

Beijing Forest Studio
北京理工大学信息系统及安全对抗实验中心



差分隐私原理及应用

硕士研究生 郜森

2020年05月17日

- 背景介绍
- 基本概念
 - 差分隐私基本原理
 - 差分隐私变体(dx-privacy)
- 算法原理
 - geo-indistinguishability(地理位置不可分辨)
- 应用总结
- 参考文献

- 预期收获
 - 1. 了解数据脱敏的方法
 - 2. 了解差分隐私的基本原理与思想
 - 3. 了解差分隐私在地理位置脱敏中的应用

- 在大数据时代背景下数据滥用、数据窃取、隐私泄露以及“大数据杀熟”等数据安全问题呈徒增和爆发趋势。
- 各国相关法规
 - 如欧盟保护个人数据的《General Data Protection Regulation》
 - 美国的《California Consumer Privacy Act》
 - 中国实施的《中华人民共和国网络安全法》。
- 在法规指导下，如何降低法律风险和隐私泄露风险，同时也能满足业务场景需求就成为一个很关键问题！

- 什么是隐私呢？

信息统计

姓名	性别	外卖次数	地点
小红	女	1	北理
小张	男	2	北理
小刚	男	3	北外
小李	男	4	北理

- 1 北理比北外的更喜欢吃外卖
- 2 北理比北外的更喜欢吃外卖
- 3 北理只有一个女生喜欢吃外卖

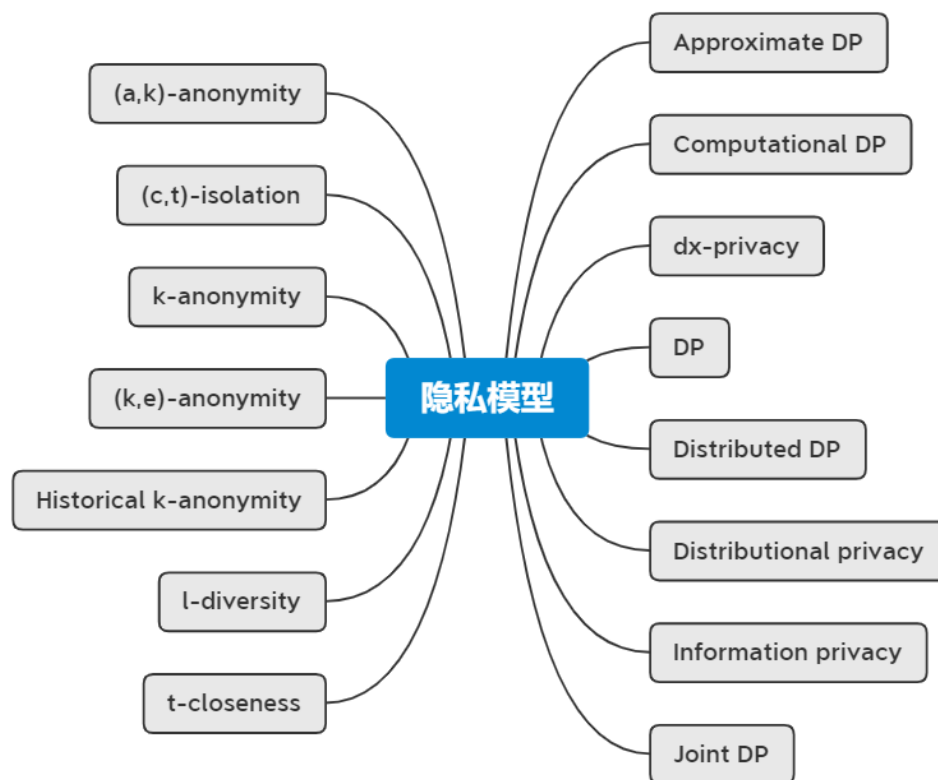
- 从隐私保护的角度来说，隐私是针对单个用户的概念，公开群体用户的信息不算是隐私泄漏，但是如果从数据中能准确推测出个体的信息，那么就算是隐私泄漏。

- 表格中的隐私属性分类：

标识符	个体的唯一标示，如姓名、电话等，这些内容需要在公开数据的时候删掉
准标识符	类似邮编、年龄、生日、性别等不是唯一的，但是能帮助研究人员关联相关数据的标示
敏感数据	比如说购买偏好、薪水等，这些数据是研究人员最关心的。一般都直接公开。

- 隐私保护通常指最小化个人身份或属性泄露的风险。
 - 1977年，提出第一个隐私保护模型
攻击者对数据的访问不应增加攻击者对个人隐私的了解。
 - 2002年，提出了k-anonymity
除敏感数据外，其他属性组合相同的值至少有K个记录。
 - 2006年，提出了l-diversity
 - 2006年，提出了Differential Privacy（差分隐私）
 - 2007年，提出了t-Closeness

- 隐私保护机制可以分为两类：
 - 差分隐私及其变种
 - k-anonymity及其变种



- K-anonymity类缺点
 - k-anonymity, l-diversity, t-closeness 不能抵御差分攻击。

姓名	性别	外卖次数	地点
小红	女	1	北理
小张	男	2	北理
小刚	男	3	北外
小李	男	4	北理

– $C(3)=6, C(4)=10$  $N(4)=4$

- 差分隐私定义：
 - 对于一对相邻数据集 D 和 D' ，查询获得相同结果的概率非常接近。 ϵ -差分隐私定义如下：

随机化算法 M 是在 D 上做任意查询操作，查询后对数据结果添加噪声

两个数据库加上统一随机噪声之后查询结果为 c 的概率

$$e^{-\epsilon} \leq \frac{\Pr(M(D) = c)}{\Pr(M(D') = c)} \leq e^{\epsilon}$$

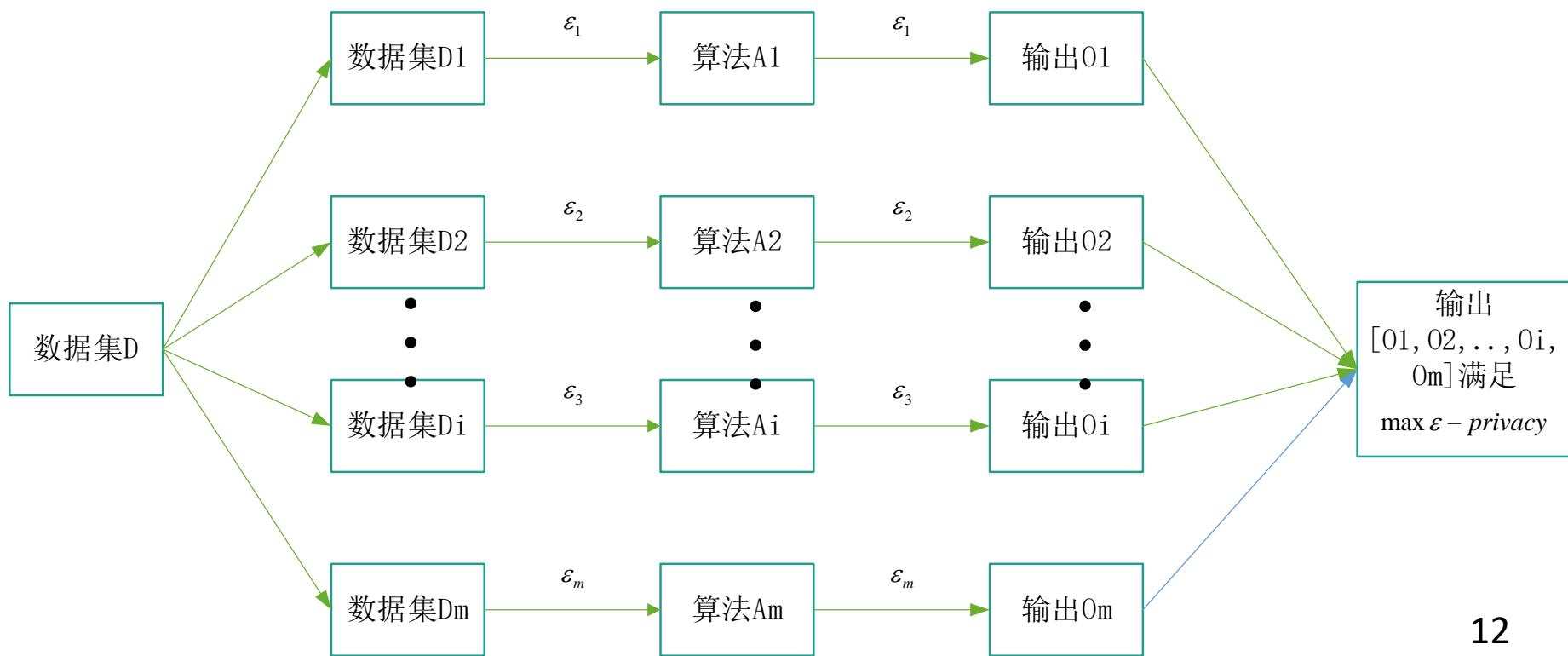
D 和 D' 是相邻数据集

隐私预算，反映差分隐私保护程度。

- 差分隐私含义解释：
 - 无论一个个体数据是否在数据库中，攻击者通过输出获得近乎相同的信息。
 - 无论攻击者拥有多少辅助信息，他都没办法从输出中获得更多的隐私信息（语义安全性）。
 - 仅意味无论数据是否在数据库中，隐私泄露程度几乎是一样的。
- 相邻数据集（汉明距离小于等于1）
 - 1、D删除一条数据
 - 2、D修改一条数据

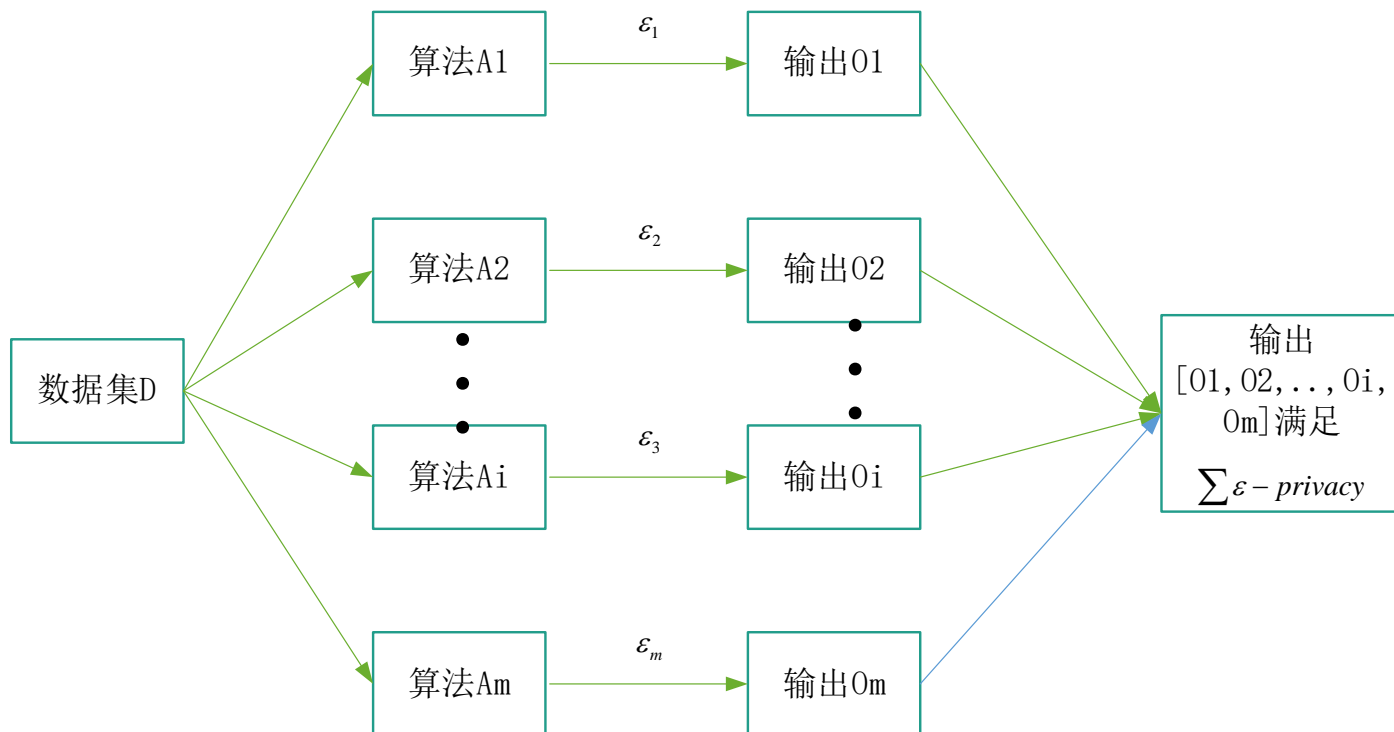
- 差分隐私并行性质:

- 若 D_1, D_2, \dots, D_m 分别表示输入数据集 $A_1(D_1), A_2(D_2), \dots, A_m(D_m)$ 为一系列满足 ϵ -差分隐私算法, 且算法间相互独立, 则组合算法也满足 ϵ -差分隐私。



- 差分隐私串行性质

- 给定数据集D以及关于D的一组差分隐私算法，算法间随机过程相互独立，则组合算法满足 $\sum_{i=1}^m \epsilon_i$ - 差分隐私



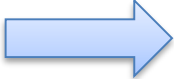
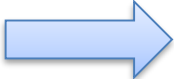
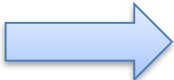
* 每访问一次数据，就会扣除一些预算，隐私保护力度下降³

- 差分隐私优缺点
 - 优点
 - 不依赖于背景知识（假定敌手拥有除了某个体之外的所有其他知识。）
 - 隐私级别可控制，隐私泄露程度可量化
 - 缺点
 - 对数据可用性损害严重
 - 适用范围窄

- 变体 d_x -privacy 的定义:

$$e^{-\varepsilon d(D,D')} \leq \frac{\Pr(M(D) = c)}{\Pr(M(D') = c)} \leq e^{\varepsilon d(D,D')}$$

- 当 $d(D,D')$ 为汉明距离时, 即为标准差分隐私, 该机制是对标准差分隐私的扩展。

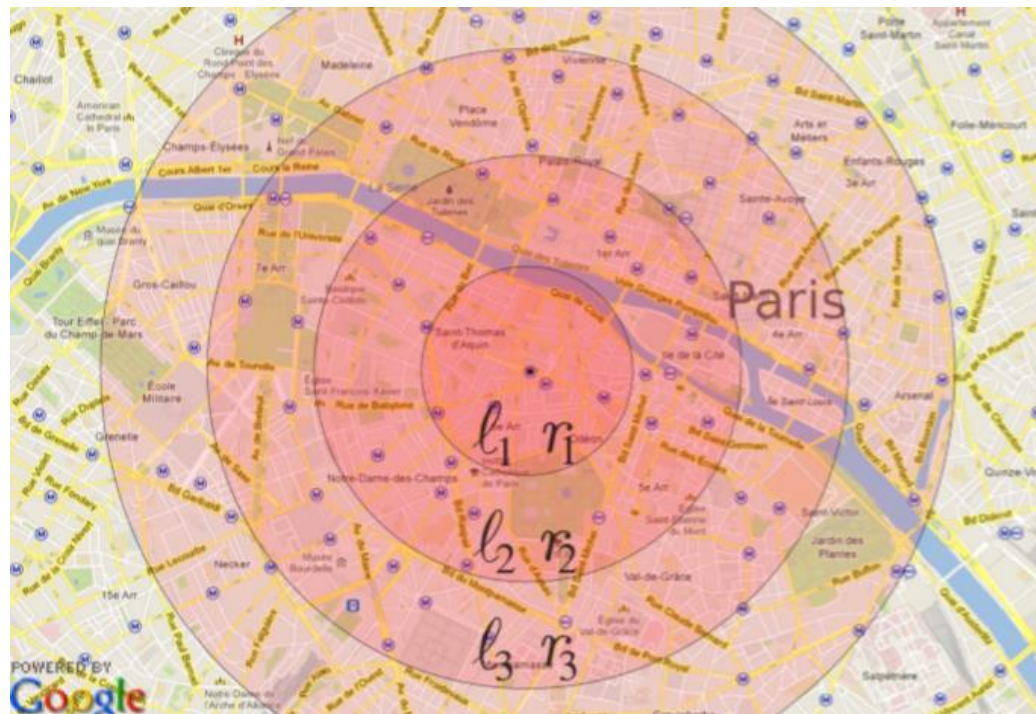
- 差分隐私包括两类：中心化(DP)和本地化差分隐私(LDP)
 - DP假设有个可信任的服务器
 - LDP假设只有自己可信
- 差分隐私实现方式：
 - 随机响应  LDP
 - 拉普拉斯机制  DP
 - 指数机制  DP

T	使用dx-privacy对地理位置进行脱敏
I	用户原始位置
P	For { 1. 获取坐标点 2. 拉普拉斯分布抽样, 并添加至坐标点中 3. 获取新的坐标点 }
O	经过脱敏后的地理位置

P	使输出保留更多原始坐标信息
C	存在准确的地理位置数据
D	如何解决地理位置离散及噪声采样问题
L	ACM SIGSAC 2014

- 相邻概念定义
 - 汉明距离保护力度太强，因此使用欧氏距离 $d(\cdot)$
 - 存在两点A与B，若 $d(A,B) < r$, 则A与B两点相邻

$$e^{-\varepsilon d(D,D')} \leq \frac{\Pr(M(D) = c)}{\Pr(M(D') = c)} \leq e^{\varepsilon d(D,D')}$$



- 噪声采样方式
 - 二元拉普拉斯分布
 - 其联合分布概率密度函数的极坐标形式为：

$$f(r, \theta) = \frac{\varepsilon^2}{2\pi} r e^{-\varepsilon r}$$

- 边缘分布为：

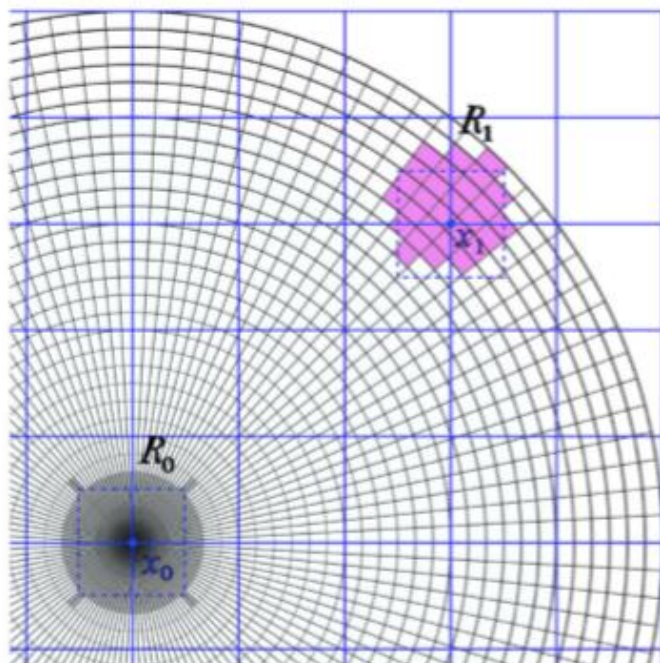
$$f(r) = \varepsilon^2 r e^{-\varepsilon r}$$

$$f(\theta) = \frac{1}{2\pi}$$

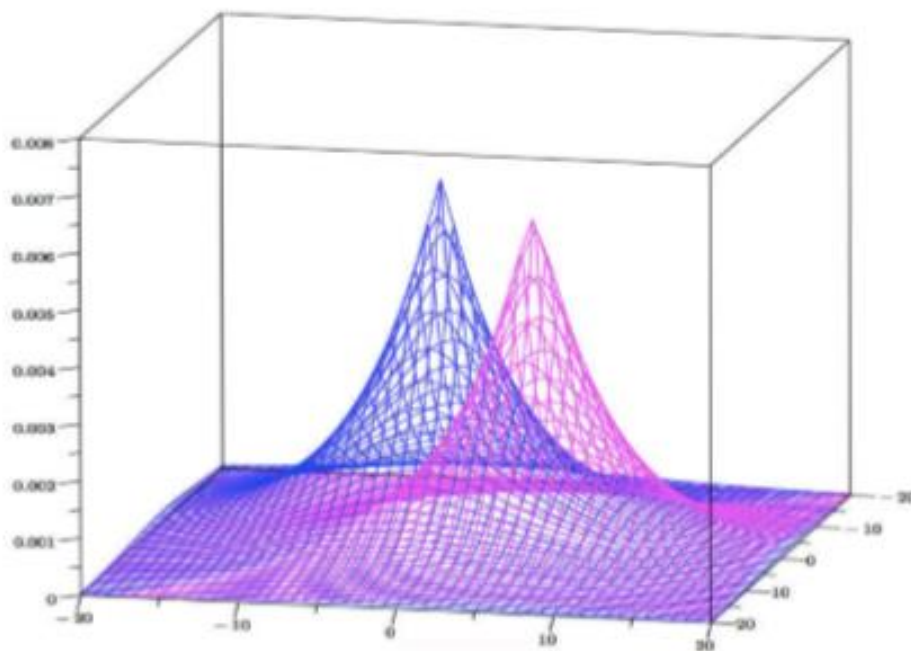
- 采样步骤

- 在 $[0, 2\pi)$ 内随机取一个值
- 假设 $f(r)$ 的分布函数为 $F(r)$,我们在 $[0, 1)$ 内随机取值 x ,则 $r = F^{-1}(x)$

- 离散
 - 对于地理位置 $A(x,y)$,注入噪声之后的地理坐标为 $B(x+r\cos(\Theta),y+r\sin(\Theta))$ 。
 - 选择地图上离 B 最近的位置作为 A 的替代位置。



- 截断
 - 将输出围绕真实点的周围。
 - 预设一个集合C,将B点映射到C中最近的点。



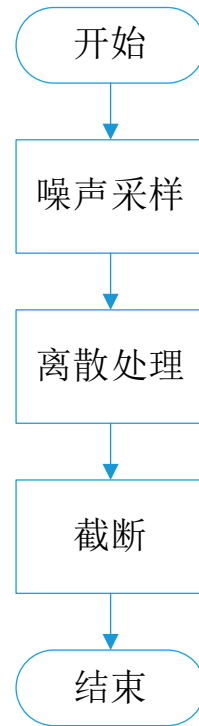
• 算法基本步骤

Input: x //point to sanitize ϵ
..... \mathcal{E} //privacy parameter ϵ
..... $u, v, \delta_\theta, \delta_r$ //precision parameters ϵ
 A //acceptable locations ϵ

Output: Sanitized version Z of input x ϵ

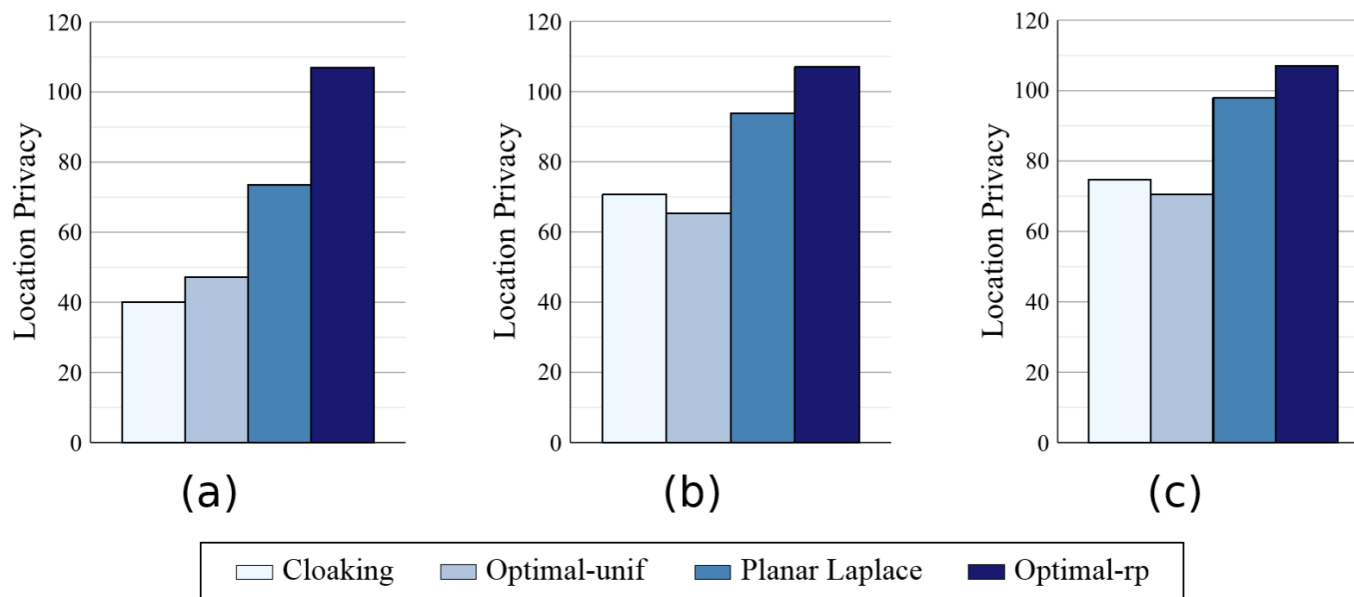
1. $\epsilon' \leftarrow \max \epsilon'$ satisfying Thm 1 for $r_{\max} = \text{diam}(A)$ ϵ
2. **draw** θ unif. In $[0, 2\pi)$ ϵ
3. **draw** p unif. In $[0, 1)$, set $r \leftarrow F_{\epsilon'}^{-1}(p)$ ϵ
4. $z \leftarrow x + \langle r \cos(\theta), r \sin(\theta) \rangle$ ϵ
5. $z \leftarrow \text{closest}(z, A)$ ϵ
6. **Return** z ϵ

算法伪代码



算法流程图

- 实验对比分析
 - 对比方法: obfuscation mechanism、cloaking mechanism
 - 数据: Google Places API
 - 评价方法: Expected Distance Error



- 算法优劣分析
 - 优点
 - 地理位置可用性强
 - 缺点
 - 没有考虑到 r 内地点数量
 - 没有考虑到 r 内关键建筑

- 算法的应用领域
 - 文本、地理等脱敏领域
 - 机器学习领域
- 未来的发展
 - 推荐系统
 - 问答领域

- [1]. Sangeetha S., Sudha Sadasivam G. (2019) Privacy of Big Data: A Review. In: Dehghantanha A., Choo KK. (eds) Handbook of Big Data and IoT Security. Springer, Cham.
- [2]. Rassouli, B. and D. Gunduz, Optimal Utility-Privacy Trade-Off With Total Variation Distance as a Privacy Measure. IEEE Transactions on Information Forensics and Security, 2020. 15: p. 594-603.
- [3]. Andrés, M., et al. Geo-indistinguishability: differential privacy for location-based systems. 2013: ACM.
- [4]. Konstantinos Chatzikokolakis, M.A.N.B., Broadening the Scope of Differential Privacy Using Metrics, in The 13th Privacy Enhancing Technologies Symposium. 2013.



知人者智，自知者明。
胜人者有力，自胜者
强。知足者富。强行
者有志。不失其所者
久。死而不亡者，寿。

谢谢!

