FLDP: FLEXIBLE STRATEGY FOR LOCAL DIFFERENTIAL PRIVACY

Dan Zhao, Hong Chen, Suyun Zhao, Ruixuan Liu, Cuiping Li, Xiaoying Zhang*

Key Laboratory of Data Engineering and Knowledge Engineering of Ministry of
Education (Renmin University), Beijing 100872, China

ABSTRACT

Local differential privacy (LDP), a technique applying unbiased statistical estimations instead of real data, is often adopted in data collection. In particular, this technique is used in frequency oracles (FO) because it can protect each user's privacy and prevent leakage of sensitive information. However, the definition of LDP is so conservative that it requires all inputs to be indistinguishable after perturbation. Indeed, LDP protects each value; however, it is rarely used in practical scenarios owing to its cost in terms of accuracy. In this paper, we address the challenge of providing weakened but flexible protection where each value only needs to be indistinguishable from part of the domain after perturbation. First, we present this weakened but flexible LDP (FLDP) notion which splits the domain. We then prove the association with LDP. Second, we design a Flexible Hadamard Response (FHR) approach for the common FO issue while satisfying FLDP. The proposed approach balances communication cost, computational complexity, and estimation accuracy. Finally, experimental results using practical and synthetic datasets verify the effectiveness and efficiency of our approach.

Index Terms— Local Differential Privacy, Frequency Oracle, Frequency Estimation

1. INTRODUCTION

In data collection, which is the foundation of signal processing and data mining, there are several threats of invading users' privacy and leaking sensitive information. Moreover, data monitor technology or collusion between nodes can lead to severe information leakage at the time of statistic data release, even cause trust crisis and financial losses. Differential privacy as the *de facto* standard for private data release was first introduced by Dwork [1] and has attracted considerable attention, including theoretical treatments [2, 3] and practical perspectives [4, 5].

In recent years, local differential privacy (LDP) has been proposed to avoid the requirement of a central trusted authority [6, 7, 8]. This mechanism uses the concept of differential privacy in the data collection stage, which means that the perturbed mechanism exists on the local side. LDP applies unbiased statistical estimations instead of real data to protect each user's privacy and prevent the leakage of sensitive information. Moreover, LDP has been widely adopted in industry, for example, companies such as Google [6], Apple [9] and Microsoft [10].

Under the notion of LDP, each data owner encodes their values and sends this information to an aggregator after perturbation: An adversary cannot distinguish any pair of values in the domain with high confidence (this is controlled by a privacy budget ε). However, LDP is overly conservative in that it requires each input to be indistinguishable from any other input after perturbation, which requires substantial noise. For example, the classical mechanism RAPPOR hashes each value into h bits on a vector of length d. To ensure that the outputs of any two different values are similar, each bit of the vector must be disturbed under the uniform privacy budget. From another perspective, only h bits are valid, and the remaining $d - h$ bits are added noise required to satisfy LDP. Indeed, excessive noise does better in protecting each value, but practical scenarios can rarely use this notion due to the accuracy reduction it entails.

Motivation. Based on this motivation, we define a weak but flexible version of LDP (FLDP) that does not need all the outputs to have the same range after perturbation but requires intersection of the output.

To satisfy LDP, the input should be hidden over the entire domain; however, to satisfy FLDP, any input should be hidden in part or all of the domain. Although FLDP weakens privacy, it makes mechanism design more flexible in addressing different practical application scenarios. Under the condition that weak privacy protection is required, the intersection of different ranges could decrease rather than increase the privacy budget, which may lead to easy leakage of the real value. Under conditions of strong privacy protection, the intersection can be enlarged to include the entire universe (satisfying the original LDP). That is, FLDP relaxes LDP such that the perturbation mechanism does not require overly strong privacy protection, where LDP would overprotect the inputs that are less sensitive.

In this study, we design an effective notion called FHR for a frequency oracle (FO) on single-item data while satisfying

This workd is supported by the National Key RD Program of China (grant no. 2018YFB1004401), National Natural Science Foundation of China (61772537, 61772536, 62072460, 62076245, 62172424), Beijing Natural Science Foundation (4212022).

978-1-6654-0540-9/22/$31.00 ©2022 IEEE

FLDP. From a practical perspective, we consider three aspects: privacy protection, query accuracy, and communication and calculation. This paper then introduces and analyzes the FHR mechanism in detail and presents results of its application in an experiment.

The main contributions are summarized as follows:

(i) We introduce a new privacy notion called FLDP, which allows a more flexible design mechanism than that of LDP for different real scenarios.

(ii)We design a notion called FHR for a FO with an unbiased estimator that satisfies FLDP. We introduce and analyze the proposed FHR mechanism in detail.

(iii)We validate the flexibility and effectiveness of our notion and mechanism and empirically demonstrate the effectiveness of the FHR mechanism using synthetic and practical datasets.

2. RELATED WORK

Differential privacy (DP) as the *de facto* standard of data privacy was first introduced by Dworkd [1]. Without a trust aggregator, LDP provides more strong protection than DP, but has more noise [11, 12]. Thus, several variants of LDP and their corresponding mechanisms have been studied. For example, personalized LDP (PLDP) [13], condensed LDP [14], utility-optimized LDP [15], and input-discriminative LDP (IN-LDP) [16]. The purpose of these definitions is to define new LDP notion and makes it less conservative, as there is no real need for such noise in practical scenarios.

3. PRELIMINARY KNOWLEDGE

3.1. LDP Notion

LDP is a local model of DP for collecting user data without a credible aggregator [17, 18]. An LDP mechanism $\mathcal{M}$ ensures that the probability of one value being sent to the server approximates the probability of any other values being sent. The formal privacy requirement satisfies ε-LDP as follows:

Definition 1 (ε-Local Differential Privacy)**.** *Given a mechanism $\mathcal{M}$ with domain $\mathcal{X}$ and range $\mathcal{R}$, if any two items t and $t'(t, t' \in \mathcal{X})$ output the same value s($s \in \mathcal{R}$) through mechanism $\mathcal{M}$, which satisfies the inequality 1, then we say that $\mathcal{M}$ satisfies ε-LDP [19].*

$$Pr[\mathcal{M}(t) = s] \leq e^{\varepsilon} \cdot Pr[\mathcal{M}(t') = s] \tag{1}$$

3.2. Problem Statement

System Model. Our system model involves one aggregator and n users $\mathcal{U} = \{u_1, u_2, ..., u_n\}$. Each user possesses one item t in a finite universe $\mathcal{I} = \{1, 2, ..., d\}$ and adds random noise independently before sending information to the aggregator. Then, the aggregator collects users' data and learns statistical information based on all users while protecting the privacy of each individual user. However, the LDP notion is conservative. We assume that any two items $t_i \in I$ and $t_j \in I$ have ranges $\mathcal{R}(t_i)$ and $\mathcal{R}(t_j)$ via perturbation mechanism $\mathcal{M}$. If $\mathcal{M}$ satisfies LDP, then $\mathcal{R}(t_i) = \mathcal{R}(t_j)$.

Thus, each item t is hidden in the domain $\mathcal{I}$. In a practical application scenario, such a strict requirement is not needed: It is sufficient for item t to be hidden in part or all of the domain $\mathcal{I}$. This is a weaker but more flexible privacy strategy than LDP that also satisfies centralized DP on the server side. In particular, a FO represents common statistical information and is the core issue in LDP research. An FO represents the protocols enabling estimation of the frequency of any value in the domain $\mathcal{I}$ [20].

Threat Model. We assume that each user is honest and not malicious, but that different users or a user and the aggregator can collude and that transmission can be monitored. Therefore, the most stringent attacker has access to all users' transmitted data and to the perturbation mechanism and its parameters, except for the target user.

Utility Goals. The utility goals of the mechanisms are that the attacker cannot infer the true user's value from the disturbance information sent by the user while still allowing the aggregator to obtain an estimated statistic. This enables the availability of user data while protecting the privacy of each user. Thus, the objective of the designed mechanism is to obtain unbiased estimates and low variances.

3.3. Frequency Oracle protocols

Frequency Oracle is a core issue under the LDP framework, which has attracted a lot of theoretical and practical attention. The protocols that are enabling to estimate the frequency of any item/itemset in the domain I are called Frequency Oracle (FO). We review the state-of-the-art LDP protocols on FO. Our mechanism is inspired by previous approaches but also contrasts with them. We want to define a more flexible LDP, and we can design algorithms to achieve more accurate FO. The classical FO algorithms are OUE [11] and OLH [11].

4. FLEXIBLE LDP AND ALGORITHM

This section introduces a new privacy notion called FLDP, which provides weaker protection but is more flexible than LDP.

4.1. Definition

LDP requires any input to be indistinguishable from any other input after perturbation. However, LDP is too conservative to use in practical applications because the definition is based on the worst-case scenario. Intuitively, we need any item in the partial domain $\mathcal{I}$ to be indistinguishable after a perturbation mechanism.

Definition 2 ((ε, η)-FLDP). *Given a mechanism $\mathcal{M}$ with domain $\mathcal{I}$ and range $\mathcal{R}$, and any two items t, t' have the range $R(t), R(t')$ via mechanism $\mathcal{M}$ respectively. If the range satisfies the inequality 2 and the output satisfies the inequality 3, then we say that $\mathcal{M}$ satisfies (ε, η)-FLDP.*

$$\min_{t,t' \in \mathcal{I}} \frac{|R(t) \cap R(t')|}{\max\{|R(t)|, |R(t')|\}} \geq \eta \tag{2}$$

$$\max_{s \in R(t) \cap R(t')} \frac{Pr[\mathcal{M}(t) = s]}{Pr[\mathcal{M}(t') = s]} \leq e^{\varepsilon} \tag{3}$$

In Definition 2, any output s can be transformed by the element of a group $\mathcal{G} \subseteq \mathcal{I}$ via perturbation mechanism $\mathcal{M}$; thus, any element $t \in \mathcal{G}$ can be hidden in the group $\mathcal{G}$. We can control the privacy protection intensity not only via the perturbation parameter ε but also by changing the size of the group $\mathcal{G}$. Therefore, the easiest way to satisfy FLDP is to divide the input into groups, which can be done using the current LDP protocols. We assume that the size of the group is m; then the number of groups is d/m. If the size of the groups in mechanism $\mathcal{M}$ satisfies $m = d$, which means that $\mathcal{G} = \mathcal{I}$, then $\mathcal{M}$ satisfies ε-LDP.

This notion is equivalent to collecting the statistic for each group separately; thus, we define a stronger FLDP to protect privacy.

There exist some other notions that are loosely considered as versions of LDP, such as PLDP [21], CLDP [14], and ID-LDP [16]. In brief, the classification of the privacy budget in the PLDP notion is based on different user requirements: to provide discriminative privacy for inputs, CLDP and IDLDP set the privacy budget for each possible pair based on defined rules. If viewed from this perspective, FLDP also sets rules on the input. Thus, CLDP and In-LDP are special cases of FLDP. FLDP sets the transition probability of two inputs with weak correlation to zero and sets the conversion parameter of any two inputs in a group to ε. This mechanism improves the efficiency and accuracy of each transmission.

4.2. Mechanism Design

Our goal is to design a framework with perturbation mechanism $\mathcal{M}$ and a FO protocol that satisfies the proposed FLDP notion. The challenge of design mechanism $\mathcal{M}$ is that the ranges $\mathcal{R}(t)$ and $\mathcal{R}(t')$ feature subtraction and intersection for any two different inputs t and t' ($\mathcal{R}(t) \cap \mathcal{R}(t') \neq \emptyset$, $\mathcal{R}(t) \backslash \mathcal{R}(t') \neq \emptyset$ and $\mathcal{R}(t') \backslash \mathcal{R}(t) \neq \emptyset$). Furthermore, to reduce communication and computation, we propose the FHR notion by introducing the Hadamard matrix to encode the input. Algorithm 1 presents the pseudocode for FHR, which is divided into three phases: *encoding*, *perturbation*, and *frequency estimation*.

Algorithm 1 Algorithm: Flexible Hadamard Response(FHR).

```
Require: The privacy budget $\varepsilon$; Hadamard matrix $H$.
Ensure: The estimated frequency $f_t$;
 1: for each user $u_i$ with value $t_i$, $i = 1$ to $n$ do
 2:     Obtain the vector $H(t_i)$ from $H$ ;
 3:     Randomly select one $+1$ and one $-1$ from $H(t_i)$ to generate $b_i^*$;
 4:     Sample a Bernoulli variable $u$ that equals 1 with probability $p = \frac{e^\varepsilon}{e^\varepsilon+1}$;
 5:     if $u = 1$ then
 6:         $b_i = b_i^*$
 7:     else
 8:         $b_i = -b_i^*$
 9:     end if
10:     Send $b_i$ to the server;
11: end for
12: The server corrects the summation $\hat{z} \leftarrow \frac{(e^\varepsilon+1)}{2(e^\varepsilon-1)} \cdot (\sum_{i=1}^{n} b_i)$;
13: Obtain the FO for any item $t$ with $\hat{f}_t = \hat{b} \leftarrow \hat{z} \cdot H(t)$;
14: return $\hat{f}_t$;
```

Encoding. (step 2) We first encode each item into a 2^r-length vector mapping from Hadamard matrix H_r, except for the first row, where $r = \lceil 2^{log_2(|\mathcal{I}|+1)} \rceil$ (unless otherwise stated, all the symbols H below represent H_r). Thus, user u_i with item t_i can obtain a vector $H(t_i)$ via Hadamard matrix mapping. The Hadamard matrix is generated by $H_{r+1} = \begin{bmatrix} H_r & H_r \\ H_r & -H_r \end{bmatrix}$.

The property of the encoding is that half of the encoded vector values are $+1$ and the others are -1. In addition, the positions of a vector at $+1$ or -1 have half values $+1$ and half values -1 in any other vector.

Perturbation. (steps 3∼11) In this phase, we randomly select one $+1$ value and one -1 value from the encoding vector, which are denoted as (x, y). We then use a random response to perturb this pair. Each user holds the true values (x, y) with probability p and the reverse values $(-x, -y)$ with probability $1-p$. In theory, user u_i has a vector b_i of length 2^r that has values only at the positions of x and y and has value 0 at all other positions. Thus, each user sends $(index_x, x, index_y)$ to the server, and the communication cost is $2r + 1$. Because y is the inverse of x, we do not need to send the value of y. Moreover, on the user side, the value of the corresponding position can be rapidly obtained based on the row and column of the Hadamard matrix. We can obtain the value in the i-th row and j-th column $[i, j]$ by $H[row, col] = (-1)^{Count\mathbb{1}(bin(i\&j))}$, which means that we first conduct the binary operation between i and j and then count the number of 1s.

Frequency Estimation. (steps 12∼14) The aggregator obtains all the vectors sent by the users and an unbiased frequency estimate is obtained by the calculation. First, the aggregator adds vectors sent by the user to obtain a summation vector $\hat{z} = \sum_i^n b_i$ of length 2^r. Thereafter, the corresponding Hadamard vector $H(t)$ of any item t can be calculated with $\hat{z}$ to obtain the corresponding frequency $\hat{f}_t \frac{(e^\varepsilon+1)}{2(e^\varepsilon-1)} \cdot (\hat{z} \cdot H(t))$

4.3. Privacy and Analysis

4.3.1. Privacy Guarantee

Two theorems establish the privacy and accuracy guarantee of FHR: Theorem 3 proves that FHR satisfies (ε, η=0.5)-FLDP while Theorem 2 proves that the frequency estimation is unbiased. These two theorems guarantee the privacy and availability of FHR.

Theorem 1. *ε-LDP is a special definition of (ε, η=1)-FLDP.*

Proof 1. *If η=1, $R(t) = R(t') = R(t) \cap R(t')$. It means that if an output s is from t, then $Pr[\mathcal{M}(t') = s] \neq 0$. The domains of output are the same between the definition of ε-LDP and (ε, η=1)-FLDP.*

Theorem 2. *The correction $\hat{f}_t$ of FHR is unbiased.*

Proof 2. *The theorem can be proved mathematically and the process omitted.*

Theorem 3. *Algorithm FHR satisfies (ε, η=0.5)-FLDP.*

Proof 3. *Suppose there are two users u_1 and u_2 with values t_1 and t_2. Then, the ranges are $R(t)$ and $R(t')$. Let d denote the length of the vector. There are a total of $|R(t)| = |\mathcal{R}(t')| = \frac{d^2}{4}$ outputs for each item via mechanism $\mathcal{M}$, and the intersection between the two ranges is $|\mathcal{R}(t) \cap \mathcal{R}(t')| = \frac{d^2}{8}$. For any output $s \in \mathcal{R}(t) \cap \mathcal{R}(t')$, the sensitivity is maximized when $\mathcal{M}_t(x, y)$ and $\mathcal{M}_{t'}(x, y)$ are opposites.*

$$
\begin{aligned}
RC_{max} &= \max \frac{P(s|t)}{P(s|t')} \\
&= \frac{P((x,y)_s = (1,-1)|(x,y)_t = (1,-1))}{P((x,y)_s = (1,-1)|(x,y)_{t'} = (-1,1))} = e^{\varepsilon}
\end{aligned}
$$

The output $s \in \mathcal{R}(t) \backslash \mathcal{R}(t')$ or $\mathcal{R}(t') \backslash \mathcal{R}(t)$, we can calculate $\eta = \frac{|\mathcal{R}(t) \cap \mathcal{R}(t')|}{|R(t)|} = 0.5$. Then FHR satisfies ($\varepsilon, \eta$=0.5)-FLDP.

4.3.2. Accuracy Analysis

We analyze the accuracy improvement of the proposed framework by evaluating the error bound in Theorem 2, which is independent of the value of the perturbation mechanism. We first calculate the variance of the estimated frequency. The upper bound noise of each item is then $O(\frac{\sqrt{\log(1/\beta)}}{\varepsilon\sqrt{n}})$ in Lemma 1. Finally, we compare the variance obtained with FHR with that obtained using other LDP mechanisms.

Variance. The variance in the frequency of item t in FHR is denoted as follows: $\hat{f}_t = \frac{(e^{\varepsilon}+1)}{2(e^{\varepsilon}-1)} \cdot (\hat{z} \cdot H(t))$

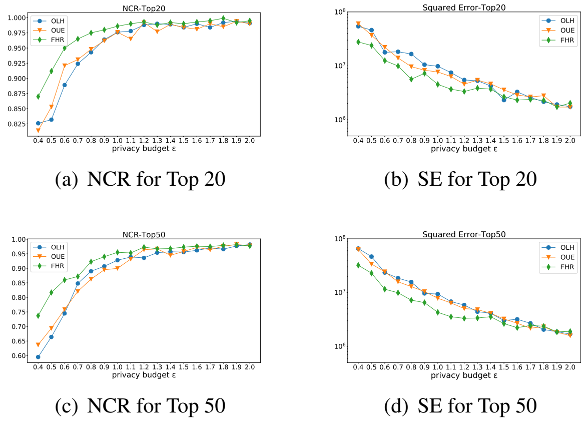
Lemma 1 (Upper bound). *Let $d = 2^r$. As described earlier, $f_t = \frac{1}{d}\sum_{i=1}^{n}(H(t) \cdot H(t))$ and $\hat{f}_t = \frac{e^{\varepsilon}+1}{2(e^{\varepsilon}-1)}\sum_{i=1}^{n}(b_i \cdot H(t))$. With at least $1-\beta$ probability,*

$$
\max|f_t - \hat{f}_t| = O(\frac{\sqrt{\log(1/\beta)}}{\varepsilon\sqrt{n}})
$$

4.4. Method comparisons

We then provide the experiment on *Online dataset* [1] from online retail. Figure 1 shows the performances of FO on different algorithms.

Figures show that: (a)The trendlines in these figures 1(a) 1(c) increase as the privacy budget ε increases, and the trendlines for FHR (green) are the highest. According to the definition of NCR [22], the higher the trendline, the more high-frequency items the FO collects. Thus, FHR performs best. (b)The trendlines decrease as the privacy budget ε increases, and the trendlines for FHR (green) are the lowest overall in figures 1(b) 1(d). According to the definition of SE [22], the lower the trendline, the more accurate the frequencies. FHR also performs best. Therefore, under the definition of FLDP, FHR algorithm can obtain more accurate estimation results while protecting users' privacy.



(a) NCR for Top 20

(b) SE for Top 20

(c) NCR for Top 50

(d) SE for Top 50

Fig. 1: Metric for Online Data

Note: NCR (Normalized cumulative rank) is the quality of a ranked function according to the top-k value. And NCR is used to evaluate the identification of top-k values. SE (Squared error) is used to evaluate the error of the frequencies of top-k values.

5. CONCLUSION

LDP is overly conservative, preventing its application in practice. In this study, we proposed FLDP to weaken LDP in order to provide more flexible notions. Under our notion, we classify the input based on the output in order to obtain more accurate statistics without decreasing disturbance parameters while ensuring LDP. Based on practical applications, this study presented the new FHR mechanism, which considers privacy protection, communication, and computational complexity. Finally, theoretical analysis and experiments demonstrated that FHR can obtain an accurate FO.

[1] http://fimi.uantwerpen.be/data/

6. REFERENCES

[1] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography Conference*. Springer, 2006, pp. 265–284.

[2] Cynthia Dwork, Aaron Roth, et al., "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[3] Salil Vadhan, "The complexity of differential privacy," in *Tutorials on the Foundations of Cryptography*, pp. 347–450. Springer, 2017.

[4] Ninghui Li, Min Lyu, Dong Su, and Weining Yang, "Differential privacy: From theory to practice," *Synthesis Lectures on Information Security, Privacy, & Trust*, vol. 8, no. 4, pp. 1–138, 2016.

[5] Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Úlfar Erlingsson, "Scalable private learning with pate," *arXiv preprint arXiv:1802.08908*, 2018.

[6] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra 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.

[7] Zitao Li, Tianhao Wang, Milan Lopuhaä-Zwakenberg, Ninghui Li, and Boris Škoric, "Estimating numerical distributions under local differential privacy," in *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, 2020, pp. 621–635.

[8] Ruixuan Liu, Yang Cao, Hong Chen, Ruoyang Guo, and Masatoshi Yoshikawa, "Flame: Differentially private federated learning in the shuffle model," *arXiv preprint arXiv:2009.08063*, 2020.

[9] ADP Team, "Learning with privacy at scale," *Apple Mach. Learn. J*, vol. 1, no. 9, 2017.

[10] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin, "Collecting telemetry data privately," in *Advances in Neural Information Processing Systems*, 2017, pp. 3571–3580.

[11] Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha, "Locally differentially private protocols for frequency estimation," in *Proc. of the 26th USENIX Security Symposium*, 2017, pp. 729–745.

[12] Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha, "Locally differentially private protocols for frequency estimation," in *26th {USENIX} Security Symposium ({USENIX} Security 17)*, 2017, pp. 729–745.

[13] Rui Chen, Haoran Li, AK Qin, Shiva Prasad Kasiviswanathan, and Hongxia Jin, "Private spatial data aggregation in the local setting," in *2016 IEEE 32nd International Conference on Data Engineering (ICDE)*. IEEE, 2016, pp. 289–300.

[14] Mehmet Emre Gursoy, Acar Tamersoy, Stacey Truex, Wenqi Wei, and Ling Liu, "Secure and utility-aware data collection with condensed local differential privacy," *IEEE Transactions on Dependable and Secure Computing*, 2019.

[15] Takao Murakami and Yusuke Kawamoto, "Utility-optimized local differential privacy mechanisms for distribution estimation," in *28th {USENIX} Security Symposium ({USENIX} Security 19)*, 2019, pp. 1877–1894.

[16] Xiaolan Gu, Ming Li, Li Xiong, and Yang Cao, "Providing input-discriminative protection for local differential privacy," in *2020 IEEE 36th International Conference on Data Engineering (ICDE)*. IEEE, 2020, pp. 505–516.

[17] Raef Bassily, Kobbi Nissim, Uri Stemmer, and Abhradeep Guha Thakurta, "Practical locally private heavy hitters," in *Advances in Neural Information Processing Systems*, 2017, pp. 2288–2296.

[18] Mark Bun, Jelani Nelson, and Uri Stemmer, "Heavy hitters and the structure of local privacy," in *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*. ACM, 2018, pp. 435–447.

[19] Cynthia Dwork, "Differential privacy: A survey of results," in *International Conference on Theory and Applications of Models of Computation*. Springer, 2008, pp. 1–19.

[20] Graham Cormode, Somesh Jha, Tejas Kulkarni, Ninghui Li, Divesh Srivastava, and Tianhao Wang, "Privacy at scale: Local differential privacy in practice," in *Proceedings of the 2018 International Conference on Management of Data*. ACM, 2018, pp. 1655–1658.

[21] NIE Yiwen, Wei Yang, Liusheng Huang, Xike Xie, Zhenhua Zhao, and Shaowei Wang, "A utility-optimized framework for personalized private histogram estimation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 4, pp. 655–669, 2018.

[22] Tianhao Wang, Ninghui Li, and Somesh Jha, "Locally differentially private frequent itemset mining," in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 127–143.