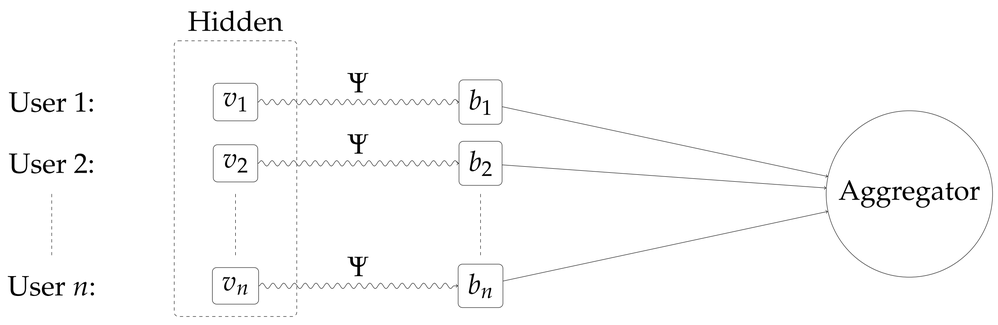
NSF NWO PI Meeting 2018

Bridging the gap between theory and practice in data privacy

Setting

There are many users and one aggregator. Each user has a value $v_i \in D$ representing its private data. The aggregator wants to know the distribution of the different elements $v \in D$ over the user population, while the users want their sensitive data to remain unknown.

Tool: The aggregator publishes a second domain B and a random function $\Psi\colon D \to B$ which obfuscates the original data. Each user sends $b_i := \Psi(v_i)$ to the aggregator.



Two competing concerns:

- *Utility*: The b_i should allow the aggregator to obtain as much information about the distribution of the v_i as possible.
- *Privacy*: b_i should contain as few information about v_i as possible. Measured via Local Differential Privacy (LDP).

Definition 1 (Local Differential Privacy) *Let $\varepsilon > 0$. The algorithm Ψ satisfies ε-differential privacy if for all $v, v' \in D$ and $T \subset B$ one has*

$$\Pr\left[\Psi(v) \in T\right] \leq e^{\varepsilon} \Pr\left[\Psi(v') \in T\right].$$

Intuitively, since the hidden values v and v' have almost equal probability of ending up in $T \subset B$ after applying Ψ, the aggregator, after observing $b_i \in T$, cannot decide whether $v_i = v$ or $v_i = v'$.

Main topic: Frequency Oracles [3]

A basic protocol under LDP is to enable **estimation of the frequency of any given value $v \in D$ among n users. Let $d = |D|$ denote the domain size.** Such a protocol is specified by:

- Ψ, which is used by each user to perturb her input value.
- Φ, which computes estimated frequency of v.

For a good utility, we want the following:

- $\Phi(v)$ should be an unbiased estimator, i.e. $\mathbb{E}[\Phi(v)] = \#\{i : v_i = v\}$.
- The variance $\mathrm{Var}(\Phi(v))$ should be as low as possible.

Question: *For a given level of differential privacy, which protocol gives the best utility?*

Existing Protocols

1. Generalized Randomized Response (GRR) [6] Take $B := D$ and choose a probability p. Define Ψ by having $\Psi(v) = v$ with probability p, and $\Psi(v)$ is a (uniformly picked) $v' \neq v$ otherwise. If p decreases the privacy increases, but the utility decreases. One can calculate that an estimator for the frequency of v is given by

$$\Phi(v) := \frac{\#\{i : b_i = v\} - \frac{n(1-p)}{d-1}}{p - \frac{1-p}{d-1}}.$$

2. Basic RAPPOR [2], Symmetric Unary Encoding (SUE) Take $B = \{0,1\}^D$ and choose a probability p. For $v \in D$ we define the vector $\Psi(v)$ by taking in the v-th coordinate 1 with probability p and 0 with probability $1 - p$. On the other coordinates, the probabilities are the other way around. This is developed by researchers in Google and included in the Chrome browser.

3. Random Matrix Projection [1], Binary Local Hash (BLH) Each user randomly selects a binary hash function, and then transmits the hashed value using GRR.

Abstraction: Pure LDP Protocol Framework

To improve on existing protocols, we define the class of *pure LDP protocols*. This class is big enough to include a wide variety of protocols, while at the same time the utility properties of such protocols can be explicitely described.

Definition 2 (Pure LDP Protocols) *A protocol is pure if there exist a function $\mathsf{Support}\colon B \to \mathscr{P}(D)$ and two probability values $p^* > q^*$ s.t. for all $v_1 \in D$,*

$$\Pr\left[\Psi(v_1) \in \{b \in B \mid v_1 \in \mathsf{Support}(b)\}\right] = p^*,$$
$$\forall_{v_2 \neq v_1} \Pr\left[\Psi(v_2) \in \{b \in B \mid v_1 \in \mathsf{Support}(b)\}\right] = q^*.$$

Each input v is mapped to an output that supports itself with probability p^*. Furthermore, any input other than v can be mapped to an output that supports v with probability q^*. Intuitively, we want p^* to be as large as possible, and q^* to be as small as possible. However, satisfying ϵ-LDP requires that $\frac{p^*}{q^*} \leq e^{\epsilon}$. To estimate the frequencies we take

$$\Phi(v) = \frac{\sum_{i=1}^{n} 1_{\mathsf{Support}(b_i)}(v) - nq^*}{p^* - q^*}. \tag{1}$$

Theorem 1 *For a pure LDP protocol Ψ and $\mathsf{Support}$, Φ is an unbiased frequerncy estimator, and the variance of $\Phi(v)$ in (1) is given by the following approximation:*

$$\mathrm{Var}[\Phi(v)] \approx \frac{nq^*(1-q^*)}{(p^*-q^*)^2}. \tag{2}$$

Newly Optimized LDP Protocols

The theory of pure LDP protocols allows us to find the following two improvements on SUE and BLH:

1. Optimized Unary Encoding (OUE) In SUE the chance of a 'flip' of a bit from 0 to 1 is the same as the other way around. We get better utility by taking different probabilities, and we can use theorem 1 to find the optimal choice.

2. Optimized Local Hash (OLH) Instead of taking hash functions with codomain $\{0,1\}$, we use theorem 1 to optimize the size of the codomain.

Results: OUE and OLH outperform SUE and BLH, especially for large ε. They give the same variance. While OUE is easier to implement, OLH has lower communication cost for large d.

Other topics

Large Domain Setting [4]: When D is large ($\approx 2^{128}$) identifying the 'heavy hitters' (frequently occuring D) by estimating the frequency for *all* $v \in D$ is computationally infeasible. Our new methods are $5\times$ as effective as existing methods in the LDP setting.

Set-Value Setting [5]: In the case that $D = \mathscr{P}(I)$ (i.e. each user has a set of items $v \subset I$). We improve on existing methods to identify frequent items and itemsets.

Multiple Attribute Setting [7]: Users can have multiple attributes, and we are interested in the marginal distributions. There is a tradeoff between the number of estimated marginals, and the accuracy in each estimations. Our methods can handle more marginals and can train machine learning models whose accuracy approaches the non-perturbed version.

Future Project Plan: Information Theory

Information theory provides a mathematical model for information and communication. Since utility and privacy are both essentially about to what extent information is retained (which we want to minimise for privacy and maximise for utility), information theory is the natural language to describe the LDP setting.

Goals:

- Establish an information-theoretical framework for the LDP setting.
- Prove meaningful results about the tradeoff between utility and privacy.
- Use this mathematical model to find optimal LDP protocols.

Other Future Project Plans

Density Estimation for Continuous Setting: Instead of a finite set we take D to be the interval $[0,1]$, and the aggregator wishes to recover the distribution from the randomized responses.

Handling Evolving Data: We want to consider the setting where a user's value v_i changes over time. We need to adapt our protocols so that both privacy and utility are preserved.

Understanding the Privacy Risks of LDP: How secure is LDP in practice, i.e. what are the success rates of re-identification and inference attacks on different LDP protocols?

Reducing Noise in the LDP setting: LDP differs from standard Differential Privacy (DP) in that DP relies on a trusted server that collects the data (and adds random noise before publishing). This trusted party allows one to use protocols that offer greater utility at the same privacy. We aim to develop protocols that rely on a non-trusted party to partly achieve this increase in utility.

References

[1] R. Bassily and A. Smith, "Local, private, efficient protocols for succinct histograms," in *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*. ACM, 2015, pp. 127–135.

[2] Ú. Erlingsson, V. Pihur, and A. Korolova, "Rappor: Randomized aggregatable privacy-preserving ordinal response," in *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*. ACM, 2014, pp. 1054–1067.

[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.

[4] T. Wang, N. Li, and S. Jha, "Locally differentially private heavy hitter identification," *arXiv preprint arXiv:1708.06674*, 2017.

[5] ——, "Locally differentially private frequent itemset mining," in *IEEE Symposium on Security and Privacy*. IEEE, 2018, pp. 578–594.

[6] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.

[7] Z. Zhang, T. Wang, N. Li, S. He, and J. Chen, "Calm: Consistent adaptive local marginal for marginal release under local differential privacy," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2018, p. 0.