Differentially private parameter estimation

Xuanyu Cao and Tamer Başar, University of Illinois at Urbana-Champaign
(Task 1.2, 2019-2020)

Objectives

- In parameter estimation, we preserve the differential privacy of the data owners while achieving satisfactory estimation accuracy.
- Privacy is important for battlefield applications, where adversaries try to overhear the signals and infer the sensitive information of soldiers.

Approach

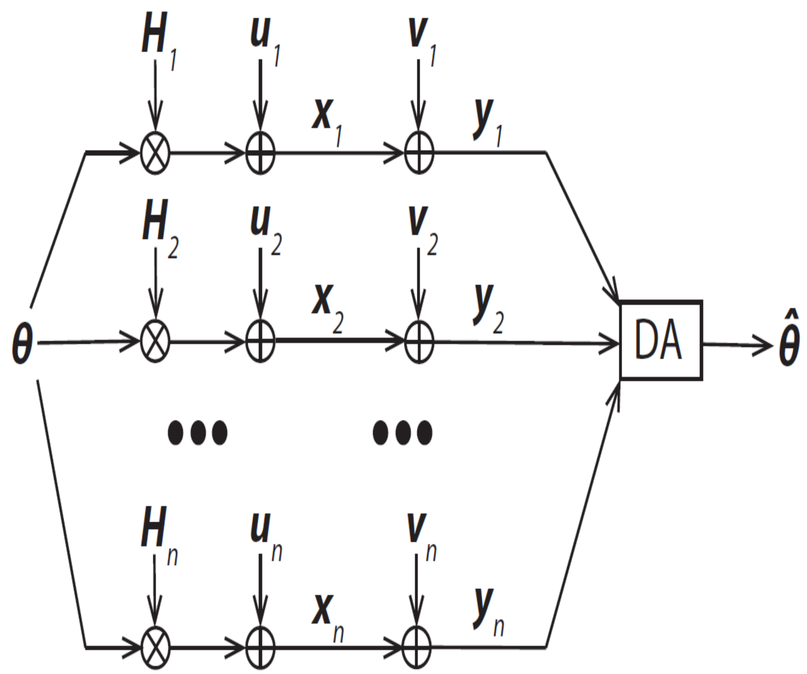
Data model: $\boldsymbol{y}_i = \boldsymbol{x}_i + \boldsymbol{v}_i$

Noise injection: $\boldsymbol{x}_i = \boldsymbol{H}_i\boldsymbol{\theta} + \boldsymbol{u}_i$

Definition 1: Two realizations of the aggregated signals $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^m$ are *neighbors* if there is some $i \in \{1, ..., n\}$ such that $\boldsymbol{x}_j = \boldsymbol{x}'_j$ for any $j \neq i$, and $\|\boldsymbol{x}_i - \boldsymbol{x}'_i\|_2 \leq \Delta_i$.

Definition 2: The added noise $\boldsymbol{v}$ preserves ϵ-differential privacy of the aggregated private signal $\boldsymbol{x}$ if for any measurable $\mathcal{A} \subset \mathbb{R}^m$ and neighboring realizations of the signals $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^m$, we have

$$\mathbb{P}(\boldsymbol{x} + \boldsymbol{v} \in \mathcal{A}) \leq e^{\epsilon}\mathbb{P}(\boldsymbol{x}' + \boldsymbol{v} \in \mathcal{A})$$



Conclusion

The optimal design of differentially private parameter estimator is studied and the optimal level of inserted noise is derived by solving convex optimization problems.

Path Forward

In the future, we can study more realistic communication models between the data owners and the data acquisitor, where the channel distortion and the (natural) noise will be taken into account.

Publication

Xuanyu Cao and Tamer Başar, "Differentially private parameter estimation: optimal noise insertion and data owner selection", *Proc. 59th IEEE CDC*, Dec 14-18, 2020, pp. 2887-2893.

Results

- To minimize the MSE of the estimate subject to privacy constraint, we can solve a **convex SDP** via off-the-shelf optimization tool.
- In the special case of uncorrelated signals, we solve the problem in almost **closed-form**. The optimal solution inserts **more noise** to signals that are **less important** for estimation purpose.
- When we can only access a limited number of data owners' signals, a relaxed SDP problem can be solved to find a heuristic solution for joint data owner selection and noise injection.

POINT OF CONTACT: Xuanyu Cao and Tamer Başar
xyc@illinois.edu, basar1@illinois.edu