差分隐私的数学原理

柴健喆（2020090917002）

信息与软件工程学院

成都，四川，中国

Chaijainzhe@gmail.com

摘要：我们从噪声算法的层面将差分隐私技术分为三个部分：拉普拉斯机制，高斯机制以及指数机制。我们从数学的层面深入探究了其原理并给出了证明。此外，我们还简明的定义了差分攻击与ε-差分隐私、ε-δ差分隐私，并探讨了参数ε，δ对隐私与准确性的影响。

关键词：差分隐私，直方图，噪声

Mathematics in Differential Privacy

Jianzhe Chai(2020090917002)

College of Information and Software Engineering

Chengdu,Sichuan,China

Chaijainzhe@gmail.com

ABSTRACT :We divide Differential Privacy into three parts: Laplace scheme, Gauss scheme and Exponential scheme and deeply explore them from the mathematical level and given their proof. In addition, we concisely define differential attack and ε- Differential Privacy, ε-δ Differential privacy and show how parameters ε，δ have an impact on privacy and accuracy.

KEYWORDS: Differential Privacy,Histograms,Dummy Contributions

1. 引言

差分隐私[1]是为了在敏感数据上进行数据分析而发展起来的一套机制，通过混淆数据库查询结果，来实现数据在个人层面的隐私性，并且保证查询结果近似正确。它可以量化用户隐私保护程度，并且能抵御攻击者发起的背景知识攻击和合成攻击[23]。在最好的情况下，不同的差分隐私算法可以使被保护数据广泛用于准确的数据分析，而无需借助于其他数据保护机制。尽管如此，数据的有效性最终还是会被消耗掉：信息恢复基本定律指出，对太多问题过于准确的答案将以一种惊人的方式破坏数据隐私。差分隐私算法研究的目标就是尽可能地避免这种对数据隐私的破坏。[4]如今，差分隐私算法连同其变体已经较为广泛的应用于医疗，金融等对数据安全有高合规要求的行业，例如：2016 年 6 月 13 日，苹果公司宣布其在 i0S10

中使用差异隐私，以改进其虚拟助理和建议技术。

从网络拓扑结构来看，差分隐私技术主要分为中心化的差分隐私机制与本地差分隐私机制。中心化的差分隐私机制[5]提供了隐私保证和准确性之间最著名的权衡。然而，它们依赖于一个强假设，即一个可信第三方可以访问整个数据集。本地差分隐私机制[6]通过将隐私机制分配给客户端，减轻了中央管理员的隐私影响，但这在精度上的成本很高[7]。从噪声产生算法来看，差分隐私技术主要分为拉普拉斯机制，高斯机制以及指数机制。本文从噪声算法方面探究了差分隐私的数学原理以及其隐私保护。

2. 差分隐私数学定义

2.1 差分攻击

假设一种情景：在数学文化课的期末考试后，老师拥有一张成绩表（Figure 1）

Name	Gender	Grade
Alice	Female	PASS
Bob	Male	FAIL
Johan	Male	PASS
Vuk	Female	PASS
Yoshi	Female	PASS
Abe	Male	FAIL

Figure 1：数学文化考试通过情况

其中，姓名与性别是公开信息，考试通过情况是隐私信息。我们开放这张表的查询而不想暴露任何人的隐私。比如我们去查询 Alice 的考试通过情况，这当然暴露了 Alice 的隐私，所以我们拒绝这种查询。而如果我们查询所有人员考试通过的人数并得到一个数对（通过人数，未通过人数），我们认为此查询为一个宏观的查询，并不会暴露任何个人的隐私，我们将允许此查询，并得到结果为（4，2）在此基础上，如果我们再去查询所有男生的通过情况，直观上此查询也并没有暴露任何人的隐私，但如果我们允许此查询并得到一个数组（1，2），若有一个诚实但好奇（honest but curious）1的客户端将以上两个数组相减，便可以显然的得出女生的通过情况（3，0）。这个差值结果暴露了 Alice 的考试通过情况。

为叙述方便，我们宏观的定义本文提到的的差分攻击：

定义 1：假设在数据集 Φ 上存在一个查询过程 F(X)，其中 $X \subseteq \Phi$。我们称（X,F(X)）为在 Φ 上对 X 的一次查询。在两次查询（X,F(X)）与（$X+\Delta$,F($X+\Delta$)），$\Delta \subseteq \Phi$ 中，我们将

[1] 诚实但好奇（honest but curious），指一个协议参与方虽然会严格遵守本协议，但会做出本协议规定之外的行为。

求（X,F(X)）与（X+Δ,F(X+Δ)）的 L1 距离（Δ,F(X+Δ)-F(X)）定义为差分。若过程 F(Δ)在Φ中不被允许而 F(X+Δ)-F(X)=F(Δ)，我们称这次差分构成一次对Φ的差分攻击。其中，F(X+Δ)-F(X)被称为额外知识。

在上述情境下，我们很容易对此数据集进行差分攻击。所以我们需要一种机制，能够对用户查询提供有用的响应，又能够保护数据中的个人隐私。然而现实是，又提供完全精确的查询响应，又能够完全保护隐私，这种条件成立的场景很少见。但是，我们可以通过随机化或近似查询的响应，放松一点精确性，但是数据有较好的隐私性。我们设计一种机制，当客户端对Φ进行一次查询过程 F(X)时，我们不返回(X,F(X)),而是返回(X,F(X)+κ)，其中κ为一个随机变量，服从概率质量函数 K。此时我们对Φ进行差分攻击，设第一次查询为(X,F(X)+κ_1)，第二次查询为(X+Δ,F(X+Δ)+κ_2)，此时差分结果为 F(X+Δ)-F(X)+κ_2-κ_1，显然不等于 F(Δ)。

由此我们认为，此机制在牺牲一定的查询精确度的条件下，显著的提高了对隐私的保护。

2.2 隐私定义

在查询前后对单个人的隐私认知几乎不发生变化”（semantic privacy），即语义隐私。在此节我们将较为严格的定义语义隐私，并给出一种差分隐私的定义，并说明它们的联系

在此之前，我们先约定以下记号：

Parameter	Description
D	一个集合，数据库的每一行在这个集合里取值
n	数据库的行数
x	数据库
f()	查询函数
M(x)	附加到查询 q 的随机化机制
X_i	X 为数据库随机变量，下标 i 表示 X 的第 i 行
P	在数据库行上的映射
Pr[]	概率分布

Figure 2 参数约定

定义 2：一个定义域为D^n的机制 M，如果对于每个 i∈[n]，每个数据库 X，每种 P，每种可能的 M(x)的输出 y，都满足：

$$e^{-\varepsilon}Pr[P(X_i)] \le Pr[P(X_i)|M(X)=y] \le e^{\varepsilon}Pr[P(X_i)]$$

则称机制 M 满足ε-语义隐私[8]

考虑一种极端情况，当 ε =0 时，会出现

$$Pr[P(X_i)] \leq Pr[P(X_i)|M(X) = y] \leq Pr[P(X_i)]$$

由夹逼定理得：

$$Pr[P(X_i)] = Pr[P(X_i)|M(X) = y]$$

也就是说在得到 M 的查询响应 y 前后，P 在第 i 行上的概率是完全相同的，也可以说在第 i 行上的先验概率和后验概率是相同的，观察 M 的响应后不会揭露第 i 行任何的附加信息。因此， ε 又被称为隐私预算。

下面我们定义 ε -差分隐私，差分隐私比语义隐私在使用上更方便，并且差分隐私蕴含语义隐私。

对于两个分布的相似情况，我们使用相对熵（KL-Divergence）来衡量：

$$\mathrm{D(Y||Z)} = \mathrm{E_{y\sim Y}}[\ln\frac{\mathrm{Pr[Y=y]}}{\mathrm{Pr[z=y]}}]$$

但是我们并不关心这两个分布的整体差异，我们只需要两个分布在差距最大的情况下能够被 ε约束，所以引入了 MAX-Divergence，并且使得它小于 ε:

$$\mathrm{D_\infty(Y||Z)} = \max_{S \subset \mathrm{Supp(y)}} [\ln\frac{\mathrm{Pr[Y=y]}}{\mathrm{Pr[z=y]}}] = \max_{y \in Y} [\ln\frac{\mathrm{Pr[Y=y]}}{\mathrm{Pr[z=y]}}] \leq \boldsymbol{\varepsilon}$$

化简得：

定义 3：如果对于每一对只有一行不相同的数据库 X 和X'以及每种可能的 M(X)的输出 y，都满足：

$$Pr[M(X) = y] \leq e^{\varepsilon} Pr[M(X') = y]$$

则说机制 M 满足 ε -差分隐私。

当 ε =0 时，会出现：

$$Pr[M(X) = y] \leq Pr[M(X') = y]$$

再令 X=X',X'=X,得到：

$$Pr[M(X) = y] \geq Pr[M(X') = y]$$

综合得到：

$$Pr[M(X) = y] = Pr[M(X') = y]$$

只有 M(X)与 X 无关，此等式才会成立。因此当 ε =0 时，可以完美的保护数据的隐私。但代价是，M（X）与 X 无关。也就是说牺牲了所有的查询精确度。如图 3，一般而言， ε 越小，隐私保护越好，但是加入的噪声就越大，数据可用性就下降了。实际上，我们不会取 ε =0，而是一个很小的数，以此来在隐私保护与查询精确度之间折中。

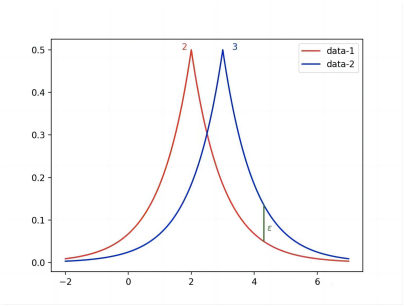

Figure 3-1 ε −差分隐私

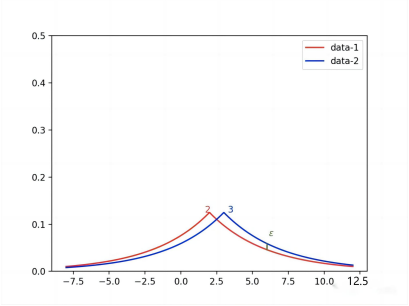

Figure 3-2 更大噪声水平下的 ε −差分隐私

Figure 3-3

可以证明的是，如果 M 是满足 ε −差分隐私的，则 M 满足 ε −语义隐私。证明过程参考[8]

2.3 ε − δ 差分隐私

ε −差分隐私太过严格，在实际的应用中需要很多的隐私预算。因此为了算法的实用性，我们引入松弛版本的差分隐私： ε − δ 差分隐私。

定义 4：如果对于每一对只有一行不相同的数据库 X 和X'以及每种可能的 M(X) 的输出 y，都满足：

$$Pr[M(X) = y] \leq e^{\varepsilon} Pr[PM(X') = y] + \delta$$

则说机制 M 满足 ε − δ 差分隐私。

此时我们来考察对应的 MAX-Divergence：

$$D_{\infty}^{\delta}(Y||Z) = \underset{S \subset Supp(y);\, Pr[Y \in S] \geq \delta}{MAX} [\ln \frac{Pr[Y = y]}{Pr[z = y]}] = \underset{y \in Y}{MAX} [\ln \frac{Pr[Y = y] - \delta}{Pr[z = y]}] \leq \varepsilon$$

对比上文，我们发现分子上减小了一个δ，是一个可以容忍的小差距。直观形式如图 4，在图中标注的位置差值大于 ε，但是我们考虑松弛项 δ，整体依旧满足差分隐私。一般而言， δ 设置的较小。δ 也被称为失败概率。

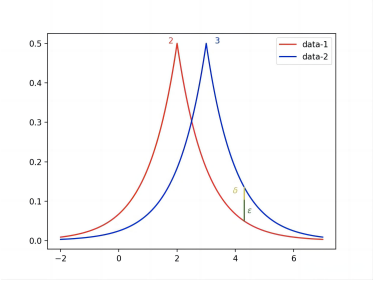

Figure 4 ε－δ 差分隐私

3. 噪声机制

在上节中，我们探讨了 ε 差分隐私与 ε－δ 差分隐私，但未对附加到查询 q 的随机化机制 M()做出显式的说明。在本节，我们将讨论 M()的实现方式。

随机化机制 M()的实现与差分隐私的数据有密切的关系。一类是数值型的数据，如数据集中已婚人士的数量；另一类是非数值型的数据,如喜欢人数的最多的颜色[9]。数值型的数据一般采用 Laplace 或者 Gauss 机制，对得到数值结果加入随机噪声即可实现差分隐私；而对于非数值型的数据，一般采用指数机制并引入一个打分函数，对每一种可能的输出都得到一个分数，归一化之后作为查询返回的概率值。

以下对这三种机制进行讨论

3.1 Laplace 机制

我们为了讨论方便，可以表示数据库 x 为一个直方图向量， 每个分量 xi 代表全集 D 中的一种记录出现的次数，所以 $x \in \mathbb{N}^{|D|}$，$\mathbb{N}$表示所有非负整数的集合。

用直方图表示数据库可以在数学上方便地定义数据库的距离，比如 L1 距离。

定义 5 ：数据库之间的距离 L1

一个数据库 x 的 L1 范数表示为 $||x||_1$ ，其定义为

$$||x||_1 = \sum_{i=1}^{|D|} |x_i|$$

数据库 x 和 y 的 L1 距离为$||\boldsymbol{x} - \boldsymbol{y}||_{\mathbf{1}}$

定义 6 ：L1 敏感度

$$\Delta f = \max_{x,y \in \mathbb{N}^{|D|}, ||\boldsymbol{x}-\boldsymbol{y}||_{\mathbf{1}} = \mathbf{1}} ||f(D) - f(D')||_1$$

敏感度刻画了：单个记录改变 f 的输出，最大能改变多少。

Laplace 机制依靠于 Laplace 分布。Laplace 分布是一个连续分布，均值为 μ=0 的拉普拉斯分布的概率密度函数为：

$$Lap(x|b) = \frac{1}{2b} e^{-\frac{|x|}{b}}$$

理论 1 ：Laplace 机制

$$M_L(x, f(\cdot), \varepsilon) = f(x) + \sum_i Yi$$

其中 Yi 是从 Yi~Lap(Δf/ε)采样的独立同分布（i.i.d.）的随机变量。

定理 1：Laplace 机制满足（ε，0）差分隐私，即

$$\frac{Pr[M(x) \in S]}{Pr[M(y) \in S]} \leq e^{\varepsilon}$$

证明：令p_x为$M_L(x, f(\cdot), \varepsilon)$的概率密度函数，令$p_y$为$M_L(y, f(\cdot), \varepsilon)$的概率密度函数；x、y$\in \mathbb{N}^{|D|}$且满足$||\boldsymbol{x}-\boldsymbol{y}||_1=1$；f(·)是 f:$\mathbb{N}^{|D|} \rightarrow \mathbb{R}^k$。我们在任意点 z$\in \mathbb{R}^k$上比较：

$$\frac{p_x(z)}{p_y(z)}$$

$$= \prod_{i=1}^{k} \left(\frac{e^{-\frac{\varepsilon|f(x)_i - z_i|}{\Delta f}}}{e^{-\frac{\varepsilon|f(y)_i - z_i|}{\Delta f}}}\right)$$

$$= \prod_{i=1}^{k} \left(e^{\frac{\varepsilon(|f(x)_i - z_i| - |f(y)_i - z_i|)}{\Delta f}}\right)$$

$$\leq \prod_{i=1}^{k} \left(e^{\frac{\varepsilon(|f(x)_i - f(y)_i|)}{\Delta f}}\right)$$

$$= \prod_{i=1}^{k} \left(e^{\frac{\varepsilon||f(x)-f(y)||_1}{\Delta f}}\right)$$

$$\leq e^{\varepsilon}$$

其中，第一个不等式由绝对值不等式容易得到，第二个不等式由敏感度 Δf 的定义也容易得到。

最后，利用概率密度函数在任意 S$\in \mathbb{R}^k$上积分(若是离散变量则为求和)，可得

$$\frac{Pr[M(x) \in S]}{Pr[M(y) \in S]} \leq e^{\varepsilon}$$

证毕。

3.2 Gaus 机制

Laplace 机制为严格的（ε，0）差分隐私，与此相对的，Gaus 机制为带有松弛项的（ε，δ）差分隐私。

定义 7：L2 敏感度

$$\Delta f = \max_{x,y \in \mathbb{N}^{|D|}, ||\boldsymbol{x}-\boldsymbol{y}||_2 = 1} ||f(x) - f(y)||_2$$

在 Gaus 机制中，相对的，L1 范数被替换成为 L2 范数，这与噪声的分布不同有关。

Gaus 机制依靠于 Gaus 分布。Gaus 分布是一个连续分布，均值为 μ=0 的拉普拉斯分布的概率密度函数为：

$$N(0, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

理论 2 :Gaus 机制

$$M_G(x, f(\cdot), \varepsilon, \delta) = f(x) + \sum Yi$$

其中 Yi 是 Yi~$N(0, \sigma^2)$采样的独立同分布（i.i.d.）的随机变量；δ$\in(0,1)$,σ >

$\frac{\sqrt{2\ln(1.25/\delta)}\,\Delta f}{\varepsilon}$

高斯机制的定义明显比 Laplace 要复杂，这里主要有三个参数，

高斯分布的标准差 σ，这决定了噪声的尺度；

ϵ 表示隐私预算，和噪声成负相关；

δ 表示松弛项，比如设置为 10^{-5}，就表示只能容忍 10^{-5} 的概率违反严格差分隐私。

定理 2：Gaus 机制满足（ε，δ）差分隐私。即

$$\Pr[\mathrm{M}(\mathrm{D}) \in S] \leq e^{\varepsilon} Pr[PM(D') \in S] + \delta$$

证明：本证明较复杂，严格证明请参考[10]。本文只指出证明思路:

在松弛差分隐私中，输出可以分为两部分，一部分是严格遵守差分隐私的，另一部分是违反了严格差分隐私的。如下，S1 表示准守严格差分隐私的部分，S2 表示违反严格差分隐私的部分。因此我们需要将输出集合分隔成两部分，证明第一部分是被 ϵ 约束住，而第二部分小于 δ 。

$\Pr[\mathrm{f}(\mathrm{x}) + \mathrm{x} \in S | x \sim \mathrm{N}(0, \sigma^2)]$

$= \Pr[\mathrm{f}(\mathrm{x}) + \mathrm{x} \in S1 | x \sim \mathrm{N}(0, \sigma^2)] + \Pr[\mathrm{f}(\mathrm{x}) + \mathrm{x} \in S2 | x \sim \mathrm{N}(0, \sigma^2)]$

$\leq \Pr[\mathrm{f}(\mathrm{x}) + \mathrm{x} \in S1 | x \sim \mathrm{N}(0, \sigma^2)] + \delta$

$\leq \mathrm{e}^{\varepsilon} \Pr[\mathrm{f}(\mathrm{y}) + \mathrm{x} \in S1 | x \sim \mathrm{N}(0, \sigma^2)] + \delta$

3.3 指数机制

与前两种机制不同，前面两种都是简单地对输出的数值结果加入噪声实现差分隐私。而对于非数值型数据而言，它的输出是一组离散数据{R1,R2,...,RN}中的元素。

指数机制整体的思想就是，当接收到一个查询之后，不是确定性的输出一个 Ri 结果，而是以一定的概率值返回结果，从而实现差分隐私。而这个概率值则是由打分函数确定，得分高的输出概率高，得分低的输出概率低。

定义 8：L1 敏感度

$$\Delta \mathrm{q} = \underset{\mathrm{x,y} \in \mathbb{N}^{|D|}, ||\boldsymbol{x}-\boldsymbol{y}||_1 = \mathbf{1}}{\mathrm{MAX}} ||\mathrm{q}(\mathrm{x}, \mathrm{R_i}) - \mathrm{q}(\mathrm{y}, \mathrm{R_i})||_1$$

理论 3：指数机制

$$\mathbf{M(D, q, R_i) \sim e^{\frac{\varepsilon q(D, R_i)}{2\Delta q}}}$$

注意其中$\mathbf{e^{\frac{\varepsilon q(D, R_i)}{2\Delta q}}}$**不表示概率值，需要进一步归一化，得到概率值。**

$$\mathbf{Pr[R_i] = \frac{e^{\frac{\varepsilon q(D, R_i)}{2\Delta q}}}{\sum_j e^{\frac{\varepsilon q(D, R_j)}{2\Delta q}}}}$$

定理 3:指数机制满足(ε,0)−差分隐私

证明：

$$\frac{Pr[M_e(x,u,R)=r]}{Pr[M_e(y,u,R)=r]}=\frac{\frac{e^{\frac{\varepsilon q(x,r)}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}}{\frac{e^{\frac{\varepsilon q(y,r)}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(y,r')}{2\Delta q}}}}$$

$$=\frac{e^{\frac{\varepsilon q(x,r)}{2\Delta q}}}{e^{\frac{\varepsilon q(y,r)}{2\Delta q}}}\times\frac{\sum_{r'\in R}e^{\frac{\varepsilon q(y,r')}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}$$

$$=e^{\frac{\varepsilon(q(x,r)-q(y,r))}{2\Delta q}}\times\frac{\sum_{r'\in R}e^{\frac{\varepsilon q(y,r')}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}$$

$$\leq e^{\frac{\varepsilon}{2}}\times e^{\frac{\varepsilon}{2}}\times\frac{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}$$

$$=e^{\varepsilon}$$

其中，$e^{\frac{\varepsilon(q(x,r)-q(y,r))}{2\Delta q}}$项由敏感度定义可推出其等于$e^{\frac{\varepsilon}{2}}$，$\frac{\sum_{r'\in R}e^{\frac{\varepsilon q(y,r')}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}\leq\frac{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')+\Delta}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}=e^{\frac{\varepsilon}{2}}\times\frac{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}{\sum_{r'\in R}e^{\frac{\varepsilon q(x,r')}{2\Delta q}}}$

证毕

4. 结语

本文主要陈述的是基础的三种差分隐私的机制，但是差分隐私的机制本质上是通过加噪声实现的[11]，但是通过对同一数据集多次频繁的查询，利用平均也可以大致推断出一些隐私信息。这种问题本质上就是对同一数据集进行多次查询花费了大量的隐私预算，在 2.2 中得到的结论是，隐私预算和可用性成正比，和隐私保护成反比，大量的隐私预算必然造成隐私保护能力下降。然而在实际的算法应用中，比如决策树或者神经网络，我们必须要频繁的访问数据，因此为了解决这个问题，人们提出了组合定理（Composition Theorem）[10]。利用强组合定理用很小的δ来换取较大的ε。在差分隐私领域的经典论文[12]提出了 Moments Accountant 算法大大减少了隐私预算ε。

与此同时，差分隐私技术也与密码学领域的其他方向产生了融合，如（全）同态加密，多方安全计算体制中的混淆电路[7]，在业界也已经有了众多的开源模型。需要注意的是，这些尝试在提高了隐私保护性能的同时，也大大增加了算法的通讯开销与计算开销。如何减少这些开销已经成为差分隐私领域的热点问题。

参考文献

1 DWORK C, ROTH A. The algorithmic foundations of differential privacy[J]. Foundations and Trends in Theoretical Computer Science, 2013, 9(3/4): 211-407
2 DEWRI R, THURIMELLA R. Exploiting service similarity for privacy in location-based search queries[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25 (2): 374-383
3 MEDKOVÁ J. Composition attack against social network data[J]. Computers & Security, 2018, 74: 115-129
4 https://zhuanlan.zhihu.com/p/522721297
5 Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrat ing Noise to Sensitivity in Private Data Analysis. In TCC
6 Alexandre Evfimievski, Johannes Gehrke, and Ramakrishnan Srikant. 2003. Lim iting privacy breaches in privacy preserving data mining. In PODS. 211–222.
7 James Bell, Adrià Gascón, Badih Ghazi, Ravi Kumar, Pasin Manurangsi, Mariana Raykova, and Phillipp Schoppmann. 2022. Distributed, Private, Sparse Histograms in the Two-Server Model. In Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22), November 7–11, 2022, Los Angeles, CA, USA. ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3548606.3559383
8 C. Dwork and A. Roth. The Algorithmic Foundations of Differential Privacy. Foundations and Trends R in Theoretical Computer Science, vol. 9, nos. 3–4, pp. 211–407, 2014
9 https://zhuanlan.zhihu.com/p/144318152
10 Foundations and Trends R in Theoretical Computer Science Vol. 9, Nos. 3–4 (2014) 211–407 c 2014 C. Dwork and A. Roth DOI: 10.1561/0400000042
11 https://aircloak.com/explaining-differential-privacy/
12 Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (ACM CCS), pp. 308-318, 2016