面向多敏感值的个性化随机响应机制设计与分析

宋海娜[1,2],罗 涛[1,2],韩新宇[1,2],李剑峰[1,2]
(1. 北京邮电大学北京先进信息网络实验室,北京 100876;
2. 北京邮电大学网络体系构建与融合北京市重点实验室,北京 100876)

摘 要: 在实际数据收集中,不同敏感值的敏感度有很大差异,隐私保护需求也不相同. 然而,现有的基于随机响应的本地化隐私保护模型针对所有敏感值都执行同样程度的隐私保护,从而可能造成某些低敏感度的敏感值过度保护,而某些高敏感度的敏感值却保护不足. 基于此,本文在常规随机响应(Conventional Randomized Response,CRR)模型的基础上,考虑个性化的隐私需求,引入敏感值权重,并将其引入到随机响应的决策中,提出一种面向多敏感值的个性化随机响应(Personalized Randomized Response,PRR)机制,该机制能够确保不同的敏感值群体均能达到各自期望的隐私保护程度,实现个性化的隐私保护. 理论分析和仿真实验表明,在机制的主观隐私泄露程度一定时,相比于CRR模型,本文所提的PRR机制统计估计误差更小,即获得的统计数据的质量更高,同时又保证了个性化的隐私保护.

关键词: 随机响应; 敏感值权重; 主观隐私泄露程度; 数据质量; 个性化隐私保护
中图分类号: TP309.2 **文献标识码:** A **文章编号:** 0372-2112 (2019)06-1236-08
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2019.06.008

Design and Analysis for Multiple Sensitive Values-Oriented Personalized Randomized Response

SONG Hai-na[1,2], LUO Tao[1,2], HAN Xin-yu[1,2], LI Jian-feng[1,2]
(1. *Beijing Laboratory of Advanced Information Networks, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*;
2. *Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*)

Abstract: In actual dada collection, the sensitivity of different sensitive information is different so that the concrete privacy need is different, too. However, the existing local privacy preservation model based on randomized response (RR), which is called conventional randomized response (CRR) for convenience, focuses on a universal approach that exerts the same amount of preservation for all sensitivity values, without catering for their concrete privacy needs. As a result, it may be offering insufficient protection to a subset of people with relatively higher privacy needs, while applying excessive privacy control to another subset with relatively lower privacy needs. Based on this, a new framework which is called personalized randomized response (PRR) is proposed based on the concept of CRR for multiple sensitive values-oriented personalized privacy preservation. The PRR technique considers personalized privacy needs, introduces sensitive value weights for different sensitive values, and then introduces the weights into the decision of RR for satisfying all sensitivity values' privacy needs, and thus, attains personalized privacy preservation. Both theoretical derivation and simulation experiment reveal that the estimate error of statistics of PRR mechanism is smaller than that of the CRR mechanism for a certain subjective degree of privacy leakage, that is, the quality of statistics obtained by PRR mechanism is higher than that of the CRR model while guaranteeing personalized privacy protection for a given subjective degree privacy preservation.

Key words: randomized response; sensitive value weight; subjective degree of privacy leakage; data quality; personalized privacy preservation

收稿日期:2018-11-23;修回日期:2019-04-02;责任编辑:覃怀银
基金项目:国家重点研发计划重点专项(No. 2016YFF0201003);国家自然科学基金(No. 61571065)

1 引言

随机响应(Randomized Response,RR)是基于数据失真的本地化隐私保护技术的主流扰动机制,模型简洁直观且易于实现,并且其扰动程度可直接量化,在统计特性方面性能优良,因此受到广泛的关注[1~4]. RR 采用依概率作答的方式来保护受访者的隐私[4],它保证敏感问题作答具有很强的可否认性,已经在 Google Chrome 的隐私保护工具和 Apple 系统中应用. 同时,RR 充分考虑了数据采集过程中数据收集者窃取或泄露用户隐私的可能性,该模型中受访者能够独立地对个体数据进行隐私化处理,大大激发其参与数据收集的积极性,并且不再需要可信第三方的介入,也免除了不可信第三方数据收集者可能带来的隐私泄露与攻击[5].

Warner 于 1965 年首次提出利用 RR 技术[1](简称 W-RR)进行敏感数据的收集,借鉴统计学研究中的经典方法,模拟调查者在尽量不侵犯被访者隐私的前提下收集到有价值的统计数据的过程. 基于 W-RR 模型,文献[6~8]使用离散无记忆二进制对称信道对其进行建模,并有学者将其推广到多元离散对称隐私信源提出 K-RR[9]、O-RR[10] 等,得到了广泛的应用. 差分隐私[11]作为一种公认的强健的隐私保护模型被提出,受到了很多学者的青睐[12~15]. 顺应这一发展趋势,很多学者对 RR 模型下的本地化差分隐私机制进行研究,通过单一隐私预算参数 ε 来衡量机制的隐私保护程度,主要集中在最优的差分隐私机制的设计与分析,重点关注机制客观的隐私-效用的折中[14,15].

以上随机响应模型(简称 CRR 模型)及其相关研究均默认不同隐私数据同等重要,隐私保护需求相同,进而对所有的敏感值都执行相同的隐私保护操作,主要关注的是整个体系的客观隐私保护程度. 然而,这可能会造成某些敏感度低的敏感值过度保护,而某些敏感度高的敏感值欠缺保护. 这是因为,实际中隐私信息并非同等重要,隐私保护的需求也不尽相同. 换言之,受访者对不同敏感值的敏感度的主观感受存在很大差异,因而隐私保护需求也就不同,此时隐私信息的敏感度在一定程度上反映了隐私需求的高低[16].

实际中,不同敏感值的敏感度有很大差异,隐私保护需求也不同. 如果硬性地对所有敏感值进行相同等级的隐私保护,直观上会降低统计数据的质量. 另一方面,对高敏感度的敏感值进行足够高的隐私保护,而对低敏感度的敏感值进行相对较小的隐私保护(达到其隐私需求即可),并将各个敏感值的主观隐私泄露程度直观地反馈给受访者,能够激发受访者的积极性,且激励受访者更加客观地提交扰动后的数据,进而使得收集到的数据更加客观,数据可用性更高. 根据课题组的调研来看,针对 RR 机制的相关研究,截止目前还未找到解决上述问题类似的研究. 因此,本文在 CRR 模型的基础上,针对多元离散隐私信源,考虑个性化的隐私保护需求,引入敏感值权重的概念,并将其引入到随机响应决策中,提出一种面向多敏感值的个性化随机响应(Personalized Randomized Response,PRR)机制. 分析表明,相比于 CRR 机制,本文所提的 PRR 机制收集到的统计数据的质量会更高,同时实现了个性化的隐私保护,具有很强的实际意义.

2 系统模型

2.1 RR 模型

对于 m 元离散有限隐私信源 $X \in \mathcal{X}$(其中,$\mathcal{X} = \{x_i\}_{i=1}^{m}$,$|\mathcal{X}| = m \geq 2$),采用扩展的 m 元 RR 模型来进行本地化隐私保护数据的收集,其 RR 过程可建模为多元离散信道的传输过程,其输出为 $Y \in \mathcal{Y}(\mathcal{Y} = \{y_j\}_{j=1}^{m})$,如图 1 所示. 隐私保护机制 $\boldsymbol{P}_{Y|X}$ 可描述为 $X = x$ 映射到 $Y = y$ 的条件概率 $P_{Y|X}(Y = y \mid X = x) = P_{Y|X}(y \mid x)$. 若多元离散信道是对称的,则 $\mathcal{X} = \mathcal{Y}$,本文主要考虑该种对称隐私信道. 此时,当 $i = j$ 时,$x_i = y_j (i,j = 1,2,\cdots,m)$.

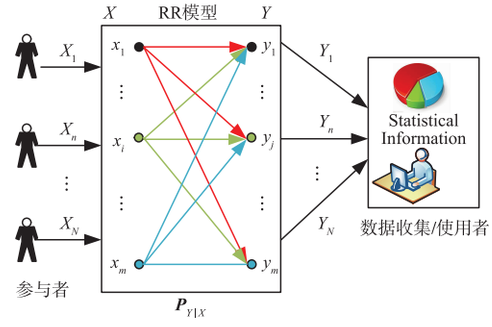


图1　基于RR的m元隐私信道建模及其本地化隐私保护数据收集模型

隐私分布估计问题描述为:隐私信源的先验分布为 $\boldsymbol{P}_X = [P_1, P_2, \cdots, P_m]$,参与者的样本量为 $N(N \gg m)$,样本 $X_1, X_2, \cdots, X_n, \cdots, X_N (n = 1,2,\cdots,N)$ 是来自于分布 $\boldsymbol{P}_X$,不同个体间隐私数据相互独立且同分布,且假设每个个体仅拥有一种敏感值. 样本 X_n 经过信道 $\boldsymbol{P}_{Y|X}$ 的随机扰动处理,得到扰动数据 Y_n,如图 1 所示. 其中,Y_n 的分布是根据 $\boldsymbol{P}_Y = \boldsymbol{P}_X \boldsymbol{P}_{Y|X}$ 来确定的,由 X_n 到 Y_n 的过程称为 RR 过程,信道转移概率 $\boldsymbol{P}_{Y|X}$ 称为隐私保护机制. 本文基于 RR 提出个性化隐私保护的目的是:在本地化数据收集中,使个体对自己的隐私可控,高隐私需求的个体执行较高强度的扰动,低隐私需求的个体执行较低强度的扰动,从而实现个性化的隐私保护,而且能够激发受访者参与的积极性,激励受访者客观地提供扰动数据,提高统计数据的质量. 另外,在一定的隐私泄露条件下,数据收集者由 $Y^N = \{Y_1, Y_2, \cdots, Y_n, \cdots, Y_N\}$ 和

$\boldsymbol{P}_{Y|X}$进行统计分析,以得到有效的估计结果.

2.2 CRR 模型

CRR 模型中默认隐私信息同等重要,隐私保护需求相同. 设置 CRR 模型的隐私保护机制为 $\boldsymbol{P}_{\mathrm{CRR}}$,可由 $m\times m$ 的行随机矩阵表示:

$$P_{\mathrm{CRR}}(y_j|x_i)=\begin{cases}P, & \text{if } j=i\\ \dfrac{1-P}{m-1}, & \text{if } j\neq i\end{cases} \tag{1}$$

即对每种敏感值都以 P 的概率保持原值,以$(1-P)/(m-1)$的概率响应输出其他 $m-1$ 种敏感值中的一种,称之为 m-CRR 模型,其中 $\boldsymbol{P}_{\mathrm{CRR}}$为 $m\times m$ 的对称矩阵,且满足$|\boldsymbol{P}_{\mathrm{CRR}}|\neq 0$,即 $\boldsymbol{P}_{\mathrm{CRR}}$可逆.

由式(1)可知,CRR 模型对所有敏感值都执行相同的隐私保护操作,这样会造成某些敏感度低的敏感值过度保护,而某些敏感度高的敏感值欠缺保护. 然而,在很多情况下隐私信息并非同等重要,隐私保护需求也不同. 基于此,本文研究面向敏感值的 PRR 机制.

2.3 PRR 模型

PRR 模型考虑不同敏感值的敏感度不同,隐私保护需求亦不同. 因此,根据敏感度的不同为各个敏感值赋予一定的权重因子,即敏感值权重,以便对不同敏感值进行个性化的设计与分析.

定义 1 敏感值权重 w_i:对于敏感值 x_i,通过权重 $w_i(i=1,2,\cdots,m)$来刻画用户对不同敏感值的主观敏感度,即敏感值权重,也称为主观敏感度.

注:隐私信息的敏感度与具体的应用场景、个体的主观感受与偏好、经济效益等诸多因素有关,因此本文重点不在敏感值权重的设置上,重点关注 PRR 机制本身的设计与分析. 但是,敏感值权重 w_i 的设定应遵循以下两个原则:

(1)原则一,敏感值的敏感度越高,相应的敏感值权重越大. 比如:在“AIDS”和“Flu”两种敏感值中,“AIDS”更加敏感,其隐私泄露所带来的影响或损失更大,许多受访者会采取不合作的态度,不予回答,这给抽样结果的分析造成极大的困难,因此设定“AIDS”的敏感值权重相对更大是合理的.

(2)原则二,敏感值的频率越低,相应的敏感值权重越大. 借鉴信息论中信息熵概念,可知:敏感值 x_i 所占的比例越小,其信息量越多.

对于敏感值 x_i,其主观敏感度越高,相应的敏感值权重 w_i 越大,隐私保护的需求也越高. PRR 模型考虑个性化的隐私保护需求,将 w_i 引入到敏感值 x_i随机响应的决策中,其隐私保护机制 $\boldsymbol{P}_{\mathrm{PRR}}$为

$$P_{\mathrm{PRR}}(y_j|x_i)=\begin{cases}P_{i,j}(w_i), & \text{if } j=i\\ \dfrac{1-P_{i,j}(w_i)}{m-1}, & \text{if } j\neq i\end{cases} \tag{2}$$

称之为 m-PRR 模型. 其中,$P_{i,i}(w_i)$可形象地描述为拥有敏感值 x_i 的个体讲真话的概率,且 $P_{i,i}(w_i)$是关于 w_i 的减函数. 这是因为,$P_{i,i}(w_i)$从侧面反映了敏感值 x_i 的隐私保护程度:$P_{i,i}(w_i)$越小,敏感值 x_i 的隐私保护程度越高,相应的隐私泄露风险越小,反之亦然. 因此,敏感值 x_i 的敏感度越高,保持原值的概率 $P_{i,i}(w_i)$应该设置的越小. 此时,PRR 模型能够激励用户更加客观地提交数据,进而使得收集的数据更加客观,特别是高敏感度的敏感值.

3 基于 RR 机制的隐私泄露程度与统计数据质量分析

3.1 数据隐私泄露程度分析

3.1.1 CMAP 准则下 CRR 机制的客观隐私泄露程度

若仅已知隐私信息的先验分布 $\boldsymbol{P}_X$,则此时正确推测出隐私信息的平均概率为

$$P_{c0}=\max_i\{P_X(x_i)\} \tag{3}$$

其中,$P_X(x_i)$表示先验分布中 $X=x_i$ 的概率,$P_{c0}\leq 1$ 反映的是攻击者仅已知先验信息条件下的最大隐私泄露程度,其值越大,说明此时的隐私泄露程度越大.

定义 2 CMAP 准则:选择重构函数 $F(y_j)=\hat{x}(\hat{x}\in X,y_j\in Y)$,使之满足条件

$$P_{X|Y}(\hat{x}|y_j)\geq P_{X|Y}(x_i|y_j),\quad x_i\in X,y_j\in Y \tag{4}$$

则称该准则为常规最大后验概率(Conventional Maximum A Posteriori, CMAP)准则”或“最小错误概率准则”.

已知后验信息,CMAP 准则下平均错误概率为

$$P_e=\sum_{j=1}^{m}P_Y(y_j)\sum_{i=1}^{m}P_{X|Y}(x_i|y_j)d_{\mathrm{H}}(k_j,i) \tag{5}$$

其中,$d_{\mathrm{H}}(k_j,i)$表示汉明距离,P_e 也称为非加权错误率或贝叶斯风险,能够客观地从侧面反映数据的安全性. 其中,

$$k_j=\arg\max_{i\in\{1,2,\cdots m\}}\{P_{X|Y}(x_i|y_j)\},j\in\{1,2,\cdots m\} \tag{6}$$

同理,令 CMAP 准则下平均正确概率为 P_c,由 $P_e+P_c=1$,则 P_c 为

$$\begin{aligned}P_c&=\sum_{j=1}^{m}\max_i\{P_Y(y_j)P_{X|Y}(x_i|y_j)\}\\&=\sum_{j=1}^{m}P_X(x_{k_j})P_{Y|X}(y_j|x_{k_j})\\&=1-P_e\end{aligned} \tag{7}$$

其中,P_c 是关于 $\boldsymbol{P}_X$ 和 $\boldsymbol{P}_{Y|X}$的函数,也称为非加权(平均)正确率或贝叶斯效用,直观衡量了数据客观上的安全性:P_c 越大,数据客观上的安全性越低. 在一次猜测中,已知后验信息下正确猜测出个体真实信息的平均最大概率为 P_c.

定义 3 客观隐私泄露程度:CRR 机制的客观隐私

泄露程度定义为已知后验信息后的正确率与仅已知先验信息条件下的正确率的比值[17]

$$L_{\mathrm{c}} \triangleq \frac{P_{\mathrm{c}}}{P_{\mathrm{c0}}} \tag{8}$$

其中,$L_{\mathrm{c}} \geq 1$ 表示 CMAP 准则下由 CRR 机制带来的相对隐私泄露程度:L_{c} 越大,相应的隐私保护程度越低.

对于敏感值 x_i 的群体而言,其更加关注的是敏感值 x_i 的隐私保护程度,而并非整个隐私保护体系的客观隐私保护程度,CMAP 准则下正确推测出敏感值 x_i 的平均概率为

$$P_{ci} = \sum_{j=1}^{m} P_Y(y_j) P_{X|Y}(x_{k_j} | y_j) \bar{d}_{\mathrm{H}}(k_j, i) \tag{9}$$

其中,$\bar{d}_{\mathrm{H}}(k_j, i)$ 是对 $d_{\mathrm{H}}(k_j, i)$ 进行取反操作. P_{ci} 反映的是 CMAP 准则下原始敏感值为 x_i 且能够正确推测出 x_i 的平均概率. 此时,所有敏感值平均正确率之和为机制的平均正确率,即 $\sum_{i=1}^{m} P_{ci} = P_{\mathrm{c}}$.

定义 4　敏感值 x_i 的客观隐私泄露程度:针对敏感值 x_i,其客观隐私泄露程度 L_{ci} 定义为 CMAP 准则下原始隐私信息为 x_i 且能够正确推测出 x_i 的平均概率与 x_i 的先验概率的比值:

$$L_{ci} \triangleq \frac{P_{ci}}{P_X(x_i)} \tag{10}$$

其中,L_{ci} 满足:$0 \leq L_{ci} \leq 1$. 当 $L_{ci} = 0$ 时,说明 CRR 机制对敏感值 x_i 起到了完全的保护.

3.1.2　PMAP 准则下 PRR 机制的主观隐私泄露程度

考虑到不同隐私信息并非同等重要,隐私保护需求亦不同,因此 PRR 机制将充分考虑敏感值权重带来的影响. PRR 机制下,若仅已知 $\boldsymbol{P}_X$,根据加权隐私效益最大化原则,此时正确推测出隐私信息的平均加权概率为

$$P_{\mathrm{cw0}} = \max_i \{ w_i P_X(x_i) \} \tag{11}$$

其中,P_{cw0} 反映的是仅已知 $\boldsymbol{P}_X$ 下,根据隐私加权概率最大化原则,所获得的最大加权效益,其值越大,说明隐私泄露程度越大.

定义 5　PMAP 准则:选择重构函数 $F(y_j) = \hat{x}(\hat{x} \in X, y_j \in Y)$,使之满足

$$w_{i^*} P_{X|Y}(\hat{x} | y_j) \geq w_i P_{X|Y}(x_i | y_j), \quad x_i \in X, y_j \in Y \tag{12}$$

其中,$i^* = \arg_{i \in \{1,2,\cdots,m\}} \{ \hat{x} = x_i \}$,则称该准则为“隐私加权最大后验概率(Privacy-weighted Maximum A Posteriori,PMAP)准则”或“隐私加权最小错误概率准则”.

定义 6　隐私加权正确率/贝叶斯效用:通过权重 $w_i (i = 1, 2, \cdots, m)$ 刻画隐私谋取者获取后验信息的条件下,采用 PMAP 准则,令

$$k_j = \underset{i \in \{1,2,\cdots m\}}{\operatorname{argmax}} \{ w_i P_{X|Y}(x_i | y_j) \}, j \in \{1, 2, \cdots, m\} \tag{13}$$

则能够正确推测出带主观感受隐私信息的加权概率为

$$P_{\mathrm{cw}} \triangleq \sum_{j=1}^{m} w_{k_j} P_Y(y_j) P_{X|Y}(x_{k_j} | y_j) \tag{14}$$

称该概率 P_{cw} 为隐私加权正确率或隐私加权贝叶斯效用. P_{cw} 能够衡量机制主观数据安全性:P_{cw} 越大,机制主观上的安全性越低.

定义 7　主观隐私泄露程度:PRR 机制的主观隐私泄露程度定义为已知后验信息条件下的隐私加权正确率与仅已知先验信息条件下的隐私加权正确率的比值:

$$L_{\mathrm{cw}} \triangleq \frac{P_{\mathrm{cw}}}{P_{\mathrm{cw0}}} \tag{15}$$

其中,$L_{\mathrm{cw}} \geq 1$ 反映的是 PMAP 准则下由 PRR 机制带来的相对隐私泄露程度,其值越大,相应的隐私泄露程度越高.

同理,针对 PRR 机制,PMAP 准则下敏感值为 x_i 且正确推测出 x_i 的隐私加权概率为

$$P_{\mathrm{cw}i} = \sum_{j=1}^{m} w_{k_j} P_Y(y_j) P_{X|Y}(x_{k_j} | y_j) \bar{d}_{\mathrm{H}}(k_j, i) \tag{16}$$

其中,$P_{\mathrm{cw}i}$ 满足:$\sum_{i=1}^{m} P_{\mathrm{cw}i} = P_{\mathrm{cw}}$.

定义 8　敏感值 x_i 的主观隐私泄露程度:PRR 机制下,敏感值 x_i 的主观隐私泄露程度 $L_{\mathrm{cw}i}$ 定义为 PMAP 准则下敏感值为 x_i 且正确推测出 x_i 的隐私加权概率与敏感值 x_i 的隐私加权先验概率的比值:

$$L_{\mathrm{cw}i} \triangleq \frac{P_{\mathrm{cw}i}}{w_i P_X(x_i)} \tag{17}$$

其中,$L_{\mathrm{cw}i}$ 满足:$0 \leq L_{\mathrm{cw}i} \leq 1$. $L_{\mathrm{cw}i}$ 越大,说明敏感值 x_i 主观上的隐私泄露程度越高,意味着 PRR 机制对敏感值 x_i 保护程度越低.

为了兼容 CRR 模型,敏感值权重 w_i 的设置应满足“归 m 化”条件,即 $\sum_{i=1}^{m} w_i = m$,且 $0 < w_i < m$. CRR 模型在 PRR 模型中被视为敏感值权重相等的一种情况,即对所有的敏感值 x_i,其敏感值权重均为 $w_i \equiv 1 (i = 1, 2, \cdots, m)$,此时基于 PRR 模型的所有理论推导与分析均适用于 CRR 模型. 因此,采用归 m 化处理后,CRR 模型为 PRR 模型的一种特例.

3.2　统计数据质量分析

基于 RR 的本地化隐私保护机制旨在将隐私泄露控制在源端,同时关注统计数据的质量,本小节使用估计误差来评估两种模型的隐私分布估计问题[10]. 将敏感值权重按照从大到小的顺序排列:$w_{i1} \geq w_{i2} \geq \cdots \geq w_{im}$,相应敏感值的隐私保护需求依次减小,由 w_i 与 $\boldsymbol{P}_{Y|X}$ 的关系,相应经过 RR 模型输出真实值的概率应该满足:$P_{i1,i1} \leq P_{i2,i2} \leq \cdots \leq P_{im,im}$,其中敏感值权重 w_{i1} 对应的敏感值 x_{i1} 群体隐私保护需求最大.

3.2.1　CRR 模型下的隐私分布的估计误差分析

CRR 模型下,不同敏感值同等重要,隐私需求相

同. 此时,对各个敏感值均执行相同的隐私保护操作,设置 $P_{i1,i1}=P_{i2,i2}=\cdots=P_{im,im}=P$. 敏感值 x_i 的先验概率 $P_X(x_i)$ 简记为 P_i,其经验估计值为 $\hat{P}_i$,根据最大似然估计准则:

$$\hat{P}_i=\frac{P-1}{mP-1}+\frac{m-1}{mP-1}\frac{N_i}{N},P\neq\frac{1}{m} \tag{18}$$

其中,$N_i=\sum_{n=1}^{N}\bar{d}_{\mathrm{H}}(Y_n,y_i)$ 表示 CRR 模型下输出数据 $Y^N=\{Y_1,Y_2,\cdots,Y_n,\cdots,Y_N\}$ 中满足 $Y_n=y_i$ 的数量. $\hat{P}_i$ 是其真实值 P_i 的无偏估计,该过程中 P_i 的估计误差为

$$\mathrm{Var}(\hat{P}_i|P_i)=\frac{(m-2+P)(1-P)}{N(mP-1)^2}+\frac{(m-3+2P)P_i}{N(mP-1)}-\frac{P_i^2}{N} \tag{19}$$

CRR 模型下 $\boldsymbol{P}_X$ 的经验估计记为 $\hat{\boldsymbol{P}}_X$,由 $\boldsymbol{P}_X\boldsymbol{P}_{\mathrm{CRR}}=\boldsymbol{P}_Y$,$\hat{\boldsymbol{P}}_X$ 可描述为

$$\hat{\boldsymbol{P}}_X=\hat{\boldsymbol{P}}_Y\boldsymbol{P}_{\mathrm{CRR}}^{-1} \tag{20}$$

其中,$\hat{\boldsymbol{P}}_Y$ 为 $\boldsymbol{P}_Y$ 的经验估计,$\boldsymbol{P}_{\mathrm{CRR}}^{-1}$ 为 $\boldsymbol{P}_{\mathrm{CRR}}$ 的逆矩阵. 根据谢尔曼-莫里森公式,对于任意 $m\geq2$,$N\gg m$ 和 $P\neq1/m$,则 $\boldsymbol{P}_{\mathrm{CRR}}^{-1}$ 为

$$(\boldsymbol{P}_{\mathrm{CRR}}^{-1})_{i,j}=\frac{1}{mP-1}\begin{cases}P+m-2, & \text{if } j=i\\ P-1, & \text{if } j\neq i\end{cases} \tag{21}$$

其中,$(\cdot)_{i,j}$ 表示矩阵的第 i 行,第 j 列的元素. 采用 $E\|\hat{\boldsymbol{P}}_X-\boldsymbol{P}_X\|_2^2$ 来评估 CRR 模型的隐私分布估计问题[10],称为估计误差,则

$$\begin{aligned}E\|\hat{\boldsymbol{P}}_X-\boldsymbol{P}_X\|_2^2&=\sum_{i=1}^{m}\mathrm{Var}\Big(\sum_{j=1}^{m}\hat{P}_{y_j}(\boldsymbol{P}_{\mathrm{CRR}}^{-1})_{j,i}\Big)\\&=\frac{1}{N}\sum_{i=1}^{m}\Big(\sum_{j=1}^{m}((\boldsymbol{P}_{\mathrm{CRR}}^{-1})_{j,i})^2P_{y_j}(1-P_{y_j})\\&\quad-2\sum_{j,k=1,j<k}^{m}(\boldsymbol{P}_{\mathrm{CRR}}^{-1})_{j,i}(\boldsymbol{P}_{\mathrm{CRR}}^{-1})_{k,i}P_{y_j}P_{y_k}\Big)\end{aligned} \tag{22}$$

其中,$\hat{P}_{y_j}=N_j/N=\sum_{n=1}^{N}\bar{d}_{\mathrm{H}}(Y_n,y_j)/N$ 表示 CRR 模型下输出数据 Y^N 中 y_j 所占的实际比例,P_{y_j} 是 $\hat{P}_{y_j}$ 的期望值. 由式(19)和式(21),式(22)亦可描述为

$$\begin{aligned}E\|\hat{\boldsymbol{P}}_X-\boldsymbol{P}_X\|_2^2&=\sum_{i=1}^{m}\mathrm{Var}(\hat{P}_i|P_i)\\&=\frac{m(m-2+P)(1-P)}{N(mP-1)^2}\\&\quad+\frac{(m-3+2P)}{N(mP-1)}-\frac{\sum_i P_i^2}{N}\end{aligned} \tag{23}$$

$E\|\hat{\boldsymbol{P}}_X-\boldsymbol{P}_X\|_2^2$ 越小,说明估计误差越小,意味着隐私分布估计的准确性越高.

考虑到敏感值 x_{i1} 的敏感值权重 w_{i1} 最大,其隐私保护需求最高,则 $P_{i1,i1}=P$ 最小. 本文借鉴 K-RR 模型的参数设计,引入参数 $\lambda>0$,为了满足所有敏感值的隐私需求,同时要求隐私泄露风险最低,不妨构造式(1)中 $\boldsymbol{P}_{\mathrm{CRR}}$ 为

$$P_{\mathrm{CRR}}(y_j|x_i)=\frac{1}{m-1+\mathrm{e}^{\lambda/w_{i1}}}\begin{cases}\mathrm{e}^{\lambda/w_{i1}}, & \text{if } j=i\\ 1, & \text{if } j\neq i\end{cases} \tag{24}$$

即对所有敏感值 x_i 均以 $P=\mathrm{e}^{\lambda/w_{i1}}/(m-1+\mathrm{e}^{\lambda/w_{i1}})$ 的概率响应输出其真实值,这是满足 CRR 模型隐私需求的参数设置中的一个特例. 实际中,λ 可由敏感值 x_{i1} 具体的隐私保护需求进行设置. 另外,L_{c} 和 L_{ci} 均与 λ 有关,实际中可通过调节 λ 的大小来实现相应的隐私保护程度.

3.2.2 PRR 模型下的隐私分布的估计误差分析

PRR 模型中考虑到不同敏感值的敏感度不同,造成其隐私保护需求不同,因此进行个性化的隐私保护. PRR 模型下 P_i 的经验估计值记为 $\hat{\hat{P}}_i$,$\boldsymbol{P}_X$ 的经验估计记为 $\hat{\hat{\boldsymbol{P}}}_X=\{\hat{\hat{P}}_i\}_{i=1}^{m}$:

$$\hat{\hat{\boldsymbol{P}}}_X=\hat{\boldsymbol{P}}_Y\boldsymbol{P}_{\mathrm{PRR}}^{-1} \tag{25}$$

其中,$\boldsymbol{P}_{\mathrm{PRR}}^{-1}$ 为 $\boldsymbol{P}_{\mathrm{PRR}}$ 的逆矩阵. 同理,PRR 模型下隐私分布的估计误差为

$$\begin{aligned}E\|\hat{\hat{\boldsymbol{P}}}_X-\boldsymbol{P}_X\|_2^2&=\sum_{i=1}^{m}\mathrm{Var}(\hat{\hat{P}}_i|P_i)\\&=\sum_{i=1}^{m}\mathrm{Var}\Big(\sum_{j=1}^{m}\hat{\hat{P}}_{y_j}(\boldsymbol{P}_{\mathrm{PRR}}^{-1})_{j,i}\Big)\\&=\frac{1}{N}\sum_{i=1}^{m}\Big(\sum_{j=1}^{m}((\boldsymbol{P}_{\mathrm{PRR}}^{-1})_{j,i})^2P_{y_j}(1-P_{y_j})\\&\quad-2\sum_{j,k=1,j<k}^{m}(\boldsymbol{P}_{\mathrm{PRR}}^{-1})_{j,i}(\boldsymbol{P}_{\mathrm{PRR}}^{-1})_{k,i}P_{y_j}P_{y_k}\Big)\end{aligned} \tag{26}$$

其中,$\hat{\hat{P}}_{y_j}=N_j/N=\sum_{n=1}^{N}\bar{d}_{\mathrm{H}}(Y_n,y_j)/N$ 表示 PRR 模型下输出数据 Y^N 中 y_j 所占的实际比例,P_{y_j} 表示 $\hat{\hat{P}}_{y_j}$ 的期望值.

同理,对于敏感值 x_i,根据 w_i 与 $\boldsymbol{P}_{\mathrm{PRR}}$ 的关系,不妨构造式(2)中的 $\boldsymbol{P}_{\mathrm{PRR}}$ 为

$$P_{\mathrm{PRR}}(y_j|x_i)=\frac{1}{m-1+\mathrm{e}^{\mu/w_i}}\begin{cases}\mathrm{e}^{\mu/w_i}, & \text{if } j=i\\ 1, & \text{if } j\neq i\end{cases} \tag{27}$$

即敏感值 x_i 群体以 $\mathrm{e}^{\mu/w_i}/(m-1+\mathrm{e}^{\mu/w_i})$ 的概率响应其真实值,同时需满足:$P_{i1,i1}\leq P_{i2,i2}\leq\cdots\leq P_{im,im}$. 其中,$\mu_i$ 可由敏感值 x_i 群体的具体隐私需求进行设置,实现了面向多敏感值的个性化隐私保护. 当所有敏感值的敏感度均相同时,即对于任意 $i,j\in\{1,2,\cdots,m\}$,$\mu_i=\mu_j=\lambda$,$w_i=w_j=1$ 时,式(27)与 CRR 模型的表达形式相同,因此式(27)是一种更为广泛的定义形式. 对于敏感值 x_i 而言,其 w_i 越大,意味着其隐私保护需求越高,倾向响应输出其真实值的概率相对较低,而式(27)的约束规则刚好符合这一趋势.

PRR 模型的隐私分布估计值$\hat{\hat{\boldsymbol{P}}}_X$相对于 CRR 模型的隐私分布估计值$\hat{\boldsymbol{P}}_X$的效率定义为相对效率$R_e$[1],描述为

$$R_e=\frac{E\parallel\hat{\boldsymbol{P}}_X-\boldsymbol{P}_X\parallel_2^2}{E\parallel\hat{\hat{\boldsymbol{P}}}_X-\boldsymbol{P}_X\parallel_2^2}\times100\% \tag{28}$$

在一定的主观隐私泄露程度下,最大化数据的统计质量,满足:$R_e\geq1$. 这是因为 CRR 模型对所有的敏感值都采用相同的隐私保护操作,为了满足各自的隐私需求同时最小化隐私泄露程度,一般采用隐私保护需求最大的敏感值所需要的隐私保护参数设置,如此会造成其他敏感度低的敏感值过度保护,而 PRR 模型是根据具体敏感值的个性化隐私需求设置其隐私参数,实现了个性化的隐私保护. 另一方面,低敏感度的敏感值群体对隐私保护的需求较低,倾向响应输出其真实值的概率相对较大,如此为统计数据的可用性做出较大贡献. 显然,本文 PRR 机制下的估计值$\hat{\hat{\boldsymbol{P}}}_X$相比于 CRR 模型更加有效.

4　实例分析

为了验证本文所提 PRR 方案的有效性和可行性,本小节通过 MATLAB 仿真模拟调查者使用图 1 的 RR 模型来进行隐私分布估计的过程. 设置参与样本数 $N=10^5$,样本取值集合 $X=\{x_1,x_2,x_3\}=\{\text{AIDS},\text{Cancer},\text{Flu}\}$,即 $m=3$,每种疾病代表一种敏感值,样本来源于分布 $\boldsymbol{P}_X=[P_1,P_2,P_3]=[0.2,0.3,0.5]$(注:实际中 $\boldsymbol{P}_X$ 是未知的,且本小节设置 $m=3$,$\boldsymbol{P}_X=[0.2,0.3,0.5]$只是一个特例,本文 PRR 机制适用于任意 $m\geq2$ 元离散有限敏感值下基于 RR 的个性化隐私保护). 其中,根据 w_i 的设置原则,三种敏感值权重如表 1 所示. 可知,“AIDS”的敏感度最高,其敏感值权重最大;“Flu”的敏感度最低,其敏感值权重最小. 同时,3-CRR 和 3-PRR 模型分别采用式(24)和(27)的隐私保护机制. 为便于分析,PRR 模型中设置 $\mu_i=\mu_j=\mu(i,j=1,2,3)$.

表 1　三种敏感值权重及其先验概率分布

敏感值 X	x_1	x_2	x_3
疾病类型	AIDS	Cancer	Flu
敏感值权重 $w_i(\times3)$	0.6	0.3	0.1
概率	0.2	0.3	0.5

图 2 分别展示了 3-CRR 模型在 CMAP 准则下机制的客观隐私泄露程度 L_c 和 3-PRR 模型在 PMAP 准则下机制的主观隐私泄露程度 L_{cw}. 可知,CRR(或 PRR)模型下机制的客观(或主观)隐私泄露程度随着参数 λ(或 μ)的增大而增大,意味着机制的隐私保护程度越来越低,即隐私保护程度与 λ(或 μ)呈负相关的关系,这与 3.1 小节的分析一致. 在实际应用中,应该综合考虑隐私保护和数据质量的需求来设置合适的 λ 和 μ 值,进而将机制的隐私泄露程度控制在一定的安全范围,同时又能保证一定的数据可用性.

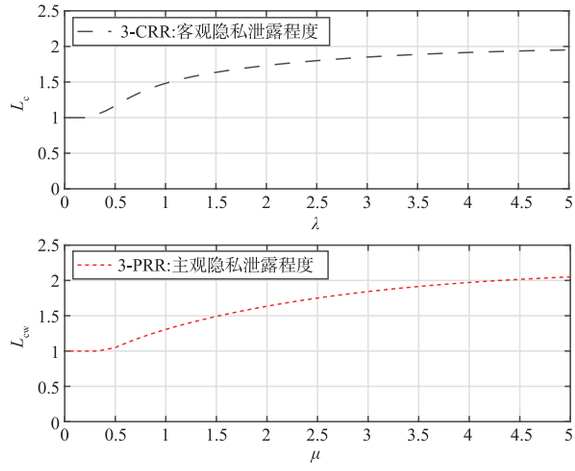


图2　3-CRR/3-PRR机制的客观/主观隐私泄露程度

图 3 给出了 3-CRR 模型在 CMAP 准则下各个敏感值的客观隐私泄露程度 L_{ci}. 可知,当 $\lambda\leq1.65$ 时(该临界值与具体的先验分布有关),隐私泄露程度由底到高依次为 x_1、x_2、x_3,表明敏感度最高的敏感值 x_1 的隐私保护程度最高,基本对其实现了完全的保护. 但是,当 $\lambda>1.65$ 时,三条曲线基本重合,说明此时 CRR 机制对三种敏感值客观上的隐私保护程度相同. 实际中,各个敏感值的隐私需求并不同,但 CRR 模型对所有敏感值的隐私保护程度相同,这是其不足之处.

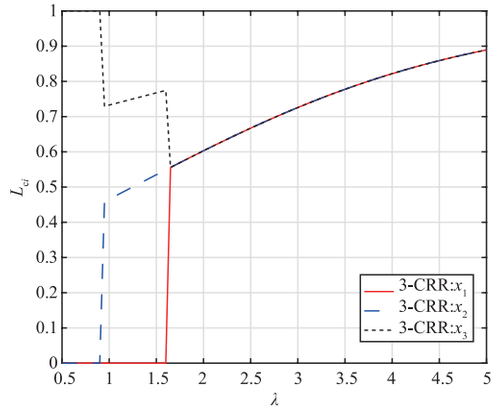


图3　CRR机制下各个敏感值的客观隐私泄露程度

图 4 给出了 3-PRR 模型在 PMAP 准则下各个敏感值的主观隐私泄露程度 L_{cwi}. 可知,三个敏感值的主观隐私泄露程度均随 μ 的增大而增大. 当 μ 一定时,x_1、x_2、x_3的主观隐私保护泄露程度依次增大,意味着相应的主观隐私保护程度依次增大,这也是本文 PRR 模型相比于 CRR 模型的优势所在. 换言之,PRR 机制对敏感

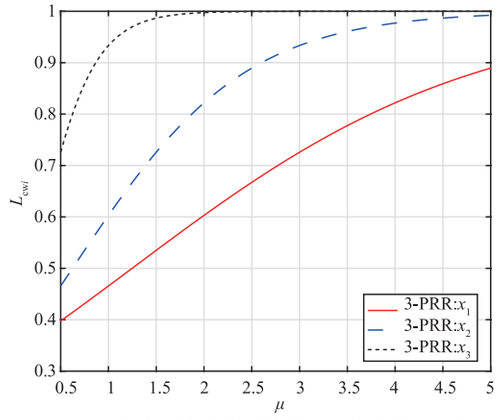


图4 PRR机制下各个敏感值的主观隐私泄露程度

度最高的敏感值 x_1 的隐私保护程度最高，对敏感度最低的敏感值 x_3 的隐私保护程度相对较低，这是因为其相应的隐私保护需求较低.

为了验证所提 PRR 机制在统计质量方面优于 CRR 机制，本小节设置两种模型下的主观隐私泄露程度一定. 图 5 给出 3-CRR 和 3-PRR 模型隐私分布估计误差的仿真值，并给出其理论值作为参考（蒙特卡洛仿真设置 10^4 次）. 可知，两种模型下隐私分布估计误差的仿真值与理论值基本一致，验证了本文理论分析的正确性. 另外，当主观隐私泄露程度 L_{cw} 一定时，CRR 模型隐私分布的估计误差均高于 PRR 模型，即 PRR 模型统计估计的准确性更高，性能更佳. 另一方面，两种模型下的估计误差均随着主观隐私泄露程度 L_{cw} 的增大而减小. 但是，实际中，受访者一般希望其主观隐私泄露程度越小越好，但如此会造成数据可用性降低，说明数据可用性与隐私之间是相互矛盾的. 实际中，应该根据数据质量和隐私需求进行良好的折中.

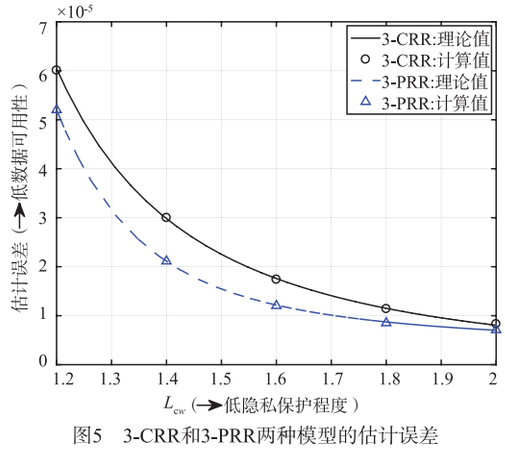


图5 3-CRR和3-PRR两种模型的估计误差

表 2 展示了五种主观隐私泄露程度下两种模型各自因此参数设置以及隐私分布估计的相对效率 R_e 的仿真值，并给出理论值作为参考（蒙特卡洛仿真设置 10^4 次）. 可知，当两种模型的主观隐私泄露程度相同时，R_e 的理论值和仿真值基本一致，验证了本文理论分析的正确性和有效性（与图 5 结论一致）. 另外，R_e 均满足：$R_e \geqslant 1$，意味着主观隐私泄露程度一定时，PRR 模型在统计数据可用性方面优于 CRR 模型，并且能够针对不同敏感值灵活调控参数 μ，实现个性化的隐私保护，这与 3.2 小节的理论分析一致. 再者，两种模型的主观隐私泄露程度均随着 λ 和 μ 的增大而增大，此时隐私数据的安全性在不断降低. 因此，实际中需要根据具体隐私保护需求设置合适的 λ 和 μ 值，进而将隐私泄露程度控制在合理的范围.

表 2 主观隐私泄露程度一定的条件下两种模型的隐私参数设置及其相对效率 R_e 的仿真值与理论值

主观隐私泄露程度 L_{cw}	CRR 参数：λ	PRR 参数：μ	相对效率 R_e	
			理论值	仿真值
1.2000	1.6369	0.7688	1.1612	1.1821
1.4000	2.3316	1.2367	1.4143	1.3002
1.6000	3.1161	1.8685	1.4457	1.4440
1.8000	4.1117	2.7536	1.3197	1.3201
2.0000	5.7205	4.3190	1.1612	1.1470

5 小结与展望

实际中，隐私信息并非同等重要，隐私需求也不尽相同. 而现有的 CRR 模型对所有的敏感值都执行相同的隐私保护操作，如此会造成某些低敏感信息过度保护，而某些高敏感信息欠缺保护. 基于此，本文考虑个性化的隐私需求，引入敏感值权重的概念，提出了一种面向多敏感值的 PRR 机制，实现了个性化隐私保护，具有重要的实际意义. 本文针对 CRR 和 PRR 两种机制下的隐私泄露程度和数据质量进行了充分的阐述与分析. 分析结果表明，在主观隐私泄露程度一定时，相比于 CRR 机制，PRR 机制所获得的统计估计的准确性较高. 另外，为了便于讨论，本文仿真将 PRR 模型的隐私参数设置为 $\mu_i = \mu_j = \mu$，这样只能对各敏感值的隐私保护程度进行粗粒度的调节. 因此，接下来的研究可对各敏感值设置合理的 μ_i 值，旨在对各敏感值进行灵活且细粒度的隐私保护调节，实现个性化精准的隐私保护，并在一定主观隐私泄露程度的限制下，通过优化设计，最大化统计数据的可用性.

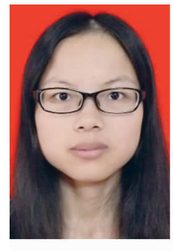
参考文献

[1] Warner S L. Randomized response: A survey technique for eliminating evasive answer bias[J]. Journal of the American Statistical Association, 1965, 60(309): 63 – 69.

[2] 罗永龙，黄刘生，荆巍巍，等. 一个保护私有信息的布尔关

联规则挖掘算法[J]. 电子学报,2005,33(5):133 - 136.
Luo Y L, Huang L S, Jing W W, et al. An algorithm for privacy-preserving boolean association rule mining[J]. Acta Electronica Sinica, 2005, 33(5): 133 - 136. (in Chinese)
[3] Hsieh S H, Lee S M, Tu S H. Randomized response techniques for a multi-level attribute using a single sensitive question[J]. Statistical Papers, 2018, 59(1): 291 - 306.
[4] Tian X, Taylor J. Selective inference with a randomized response[J]. The Annals of Statistics, 2018, 46(2): 679 - 710.
[5] 叶青青,孟小峰,朱敏杰,等. 本地化差分隐私研究综述[J]. 软件学报,2018,29(7):159 - 183.
Ye Q Q, Meng X F, Zhu M J, et al. Survey on local differential privacy[J]. Journal of Software, 2018, 29(7): 159 - 183. (in Chinese)
[6] Lin B C, Wu S H, Tsou Y T, et al. PPDCA: Privacy-preserving crowdsensing data collection and analysis with randomized response[A]. Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)[C]. Barcelona, Spain: IEEE, 2018. 1 - 6.
[7] Aoki S, Iwai M, Sezaki K. Privacy-aware community sensing using randomized response[A]. Proceedings of the 37th Annual Computer Software and Applications Conference Workshops[C]. Japan: IEEE, 2013. 127 - 132.
[8] Xiao X K, Tao Y F, Chen M H. Optimal random perturbation at multiple privacy levels[J]. Proceedings of the VLDB Endowment, 2009, 2(1): 814 - 825.
[9] Kairouz P, Oh S, Viswanath P. Extremal mechanisms for local differential privacy[A]. Advances in Neural Information Processing Systems (NIPS)[C]. Red Hook, NY, USA: Curran Associates, Inc, 2014. 2879 - 2887.
[10] Kairouz P, Bonawitz K, Ramage D. Discrete distribution estimation under local privacy[A]. Proceedings of the 33rd International Conference on Machine Learning[C]. New York, NY, USA: 2016. 2436 - 2444.
[11] Dwork C. Differential privacy[A]. Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)[C]. Venice, Italy: Springer, 2006. 1 - 12.
[12] Wang W N, Ying L, Zhang J S. A game-theoretic approach to quality control for collecting privacy-preserving data[A]. Proceedings of the 53rd Annual Allerton Conference on Communication, Control, and Computing[C]. Allerton House, UIUC, Illinois, USA: IEEE, 2016. 474 - 479.
[13] Kim J W, Kim D, Jang B. Application of local differential privacy to collection of indoor positioning data[J]. IEEE Access, 2018, 6: 4276 - 4286.
[14] Holohan N, Leith D J, Mason O. Optimal differentially private mechanisms for randomised response[J]. IEEE Transactions on Information Forensics and Security, 2017, 12(11): 2726 - 2735.
[15] Ye M, Barg A. Optimal schemes for discrete distribution estimation under local differential privacy[J]. IEEE Transactions on Information Theory, 2018, 64(8): 5662 - 5676.
[16] Xiao X, Tao Y. Personalized privacy preservation[A]. Proceedings of the ACM SIGMOD International Conference on Management of Data[C]. Chicago, IL, USA: ACM, 2006. 229 - 240.
[17] Braun C, Chatzikokolakis K, Palamidessi C. Quantitative notions of leakage for one-try attacks[J]. Electronic Notes in Theoretical Computer Science, 2009, 249: 75 - 91.

作者简介



宋海娜　女,1990 年出生,湖北襄阳人,现为北京邮电大学信息与通信工程学院博士研究生,主要从事无线通信、信息安全、隐私保护等相关研究.
E-mail:songhn_cqupt@163.com

罗　涛　男,1971 年出生,陕西宝鸡人,博士,现为北京邮电大学信息与通信工程学院教授,博士生导师,主要从事移动通信、认知无线电、车联网、机器学习和隐私保护等相关研究.
E-mail:tluo@bupt.edu.cn

韩新宇　男,1996 年出生,山东烟台人,现为北京邮电大学信息与通信工程学院硕士研究生,主要从事隐私保护等相关研究.
E-mail:hanxinyucat@126.com

李剑峰　男,1960 年出生,内蒙人,现为北京邮电大学信息与通信工程学院研究员,硕士生导师,主要研究方向为智慧医疗与数据分析.
E-mail:lijf@bupt.edu.cn