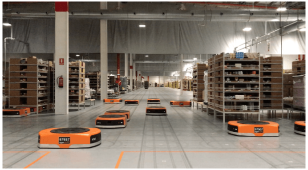
Hypothesis assignment and partial likelihood averaging for cooperative estimation

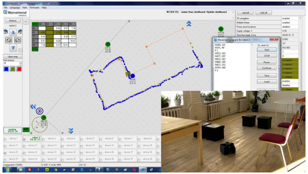
Presented by: Parth Paritosh
Supervised by: Nikolay Atanasov, Sonia Martínez

University of California San Diego (UCSD)
Mechanical and Aerospace Engineering

Dec 13 2019
IEEE CDC 2019 in Nice FR

The ubiquity of distributed sensing infrastructure





Sensor network operations

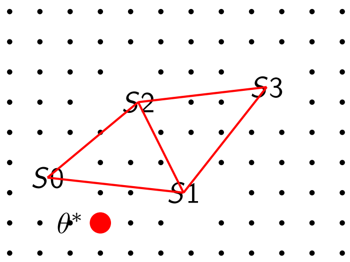
Physical	Sensor selection, communication infrastructure
Data	Storage and retrieval, Clustering
Inference	Estimation, Optimization and Control

Requirements

- Reliability, Security and Robustness
- Computational, storage and energy efficiency
- Temporal variations and Network size

Problem setup

- Discrete hypothesis space: Θ
- True hypothesis location(s): θ^*
- Sensing agents $(S0, \cdots, S3)$
- Local communication, (Weighted adjacency matrix: A)



- How to find the probability of true source location with decentralized communication network?

Existing research

- Jadbabaie et al.[1](2012): Opinion pooling with weighted sum updates in communication graphs
- Nedich et al.[2](2017): Convergence rates of geometric averaging of inferences for the decentralized communication problem
- Atanasov et al.[3](2014): Gaussian filtering algorithm based on geometric averaging

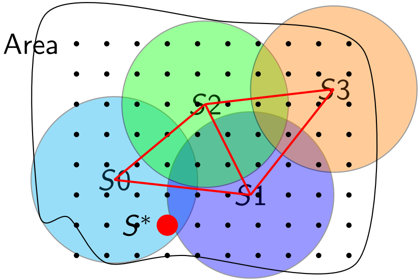
[3] Nedić, Angelia, Alex Olshevsky, and César A. Uribe. "Fast convergence rates for distributed non-bayesian learning." IEEE Transactions on Automatic Control 62.11 (2017): 5538-5553.

[3] Jadbabaie, Ali, et al. "Non-Bayesian social learning." Games and Economic Behavior 76.1 (2012): 210-225.

[3] Atanasov, Nikolay, et al. "Joint estimation and localization in sensor networks." 53rd IEEE Conference on Decision and Control. IEEE, 2014.

Problem setup

- Discrete hypothesis space: Θ
- True hypothesis location(s): θ^*
- Sensing agents $(S0, \cdots, S3)$
- Local communication
- Restricted observation space
- Local storage: Θ_i



Sensors $S0, S1, S2, S3$ learning the source S^* location as a probability over the Area.

The research questions

- Estimating source location probability in distributed communication and storage settings:

RQ1 Distributed storage: How to assign the discrete space (Θ_i) tracked by each agent?

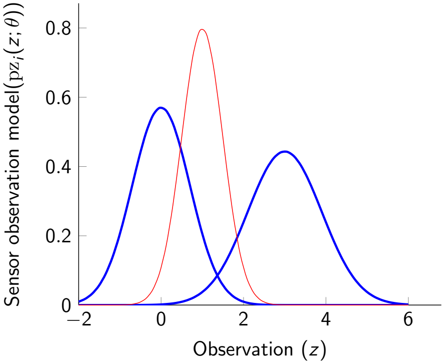
RQ2 Local communication: How to find true source location probability estimates $(p_i(\theta), \forall \theta \in \Theta_i)$ for each agent?

RQ1: Agent space assignment

Assignment objective: Intuition

Maximize the diversity in sensor observation models($\mathrm{pz}_i(z;\theta)$) at each hypothesis.

How to choose the sensors observing any hypothesis?



Assignment objective: Formulation

Diversity maximization

- Choosing sensor observation models over larger observation domain.
- Pairs of observation models with high divergence and entropy term.

$$F(\mathcal{G}_{\boldsymbol{\theta}}) = \sum_{\substack{i,j \in \mathcal{V}(\boldsymbol{\theta}) \\ (i,j) \in \mathcal{E}}} \mathrm{D_{KL}}(\mathrm{pz}_i(\cdot|\boldsymbol{\theta}), \mathrm{pz}_j(\cdot|\boldsymbol{\theta})) + \sum_{i \in \mathcal{V}(\boldsymbol{\theta})} \mathrm{H}(\mathrm{pz}_i(\cdot|\boldsymbol{\theta})). \tag{1}$$

Objective function defined over sets of subgraphs $\mathcal{G}_{\boldsymbol{\theta}}, \forall \boldsymbol{\theta} \in \Theta$ as

$$\max_{\{\mathcal{G}_{\boldsymbol{\theta}}\}} \sum_{\boldsymbol{\theta} \in \Theta} F(\mathcal{G}_{\boldsymbol{\theta}})$$

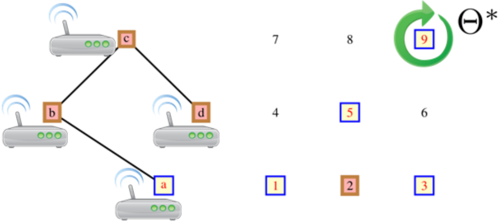
Sub-network assignment constraints

Spatial Coverage

Every hypothesis is observed by one of the sensors.

Limited observation space

Limit the observation space for each agent.



Sub-network assignment constraints

Spatial Coverage

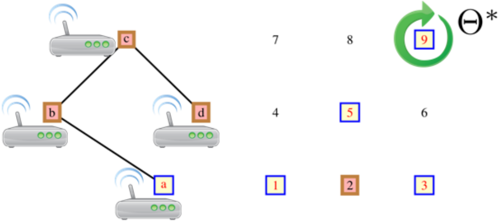
Every hypothesis is observed by one of the sensors.

Limited observation space

Limit the observation space for each agent.

Subgraph connectivity

- Assign connected graphs for learning source probability at each hypothesis.
- NP-hard constraint.



Sub-network assignment as integer optimization

$y_{i,v}, b_{ij,v}$: Inclusion of sensor i and edge (i,j) in sub-network observing θ_v
$f^i_{ij,v}$: Flow variable for sensor i on edge (i,j) in sub-network observing θ_v

$$\sum_{v=1}^{m}\left[\max_{\mathbf{y}_v,\mathbf{b}_v}\sum_{(i,j)\in\mathcal{E}} b_{ij,v}\,\mathrm{D_{KL}}(\mathrm{pz}_i(z|\boldsymbol{\theta}_v),\mathrm{pz}_j(z|\boldsymbol{\theta}_v)) + \sum_{i=1}^{n} y_{iv}\,\mathrm{H}(\mathrm{pz}_i(z|\boldsymbol{\theta}_v))\right]$$

$$\sum_{v=1}^{m} y_{iv} \le m_i, \quad \forall i \in \{1,\ldots,n\}, \qquad \text{(Cardinality)}$$

$$\sum_{i} y_{iv} \ge 1, \quad \forall v \in \{1,\ldots,m\}. \qquad \text{(Coverage)}$$

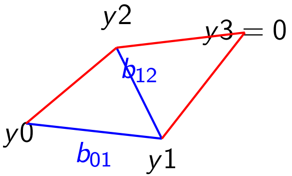
Tackling connectivity constraint

$$\sum_{ij \in \mathcal{E}} b_{ij,v} = \sum_{i=1}^{n} y_{iv} - 1, \qquad \text{(Tree selection)}$$

$$b_{ij,v} \leq y_{iv}, y_{jv}, \forall ij \in \mathcal{E}, \qquad \text{(Edge-node coupling)}$$

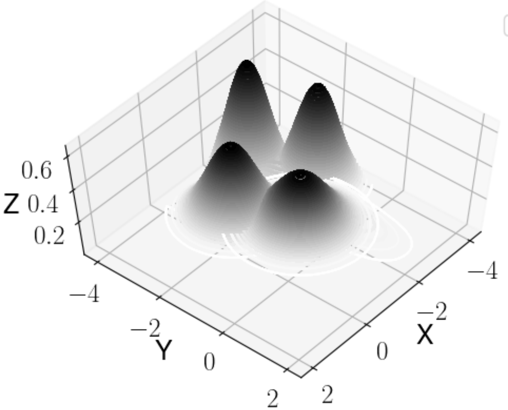
$$f^{i}_{ij,v} + f^{j}_{ij,v} = 2, \qquad \text{(Edge flow)}$$

$$\sum_{j \in \mathcal{N}_i} f^{j}_{ij,v} \leq 2 - \frac{2}{\sum_{i=1}^{n} y_{iv}}. \qquad \text{(Maximum average degree)}$$



Maximum average degree: $\left(\max_{H} \frac{\#edges}{\#nodes}\right)$, where H is subgraph of G. $2(k-1)/k$ for a connected tree of k nodes.

Illustration



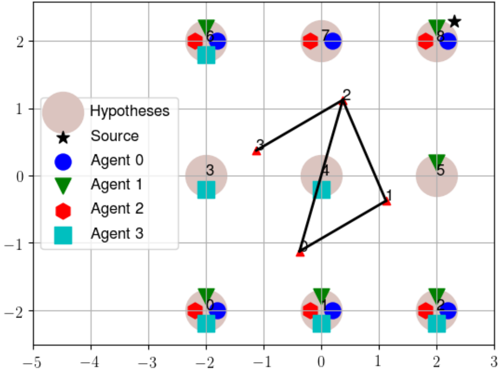


Figure: (a) True distance based Gaussian observation models $pz_i(\cdot|\boldsymbol{\theta}^*)$ for 4 agents. (b) Sensor-hypothesis assignment with number of hypotheses at each sensor limited to 6.

RQ2: Partial likelihood averaging algorithm

Algorithm: Partial likelihood averaging

Sensor i receiving observations $\{\boldsymbol{z}_{i,t}\}_{t=1}^{T}$ to combine inferences with distributed communication matrix $(\boldsymbol{A}(\boldsymbol{\theta}))$

Output: posterior probability $p_{i,T}$ over Θ_i

$\mu_{i,0}(\boldsymbol{\theta}) \leftarrow p_{i,0}(\boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \Theta_i$ *% Initialization*

for $t \in \{1, \ldots, T-1\}$ **do**

  for $\boldsymbol{\theta} \in \Theta_i$ **do**

    % Geometric update

    $\mu_{i,t+1}(\boldsymbol{\theta}) = \prod_{j \in \mathcal{N}_i} \mu_{j,t}(\boldsymbol{\theta})^{A(\boldsymbol{\theta})_{ij}} \, \mathrm{pz}_i(\boldsymbol{z}_{i,t} | \boldsymbol{\theta})$

% Normalization factor

$Z_{i,T+1} = \sum_{\boldsymbol{\theta} \in \Theta_i} \mu_{i,T+1}(\boldsymbol{\theta}) + \sum_{\boldsymbol{\theta} \in \Theta \setminus \Theta_i} \mu_{j,T+1}(\boldsymbol{\theta})$

$p_{i,T}(\boldsymbol{\theta}) = \mu_{i,T}(\boldsymbol{\theta}) / Z_{i,t+1} \, \forall \boldsymbol{\theta} \in \Theta_i$

Algorithm: Partial likelihood averaging

Sensor i receiving observations $\{\boldsymbol{z}_{i,t}\}_{t=1}^{T}$ to combine inferences with distributed communication matrix ($\boldsymbol{A}(\boldsymbol{\theta})$)

Output: posterior probability $p_{i,T}$ over Θ_i

$\mu_{i,0}(\boldsymbol{\theta}) \leftarrow p_{i,0}(\boldsymbol{\theta}),\ \forall \boldsymbol{\theta} \in \Theta_i$ *% Initialization*

for $t \in \{1, \dots, T-1\}$ **do**

  for $\boldsymbol{\theta} \in \Theta_i$ **do**

    % Geometric update

    $\mu_{i,t+1}(\boldsymbol{\theta}) = \prod_{j \in \mathcal{N}_i} \mu_{j,t}(\boldsymbol{\theta})^{A(\boldsymbol{\theta})_{ij}}\, \mathrm{pz}_i(\boldsymbol{z}_{i,t}|\boldsymbol{\theta})$

% Normalization factor

$Z_{i,T+1} = \sum_{\boldsymbol{\theta} \in \Theta_i} \mu_{i,T+1}(\boldsymbol{\theta}) + \sum_{\boldsymbol{\theta} \in \Theta \setminus \Theta_i} \mu_{j,T+1}(\boldsymbol{\theta})$

$p_{i,T}(\boldsymbol{\theta}) = \mu_{i,T}(\boldsymbol{\theta}) / Z_{i,t+1}\ \forall \boldsymbol{\theta} \in \Theta_i$

Assumptions

Sensor data reception

Assume that if the true data generating pdf for agent i, $f_i^*(\boldsymbol{z}) > 0$ for some $z \in \mathbb{R}^{d_z}$, then $1 \geq \bar{\alpha} \geq \mathrm{pz}_i(z|\boldsymbol{\theta}) \geq \underline{\alpha} > 0$, for all $\boldsymbol{\theta} \in \Theta_i$ and some constants $\bar{\alpha}$, $\underline{\alpha}$. Note that $\bar{\alpha}$ exists for any pdf.

Other assumptions

Static graph The undirected graph $\mathcal{G}$ describing the agent communication is static and time-invariant.

Coverage Each hypothesis $\boldsymbol{\theta} \in \Theta$ is observed by at least one agent, i.e., $|\mathcal{V}(\boldsymbol{\theta})| \geq 1$.

Initial probabilities Every agent i has an initial likelihood $p_{i,0}(\boldsymbol{\theta}) > 0$, $\forall \boldsymbol{\theta} \in \Theta_i$.

Proof sketch

1. Consensus: All agents probability estimates converge to same value at all hypotheses.

$$\lim_{t\to\infty} A(\theta)^t = \frac{1}{n}\mathbb{1}\mathbb{1}^\top \qquad \text{(Doubly stochastic matrix)}$$

$$\lim_{t\to\infty} p_{i,t}(\theta) = p^*(\theta) \qquad \text{(For agents observing } \theta\text{)}$$

2. Convergence: The probability of an incorrect hypothesis is almost surely zero.

$$\lim_{t\to\infty} \frac{p_{i,t}(\theta_1)}{p_{i,t}(\theta_2)} \to 0 \qquad \text{(When only } \theta_2 \text{ is optimal)}$$

Proof sketch

1. Consensus: All agents probability estimates converge to same value at all hypotheses.

$$\lim_{t\to\infty} A(\theta)^t = \frac{1}{n}\mathbb{1}\mathbb{1}^\top \qquad \text{(Doubly stochastic matrix)}$$

$$\lim_{t\to\infty} p_{i,t}(\theta) = p^*(\theta) \qquad \text{(For agents observing } \theta\text{)}$$

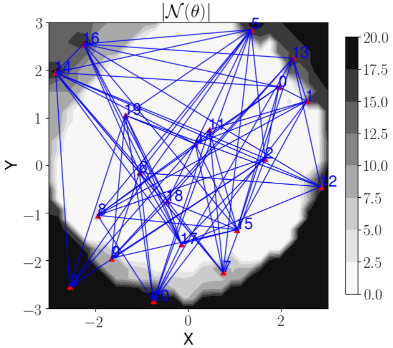
2. Convergence: The probability of an incorrect hypothesis is almost surely zero.

$$\lim_{t\to\infty} \frac{p_{i,t}(\theta_1)}{p_{i,t}(\theta_2)} \to 0 \qquad \text{(When only } \theta_2 \text{ is optimal)}$$

Assignment in 20-node network

For n agents, E edges and m hypotheses

- #variables : $\mathcal{O}(m(n+3|E|))$
- #constraints : $n+2m+4m|E|$

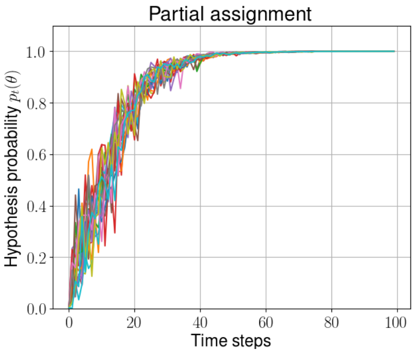


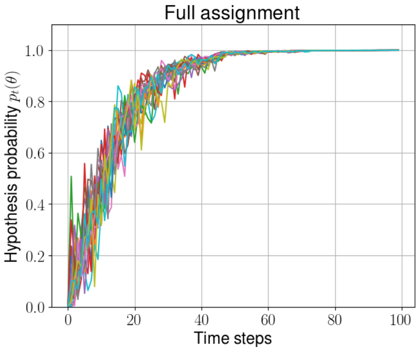
Number of agents
assigned at each hypothesis

* 20 agent connected graph in space with 300 discrete hypotheses
* Distance based observation model assigns more agents to low coverage areas

Convergence in 20-node network

* Comparing estimated probability at source with almost 1/3 communication and memory at each sensor:





Summary

- Matching each hypothesis to connected communication sub-networks based on diversity in observation models
- Geometric update based algorithm for finding inference
- Consensus among agents estimates at all hypotheses
- Proof based on comparing estimates at true and false hypotheses

Thank you