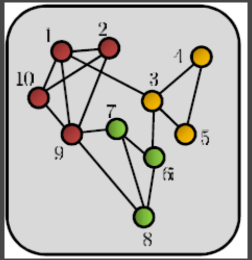



# INFERENCE OVER NETWORKS

## LECTURE #1: Motivation & Examples

Professor Ali H. Sayed
UCLA Electrical Engineering

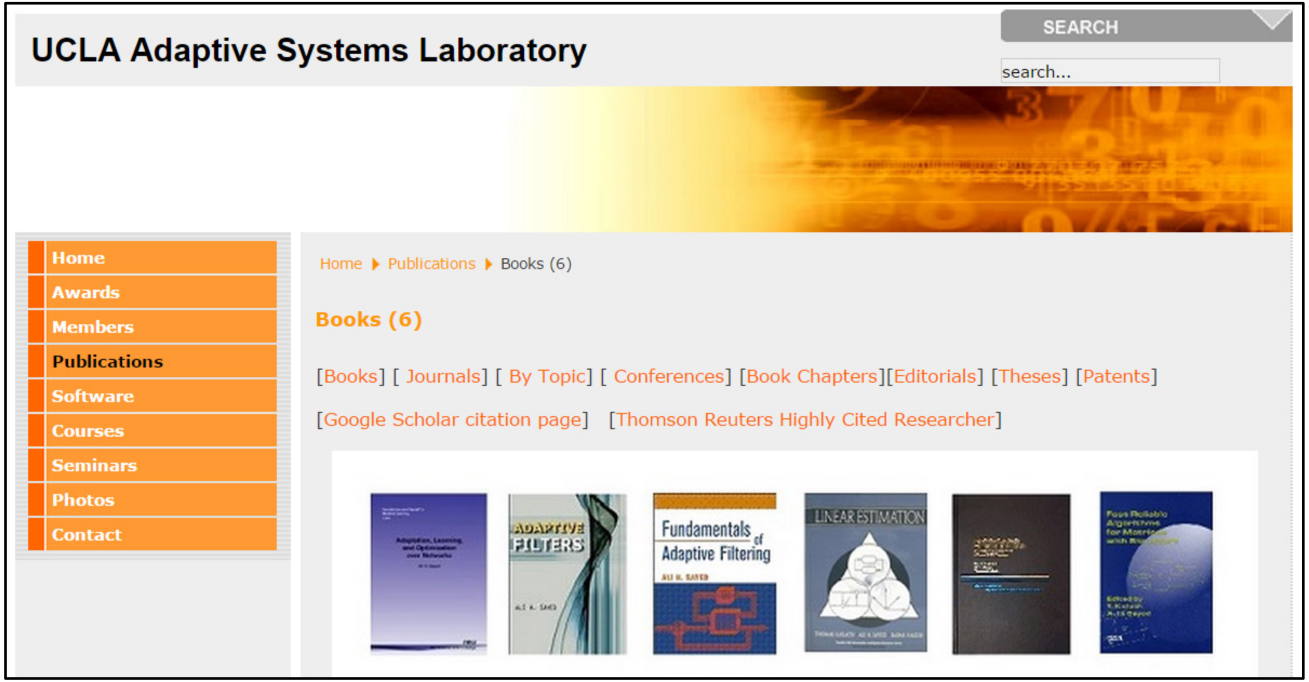
Course EE210B
Spring Quarter 2015

**Proc. IEEE**, vol. 102, no. 4, pp. 460-497, April 2014.
**Foundations and Trends in Machine Learning**, vol. 7, no. 4-5, pp. 311-801, July 2014.

# Contact

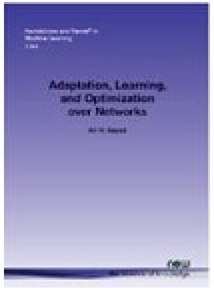


**Email:** sayed@ucla.edu **Web:** http://www.ee.ucla.edu/asl



# References






A. H. Sayed, ``Adaptation, learning, and optimization over networks,'' ***Foundations and Trends in Machine Learning***, vol. 7, issue 4-5, pp. 311-801, NOW Publishers, 2014.

A. H. Sayed, ``Adaptive networks,'' ***Proceedings of the IEEE***, vol. 102, no. 4, pp. 460-497, April 2014.

A. H. Sayed et al, ``Diffusion strategies for adaptation and learning over networks,'' ***IEEE Signal Process. Mag.***, vol. 30, no. 3, pp. 155-171, May 2013.

A. H. Sayed, ``Diffusion adaptation over networks,'' in ***Academic Press Library in Signal Processing***, vol. 3, pp. 323-454, Academic Press, Elsevier, 2014.

# Pre-Requisites



**Prior training expected in:**

**Probability theory:** scalar and vector random variables, pdf, mean, variance, correlation, independence.

**Linear algebra:** matrices, inverses, range and null-spaces, rank, vector and matrix norms, Kronecker products, linear systems of equations, eigen-decomposition, singular-value decomposition, Jordan decomposition.

# Organization



**The course consists largely of five parts:**

**1. Background Material:** Linear Algebra and Matrix Theory Results, Complex Gradients and Complex Hessian Matrices, Convexity, Strict Convexity, and Strong Convexity, Mean-Value Theorems, Lipschitz Conditions.

**2. Single-Agent Adaptation and Learning:** Single-Agent Optimization, Stochastic-Gradient Optimization, Convergence and Stability Properties, Mean-Square-Error Performance.

# Organization



**3. Centralized Adaptation and Learning:** Batch and Centralized Processing, Convergence, Stability, and Performance, Comparison to Single-Agent Processing.

**4. Multi-Agent Network Model:** Graph Properties. Connected and Strongly-Connected Networks, Multi-Agent Inference Strategies, Limit Point and Pareto Optimality, Evolution of Network Dynamics.

# Organization

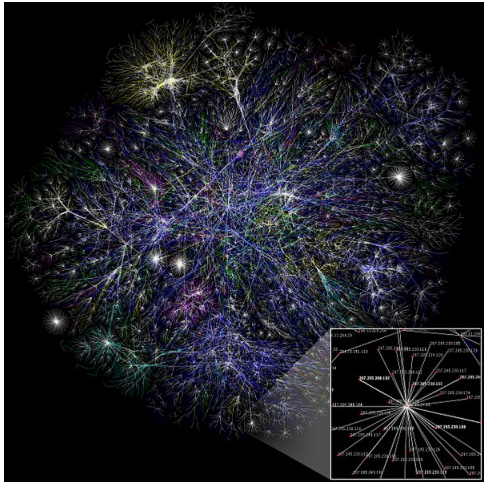


**5. Multi-Agent Network Stability and Performance:** Stability of Network Dynamics, Long-Term Error Dynamics, Performance of Multi-Agent Networks, Benefits of Cooperation, Role of Informed Agents, Adaptive Combination Strategies, Gossip and Asynchronous Strategies, Constrained Optimization, Proximal Strategies, ADMM Strategies, Clustering.

# Motivation#1



**Course deals with the topic of information processing over graphs.**



Internet map (2005). Wikimedia commons.

→ Multi-agent networks for the distributed solution of adaptation, learning, and optimization problems from **streaming** data through **localized** interactions.

# Motivation#2

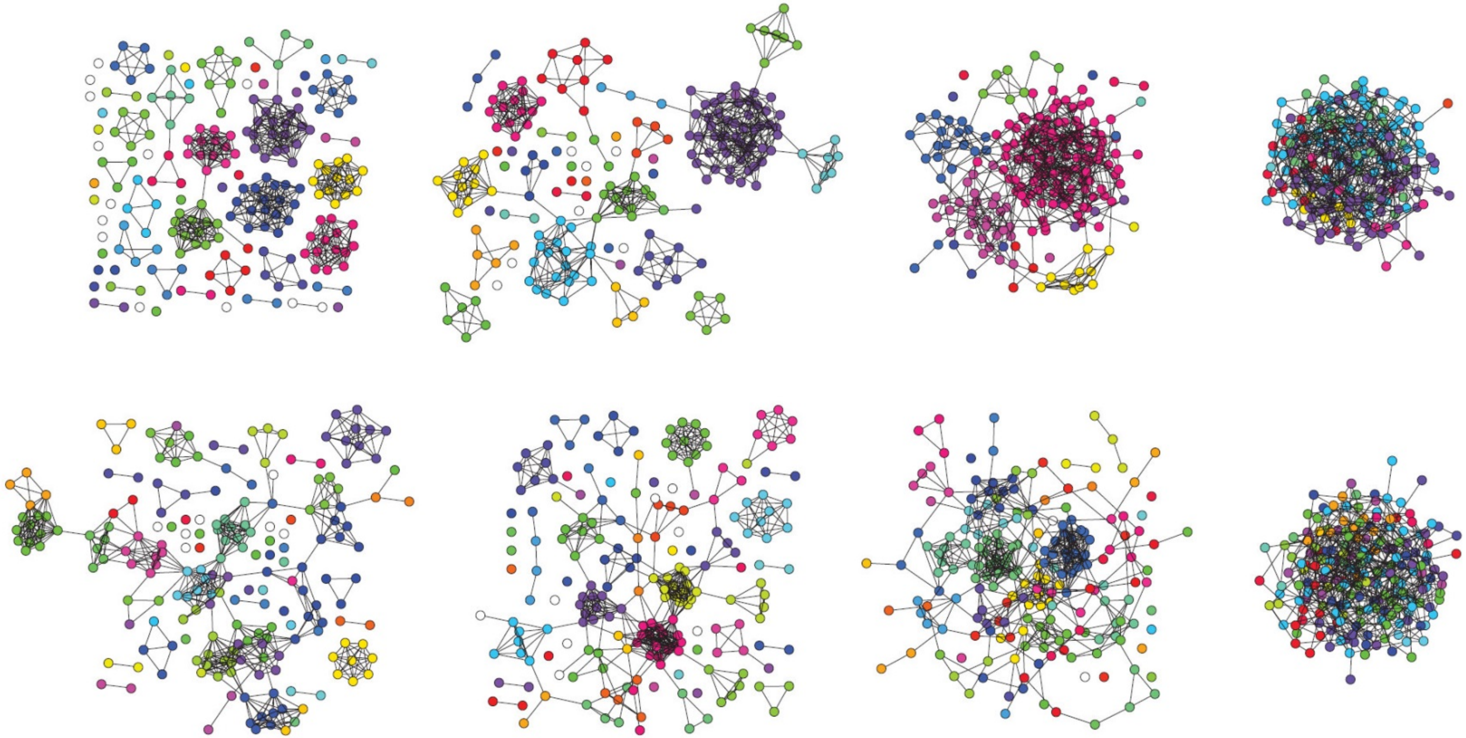


## The results derived here are useful in:

- Comparing network configurations against each other;
- Comparing networks against batch solutions;
- Understanding limits of performance;
- Understanding benefits & pitfalls of cooperation;
- Highlighting interesting phenomena over networks.

# Motivation#3





**Source:** Image from *Agents, Interaction, and Complexity* Research Group website. University of Southampton.

# Motivation#4



**We will examine the analysis and design of networked solutions plus applications in:**

- distributed sensing • intrusion detection • distributed estimation • online learning • pattern classification
- clustering • distributed optimization • multi-agent systems

# Concepts & Examples

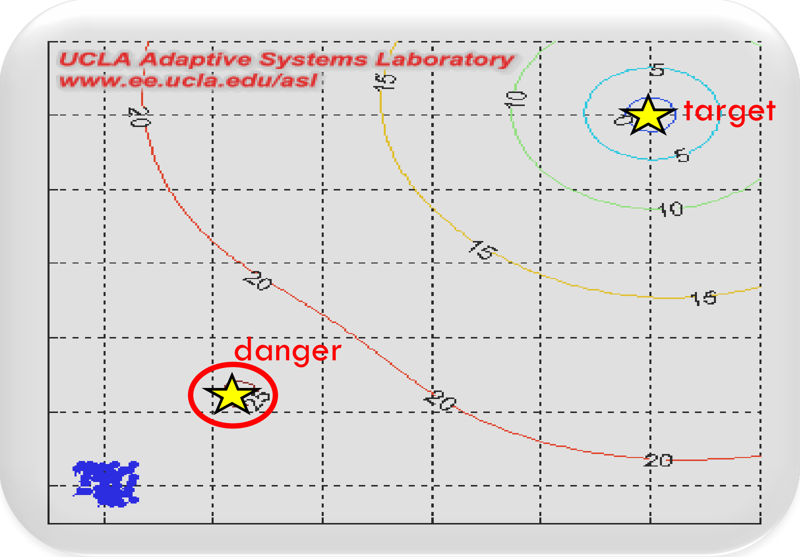


**Proc. IEEE**, vol. 102, no. 4, pp. 460-497, April 2014.
**Foundations and Trends in Machine Learning**, vol. 7, no. 4-5, pp. 311-801, July 2014.

# Distributed Processing



→ **Deals with the discovery of global information from local interactions among dispersed agents.**




## Features:

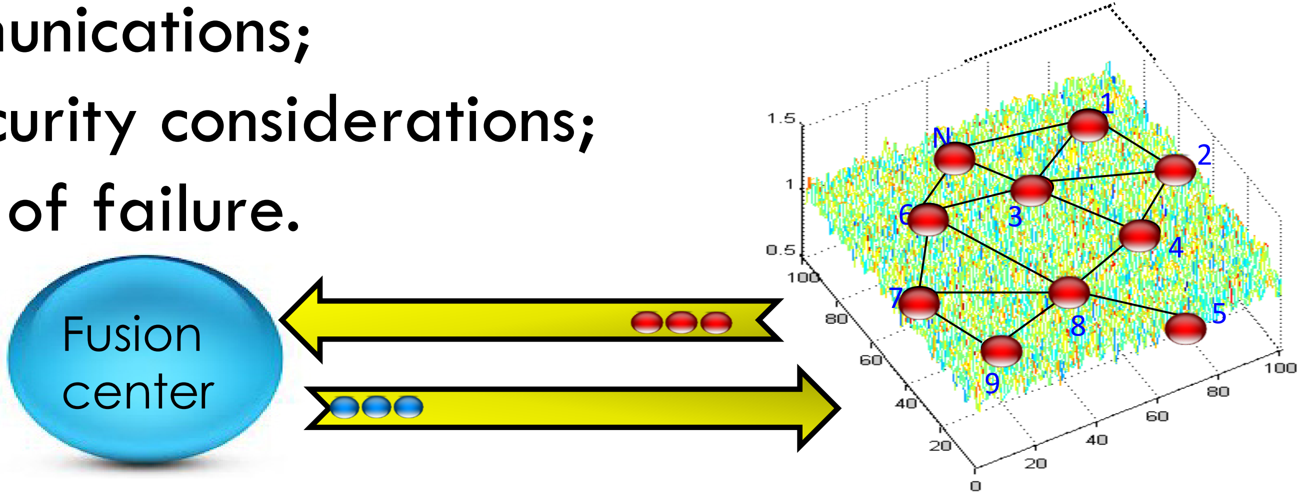
- Common objective(s);
- In-network processing;
- Dispersed agents.

# Centralized Processing



**→ Exchange of data between the dispersed agents and a fusion center.**

- Cost of communications;
- Privacy & security considerations;
- Critical point of failure.



# Why Distributed Processing?



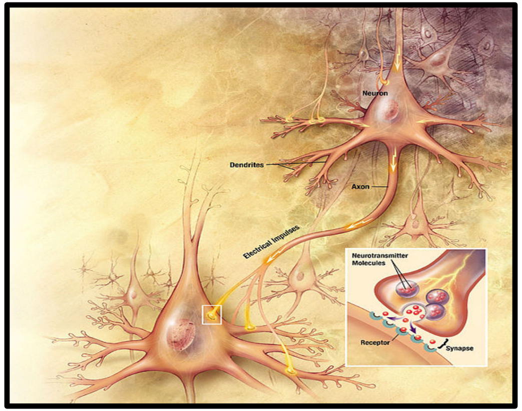
- Data already available at dispersed locations (cloud).
- Power of cooperation → mining of Big Data sets.
- Privacy and security considerations.
- Robustness and resilience (biological networks).
- Robotic swarms (disaster areas).
- Network science (social networks).

**Source:** IEEE SPM May 2013

# Biological Networks



Nature provides splendid examples of real-time decentralized learning & adaptation.




**Source:** Wikimedia.

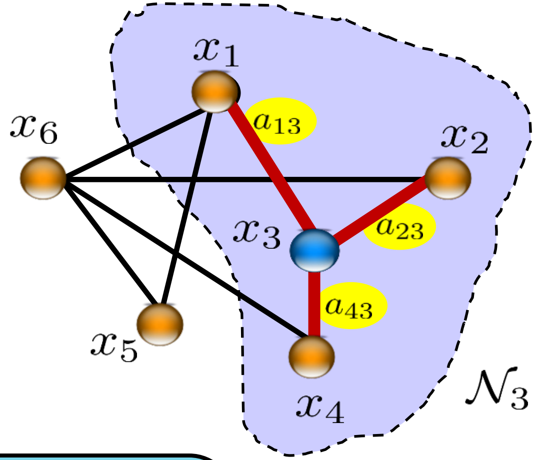
**Source:** Wikimedia; Creative Commons License.

**Source:** Professor S. Pratt Lab, ASU.

# One Useful Result



## Consensus Construction (1974)




Each agent $k$ has a measurement $x_k$.

**Objective:** Compute average value.

$$x_3 \longleftarrow \sum_{\ell \in \mathcal{N}_3} a_{\ell 3}\, x_\ell$$

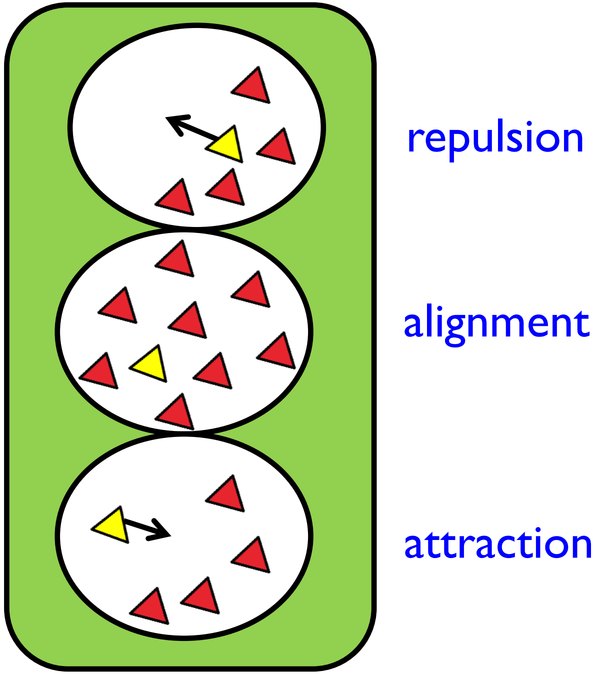
$$\boxed{x_k \longrightarrow \frac{1}{N}\sum_{\ell=1}^{N} x_\ell}$$

# One Useful Application



Technique used in **The Lion King** (1994) and **Batman Returns** (1992) to produce swarming effects.

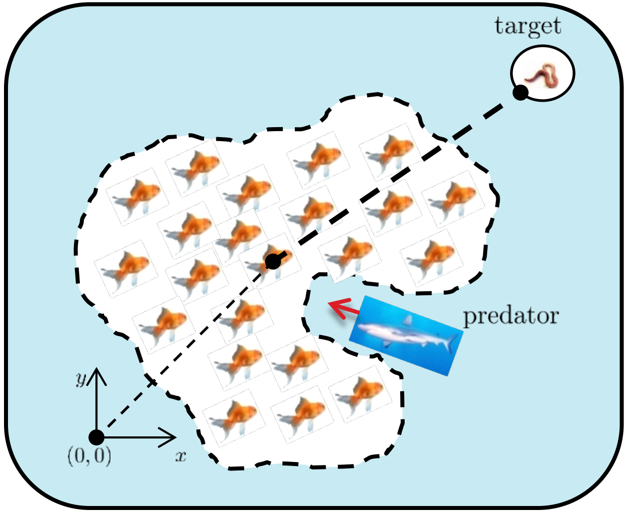

YouTube: https://www.youtube.com/watch?v=2m-42ek85G4



# Issue#1: Cognition



Biological networks have **more complex objectives** such as tracking food sources or evading predators.



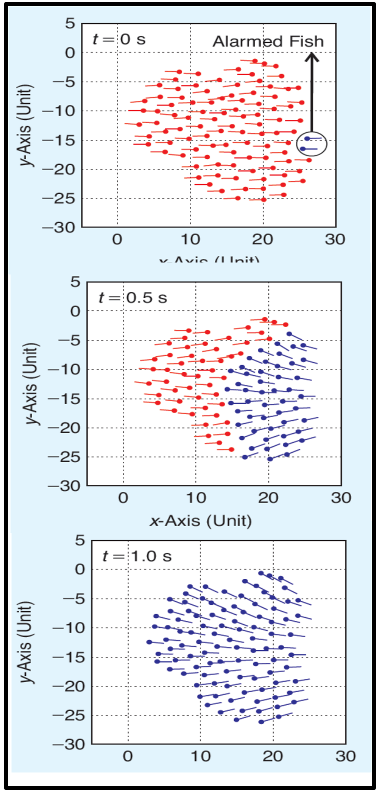

**Source:** http://youtu.be/zvfY8-3ktNA

# Issue#2: Interactions



Interactions are **information-aware** → informed vs. uninformed agents

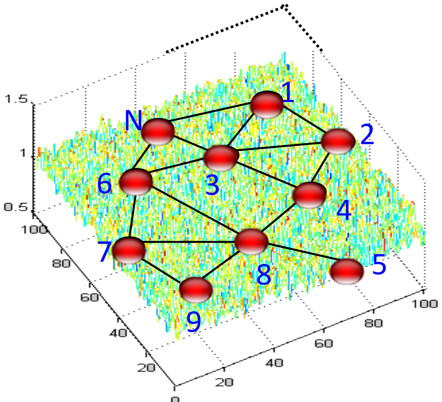
**Source:** Collective Animal Behavior Lab (I. D. Couzin, Princeton University)



# Adaptive Networks



- **Adaptive agents:** learn from streaming data.
- **Cooperative agents:** interact locally.
- **Adaptive topology:** re-wire the graph.
- **Distributed optimization:** solve meaningful problems.



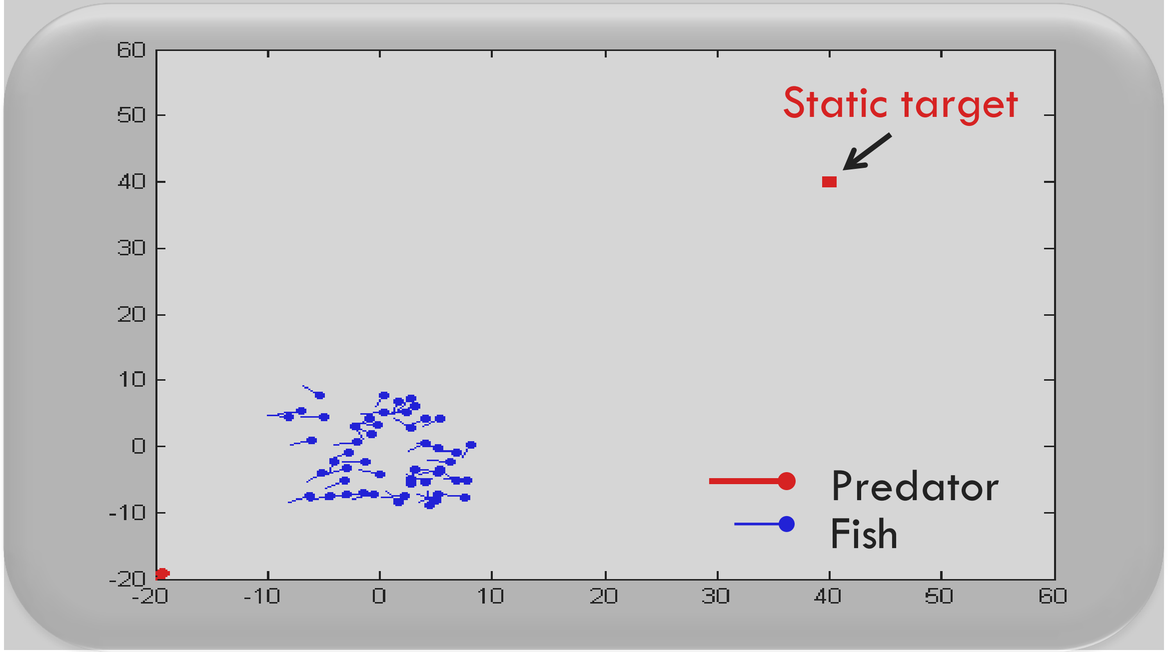

$$
\min_{w} \sum_{k=1}^{N} J_k(w)
$$

$$
\text{subject to} \begin{cases} g_k(w) \leq 0 \\ h_k(w) = 0 \end{cases}
$$

# Example#1: Cooperation



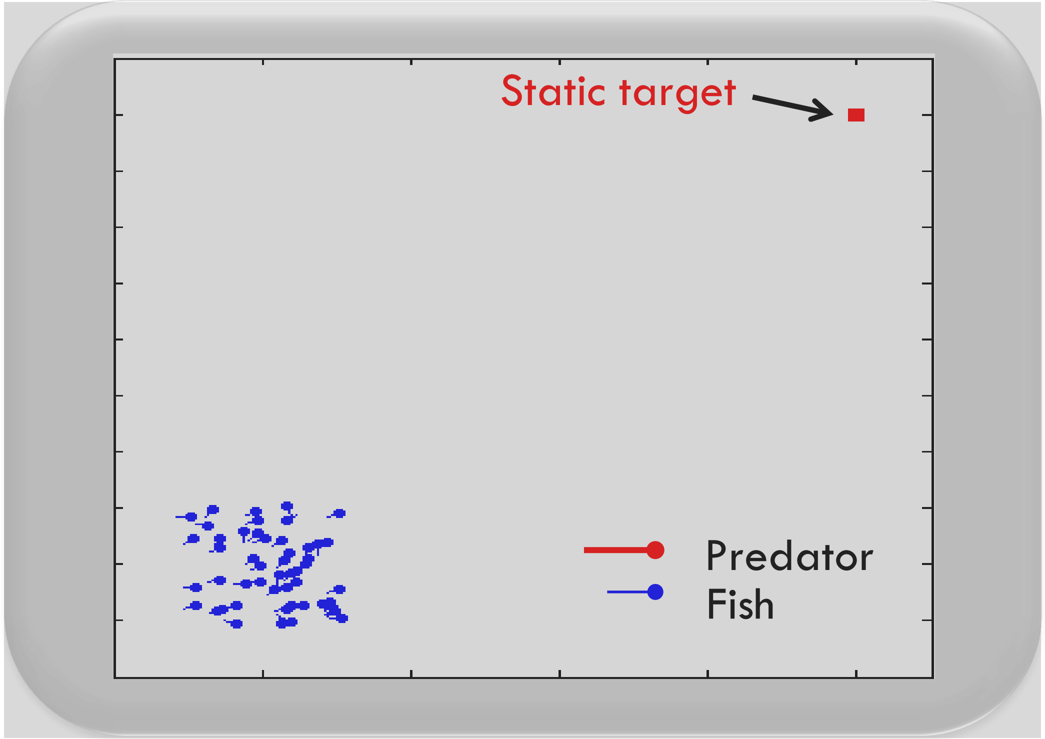
Emulating fish schooling and prey-predator behavior → **Cooperative Networks.**



# Example#2: Competition

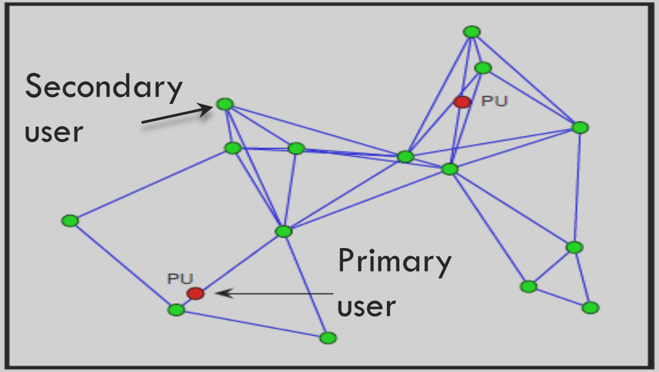


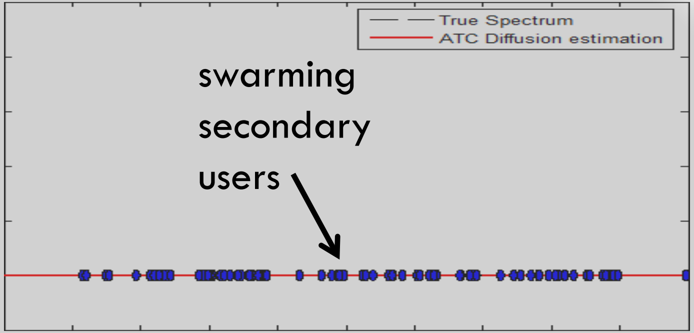
Emulating prey-predator behavior → **Network Competition**

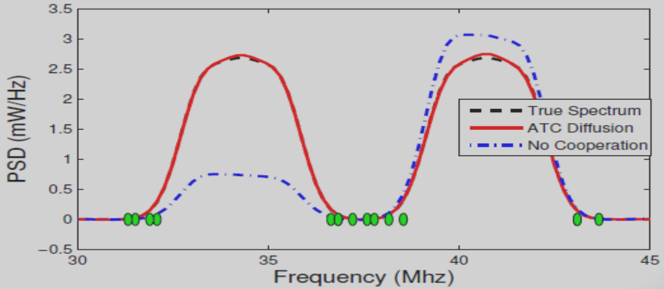


# Example#3:Cognitive Radios



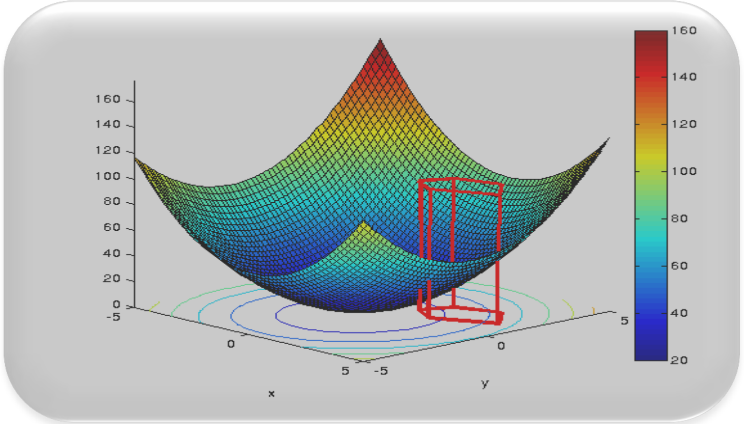


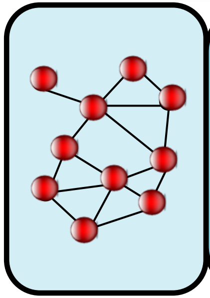



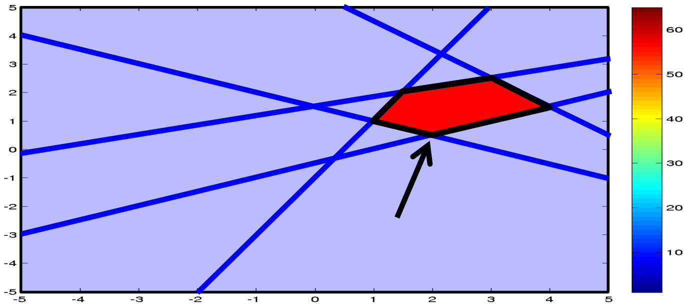



# Example#4: Optimization



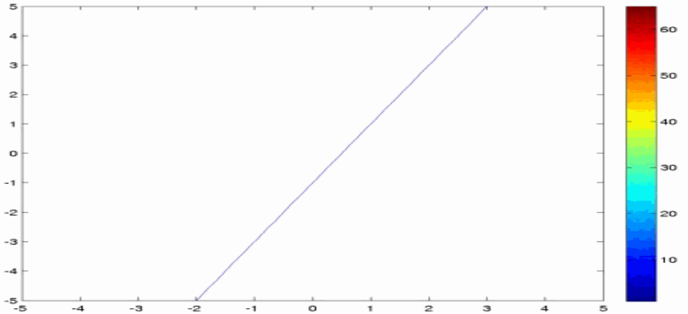




$$\min_{w} \sum_{k=1}^{N} J_k(w)$$

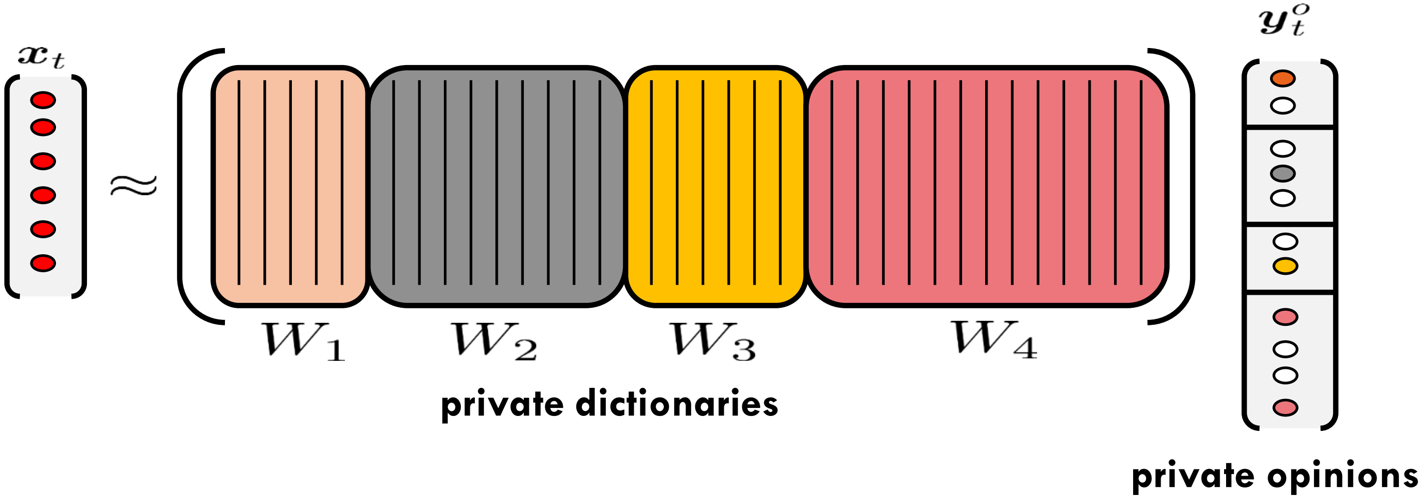
$$\text{subject to} \begin{cases} g_k(w) \leq 0 \\ h_k(w) = 0 \end{cases}$$



# Example#5: Dictionary Learning

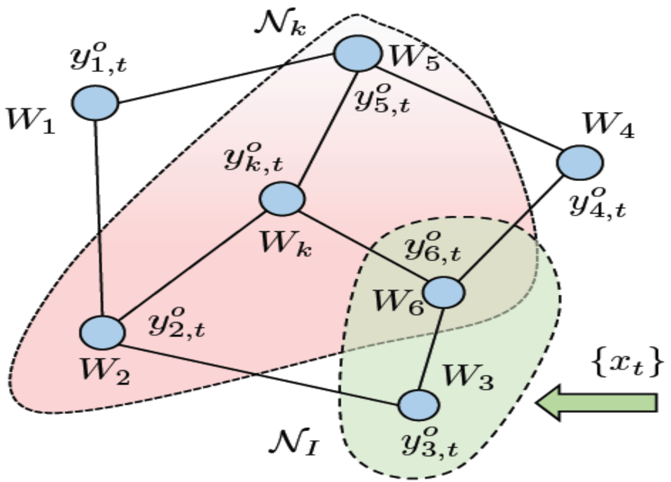


Sparse representation of a signal $\boldsymbol{x}_t \in \mathbb{R}^M$ using **atoms** from a dictionary, $W \in \mathbb{R}^{M \times K}$.



# Example#5: Dictionary Learning






**(distributed dictionary learning)**

$$\boldsymbol{x}_t \approx \sum_{k=1}^{N} W_k \boldsymbol{y}^o_{k,t}$$

$$\begin{aligned} W &= \begin{bmatrix} W_1 & W_2 & \ldots & W_N \end{bmatrix} \\ \boldsymbol{y}^o_t &= \text{col}\{\boldsymbol{y}^o_{1,t}, \boldsymbol{y}^o_{2,t}, \ldots, \boldsymbol{y}^o_{N,t}\} \end{aligned}$$

# Relevant Questions



- What strategies enable distributed learning?
- How to ensure stable behavior?
- What are the limits of performance?
- Can we match centralized processing?
- Does cooperation always help?
- Does it help to have more agents?

# End of Lecture



**Proc. IEEE**, vol. 102, no. 4, pp. 460-497, April 2014.
**Foundations and Trends in Machine Learning**, vol. 7, no. 4-5, pp. 311-801, July 2014.