Inferring Private Information in Wireless Sensor Networks

Daniel A. Burbano-L.[1], Jemin George[2], Randy A. Freeman[1], and Kevin M. Lynch[1]
[1]Northwestern University, Evanston, IL 60208, USA
[2]CCDC Army Research Laboratory, Adelphi, MD 20783, USA

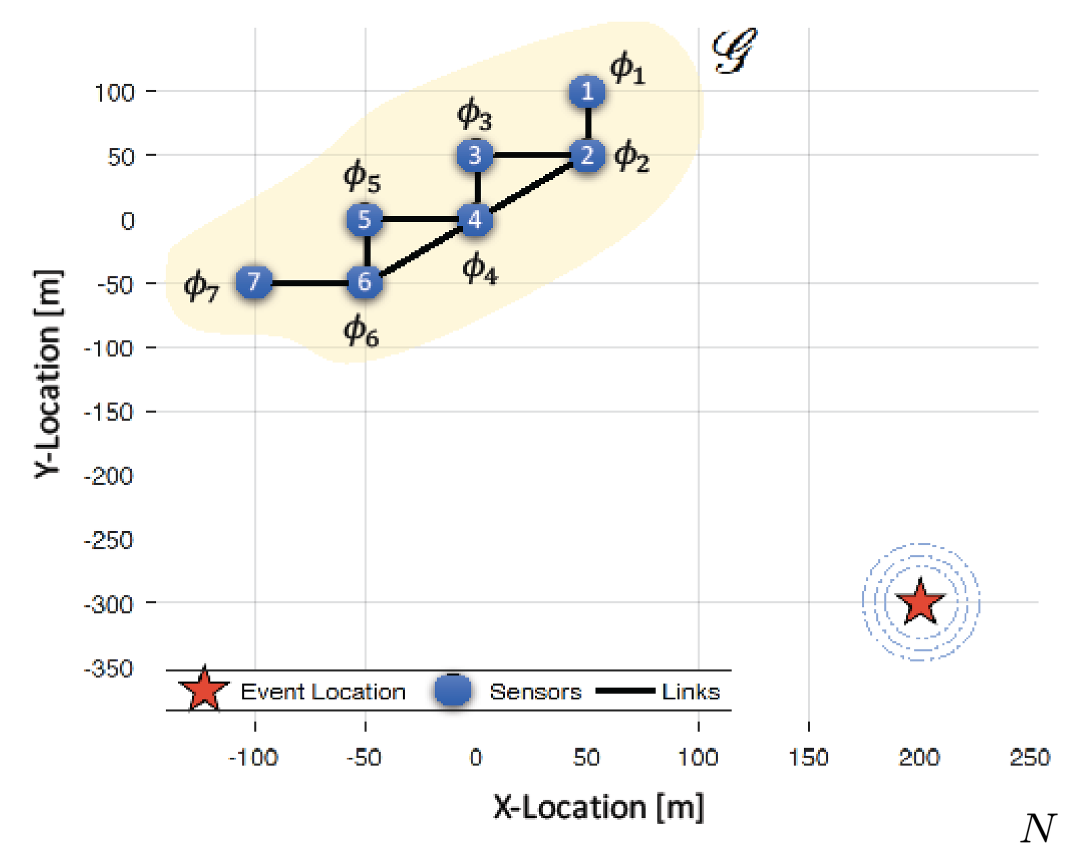
Problem Formulation

- Sensor network as an undirected graph $\mathscr{G}$ of order N
- Measurements are $\boldsymbol{\phi}_i \in \mathbb{R}^n$

$$\boldsymbol{\phi}_i = \boldsymbol{h}_i(\boldsymbol{\theta}, \boldsymbol{p}_i) + \boldsymbol{\xi}_i, \quad i \in \mathcal{N} := \{1, \cdots, N\}$$

  - Unknown variable: $\boldsymbol{\theta} \in \mathbb{R}^n$
  - Private parameters: $\boldsymbol{p}_i \in \mathbb{R}^q$
  - Sensor mapping: $\boldsymbol{h}_i : \mathbb{R}^{m=n+q} \mapsto \mathbb{R}^n$
  - Gaussian noise: $\boldsymbol{\xi}_i \sim G(\mathbf{0}, \mathbf{R}_i)$
- **Goal**: Estimate θ distributively via local interactions

Distributed Localization



- Localization problem $\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta})$, $J(\boldsymbol{\theta}) := \sum_{i=1}^{N} \boldsymbol{f}_i(\boldsymbol{\theta}, \boldsymbol{p}_i)$

$$\boldsymbol{f}_i(\boldsymbol{\theta}, \boldsymbol{p}_i) := \frac{1}{2}(\boldsymbol{\phi}_i - \boldsymbol{h}_i(\boldsymbol{\theta}, \boldsymbol{p}_i))^\top \boldsymbol{R}_i^{-1}(\boldsymbol{\phi}_i - \boldsymbol{h}_i(\boldsymbol{\theta}, \boldsymbol{p}_i))$$

- Distributed algorithm (for i-th agent)

$$\dot{\boldsymbol{v}}_i = \alpha\beta \sum_{j=1}^{N} a_{ij}(\widehat{\boldsymbol{\theta}}_i - \widehat{\boldsymbol{\theta}}_j),$$

$$\dot{\widehat{\boldsymbol{\theta}}}_i = -\alpha \boldsymbol{g}_i(\widehat{\boldsymbol{\theta}}_i, \boldsymbol{p}_i) - \boldsymbol{v}_i - \beta \sum_{j=1}^{N} a_{ij}(\widehat{\boldsymbol{\theta}}_i - \widehat{\boldsymbol{\theta}}_j),$$

  - $\boldsymbol{v}_i(0) = \boldsymbol{v}_{i,o} \in \mathbb{R}^n$ and $\sum_{i=1}^{N} \boldsymbol{v}_{i,o} = \mathbf{0}$
  - $\widehat{\boldsymbol{\theta}}_i(0) = \widehat{\boldsymbol{\theta}}_{i,o} \in \mathbb{R}^n$
  - α, β are positive constants
  - *Adjacency matrix* $\boldsymbol{\mathcal{A}} \triangleq a_{ij}$ $(\forall\, i,j \in \mathcal{N})$

- Distributed algorithm (for i-th agent)

$$\dot{\boldsymbol{v}}_i = \alpha\beta d_i \widehat{\boldsymbol{\theta}}_i - \alpha\beta \boldsymbol{u}_i$$

$$\dot{\widehat{\boldsymbol{\theta}}}_i = -\alpha \boldsymbol{g}_i(\widehat{\boldsymbol{\theta}}_i, \boldsymbol{p}_i) - \beta d_i \widehat{\boldsymbol{\theta}}_i - \boldsymbol{v}_i + \beta \boldsymbol{u}_i$$

  - Node degree: $d_i = \sum_{i=1}^{N} a_{ij}$
  - Incomming communication: $\boldsymbol{u}_i := \sum_{j=1}^{N} a_{ij} \widehat{\boldsymbol{\theta}}_j$

Problem 1. *For $k \in \mathcal{N}$, infer (or reconstruct) the k-th gradient $\boldsymbol{g}_k(\widehat{\boldsymbol{\theta}}_k, \boldsymbol{p}_k)$ and the private parameters $\boldsymbol{p}_k$ by listening to (or intercepting) both $\widehat{\boldsymbol{\theta}}_k$ and $\boldsymbol{u}_k$.*

Reconstruction Strategy

Assumption 1. *Parameters α, β, and d_i are known to adversary.*

Gradient Reconstruction

Gradient estimator

$$\dot{\widehat{\boldsymbol{v}}} = \alpha\beta d_i \widehat{\boldsymbol{\theta}}_k - \alpha\beta \boldsymbol{u}_k, \quad \widehat{\boldsymbol{v}}(0) = \mathbf{0},$$

$$\dot{\boldsymbol{z}} = -\widehat{\boldsymbol{v}} - \widehat{\boldsymbol{a}} - \beta d_i \widehat{\boldsymbol{\theta}}_k + \beta \boldsymbol{u}_i, \quad \boldsymbol{z}(0) = \widehat{\boldsymbol{\theta}}_{k,o},$$

$$\dot{\widehat{\boldsymbol{a}}} = -\frac{1}{\tau}\widehat{\boldsymbol{a}} - \frac{\kappa}{\tau}\mathrm{sgn}\{\widehat{\boldsymbol{\theta}}_k - \boldsymbol{z}\}, \quad \widehat{\boldsymbol{a}}(0) = \widehat{\boldsymbol{a}}_o$$

- $\widehat{\boldsymbol{v}} \in \mathbb{R}^n$ and $\widehat{\boldsymbol{z}} \in \mathbb{R}^n$ are the estimates of $\boldsymbol{v}_k$ and $\widehat{\boldsymbol{\theta}}_k$
- $\widehat{\boldsymbol{a}} \in \mathbb{R}^n$ is an estimate of the gradient $\boldsymbol{g}_k$ plus a bias
- $\kappa > 0$ is a feedback gain

Theorem 1. *Given $0 < \tau \ll 1$, the estimates $\widehat{\boldsymbol{a}}$ converges to the private gradient $\alpha \boldsymbol{g}_k(\widehat{\boldsymbol{\theta}}, \boldsymbol{p}_i) + \boldsymbol{v}_{k,o}$ in finite time $t^* > 0$; i.e., $\forall\ t \geq t^*$, $\|\alpha \boldsymbol{g}_k(\widehat{\boldsymbol{\theta}}_k(t), \boldsymbol{p}_k) + \boldsymbol{v}_{k,o} - \widehat{\boldsymbol{a}}(t)\|_2 = \mathcal{O}(\tau)$, where $\mathcal{O}(\tau)$ is a residual error.*

Reconstruction of Private Parameters

For all $t \geq t^*$, we have $\alpha \boldsymbol{g}_k(\widehat{\boldsymbol{\theta}}_k(t), \boldsymbol{p}_k) + \boldsymbol{v}_{k,o} \approx \widehat{\boldsymbol{a}}(t)$.

- M measurements are taken of the signals $\widehat{\boldsymbol{a}}(sT)$ and $\widehat{\boldsymbol{\theta}}_k(sT)$ with sampling rate T and $s = \{1, \cdots, M\}$
- Assuming the functional form of the gradient and the variance of noise $\boldsymbol{R}_k$ are known

Unknwon parameters $\boldsymbol{v}_{k,o}$ and $\boldsymbol{p}_k$ along with the measurements $\boldsymbol{\phi}_k$ can be estimated by solving

$$\min_{\boldsymbol{v}_{k,o}, \boldsymbol{p}_k, \boldsymbol{\phi}_k} \sum_{s=1}^{M} \|\alpha \boldsymbol{g}_k(\widehat{\boldsymbol{\theta}}_k(sT), \boldsymbol{p}_k) + \boldsymbol{v}_{k,o} - \widehat{\boldsymbol{a}}(sT)\|_2^2$$

Numerical Results

- Distributed event localization with $N = 7$ agents
- Each agent can obtain DoA ϕ_i

$$h(\boldsymbol{\theta}, \boldsymbol{p}_i) = \arctan\left(\frac{T_y - S_i^y}{T_x - S_i^x}\right)$$

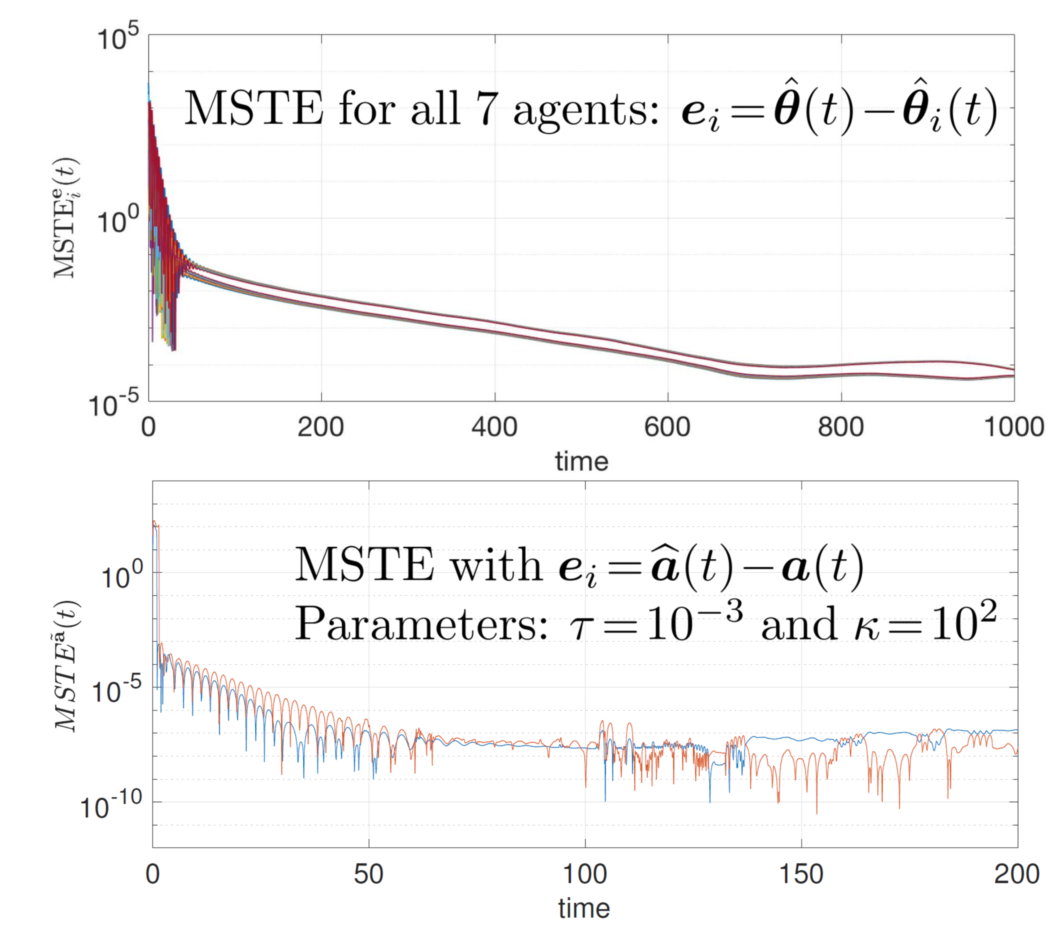
  - Target location: $\boldsymbol{\theta} = [T_x, T_y]^\top \in \mathbb{R}^2$
  - Sensor location: $\boldsymbol{p}_i = [S_i^x, S_i^y]^\top \in \mathbb{R}^2$
  - Gaussian noise: $\boldsymbol{\xi}_i \sim G(\mathbf{0}, 10^{-3})$

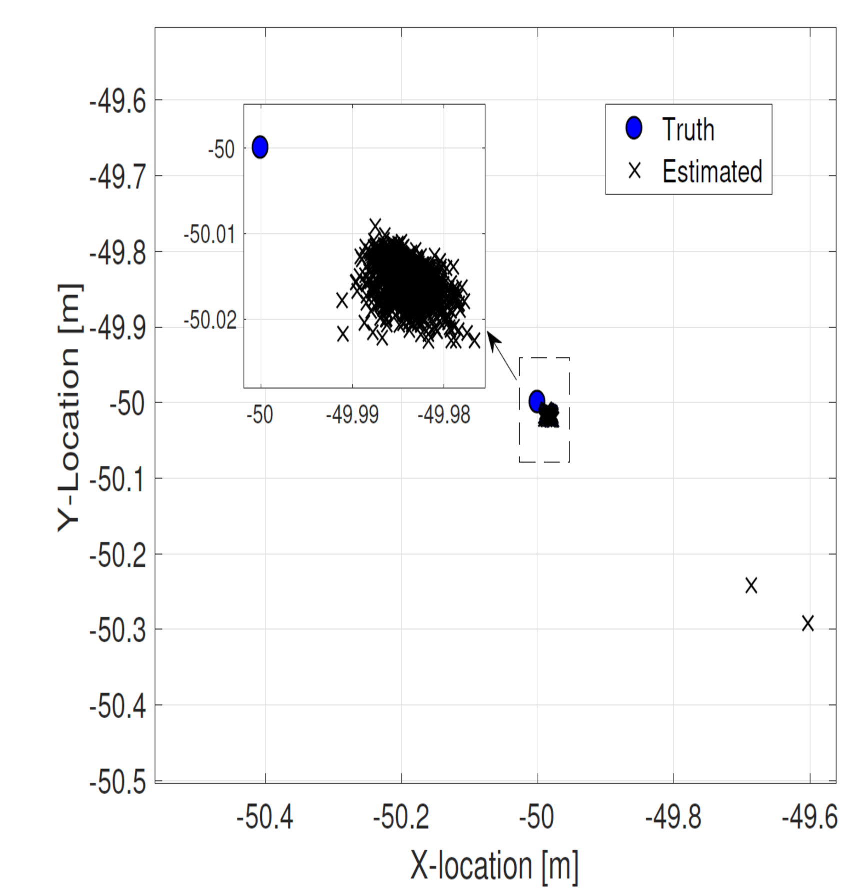
Results from 10^3 MC simulations

- Mean-square tracking error (MSTE)

$$\mathrm{MSTE}_i^{\boldsymbol{e}}(t) = (1/10^3)\sum_{l=1}^{10^3} \|\boldsymbol{e}_i\|^2$$

  - Centralized vs distributed: $\boldsymbol{e}_i = \hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}_i(t)$
  - Gradient estimation error: $\boldsymbol{e}_i = \widehat{\boldsymbol{a}}(t) - \boldsymbol{a}(t)$





Conclusion

- Two step prcess to infer private information
  - Gradient estimator using sliding mode observer
  - Parameters are inferred by solving a nonlinear least-squares problem
- Future work
  - Consider dynamic (tracking) problems
  - Extend the results to discrete-time problems
  - Develop privacy preserving/secure distributed estimation algorithms