Collecting Preference Rankings under Local Differential Privacy (technical report)

Jianyu Yang †1, Xiang Cheng †2*, Sen Su †3, Rui Chen ‡, Qiyu Ren †4, Yuhan Liu †5

†State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications, Beijing, China
{jyyang[1], chengxiang[2], susen[3], qyren[4], liuyuhan[5]}@bupt.edu.cn

‡Samsung Research America, Mountain View, USA
rui.chen1@samsung.com

Theoretical Analysis of SAFA. We theoretically analyze the privacy and utility guarantees of SAFA. We first establish the privacy guarantee of SAFA in the following theorem.

Theorem 1. *For any user u_i with privacy budget ε', SAFA is ε'-LDP for u_i.*

Proof. By definition, for any two different tuples $t_i, t_i{}'$, and any perturbed value index $\widetilde{k}_i^j$ where $j \in \{1, \ldots, |\mathcal{A}|\}$ is the attribute index selected by the data collector, we need to prove that

$$\frac{\Pr[SAFA(t_i, \varepsilon') = \widetilde{k}_i^j]}{\Pr[SAFA(t_i{}', \varepsilon') = \widetilde{k}_i^j]} \leq e^{\varepsilon'}.$$

Due to the random sampling of the attribute index j, we have

$$\begin{aligned}
&\frac{\Pr[SAFA(t_i, \varepsilon') = \widetilde{k}_i^j]}{\Pr[SAFA(t_i{}', \varepsilon') = \widetilde{k}_i^j]} \\
&= \frac{\Pr[j \ is \ sampled] \cdot \Pr[LR(j, t_i, \varepsilon') = \widetilde{k}_i^j]}{\Pr[j \ is \ sampled] \cdot \Pr[LR(j, t_i{}', \varepsilon') = \widetilde{k}_i^j]} \\
&= \frac{\Pr[LR(j, t_i, \varepsilon') = \widetilde{k}_i^j]}{\Pr[LR(j, t_i{}', \varepsilon') = \widetilde{k}_i^j]} \\
&= \frac{\Pr[\widetilde{k}_i^j | I(t_i[A_j])]}{\Pr[\widetilde{k}_i^j | I(t_i{}'[A_j])]}.
\end{aligned} \tag{1}$$

We discuss (1) in all four possible cases:

Case 1: if $I(t_i[A_j]) = \widetilde{k}_i^j$ and $I(t_i{}'[A_j]) = \widetilde{k}_i^j$, $\frac{\Pr[\widetilde{k}_i^j | I(t_i[A_j])]}{\Pr[\widetilde{k}_i^j | I(t_i{}'[A_j])]} = \frac{e^{\varepsilon'}}{e^{\varepsilon'} + |dom(A_j)| - 1} \Big/ \frac{e^{\varepsilon'}}{e^{\varepsilon'} + |dom(A_j)| - 1} = 1;$

Case 2: if $I(t_i[A_j]) \neq \widetilde{k}_i^j$ and $I(t_i{}'[A_j]) = \widetilde{k}_i^j$, $\frac{\Pr[\widetilde{k}_i^j | I(t_i[A_j])]}{\Pr[\widetilde{k}_i^j | I(t_i{}'[A_j])]} = \frac{1}{e^{\varepsilon'} + |dom(A_j)| - 1} \Big/ \frac{e^{\varepsilon'}}{e^{\varepsilon'} + |dom(A_j)| - 1} = e^{-\varepsilon'};$

Case 3: if $I(t_i[A_j]) = \widetilde{k}_i^j$ and $I(t_i{}'[A_j]) \neq \widetilde{k}_i^j$, $\frac{\Pr[\widetilde{k}_i^j | I(t_i[A_j])]}{\Pr[\widetilde{k}_i^j | I(t_i{}'[A_j])]} = \frac{e^{\varepsilon'}}{e^{\varepsilon'} + |dom(A_j)| - 1} \Big/ \frac{1}{e^{\varepsilon'} + |dom(A_j)| - 1} = e^{\varepsilon'};$

Case 4: if $I(t_i[A_j]) \neq \widetilde{k}_i^j$ and $I(t_i{}'[A_j]) \neq \widetilde{k}_i^j$, $\frac{\Pr[\widetilde{k}_i^j | I(t_i[A_j])]}{\Pr[\widetilde{k}_i^j | I(t_i{}'[A_j])]} = \frac{1}{e^{\varepsilon'} + |dom(A_j)| - 1} \Big/ \frac{1}{e^{\varepsilon'} + |dom(A_j)| - 1} = 1.$

Therefore, we have $\frac{\Pr[SAFA(t_i, \varepsilon') = \widetilde{k}_i^j]}{\Pr[SAFA(t_i{}', \varepsilon') = \widetilde{k}_i^j]} \leq e^{\varepsilon'}$. As such, $SAFA$ is ε'-LDP for u_i. □

In what follows, we give the utility guarantee of SAFA. In particular, we have the following theorems.

Theorem 2. *Let $\mathbf{f}_j[k]$ be the true frequency of the k-th value in $dom(A_j)$ for n users. Then, for any attribute index $j \in \{1, \ldots, |\mathcal{A}|\}$ and value index $k \in \{1, \ldots, |dom(A_j)|\}$, we have*

$$\mathbb{E}\left[\mathbf{z}_j[k]\right] = \mathbf{f}_j[k].$$

Proof. To start with, we define a function

$$\mathbb{Y}_j^k(i) = \begin{cases} 1, & if \ DC \ sends \ j \ to \ u_i \ and \ \widetilde{k}_i^j = k \\ 0, & else \end{cases}.$$

Then, we have

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{z}_j[k]\right] \\
&= \mathbb{E}\left[\frac{1}{n} \cdot \frac{|\mathcal{A}| \sum_i^n \mathbb{Y}_j^k(i) - nq_j}{p_j - q_j}\right] \\
&= \frac{1}{p_j - q_j} \cdot \left[\frac{|\mathcal{A}|}{n} \cdot \mathbb{E}\left[\sum_i^n \mathbb{Y}_j^k(i)\right] - q_j\right].
\end{aligned} \tag{2}$$

Due to the random sampling of the attribute index j, the attribute A_j is selected with probability $\frac{1}{|\mathcal{A}|}$. Hence,

*Corresponding author

we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i}^{n}\mathbb{Y}_j^k(i)\right] \\
&= \frac{n}{|\mathcal{A}|}\cdot\left[\mathbf{f}_j[k]\cdot p_j + (1-\mathbf{f}_j[k])\cdot q_j\right] \\
&= \frac{n}{|\mathcal{A}|}\cdot\left[\mathbf{f}_j[k]\cdot(p_j-q_j)+q_j\right]. \qquad (3)
\end{aligned}$$

By substituting (3) into (2), we obtain $\mathbb{E}\left[\mathbf{z}_j[k]\right] = \mathbf{f}_j[k]$. This completes the proof. □

Theorem 2 shows that SAFA is an unbiased estimator and explains why the untrusted data collector can learn useful information regarding the true frequency of every possible value of each attribute in $\mathcal{A}$. The following theorem (i.e., Theorem 3) shows the variation of the estimated frequency of every possible value of each attribute in $\mathcal{A}$.

Theorem 3. *Let $\mathbf{f}_j[k]$ be the true frequency of the k-th value in $dom(A_j)$ for n users. Then, for any attribute index $j \in \{1,\ldots,|\mathcal{A}|\}$ and value index $k \in \{1,\ldots,|dom(A_j)|\}$, the variance of $\mathbf{z}_j[k]$ is*

$$Var\left[\mathbf{z}_j[k]\right] \approx \frac{\left(e^{\varepsilon'}+|dom(A_j)|-1\right)\cdot|\mathcal{A}|-1}{n\cdot(e^{\varepsilon'}-1)^2}.$$

Proof. Initially, we have

$$\begin{aligned}
&Var\left[\mathbf{z}_j[k]\right] \\
&= Var\left[\frac{1}{n}\cdot\frac{|\mathcal{A}|\sum_{i}^{n}\mathbb{Y}_j^k(i)-nq_j}{p_j-q_j}\right] \\
&= \frac{|\mathcal{A}|^2}{n^2\cdot(p_j-q_j)^2}\cdot Var\left[\sum_{i}^{n}\mathbb{Y}_j^k(i)\right]. \qquad (4)
\end{aligned}$$

The random variable $\sum_{i}^{n}\mathbb{Y}_j^k(i)$ is the summation of n independent random variables drawn from the Bernoulli distribution. For n users, $n\cdot\mathbf{f}_j[k]$ (resp. $n\cdot\left(1-\mathbf{f}_j[k]\right)$) of these random variables are from the Bernoulli distribution with parameter $\frac{p_j}{|A|}$ (resp. $\frac{q_j}{|A|}$). Thus, we have

$$\begin{aligned}
&Var\left[\sum_{i}^{n}\mathbb{Y}_j^k(i)\right] \\
&= n\cdot\mathbf{f}_j[k]\cdot\left[\frac{p_j}{|\mathcal{A}|}\cdot\left(1-\frac{p_j}{|\mathcal{A}|}\right)\right] \\
&+ n\cdot\left(1-\mathbf{f}_j[k]\right)\cdot\left[\frac{q_j}{|\mathcal{A}|}\cdot\left(1-\frac{q_j}{|\mathcal{A}|}\right)\right]. \qquad (5)
\end{aligned}$$

By substituting (5) into (4), we obtain

$$\begin{aligned}
&Var\left[\mathbf{z}_j[k]\right] \\
&= \frac{\mathbf{f}_j[k]\cdot\left[p_j\cdot\left(|\mathcal{A}|-p_j\right)\right]+\left(1-\mathbf{f}_j[k]\right)\cdot\left[q_j\cdot\left(|\mathcal{A}|-q_j\right)\right]}{n\cdot(p_j-q_j)^2} \\
&= \frac{q_j\cdot\left(|\mathcal{A}|-q_j\right)}{n\cdot(p_j-q_j)^2}+\frac{\mathbf{f}_j[k]\cdot\left[|\mathcal{A}|\cdot(p_j-q_j)-\left({p_j}^2-{q_j}^2\right)\right]}{n\cdot(p_j-q_j)^2} \\
&\approx \frac{q_j\cdot\left(|\mathcal{A}|-q_j\right)}{n\cdot(p_j-q_j)^2} \\
&= \frac{\left(e^{\varepsilon'}+|dom(A_j)|-1\right)\cdot|A|-1}{n\cdot(e^{\varepsilon'}-1)^2}. \qquad (6)
\end{aligned}$$

This completes the proof. □

Theorem 4. *For any attribute index $j \in \{1,\ldots,|\mathcal{A}|\}$, compared with Harmony, SAFA can achieve higher accuracy of the frequency of every possible value of A_j when*

$$|dom(A_j)| < \frac{(2|\mathcal{A}|-1)\cdot(e^{\varepsilon'}+1)^2+1}{|\mathcal{A}|}-e^{\varepsilon'}+1.$$

Proof. Based on the analysis of Binary Local Hashing in [1], we can derive that the variance of $\mathbf{z}_j[k]$ collected by Harmony is

$$Var_H\left[\mathbf{z}_j[k]\right] \approx \frac{(2|\mathcal{A}|-1)\cdot(e^{\varepsilon'}+1)^2}{n\cdot(e^{\varepsilon'}-1)^2}. \qquad (7)$$

Let (6) < (7), then we have

$$|dom(A_j)| < \frac{(2|\mathcal{A}|-1)\cdot(e^{\varepsilon'}+1)^2+1}{|\mathcal{A}|}-e^{\varepsilon'}+1. \qquad (8)$$

This completes the proof. □

References

[1] T. Wang, J. Blocki, N. Li, and S. Jha, "Locally differentially private protocols for frequency estimation," in *USENIX Security*, 2017.