差分隐私下求解近似最小割问题的算法设计

Finding Approximate Minimum Cut in Differential Privacy

吉林大学 计算机科学与技术学院 唐敖庆理科试验班

答辩人：周宇恒

指导老师：刘淼

答辩时间：2025年6月6日

研究背景

吉林大学 JILIN UNIVERSITY

差分隐私

- 医疗等数据敏感场景下，数据的使用者有责任保护信息安全。因此，研究者提出差分隐私（Differential Privacy）的概念来量化隐私泄露风险。
- 差分隐私关注算法输出中的个体隐私泄露。例如，对于统计平均年龄的算法，攻击者可以通过在输入名单中添加一个人并对比两次询问的结果，来获取具体某个人的年龄这一敏感信息。
- 差分隐私要求，单个数据应当对输出结果的影响不显著。研究者往往需要向算法中加入噪声作为隐私保护方法。

图1：使用场景

$$\mathbb{P}[A(G) \in O] \leq e^{\varepsilon}\mathbb{P}[A(G') \in O] + \delta$$

图2：差分隐私

研究背景

最小割问题

- 割（Cut）是顶点集合的二划分，割的权重为两个点集之间的边权和。最小割（Minimum Cut）为图中权重最小的割，最小割问题常在拓扑结构设计与资源分配优化等场景中出现。
- 差分隐私下的最小割算法需要控制割值的误差与输出的稳定性，这给算法设计带来了挑战。
- 仙人掌图表示法（Cactus Representation）是一种特殊的数据结构，其以规模为$O(n)$ 的稀疏化图表示了规模为$O(n^2)$的最小割。

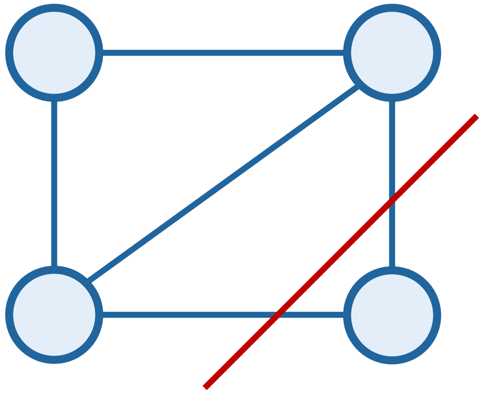


图3：最小割

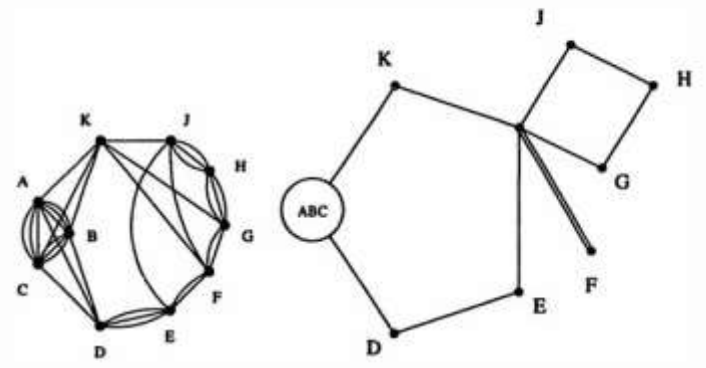


图4：仙人掌图表示法

差分隐私

拉普拉斯机制（Laplace Mechanism）

- 边相邻图输出的隐私化方法
- 对每个输出加入独立同分布的拉普拉斯分布噪声

指数机制（Exponential Mechanism）

- 选择最优值的隐私化方法
- 将估值函数的幂次作为权重输出结果

K优选择机制（Top-k Selection）

- 从m个值中选择前k优的隐私化方法
- 加入噪声后排序

$$f(x|\mu,b)=\frac{1}{2b}exp\left(-\frac{|x-\mu|}{b}\right)$$

图5：拉普拉斯分布

$$Pr[y=i]=\frac{exp(-\frac{\varepsilon}{2}x_i)}{\sum_{j\in[m]}exp(-\frac{\varepsilon}{2}x_j)}$$

图6：指数机制

最小割问题

➢ 最小割问题及相关差分隐私算法的研究已取得相当的进展，但尚未有最小割算法能够差分隐私地输出所有最小割的集合。

算法	误差	隐私性	输出	时间复杂度
Karger 收缩算法	精确值	非DP	最小割或近似最小割集	$O(n^2 \log^3(n))$
Karger 树包装算法	精确值	非DP	最小割	$O(m \log^3(n))$
隐私最小割算法	$\Theta(\log(n)/\varepsilon)$	纯DP	最小割	指数
图隐私化算法	$\tilde{O}(\frac{\sqrt{nm}}{\varepsilon})$	近似DP	差分隐私图	多项式
仙人掌图表示法构造算法	精确值	非DP	仙人掌图表示法，最小割集	$O(m \log^3(n))$

图7：现有的相关研究

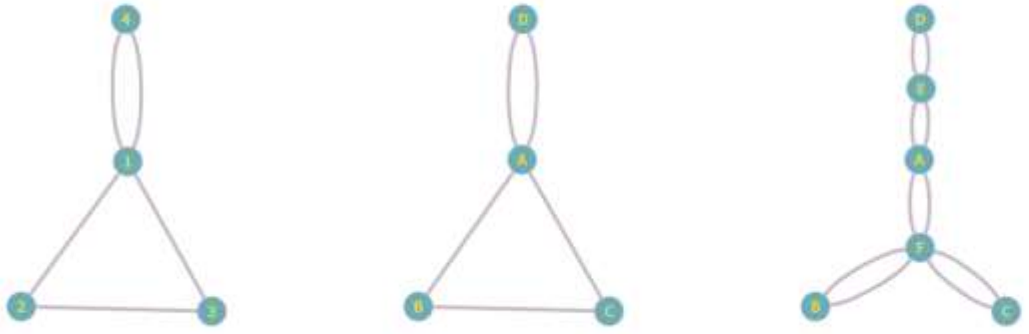
方法设计

标准仙人掌图表示法

➢ 本文发现：图的仙人掌图表示法不具有唯一性

```
output();
find_connected_components();
for(int i=1;i<=circle_number;i++){
    if(circle[i].size()==1)continue;
    for(auto k2:circle[i]){//Step5,由于映射由k2继承，因此用k2替换k
        int k1=add_node();
        for(auto e:E[k2]){
            if(e.se!=c/2)continue;//Step10，环边连向k1，其它边连向k2
            int v=e.fi;
            E[v][k2]=0;
            E[k2][v]=0;
            E[v][k1]=c/2;
            E[k1][v]=c/2;
        }
        E[k2][k1]=c;//Step7,k1,k2连边
        E[k1][k2]=c;
    }
}
```

图9：核心代码



(a) 图 G　　(b) 图 Γ　　(c) 图 Γ'

图8：仙人掌图表示法的不唯一性

a）分析P割与T割的性质，定义标准仙人掌图表示法，并给出构造算法

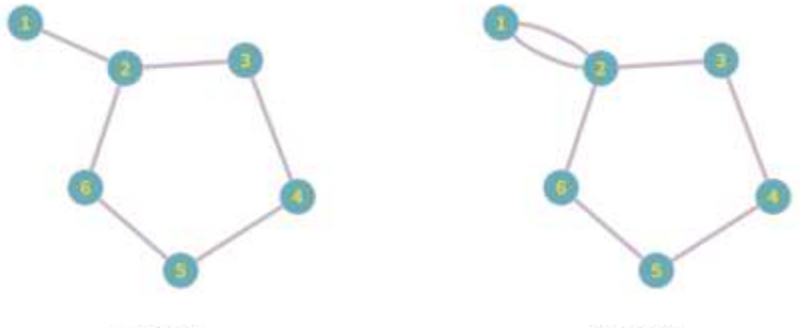
b）设计仙人掌图表示法的标准化算法

c）完成代码验证

方法设计

最小割数量敏感度分析

- 给出高敏感度边相邻图的构造
- 基于仙人掌图表示法的结构定量分析敏感度
- 基于原图最小割数量M_G定量分析敏感度



(a) 图 G　　(b) 图 G'

图10：n=6时的构造示例

$$M_G - \alpha_d - \min\{\alpha_d, \alpha_c, \frac{\alpha_g}{\alpha_r}\} \cdot \max\{0, \lfloor \frac{\alpha_r^2}{4} - 2 \rfloor\} \le M_{G'} \le M_G$$

图11：基于仙人掌图表示法的敏感度分析结果

$$\frac{M_G}{2} - 1.5n \le M_{G'} \le M_G$$

图11：基于M_G的敏感度分析结果

差分隐私最小割算法

基于差分隐私图的算法

- 发布差分隐私图后枚举割
- 近似误差为$\tilde{O}(\frac{\sqrt{nm}}{\varepsilon})$

基于k优选择机制的算法

- 枚举所有割后使用k优选择机制
- 近似误差为$\tilde{O}\left(\frac{\mathrm{n}\sqrt{n}}{\varepsilon}\right)$，提供优化可能

加法近似参数的优化

- 使用指数机制提高最小割的割值并应用Karger收缩算法缩小割的枚举范围，来降低k优选择机制的误差
- 近似误差为$\tilde{O}\left(\frac{\mathrm{n}\ln(n)}{\varepsilon}\right)$

本论文的贡献

1

标准仙人掌图表示法

- 提出了一种仙人掌图表示法的标准化算法，算法效率高，可以直接应用于现有仙人掌图表示法的构造算法中，为隐私化提供前置条件。

2

最小割数量敏感度分析模型

- 结合仙人掌图表示法分析最小割数量的敏感度，并给出最小割值相同的边相邻图的敏感度上界$0.5\mathrm{M}_G + 1.5\mathrm{n}$。

3

近似最小割集合输出算法

- 融合指数机制与 Karger 收缩算法，达成最优加性误差$O(\frac{n\,ln(n)}{\varepsilon})$。

感谢各位老师！

答辩人：周宇恒

指导老师：刘淼