

doi:10.1631/FITEE.1400395

题目: 面向优选应答的 k -匿名模型

目的: 数据采集的隐私保护问题近年来受到广泛关注。传统的数据采集过程中,采集机构假设在匿名发布数据的前提下,所有应答者对于提交各自数据的过程均是满意的。本文作者认为这一假设并不实际,因日益增长的隐私保护需求导致部分应答者拒绝或提交不准确的数据,从而将导致由这些数据得到的分析结果不可靠。因此,本文引入 k_r -匿名模型,由应答者选择自己偏好的匿名水平。

创新点: 本文所提算法其背后主要思想是允许每个应答者学习自身记录的事件数目,即仅需获取自身约束条件的满意度得分。

方法: 首先,生成唯一身份标识和约束条件。接着,检查约束条件的满意度情况。然后,计算满意度得分。最后,更新约束条件的满意度表格(图1)。

结论: 引入 k_r -匿名模型的概念,允许应答者在提交数据前能够选择自己偏好的匿名保护水平。所提算法确保应答者在数据采集过程中实现其偏好的匿名保护水平,且所采集的数据真实并有效用于数据分析。

关键词: 匿名数据采集; 优选应答的隐私保护; k -匿名