Resilient Distributed Hypothesis Testing

Bo Wu, Steven Carr, Suda Bharadwaj, Zhe Xu, and Ufuk Topcu
Motivation

A network of mobile agents collaboratively monitors a large-scale environment

- Heterogenous sensing capabilities
- Noisy sensors
- Limited communication range
- Time-varying network topology

System hypothesis

\[ \theta^* \in \Theta = \{\theta_0, \ldots, \theta_{m-1}\} \]

Unknown true hypothesis of the system
(e.g., true position of the object of interest)
Collaboration in the Presence of Adversaries

Each agent only receives partial information about the state of the system

- Requires information sharing

May contain (Byzantine) adversaries

- Requires resilience against adversaries

How to process local and shared information so that all the good (non-adversarial) agents can collaboratively detect the true hypothesis of the system?
The Modeling Framework

$N$ agents

At most $f$ adversaries ($f < N - f$)

$m$ possible hypotheses, $\Theta = \{\theta_0, \ldots, \theta_{m-1}\}$

One true hypothesis of the environment $\theta^* \in \Theta$

Each agent follows a given path

At time $t$, agent $i$ gets a local observation $s_{i,t}$ with a probability given by a local likelihood function

\[ l_i(s_{i,t} | q_{i,t}, \theta^*) \]

Position of agent $i$ at time $t$ | Unknown true hypothesis of the system
How will things work?

1. Makes frequent local observations
   • Updates its local belief

2. Shares its actual belief with neighbors

3. Receives neighbors’ actual beliefs
   • Updates its own actual belief
Distributed Hypothesis Testing Problem

Each agent maintains two beliefs (probability distributions over $\Theta$)

Local belief

$$b_{i,t+1}^l(\theta) = \frac{l_i(s_{i,t+1} | \theta, q_{i,t+1})b_{i,t}^l(\theta)}{\sum_{p=1}^{m} l_i(s_{i,t+1} | \theta_p, q_{i,t+1})b_{i,t}^l(\theta_p)}$$

Likelihood function

Prior local belief

Based on purely local information

Actual belief

$$b_{i,t+1}^a(\theta) = g(b_{i,t}^a(\theta), b_{i,t+1}^l(\theta), \{b_{j,t}^a(\theta)\})$$

Current local belief

Merges shared beliefs from neighbors

Adversarial agents can share false actual beliefs!

Problem:
Design resilient actual belief update rule $g$ such that $b_{i,t}^a(\theta^*) \to 1$. 
Related Work

Consensus-based belief update algorithms are established in [1-3]

- Asymptotic convergence is proved
- No adversaries
- Assumes (periodic) connectivity of the network topology

Byzantine-resilient belief update algorithms are established in [4-6].

- Asymptotic convergence is proved in the presence of adversarial agents
- Static network topology
- Assumes global connectivity of the network topology

Contributions

• Resilient algorithms for distributed hypothesis testing with time-varying network topology

• No requirement on the global connectivity of the network topology (the underlying communication graph does not have to be connected)

• Accommodating different sensor noise levels

What can local belief evolution tell?

Given a pair of hypotheses $\theta$ and $\theta'$, agent $i$ is a source agent, denoted as $i \in S(\theta, \theta')$, if it visits at least one position $q$ infinitely often, where $q$ satisfies

$$l_i( \cdot | q, \theta) \neq l_i( \cdot | q, \theta')$$

$i \in S(\theta, \theta') \rightarrow$ Agent $i$ can locally distinguish between $\theta$ and $\theta'$

For every $q$ along the path $s_0, s_1$

$$l_i(s | q, \theta) = \begin{bmatrix} 1/2 & 1/2 \\ 2/3 & 1/3 \\ 1/2 & 1/2 \end{bmatrix}$$

$\theta_0 \quad \theta_1 \quad \theta_2 = \theta^*$

Local belief itself may rule out of some of the hypotheses
Synchronized Distributed Hypothesis Testing (SDHT)

For agent $i$

Update actual belief $b_{i,t+1}^a(\theta)$ for each hypothesis $\theta$

Update the local belief $b_{i,t+1}^l$

$t = t + 1$

for all $\theta' \neq \theta$, $|S(\theta, \theta') \cap N_{i,t+1}| \geq 2f + 1$ ?

Neighboring agents that can distinguish between $\theta$ and $\theta'$

Yes

Neighboring agents

More good agents than (at most $f$) adversarial agents

Case one

Enough informative neighbors

1. Remove $f$ lowest shared actual beliefs on $\theta$ (reject extreme beliefs)

2. $b_{i,t+1}^a(\theta) = \min\{b_{j,t}^a(\theta)\}_{j \in N_{i,t+1}}^\theta, b_{i,t+1}^l(\theta)$

(minimum rule)

Agents remained after step 1 of case one

Normalization of $b_{i,t+1}^a$

Case two

Not enough informative neighbors

$b_{i,t+1}^a(\theta) = \min\{b_{i,t}^a(\theta), b_{i,t+1}^l(\theta)\}$

10
Convergence of SDHT

For any non-adversarial agent $i$, suppose the following conditions hold:

1. Non-zero initial local and initial actual beliefs for every $\theta \in \Theta$

2. Every agent that does not interact “often enough” can distinguish by itself.

That is, if case one in SDHT happens only finitely often for a hypothesis $\theta \in \Theta$, then $i \in S(\theta, \theta')$ for any $\theta' \neq \theta$.

Then SDHT guarantees that $b_{i,t}^a(\theta^*) \to 1$ almost surely as $t \to \infty$. 
How to Make Better Use of the Shared Information?

Recall in SDHT
Case one: If for all $\theta' \neq \theta$, $|S(\theta, \theta') \cap \mathcal{N}_{i,t+1}| \geq 2f + 1$

An agent must have enough number of informative neighboring agents to make use of the shared information.

**Enough number of neighbors at the same time instant!**

Can we make more frequent use of the shared information?

Key idea:
- Keep collecting shared actual beliefs over time
- Until enough actual beliefs collected for actual belief update
Asynchronous Distributed Hypothesis Testing (ADHT)

For agent $i$

Update actual belief $b_{i,t+1}^a(\theta)$ for each hypothesis $\theta$

Update the local belief $b_{i,t+1}^l$

$t = t + 1$

For all $\theta' \neq \theta$, $|S(\theta, \theta') \cap N_{i(t+1)}| \geq 2f + 1$?

Yes

Case one

Enough informative neighbors

1. Remove $f$ lowest shared actual beliefs on $\theta$ (reject extreme beliefs)

2. $b_{i,t+1}^a(\theta) = \min \{ b_{j,t}^a(\theta) \} \cap N_{i(t+1)}$

(minimum rule)

Agents remained after step 1 of case one

Normalization of $b_{i,t+1}^a$

No

Case two

Not enough informative neighbors

$b_{i,t+1}^a(\theta) = \min \{ b_{i,t}^a(\theta), b_{i,t+1}^l(\theta) \}$
“Case 1”:
Minimum Rule versus Averaging Rule in Actual Belief Update

Minimum rule

1. Remove \( f \) lowest shared actual beliefs on \( \theta \)
2. \( b_{i,t+1}^a(\theta) = \min \{ b_j^a(\theta) \}_{j \in \mathcal{N}_{i,t+1}} \) \( b_{i,t+1}^l(\theta) \}

Averaging rule

1. Remove \( f \) lowest and \( f \) highest shared actual beliefs on \( \theta \)
2. \( b_{i,t+1}^a(\theta) = \min \{ \text{avg} \{ b_j^a(\theta) \}_{j \in \mathcal{N}_{i,t+1}} \) \( b_{i,t+1}^l(\theta) \)

Agents remained after step 1

- Pro: Quickly rules out unlikely hypothesis
- Con: May have oscillations with noisy sensors

- Pro: Averages out the effect of noisy sensors
- Con: May take more time to suppress unlikely hypotheses
Case Study: Compromised UAV Classification

A set of agents with persistent surveillance tasks

Five agents — four good agents, one bad agent

Each agent

• Has a trajectory for prescribed surveillance task
• May follow a set of possible alternative trajectories if compromised
• Bad agent shares randomly generated beliefs

Noisy sensors

• When an agent is in range, its position is sensed from a distribution across the viewable positions

Objective: identify the bad agent

Hypothesis $\theta = \langle \theta(0), \theta(1), \theta(2), \theta(3), \theta(4) \rangle$

$\theta(i) \in \{0 \text{ (bad)}, 1 \text{ (good)}\}$

True hypothesis $\theta^* = \langle 1,1,1,0,1 \rangle$

Agent 3 is bad, $(\theta(3) = 0)$. 
Case Study: Results

- Both SDHT and ADHT converge to the true hypothesis
- ADHT converges faster

\[ \theta^* = \langle 1, 1, 1, 0, 1 \rangle \]

- ADHT uses shared beliefs more frequently than SDHT
Case Study: Simulation

Dashed box indicates sensor and communication range

Case Study: Demonstration from perspective of agent 4

Belief over true status for agent 4

Agent 4 radar plot

Edge = good
Center = bad
Minimum Rule vs Averaging Rule

Minimum rule converges faster with a low noise sensor

Averaging rule converges faster with a high noise sensor
Summary

- New distributed learning algorithms that are resilient to Byzantine adversaries
- Convergence is guaranteed without global connectivity constraints
- Minimum and average rules for different levels of sensing noises

What is next?

- How to plan the motion for each agent to guarantee convergence?
- Convergence rate analysis

\[ b_{i,t}^a(\theta^*) \to 1 \text{ almost surely as } t \to \infty \]