北京理工大學
BEIJING INSTITUTE OF TECHNOLOGY

个人隐私保护与脱敏

小组成员：林勇　管浩良　高思睿　林语欣　桑子玉　周彦哲

汇报时间：2021/11/01

北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY

目录

CONTENTS

个人隐私保护面临挑战—背景

大数据时代—“**数据革命**”

利用数据可提供更好更便捷的服务

2019年的线上广告收入

facebook 700亿美金

阿里巴巴 Alibaba.com >400亿美金

Baidu 百度 >100亿美金

个人隐私数据蕴含
巨大商业市场价值

其中相当比例都是针对个人用户的定制化广告，投其所好，高效推送来获利。

个人隐私保护面临挑战—背景

与此同时，用户数据也是危险的“**潘多拉之盒**”。一旦泄漏，用户的隐私将被侵犯

2017

10月**南非**史上最大规模数据**泄露**，**3000多万**客户信息被公开，总统都未能幸免

11月**五角大楼**AWS S3配置错误，意外暴露**18亿**公民信息

2018

3月**Facebook**被曝超过**5000万条**用户数据泄露。

2019

12月一个**Elastic search**数据库泄露，包括**27亿**个电子邮件地址，其中**10亿**个简单明文存储的密码。

51信用卡被查事件，涉嫌利用**爬虫**，在互联网上帮助催债人**违规获取**欠款人的个人通讯录、地址定位等**敏感信息**。

2020

8月不法分子与**圆通快递**多位“内鬼”勾结，导致**40万条**公民个人信息被泄漏

2021

6月据外媒报道，**大众汽车**约**330万**客户数据遭泄露。不仅暴露车主个人姓名、地址、手机号码，而且还暴露部分驾照号码、贷款号码等私密。

个人隐私保护面临挑战

隐私描述：
隐私是可确认特定个人（或团体）身份或其特征，但个人（或团 体）不愿被暴露的敏感信息

1 法律
- 没有完整的大数据隐私保护法律框架

2 意识
- 民众自我保护意识不够强

3 行业
- 行业管理不规范，个人信息滥采现象十分严重

4 标准
- 侵犯个人隐私的行为难以界定，网络隐私权的范围和界定标准不明

个人隐私保护面临挑战—应对手段

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

2021年9月1日起施行
中华人民共和国
数据安全法

2021年11月1日起施行
中华人民共和国
个人信息保护法

非技术手段

- 建立健全符合我国国情的数据治理体系和完整的大数据隐私保护法律框架
- 提高民众自我保护意识
- 规范行业的信息采集及管理
- 明确网络隐私权的范围和界定标准

&

技术手段

- 匿名化
- 数据脱敏
- 同态加密
- 安全多方计算
- 差分隐私

2 匿名化

匿名化—方法

匿名化实例

姓名	性别	学号	班级	年龄	生源地	课程	成绩
李明	男	1120171853	09111701	28	艾欧尼亚	微积分	59

↓ 去除标识

性别	班级	年龄	生源地	课程	成绩
男	09111701	28	艾欧尼亚	微积分	59

↓ 泛化

性别	班级	年龄	生源地	课程	成绩
男	091117*1	>25	艾***	微积分	5*

举例

班级	生源地	课程	成绩
09111701	艾欧尼亚	微积分	59
09111701	祖安	微积分	85
09111701	祖安	微积分	30
09111701	蒙德	微积分	80
09111702	璃月	微积分	92
09111702	璃月	微积分	100
09111703	艾欧尼亚	微积分	60

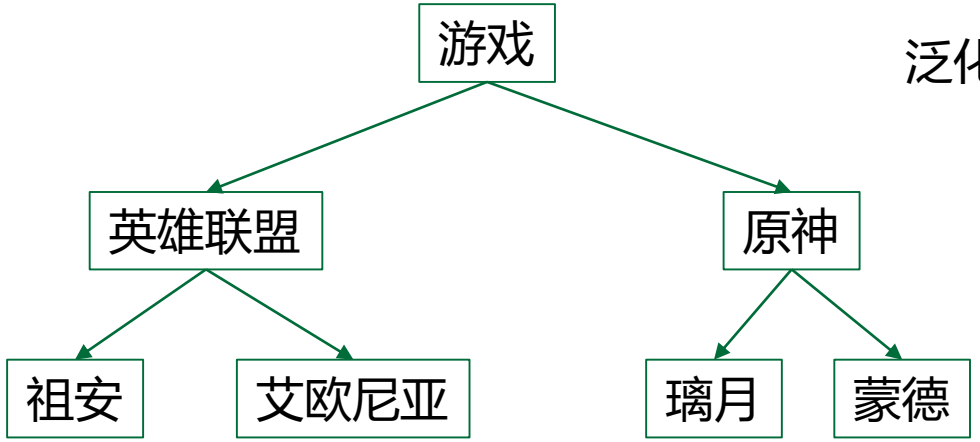
标识符：姓名，学号

准标识符：班级，生源

敏感属性：成绩

等价类：准标识符完全一致的元组

举例



泛化层间

祖安	英雄联盟	游戏	互联网
D4	D3	D2	D1

$$\omega_{j,j-1} = \frac{1}{(j-1)^{\beta}}$$

举例

k-匿名满足每一个等价类中，有至少 k 个 records，对于在等价类中的属性中，不可区分这 k 个 records

班级	生源地	课程	成绩
09111701	艾欧尼亚	微积分	59
09111701	祖安	微积分	85
09111701	祖安	微积分	30
09111701	蒙德	微积分	80
09111702	璃月	微积分	92
09111702	璃月	微积分	100
09111703	艾欧尼亚	微积分	60

DataFly →

班级	生源地	课程	成绩
0911170*	英雄联盟	微积分	59
0911170*	英雄联盟	微积分	85
0911170*	英雄联盟	微积分	30
0911170*	原神	微积分	80
0911170*	原神	微积分	92
0911170*	原神	微积分	100
0911170*	英雄联盟	微积分	60

举例

班级	生源地	课程	成绩
09111701	艾欧尼亚	微积分	59
09111701	祖安	微积分	85
09111701	祖安	微积分	30
09111703	艾欧尼亚	微积分	60
09111701	蒙德	微积分	80
09111702	璃月	微积分	92
09111702	璃月	微积分	100

KACA

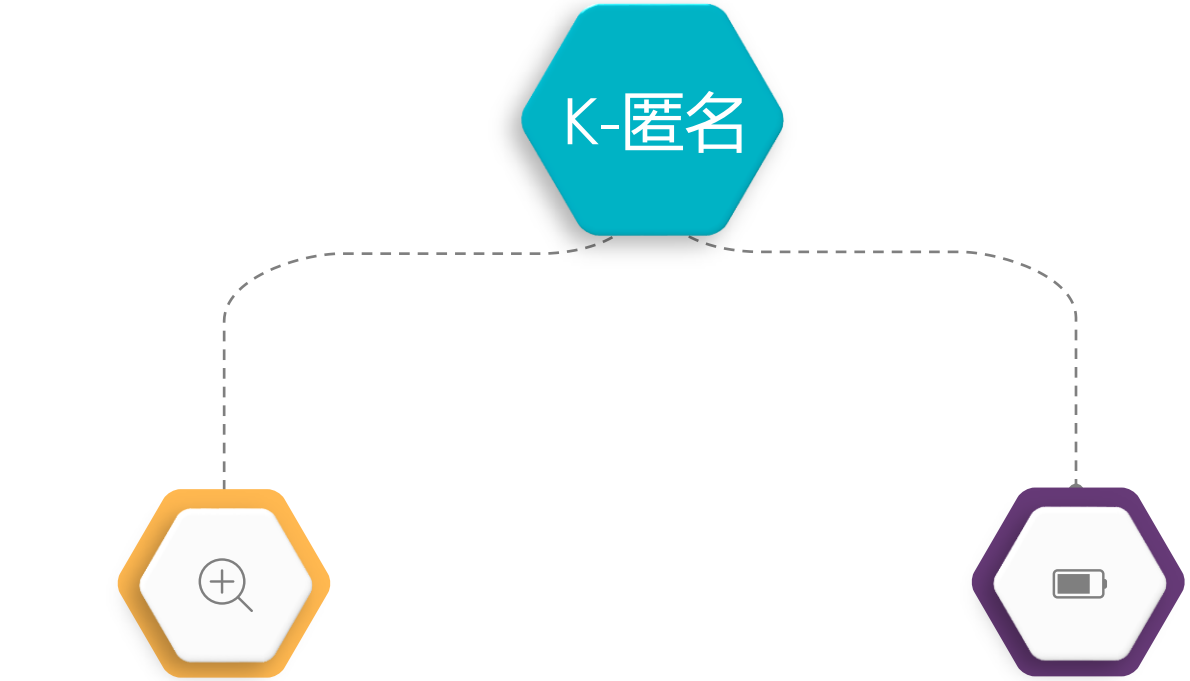
0911170*	艾欧尼亚	微积分	59
0911170*	艾欧尼亚	微积分	60

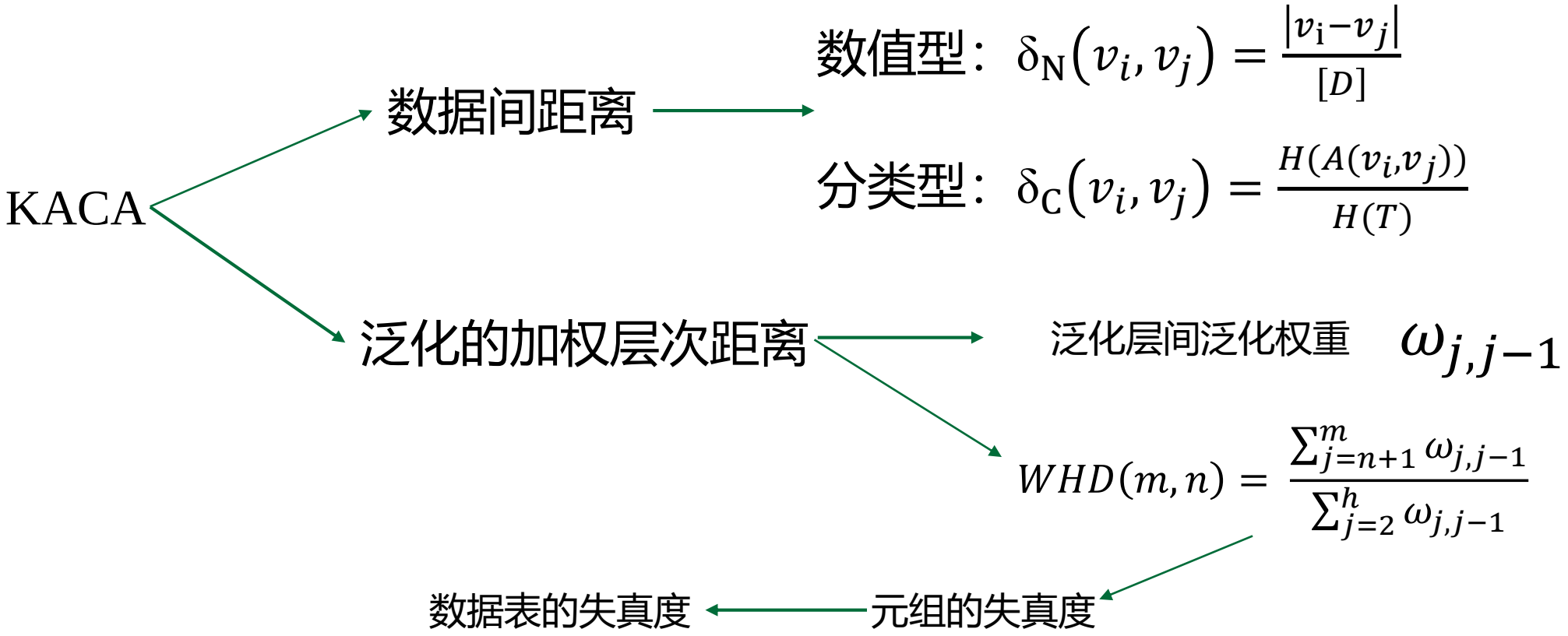
09111701	祖安	微积分	85
09111701	祖安	微积分	30

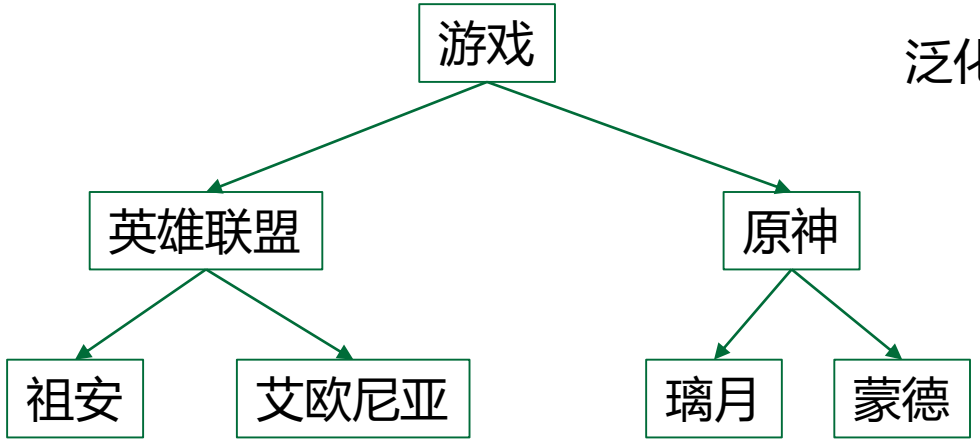
0911170*	原神	微积分	92
0911170*	原神	微积分	100
0911170*	原神	微积分	80

k-匿名满足每一个等价类中，有至少k个记录，对于在等价类中的属性 attributes 中，不可区分这k个记录

K-匿名的实施


K-匿名
泛化(Generalization)
泛化(Generalization)指对数据进行更加概括、抽象的描述，使得无法区分具体数值。
抑制(Suppression)
抑制(Suppression)指不发布某些信息


KACA
数据间距离
数值型：$\delta_N(v_i, v_j) = \frac{|v_i - v_j|}{[D]}$
分类型：$\delta_C(v_i, v_j) = \frac{H(A(v_i, v_j))}{H(T)}$
泛化的加权层次距离
泛化层间泛化权重 $\omega_{j,j-1}$
$WHD(m,n) = \frac{\sum_{j=n+1}^{m} \omega_{j,j-1}}{\sum_{j=2}^{h} \omega_{j,j-1}}$
元组的失真度
数据表的失真度

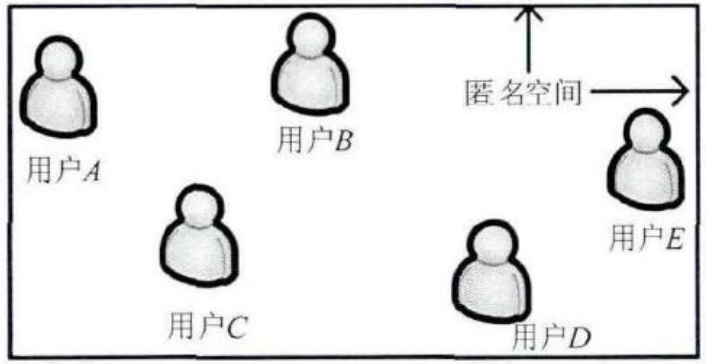


泛化层间

祖安	英雄联盟	游戏	互联网
D4	D3	D2	D1

$$\omega_{j,j-1} = \frac{1}{(j-1)^{\beta}}$$

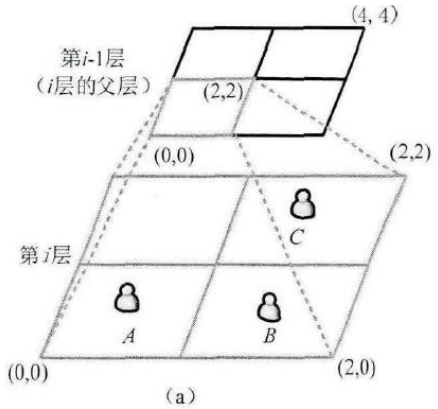
空间k匿名

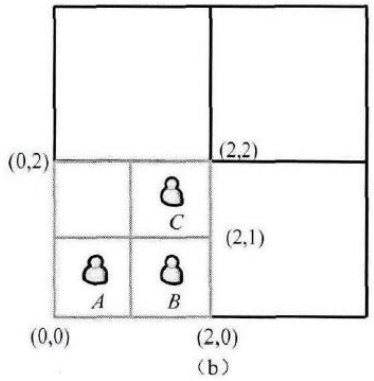


基于可信第三方

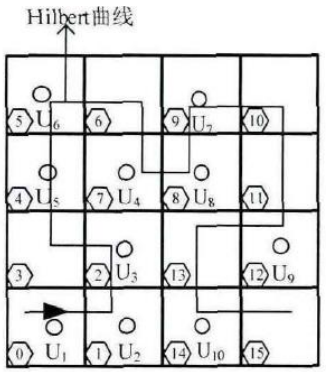
基于移动用户自组网

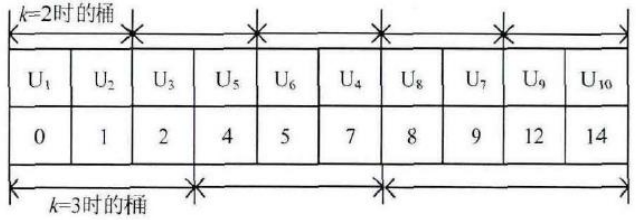
空间四分法





希伯特空间分组法





结合空间k匿名的隐私保护方法

K匿名的攻击手段

0911170*	艾欧尼亚	微积分	59
0911170*	艾欧尼亚	微积分	60

补充数据攻击：当公开的数据有多种类型，**如果他们的K-匿名方法不同，那么攻击者可以通过关联多种数据推测用户信息**

09111701	祖安	微积分	85
09111701	祖安	微积分	30

背景知识攻击：**Background Knowledge Attack**

0911170*	原神	微积分	92
0911170*	原神	微积分	100
0911170*	原神	微积分	80

同质化攻击：**Homogeneity Attack**

$\mathcal{L}$-diversity

如果一个等价类里的敏感属性至少有$\mathcal{L}$个良表示（well-represented）的取值，则称该等价类具有$\mathcal{L}$-diversity。如果一个数据表里的所有等价类都具有$\mathcal{L}$-diversity，则称该表具有$\mathcal{L}$-diversity。

同一等价类中的敏感属性要有至少$\mathcal{L}$ 个可区分的取值

Entropy -diversity

$$Entropy(E)=-\sum_{s\in S} p(E,s)\log p(E,s)\geq \log \mathcal{L}$$

Recursive (c, $\mathcal{L}$)-diversity

设等价类E中敏感属性有m种取值，记r_i为出现次数第i多的取值的频次，如果E满足：

$$r_1 < c(r_{\mathcal{L}}+r_{\mathcal{L}+1}+\cdots r_m)$$

$\mathcal{L}$-diversity的缺陷

difficult and unnecessary to achieve →

总数	10000（个）
阴性	99%
阳性	1%

insufficient to prevent attribute disclosure

- *Skewness Attack*
- *Similarity Attack* →

0911170*	艾欧尼亚	微积分	59
0911170*	艾欧尼亚	微积分	58

t-Closeness

在看到发布的表之前，观察者对个人的敏感属性值有一定的先验信念。在看到被释放的表后，观察者有一个后验信念。信息增益可以表示为后验信念与先验信念之差

将获得的信息分成两部分:关于发布数据中的整个群体的信息和关于特定个体的信息

Q:the distribution of the sensitive attribute value in the whole table

P:the distribution of the sensitive attribute value in equivalence class

t-Closeness


B_0
成绩
59
85
30
80
92
100
60
B_1
0911170* 艾欧尼亚 微积分 59
0911170* 艾欧尼亚 微积分 60
09111701 祖安 微积分 85
09111701 祖安 微积分 30
0911170* 原神 微积分 100
0911170* 原神 微积分 92
0911170* 原神 微积分 80
B_2

t-Closeness

如果一个等价类中敏感属性的分布与该属性在整个表中的分布之间的距离不超过一个阈值，则称该等价类具有**t贴近性**。如果一个表的所有等价类都具有t-close，那么这个表就具有t-close。

问题：测量两个概率分布之间的距离 → Earth Mover’ s distance (EMD)

t-Closeness 的效果如何呢

	ZIP Code	Age	Salary	Disease
1	476**	2*	3K	gastric ulcer
2	476**	2*	4K	gastritis
3	476**	2*	5K	stomach cancer
4	4790*	≥ 40	6K	gastritis
5	4790*	≥ 40	11K	flu
6	4790*	≥ 40	8K	bronchitis
7	476**	3*	7K	bronchitis
8	476**	3*	9K	pneumonia
9	476**	3*	10K	stomach cancer

保持了敏感属性的分布相对一致

	ZIP Code	Age	Salary	Disease
1	4767*	≤ 40	3K	gastric ulcer
3	4767*	≤ 40	5K	stomach cancer
8	4767*	≤ 40	9K	pneumonia
4	4790*	≥ 40	6K	gastritis
5	4790*	≥ 40	11K	flu
6	4790*	≥ 40	8K	bronchitis
2	4760*	≤ 40	4K	gastritis
7	4760*	≤ 40	7K	bronchitis
9	4760*	≤ 40	10K	stomach cancer

匿名化前沿进展

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

2019

在APACHE SPARK 平台上实现大数据k匿名，L多样，T保密。
数据集大小达到2GB时，时间开销大幅增加

2020

基于微聚集理论的k-CMVM和Constrained-CMVM，利用SOM自学习的确定k值。
面对不平衡数据集时性能有所下降

2021

对数据流进行即时匿名，可将z匿名性映射为k匿名进行管理。匿名性弱，实际效用待观察

2021

- 针对数据流匿名化后分类准确性的研究；采用加权的多目标优化公式；CUDSA
- 减小了数据匿名化后的信息损失，即时处理能力未测试

2021

采用神经网络生成人脸对图片人脸进行替换；采用MIKU匿名方法。
生成的人脸质量存在不稳定情况

2020

- 对图（k,l）匿名算法的规模和速度有大幅度提升。
- 当匿名化要求提高时，开销大幅度提高

2021

医疗领域；对客户机-服务器-用户数据路径进行基于聚类的（α,k）匿名

2021

日常生活活动；使用基于聚类的算法产生L多样的匿名；

如果保证了 k-anonymity, l-diversity 和 t-closeness, 隐私就不会泄露了么?

姓名	年龄	邮编
小明	36	102209

姓名	年龄	邮编	工资	购买偏好
*	20~30	1000**	7k	电子产品
*	20~30	1000**	10k	家用电器
*	20~30	1001**	9k	护肤品
*	20~30	1001**	11k	厨具
*	30~40	1022**	13k	电子产品
*	30~40	1022**	8k	家用电器
*	30~40	1022**	4k	图书
*	30~40	1022**	12k	家用电器

很遗憾，依然不能完全保证隐私的保护

北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY

脱敏

CONTENTS

数据脱敏的目的

个人数据（Personal data)

个人数据 是关于一个已识别或者可能识别的自然人（即数据主体）的任何信息。

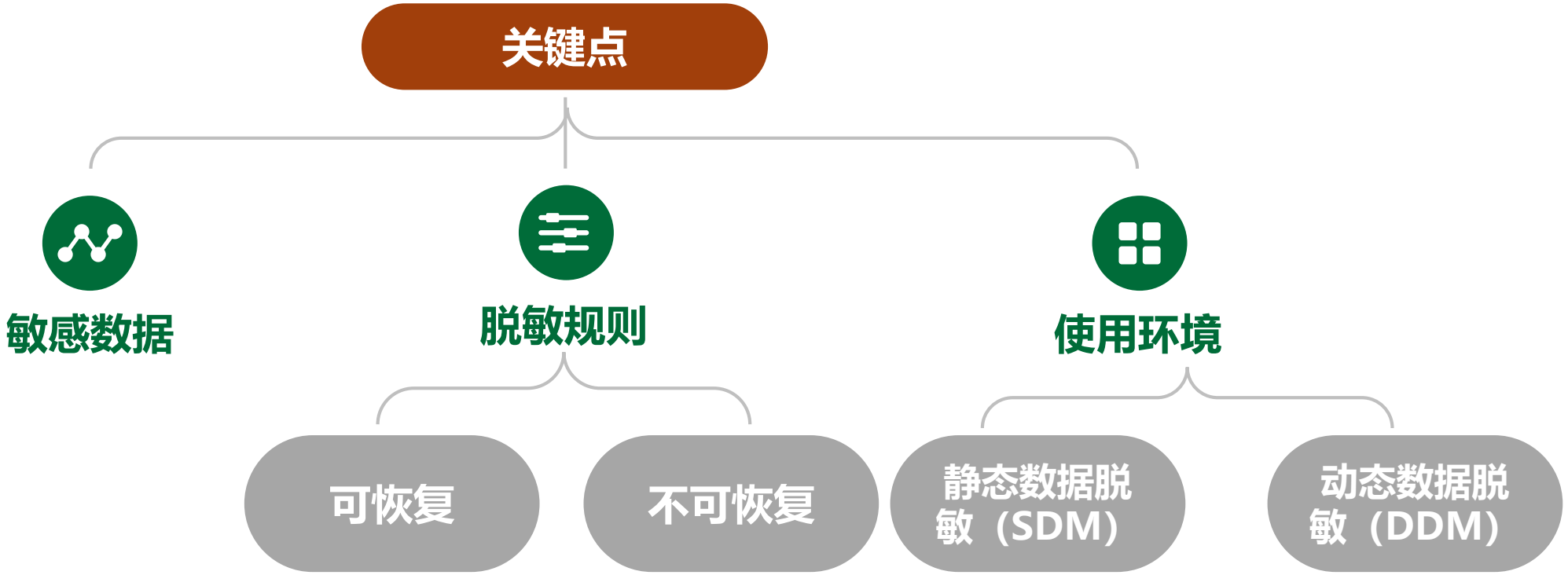
消除个人数据的“已识别”和“可能识别”属性。

数据脱敏

数据脱敏的定义

数据脱敏，也称为数据漂白(英文称为Data Masking或Data Desensitization)、数据去隐私化或数据变形，**本质是通过降低数据敏感度的方式来达到保护信息隐私的目的**。它是一个行业中常用的话术，在法律语境下很少使用。

数据脱敏关键点



使用环境

根据脱敏时段将使用环境划分为两类：

静态数据脱敏

一般用在非生产环境，**先将敏感数据从生产环境脱敏完毕后再投入非生产环境中使用**。一般用于解决测试、开发库需要生产库的数据量与数据间的关联，以排查问题或进行数据分析等，但又不能将敏感数据存储于非生产环境的问题。

动态数据脱敏

一般用在生产环境，**在访问敏感数据时进行脱敏**，一般用来解决在生产环境需要根据不同情况对同一敏感数据读取时需要进行不同级别脱敏的问题。

动态脱敏技术演进

动态脱敏技术		介绍	优点	缺点
第一代动态脱敏技术（结果集解析）	数据库层动态脱敏	结果集返回到应用系统之前即进行脱敏处理	兼容性高、模糊脱敏、易用性好	效率低、精准性差
	应用层动态脱敏	采用HTTP/HTTPS协议代理技术，在终端获取数据之前，可以根据不同权限和角色执行不同的脱敏	兼容性高、分权脱敏	效率低、适用范围窄、易用性差、配置和实施复杂
	API层动态脱敏	采用HTTP/HTTPS协议代理技术，针对API接口数据交互的脱敏	兼容性高、易用简单	适用范围窄、效率低
第二代动态脱敏技术（语句改写）	数据库层动态脱敏	改写包含敏感字段的查询语句，从数据库直接返回不包含敏感数据的结果	效率高、摆脱了性能瓶颈、针对性脱敏	兼容性低、复杂语句难应对、易用性差
第三代动态脱敏技术（混合模式）	混合模式动态脱敏	同时支持结果集解析和语句改写两种技术，可以根据需选择脱敏技术	可用性强、适用范围广、脱敏智能化	

数据脱敏过程

敏感数据发现

01

通过人工发现或者自动发现的方式识别出数据库中的敏感字段信息

脱敏方案制定

02

根据特定的应用场景对敏感字段制定具体的脱敏处理方法

脱敏任务执行

03

执行方案实现脱敏

脱敏数据集评价

04

对脱敏处理后的数据集进行评价，以确保其符合脱敏要求

数据脱敏技术

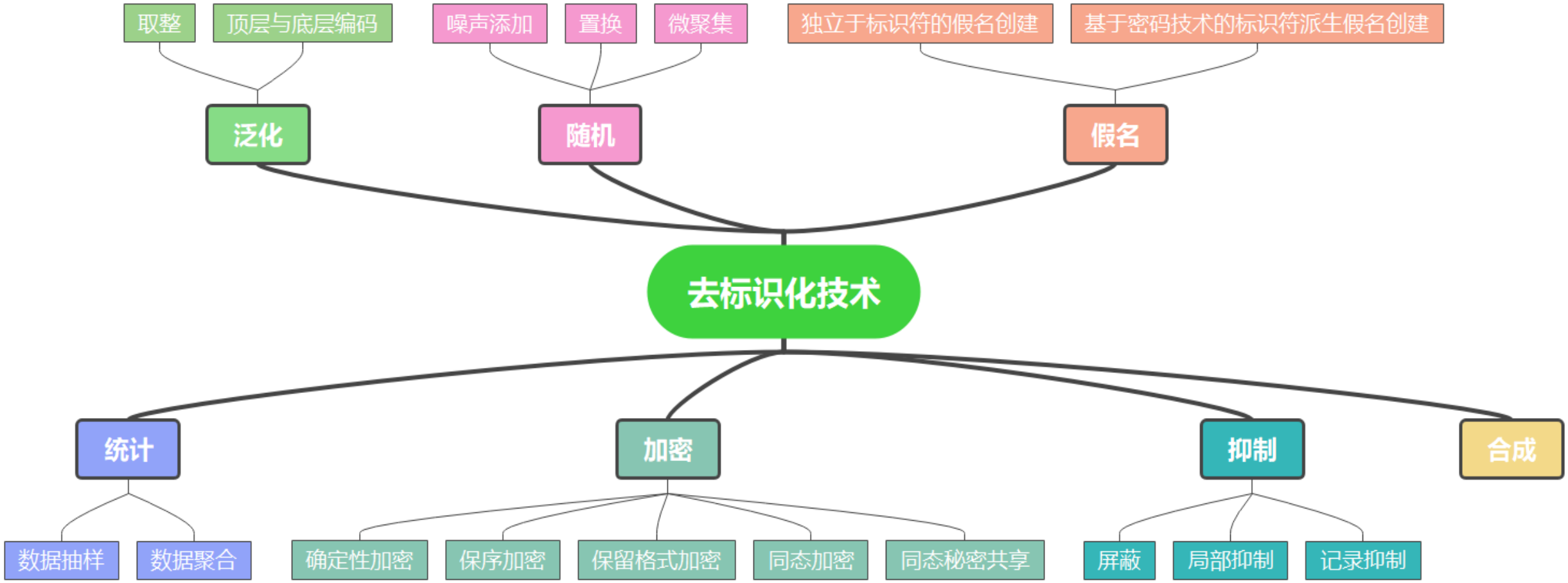
数据脱敏技术包括去标识化和匿名化，并且匿名化的门槛比去标识化门槛更高。如果一个技术属于匿名化技术，那么它一定满足去标识化技术，即去标识化包含匿名化。

去标识化

在不同的地区关于去标识化有着不同的定义和法律效果，这个概念主要出现在美国和加拿大等地的隐私法律中，并且它在中国的《个人信息安全规范》也有着相关的规定。

法规标准	去标识化定义
美国《加州消费者隐私法案》（CCPA）	“去标识”指的是信息不能合理地 (reasonably) 识别，关联，描述，被联系在一起，或者说被链接，直接地或间接地，到特定消费者。
美国《健康保险可转移及责任法案》（HIPAA）	通过处理使得数据不能识别特定个人，或者没有合理的基础能够认为该数据可以被用来识别特定个人。
中国《个人信息安全规范》	通过对个人信息的技术处理，使其在不借助额外信息的情况下，无法识别个人信息主体的过程。
NIST《De-Identification of Personal Information》	表示移除一组标识数据和数据主体之间的关联的任何过程的一个通用术语。

去标识化技术


取整
顶层与底层编码
噪声添加
置换
微聚集
独立于标识符的假名创建
基于密码技术的标识符派生假名创建
泛化
随机
假名
去标识化技术
统计
加密
抑制
合成
数据抽样
数据聚合
确定性加密
保序加密
保留格式加密
同态加密
同态秘密共享
屏蔽
局部抑制
记录抑制

匿名化

在上世纪90年代人们就提出了匿名化思想，并随着大数据时代的到来让匿名化技术迅速成为热点。匿名化这个概念被各个国家的相关立法机构所接受、采纳，尽管在描述上有些许差异，但核心思想基本相同，现选取一些不同地区的定义如下所示：

《通用数据保护条例》(GDPR)	日本的《个人信息保护法》
匿名信息是指与已经识别或可能识别的自然人不相关的信息，或者以数据主体不可或不再可识别的方式提供的信息。	匿名处理信息是指通过处理个人信息而产生的相关信息，它既不能根据采取删除个人信息包含的描述部分及全部标识符等处理措施来识别到特定个人，也无法还原成个人信息。

匿名化

国内的《个人信息安全规范》中定义：匿名化 (Anonymization)指通过对个人信息的技术处理，使得个人信息主体无法被识别，且处理后的信息不能被复原的过程。个人信息经匿名化处理后所得的信息不属于个人信息。

由上述的定义可以看出匿名化的门槛比去标识化门槛更高。它不仅要对**直接标识符**进行脱敏处理，还需要再对**间接标识符**进行泛化或者随机化。随着泛化或者随机化程度的增高，数据安全性也随之升高，但是数据的可用性也逐渐降低，所以，在对数据处理时不仅需要考虑脱敏程度还需要关注数据的可用性，在**匿名性与可用性之间达到一个平衡**。

去标识化与匿名化对比

VS

相同点

两者均是为保护用户隐私，对数据做相关处理，降低个人数据与数据主体之间的关联程度；均采用统计、加密、假名化、泛化、随机化等技术手段

应用

应用

具体应用	作者	特点
基于云计算的数据加密与脱敏	韩培义	隐私数据加密后上传至云服务，从而使得服务商只能看到密文。CloudDLP敏感数据识别与脱敏系统既保护了用户隐私也能最大限度地保留云服务原有的重要功能
基于代理重加密的数据脱敏	宋文鹏、阚海斌、李鸣	采用概率函数，保证由代理重加密密文与原始密文无法形成数据统计关系
基于深度学习的数据脱敏	郑旭如	生成的脱敏数据不仅可以抵御链接攻击，同时避免了基于匿名化技术脱敏的数据中的同质化攻击
一种基于JSON解析的数据脱敏系统及方法	唐更新、徐强、宋辉、王锋、赵卫国	支持多种复杂结构的JSON数据使用，解析所述JSON路径和所述JSON路径的值，当所述JSON路径和敏感JSON路径匹配则对值进行脱敏处理
High-end equipment data desensitization method based on improved Stackelberg GAN	Nan Xiang, Xiongtao Zhang, Yajie Dou, Xiangqian Xu, Kewei Yang, Yuejin Tan	改进Stackelberg GAN用于高端设备企业数值数据脱敏，利用无监督学习和敏感数据进行训练，最终输出任意数量的假数据，达到数据脱敏效果，并用生成的数据代替原始数据进行第三方数据交换或数据分析

数据脱敏技术发展趋势

符合行业化多样化脱敏场景

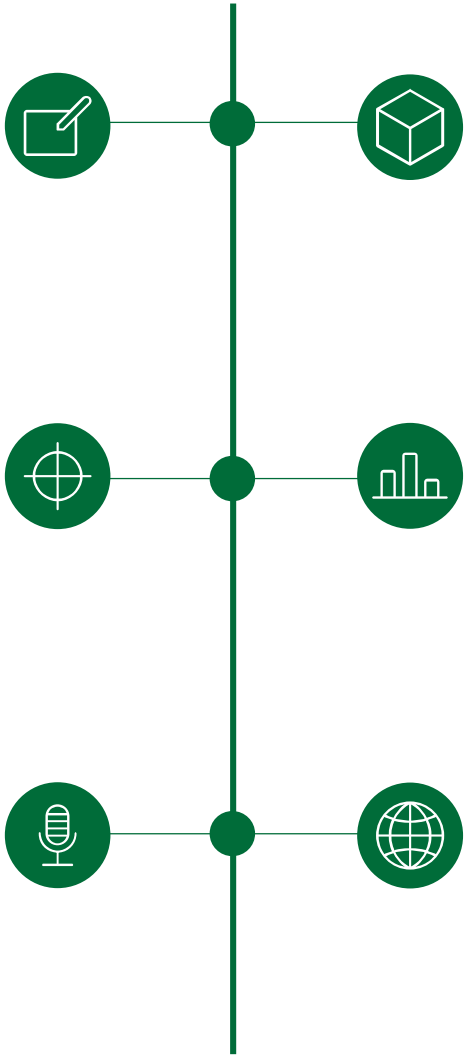
数据脱敏自动化

嵌入式的数据脱敏工具

基于人工智能的数据脱敏

数据脱敏性能升级

脱敏数据可用性提升



4 同态加密

对称加密（DES/AES）

能保证对存储的大数据隐私信息的加解密速度，但其密钥管理过程较为复杂，难以适用于有着大量用户的大数据存储系统。

非对称加密（RSA/Elgamal）

密钥易于管理，但算法计算量太大，不适用于对不断增长的大数据隐私信息进行加解密。

数据加密加重了用户和云平台的计算开销，同时限制了数据的使用和共享，造成了高价值数据的浪费。

同态加密算法可以允许人们对密文进行特定的运算，而其运算结果解密后与对明文进行相应运算所得的结果一致。

将同态加密算法用于大数据隐私存储保护，可以有效避免存储的加密数据在进行分布式处理时的加解密过程。

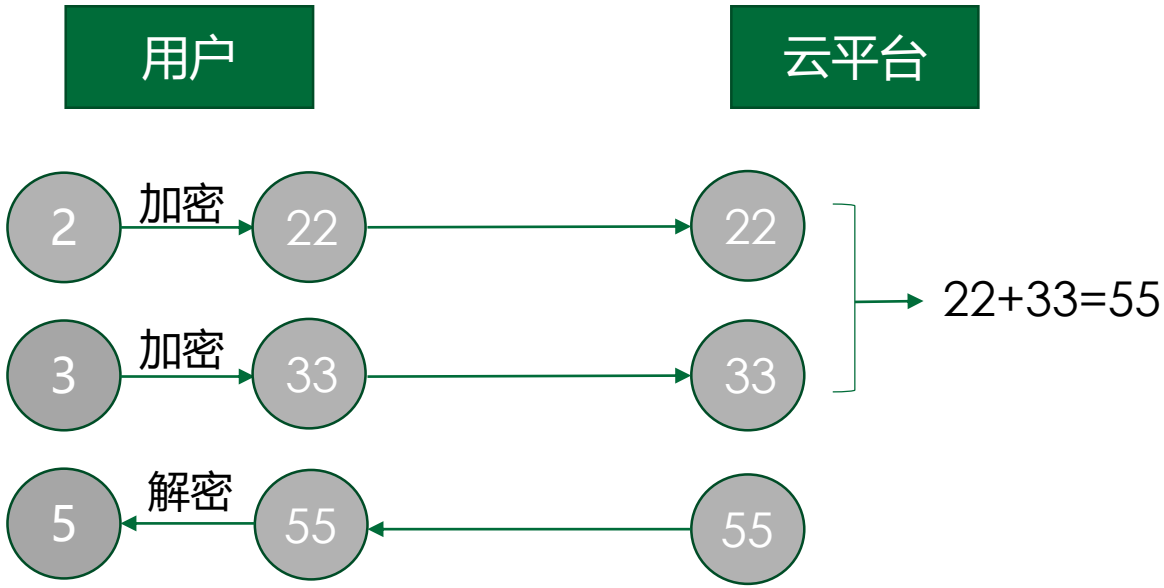
先计算后解密==先解密后计算

同态加密组成部分

- 密钥生成算法——KeyGen(λ)：输入安全系数λ ，输出解密私钥sk、加密公钥pk以及用于密文计算的公开密钥ek；
- 加密算法——Enc(pk,m)：随机化算法，输入公钥pk和明文m，输出密文c，对于相同的明文每次加密得到的密文都是不同的；
- 解密算法——Dec(sk,c)：确定性算法，输入私钥sk以及密文c，输出明文m；
- 密文计算算法——Eval(ek,$(c_1, c_2, \ldots, c_k)$,C)：输入密文计算密钥ek，电路C和密文$(c_1, c_2, \ldots, c_k)$，输出为密文计算结果c^*。

同态加密的简单例子

- 假设加密算法为：f(x)=11x，密钥k=11



2×3=?

部分同态加密

Partial Homomorphic Encryption(PHE)：只支持加法或乘法同态运算。PHE方案稍弱，但开销会变得较小，容易实现，现在已经可以在实际中使用。

类同态加密

SomeWhat Homomorphic Encryption(SWHE)：同时支持加法和乘法的密文同态计算，但其计算次数是有限的.

全同态加密

Fully Homomorphic Encryption(FHC)：支持无限次加法和乘法的运算。FHE方案计算开销极大，目前的主要问题是密钥生成效率问题，暂时还无法在实际中使用。

同态加密研究进展

类型		算法	时间	说明	实际应用
部分同态加密	乘法同态	RSA算法	1977	非随机化加密，具有乘法同态性的原始算法面临选择明文攻击	在非同态场景中应用广泛
		ElGamal算法	1985	随机化加密	DSS数字签名标准基于ElGamal数字签名算法的变体
	加法同态	Paillier算法	1999	应用最为成熟	联邦学习
类同态加密		Boneh-Goh-Nissim方案	2005	仅支持1次乘法同态运算	
全同态加密		Gentry方案	2009	第一代全同态加密，性能较差	
		BGV方案	2012	第二代全同态加密，采用模交换控制噪声	IBM HElib开源库
		BFV方案	2012	第二代全同态加密，与BGV类似，无模交换	微软SEAL开源库
		GSW方案	2013	第三代全同态加密，基于近似特征向量	TFHE开源库
		CKKS方案	2017	可实现浮点数近似计算，适合机器学习建模场景	HElib和SEAL

同态加密国际标准

4.7 同态加密库效率分析

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

SEAL 中 BFV 方案效率测试结果（μs）

Poly	Coeff	Plain	加密	解密	同态加	同态乘	重线性化
4096	109	786433	93464	32306	314	390192	63711
8192	218	786433	267727	112898	1074	1510281	319876
16384	438	786433	884862	439232	4341	6146131	1846517

SEAL 中 CKKS 方案效率测试结果（μs）

Poly	Coeff	加密	解密	同态加	同态乘	重线性化
4096	109	87557	3359	309	12476	63459
8192	218	274215	12748	1071	47599	314501
16384	438	964821	51465	4317	200248	1850888

HElib 效率测试结果（μs）

m	密钥生成	加密	解密	同态加	同态乘
4051	317037	6786	4039	97	26701
4369	571082	8041	5173	98	31477
4859	664138	10497	7554	193	41354

- HElib目前主要的问题在于生成密钥的效率
- SEAL中的加解密、同态乘以及重线性化效率都不高

同态加密在隐私保护中的应用

云存储、云计算

加密原始信息，在密文上进行计算

人脸识别、指纹识别

安全多方计算协议设计的工具

不信任其它参与者和任何第三方的前提下，通过同态加密保证隐私

机器学习

训练数据隐私、模型隐私

预测结果隐私

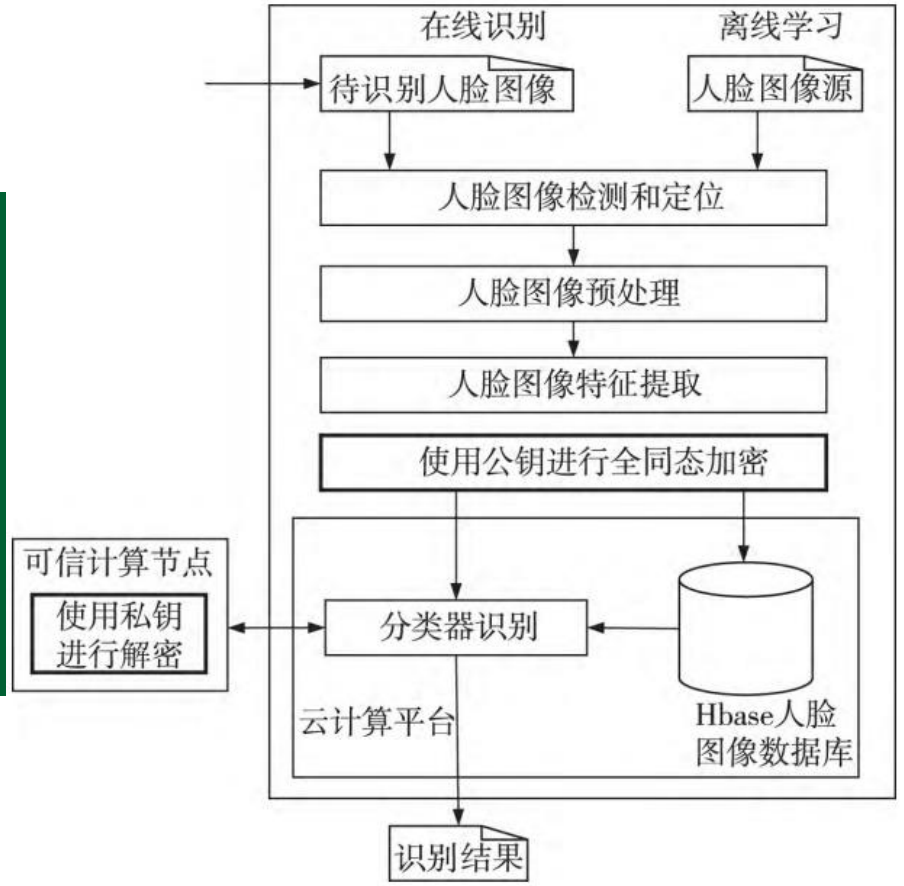
区块链

加密交易信息

验证交易合法性

同态加密在人脸识别中的应用

在全同态加密下计算人脸特征向量间的欧式距离和余弦相似度

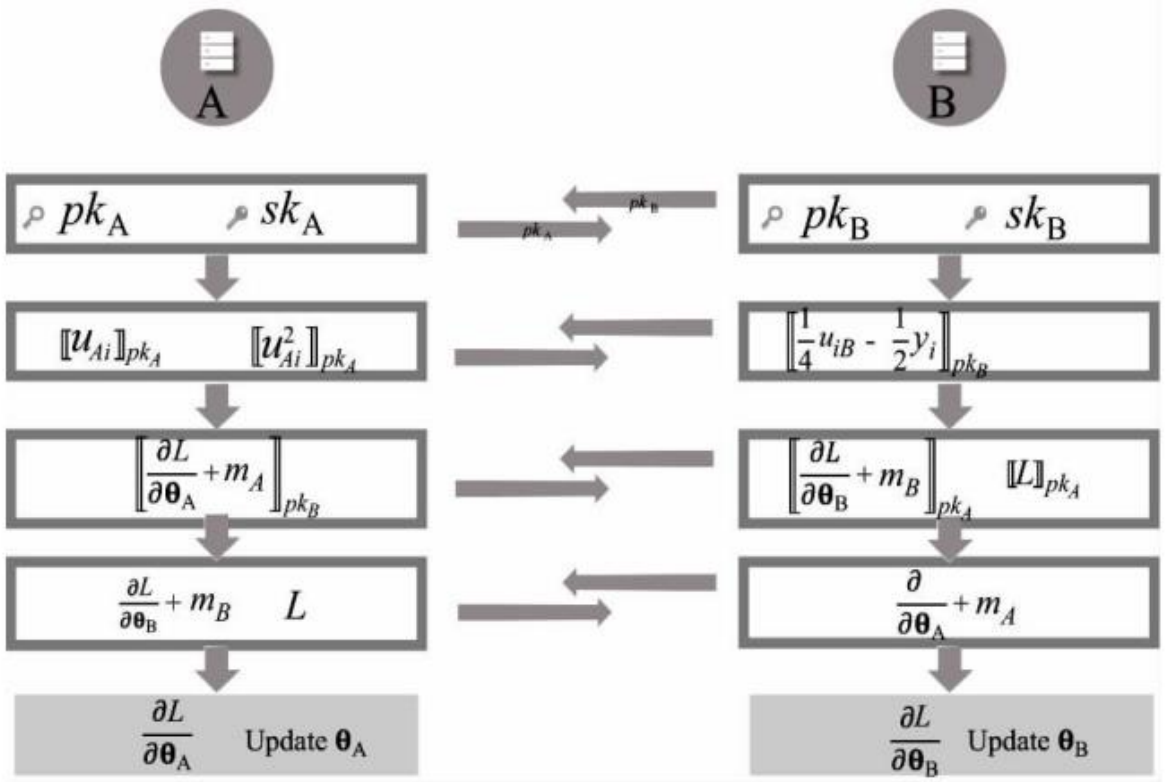


[1]徐慧华,杨雄,张晓惠.云计算环境下基于全同态加密的人脸信息保护[J].贵州大学学报(自然科学版),2021,38(03):83-91.

同态加密在联邦学习中的应用

北京理工大學 BEIJING INSTITUTE OF TECHNOLOGY

- 联邦学习：两个或两个以上参与方共同参与，在保证各数据方的原始数据不出库的前提下，协作构建并使用机器学习模型的人工智能技术。

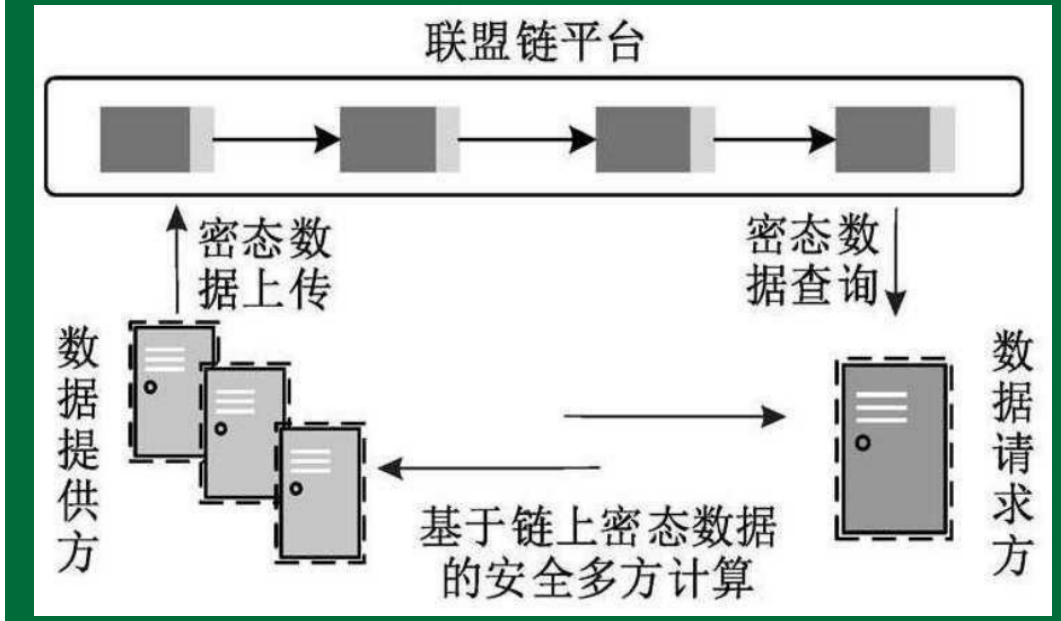


1. 使用同态密钥生成算法为A、B生成公私钥对
2. A加密自己的模型参数将其传输给B，B进行相同操作
3. B将A传送的密文参数与自己的加密参数结合，进行密文运算，得到A应更新的梯度的密文，发送给A，A解密后得到相应梯度，B进行相同操作之后同样得到相应梯度，对模型进行更新

实现了在互不知晓对方数据的情况下更新模型

[2]夏家骏,鲁颖,张子扬,张钰婷,张佳辰.基于秘密共享与同态加密的纵向联邦学习方案研究[J].信息通信技术与政策,2021,47(06):19-26.

同态加密与区块链、安全多方计算结合



1. 数据提供方同态加密数据，上传到区块链中；
2. 区块链以去中心化的方式存储，保证加密数据在机构间公开同时又不可篡改；
3. 采用安全多方计算的方法训练支持向量机模型，在数据共享的过程中保证了任何征信机构都无法从密态数据在计算期间产生的中间结果和最后结果中推测出原始的征信数据。

[3]沈蒙,张杰,祝烈煌,徐恪,张开翔,李辉忠,唐湘云.面向征信数据安全共享的SVM训练机制[J].计算机学报,2021,44(04):696-708.

5 SMC：安全多方计算

背景

医院需要共享医疗信息，但是又不想泄露单个患者的隐私；

政府机构需要统计选举数据，但是又不想公开投票选民的选举记录；

制造厂商想要以行业标准检验产品水准，但又不想让竞争对手知道他们真实的生产数据

目的

1

两方或者多方参与基于他们各自隐私或秘密数据输入的计算

2

参与一方都不愿意让其他任何第三方知道自己的输入信息。

定义&历史

安全多方计算（SMC亦简称MPC或SMPC），解决一组互不信任的参与方之间保护隐私的协同计算问题，不泄露各输入值给参与计算的其他成员。

1982

Assets:X

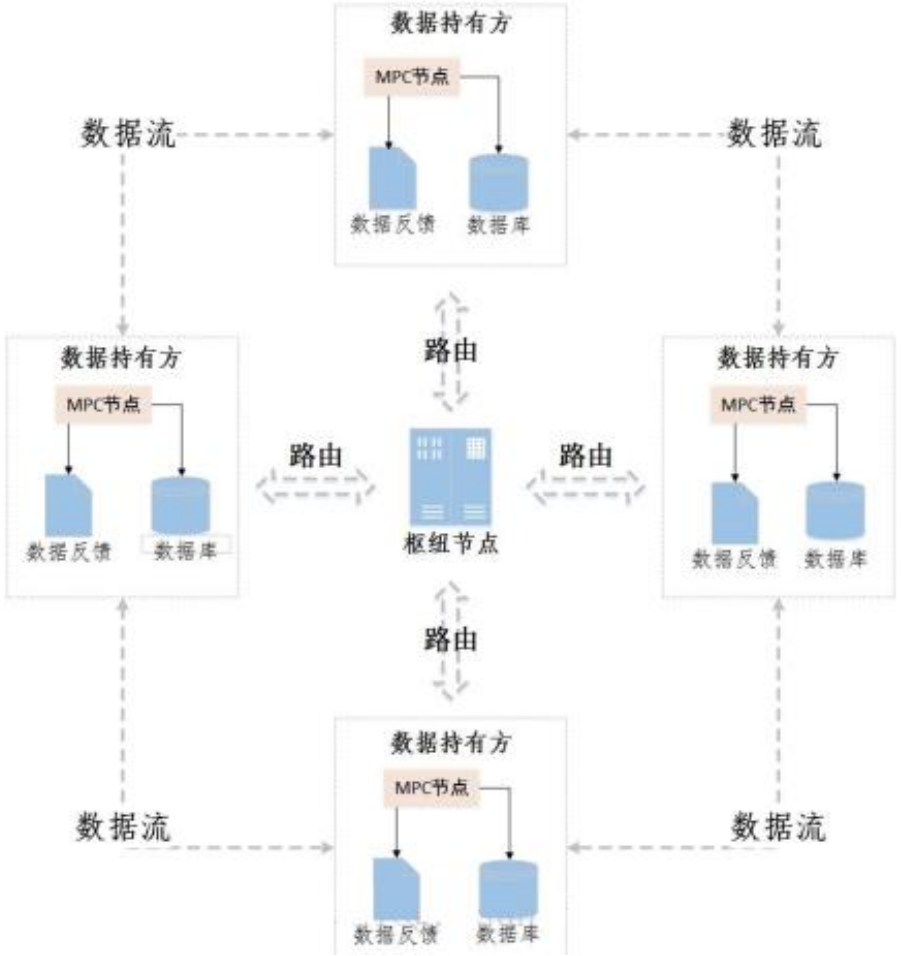
X>? Y

Assets:Y

特 征

- ◆ 去中心化
- ◆ 输入的独立性
- ◆ 计算的正确性

技术框架和模型分类



半诚实敌手模型

恶意敌手模型

MPC组成

安全多方计算技术：

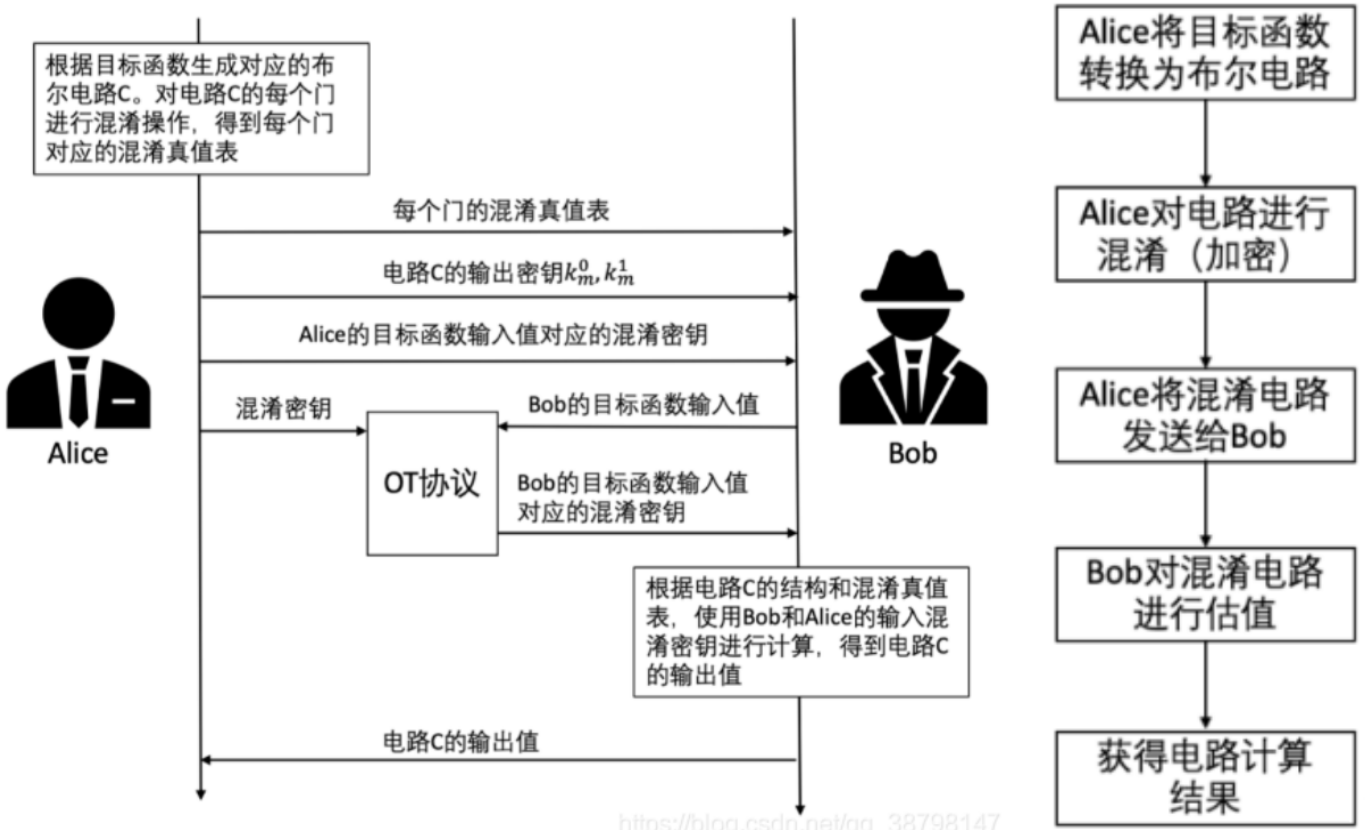
一个单一的技术　No

是由一系列技术组成的协议栈　Yes

混淆电路

两方能在互相不知晓对方数据的情况下计算某一能被逻辑电路表示的函数。

混肴电路=逻辑电路+混肴逻辑



不经意传输

在这个协议中，消息发送者从一些待发送的消息中发送一条给接收者，对接收者收到的具体是哪条信息不知道。

谷歌的开源SMC项目

PIS

Private Intersection-Sum，在不泄露用户任何隐私信息下，计算隐私交集和的问题。

PSI

Private Set Intersection，隐私集合交集问题，是一种多方安全计算加密技术，它允许持有集合的两方比较这些集合的加密版本以计算交集。

谷歌的PIS≈PSI + 对交集元素求和（在不泄露交集元素的前提下）

机器学习与安全多方计算

必要性：机器学习隐私泄露问题的必要性来源于:
(1)训练数据集中有敏感信息;
(2)机器学习训练或者预测阶段存在不可信参与方

安全多方计算局限性：
通信开销过大，如何降低通信轮数，
恶意敌手的存在，共谋问题是否能解决。

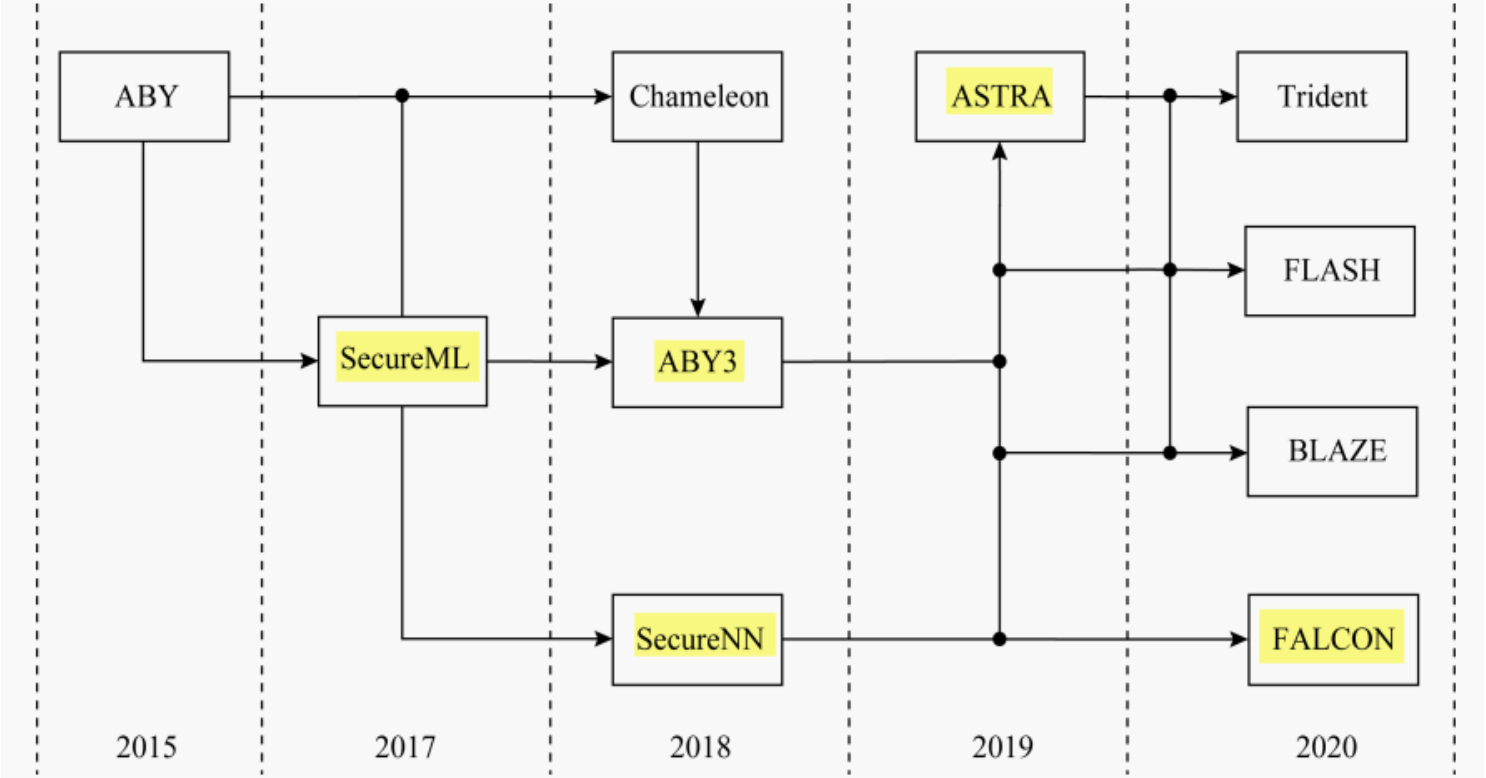


Fig. 6 Development of privacy preserving machine learning schemes involving multi-party

图 6 多方参与的隐私保护机器学习相关方案的发展历程

部分方案比较

方案	参与方数目	模型	特点
SecureML	2	Semi-honest	提出了一种新的近似点乘协议
SecureNN	3\4	honest	实施基于神经网络的PPML方案
ASTRA	3\4	Semi-honest,Malicious	提出了一种有效的点积协议
FALCON	3	Semi-honest,Malicious	支持在隐私机器学习中对批处理规范化层操作的一个安全框架；支持大容量网络的培训

前沿报告：实现以SMC为架构的完全隐私

2019，慕尼黑工业大学; Marcel von Maltitz

《Secure and Privacy-Preserving Services based on Secure Multiparty Computation》

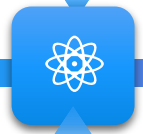
三个挑战

三个技术成果

建立对特定背景的明确的隐私含义

确定能够支持这一意图的技术手段

如何有效地应用这些手段



创建SMC服务机制，降低人工开销，提升效率

为外部客户端开发管理体系结构，实现自适应动态化管理

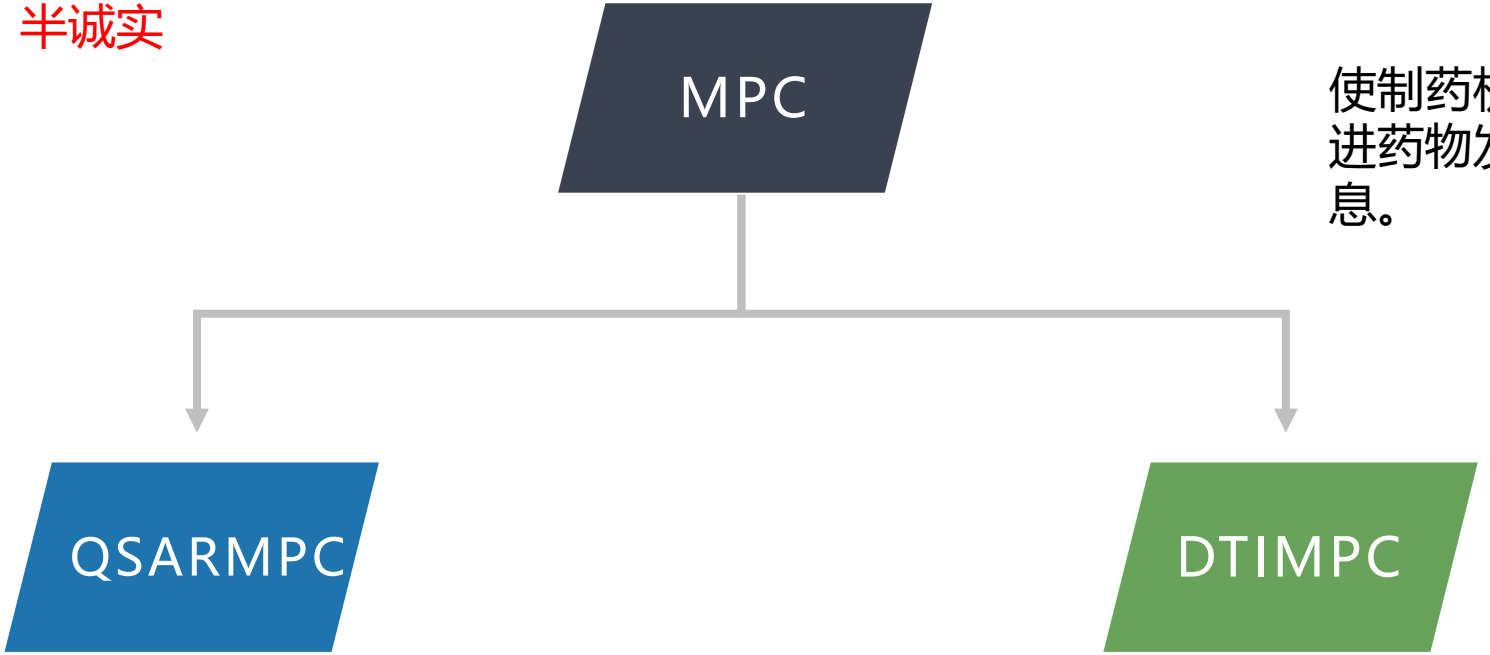
提供数据访问控制的手段来覆盖其余的需求，使拥有数据的组件的访问透明，达到完全隐私

前沿报告：保护隐私的安全多方计算药物发现

2020，清华大学 跨学科信息科学研究所

《Secure multiparty computation for privacy-preserving drug discovery》

环境：半诚实



使制药机构能够实现 高质量的合作，推进药物发现，而不泄露私人药物相关信息。

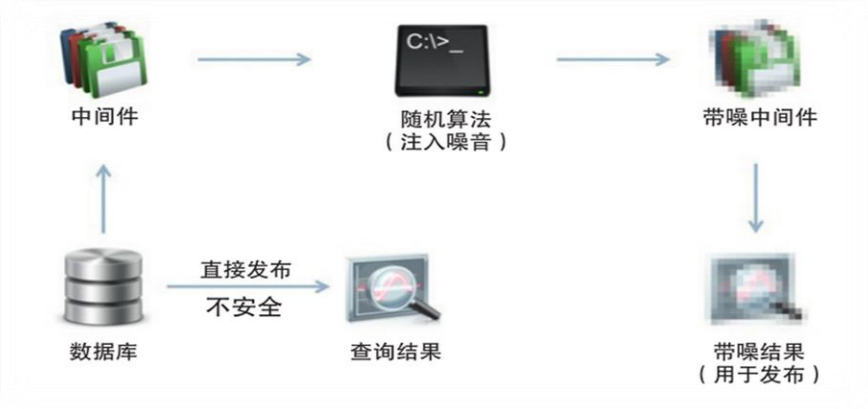
Qsarmpc是MPC下的神经网络模型，具有隐私保护协作的可行性

DTIMPC整合了与药物相关的异构网络数据，准确预测新的dti，同时对药物信息保密

6 差分隐私

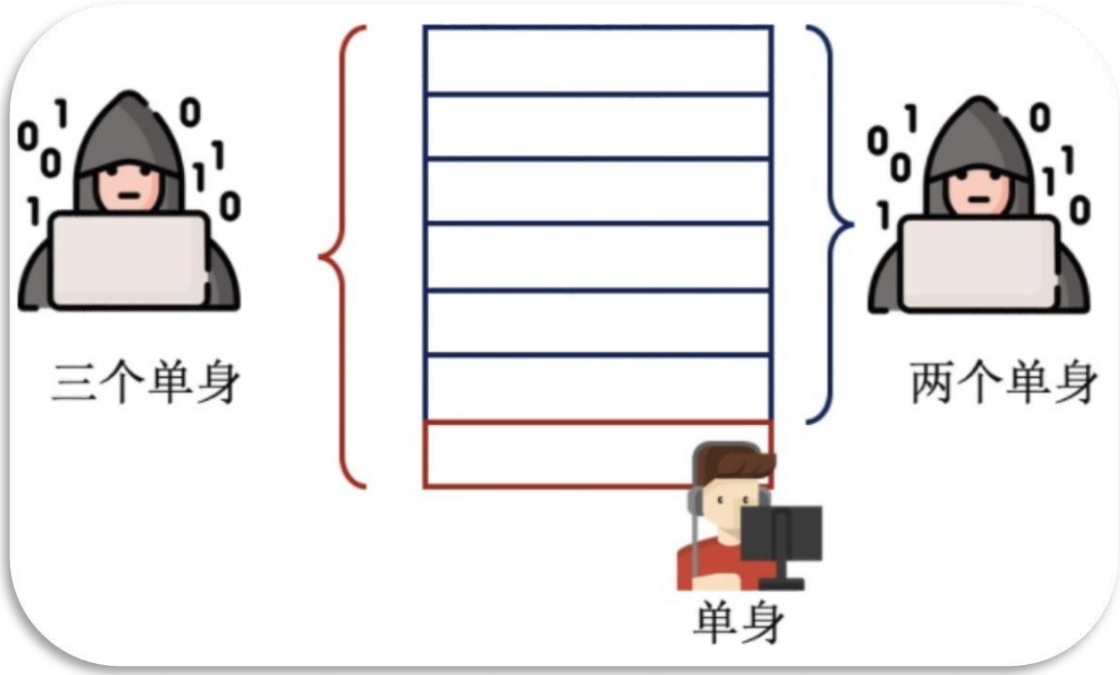
差分隐私—起源

2006年
微软的**C.Dwork**
针对数据库隐私泄露问题
提出**差分隐私**概念



✓ 差分隐私保护就是保证任一个体在或者不在数据集中，对最终发布查询结果几乎没有影响。

举例说明



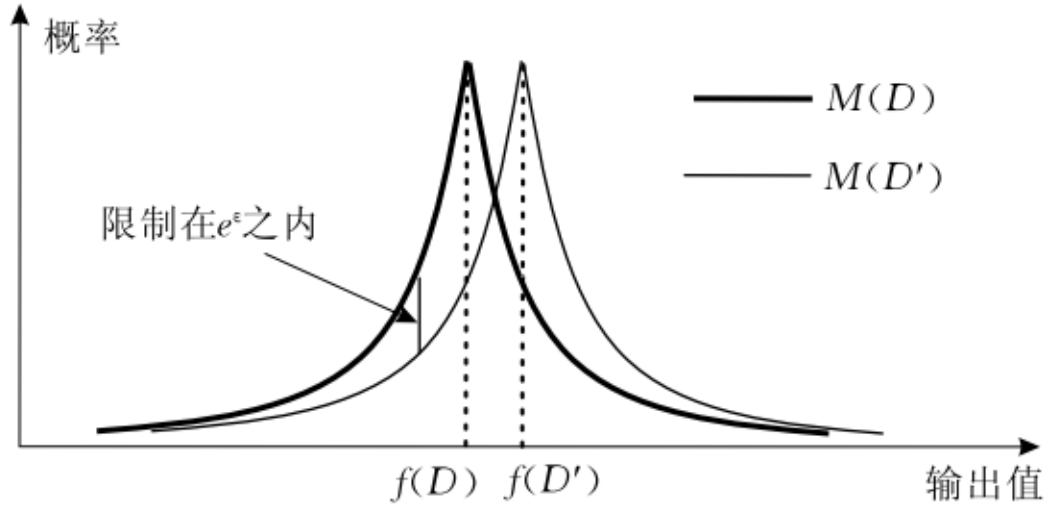
举个简单的例子，假设现在有一个婚恋数据库，2个单身8个已婚，只能查有多少人单身。

差分隐私—数学定义

设有随机算法 M，P_M 为 M 所有可能的输出构成的集合。对于任意两个邻近数据集 D 和 D'以及 P_M 的任何子集 S_M，若算法 M 满足

$$P_r[M(D) \in S_M] \leq \exp(\varepsilon) \times P_r[M(D') \in S_M]$$

则称算法 M 提供ε-差分隐私保护，其中参数ε称为隐私保护预算。



随机算法在邻近数据集上的输出概率

差分隐私—相关概念

隐私保护预算

隐私保护预算ε用来控制算法M在两个邻近数据集上获得相同输出的概率比值，它事实上体现了所能够提供的隐私保护水平。

差分隐私保护算法组合性质

- **序列组合性**

  一个差分隐私保护算法序列构成的组合算法，其提供的隐私保护水平为全部预算的**总和**

- **并行组合性**

  如果一个差分隐私保护算法序列中所有算法处理的数据集彼此不相交，那么该算法序列构成的组合算法提供的隐私保护水平取决于算法序列中的保护水平最差者，即**预算最大者**

敏感度

差分隐私保护

敏感度 —决定大小→ 加噪

- 过多会影响结果可用性
- 过少无法提供足够安全保障

全局敏感度

函数的全局敏感度由**函数本身**决定。较大时，必须在函数输出中添加足够大噪声才能保证隐私安全，导致数据可用性较差。

局部敏感度

局部敏感度一定程度上体现数据集**数据分布特征**。

差分隐私—噪音机制

01

拉普拉斯机制

对于**数值型**数据，一般采用该机制，对得到数值结果**加入拉普拉斯分布**产生的随机噪音即实现差分隐私；

02

高斯机制

对于**数值型**数据，一般采用该机制，对得到数值结果**加入高斯分布**产生的随机噪音即实现差分隐私；

03

指数机制

对于**非数值型**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分数，归一化后作为查询返回概率值。

差分隐私—噪音机制

◆拉普拉斯机制

Y ＋ F(D) ＝ M(D)

Laplace 随机噪声 —— Y

查询函数 —— F(D)

最后的返回结果 —— M(D)

噪声$Y\sim L(0,\frac{\Delta f}{\varepsilon})$满足$(\varepsilon,0)$-差分隐私

隐私预算↓ → 噪声↑ → 结果可用性↓

隐私保护↑

就像下图一样，隐私预算和可用性成正比

◆高斯机制

对**拉普拉斯机制**定义稍加**变形**，我们就能获得拉普拉斯机制的一个**特例**——**高斯机制**。

对于任意的 $\delta \in (0,1), \sigma > \frac{\sqrt{2ln(1.25/\delta)}\Delta f}{\epsilon}$ ，有噪声 $Y \sim N(0,\sigma^2)$ 瞒足 $(\epsilon,\delta)-DP.$

$$P[M(D)\in S] \le e^{\epsilon}P[M(D')\in S]+\delta$$

其中， $M(D)=f(D)+Y$ ，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了；
- 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。

差分隐私—噪音机制及工作模式

◆ 指数机制

指数机制整体的思想是，当收到一查询之后，不是确定性输出一个R_i结果，而是以一定概率值返回结果，从而实现差分隐私。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。

工作模式

Global Privacy

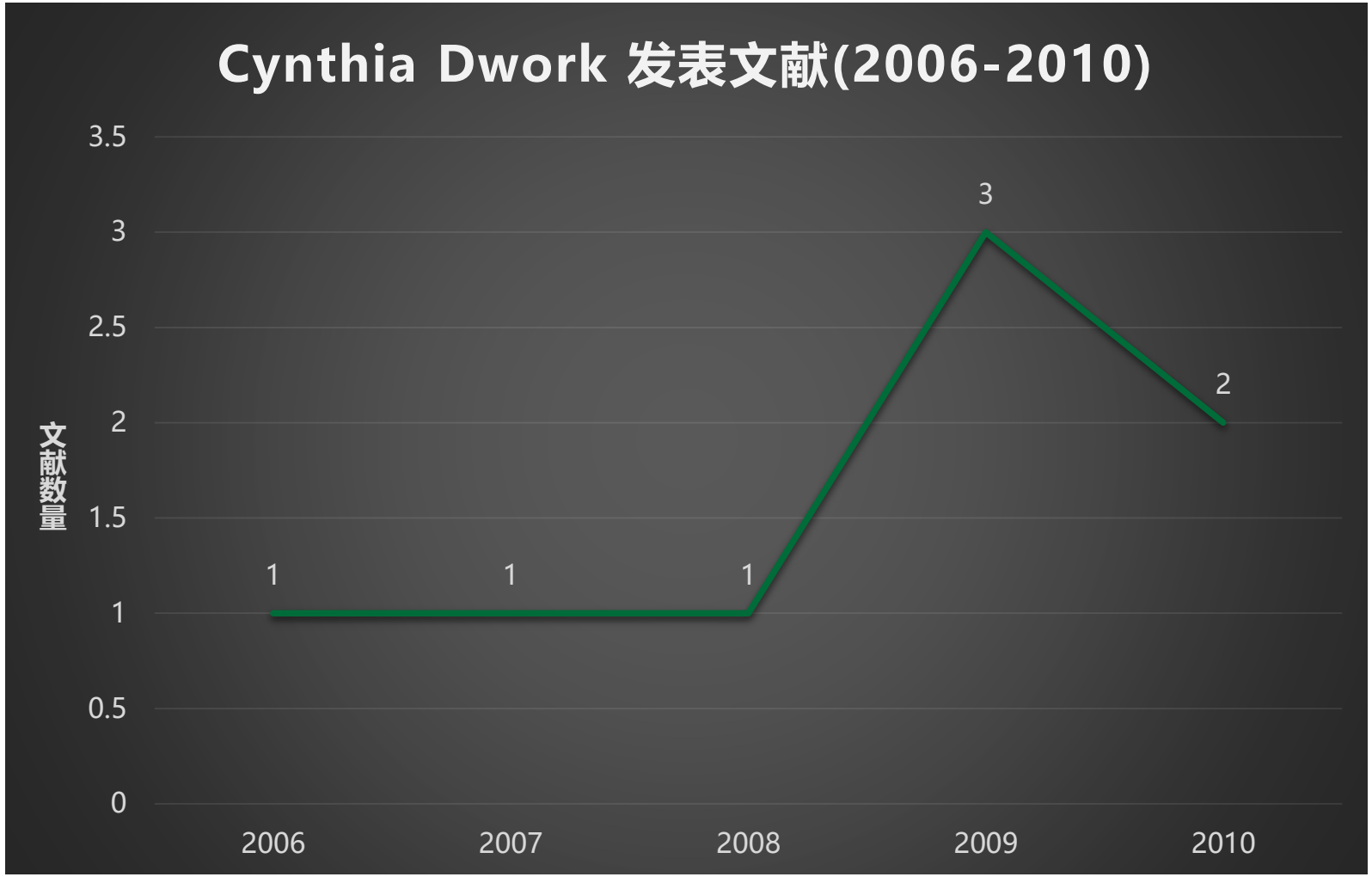
中心化差分隐私就是认为第三方是可信的，数据上传到中心统一加噪声后再对外提供服务

Local Privacy

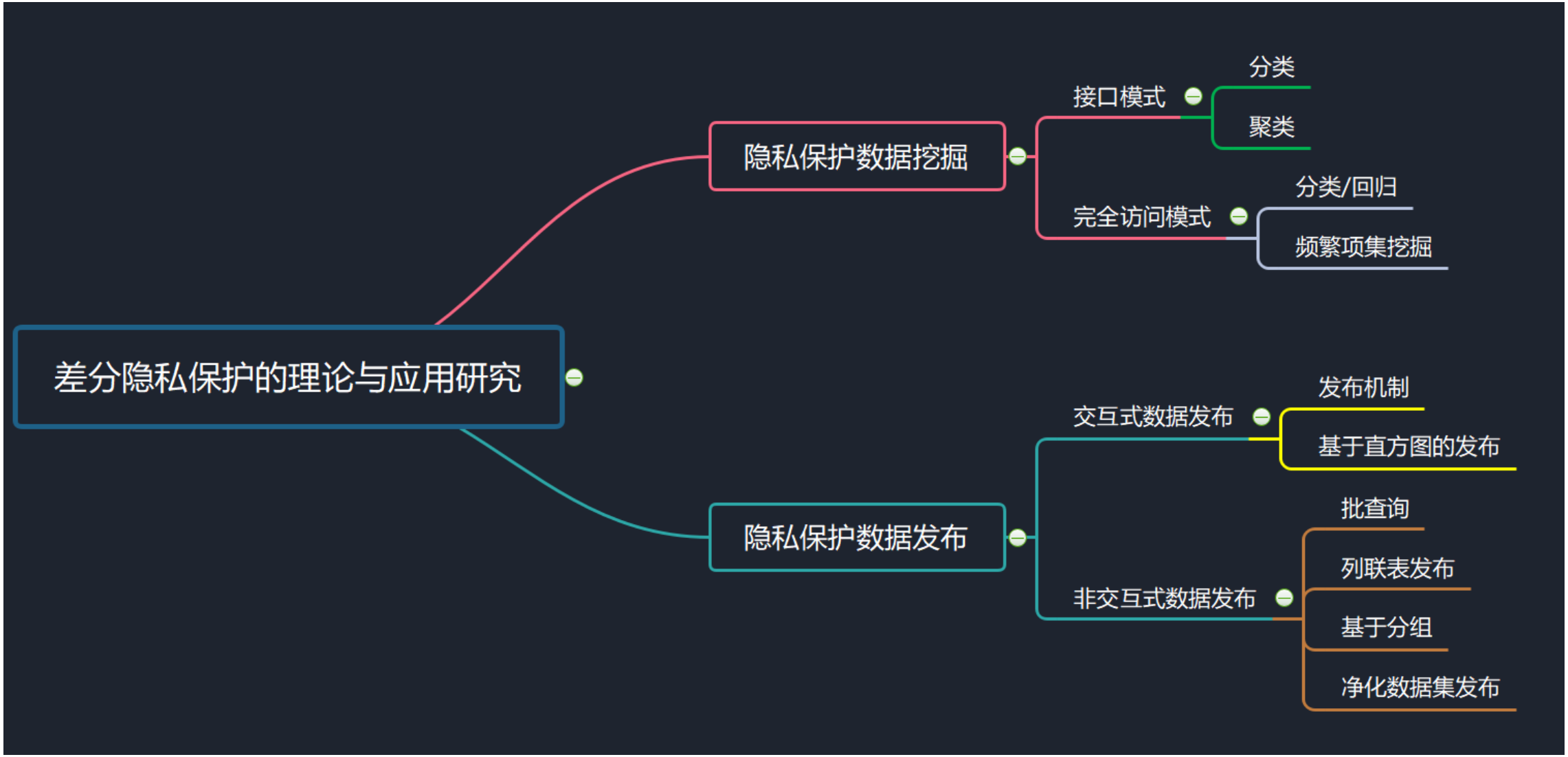
本地化差分隐私就是认为第三方是不可信的，数据在本地经过差分隐私机制后，再上传数据中心

差分隐私—研究现状

2006到2010年间，Cynthia Dwork单独或联合其他人的就发表了8篇文献，基本都是顶会，还有一篇发在自己创办的杂志上。这一系列研究使得差分隐私在理论上不断成熟。



差分隐私—研究现状


差分隐私保护的理论与应用研究
隐私保护数据挖掘
接口模式
分类
聚类
完全访问模式
分类/回归
频繁项集挖掘
隐私保护数据发布
交互式数据发布
发布机制
基于直方图的发布
非交互式数据发布
批查询
列联表发布
基于分组
净化数据集发布

差分隐私—研究现状

隐私保护数据挖掘方法比较

实现模式	挖掘方法	典型算法	优点	缺点
接口模式	分类	DiffPID3，SuLQ-based ID3	容易实现， 分类准确率高	需事先确定迭代次数， 隐私保护预算分配困难
	聚类	SuLQ-based k-means,Corset	容易实现， 方法有效	敏感度高且难以计算； 需较大隐私保护预算才能 保证精度
完全访问模式	分类/回归	DiffGen, Private-RDT	分类准确率高	计算代价高
	频繁项集挖掘	FIM,PrivBasis,TruncatedDB	精度较高， 挖掘速度快	频繁项集长度有限

差分隐私—研究现状

交互式隐私保护数据发布—发布机制

2006	2010	2010	2010	2012
最早用于交互式数据发布的差分隐私保护机制	相比拉普拉斯机制，能在相同预算下提供**更多数量查询**。缺点是算法的时间复杂度会随数据集容量增长呈指数增长	采用投票机制来**减少**隐私保护预算的**消耗**，使该机制能在给定隐私保护预算下，回答**更多查询**。	将差分隐私保护噪声复杂度作为高维凸体的**几何属性**来研究，但有较高的计算复杂度	进一步证明之前中位数和PMW机制都是**此架构的特例**
拉普拉斯机制	**中位数机制**	**PMW**	**K-norm机制**	**迭代数据集生成架构**
C.Dwork	Roth和Roughgarden	Hardt	Hardt	Gupta

差分隐私—研究现状

交互式隐私保护数据发布—基于直方图的发布方法

LP方法
响应**较长范围**的查询，会积累**大量噪声**，导致发布数据不够精确

2006年
C.Dwork

2010年
Hay M等人

Boost方法
基于**后置噪音优化**，提高发布数据准确性

基于k-d树的直方图发布算法
将**空间划分**为若干个分区，以这分区方案对原始直方图进行分区。

2010年
Xiao Xiaokui

2012年
Xiao Yonghui

DPCube方法
首次实现多维直方图的数据发布，主要是响应多维计数查询，查询精度会严重受到维度的影响。

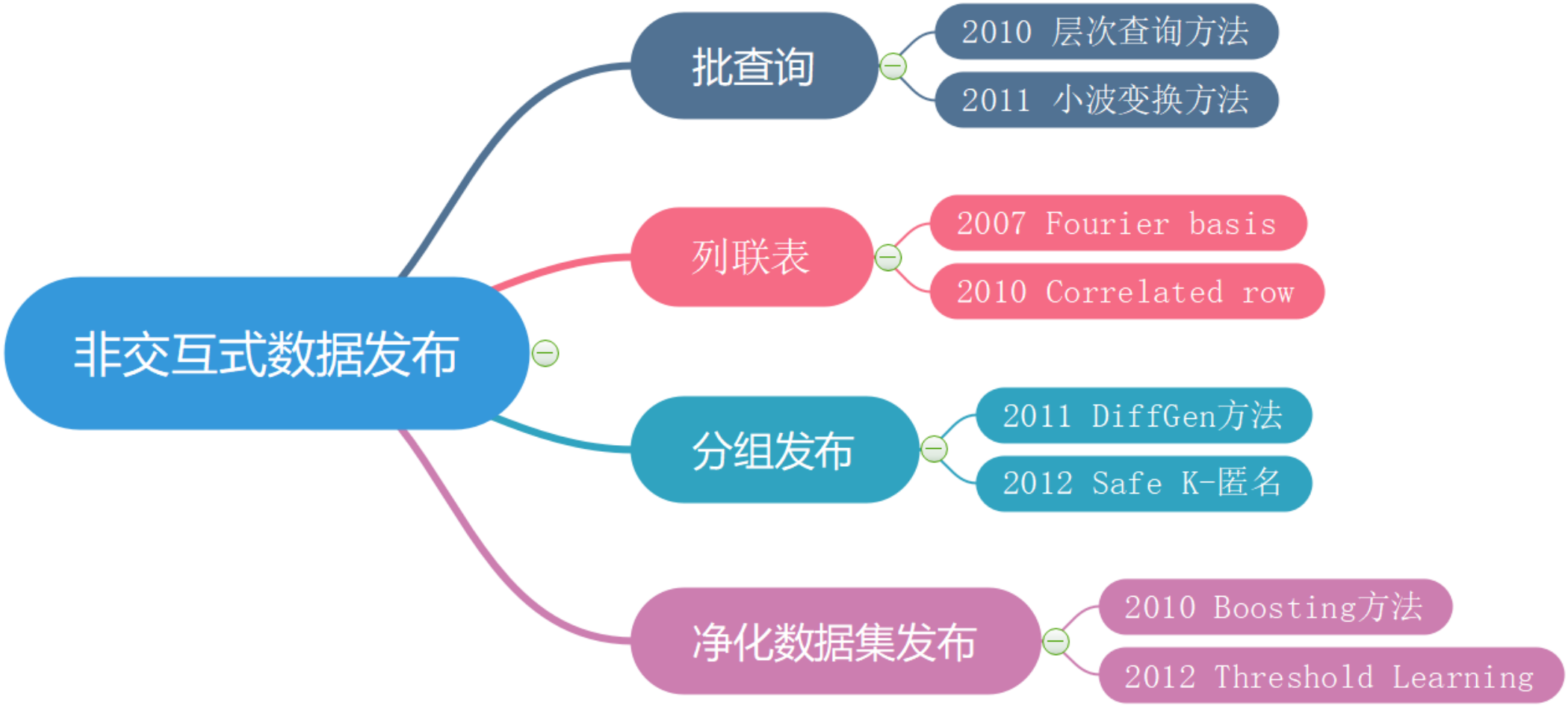
自顶向下聚类方法
在**时间效率**和**数据发布**准确性上有显著提高

2012年
Gergely Acs

2016年
Shao bo

Outlier-histopub方法
解决原始直方图中存在**离群点的问题**

差分隐私—研究现状


非交互式数据发布
批查询
2010 层次查询方法
2011 小波变换方法
列联表
2007 Fourier basis
2010 Correlated row
分组发布
2011 DiffGen方法
2012 Safe K-匿名
净化数据集发布
2010 Boosting方法
2012 Threshold Learning

差分隐私—研究现状

隐私保护数据发布方法比较

方式	方法	方法描述	优点	缺点
交互式	交互式查询	对原始数据集查询产生结果，加噪后发布	**容易实现**，可满足所有查询类型	噪声较大，**查询次数有限**
	直方图查询	由原始数据集产生加噪后的直方图，根据直方图响应查询	**敏感度小**，分析简单，噪声可以控制在小范围内	查询类型收限制，**查询次数有限**
非交互式	批查询发布	数据管理者针对所以后可能查询，一次性对外发布所有查询结果	**容易实现**，可满足所有查询类型	**噪声较大**，但可以采用不同机制降低
	列联表发布	对数据集中的记录按k个属性的排列组合产生k维频数表，加噪发布	可**满足大部分查询**类型	高纬度列联表**噪声大**
	分组发布	对原始数据集进行泛化处理并发布	**结合**泛化和差分隐私方法，**容易实现**	隐私保护**预算分配主观性大**
	净化数据集发布	对原始数据集加入噪声后产生净化数据集并对外发布	**可满足多种查询类型**，查询次数可达到n指数阶	时间复杂度高，实现困难，**噪声大**

差分隐私—前沿进展

2018 tCDP

Mark Bun，Cynthia Dwork等人提出截断的集中差分隐私。它是鲁棒的、有效的，支持强大的算法技术例如隐私放大，并且支持更精确的统计分析。

2019 高斯差分隐私

董金硕等人，受假设检验隐私制定的启发，提出f-DP来刻画隐私，高斯差分隐私(GDP)作为 f-DP类中的单参数定义，是基于两个高斯分布的假设检验。

2020 深度学习GDP

卜至祺等人将高斯差分隐私和深度学习结合，在多种类型的任务上取得了不俗的成绩，可以更精准地呈现隐私损失，从而更好地保护隐私以及提升隐私算法的性能。

2021 差分隐私区块链

差分隐私结合区块链，应用于电子健康记录的安全存储，让区块链中的医学中心既能做研究还能保护电子健康记录安全。

2021 图片混淆DP

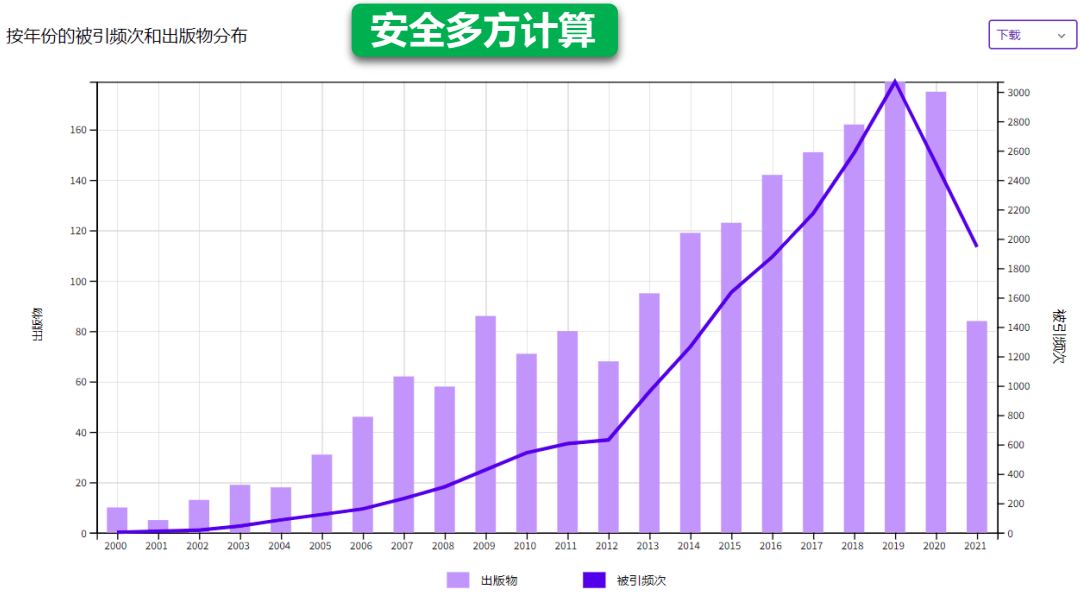
通过差分隐私获得图片的模糊结果，指数机制比拉普拉斯机制在像素间隔的混淆上有更好的视觉质量。

2021 联邦的DP

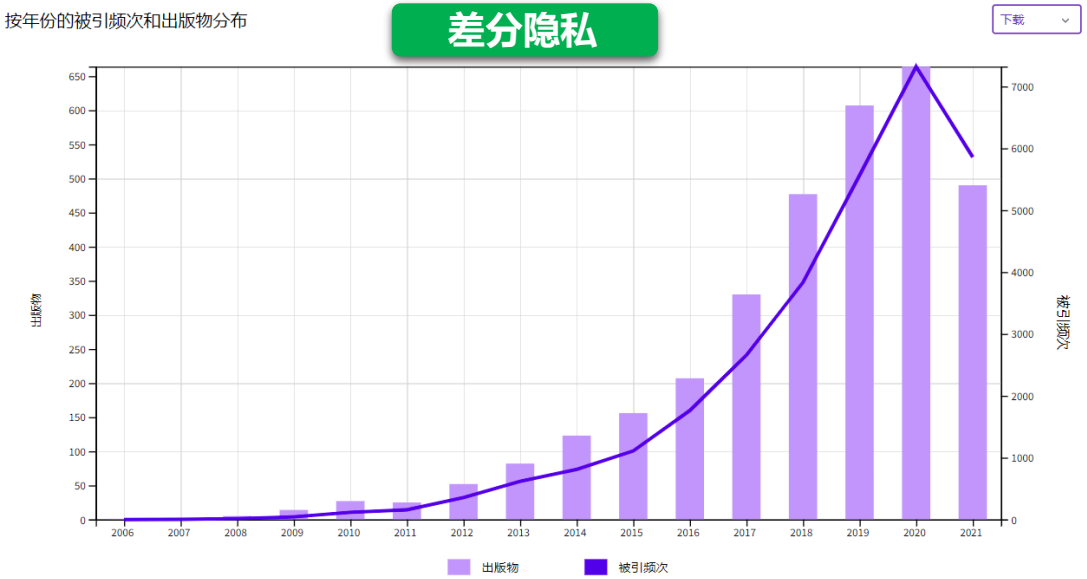
提出基于高斯差分隐私的一些列联邦学习算法，拥有联合的f-差分隐私。

总结—Web of Science 检索结果

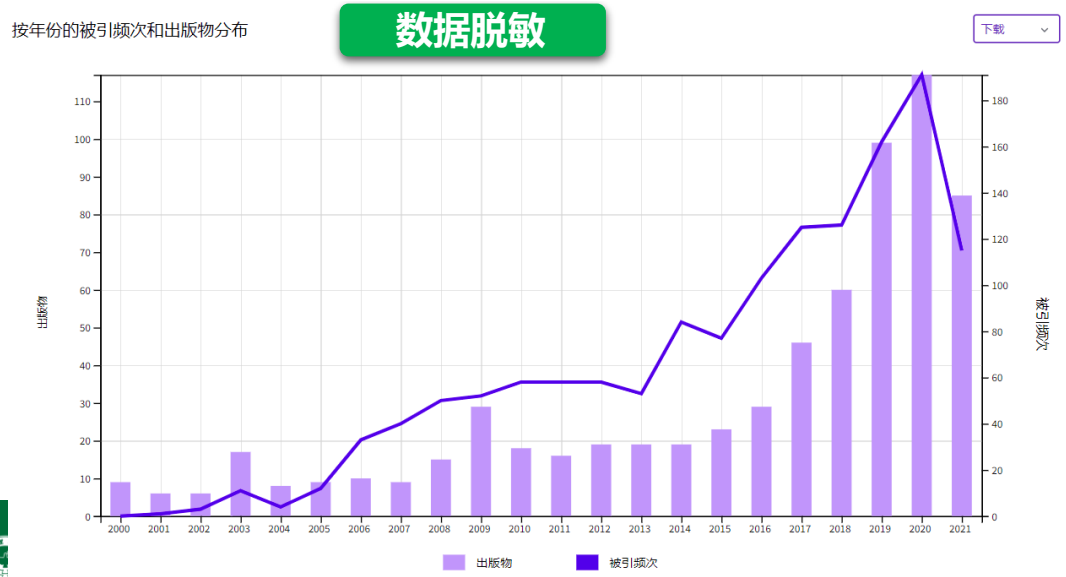
安全多方计算



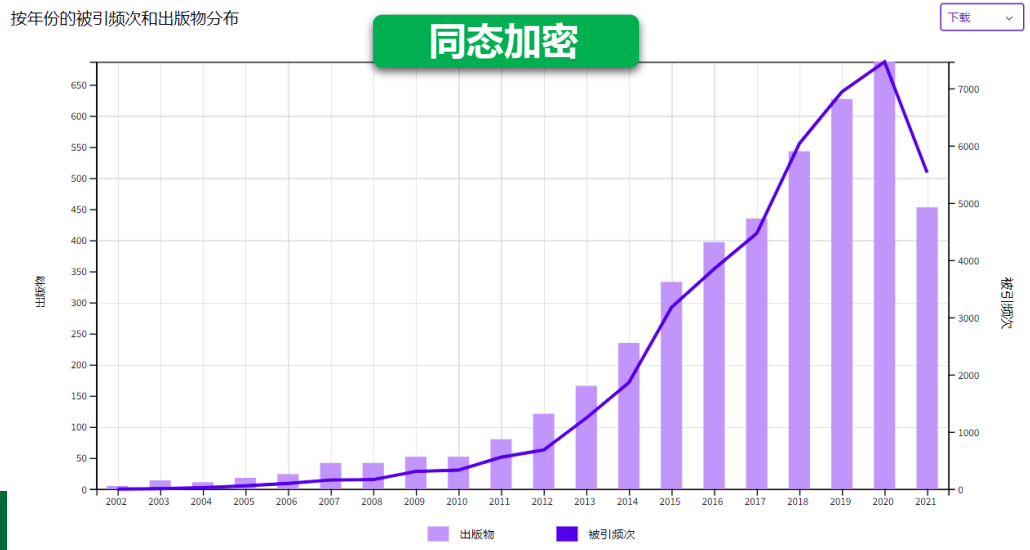
差分隐私



数据脱敏



同态加密



总结

名称	技术特点	计算	优点	缺点	隐私保护场景	模型精度
差分隐私	噪声扰动	低	隐私和效率较高	影响分类器的性能，可用性不高	算力较弱环境	有影响
同态加密	密文计算	高	隐私性高	计算、存储开销大，效率、可用性低	算力强大、隐私要求高	无影响
安全多方计算	不暴露隐私的联合计算	中	隐私性和可用性好	通信开销大，效率低	分布式协同学习环境	无影响

7 技术Demo

7 技术Demo

在对于大批量数据进行数据脱敏的操作中，首先需要定位和识别敏感信息字段

正则匹配

- 对于有明确规则的敏感信息：

如银行卡号、证件号、手机号、民族、省份等个人隐私信息。

关键字匹配

- 对于没有明确规则，可以为任意字符串的如个人姓名等的敏感信息。

NLP算法识别

- 对于没有明确规则，且可以通过自然语言算法进行识别的敏感信息：

如家庭住址、营业证件号及常出现敏感词等

7 技术Demo

基于敏感信息已被定位识别的基础上，可视化对数据进行脱敏处理，技术方案为：

生成虚拟信息表

- 由于个人隐私信息的私密性导致不可能对真实数据进行处理，为展示数据脱敏相关技术，生成虚拟数据进行演示。

数据脱敏处理

- 采用随机和加密方法对数据进行脱敏处理，包括一般简单加密、哈希散列加密、格式保留加密、对称与非对称加密等方法。

数据复原

- 对于可逆加密方法，通过密钥对密文进行解密。

- 技术Demo由python实现，对生成的个人信息数据的不同字段进行数据脱敏

如下表所示，虚拟个人信息由python Faker库生成，总数据条目初定为3000条，数据涵盖个人基本信息在内的多个隐私信息。

	A	B	C	D	E	F	G	H	I	J
1	name	gender	birthday	snn	uid	position	code	phoneNumber	email	creditCard
2	汤丽娟	女	2009/5/29	3.70112E+17	483C 96FA	省太原市沈北新哈尔滨街	902758	15297563349	tangjun@example.com	4.1227E+15
3	黄红霞	男	2016/12/31	6.32701E+17	4302 93D6	徽省洋市兴山惠州街j座	285810	13760305240	uwan@example.org	3.01334E+13
4	姬凤英	女	2004/9/10	451421198309062 66X	4216 8289	林省秀兰县萧山兴城街n	483735	18074152520	myang@example.org	4.06193E+12
5	刘倩	女	2001/2/7	3.505E+17	4709 A40	北省潜江县兴山狄路F座	117484	14535013123	ylai@example.com	4.99277E+15
6	曾玉梅	男	2015/8/31	4.40233E+17	41D4 BE82	天津市斌市兴山向街p座	135525	18530943643	swang@example.org	4.23902E+15
7	蒋俊	女	2013/4/12	6.10924E+17	47BE A3D3	族自治区桂珍县兴山周路	297012	15284861698	otan@example.org	3.59513E+15

7 技术Demo

采用PyQt5实现一个可视化界面，上表为数据表，下表为脱敏后数据表，图中是对全体数据进行Md5方法散列化得到脱敏后结果。

一般简单加密

Hiding&Floor

通过将数据置为常量，并将浮点数或日期向下取整，隐藏不重要字段，用于处理较不敏感字段。

Enumberate

将数据映射为新的值，且保留数据在序列中的大小位置不变。

Prefix Preserve

保持数据前n位不变，对剩余部分进行随机混淆。

- 一般简单加密采取方案如对数据字段进行简单替换或是随机噪声添加，实现效果为：

	Name	Gender	Birthday	Snn	Uid	Ad
1	何*	*	****-**-**	**********	1507BA73 173...	重庆市
2	杨*	*	****-**-**	**********	0FD9256A C2...	福建省
3	刘*	*	****-**-**	**********	1DFE7D70 666...	贵州省
4	刘*	*	****-**-**	**********	D7408EEF 73D...	澳门特
5	张*	*	****-**-**	**********	70F8B25D AE6...	内蒙古
6	欧*	*	****-**-**	**********	36A49DA5 ...	江苏省

对称与非对称加密

Homomorphic Encryption

根据同态性性质，可对加密后的密文信息进行处理而不泄露原始明文信息。

Data Encryption Standard(DES)

对称密钥加密算法，采用密钥对数据进行加密，可采用2至3密钥对数据进行多次加密。

RSA Algorithm

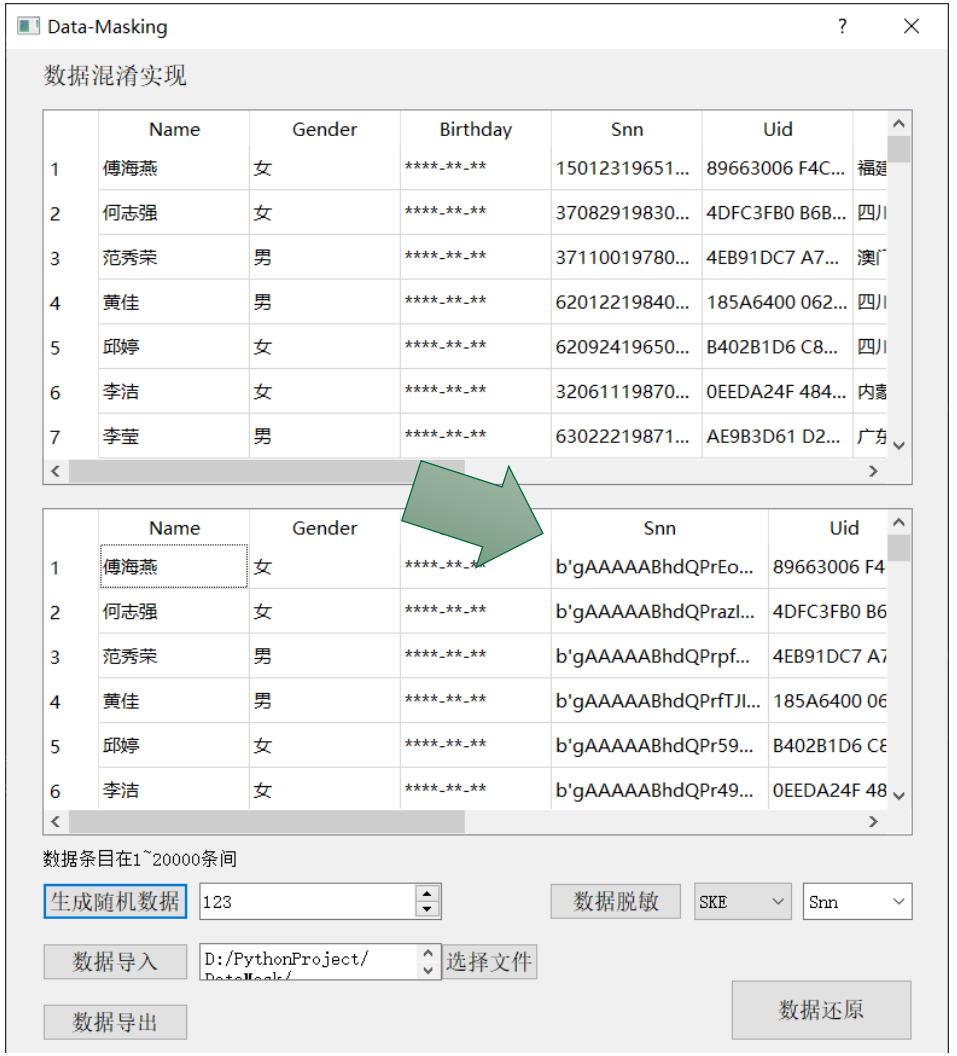
非对称密钥加密算法，将两大素数乘积公开作为加密密钥对数据进行加密，同时设置私钥。

7 技术Demo

如身份信息这一栏，采用非对称加密方法，生成公钥与私钥，并将密文显示在加密数据栏。

```
-----BEGIN PUBLIC KEY-----
MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQCv8sMy+A9yS5w75x3Zx3ozV3Rp
l1zGvGwXqsfoXV2NJBVE2G+MxLvMyvuQiPO9WBA8EPTNXi3cWXmd+sgPPgKm9C0P
gUErmVaYJ9DIlxSJKAi7Fl4S4mvmquoN9ys+WkoQtRGms5Po+CYn8fpHOyOwgvdb
b0S0VZY0OPmA6Twm5QIDAQAB
-----END PUBLIC KEY-----
```

```
-----BEGIN RSA PRIVATE KEY-----
MIICXQIBAAKBgQCv8sMy+A9yS5w75x3Zx3ozV3Rpl1zGvGwXqsfoXV2NJBVE2G+M
xLvMyvuQiPO9WBA8EPTNXi3cWXmd+sgPPgKm9C0PgUErmVaYJ9DIlxSJKAi7Fl4S
4mvmquoN9ys+WkoQtRGms5Po+CYn8fpHOyOwgvdbb0S0VZY0OPmA6Twm5QIDAQAB
AoGAFaHRuP8BmypLo5vHMkLRteU/NUo8SFzbJ0/Cc0GjYlp42PWJRXPcqRGdu0fz
O6jLmi71ghhdqBVCHBe8qA8WMS8XKacKv2mSvM1hrnhuZhycxJyxMq+rX03wCPB+
o7alh3NxeFQVt/bVPG0hBoYE4OLaqz4MPaB0J7YPMrCY2BMCQQC9dzK6bAMopah8
```

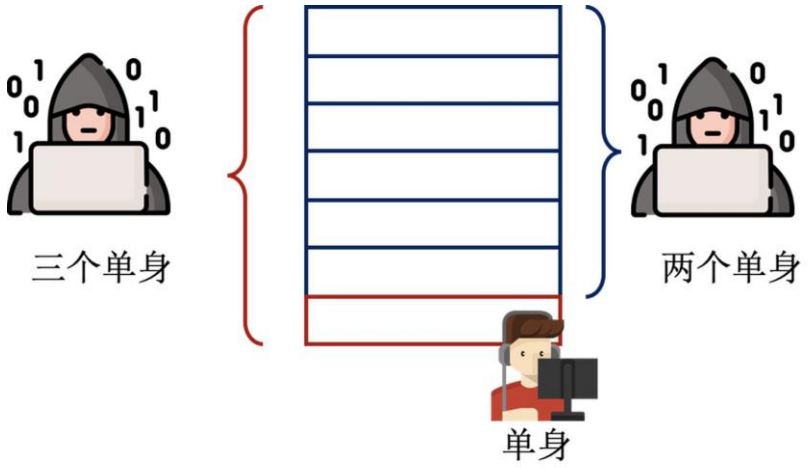


- 对于不可逆推且一一映射的加密算法如哈希散列算法，如表中得到的定长散列值，且可进一步对散列值加“盐”(salt)提高数据安全性。

	Name	Gender	Birthday	Snn	Uid	Ad
1	何晨	36a4908a557...	c904816195d...	77d9ca69a3e...	1507BA73 173...	重庆市开
2	杨龙	36a4908a557...	ab57491bc5c...	97ca1267fb45...	0FD9256A C2...	福建省长
3	刘平	87c835a6b17...	593d98d35f1...	28c4aba35dc...	1DFE7D70 666...	贵州省新
4	刘琳	36a4908a557...	58a22007592...	c1fbad8b3317...	D7408EEF 73D...	澳门特别
5	张欣	87c835a6b17...	b2ea6e6013b...	c8358fb13b0...	70F8B25D AE6...	内蒙古自
6	欧鹏	36a4908a557...	a7d288a7109...	7082135d849...	36A49DA5 ...	江苏省历

7 技术Demo

考虑通过比对不同公开数据表可能精确定位已脱敏数据(如图中年龄表格)的问题，采用Epsilon-差分隐私对可能的查询结果进行处理。



数据混淆实现

	Name	Gender	Birthday	Age
1	杨*	*	****-**-**	49
2	吴**	*	****-**-**	19
3	段**	*	****-**-**	44
4	万**	*	****-**-**	55
5	朱*	*	****-**-**	55
6	贺**	*	****-**-**	34
7	陈**	*	****-**-**	53

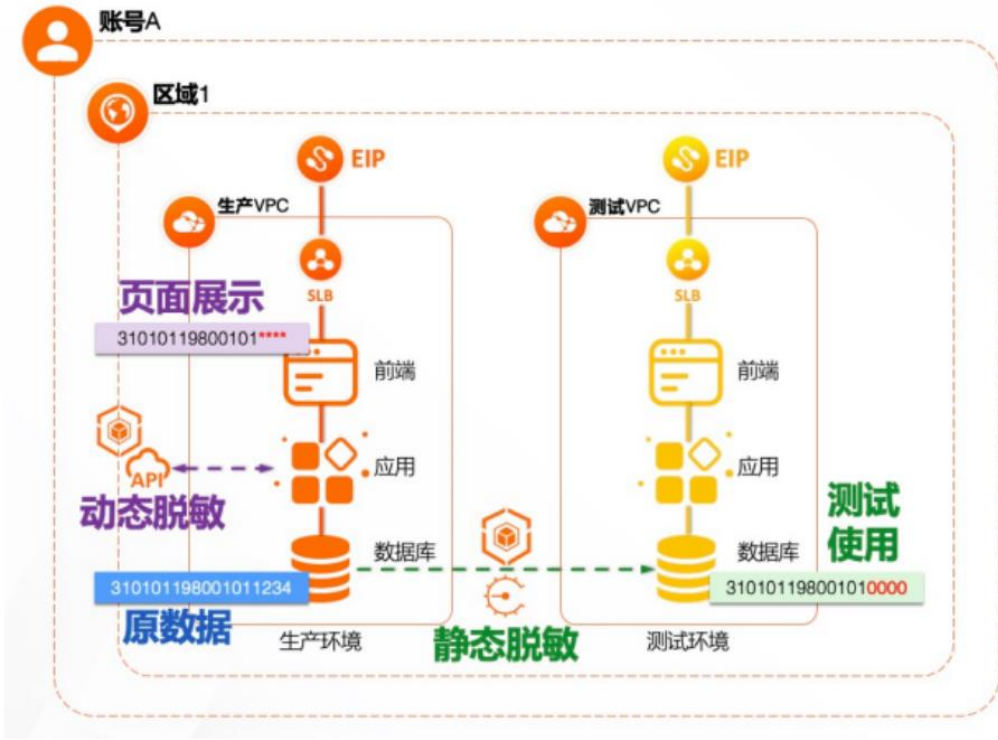
	Name	Gender	Birthday	Age
1	杨*	*	****-**-**	63
2	吴**	*	****-**-**	16
3	段**	*	****-**-**	34
4	万**	*	****-**-**	56
5	朱*	*	****-**-**	57
6	贺**	*	****-**-**	27

7 技术Demo

当前市场已有成熟数据脱敏框架结构简图大致如下：

在数据库及应用服务器与终端进行数据传输过程中，数据脱敏系统对传输信息进行脱敏处理，确保数据安全。

举例而言：



7 技术Demo

- 开源数据脱敏框架中，Apache ShardingSphere支持对数据表中某个或多个列进行数据加密&解密，且兼容所有常用SQL。

- 阿里云数据安全中心提供了大量的数据匿名化和去标识化算法以及相应脱敏模板，可以根据实际业务场景灵活选择，自定义参数，做到个性化数据脱敏。

- 其动态数据脱敏产品能够实现敏感数据自动检测，自定义脱敏算法，且脱敏规则细化到用户级别，并采用多级管控手段保护敏感数据。