SITB2019

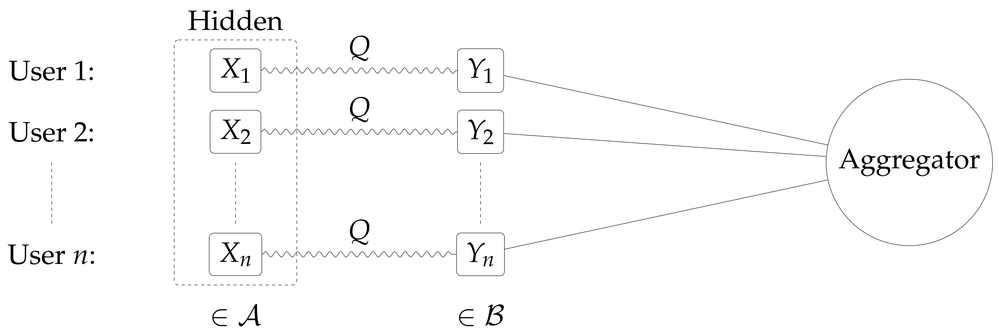
Information-theoretic metrics for Local Differential Privacy protocols

Milan Lopuhaä-Zwakenberg (TUE), Boris Škorić (TUE) & Ninghui Li (Purdue)

1 – Setting: Local Differential Privacy

There are many users and one aggregator. Each user has a value $X_i \in \mathcal{A}$ representing their private data. The aggregator wants to know the distribution of the different elements $x \in \mathcal{A}$ over the user population, while the users want their sensitive data to remain unknown.

Tool: The aggregator publishes a second domain $\mathcal{B}$ and a random function $Q\colon \mathcal{A} \to \mathcal{B}$ which obfuscates the original data. Each user sends $Y_i := Q(X_i)$ to the aggregator.



Two competing concerns:
- *Aggregator*: wants to estimate frequencies of the $x \in \mathcal{A}$, i.e. *Utility*.
- *Users*: want the Y_i to contain little information about the X_i, i.e. *Privacy*.

2 – Current metrics do not capture Privacy and Utility well

Privacy metric: Local Differential Privacy (LDP) [1]

Definition. *Define* $Q_{y|x} := \mathbb{P}(Y_i = y | X_i = x)$. *Then*

$$\mathrm{LDP}(Q) := \max_{x,x' \in \mathcal{A}} \max_{y \in \mathcal{B}} \log \frac{Q_{y|x'}}{Q_{y|x}} \in [0, \infty].$$

Problems:
- Oversensitive: even y with very low probability have big effect on LDP.
- Restrictive: Cannot handle many privacy protocols.
- Opaque: Now intuitive answer as to what LDP value users are happy with.

Utility metric: estimator accuracy [3]

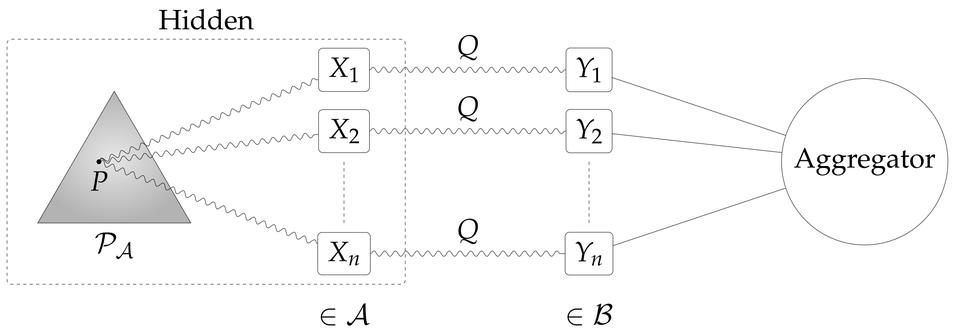
Definition. *For $x \in \mathcal{A}$, let f_x be the true frequency of x among the users, and let $\hat{f}_x\colon \mathcal{B}^n \to \mathbb{R}$ be an estimator for f_x based on the Y_i. Then*

$$\mathrm{Acc}(Q) = \sum_{x \in \mathcal{A}} \mathbb{E}(\hat{f}_x(\vec{Y}) - f_x)^2.$$

Problems:
- Estimator-dependent: measures the utility of the pair $(Q, \hat{f})$ rather than that of just Q.
- Negativity: Many estimators given negative values for $\hat{f}_x$, no good way to account for this.

3 – New metrics based on Information Theory

Extended Model
- The X_i are drawn from some probability distribution $P = (P_1, \cdots, P_a) \in \mathcal{P}_\mathcal{A}$ (the space of probability distributions on $\mathcal{A}$)
- P is itself a random variable drawn from some probability distribution Δ on $\mathcal{P}_\mathcal{A}$.
- Aggregator wants to learn P.
- Δ reflects the aggregator's prior knowledge.



Information theory gives us:

$$\underbrace{\mathrm{H}(\vec{X})}_{\text{total info}} = \underbrace{\mathrm{I}(\vec{X}; P)}_{\text{nonprivate info}} + \sum_i \underbrace{\mathrm{H}(X_i|P)}_{i\text{'s private info}}$$

$$\underbrace{\mathrm{I}(\vec{Y}; \vec{X})}_{\substack{\text{total info} \\ \text{avail. to aggregator}}} = \underbrace{\mathrm{I}(\vec{Y}; P)}_{\substack{\text{nonprivate info} \\ \text{avail. to aggregator}}} + \sum_i \underbrace{\mathrm{I}(Y_i; X_i|P)}_{\substack{i\text{'s private info} \\ \text{avail. to aggregator}}}$$

Definition: New Metrics

$$\mathrm{Uti}_{n,\Delta}(Q) = \frac{\mathrm{I}(\vec{Y}; P)}{\mathrm{I}(\vec{X}; P)}$$

$$\mathrm{Priv}_\Delta(Q) = 1 - \frac{\mathrm{I}(Y_i; X_i|P)}{\mathrm{H}(X_i|P)} = \frac{\mathrm{H}(X_i|Y_i, P)}{\mathrm{H}(X_i|P)}$$

Intuition:

$$\mathrm{Uti}_{n,\Delta}(Q) = \text{Part of nonprivate info available to aggregator,}$$
$$\mathrm{Priv}_\Delta(Q) = \text{Part of private info hidden from aggregator.}$$

4 – The new metrics do not have old metrics' problems

Privacy metric problems:
- Oversensitive: *Solved – by averaging over y;*
- Restrictive: *Solved – also applies in settings of k-anonymity etc.;*
- Opaque: *Solved – privacy is now a percentage with clear meaning.*

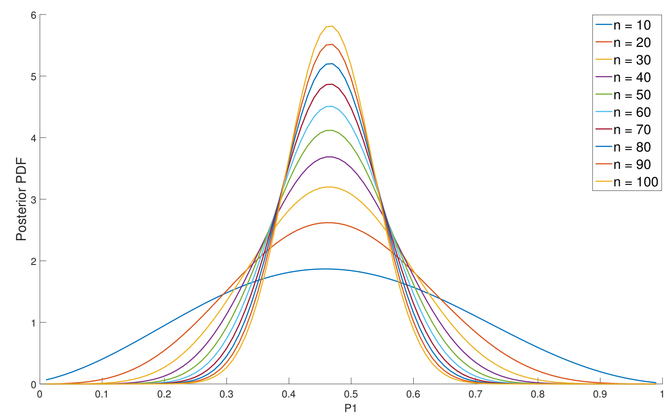
Utility metric problems:
- Estimator-dependent: *Solved – no estimator involved;*
- Negativity: *Solved – metric relates to posterior distribution rather than (negative) estimators.*

5 – Learning about P

Theorem (*Accuracy of aggregator's knowledge*)
Let $P_{\langle d \rangle}$ be the d-digit discretisation of P. Then there exists a constant $c(Q)$ such that as $n \to \infty$,

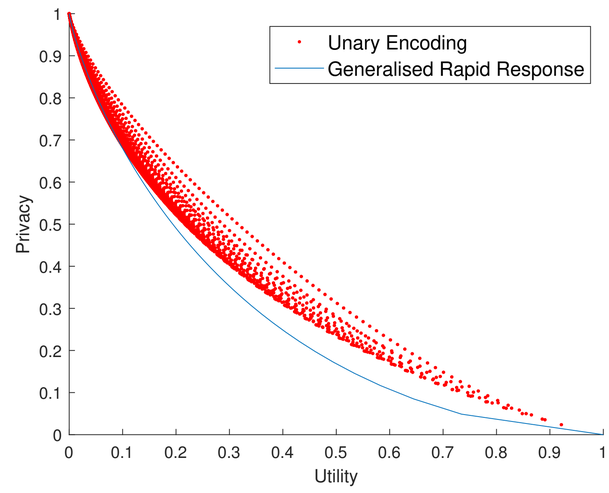
$$\mathrm{I}(Y_1, \cdots, Y_n; P) \approx \mathrm{H}(P_{\langle \frac{1}{2} \log n + c(Q) \rangle}).$$

- The aggregator knows P up to $\approx \frac{1}{2} \log n + c(Q)$ digits.
- One can view $c(Q)$ as an asymptotic utility measure.
- $c(Q)$ is explicitely computable for a given Q.
- The aggregator's knowledge about P manifests in the posterior distribution:

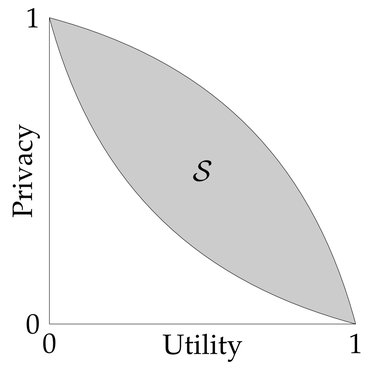


6 – Privacy-Utility Tradeoff (work in progress)

Two privacy protocols, GRR (blue) [2] and UE (red) [3]:



Define $\mathcal{S}$ to be the *feasible region* of privacy protocols, i.e.



$$\mathcal{S} = \Big\{ (\mathrm{Uti}(Q), \mathrm{Priv}(Q)) : Q \text{ privacy protocol} \Big\}$$

To do:
- Can we describe the boundary of $\mathcal{S}$?
- Can we find protocols on the boundary?

References

[1] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local privacy and statistical minimax rates," in *FOCS*, 2013, pp. 429–438.

[2] S. Wang, L. Huang, P. Wang, H. Deng, H. Xu, and W. Yang, "Private weighted histogram aggregation in crowdsourcing," in *International Conference on Wireless Algorithms, Systems, and Applications*. Springer, 2016, pp. 250–261.

[3] T. Wang, J. Blocki, N. Li, and S. Jha, "Locally differentially private protocols for frequency estimation," in *USENIX'17: Proceedings of 26th USENIX Security Symposium on USENIX Security Symposium*. USENIX Association, 2017.