**The Algorithmic Foundations of Differential Privacy**

**差分隐私的算法基础**

**前言**

差分隐私可以被描绘成是一种”允许您的数据用于任何研究或分析，不会对您造成不利或其他影响”的技术。最佳情形下差分隐私允许机密数据直接使用，而不用对数据进行清洗或受限视图等，然而基于此的差分隐私的预算(data utility)可能会被很快的消耗完毕。差分隐私解决了”未知”和”已知”的矛盾，如医学数据库中得知吸烟者增加致癌风险从而提高保费，但差分隐私认为其不触犯病人的隐私，因为无论病人在数据库出席与否，结论都是一致的。具体来说，它确保任何输出序列（对查询的响应）本质上是同样可能发生，独立于任何个人，独立于是否在数据库中出现。而“本质上”这个术语是由一个参数ε捕捉的。较小的ε会产生更好的隐私（和更不准确的回应）。

**差分隐私是一个定义而非一种特定的算法实现**，其可以拥有多种算法实现，但是往往很难在ε很小时，实现一个高精度的ε-差分隐私算法。

**第一章 差分隐私的承诺The Promise of Difffferential Privacy**

1.1隐私保护数据分析Privacy-preserving data analysis

差异隐私是针对问题定制的隐私定义保护隐私的数据分析。一般说来，数据越丰富，就越有趣和有用。数据的丰富性使得“命名( Naming)”一个个体成为可能性，即使用一些特征(如邮政编码、出生日期、性别和最近收看的电影)，**将不同数据集中的匿名记录与非匿名记录匹配，从而进行链接攻击**。差异隐私**抵消了链接攻击**：因为差异隐私是数据访问机制的一个属性，并且与辅助信息的存在或不存在无关。

**一般认为(直接查询)大型的数据库是不安全的**，同上，攻击者可以通过得知多种属性后推断到需要得知的信息。又可知**针对数据库的审查是无法保证安全的，**其一有可能拒绝通过的数据库访问，其需要得到的数据集本身已经泄露；其二审计数据库访问在计算上不可证明有效，如果一个查询语言足够丰富，则数据审计一定不安全。**摘要统计(Summary Statistics)不是安全的，**理由同刚才描述过的攻击模式，如在人类基因组计划中，最早公众可以进行对数据库的部分有限访问，根据定义，一些受调查的等位基因频率只是聚合的统计数据，而（错误的）假设是，通过这种聚合，它们可以保护隐私。

**普通的事实不一定不会泄露隐私，**如T先生天天买面包，有一天他突然改变了习惯很少买面包了，可以推测出可能T先生患了糖尿病，无论其成立与否都对T先生的隐私产生了损失。**”少数人”隐私被损害不适用于差分隐私，**如果一种技术可以对大多数人的隐私起到保护作用，一些人可能认为其是足够的了，但这对于少数未受保护的人的隐私可以产生足够损害(这更多的是个社会学问题)，一般基于此伦理的隐私保护方法通过随机抽样来确定其算法的保护程度，但差分隐私并不适用于此伦理(Difffferential privacy

provides an alternative when the “just a few” philosophy is rejected

)。

1.2 书目注释 Bibliographic notes

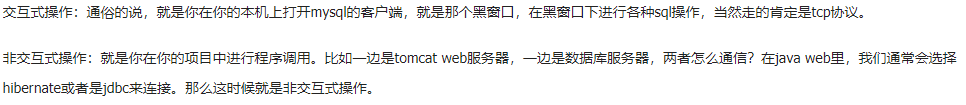
无

**第二章 基本术语 Basic Terms**

2.1 计算模型 The model of computation

假定有个**可信任的数据保管员**将个人数据保存在数据库D中，隐私保护的目的即对于对数据库D的统计分析中同时保护每一个查询行。离线数据库一般只用于普通查询，在项目结束后认为将对数据进行销毁；在线数据库该模型允许进行外部查询和分析。可信的数据保管员可以被集合运行的协议所替代，如使用加密技术来实现安全多方协议，但其不能推断到密码学安全上。当将要进行的查询已知时可提供最良好的查询结果，但当将要进行的查询未知时，数据库必须为所有可能的问题提供答案，(为了实现隐私保护)准确性必然会随着问题的数量而降低。

**非交互式情况**的定义为针对于当前模型的查询未提出任何问题时(If no queries are presented then we are in the non-interactive case)，此时希望查询结果可以回答未来的问题。



在某些情况下，我们可能进行多重查询并获得合成的查询集，在这种情况下，解码方法是执行查询合成数据库，然后应用某种转换，如乘以比例因子，以获得对查询的真正答案的近似。

2.2 定义私有数据分析 Towards defining private data analysis

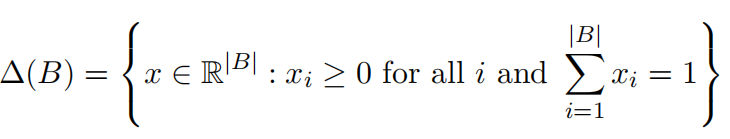
差分隐私要求对抗者(Adversary)对查询的前后视图（即访问数据库之前和之后的视图）不应“太不同了”，但如果数据库包含的数据足够丰富，隐私的概念是不可能实现的。**前后视图未改变也不能保证隐私未产生侵害，**如前面吸烟致癌的例子中辅助信息(X先生吸烟)导致了隐私泄露。“什么都没学到(Nothing is learned)”可以对应于密码学中的语义安全，**即只会从密文中知道密文对应的明文知识，而不会获得更多的知识，**但密文的接收者必须能够从密文中正确解密，否则只为了确保语义安全而完全丢失密文的解读性是微不足道的。

在标准的计算假设下，语义安全密码系统存在，但不意味着可以构建基于此安全的数据库系统，如语义安全密码系统中有三方，而私有数据系统中有两方等。不应质疑所有的辅助信息，有一些辅助信息可以提供帮助，但很难界定这种辅助信息的种类，即一种可以使使用者获得更多信息，而不会泄露隐私给恶意攻击者的辅助信息。

2.3 形式化的差分隐私 Formalizing differential privacy

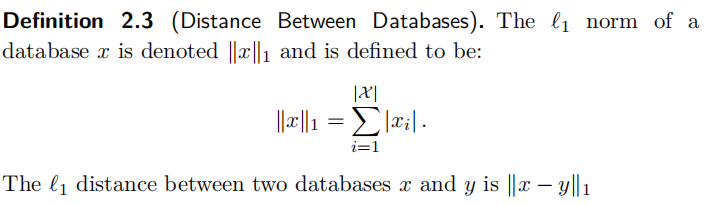
差异隐私将通过过程的进行提供隐私保护，特别是它会引入随机性。早期的随机化隐私保护通过**随机响应**机制来实现，举例如**随机投掷硬币并作出回答**。给出定义如下:

定义2.1（概率单纯形）。给定一个离散集B，B上的概率单纯形，表示为△（b）定义为：



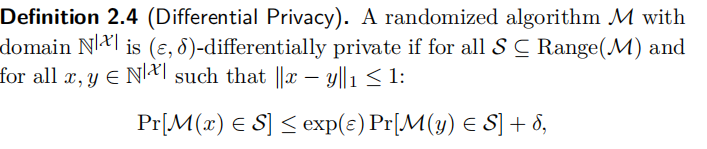
定义2.2（随机算法）

对于给定的非连续域A和非连续域B，有M:A→△（B），对于给定的输入a属于A，以概率((M(a))b)输出M(a)=b，其概率分布符合硬币概率。

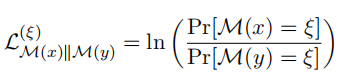
定义**l1距离**为下图，其中Xi表示数据集中type为i的数量，其中i属于

注意是数据库x大小的度量（即它包含的记录数）和是衡量X和Y之间的记录不同。

定义了差分隐私如下，Pr代表概率值，exp指自然常数e为底的幂函数，其中M仍然指符合硬币概率的概率方法，如果值为0时，称M为-差分隐私：



如果的值相近于1/，则是非常危险的，我们可以认为这类似于上面的”少数人”隐私模式。**对于差分隐私来说，保证在M(x)机制运行后在相邻的数据集中观测到的结果集相同，差分隐私则与其不同。**我们将视作输出结果集，则有定义**隐私损失**为：



**差异隐私不受后处理的影响**：数据分析师，如果没有关于私有数据库的额外知识，就无法计算私有算法m的输出函数，并使其具有较小的差异私有性。也就是说，如果一个算法保护了个人的隐私，那么数据分析师也不能增加隐私损失。

群组隐私(Group privacy) 也符合上述给出的差分隐私的定义，定义不再详述，**隐私保护的隐私(privacy guarantee)担保随着数据组规模的增大，而线性减小。**

更普遍地说，组合composition和群组隐私group是不一样的。

2.3.1 差分隐私的许诺 What differential privacy promises

**经济观点，**差分隐私在差分隐私机制M(x)未泄露的情形下，**差分隐私承诺无论是否将数据存入数据库，其被隐私泄露的风险是相同的。**基于此，对于被调查者来说，给予任何激励-从利他主义到金钱奖励-差别隐私可能会说服他允许他的数据被利用。通过保证ε-微分隐私，数据分析师可以向个人保证他预期的未来效用不会受到超过exp（ε）≈（1+ε）因子的损害。

2.3.2 差分隐私没有许诺什么 What difffferential privacy does not promise

隐私是一个极其强大的保证，但它**不承诺无条件的免受伤害**。更普遍的说，差**别隐私并不能保证一个人所相信的秘密仍然是秘密**。

**差异隐私的定性性质**：①防止任意风险(Protection against arbitrary risks)②中和关联攻击③可以量化的隐私损失，这允许在不同的技术之间进行比较。

④组合(Composition)(什么意思**？**)⑤集体隐私，差分隐私允许分析和控制家庭等群体所遭受的隐私损失。⑥差分隐私免疫数据发布后处理，一个没有额外知识的数据分析师无法计算差分隐私的函数，这无关于辅助信息(额外知识 additional knowledge和辅助信息auxiliary information的区别？)是否可用。

2.3.3定义的最后说明 Final remarks on the definition

**隐私的粒度，**如把社交网络中每个人拟定为图论中的顶点，个体之间的关系由边表示，我们可以在与个体对应的粒度级别上考虑差异隐私：也就是说，我们可以要求差异隐私算法对从图中添加或删除任何顶点不敏感，这提供了很强的隐私保证。另一方面，我们可以在一定程度上考虑差异隐私与边相对应的粒度，并要求我们的算法只对单个或小数字的添加或删除不敏感图的边，这种隐私保护能力较弱。在某些情况下，大量的社交关系可能不认为是敏感信息：例如，一个人可能觉得没有必要隐藏这样一个事实他大部分的接触对象都是他所在城市或工作场所的人，另一方面，可能有一小部分社会交往者存在高度敏感（例如未来的新雇主）。在这种情况下，边缘隐私应该足够以保护敏感信息，同时仍允许对数据进行更细致的分析。边缘隐私将保护此类个人的敏感信息，前提是其拥有少于1/ε的朋友。

**所有小的ε 是类似的，**当ε很小时，（ε，0）-差分隐私断言对于所有相邻数据库对x、y和all输出o时，攻击者无法根据观察o来区分哪个是真正的数据库。相反ε较大时，则可认为保护能力较弱。

**一些额外的形式化说明。**如可以增加一些辅助参数W作为输入。

2.4 书目注释 Bibliographic notes

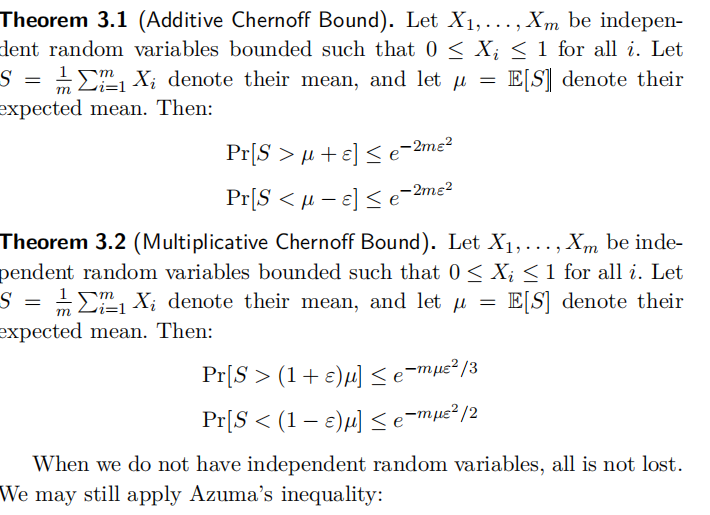
略

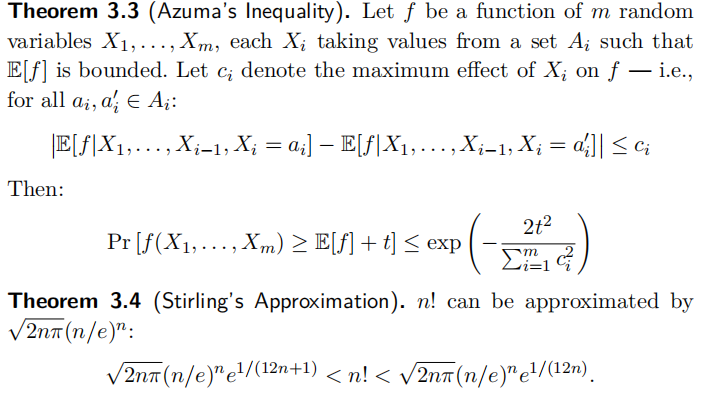
**第三章 基本技术与合成定理 Basic Techniques and Composition Theorems**

在本节中，我们将介绍差异隐私，如Laplace机制等。

3.1有用的概率工具

下面介绍几个不等式：



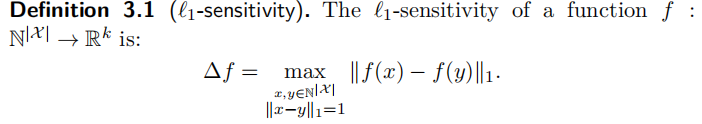


3.2 随机反应

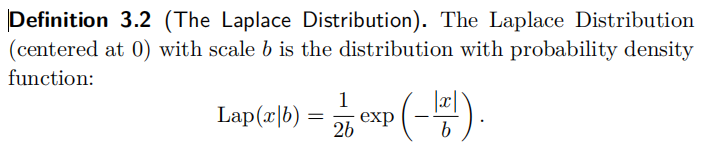
在第一章的例子中，随机化反应背后的直觉是它提供了“似是而非的可否认性”。因为第一个和第二个硬币都是正面的以1/4的概率发生。换句话说，**隐私是通过过程获得的**，没有“好”或“坏”的反应。可证明上述随机应答的版本是（ln 3，0）-差分隐私。

3.3 拉普拉斯机制

数值查询，即函数是最为常见的数据库查询方式， ℓ1 敏感度可以决定数据库将以多准确的方式来回答查询，定义ℓ1 敏感度如下:



ℓ1 敏感度表示在最坏的情形下f函数可以改变的幅度，直观地说，我们必须在反应中引入不确定性，以隐藏单个个体的参与，函数的灵敏度给出了我们必须对其输出进行多少扰动以保护隐私的上限，即拉普拉斯扰动。



我们现在定义拉普拉斯机制。顾名思义，拉普拉斯机制将简单地计算f，并用拉普拉斯分布的噪声扰动每个坐标。

