solves a local decision problem and makes a tentative (local) decision for *x*



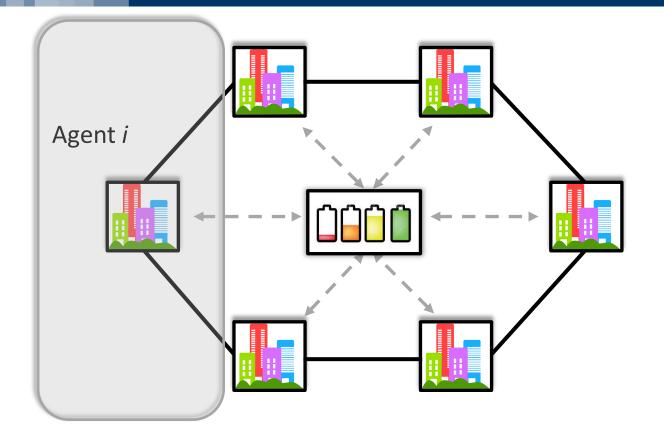




Step 2: neighbouring agents communicate their tentative decisions to agent *i*



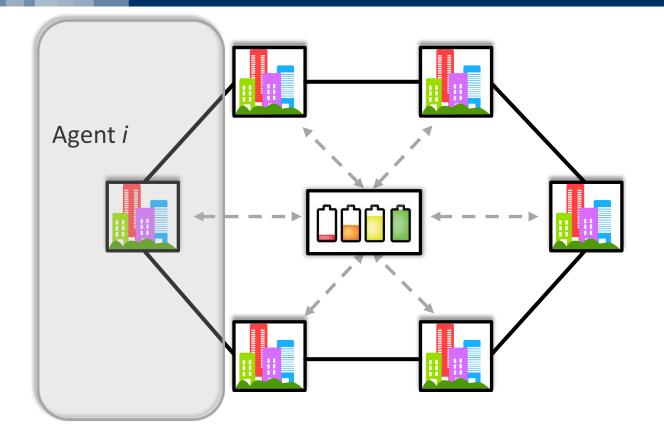




Step 3: Agent *i* weights the received information, solves a refined problem and makes a new decision for *x*







Step 3: Agent *i* weights the received information, solves a refined problem and makes a new decision for *x* until convergence to some consensus solution



Local problem of agent *i* at iteration k + 1



$$z_{i}(k) = \sum_{j} a_{j}^{i}(k) x_{j}(k)$$
$$x_{i}(k+1) = \arg\min_{x_{i} \in X_{i}} f_{i}(x_{i}) + \frac{1}{2c(k)} \|x_{i} - z_{i}(k)\|^{2}$$

Information vector

 $z_i(k) = \sum_j a_j^i(k) x_j(k)$ $a_j^i(k): \text{ how agent } i \text{ weights info of agent } j$

Proxy term

 $\frac{1}{2c(k)} ||x_i - z_i(k)||^2$: deviation from (weighted) average c(k): trade-off between optimality and agents' disagreement





Theorem (consensus and optimality)

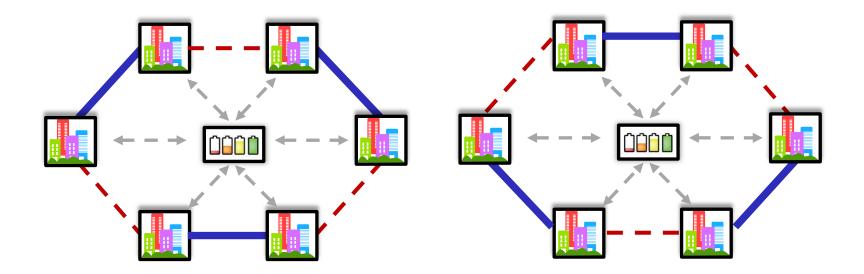
Under suitable assumptions on the weights and connectivity, if the problem is convex, then, all the agents asymptotically converge to the optimal solution of the global problem.

Distributed constrained optimization and consensus in uncertain networks via proximal minimization. IEEE TAC, vol. 63, May 2018.



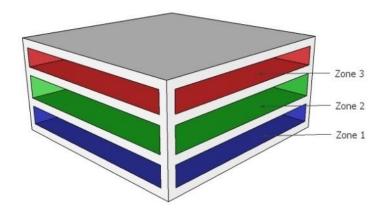
Assumption on connectivity

- Any pair of agents communicates infinitely often, possibly through a communication graph that changes through iterations
- The intercommunication time is bounded

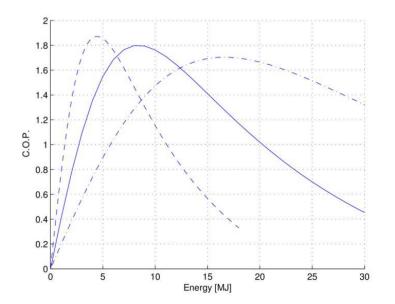




Network configuration: 3 identical buildings with three zones and with different chillers



Chiller types: 'small' for building 2 'medium' for building 1 'large' for building 3



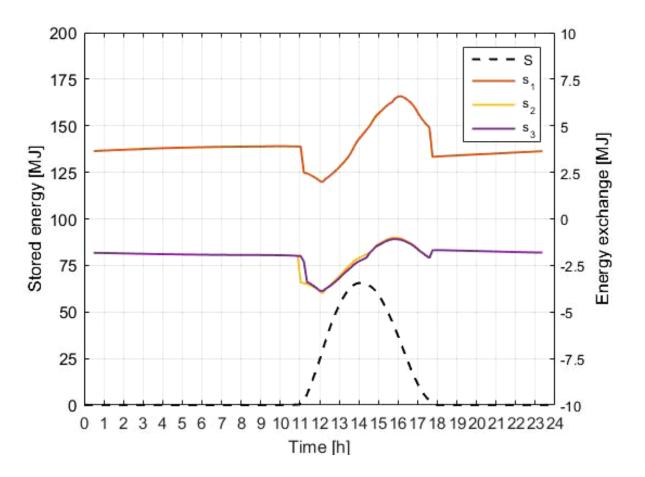
Numerical example: convergence



Chiller types: 'small' for building 2 'medium' for building 1 'large' for building 3

solution computed by

building 1



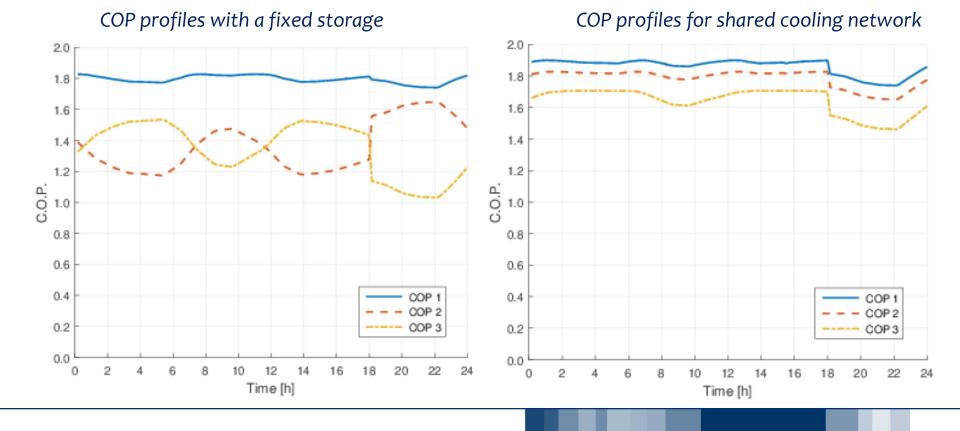


If fixed storage, then building 2 and building 3 chillers work at suboptimal efficiency

- Building 2 chiller (small) is overloaded
- Building 3 chiller (large) is underloaded

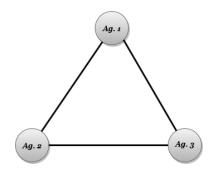
Effect of the "shared cooling system"

All chillers work close to maximum COP

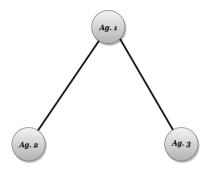




Complete network graph For each pair of agents (i, j) there exists a direct arc

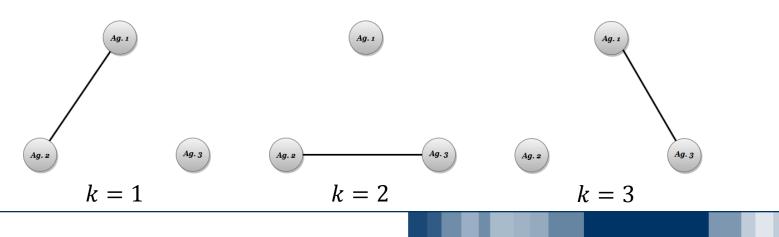


(Strongly) Connected network graph For each pair of agents (i, j)There exists a direct path



Time – varying network graph

The network graph changes at each iteration k, cyclically with 3 iteration period





Communication network

- Does affect the speed of convergence and # of iterations required
 - Complete network is the fastest
 - Time varying is the slowest
- Does **not** affect convergence properties or local problem computational time

	Complete	Connected	Time-varying
	Network	Network	Network
# of iterations	194	354	878
Average time			
Building 1	2.743s	2.747s	2.693s
Building 2	2.753s	2.753s	2.687s
Building 3	2.738s	2.742s	2.686s
Overall time			
Building 1	532.160s	972.692s	2407.32s
Building 2	534.164s	968.333s	2399.91s
Building 3	531.219s	970.668s	2398.38s

$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

- local costs
- Iocal constraints
- coupling decision vector



$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

Each agent knows

- local cost
- local constraint

Goal:

agree on a global minimizer

- local costs
- Iocal constraints
- coupling decision vector



$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

Each agent knows

- local cost
- local constraint

Goal:

agree on a global minimizer

- local costs
- local constraints
- coupling decision vector

decision coupled problem

$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

- local costs
- local constraints
- coupling decision vector

At each iteration $\boldsymbol{k},$ agent \boldsymbol{i}

$$\begin{aligned} z_i(k) &\leftarrow \sum_j \alpha_j^i x_j(k) \\ x_i(k+1) &\leftarrow \arg\min_{x_i} \tilde{f}_i(x_i) \\ \text{s.t. } x_i \in \mathcal{X}_i \end{aligned}$$

where

$$\tilde{f}_i(x_i) = f_i(x_i) + \frac{1}{c(k)} ||x_i - z_i(k)||_2^2$$

$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

- local costs
- local constraints
- coupling decision vector

At each iteration $\boldsymbol{k},$ agent \boldsymbol{i}

$$z_i(k) \leftarrow \sum_j \alpha_j^i x_j(k)$$
$$x_i(k+1) \leftarrow \arg\min_{x_i} \tilde{f}_i(x_i)$$
s.t. $x_i \in \mathcal{X}_i$

where

$$\tilde{f}_i(x_i) = f_i(x_i) + \frac{1}{c(k)} ||x_i - z_i(k)||_2^2$$

primal-based distributed algorithms

$$\min_{x_1, \dots, x_m} \sum_{i=1}^m f_i(x_i)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$
$$\sum_{i=1}^m g_i(x_i) \le 0$$

- Iocal costs
- Iocal constraints
- local decision vectors
- coupling constraint



Structured optimization problems

Problem 2

$$\min_{x_1, \dots, x_m} \sum_{i=1}^m f_i(x_i)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$
$$\sum_{i=1}^m g_i(x_i) \le 0$$

- local costs
- local constraints
- local decision vectors
- coupling constraint

Each agent knows

- local cost
- local constraint
- local coupling contributionGoal:

minimize global cost



Structured optimization problems

Problem 2

$$\min_{x_1, \dots, x_m} \sum_{i=1}^m f_i(x_i)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$
$$\sum_{i=1}^m g_i(x_i) \le 0$$

- local costs
- local constraints
- local decision vectors
- coupling constraint

Each agent knows

- local cost
- local constraint
- local coupling contributionGoal:

minimize global cost

constraint coupled problem

Introduce the Lagrangian function

$$L(x,\lambda) = \sum_{i=1}^{m} L_i(x_i,\lambda) = \sum_{i=1}^{m} \left\{ f_i(x_i) + \lambda^\top g_i(x_i) \right\}$$

where $\boldsymbol{\lambda}$ is a vector of Lagrange multipliers and define the dual function

$$\varphi(\lambda) = \min_{x \in X} L(x, \lambda) = \sum_{i=1}^{m} \varphi_i(\lambda) = \sum_{i=1}^{m} \min_{x_i \in X_i} L_i(x_i, \lambda)$$

Then, if the problem is convex, the dual problem

$$\max_{\lambda \ge 0} \sum_{i=1}^m \varphi_i(\lambda)$$

provides the optimal value for the primal problem.

Lectures on Math tools: duality theory

Structured optimization problems



Problem 2

$$\min_{x_1, \dots, x_m} \sum_{i=1}^m f_i(x_i)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$
$$\sum_{i=1}^m g_i(x_i) \le 0$$

- local costs
- Iocal constraints
- local decision vectors
- coupling constraint

At each iteration k, agent i

$$\ell_i(k) \leftarrow \sum_j \alpha^i_j \lambda_j(k)$$

$$\lambda_i(k+1) \leftarrow \arg\max_{\lambda_i \ge 0} \varphi_i(\lambda_i)$$

where

$$\varphi_i(\lambda_i) = \lambda_i^\top g_i(x_i(k+1)) \\ - \frac{1}{c(k)} \|\lambda_i - \ell_i(k)\|_2^2$$

Structured optimization problems



Problem 2

$$\min_{x_1, \dots, x_m} \sum_{i=1}^m f_i(x_i)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$
$$\sum_{i=1}^m g_i(x_i) \le 0$$

- local costs
- local constraints
- local decision vectors
- coupling constraint

At each iteration k, agent i

$$\ell_i(k) \leftarrow \sum_j \alpha_j^i \lambda_j(k)$$
$$x_i(k+1) \leftarrow \arg\min_{\substack{x_i \in \mathcal{X}_i \\ \lambda_i(k+1)}} \tilde{f}_i(x_i)$$
$$\lambda_i(k+1) \leftarrow \arg\max_{\substack{\lambda_i \ge 0}} \varphi_i(\lambda_i)$$

where

$$\tilde{f}_i(x_i) = f_i(x_i) + \ell_i(k)^\top g_i(x_i)$$
$$\varphi_i(\lambda_i) = \lambda_i^\top g_i(x_i(k+1))$$
$$- \frac{1}{c(k)} \|\lambda_i - \ell_i(k)\|_2^2$$

duality-based distributed algorithms



Theorem (consensus and optimality)

Under suitable assumptions, all the agents asymptotically converge to the optimal solution of the dual of global problem 2. Moreover, they are able to recover the optimal primal solution.

Dual decomposition for multi-agent distributed optimization with coupling constraints. Automatica, vol. 84, pp. 149-158, October 2017.

Structured optimization problems

Problem 1

$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

decision coupled

$$\min_{x_1, \dots, x_m} \sum_{i=1}^m f_i(x_i)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$
$$\sum_{i=1}^m g_i(x_i) \le 0$$

Structured optimization problems

Problem 1

$$\min_{x} \sum_{i=1}^{m} f_i(x)$$
s.t. $x \in \bigcap_{i=1}^{m} \mathcal{X}_i$

decision coupled

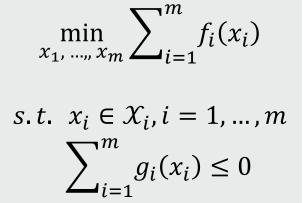
primal-based algorithm

constraint coupled

duality-based algorithm

Math tools from convex analysis, optimization and duality theory to study convergence and optimality, extension to the uncertain case through a data-driven approach.





$$\min_{\substack{x_1, \dots, x_m \\ s.t. }} \sum_{i=1}^m f_i(x_1, \dots, x_m)$$

- Iocal constraints
- local decision vectors
- coupled local costs





$$\min_{\substack{x_1, \dots, x_m \\ s.t. }} \sum_{i=1}^m f_i(x_1, \dots, x_m)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$

- Iocal constraints
- local decision vectors
- coupled local costs

Each agent knows

- local constraint
- local cost

Goal: minimize global cost



$$\min_{\substack{x_1, \dots, x_m \\ s.t. }} \sum_{i=1}^m f_i(x_1, \dots, x_m)$$

s.t. $x_i \in \mathcal{X}_i, i = 1, \dots, m$

- Iocal constraints
- local decision vectors
- coupled local costs

Each agent knows

- local constraint
- local cost

Goal: minimize global cost

cost coupled problem





Design a charging strategy for a fleet of m electric vehicles over a finite horizon





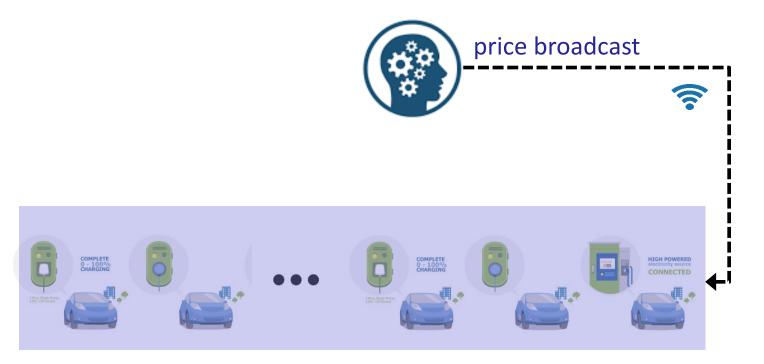
Design a charging strategy for a fleet of m electric vehicles over a finite horizon







Design a charging strategy for a fleet of m electric vehicles over a finite horizon



- A central authority broadcast an energy price incentive based on the total energy consumption
- Agents (EVs) optimize their charging strategy based on the energy price incentive

EV charging – Cooperative set-up

Agents cooperate to achieve a **social welfare** optimizing solution

Agents cooperate to achieve a **social welfare** optimizing solution

Objective

• design a charging strategy over a finite horizon t = 0, ..., T - 1, so as to minimize the cost for the whole fleet

Decision variables

• x^{it} represents the charging rate of agent i at time t

EV constraints

- \underline{x}^{it} and \overline{x}^{it} lower and upper bounds on x^{it}
- γ^i is the charging level to be reached by agent *i* at the end of the time horizon

Energy cost

- $\sum_{i=1}^{m} x^{it}$ total demand at time t
- $p_t(\sum_{i=1}^m x^{it})$ price of energy depends on the total demand

EV charging problem can be formulated as a constrained optimization:

$$\min_{\{x^{1},...,x^{m}\}} \sum_{i=1}^{m} f_{i}(x^{1},...,x^{m}) = \sum_{i=1}^{m} \sum_{t=0}^{T-1} p_{t}\left(\sum_{i=1}^{m} x^{it}\right) x^{it}$$

s.t.
$$\frac{x^{it} \le x^{it} \le \overline{x}^{it}}{\sum_{t=0}^{T-1} x^{it} \le \overline{x}^{it}} \quad \forall t = 0, ..., T-1 \quad \forall i = 1, ..., m$$

- $x^{i} = (x^{i0}, x^{i1}, ..., x^{iT-1})$ is the decision vector of agent i
- the objective function represents the total cost given by the sum of the costs $f_i(x^1, ..., x^m)$ over all the agents
- price of energy per time slot depends on the total demand: $p_t(\sum_{i=1}^m x^{it})$
- constraints are decoupled and ensure that each EV is fully charged, and that power limitations are respected





$$\min_{x^{1},\dots,x^{m}} \sum_{i=1}^{m} f_{i}(x^{1},\dots,x^{m})$$
s.t. $x^{i} \in \mathcal{X}^{i}, i = 1,\dots,m$





$$\min_{\substack{x^1,\dots,x^m \\ s.t. x^i \in \mathcal{X}^i, i = 1,\dots,m}} f_i(x^1,\dots,x^m)$$

- electric vehicles need to charge their battery and decide how much to charge along some time horizon
- price depends on the overall demand and is set by some authority



$$\min_{x^1,\dots,x^m} \sum_{i=1}^m f_i(x^1,\dots,x^m)$$

s.t. $x^i \in \mathcal{X}^i, i = 1,\dots,m$

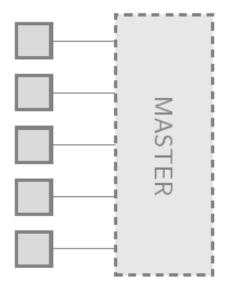
- electric vehicles need to charge their battery and decide how much to charge along some time horizon
- price depends on the overall demand and is set by some authority



decision making problem naturally addressed according to a decentralized optimization scheme where the authority is the central unit collecting info from all agents broadcasting the new price

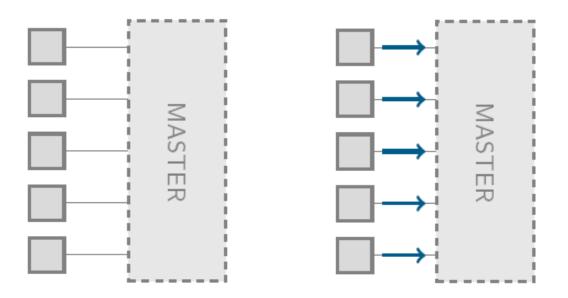
Decentralized, parallel, master/slave scheme





Agents make local computations

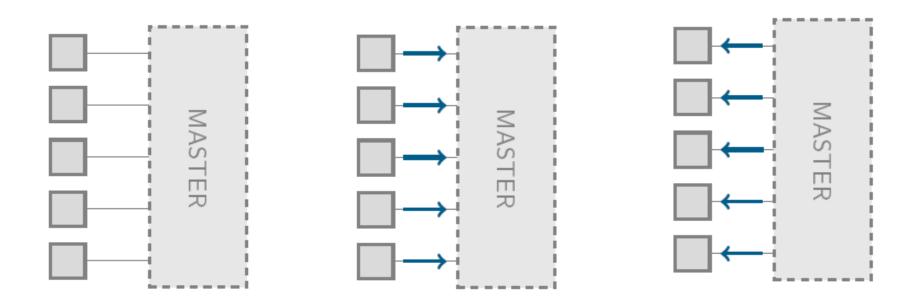




- Agents make local computations
- They communicate the outcome to the central unit (master)

Decentralized, parallel, master/slave scheme

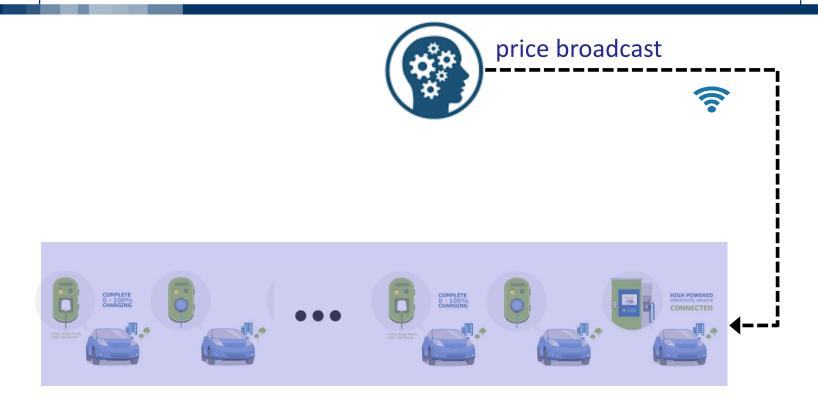




- Agents make local computations
- They communicate the outcome to the central unit (master)
- The master provide back some updated info based on the ones received from the agents

EV charging – Cooperative set-up

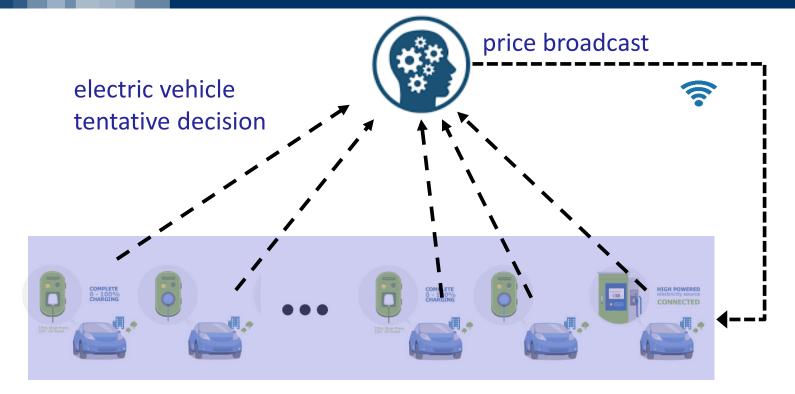




Aggregator/central authority sends a price incentive to all agents, which depends on their consumption levels $x^1, ..., x^m$

EV charging – Cooperative set-up





Agents solve some local problem and send to the aggregator/central authority an update of their consumption levels $x^1, ..., x^m$

A new price is calculated and the process is repeated







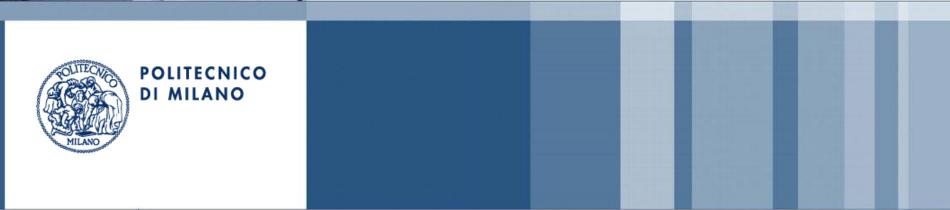




If limit point of the resulting iterative method can be characterized as a fixed point of an operator, then, decentralized optimization rewrites as a fixed point algorithm

Lecture on Decomposition optimization algorithms







Distributed algorithms for optimization and control over networks

DEIB PhD Course, Politecnico di Milano February 10-14, 2020