

隐私计算的学术内涵与研究趋势

李风华¹, 李晖², 牛犇¹, 邱卫东³

(1. 中国科学院信息工程研究所, 北京 100085; 2. 西安电子科技大学网络与信息安全学院, 陕西 西安 710126;
3. 上海交通大学网络空间安全学院, 上海 200240)

摘要: 笔者在国际上率先提出隐私计算的概念、定义及学术内涵, 并形成了较为成熟的理论与技术体系。为了持续推动隐私计算的学术研究和产业应用, 详细诠释了隐私计算的学术内涵, 包括如何理解全生命周期、延伸控制、隐私量化与映射、脱敏效果评估, 为什么要做迭代按需脱敏、为什么要研究隐私计算语言以及自存证在泛在共享中的作用等, 并对一些被曲解的学术概念予以澄清; 给出了隐私计算九大方面 37 个研究点, 以及数据安全八大方面 40 个研究点, 并从 18 个维度将隐私计算与数据安全等技术进行了全面对比, 以帮助读者更好地理解隐私计算的研究范畴, 正确区分隐私计算与数据安全。

关键词: 隐私计算; 延伸控制; 动态度量; 迭代按需脱敏; 保护效果评估

中图分类号: TP309

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2022078

Academic connotation and research trends of privacy computing

LI Fenghua¹, LI Hui², NIU Ben¹, QIU Weidong³

1. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100085, China

2. School of Cyber Engineering, Xidian University, Xi'an 710126, China

3. School of Cyber Science and Engineering, Shanghai Jiaotong University, Shanghai 200240, China

Abstract: The authors of this paper first proposed the concept, definition and academic connotation of privacy computing, and formed a relatively mature theoretical and technical system accordingly. In order to continuously promote the academic research and industrial application of privacy computing, the academic connotation of privacy computing was elaborated, including how to understand the full-life cycle, extended control, privacy quantification and mapping, desensitization effect evaluation. Besides, the necessity of iterative on-demand desensitization and the motivation to study the language of privacy computing were presented. The role of audit log self-storage in ubiquitous sharing was explained and some distorted academic concepts were also clarified. Moreover, 37 research points in 9 aspects of privacy computing and 40 research points in 8 aspects of data security were given. It helps to better understand the

收稿日期: 2022-11-09; 修回日期: 2022-11-30

通信作者: 李晖, lihui@mail.xidian.edu.cn

基金项目: 国家重点研发计划 (2021YFB3101301); 国家自然科学基金 (61932015)

Foundation Items: The National Key R&D Program of China (2021YFB3101301), The National Natural Science Foundation of China (61932015)

引用格式: 李风华, 李晖, 牛犇, 等. 隐私计算的学术内涵与研究趋势[J]. 网络与信息安全学报, 2022, 8(6): 1-8.

Citation Format: LI F H, LI H, NIU B, et al. Academic connotation and research trends of privacy computing[J]. Chinese Journal of Network and Information Security, 2022, 8(6): 1-8.

research scope of privacy computing and correctly distinguish between privacy computing and data security.

Keywords: privacy computing, extended control, dynamic measurement, iterative on-demand desensitization, protection effect evaluation

0 引言

信息时代之前, 由于信息在小范围内传播或在封闭信息系统内使用, 隐私泄露并没有成为大众关注的焦点。然而, 移动通信、网络和信息等技术的迭代演进推动人类从 IT (information technology) 时代进入 DT (data technology) 时代, DT 时代的核心是面向数据流通的信息广泛传播和受控共享, 共享数据中包含大量个人隐私信息, 因此隐私信息的有效保护是数据有序共享、释放数据价值的前提条件。当前隐私保护面临的问题与日俱增, 如 App 频繁超范围采集个人信息, 后台信息系统中的隐私信息越权使用、大数据杀熟、个人画像结果滥用、个人信息过度留存, 生态圈之间信息共享缺乏延伸控制来抑制非授权共享, 缺乏抗隐私挖掘的迭代按需脱敏, 多副本留存和保护短板效应凸显, 删除权无法保障等。

各国对隐私保护的重视程度日益提高。欧盟颁布的《通用数据保护条例 (GDPR, general data protection regulation)》强化了对知情权、被遗忘权、删除权的要求; 我国颁布的《中华人民共和国民法典》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等将隐私保护的要求提高到了法律和合规审查的高度。

2015 年, 笔者在国际上率先提出了隐私计算的概念、定义及学术内涵, 并于 2016 年在《通信学报》上正式发表^[1]。自隐私计算被提出至今, 已形成较为成熟的理论与技术体系, 得到学术界和工业界的广泛关注和认可。然而, 一些机构并没有真正理解隐私计算的学术内涵, 作为原创者, 有必要对隐私计算的真正学术内涵做进一步澄清, 避免在隐私计算应用中误导企业界、投资界、主管部门和研究人员, 以进一步促进隐私计算生态的健康发展。

1 隐私计算的学术内涵

(1) 隐私计算的定义

隐私计算是面向隐私信息全生命周期保护的

计算理论和方法, 是隐私信息的所有权、管理权和使用权分离时隐私度量、隐私泄露代价、隐私保护与隐私分析复杂性的可计算模型与公理化系统。隐私计算具体是指在处理视频、音频、图像、图形、文字、数值、泛在网络行为信息流等信息时, 对所涉及的隐私信息进行描述、度量、评价和融合等操作, 形成一套符号化、公式化且具有量化评价标准的隐私计算理论、算法及应用技术, 支持多系统融合的隐私信息保护。隐私计算涵盖了信息搜集者、发布者和使用者在信息产生、感知、发布、传播、存储、处理、使用、销毁等全生命周期过程的所有计算操作, 并包含支持海量用户、高并发、高效能隐私保护的系统设计理论与架构。

如图 1 所示, 隐私计算框架在隐私信息全生命周期的各个环节中建立了应用场景、保护需求与计算模型等之间的映射关系, 并基于场景描述和保护需求, 适应性地选择相应环节的计算方法实现相应的计算功能。隐私计算框架包括隐私信息抽取、场景描述、隐私控制、隐私操作、隐私效果评估等 5 个步骤。隐私信息抽取根据多模态文档的格式、语义等抽取隐私信息, 并得到隐私信息向量; 场景描述根据各隐私信息分量的类型、语义等, 对应用场景进行定义与抽象; 隐私控制是根据主体意愿、使用者的保护能力决定对隐私信息分量的操作控制, 并生成传播控制操作集合; 隐私操作面向各隐私信息分量选取其对应的隐私保护算法或信息处理动作; 隐私保护效果评估根据相关评价准则, 确定所选择隐私保护方案的隐私保护效果。效果评估还为隐私控制方案的迭代优化提供支撑, 如效果达不到预期要求, 则分别从场景描述、重新调整控制策略、重新定义操作等环节进行反馈迭代, 直至达到期望的保护效果。

(2) 如何理解全生命周期

隐私信息在单一信息系统内或者一个使用者控制范围内的全生命周期不是真正的全生命周期, 同一隐私信息跨系统流转, 其在所有流经的信息系统或者在所有使用者控制范围内的全生命

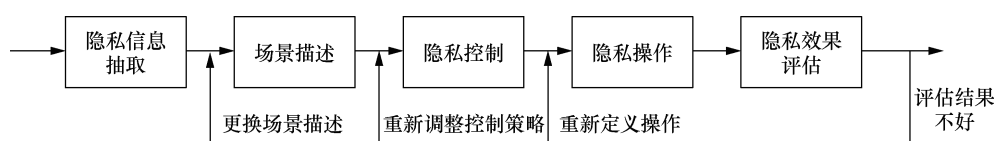


图1 隐私计算框架

Figure 1 The framework of privacy computing

周期才是隐私计算内涵中重点关注的全生命周期。

（3）如何理解延伸控制

延伸控制是指泛在共享环境下隐私信息在跨域受控交换过程中全生命周期各环节隐私操作的迭代控制、控制策略的动态调整、控制策略的可控传递、控制策略执行的可信审计等。延伸控制根据信息主体或数据提供方的控制意图、当前使用者控制约束和数据接收者保护能力生成控制策略，使其随信息流转过程同步传递且不可分割，并根据使用场景、延伸控制要求不断动态变化并向前可信可控传递，从而实现迭代控制直至数据的所有副本销毁为止。

延伸控制机制包括控制意图、控制策略、隐私操作等。控制意图由信息主体设置、场景适应的自调整、信息来源的迭代传递等方式多源获取，通常由信息所有者、搜集/发布者、使用者实施；控制策略依据运行环境信息、接收者保护能力、上级传递的主体控制意图、当前延伸控制信息等要素迭代生成或调整，在全生命周期过程中由不同使用者执行延伸控制；依据应用场景和延伸控制信息，优选脱敏和删除算法，实现 QoS 与差异化脱敏、删除的效果平衡。

（4）如何理解隐私量化与映射

在隐私信息的形式化描述^[3]中，隐私属性分量用于量化隐私信息分量及分量组合的敏感度或者期望保护程度。隐私属性分量的划分及其量化需要从本质特征上研究学术分类，并研究场景适应的分级方法。例如，GPS 数据、门牌号、邮政编码、小区名称、移动基站标识本质上都属于位置隐私，但敏感度分级各不相同。GPS 数据、门牌号、移动基站标识精度和敏感度高；而邮政编码、小区名称覆盖的位置范围更大，精度和敏感度较低。

隐私信息跨系统交换和传播时，不同信息系统的隐私属性分量的量化标准可能不同，因此需

要在不同系统之间建立隐私量化标准的映射关系，使得相同隐私信息在不同的隐私信息系统中保护效果具有一致性。

（5）如何理解脱敏效果评估

在隐私信息泛在传播的迭代按需脱敏过程中，对隐私信息的脱敏效果评估主要体现在 3 个方面。一是用于信息发布的单个隐私信息的脱敏效果评估，如果脱敏效果评估达到预定要求，就可以发布；二是用于抗大数据隐私挖掘的脱敏效果评估，具体是对同一主体关联的所有当前信息和历史信息进行大数据隐私挖掘分析，评价不同时期、不同算法脱敏后的信息是否能抗大数据隐私挖掘；三是通过对不同主体的同一类信息进行脱敏效果评估，主要用于对算法的脱敏能力、算法选择控制的正确性进行评估，以支撑算法的迭代修正，以及算法管理方案的迭代修正。

（6）为什么要做迭代按需脱敏

隐私信息在每一次跨系统共享过程中，需要根据所处时空场景、隐私信息中的主体关系、隐私脱敏需求、传播过程中接收方的保护能力等要素差异，以及不同保护算法的特征、适用范围、保护效果，对隐私信息实现场景适应的差异化保护。因此，同一隐私信息的脱敏不是在首次传播时一次性脱敏就能解决问题，而是需要在每次传播的过程中做迭代按需脱敏。例如，在导航应用场景下的服务过程中，信息系统需要相对精准的起始地、目的地和轨迹信息；在服务结束后，隐私信息留存时应该做泛化脱敏操作；留存信息在后台转移到其他信息系统进行利用时，还要做进一步的泛化脱敏操作；企业在跨生态圈共享导航信息时，则要再次做进一步的迭代泛化脱敏操作。这个典型场景的脱敏应用需求可抽象为迭代按需脱敏。

（7）为什么要研究隐私计算语言

隐私计算语言（PCL，privacy computing

language) 用于高效简洁地形式化描述隐私信息定义、脱敏、控制等操作, 包括隐私定义语言、隐私操作语言、隐私控制语言等。隐私计算语言能够便捷地支持隐私信息跨平台交换与延伸控制; 还可对开发者屏蔽复杂的理论细节, 降低程序开发者的技术门槛, 提升系统开发效率, 从而快速构建隐私保护信息系统。隐私计算语言能够准确地描述隐私计算各个环节的操作, 便于隐私计算理论的准确表达, 易于学者之间交流以及开发者理解, 确保没有二义性。

(8) 自存证在泛在共享中的作用

在隐私信息泛在共享过程中, 隐私信息的延伸控制策略随隐私信息一起可信可控传递, 各种主体对隐私信息的各类操作应该进行不可篡改的存证记录, 并对操作与延伸控制策略的一致性进行及时判定及判定结果存证。这样, 当违反延伸控制策略的行为发生时, 可以实现泛在传播过程中随遇、实时的违规、侵权判定, 并支撑溯源取证。

2 如何区分隐私计算与数据安全

数据安全主要指保证数据的机密性、完整性、不可否认性等, 确保被保护的数据具有可恢复性, 即强调信息的无损性, 大多使用密码学、访问控制等技术实施。隐私保护可分为两种情况: 一是保障信息不受损失的前提下隐私不被非授权者获取及处理, 称之为隐私防护, 即防护是在单一信任域中确保信息不泄露; 二是在隐私交换过程中信息接收者得到隐私的信息量小于信息发送方同一隐私的信息量, 接收方不能完全获得发送方的全部信息, 称之为隐私脱敏。

单一有界信息系统中单一环节的数据安全和隐私防护技术有高度的原理相近性和使用互换性, 而用于跨系统交换的隐私脱敏与数据安全技术则有明显的差异。此外, 数据脱敏与隐私脱敏也存在差异, 数据脱敏通常针对国家秘密和企业的商业秘密而言, 不能交换的敏感数据以删除为主要手段, 提供部分数据子集, 不以提供假数据方式进行数据脱敏; 隐私脱敏则针对个人信息, 数据脱敏的方法可以用于隐私脱敏, 但隐私脱敏还有其他替代、泛化、加扰等方式, 使得脱敏后的信息存在失真的情况。

在现实社会中, 目前对数据安全的保护力度大于对隐私的保护力度。数据安全主要针对国家和企业部门, 数据使用部门的管理制度严格, 工作人员数据安全的自觉性强, 数据泄露很多情况下要承担刑事责任。隐私属于个人信息, 个人信息的泄露大多为民事纠纷, 公众保护意识普遍不足, 法律处罚措施相对较轻, 企业泄露个人隐私也以罚款为主。因此, 对隐私信息如果不脱敏, 在泛在共享的环境下隐私无从保护, 故数据安全和隐私防护的相关技术不适应跨系统交换的隐私保护。

值得强调的是, 密文计算、安全多方计算、机密计算、可信计算、访问控制等属于数据安全范畴, 密文计算、机密计算、可信计算等保护计算环节的数据安全, 安全多方计算可以保护交换环节的数据安全, 它们可以用于单一信息系统、局部环节的隐私防护, 隐私信息并没有被脱敏, 具有可逆性, 因此某一系统的某一环节的隐私泄露会导致其他系统保护的失效, 具有“一损俱损”的短板效应缺陷。在此, 对学术界、产业界容易混淆的若干概念澄清如下。

(1) 密文计算

密文计算是指计算过程中的数据不被计算参与方所获取, 主要用于外包计算场景。同态加密是密文计算的代表性技术, 是在事先确定转换规则的前提下, 所有参与运算的明文数据使用该规则转换为密文, 在密文空间中进行特定形式的运算并得到密文运算的结果, 再通过相应的转换规则转换为明文运算结果, 该结果与直接对明文运算得到的结果一致。本质上, 密文计算参与运算的明文及明文结果都没有信息损失, 因此密文计算用于隐私保护时, 仅能解决计算过程中的隐私防护, 不适用于信息泛在共享的隐私防护。

(2) 安全多方计算

在事先确定参与方数目范围及交互协议的前提下, 所有参与方以密文形式交互参与运算的信息并完成预先约定的运算任务, 所有参与方都能得到运算结果的明文, 但不能得到相互交互参与运算的明文信息。安全多方计算在有恶意参与者的情况下, 诚实参与者仍能得到正确的结果, 不泄露参与方的原始信息。现阶段, 参与方的数目一般是两方和三方。秘密共享、不经意传输、同

态加密等是构造安全多方计算的重要机制。本质上, 安全多方计算没有信息损失, 主要用于计算环节, 原始参与方的信息不泄露, 但运算结果具有隐私防护的等价效果。因此, 安全多方计算用于隐私保护时, 仅能解决计算过程中的隐私防护, 不适用于信息泛在共享的隐私防护。

(3) 联邦学习

人工智能模型训练时在全量完整数据上训练才能达到最佳效果，但由于收集数据量受限或者全量数据训练运算量大，通常在有限的样本数据上训练，而样本数据与全量完整数据的特征偏离程度决定了训练的效果。

将数据集中起来进行全局训练涉及两个问题：① 算力集中导致投资巨大且算力可能得不到充分利用；② 数据集中导致数据出域，在安全保护和使用控制没有得到有效解决的情况下数据所有者不愿意分享数据。联邦学习是在这两个背景下提出的一种分布式模型训练架构，首先，可以充分利用分布式算力减少最终模型需求方的算力投入；其次，通过本地样本数据的局部训练，以及训练结果的迭代聚合，在牺牲少量训练结果质量的条件下，迎合数据不愿意出域共享的现状，间接地减少数据泄露的机会。然而，分布模式模型训练仍然需要交换中间结果和模型参数，存在数据泄露的问题，当然也包含隐私泄露问题。

综上所述,联邦学习的本质是一个分布式模型训练架构,因不进行数据集中训练,间接地减少了数据泄露,但交换的信息仍然存在数据泄露,因此,从学术本质上联邦学习属于人工智能的范畴,不属于数据安全和隐私保护学科范畴,也不属于隐私计算的范畴。

(4) 隐私增强计算

Gartner 发布的 2021 年前沿科技战略趋势^[5]中提到了隐私增强计算 (privacy enhancing computation), 但笔者认为其命名并不妥当, 隐私保护的的根本目的是不让隐私本身增强, 但“隐私增强计算”的词义理解为隐私的增强计算技术, 相应地属于挖掘隐私信息的技术领域, 即让隐私特征信息更加凸显出来。若要表达用于隐私保护的技术, “隐私保护能力增强计算 (capability enhancing computation for privacy preservation)”的计算技术更为恰当。此外, 目前学术界研究的

“隐私增强计算”仍是针对单一环节、单一场景的保护方案，属于传统的隐私保护技术，只是零散的技术点，没有形成体系，更不能替代隐私计算。

(5) 广义隐私计算

有些研究机构将笔者对隐私计算定义中的“搜集者、发布者和使用者”改为“所有者、转发者和接收者”，并称其为广义隐私计算，实际上没有正确理解隐私计算的内涵，不但不是广义隐私计算，反而是更为狭义地理解隐私计算；将“搜集者”改为“所有者”是概念错误，信息所有者对隐私具有任意的处置权；而笔者对隐私计算定义中的“搜集者”包括搜索服务商（搜索引擎）、信息收集加工服务商，它们获取信息时应保障信息主体的知情权。将“发布者”改为“转发者”是没有认识到平台发布隐私信息需要承担隐私保护的责任，“转发”的语义通常指传输，而传输不承担内容侵权的法律责任，发布平台则要承担法律责任。将“使用者”改为“接收者”是缩小了范围，使用者包括数据接收和数据使用处理两个方面，而数据接收在语义上只涵盖数据留存，只涉及数据删除权。数据使用则涉及隐私信息的加工、分析、交易等广泛的行为，对隐私信息侵权体现在更广义的范围，对信息主体造成经济损失和社会影响。

(6) 隐私计算与数据安全学术内涵对比

隐私计算和数据安全的学术内涵如图 2 和图 3 所示。



图2 隐私计算的学术内涵
Figure 2 Academic connotation of privacy computing

隐私计算学术内涵具体分为九大方向：隐私计算框架、延伸控制、隐私感知、动态度量、迭代按需脱敏、保护效果评估、多副本完备删除、溯源取证、隐私计算语言。在此基础上，可分为

37 个研究点: 控制迭代传递、操作约束条件、保护能力量化映射、隐私特征提取、隐私分量、隐私属性向量、场景识别、泄露风险评估、算法通用框架、脱敏原语、组合规则、差分隐私、本地化差分、个性化差分、隐私预算、 k -匿名、 l -多样性、 t -邻近性、去标识、混淆、加扰、置乱、泛化、替换、抑制、数据合成、可用性、算法复杂性、脱敏效果评估、删除效果评估、隐私挖掘、操作自存证、权属转移、侵权行为判定、侵权取证、证据交叉认证、线索挖掘。



图 3 数据安全的学术内涵

Figure 3 Academic connotation of data security

数据安全的学术内涵具体分为八大方向: 机密计算、可信计算、密文计算、安全多方计算、

访问控制、数据灾备、数据治理、身份认证。在此基础上, 可细分为 40 个研究点: 可信执行环境、同态加密、可搜索加密、可交换加密、性质保持加密、远程验证、完整性可信度量、可信迁移、不经意传输、秘密共享、门限密码、混淆电路、零知识证明、承诺协议、隐私求交、不经意随机预言机、权限管理、自主访问控制、强制访问控制、基于角色的访问控制、基于属性加密的访问控制、基于行为的访问控制、网络空间的访问控制、身份鉴别、基于身份加密、数字签名、多因子认证、交叉认证、异地容灾、安全存储、纠删码、安全删除、数据清洗、分类分级、合规性检测、消息鉴别、数据确权、追踪溯源、数据审计、流转管控。

(7) 隐私计算与数据安全等相关解决方案对比

为了明确科学地界定隐私计算的学术内涵, 刻画不同解决方案之间学术内涵的差异, 帮助理解和判断什么样的技术才是真正的隐私计算, 笔者提出了 18 个维度的对比标准。隐私计算与传统隐私保护、数据安全等方案的对比如表 1 所示。

相较于隐私保护的傳統方法, 以及密文计算、机密计算、可信计算、安全多方计算、联邦学习和访问控制技术, 隐私计算的计算开销和通信开销低, 支持全生命周期的保护、延伸控制、迭代脱敏, 既可适用于有界系统不出域也可适用于无

表 1 隐私计算与传统隐私保护、数据安全等方案的对比

Table 1 Comparison of privacy computing traditional privacy protection and data security

保护技术 对比项	计算 (使用)	计算 开销	通信 开销	需要专 用硬件	交换	隐私 防护	脱敏	迭代 脱敏	有界 系统 (不出域)	无界 系统 (出域)	全生命 周期	延伸 控制	差异 保护	量化 与映射	评估	取证	用户数	计算 粒度
隐私计算	√	低	低	×	√	√	√	√	√	√	√	√	√	√	√	√	海量	对象级
传统方法 (k -匿名、 差分等)	√	中	低	×	√	×	√	×	√	√	×	×	×	×	×	×	海量	单等级
密文计算 (同态加密)	√	高	高	×	交换结果, 有隐私泄露	√	×	×	√	×	×	×	×	×	×	×	少量	单等级
安全 多方计算	√	高	高	×	√	√	×	×	√	×	×	×	×	×	×	×	2~3	单等级
联邦学习	√	中	高	×	交换结果及 参数, 有隐 私泄露	√	×	×	√	×	×	×	×	×	×	×	少量	单等级
机密计算	√	低	低	√	√	√	×	×	√	×	×	×	×	×	×	×	少量	单等级
可信计算	√	低	低	√	√	√	×	×	√	×	×	×	×	×	×	×	少量	单等级
访问控制	√	低	低	×	√	√	×	×	√	×	√	部分支 持	×	部分支 持	×	×	海量	单等级

界系统出域的场景, 同时支持同一隐私信息在不同场景、不同约束条件下的差异化脱敏保护, 以及保护能力在不同信息系统中的量化映射, 并包含对隐私侵权行为的判定和取证溯源, 可适用于海量用户的细粒度隐私保护。

3 未来发展趋势与应用

(1) 隐私计算的基础理论

从隐私感知与动态度量、隐私保护算法、隐私保护效果评估、隐私信息延伸控制、隐私侵权行为存证和溯源等方面进一步研究并完善隐私计算框架及其数学基础, 细化全生命周期不同环节间的关联机制、操作控制及其传递, 研究全流程隐私信息的流转控制模型、脱敏延伸控制模型、删除延伸控制模型等内容, 研究业务服务与隐私计算深度融合的高效隐私信息保护系统技术架构, 并针对典型应用场景的隐私信息保护提供解决方案。

(2) 隐私感知与动态度量

从隐私信息知识表示模型、学术分类、场景分级、原子抽象建模、特征分析与隐私分量抽取、智能感知、隐私分量关联关系挖掘等角度入手, 研究隐私分量与场景关联模型、隐私分量量化与动态调整、隐私分量组合与动态度量、隐私度量的量化指标体系等内容, 解决时空差异和主体动态下隐私动态交换的精准度量问题, 支撑泛在共享下隐私信息交换控制与按需脱敏。

(3) 隐私保护算法

在不同环节、不同场景下研究基于不同数学基础的隐私脱敏原语及其等价或映射关系; 研究隐私保护算法通用框架与设计准则、算法选择和优化组合设计、算法前后台任务动态调度、算法保护能力量化指标之间的等价关系等内容, 支撑泛在共享下隐私信息跨系统交换控制、隐私信息保护系统的柔性重构和隐私脱敏功能的动态编排、隐私保护算法的设计与能力评估。

(4) 隐私保护效果评估

从延伸控制性、可逆性、复杂性、偏差性、信息损失性、合规性等维度入手, 研究保护算法及其组合的效果评估量化指标、量化指标的关联关系和动态权值、效果评估系统的计算模型、隐

私挖掘等内容, 构建效果评估指标体系, 支撑隐私保护的效果反馈和隐私保护方案的迭代优化、隐私信息保护系统能力评估。

(5) 隐私侵权行为判定与溯源

以隐私侵权行为判决规则与约束表示为基础, 研究延伸控制策略判定、全流程隐私侵权线索存证、侵权行为的场景与内容的存证、审计信息可信存证、隐私操作行为和策略声明的一致性与合规性检测、侵权事件识别与判定、侵权场景构建与行为重构等内容, 支撑隐私信息受控共享。

(6) 隐私信息的完备删除

从传播路径发现、通知与确认拓扑生成、删除方案选择、删除操作行为可验证等方面, 研究自动/指定删除机制、删除粒度协商机制、信息多副本检索、删除粒度控制、自主/自动删除触发、删除效果远程验证机制、删除不可恢复性评估、删除操作行为审计等内容, 支撑删除可信验证。

4 结束语

隐私计算因需而生, 其核心是泛在环境下隐私信息全生命周期的保护, 其灵魂是迭代延伸控制和按需保护。从严谨的学术定义角度来讲, 如果一种技术架构不存在信息泄露, 就不需要利用数据安全技术来解决信息泄露、利用隐私保护技术来解决隐私泄露。针对有隐私泄露的应用场景, 能解决隐私泄露问题的技术才是隐私保护技术。隐私计算是解决数据泛在共享过程中隐私泄露问题的完整理论框架和技术体系, 当然针对不同场景还需要不断具体细化、不断丰富, 需要广大学者共同努力。作为隐私计算的提出者, 怀抱历史责任感来写这篇文章, 希望能正本清源, 促进大家正确理解隐私计算的学术内涵, 一起为数字经济发展保驾护航。

参考文献:

- [1] 李风华, 李晖, 贾焰, 等. 隐私计算研究范畴及发展趋势[J]. 通信学报, 2016, 37(4): 1-11.
LI F H, LI H, JIA Y, et al. Privacy computing: concept, connotation and its research trend[J]. Journal on Communications, 2016, 37(4): 1-11.

- [2] LI F, HUI L, NIU B, et al. Privacy computing: concept, computing framework, and future development trends[J]. Engineering, 2019, 5(6): 14.
- [3] 李凤华, 李晖, 牛犇. 隐私计算理论与技术[M]. 北京: 人民邮电出版社, 2021.
- LI F H, LI H, NIU B. Privacy computing theory and technology[M]. Beijing: Posts & Telecom Press, 2021.
- [4] 尤为. 专访李凤华: 隐私数据共享和泄露间的矛盾永恒存在, 隐私计算必将越来越成熟[EB].
- YOU W. Exclusive interview with Li Fenghua: the contradiction between private data sharing and disclosure is eternal, and private computing will become more and more mature[EB].
- [5] Gartner top strategic technology trends for 2021[EB].

[作者简介]



李凤华 (1966—), 男, 湖北浠水人, 博士, 中国科学院信息工程研究所研究员、博士生导师, 主要研究方向为网络与系统安全、隐私计算、密码应用。



李晖 (1968—), 男, 河南灵宝人, 博士, 西安电子科技大学教授、博士生导师, 主要研究方向为网络与信息安全、隐私计算、信息论。



牛犇 (1984—), 男, 陕西西安人, 博士, 中国科学院信息工程研究所副研究员、博士生导师, 主要研究方向为隐私计算、网络安全防护。



邱卫东 (1973—), 男, 江西修水人, 上海交通大学教授、博士生导师, 主要研究方向为计算机取证、密码分析、人工智能安全、大数据隐私保护。