Communication-Aware and Decentralized Strategic Learning in Networked Multiagent Systems

Ceyhun Eksin

TEXAS A&M UNIVERSITY

Wm Michael Barnes '64 Department of Industrial & Systems Engineering

September 6, 2021

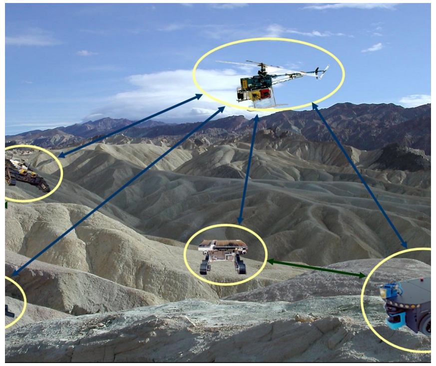
netmas.engr.tamu.edu

Networked systems are everywhere

Autonomous Robots

Unmanned vehicles

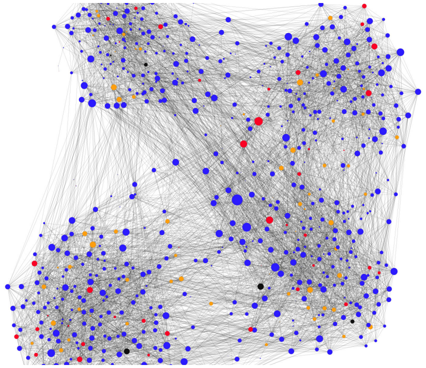
Search & Rescue



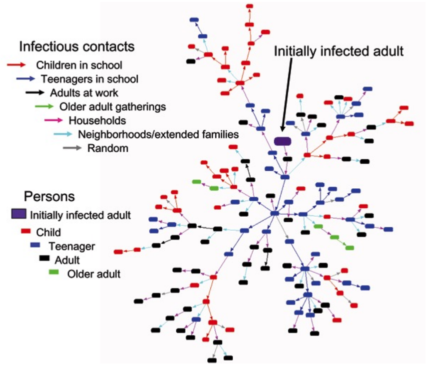
Swarms

Epidemics

Contact network

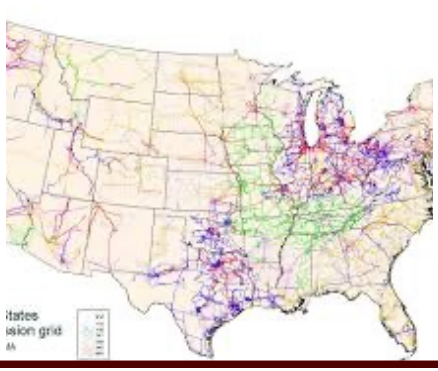


STD

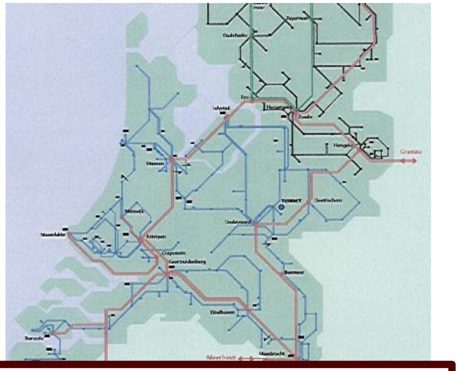


Energy Systems

US grid



Netherlands grid

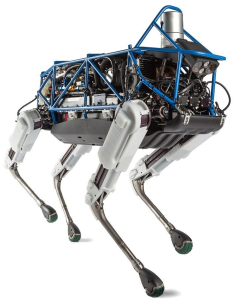


Modeling and designing decentralized decision making

Autonomous robots – teams and collaborations

- Ranging size and capabilities
  - Scale cm to meters
  - UAVs, terrain locomotion
- Wide variety of objectives
  - Active sensing, containment, surveillance, formation
- Vision:
  - Plan, learn and coordinate in novel environments

DesertRhex - KodLab



Wildcat - Boston Dynamics

Swarm of 1K robots 2D shapes

Fundamental challenges in autonomous teams

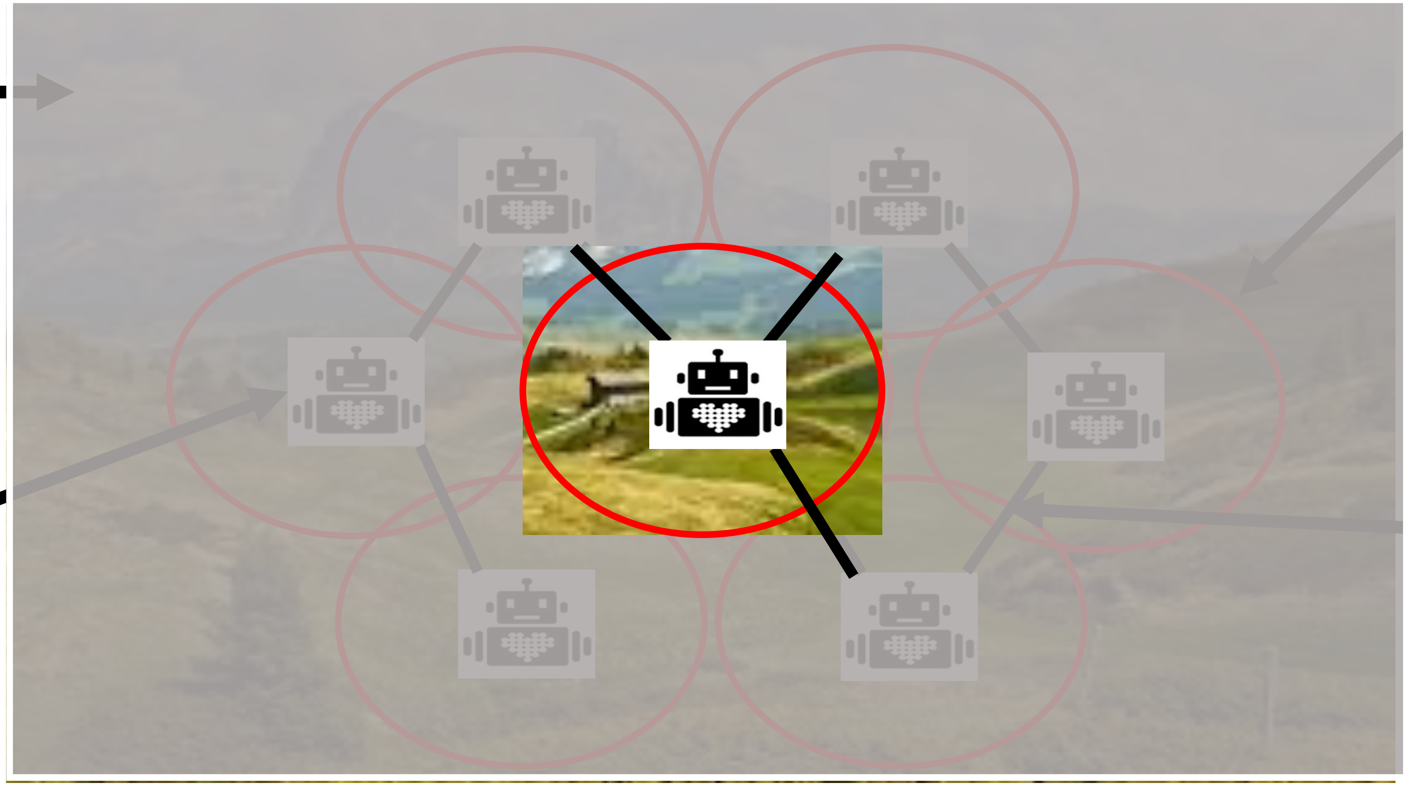
Vision: Plan, learn and coordinate in novel environments

Environment:
Actor's surrounding influencing goals

Topology

Actor:
Decision-maker with a goal

Robot



Partial information:
Noisy information about the environment

Local sensors

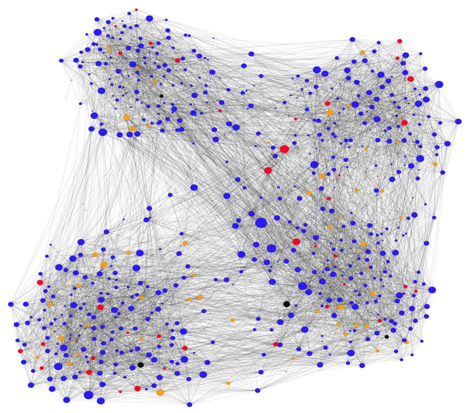
Local information:
Information exchange and observations

Communication

Common challenge in networked multiagent systems with uncertainty

Actors given partial information reason about information & goals of others

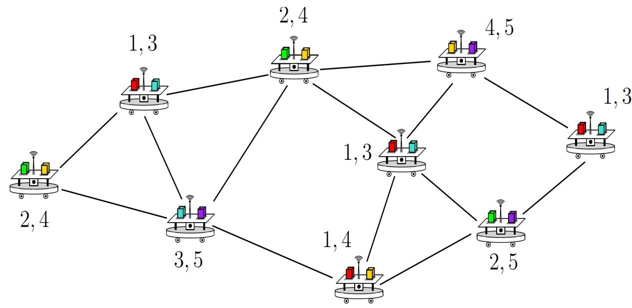
Today's theme: communication-aware game-theoretic algorithms



Networked multi-agent systems

1. Bayesian network games

[IEEE SPM 2013, IEEE TSP 2014, OR 2015]



Communication-aware autonomous teams

2. DFP with voluntary communication

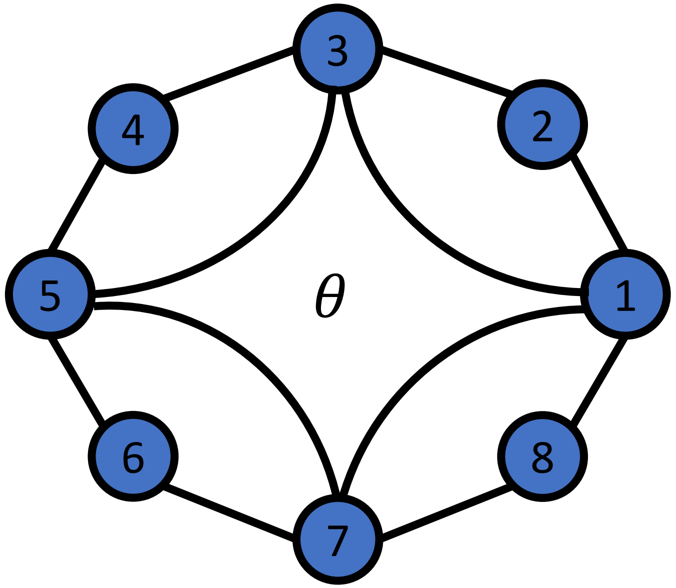
[IEEE TAC 2018, IEEE TAC 2018, Aydin & Eksin 2020]

DFP in unknown environments

3. Learning in near-potential games

[Aydin et al 2021]

Networks and uncertain environments



- Graph with $N = 8$ agents and connections
- Agent 1 has 4 neighbors

$$n(1) = (2,3,7,8)$$

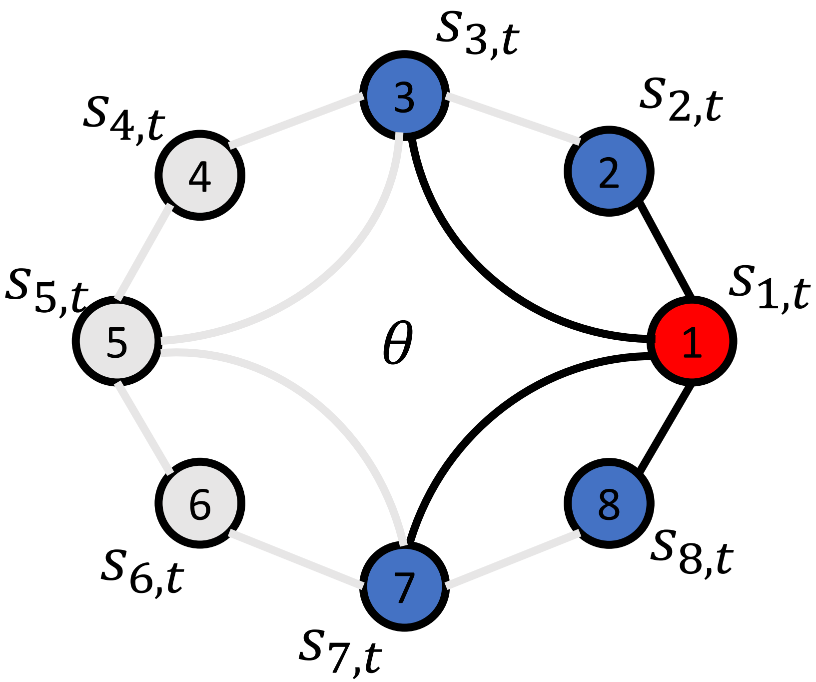
Time-varying, and random graphs

- Environment θ
- Agent 1 has signal s_1 about environment, e.g.,

$$s_1 = \theta + \epsilon_1$$

with noise ϵ_1

Interactive decision-making problem



- Repeated decision-making $a_{1,t}$
- Agents act with respect to objectives (payoff)

$$u_1(a_{1,t}, a_{-1,t}, \theta)$$

  - Payoff depends on actions of others $-1 = N \setminus 1$

- Messages from neighbors $m_{2,t}, m_{3,t}, m_{7,t}, m_{8,t}$
- Information of agent 1

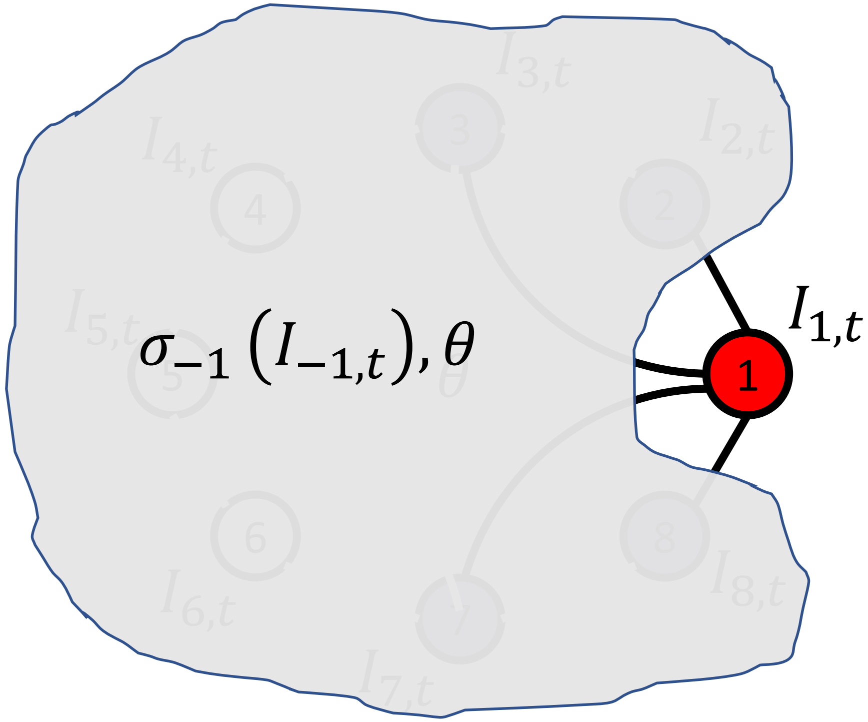
$$I_{1,t+1} = \{I_{1,t}, s_{1,t}, m_{n(1),t}\}$$

$$a_{1,t+1} = \arg\max_{a_1} E\left[u_1(a_1, a_{-1,t+1}, \theta) \mid I_{1,t+1}\right]$$

- Strategic reasoning by considering motives of others based on information ➡ Model

Game theory in networked systems

- Game theory is concerned with decision-making of self-interested individuals



- Information of agent 1: $I_{1,t+1} = \{I_{1,t}, s_{1,t}, m_{n(1),t}\}$

- Strategy of agent 1

  $\sigma_{1,t}\left(I_{1,t}\right): Information \rightarrow Action$

- Strategy profile of all agents

  $\sigma := \{\sigma_{i,t}\}_{i=1,\dots,N,t=1,\dots,\infty}$

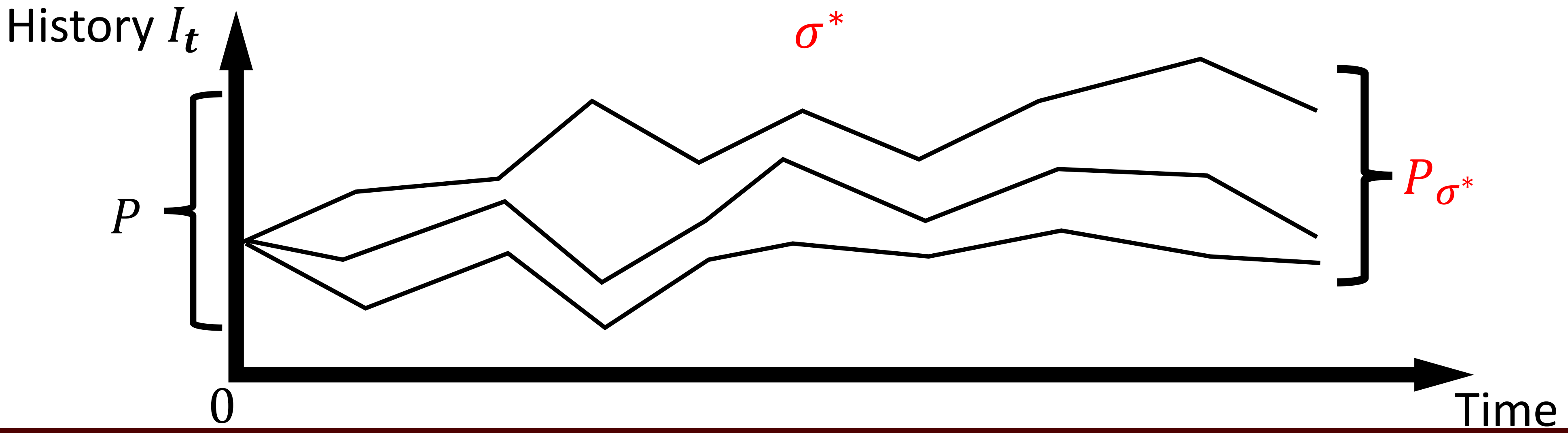
- Agents have different information

  $I_{1,t} \neq I_{2,t}$

- Reason about the information of others based on strategy profile σ

Equilibrium behavior

- Space of play $\Omega = \Theta \times (S \times A)^{\mathbb{N}}$
- Prior P on $\Theta \times S^{\mathbb{N}}$ induce a distribution on Ω with strategy profile σ



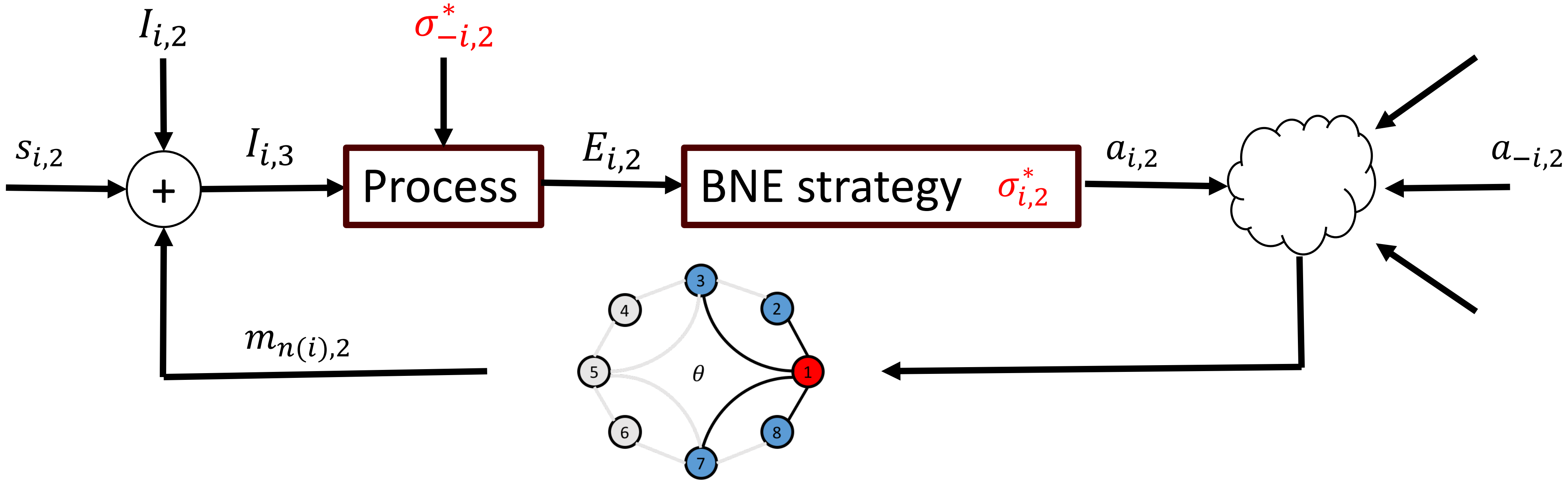
Definition (Bayesian-Nash equilibrium)

Strategy profile σ^* and beliefs $P_{\sigma^*}(\cdot | I_{i,t})$ that satisfy

$$E_{\sigma^*}\left[u_i(\sigma^*_{i,t}, \sigma^*_{-i,t}, \theta) \middle| I_{i,t}\right] \geq E_{\sigma^*}\left[u_i\left(a_i, \sigma^*_{-i,t}, \theta\right) \middle| I_{i,t}\right] \quad I_{i,t} \in \mathcal{J}_i, \forall i$$

BNE is optimal with respect to strategies of other agents given local information.

Equilibrium behavior – block diagram



Bayesian Network Games (BNG)

Asymptotic learning in BNG

- **What happens when individuals act according to BNE (optimal) at each step?**

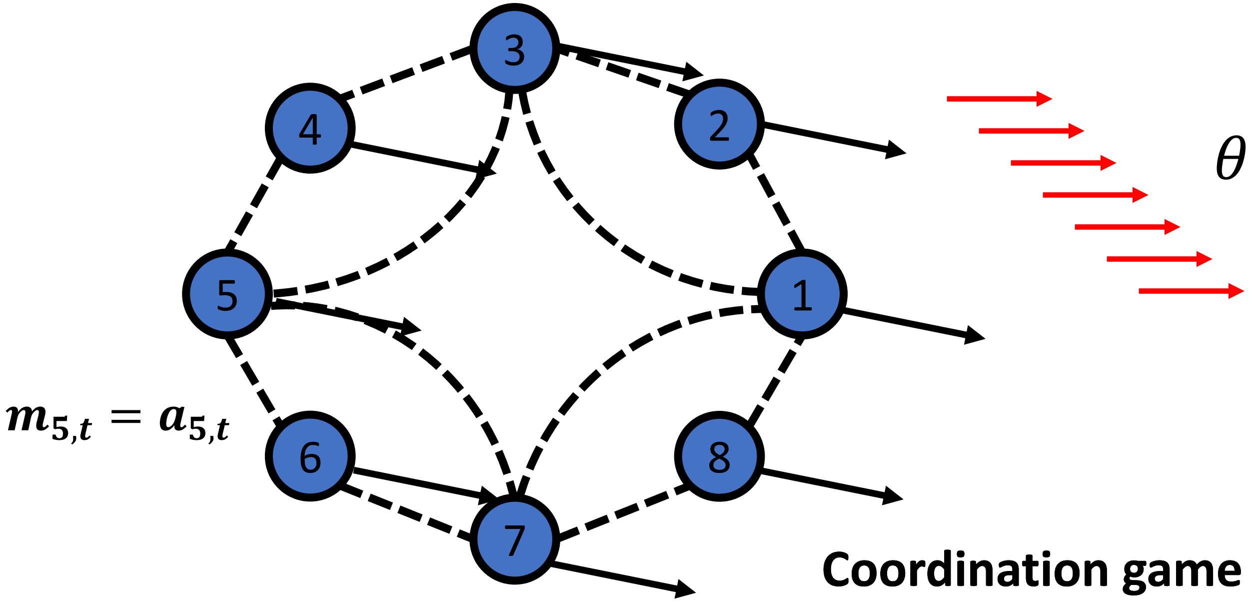
Theorem (w/Molavi, Ribeiro, Jadbabaie, *Operations Research* 2015)

If agents play **a coordination game** and **message their actions**, then they

a) Consensus: $\lim_{t\to\infty} \sigma^*_{i,t} - \sigma^*_{j,t} = 0$ almost surely

b) Information aggregation: $\lim_{t\to\infty} E_{\sigma^*}[\theta | I_{i,t}] = E_{\sigma^*}[\theta | I_{\infty}]$ almost surely

- Reaching consensus in actions implies
  - Ex-ante identical payoffs
  - $E_{\sigma^*}\left[\theta \middle| I_{i,t}\right] = E_{\sigma^*}[\theta | I_{j,t}]$
- Not necessarily aggregate information
  - True generically



Coordination game

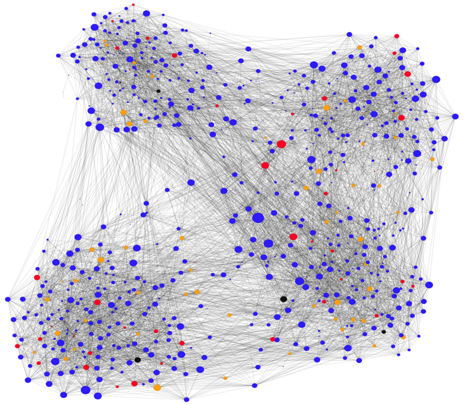
Bayesian social learning: Borkar & Varaiya 1982 IEEE TAC, Jadbabaie et al 2012 GEB, Jackson & Kalai 1997 GEB

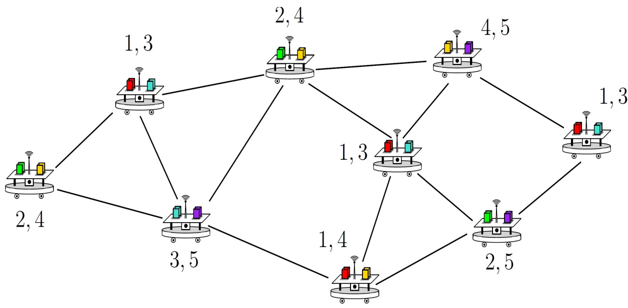
Communication-aware game-theoretic learning algorithms

Networked multi-agent systems

1. Bayesian network games

[IEEE SPM 2013, IEEE TSP 2014, OR 2015]





Communication-aware autonomous teams

2. DFP with voluntary communication

[IEEE TAC 2018, IEEE TAC 2018, Aydin & Eksin 2020]

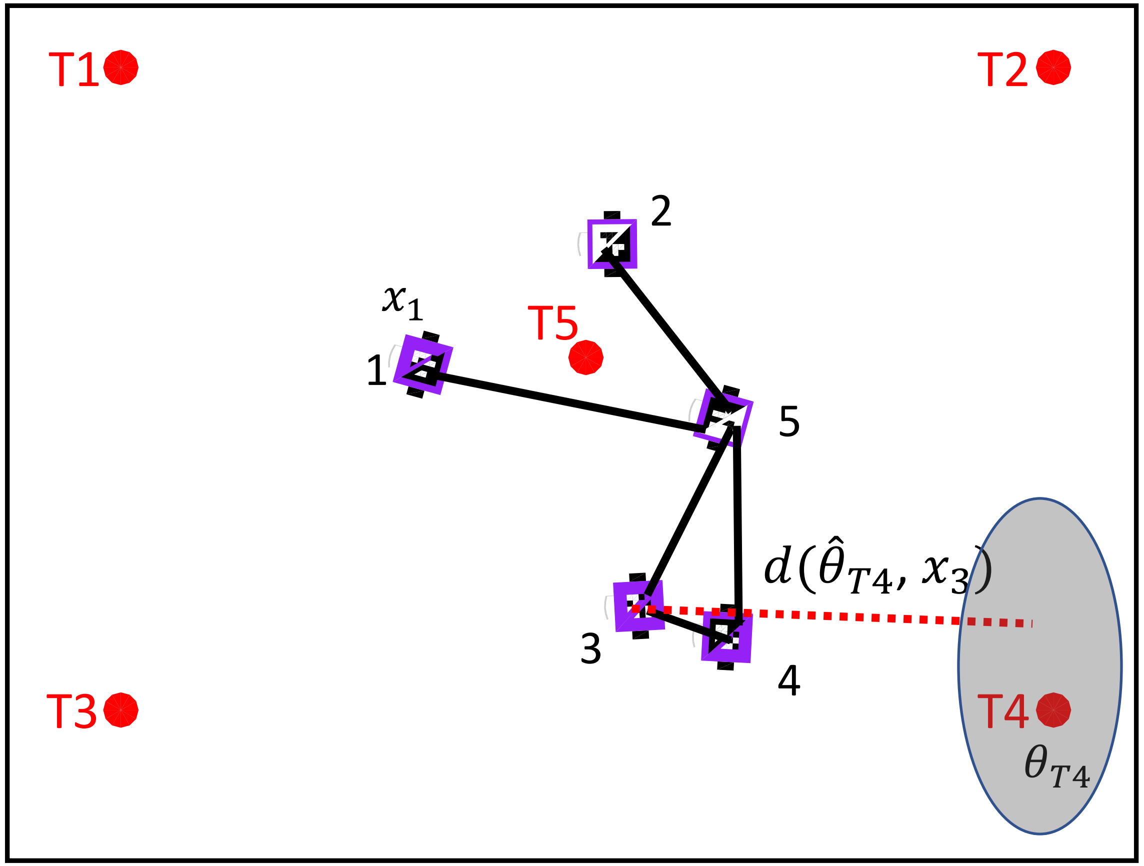
Persisting uncertainty among teams

3. Learning in Near Potential Games

[Aydin et al 2021]

Task assignment problem

- Communication network of 5 robots
- 5 Targets



- Global Objective:

  Cover all targets with minimum effort

$$u_i(\mathrm{T}, others, \theta) = \begin{cases} 0 & \text{if another pick T} \\ d(\theta_T, x_i)^{-1} & \text{o.w.} \end{cases}$$

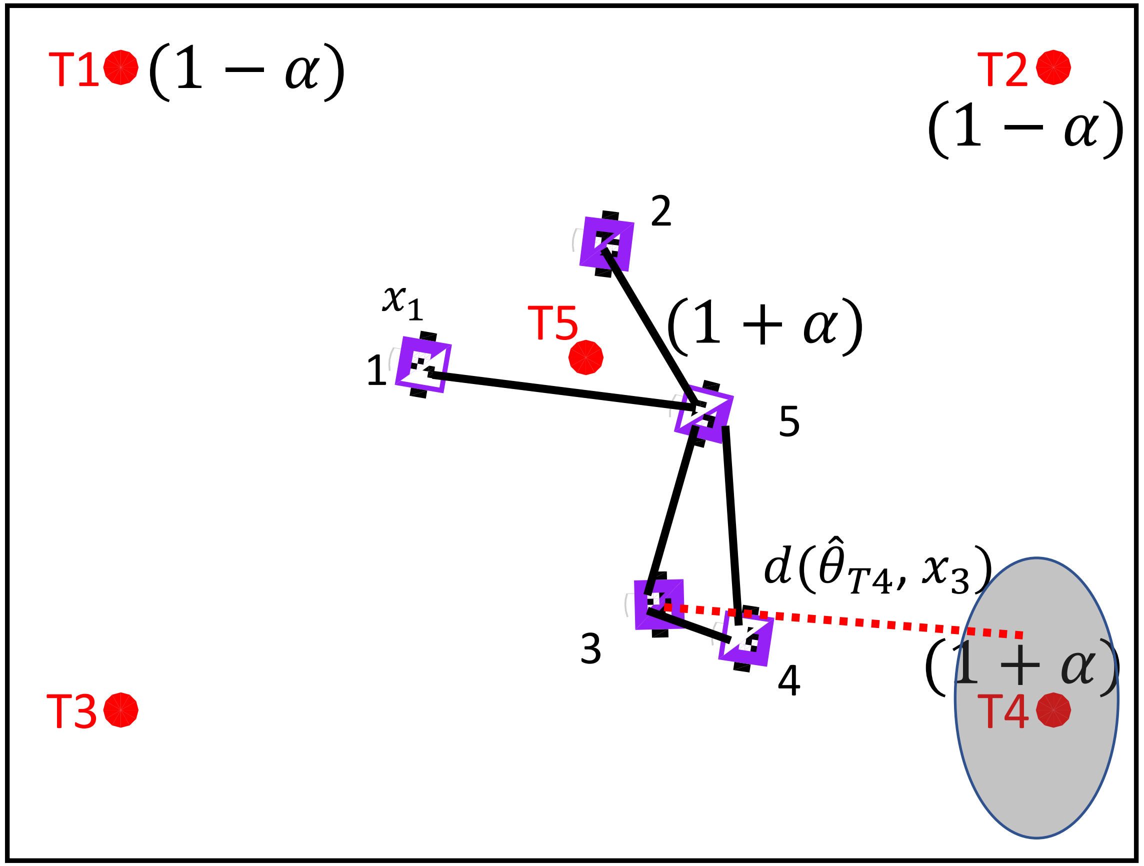
- State θ is the location of targets
- Noisy information about the locations

$$signal = f(\theta_{T1}, noise)$$

Arslan, Marden & Shamma 2007 ASME JDSMC
Richards, Bellinghem, Tillerson & How 2002 AIAA GNC

Task assignment – local algorithm

- Consider agent 3 with location x_3
- Neighbors $n(3) = \{4, 5\}$

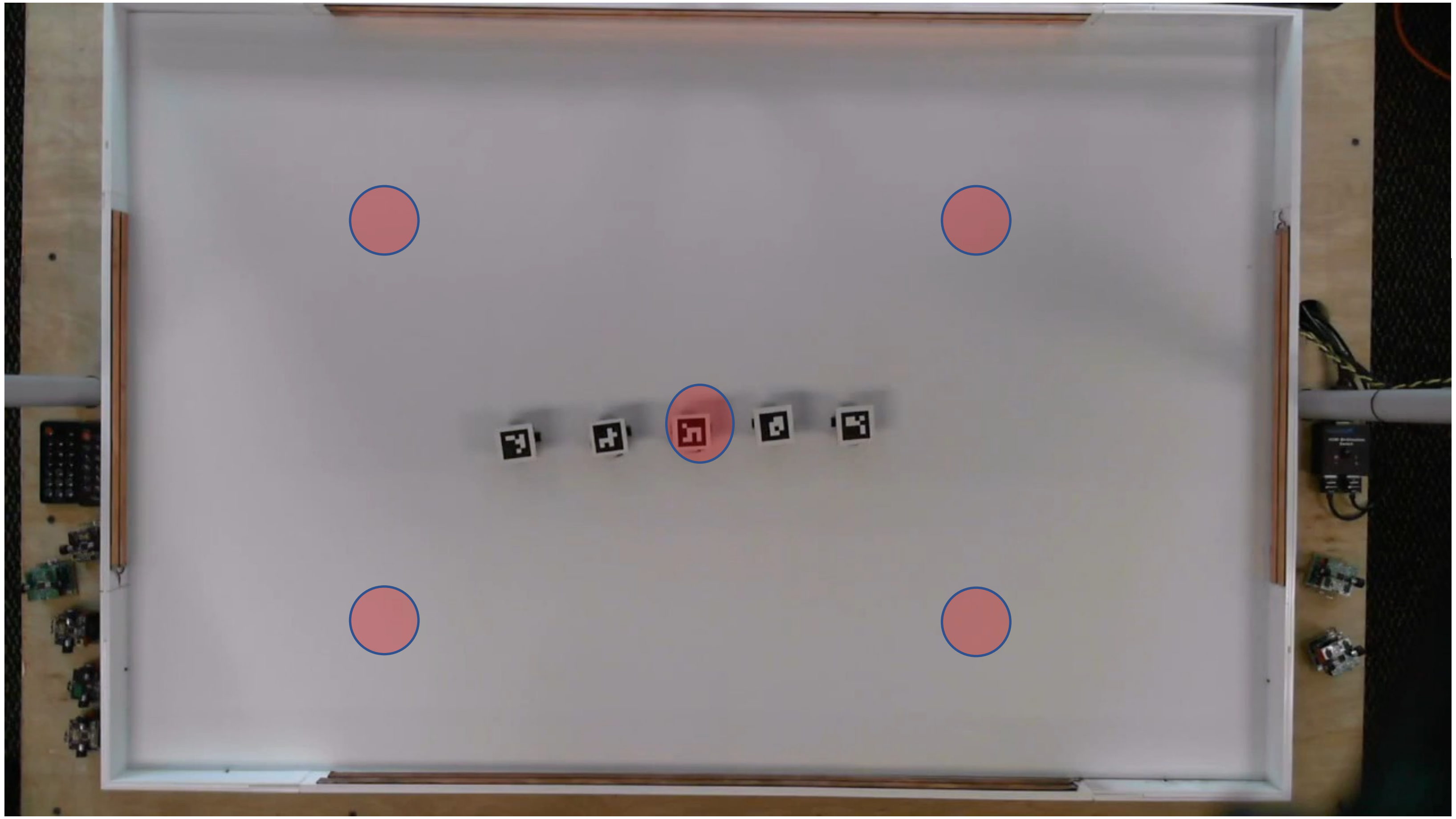


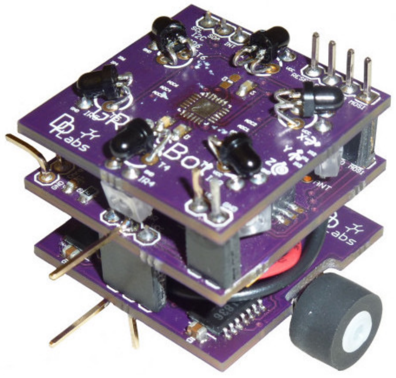
1. Initial estimate *congestion* at each target
2. Subtract congestion caused by self
3. Compute *expected utility* from each T
   $u_i(T4, estimate, \hat{\theta}) = d(\hat{\theta}_{T4}, x_3)^{-1}$
4. Pick target $a_{3,1}$= T4
5. Move in the estimated direction
6. Update congestion caused at each target
   $cong_{i,t}(T4) = \alpha \times cong_{i,t}(T4) + (1-\alpha)$
7. Share congestion estimate with $n(3)$
8. Update congestion estimates averaging
9. Update location estimates $\hat{\theta}$
10. Go to step 2 and repeat

Decentralized Fictitious Play (D-FP) algorithm

Decentralized fictitious play in action

- Common initial estimates
- Gaussian independent signals about target locations





GRITSBot
[Pickem et al.,2015]

- Underlying communication network, and constant (successful) communication attempts

Standard fictitious play with inertia

- Agents assume other agents use stationary strategies (even though they are not)
- Empirical frequency of past actions: $f_{j,t} \in \Delta A$, A is the set of finite actions

$$\sigma_{j,t} := f_{j,t} = (1-\alpha) f_{j,t-1} + \alpha \Phi(a_{j,t})$$

- Fading constant: α (can depend on time, e.g., $\alpha = \frac{1}{t}$)

- Mapping from action to the probability over the action space: $\Phi(a_{j,t}) \in \Delta A$

- Expected utility of agent given the empirical frequencies of other agents $f_{-i,t} = \{f_{j,t}\}_{j \in N \setminus i}$

$$u_i(a, f_{-i,t}, \theta) = \sum_{a_{-i}} u_i(a, a_{-i}, \theta) f_{-i,t}(a_{-i})$$

- Best respond with inertia: $a_{i,t} = \begin{cases} \arg\max_{a \in A} u_i(a, f_{-i,t}, \theta) & w.prob.\ 1-\rho \\ a_{i,t-1} & w.prob.\ \rho \end{cases}$

Standard fictitious play with inertia

- Agents assume other agents use stationary strategies (even though they are not)
- Empirical frequency of past actions: $f_{j,t} \in \Delta A$, A is the set of finite actions

$$\sigma_{j,t} := f_{j,t} = (1-\alpha) f_{j,t-1} + \alpha \Phi(a_{j,t})$$

  - Fading constant: α (can depend on time, e.g., $\alpha = \frac{1}{t}$)
  - Mapping from action to the probability over the action space: $\Phi(a_{j,t}) \in \Delta A$

- Expected utility of agent given the empirical frequencies of other agents $f_{-i,t} = \{f_{j,t}\}_{j \in N \setminus i}$

$$u_i(a, f_{-i,t}, \theta) = \sum_{a_{-i}} u_i(a, a_{-i}, \theta) f_{-i,t}(a_{-i})$$

- State and empirical frequencies are unknown

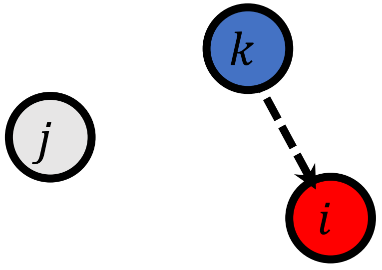
Decentralized-FP with inertia

- Agent i's *belief* of agent j's strategy: $f^i_{j,t} \in \Delta A$
- Agent i's *belief* of others' future actions: $f^i_{-i,t}$
- Agent i's *belief* of the environment state θ: $\hat{\theta}^i_t$

1) **Inertial best-response**: $$a_{i,t} = \begin{cases} \arg\max_{a \in A} u_i\left(a, f^i_{-i,t-1}, \hat{\theta}^i_{t-1}\right) & w.prob.\ 1-\rho \\ a_{i,t-1} & w.prob.\ \rho \end{cases}$$

2) **Update local empirical frequency**: $$f_{i,t} = (1-\alpha) f_{i,t-1} + \alpha \Phi(a_{i,t})$$

3) **Share beliefs** $\{f^i_{j,t-1}\}_{j \in N}$ where $f_{i,t} = f^i_{i,t-1}$ with current neighbors



Update agent i's belief on j's frequency: $$f^i_{j,t} = \sum_{k \in N} w^i_{j,k}(t) f^k_{j,t}$$

- Agent i's weight on k's belief on j's frequency: $w^i_{j,k}(t)$

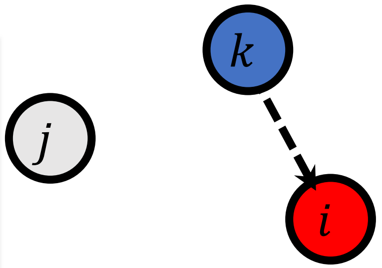
Decentralized-FP with inertia

- Agent i's *belief* of agent j's strategy: $f^i_{j,t} \in \Delta A$
- Agent i's *belief* of others' future actions: $f^i_{-i,t}$
- Agent i's *belief* of the environment state θ: $\hat{\theta}^i_t$

1) **Inertial best-response**: $$a_{i,t} = \begin{cases} \arg\max_{a \in A} u_i\left(a, f^i_{-i,t-1}, \hat{\theta}^i_{t-1}\right) & w.prob.\ 1-\rho \\ a_{i,t-1} & w.prob.\ \rho \end{cases}$$

2) **Update local empirical frequency**: $$f_{i,t} = (1-\alpha) f_{i,t-1} + \alpha \Phi(a_{i,t})$$

Action-based learning in games:
Marden, Arslan & Shamma 2009 IEEE TAC, Monderer & Shapley 1996 JET
Distributed convex optimization:
Nedic & Ozdaglar 2009 TAC, Chen & Sayed 2012 TSP



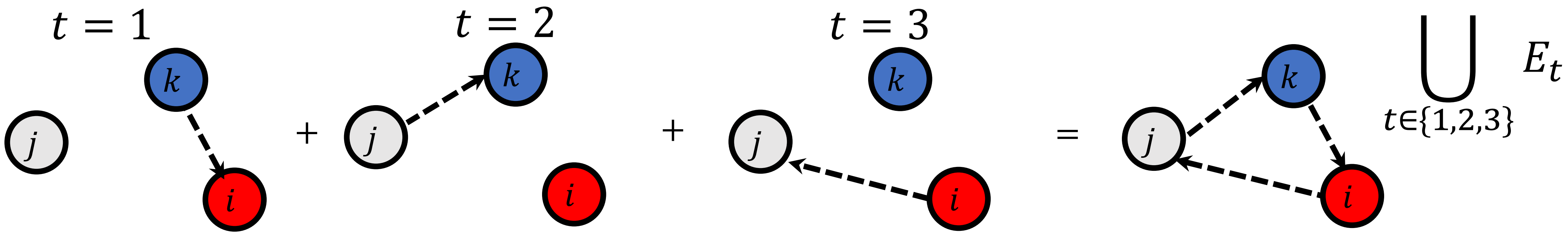
- Agent i's weight on k's belief on j's frequency: $w_{j,k}(t)$

DFP converges in potential games

Theorem (w/Swenson, Kar, Ribeiro, *IEEE TAC* 2018; Arefizadeh, Asilomar 2019)

1) *Potential games (a function $u(\cdot)$ that captures incentives of agents)*
2) *Asymptotically **weakly agree** on the state estimate $(\hat{\theta}_t^i \xrightarrow{w} \hat{\theta})$,*
3) *Time-varying network:*
   a. *Union of edges over a finite horizon constitute a strongly connected network*
   b. *Weights sum to 1 (row stochastic), and positive only when neighbors*

⟹ $\{a_t\}_{t\geq 1}$ ***converges to a pure-strategy Nash equilibrium almost surely***.



- Union of edges over times **1, 2 and 3** yields a connected network!
- **Information needs to be able to travel from one node to another in finite time**

DFP converges in potential games

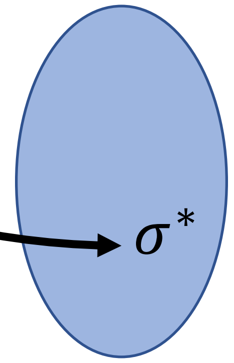
Theorem (w/Swenson, Kar, Ribeiro, *IEEE TAC* 2018; Arefizadeh, Asilomar 2019)

1) *Potential games (a function $u(\cdot)$ that captures incentives of agents)*
2) *Asymptotically **weakly agree** on the state estimate ($\hat{\theta}_t^i \xrightarrow{w} \hat{\theta}$),*
3) *Time-varying network:*
   a. *Union of edges over a finite horizon constitute a strongly connected network*
   b. *Weights sum to 1 (row stochastic), and positive only when neighbors*

$\Rightarrow$ $\{a_t\}_{t\geq 1}$ ***converges to a pure-strategy Nash equilibrium almost surely.***

	$t=1$	2	3	…	$t=\infty$
1	A	A	B		$a_{1,\infty}=B$
2	B	A	B	…	$a_{2,\infty}=B$
3	B	B	B		$a_{3,\infty}=B$
4	B	B	A		$a_{4,\infty}=A$

Nash equilibria $K(\hat{\theta})$



Any covering is a NE

DFP converges in potential games

Theorem (w/Swenson, Kar, Ribeiro, *IEEE TAC* 2018; Arefizadeh, Asilomar 2019)

1) *Potential games (a function $u(\cdot)$ that captures incentives of agents)*
2) *Asymptotically **weakly agree** on the state estimate ($\hat{\theta}_t^i \xrightarrow{w} \hat{\theta}$),*
3) *Time-varying network:*
   a. *Union of edges over a finite horizon constitute a strongly connected network*
   b. *Weights sum to 1 (row stochastic), and positive only when neighbors*

$\Rightarrow$ $\{a_t\}_{t\geq 1}$ ***converges to a pure-strategy Nash equilibrium almost surely***.

	$t=1$	2	3	$\cdots$	$t=\infty$
1	A	A	B		$a_{1,\infty}=B$
2	B	A	B	$\cdots$	$a_{2,\infty}=B$
3	B	B	B		$a_{3,\infty}=B$

Nash equilibria $K(\hat{\theta})$

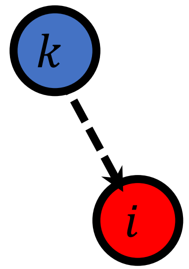
σ^*

Each agent communicates its estimates to all its neighbors at each time step.

is a NE

Toward communication efficient decentralized algorithms

- Communicating at each step is costly (energy, bandwidth)
  - In DFP, there are $N \times |n(i)|$ messages shared at each step
  - Each message is matrix with $N \times |A|$ values
- Communication may be subject to failures, thus it may not be possible
  - Wireless communication is subject to failures
- Existing efficient communication protocols in distributed optimization:
  - One-bit communication, Communication censoring
  - Slows down convergence but reduces communication



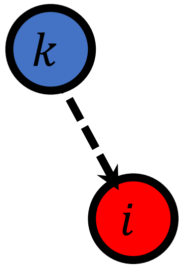
- A general condition for the communication protocol $\Longrightarrow$ convergence of DFP
  - First and second order belief-based communication protocols
- The setup is for random communication networks

Random communication networks

- Random point-to-point communication

$$c_{ki}(t) \sim Bernoulli(p_{ki}(t))$$

where $0 < p_{ki}(t) \leq 1$



- Positive probability of communication between each agent

Algorithm 1 DFP with voluntary communication

Input: ρ, ϵ and communication specific parameters.
for $t = 1, 2, \cdots,$ **do**
    Best-response with inertia
    Communication protocol
    Belief update
end for

DFP with Voluntary Communication

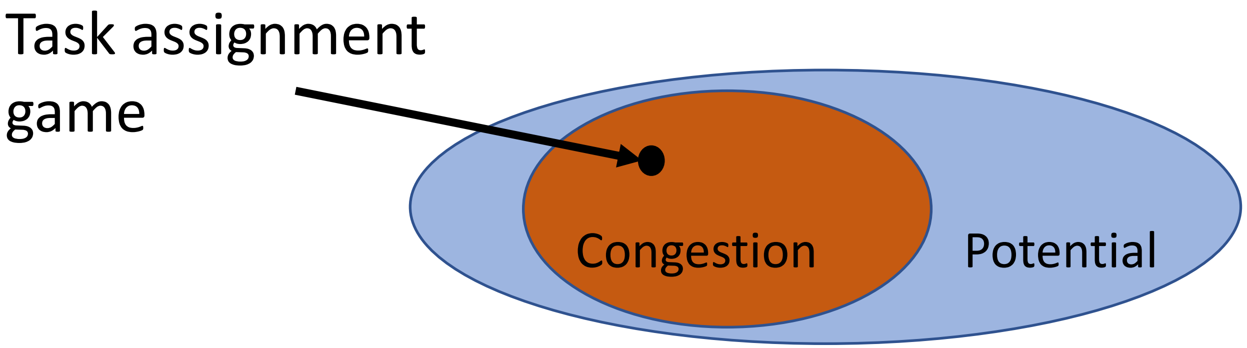
- Agnostic to the specific communication protocol as long as the following condition holds:

Condition (Prediction convergence under static environments)

If agent j repeats an action $a_j \in A$ for $T \geq \bar{T}$ times after $t > 0$, then agent i learns agent j's action a_j with finite probability, i.e. $P\left(\left|a_{j,t+T} - f^i_{j,t+T}\right| \leq \eta \middle| H(t+T)\right) \geq \hat{\epsilon}$.

- Any agent is able to learn the repeated action of another agent with positive probability

Weakly-acyclic games



- Congestion games: m resources, value of a resource decreases with number of players using
- Autonomous systems with global objectives are **potential games**

Definition (Weakly-acyclic games)

A game Γ is weakly acyclic if from any joint action profile, there exists a best-response path ending at a pure Nash equilibrium.

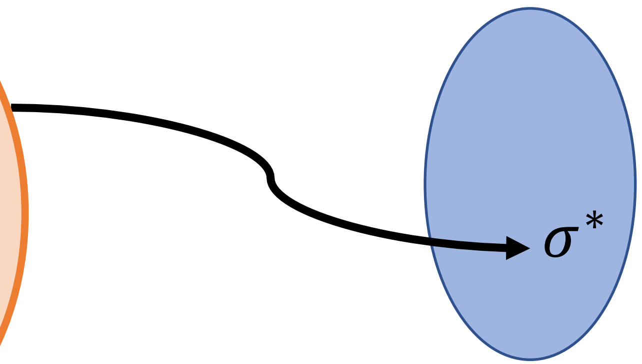
Convergence of DFP type algorithms

Theorem (w/Aydin, *Arxiv* 2020)

Given the prediction convergence under static environments is satisfied, the actions in DFP with voluntary communication converge to a pure NE of any weakly acyclic game almost surely in finite time.

	$t=1$	2	3	…	$t=\infty$
1	A	A	B		$a_{1,\infty}=B$
2	B	A	B	…	$a_{2,\infty}=B$
3	B	B	B		$a_{3,\infty}=B$
4	B	B	A		$a_{4,\infty}=A$

Nash equilibria $K(\hat{\theta})$



Any covering is a NE

Proof method: absorbing Markov Chain

Condition (Prediction convergence under static environments)

If action profile a is repeated T times after $t > 0$ then

$$\arg\max_{\alpha} u_i(\alpha, f^{i}_{-i,t+T}) = \arg\max_{\alpha} u_i(\alpha, a_{-i})$$

Let $a = \{B, B, B, A\}$ be the repeated action profile

	1	2	…	t	t+1	…	$t+T$
1	A	A		B	B		B
2	B	A		B	B		B
3	B	B	…	B	B	…	B
4	B	B		A	A		A

Estimated best response = Best response to actual

Proof method: absorbing Markov Chain

Lemma (positive probability of reaching NE)

Let a_t for $t \geq 1$ be generated by FP algorithm with inertia. For all $t \geq \bar{t}$,

$$P(\text{Pure Nash equilibrium } a^* \text{ is reached starting from time } t) > \epsilon$$

- Weakly-acyclic ➡ From any action a there exists a finite best-response path to a^*

Start from action $a_t = \{B, A, B, A\}$, let $a^* = \{B, B, B, A\}$ be a NE of the game

	$t \geq \bar{t}$	…	$t + T - 1$		$t + T$
1	B		B	w. prob. ρ	B
2	A	…	A	$\arg\max_{\alpha} u_2(\alpha, a_{-i}) =$	B
3	B		B	w. prob. ρ	B
4	A		A	w. prob. ρ	A

Proof method: absorbing Markov Chain

Lemma (absorption property of pure Nash equilibria)

$\exists\, T$ such that after $t > 0$ if Nash equilibrium action a^* is repeated for T periods, then future actions will be Nash equilibrium, i.e., $a_{t+\tau} = a^*$ for all $\tau > T$.

- Use the condition

Let $a^* = \{B, B, B, A\}$ be a NE of the game with common belief

	1	2	…	$t > \bar{t}$	$t+1$	…	$t+T-1$	$t+\tau$	…
1	A	A		B	B		B	B	
2	B	A	…	B	B	…	B	B	…
3	B	B		B	B		B	B	
4	B	B		A	A		A	A	

Learning-aware communication protocols for DFP

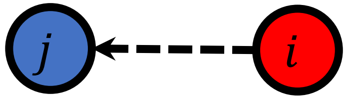
- Protocols that rely on these two metrics

  Novelty of information ➡ $H_{ii}(t) = |f_{i,t} - \Phi(a_{i,t})|$

  Belief mismatch ➡ $H_{ij}(t) = |f_{i,t} - f_{i,t}^{j}|$

- **(CC) - Communication censoring** *based on novelty of information:*
  - Agent i attempts to send empirical frequency to *all agents*

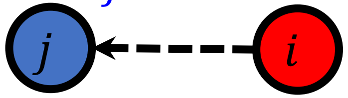
$H_{ii}(t) > \eta_1$



- **(CV) - Voluntary communication** *based on novelty and belief mismatch* **:**
  - Agent i attempts to send empirical frequency to *agent j*
  - *Provides a preference ranking of neighbors to communicate*

$H_{ii}(t) > \eta_1$

$H_{ij}(t) > \eta_2$



- **(L) - Limited bandwidth communication**
  - Agent i attempts to send leading action and its frequency to *agent j*

$\max_{a \in A} f_{it}(a)$

$\operatorname{argmax}_{a \in A} f_{it}(a)$

- Any combination of protocols **C, V or L** satisfies the condition for η_1, η_2

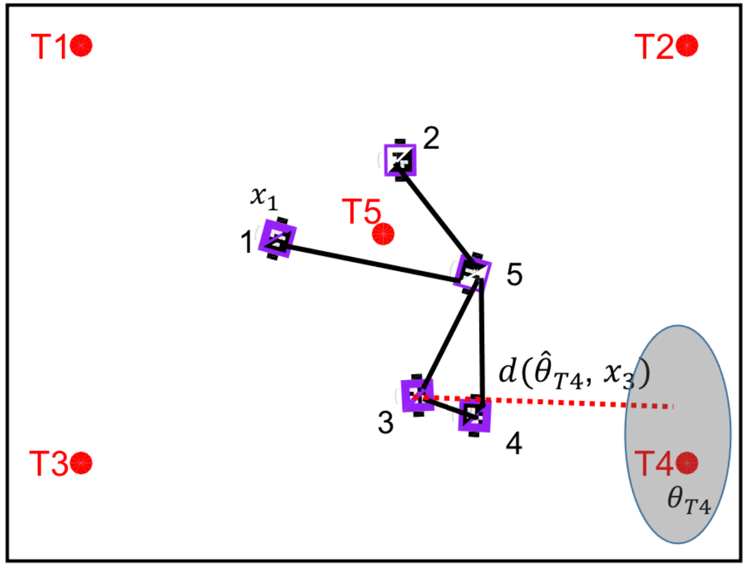
Target assignment game

- $N = 5$ agents select among $K = 5$ targets to maximize coverage
- Agents move at constant speeds in the direction
- Positions affect the probability of successful communication

$$p_{ij}(t) = \beta_{ij}(t)e^{-r\left|x_i(t)-x_j(t)\right|^2}$$

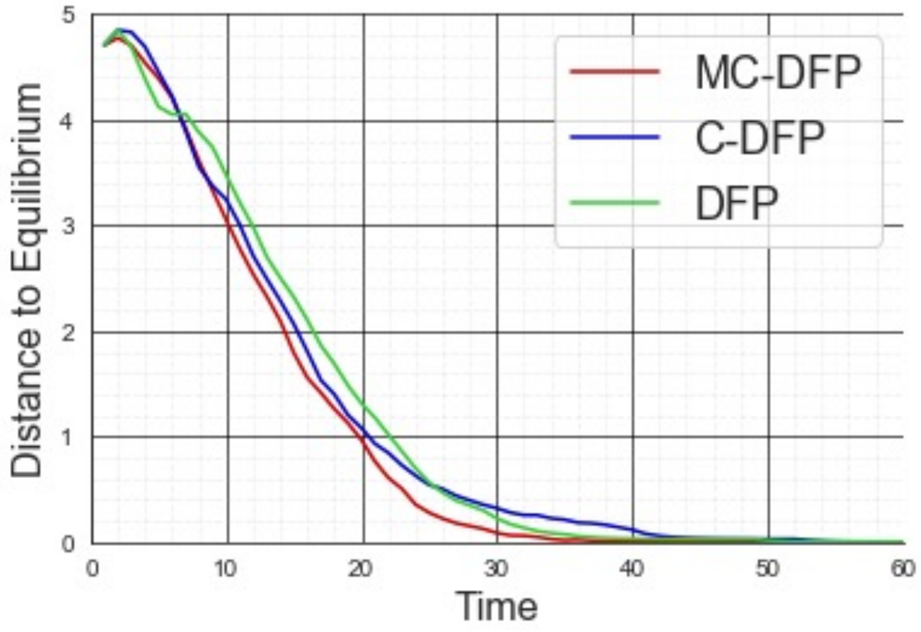
where routing rate $\beta_{ij}(t)$ and fading constant r

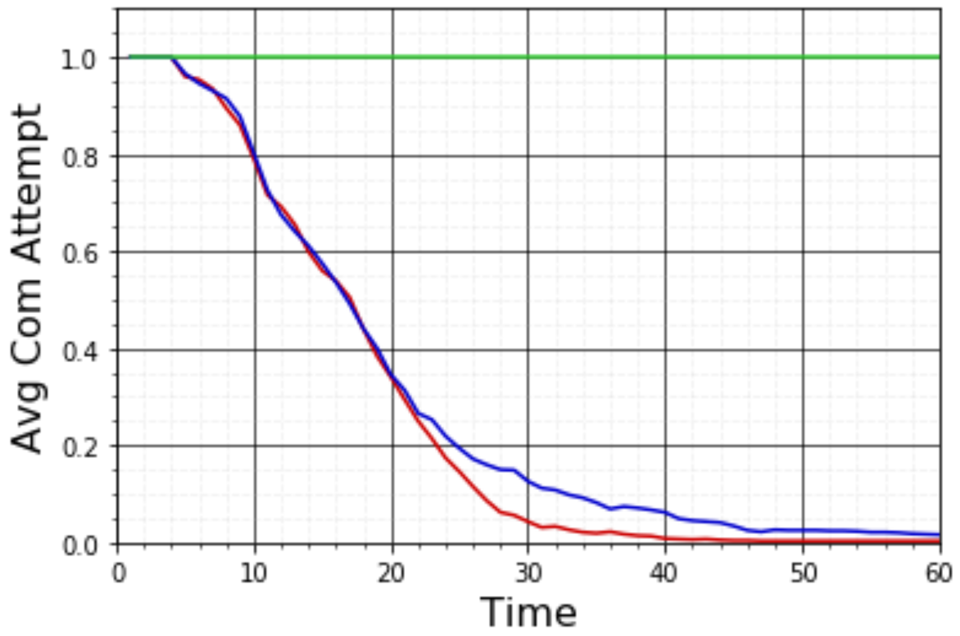
- Chance of communication drops with distance between i and j
- Routing rate $\beta_{ij}(t)$ can be determined by the sender
  - Inversely proportional to belief mismatch H_{ij}
- Communication-aware mobility: Mobility can be "*optimized*" to account for communication



Communication vs convergence rate

- Communication parameters: *censoring* $\eta_1 = 0.1$ and *mismatch* $\eta_2 = 0.4$
- DFP parameters: *fading* $\alpha = 0.05$ and *inertia* $\rho = 0.05$

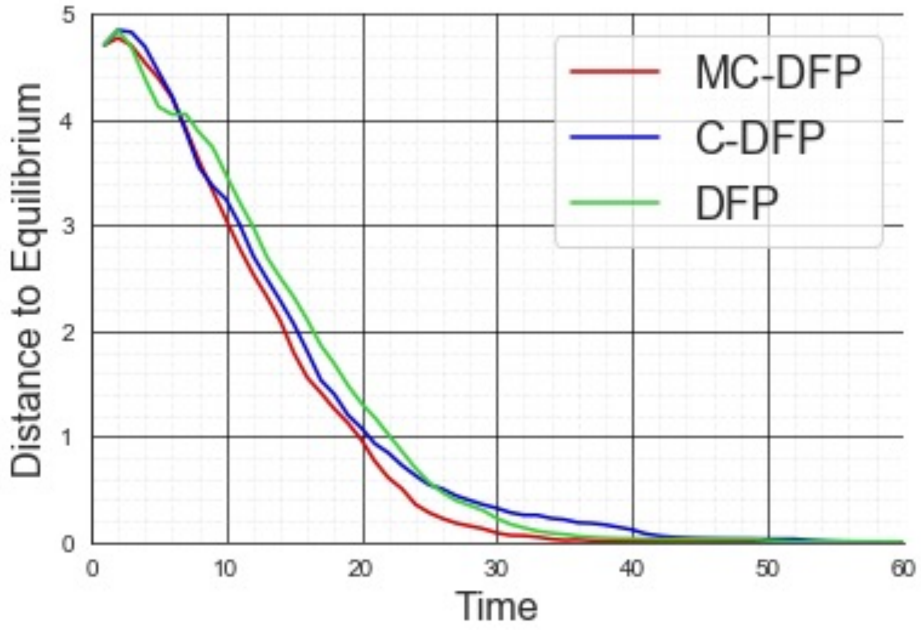


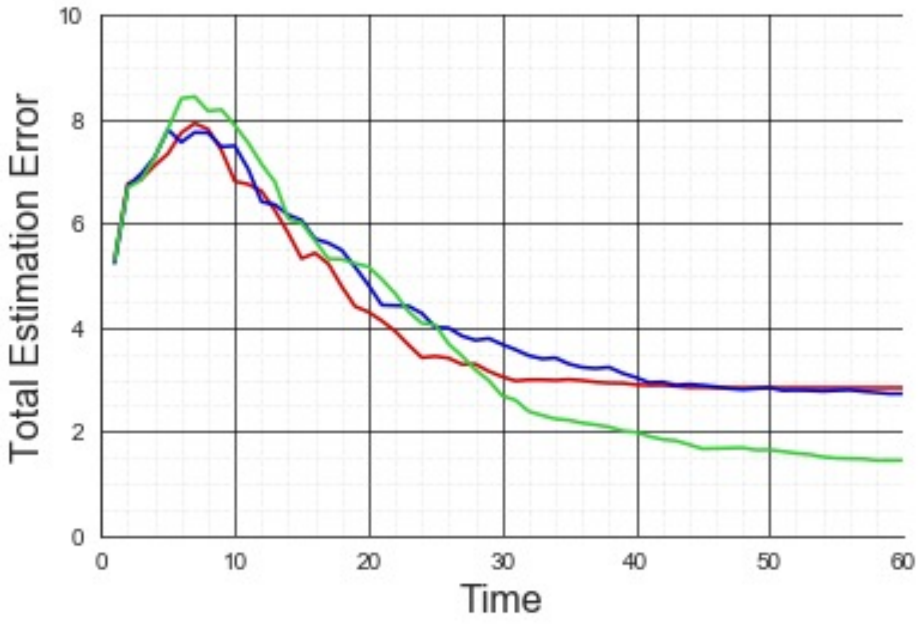


- Voluntary communication (C) and mobility-aware V-com (MC) have comparable convergence
- By $t = 20$, attempt drop below 40%, more than 60% reduction in communication attempts

Communication vs convergence rate

- Communication parameters: *censoring* $\eta_1 = 0.1$ and *mismatch* $\eta_2 = 0.4$
- DFP parameters: *fading* $\alpha = 0.05$ and *inertia* $\rho = 0.05$

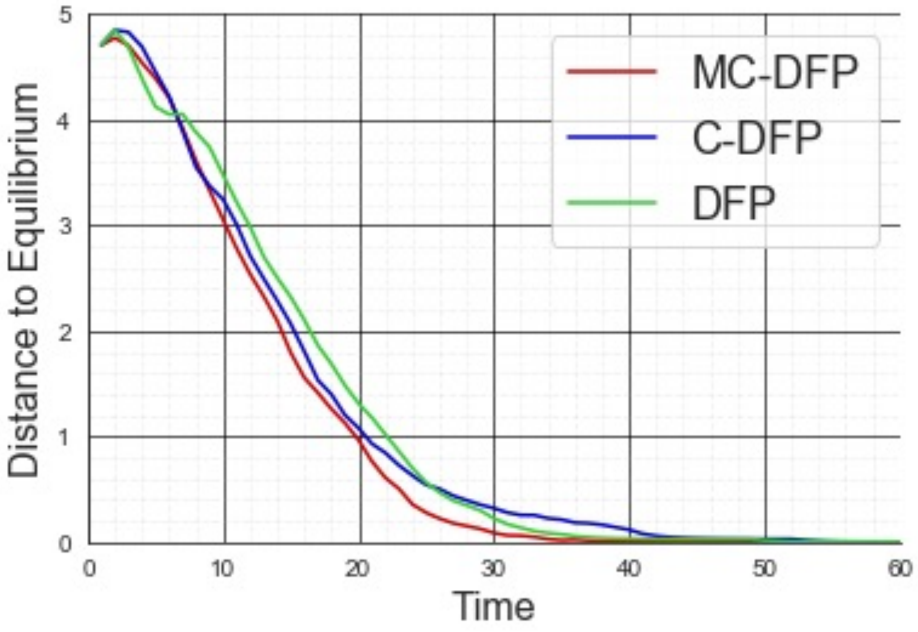


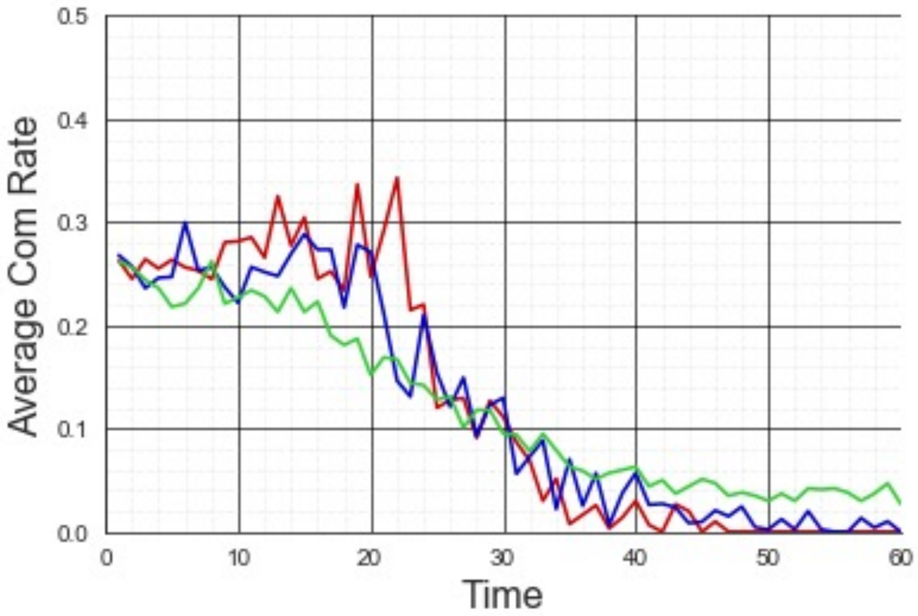


- Voluntary communication (C) and mobility-aware V-com (MC) have comparable convergence
- Total belief mismatch is lower in standard DFP

Communication vs convergence rate

- Communication parameters: *censoring* $\eta_1 = 0.1$ and *mismatch* $\eta_2 = 0.4$
- DFP parameters: *fading* $\alpha = 0.05$ and *inertia* $\rho = 0.05$

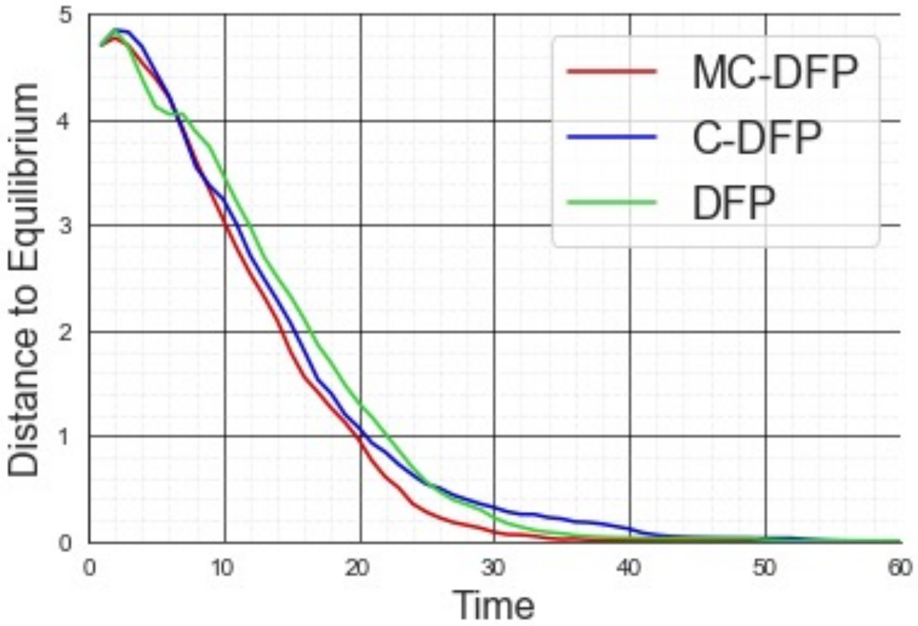


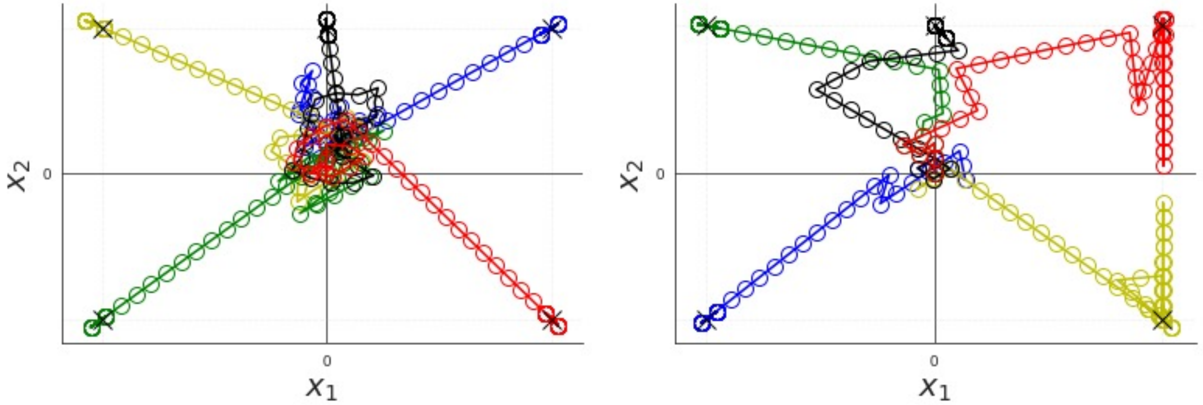


- Voluntary communication (C) and mobility-aware V-com (MC) have comparable convergence
- MC-DFP has higher rate of successful communication initially

Communication vs convergence rate

- Communication parameters: *censoring* $\eta_1 = 0.1$ and *mismatch* $\eta_2 = 0.4$
- DFP parameters: *fading* $\alpha = 0.05$ and *inertia* $\rho = 0.05$

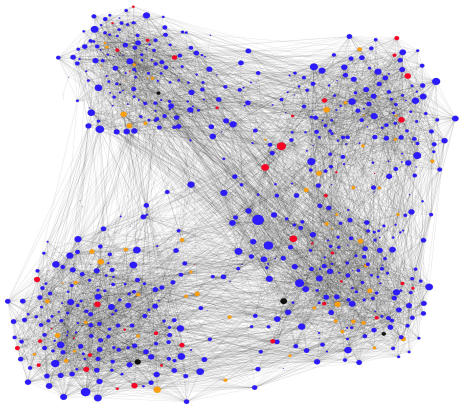




MC-DFP DFP

- Voluntary communication (C) and mobility-aware V-com (MC) have comparable convergence
- In MC-DFP, agents stick together until they resolve their differences

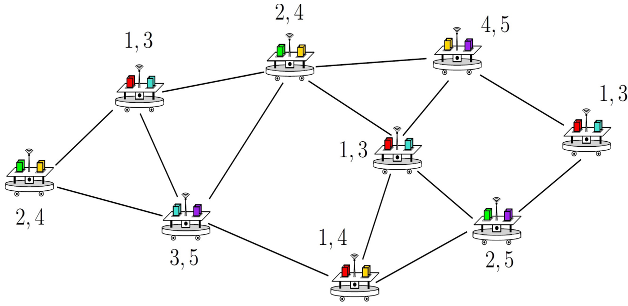
Communication-aware game-theoretic learning algorithms



Networked multi-agent systems

1. Bayesian network games

[IEEE SPM 2013, IEEE TSP 2014, OR 2015]



Communication-aware autonomous teams

2. DFP with voluntary communication

[IEEE TAC 2018, IEEE TAC 2018, Aydin & Eksin 2020]

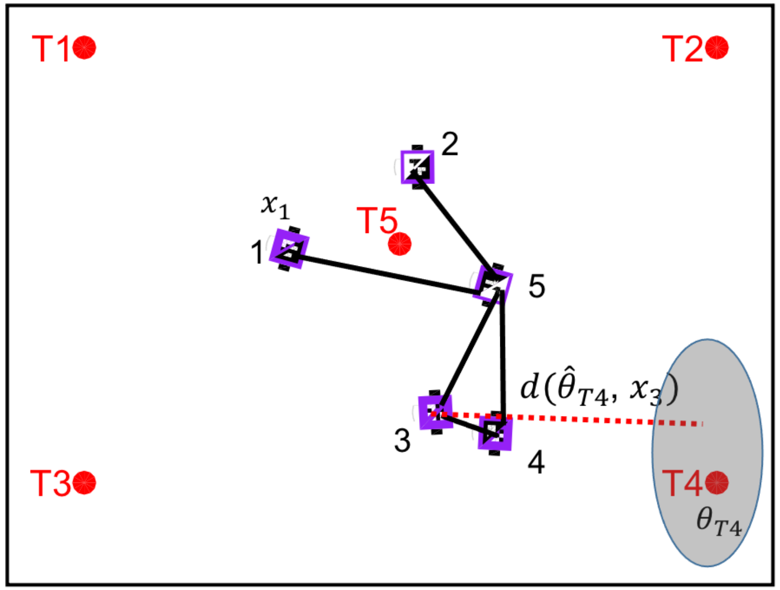
Persisting uncertainty among teams

3. Learning in Near-Potential Games

[Aydin et al 2021]

Persisting uncertainty

- Environmental uncertainty may not be resolved
  - Agents may stop receiving signals
  - Agents may stop communicating



- What can we say when uncertainty is not resolved?
  - I.e., when consensus is not possible

- Framework of *near-potential games*
  - Relaxation of potential games $\hat{\Gamma} = (N, A^N, u)$
  - Defines a game $\Gamma = (N, A^N, \{u_i\}_{i \in N})$ based on its closeness *to the nearest potential*

- Premise: when there is disagreement, game agents play is near the potential game

Near-potential games

Definition (Maximum Pairwise Difference)

Consider two games Γ *and* $\hat{\Gamma}$*. Let* $d^{\Gamma}_{a_i',a} := u_i(a_i', a_{-i}) - u_i(a_i, a_{-i})$ *difference in utility functions when agent* i *deviates from action profile* $a = (a_i, a_{-i})$ *and takes action* a_i'*. MPD is*

$$d(\Gamma, \hat{\Gamma}) = \max_{i \in N, a_i' \in A, a \in A^N} |d^{\Gamma}_{a_i',a} - d^{\hat{\Gamma}}_{a_i',a}|.$$

- MPD based on the difference in agent payoffs resulting from *unilateral changes*

Definition (δ Near-potential games)

The game Γ *is a* δ *near-potential game if there exists a potential game* $\hat{\Gamma}$ *within* $d(\Gamma, \hat{\Gamma}) < \delta$.

- Finding the nearest potential game and its potential function (Candogan et al, 2011)
- Focus on learning in near-potential games assuming potential function is known

Convergence of DFP in near-potential games

1) **Best-response**: $a_{i,t} = \arg\max_{a_i \in A} u_i\left(a_i, f^i_{-i,t-1}\right)$

2) **Update local empirical frequency**: $f_{i,t} = (1-\alpha) f_{i,t-1} + \alpha \Phi\left(a_{i,t}\right)$. where $\alpha = \frac{1}{t}$

3) **Share beliefs** $\{f^i_{j,t-1}\}_{j \in N}$ where $f_{i,t} = f^i_{j,t-1}$ with current neighbors $n(i,t)$

Update agent i's belief on j's frequency: $f^i_{j,t} = \sum_{k \in N} w^i_{j,k}(t) f^k_{j,t}$

- A few differences: no inertia, empirical frequency is computed using weight $\frac{1}{t}$
- Note: we do not have an environment state to disagree on
  - Implicit in the near potential game framework

Convergence of DFP in near-potential games

Theorem (Aydin et al, Arxiv 2021)

Γ is a δ-near-potential game with closest potential game with function $u(\cdot)$. Suppose a time-varying communication network:

a. Union of edges over a finite horizon constitute a strongly connected network

b. Weights sum to 1 (row stochastic), and positive only when neighbors

If the empirical frequencies f_t are ϵ away from NE, then given large enough $T > 0$, for $t > T$

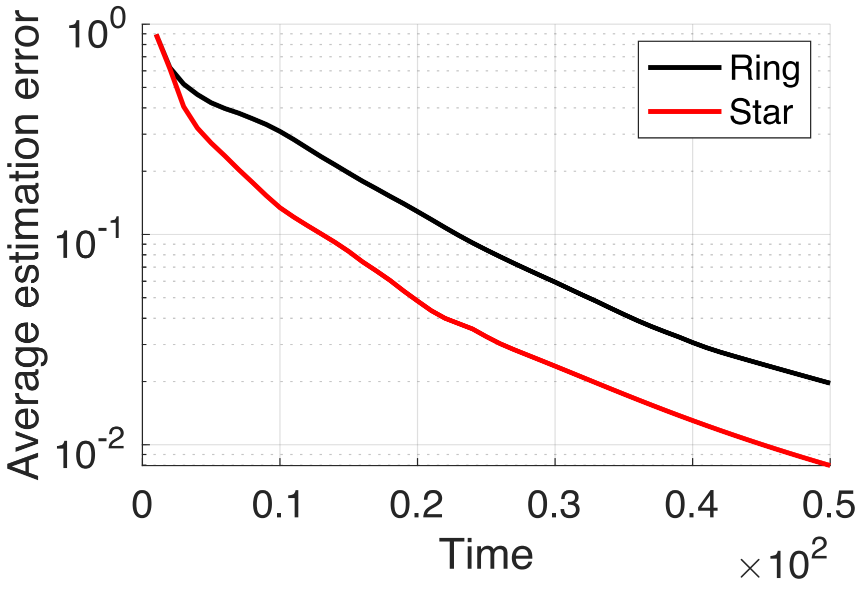
$$u(f_{t+1}) - u(f_t) \geq \frac{\epsilon - N\delta}{t+1} - O(\frac{\log t}{t^2})$$

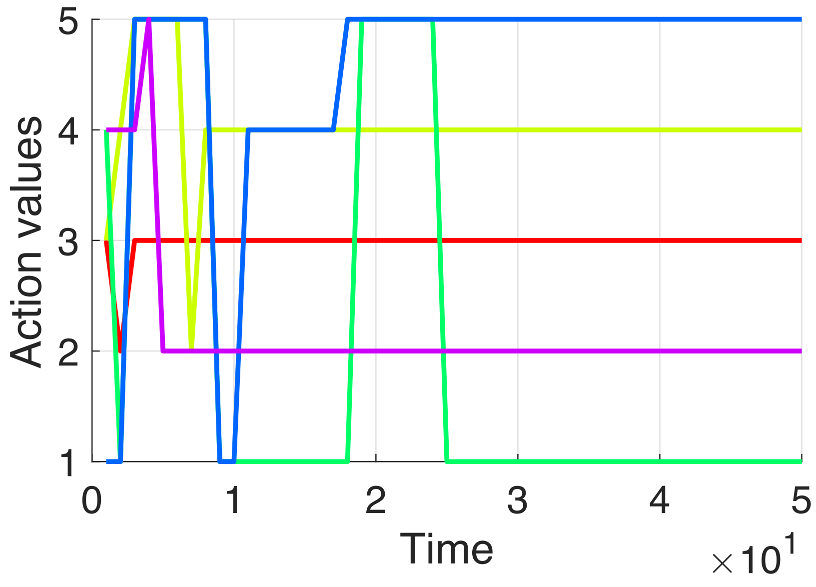
- Potential function value $u\ (f_t)$ improves until near Nash equilibrium

- **Main result:** $f_t \in \{\sigma \in \Delta A^N | u(\sigma) \geq \min_{y \in \Delta_{N\delta+\epsilon}} u(y)\}$ for large $t > 0$
  - Convergence to better potential value than potential function value of an approx. NE

Target assignment game with unknown target locations

- Agents start from random initial positions around the center & ring and star com. networks
- Target locations $\boldsymbol{\theta}$ are unknown & Agents receive noisy signals about $\boldsymbol{\theta}$ for 10 steps





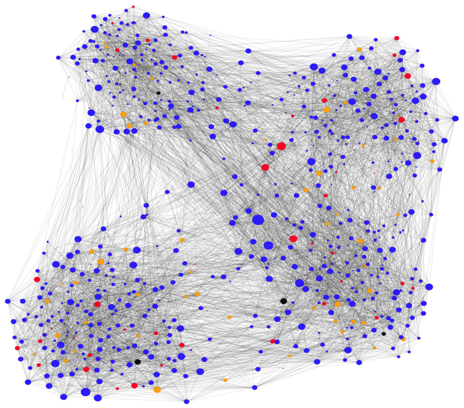
- Different estimates of target locations generates a near-potential game
- Convergence to NE of the ``closest'' potential game in finite time

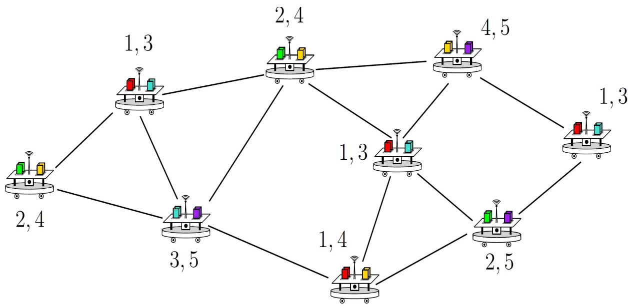
Persisting uncertainty among teams

Networked multi-agent systems

1. Bayesian network games, Distributed fictitious play (DFP)

[IEEE SPM 2013, IEEE TSP 2014, OR 2015]





Communication-aware autonomous teams

2. DFP with voluntary communication

[IEEE TAC 2018, IEEE TAC 2018, Aydin & Eksin 2020]

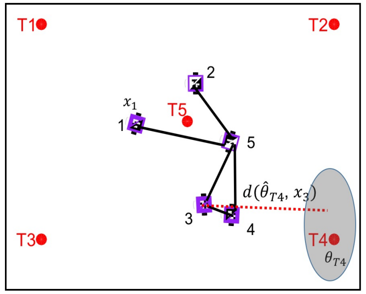
Persisting uncertainty among teams

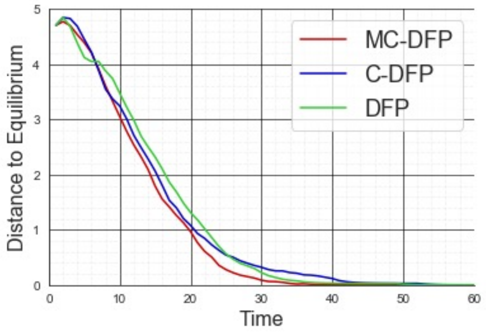
3. Learning in Near Potential Games

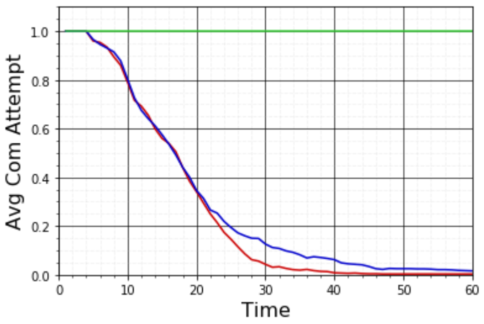
[Aydin et al 2021]

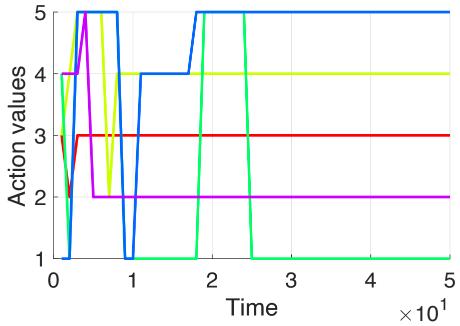
Learning-aware communication in distributed algorithms

Networked multi-agent systems, DFP, communication protocols, persisting uncertainty









Sarper Aydin:
sarper.aydin@tamu.edu

Group members:

- Dr. Hossein Khazaei
- Soham Das
- Furkan Sezer

Collaborators:

- Alejandro Ribeiro, UPenn
- Ali Jadbabaie, MIT
- Pooya Molavi, Northwestern U.
- Jeff S. Shamma, UIUC
- Keith Paarporn, UC Santa Barbara
- Brian Swenson, Princeton
- Soummya Kar, Carnegie Mellon
- Alfredo Garcia, Texas A&M

NSF - Computing and Communication Foundations (CCF)