

面向互联网数据互操作的授权技术综述

李颖^{1,2}, 李晓东^{1,3}, 费子郁^{1,2}, 彭博韬¹

1. 中国科学院计算技术研究所, 北京 100190;

2. 中国科学院大学, 北京 100190;

3. 伏羲智库互联网研究院, 北京 100192

摘要

互联网数据互操作可实现数据跨域互联和交换, 其中授权是保障数据权属和安全的关键技术之一。数据跨域流通的复杂性对授权技术提出了更高的可用和可信要求, 亟须面向互联网数据互操作展开授权技术研究。首先, 简述授权技术的基础概念和知识, 分析互联网数据互操作对授权提出的技术要求; 其次, 从信任机制、权限模型、策略管理3个方面分析现有授权技术能力; 最后, 针对现有授权技术能力的不足, 提出一些潜在的研究问题, 旨在为后续研究提供思路和参考。

关键词

数据互操作; 跨域数据交换; 授权协议; 信任机制

中图分类号: TN92

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2025008

Survey on data interoperating-oriented authorization technology

LI Ying^{1,2}, LI Xiaodong^{1,3}, FEI Ziyu^{1,2}, PENG Botao¹

1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

2. University of Chinese Academy of Sciences, Beijing 100190, China

3. FUXI Institution, Beijing 100192, China

Abstract

Data interoperating over Internet enables cross-domain connectivity and exchange of data, where authorization is a basic mechanism to protect data ownership and ensure data security. The complexity of cross-domain data exchange poses higher requirements for authorization technology in terms of utility and credibility, thus, there is an urgent need to carry out research on authorization technology for data interoperating-oriented. Firstly, the basic concepts and knowledge of authorization technology were briefly introduced, and the technical requirements of data interoperating-oriented for authorization were analyzed. Secondly, the existing authorization technical capabilities were analyzed from there aspects: trust mechanism, permission model, and policy management. Finally, in response to the limitations of existing authorization technology capabilities, some potential research questions were proposed, aiming to provide ideas and references for follow-up research.

Key words

data interoperating, cross-domain data exchange, authorization protocol, trust mechanism

0 引言

随着数字经济的快速发展,数据要素的流通需求日益凸显,数据要素亟须在不同的信任域之间开放使用,以充分发挥数据的价值。作为数字化数据传输和交换的载体,互联网进入了价值互联网阶段,互联互通的需求也从主机、网站发展到了更细的数据粒度。目前数据主要集中在互联网服务平台。随着《通用数据保护条例》(GDPR)、《中华人民共和国数据安全法》等一系列法律法规相继发布,数据安全合规压力施加到互联网服务平台。数据与应用的解耦和开放交换成为趋势,如何在尊重数据权属的前提下完成安全可信的数据互操作成为当今互联网发展的重要课题^[1-2]。

授权是授予进程(或用户)权限的过程^[3],是数据互操作中数据权属保护和交换的关键一环^[4]。授权通常采用特定的技术和协议实现(如开放授权协议^[5]),其统称为授权技术。这些技术面向特定范围的数据交换设计,具有参与角色少、管理方式单一的特点。然而,互联网数据互操作中数据开放使用,服务场景更复杂,权限要求也更多样^[6]。相比于专注特定范围内数据交换的传统授权,面向互联网数据互操作的授权对信任机制、权限模型和服务效率提出了更高的要求。

近年来,授权技术在信任机制、权限模型和服务效率等方面不断发展。一方面,在数据互操作技术领域,已有研究基于数字对象架构(digital object architecture, DOA)^[7]、社交关联数据(social linked data, SoLiD)项目^[8]以及数据基础设施

Gaia-X^[9]等不同数据互操作模型,提出了不同的授权解决方案;另一方面,授权的复杂性和可信性等问题受到了不同领域(如信息中心网络^[10]、以社区为中心的协作系统^[11]等)的关注,这些相关领域的授权技术研究具备适应互联网数据互操作场景的部分特点,对于面向互联网数据互操作的授权技术的研究和实现具有重要的借鉴意义。

目前,国内外研究尚缺乏对面向互联网数据互操作的授权技术的系统性讨论。尽管面向地理数据的开放地理数据互操作规范(open geodata interoperation specification, OpenGIS)^[12]、面向科学数据的FAIR(可发现(findable)、可访问(accessible)、可互操作(interoperable)、可重用(reusable))原则^[13]、面向大数据的互操作性框架(NIST big data interoperability framework, NBDIF)^[14]等专用领域的标准、原则和框架已形成,但其普遍专注于数据模型和接口规范的统一,尚未对授权技术进行深入探讨。因此,本文旨在面向互联网数据互操作的授权技术进行调查和研究,分析互联网数据互操作所提出的对授权技术的要求,并梳理和总结当前授权技术的研究现状,分析这些技术对互联网数据互操作的支持和不足,进一步探讨未来需要主要关注的研究问题。本文的主要贡献如下。

- 基于文献研究和场景分析,在该领域首次探讨了互联网数据互操作需求下的授权服务特点和技术要求,总结了授权技术体系。

- 概述了互联网数据互操作中的授权现状,结合互联网数据互操作对授权技术提出的要求,从信任机制、权限模型、策

略管理 3 个方面整理和归纳了现有授权技术的能力。

• 根据当前研究现状，探讨了面向互联网数据互操作的授权技术的未来研究方向。

本文的研究旨在填补互联网数据互操作领域中授权技术系统性研究的空白，提供对互联网数据互操作中授权技术要求、授权现状以及现有技术能力的更清晰认识，以促进该领域的进一步发展。

1 研究背景

1.1 授权技术概述

授权是授予进程（或用户）权限的过程^[3]，涉及资源管理者、授权服务、资源服务、资源使用者 4 个角色。资源管理者通过授权服务执行授权，授予进程或用户（即资源使用者）权限。资源服务通过授权服务验证使用者被授予的权限，以进一步控制资源使用者对资源的操作。授权过程

定义如下。假设存在资源 X 及其资源服务，由资源管理者 M 执行权限管理。当资源管理者 M 期望授予资源使用者 U 操作资源 X 的权限 P 时，资源管理者 M 通过与授权服务交互产生权属信息 I，资源使用者 U 对资源 X 的权限 P 由三元组 <U, X, P> 表示。权属信息 I 由授权服务存储，用于资源服务查询验证。上述资源管理者 M、资源使用者 U、资源 X 允许为一个或多个，权属信息三元组的语法由权限模型决定。

授权服务采用授权技术实现授权管理、存储和验证，与身份认证、访问控制技术紧密相关^[15-16]。其中，身份认证用于识别和验证用户身份，是授权的基础。访问控制是依据授权技术输出的权属信息执行对资源访问的控制过程，是授权的后置环节。授权服务的构成及相关技术如图 1 所示。

授权技术的发展主要分为本地授权和跨域授权两个阶段。

(1) 本地授权

实现同一信任域内的文件及系统权限管理，专注于权限模型的研究。权限模型最初包括适用于操作系统文件管理的自主访问控制（discretionary access control, DAC）^[17]和强制访问控制（mandatory access control, MAC）^[18]。随着计算机系统应用场景的增多，授权需求趋于复杂，权限模型出现基于角色的访问控制（role-based access control, RBAC）^[19]、基于属性的访问控制（attribute-based access control, ABAC）^[20]等方式。不同权限模型的主要区别在于采用不同的语法规则描述权属信息三元组，从用户组、角色、属性等不同粒度对用户进行权限管理，采用不同的权限集定义权限，在该阶段，授权通过对访问控制列表的读写完成。

(2) 跨域授权

解决来自不同信任域的资源使用者的

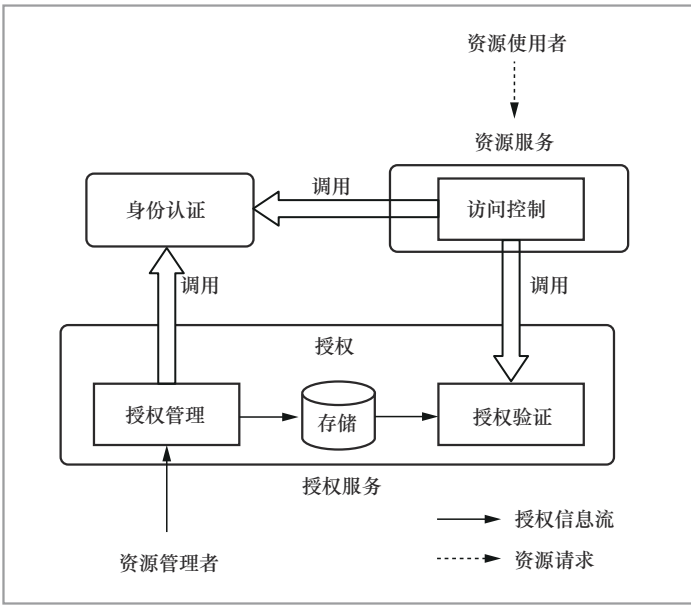


图1 授权服务的构成及相关技术

权限授予问题，专注于多个系统间的交互协议设计，包括简单公钥基础设施（simple public key infrastructure, SPKI）、简单分布式安全基础设施（simple distributed security infrastructure, SDSI）^[21-22]、开放授权（open authorization 2.0, OAuth2）^[5]、用户管理访问（user-managed access, UMA）^[23]等。不同授权协议适用于不同的服务场景，侧重于授权过程的不同交互过程设计。例如，SPKI/SDSI 类似于公钥基础设施（public key infrastructure, PKI），允许用户通过自签名证书完成分布式的权限管理和验证。OAuth2 常用于单个资源的跨域授权（如单点登录），采用令牌判定第三方信息服务作为数据使用者的权限，因此，其侧重于设计第三方信息服务从授权服务获取资源访问令牌的交互过程。UMA 则用于多个资源的统一授权管理，在 OAuth2 的基础上，采用令牌约束资源服务对授权服务应用程序接口（application programming interface, API）访问，因此，其增加了资源服务和授权服务之间资源注册、权限注册等所需令牌获取和交互过程。

回顾授权技术的发展，权限模型呈现细粒度化的趋势，各角色间的通信协议也逐渐标准化，以适应逐渐复杂的授权场景。然而，上述授权技术仍以服务有限范围的授权为目标，以资源管理者为中心，重点关注使用权限的管理，灵活性和可扩展性较差，难以适应多方参与、使用方式不可预见的互联网数据互操作场景。

1.2 面向互联网数据互操作的授权技术要求

互操作指两个或多个系统/组件交换信息并使用已交换信息的过程^[24]，数据的权属保护和安全可靠要求驱使该过程中数据

与应用解耦。这一概念源于个人数据管理系统^[25-27]，并在万维网之父蒂姆·伯纳斯·李发起的 SoLiD（social linked data）^[8]项目中正式被提出，要求数据以用户可选择、可信任、可控制的方式进行存储，经授权许可后允许应用使用。在数据与应用解耦模式下，数据脱离于应用，被存储在数据自治空间中，应用与数据的关系由传统一对一关系演变为多对多关系。数据与应用解耦模式如图2所示。

相应地，授权技术需要支持同一数据资源被来自多个应用的使用者访问，同时支持不同应用接入来自不同数据管理者的数据资源，这与传统授权技术存在以下区别。

（1）服务场景不同

数据资源与应用分离，在授权复杂度大幅提升的同时，引发了更高的安全性风险，如违背管理者授权意愿而造成的数据泄露风险，以及使用者因缺乏对授权的知情权和拒绝权而面临访问非法数据的风险^[28]。

（2）权限要求不同

数据流通过程涉及的多方（如涉及所有者、持有者、监管者等^[29]）权限要求各异，权属意愿存在冲突，使用权限出现数学运算、联合查询、扰动等各类计算及其组合^[4,30]等复杂操作，使传统固化的权限模型难以准确表达使用权限，数据权限难以被控制。

（3）效率要求不同

GDPR 等法律法规明确要求数据的使用需要“告知”用户并经用户“同意”。传统授权技术（例如应用程序跟踪透明度框架、开放授权协议^[5]等）要求数据管理者在线处理授权请求。然而，在涉及多方管理的数据互操作中，多个数据管理者难以保证同时在线，授权效率难以被保证，进

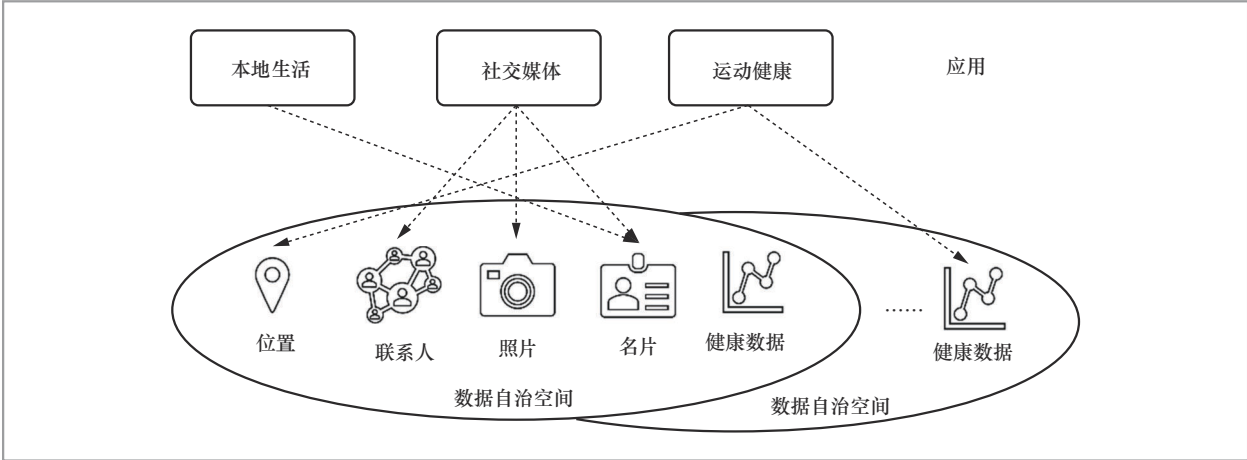


图2 数据与应用解耦模式

而影响数据跨域流通的可用性。特别地，开放的数据使用范围导致管理者需要处理的授权请求量增加，这进一步对授权管理的效率提出了更高的要求。

综上，面向互联网数据互操作的授权技术要求如下。

- R1：构建信任机制，保证授权过程的透明性和可审计性，保障权属信息的可验证性和隐私性，保护互操作中各方的权利。
- R2：通过结构化的管理权限类型定义冲突解决机制，支持互联网数据互操作过程中对管理权限进行细粒度和自定义配置。
- R3：设计可扩展的使用权限类型和生效条件定义，支持可度量的使用权限定义，刻画复杂的数据使用场景。
- R4：采用高效且准确的自动化授权策略生成方法，降低互操作中人工授权管理的成本。
- R5：采用高效的策略评估技术，快速定位匹配的规则、完成策略评估，以提升授权管理的执行效率。

基于上述授权技术要求，本文归纳面向互联网数据互操作的授权技术体系如图3所示。

2 面向互联网数据互操作的授权技术

在互联网数据互操作技术体系中，授权技术研究尚处于初步阶段。

数字对象架构^[7]提出将权限作为数字对象的元数据，但并未在协议设计中对此进行标准化定义。基于DOA构建的数联网架构^[31]提出的数据即服务（data as a service，DaaS）也存在类似的问题，仅提出数据所有者对数据授权管理的设计原则，并未给出具体解决方案。但该工作已意识到面向互联网数据互操作的授权需要解决多方管理及多方授权意愿的不一致性问题。

社交关联数据 SoLiD 项目^[8]采用基于标识 WebID（Web identity）的网络访问控制（web access control，WAC）^[32]作为权限模型，通过对符合 WAC 格式的 .acl 文件中的规则读写实现授权。文献[33–35]提出融合 WAC 和开放数字权利语言（open digital right language，ODRL），使用策略语言、形状约束语言（shapes

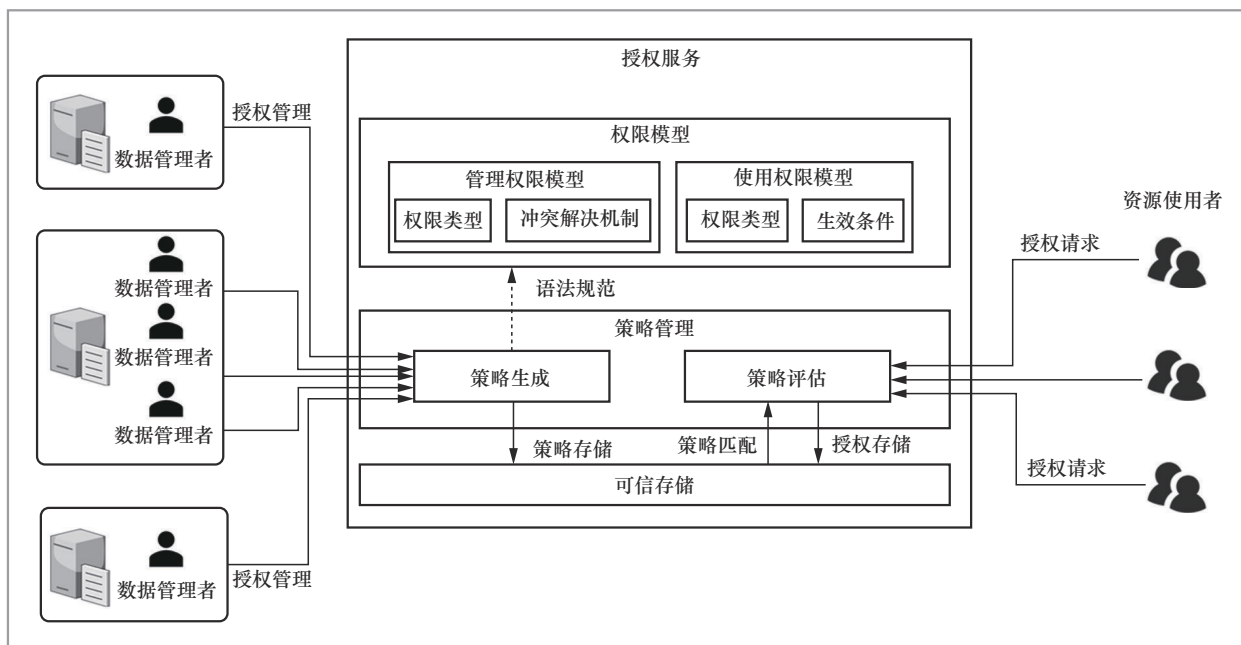


图3 面向互联网数据互操作的授权技术体系

constraint language, SHACL) 等扩展权限描述, 满足数据隐私保护和安全合规的监管需求。但这类研究仍仅关注使用权限定义, 尚未解决多方管理问题。文献[36-37]则关注到了信任机制的问题, 提出基于区块链存储权属信息, 实现权限管理过程的透明性和可审计性。

数据基础设施 Gaia-X^[9]作为数据与应用之间的连接中枢, 提供数据发现和交换服务。不同于 SoLiD 中权限由所有者自主管理, Gaia-X 提供联盟授权服务, 由多个独立于数据所有者和使用者之外的授权服务分布式管理, 允许基于区块链构建。然而, Gaia-X 并未给出关于该授权服务的进一步实现细节。

数据的跨域互联互通(即互操作)的实现方案还包括各数据交易市场。国内的数据交易市场大多基于中心化服务构建, 国外则出现了 IOTA (Internet of thing application) 等基于区块链构建的数据交

易市场。授权作为服务功能, 被数据发布者作为管理者管理数据 API、数据集等的访问权限, 以建立一对一的数据交易关系。

可以看到, 上述数据互操作技术中的授权实现仍较原型化, 这与当前互联网数据互操作技术的实施仍处于初步阶段有关。当前, 互联网数据互操作领域仍以解决一对一的跨域数据交换问题为目标, 尚未考虑多对多关系的权属保护问题、未针对性地解决涉及多方管理的可信授权问题。表1总结了互联网数据互操作技术中的授权现状。

尽管互联网数据互操作技术中的授权尚处于初步发展阶段, 但其面临的信任机制及复杂性问题已经受到信息中心网络^[10]、以社区为中心的协作系统^[11]等领域的关注。这些研究均面向跨域数据授权展开, 对于面向互联网数据互操作的授权技术研究和实现具有重要的借鉴意义。

表1 数据互操作技术中的授权现状

数字对象架构	服务架构	权限模型	R1	R2	R3	R4	R5
DOA	分布式	未定义统一模型,采用类似访问控制列表管理的方式	×	×	×	×	×
SoLiD	分布式	WAC,允许基于ODRL,SHACL等的其他语言扩展	*	×	√	×	×
Gaia-X	联盟式	未定义明确规范,仅提出采用ODRL等机器可读的模型	√	×	×	×	×
数据市场	中心化	未定义统一模型,采用类似访问控制列表管理的方式	*	×	×	×	×

注:*表示部分研究中满足该特性。

2.1 信任机制

面向互联网数据互操作的授权要求信任机制的构建（R1），本节将按照集中式授权服务和分布式授权服务两类架构展开对信任机制研究的讨论，不同类型的授权服务架构如图4所示。

2.1.1 集中式授权服务

集中式授权服务由数据管理者之外的第三方提供，统一管理数据权属信息，如图4（a）所示。现有集中式授权服务大多依赖于对授权服务提供商的信任。然而，已有研究发现授权服务商并不总是可信的，特别是在其面临中心化攻击的风险时^[55]。因此，部分研究提出通过去中心化的方式来构建信任机制，采用区块链技术解决数

据管理者对授权服务的信任问题。例如，文献[38]将OAuth2服务去中心化，将授权令牌的生成过程编码到智能合约中，并将生成的令牌存储在分布式账本中，以保证OAuth2授权服务提供商按照数据管理者的意愿进行授权管理。类似地，文献[39-41]将授权策略或权属信息记录在区块链中，保证其极难被篡改。这类方法大多采用联盟链技术，通过多个组织间的共识机制，确保权属信息的一致性和极难被篡改性，进而保证授权服务的透明性和可审计性。然而，这类研究以数据管理者为中心，忽视了数据使用者对权属信息的可验证性要求。此外，区块链中权属信息明文存储，允许授权服务提供商通过本地账本掌握全局授权信息，造成权属信息隐含的隐私面临暴露的风险。

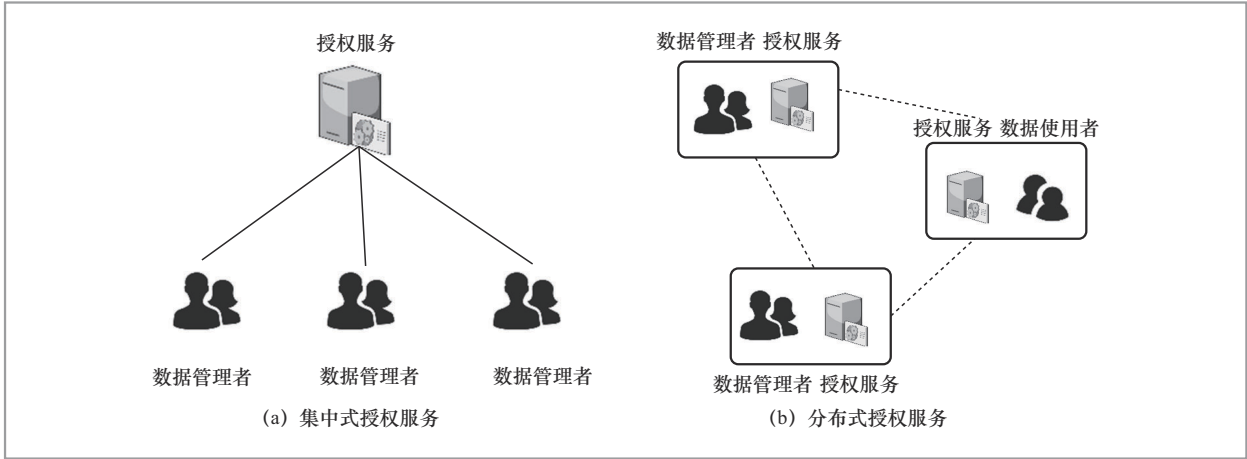


图4 不同类型的授权服务架构

2.1.2 分布式授权服务

分布式授权服务由数据管理者和数据使用者在本地分布式地完成授权管理、存储和验证,如图4(b)所示。根据权属信息的格式和授权服务的参与方,这类研究可分为基于访问控制列表的分布式授权、基于密钥分发的分布式授权、基于证书的分布式授权3类。

(1) 基于访问控制列表的分布式授权

基于访问控制列表的分布式授权是最经典的授权形式,即采用类似于文件系统的方式,由数据管理者在本地为每个数据资源独立维护访问控制列表(access control list, ACL)文件。典型的基于访问控制列表的分布式授权有语义网领域所采用的访问控制技术^[42-43],每个数据资源采用资源描述框架(resource description framework, RDF)进行结构化描述,具有与之对应的ACL文件。ACL由数据管理者自定义,同样采用结构化的语言描述。例如,可扩展标记语言(extensible markup language, XML)、网络本体语言(web ontology language, OWL)。区别于文件管理,这些技术还需要结合身份认证技术PKI、OIDC(OpenID connect)等实现,以实现跨域环境下的身份认证。这类研究主要解决数据管理者对数据权限可控的问题,因此,将权属信息看作一类特殊数据,由数据管理者在本地控制,通过数据管理者对数据授权过程的自治性,满足数据管理者对授权的可控要求。信任机制并不是这类研究的重点,尽管数据管理者本地可控的ACL满足管理者对授权过程的透明性要求,但数据使用者被排除在授权流程之外,权属信息的可验证性难以保障。

(2) 基于密钥分发的分布式授权

基于密钥分发的分布式授权^[44-49]常见

于流向不可控的信息中心网络和物联网。资源加密存储,数据管理者通过分发资源解密密钥完成授权管理,数据使用者在本地执行解密过程中完成授权验证,管理者和使用者共同参与授权过程。

这类研究通常采用密码学技术确保授权服务的可信性,具体来说,使用密钥对数据加密,被授予权限的用户通过持有数据的解密密钥来获得数据的访问权限。因此,基于密钥分发的授权服务能力与加密技术紧密相关。信息中心网络中的基于命名的访问控制^[44]提出使用对称密钥CK对数据加密,只有授权的第三方可以获取该解密密钥。因此,数据管理者用被授权的数据使用者的公钥KEK加密对称密钥CK完成权限授予,数据使用者采用公钥KEK对应的私钥KDK进行解密完成权限验证。权属信息保存在对称密钥CK的数据包中,由对称密钥CK的路径及公钥KEK的路径拼接构成,表明对称密钥CK被授权给持有公钥KEK对应私钥KDK的用户。然而,这类技术要求为新增的授权用户单独加密对称密钥,权限授予成本高。因此,部分研究提出采用属性基加密(attribute-based encryption, ABE)的方式实现可扩展性更强的授权^[45-46],将密钥与特定属性相关联,符合属性要求的密钥持有者均能够解密密文。进一步地,密文策略属性基加密(ciphertext-policy attribute-based encryption, CP-ABE)技术被引入^[47-48],采用访问控制策略作为密钥生成的输入,以完成更细粒度的权限控制。然而,属性加密的高时延、高功耗特点难以适应物联网场景。文献[49]提出采用通配的身份基加密(WKD-IBE),根据资源地址的路径信息生成密钥对资源及其子资源进行加密。WKD-IBE允许基于层级结构分发密钥,子节点的私钥由父节点根据子

节点的身份标识 (ID) 生成, 父节点的私钥可解密所有子节点的公钥加密的内容。因此, 资源管理者可对资源进行层级式管理, 将订阅资源地址及有效期表示为层级结构, 派发不同的私钥给不同订阅资源的资源使用者。基于密钥分发的分布式授权除了提供了一种授权服务信任机制, 也减轻了对资源服务的信任依赖。然而, 密钥分发引入了对密钥管理服务的依赖, 造成授权服务并非完全透明和可审计。

(3) 基于证书的分布式授权

基于证书的分布式授权^[50-51]源于 SPKI/SDSI。证书作为权属信息, 由数据管理者授予使用者, 由数据使用者自行保存, 并在访问资源时携带以验证权属。这类研究通过数据管理者和使用者共同参与授权过程, 以加密数字签名等密码学技术为基础建立起数据管理者和使用者间的信任。例如, 传统 SPKI/SDSI 的授权证书包含授予数据使用者权限的数据管理者签名, 以避免证书的伪造。文献[50]采用基于散列消息认证码 (hash-based message authentication code, HMAC) 的授权凭证确保其完整性。文献[51]则在 SPKI/SDSI 的基础上, 采用 WKD-IBE 保证权属信息的隐私性和机密性。然而, 基于证书的本地权属信息存储在允许验证的同时, 引入了多级权限委派链路中的各级证书验证的通信负载和时间成本。因此, 文献[28]采用分布式哈希表存储权属记录, 文献[51]提出基于默克尔树的可扩展日志存储方案, 对证书链进行集中维护, 在提高验证效率的同时, 确保操作日志的不可被篡改性和可审计性。综上所述, 基于证书的分布式授权由数据管理者完成授权管理, 由数据使用者完成授权验证, 授权存储则位于数据使用者本地, 结合加密和数字签名技术建立信任机制。但由于本地存储难以适应

权限委派过程中的长链路权限验证, 集中维护证书的方式被引入以解决该问题。

2.2 权限模型

针对面向互联网数据互操作的授权技术要求 R2 和 R3, 本节分别从使用权限模型和管理权限模型两个方面讨论现有研究进展。

2.2.1 使用权限模型

权限类型与使用场景紧密相关。在传统授权技术中, 权限类型包括读取、写入、创建、删除、执行等预定义的数据操作类型, 以及类似于数据库操作的权限定义。而这类预定义的权限集难以适应动态且复杂的数据交换场景。

为解决该问题, 现有研究提出对数据的动态使用过程进行建模, 典型技术为使用控制 (usage control, UCON)^[52]。UCON 将数据的动态使用过程拆分为前置、正在进行和后置 3 个环节分别定义权限^[53]。其中, 权限被划分为消费者权利、供应商权利和身份权利 3 种类型。消费者权利对应于使用权限, 分为查看和修改两大类, 以及不公开、有条件允许、公开 3 个权限级别, 并允许对使用权限的次数进行控制。文献[54]提出使用控制的形式化定义和策略规范, 将系统属性、状态、义务、条件等组成的使用策略定义为一组时序逻辑公式, 以符合使用控制的决策连续性和属性可变性的特点。然而, UCON 假设所有授权具有固定的生命周期, 无法满足数据流通中的可扩展要求。因此, 文献[55]提出一种用于持续授权和使用控制的可扩展模型, 使用确定性有限状态机对操作权限及关系进行刻画。

上述多阶段使用过程定义的方式将数

据使用过程拆解为多个操作过程，每个操作过程的权限定义采用预定义的使用权限类型，缺乏对复杂的单次操作权限的定义能力。为支持可扩展的操作权限定义，文献[56]提出对数据处理函数标识，作为使用权限定义的一部分，并对数据的预处理、后处理等过程的操作进行明确定义。数据跨域使用的不确定性引起数据管理者对数据安全和隐私的担忧。因此，在数据离域前，数据匿名化和数据扰动等技术常被用于敏感信息脱敏。数据的离域处理仅允许在经过特殊处理的非敏感信息上执行，以降低数据离域的隐私风险。为明确定义这类具有隐私要求的数据使用方式，泛化级别^[57]、匿名化方法^[58]、差分隐私算法及隐私预算^[30]等参数被引入，作为使用权限定义的一部分，约束数据使用者对数据的操作。然而，不同于传统读写操作，这类使用方式难以被度量和控制。文献[62]提出可采用静态程序分析进行权限验证，通过规定程序中基本函数和操作对应的权限，评估程序代码的剩余权限。

不同权限模型的区别除了体现在上述使用权限定义上，还体现在使用场景的建模能力上。例如，MAC、RBAC、ABAC等权限模型分别具备基于身份、角色、属性等不同粒度的使用场景刻画能力。随着授权使用场景的复杂化，现有研究从关系特征、对象属性、上下文环境3个方面扩展，以提升使用权限模型的表达能力。

关系特征多出现于基于关系的访问控制（relationship-based access control, ReBAC）^[59-60]，依赖于社交网络图模型构建，将使用者和管理者表示为节点，关系表示为边。因此，使用者的特征可建模由一条或多边构成的路径特征，包括使用者和管理者之间的关系及跳数，以表达传统权限模型难以表达的信任关系。单一的关

系特征限制了权限模型的表达能力，因此，基于用户与用户关系的访问控制（user-to-user relationship-based access control, UURAC）^[61]在ReBAC基础上允许为路径添加属性信息，支持属性的ReBAC（attribute-supporting ReBAC, AReBAC）^[62]提出一种支持资源属性描述的ReBAC模型，以实现细粒度的权限授予。上述ReBAC方法常应用于图网络结构中，在未建立用户关系的场景中难以适用。

数据权限的授予除了受用户关系影响，还受数据派生关系影响。例如，数据的二次开发在一定程度上受原始数据权限的约束。此外，数据资源、数据使用者本身的特征也是管理者决定是否授权的判断因素之一。考虑到这些因素，面向对象的属性定义方式是一种常见的解决方案。基于实体的访问控制（entity-based access control, EBAC）^[63]在权限模型中引入实体关系定义来增强ABAC模型的表达能力。面向对象的基于关系的访问控制语言（object-oriented relationship-based access-control language, ORAL）^[64]将ReBAC形式转化为面向对象的ABAC模型拓展，将关系表示为类模型、对象模型的属性。

除了上述数据资源特征、数据使用者特征、关系特征等因素，数据跨域使用的动态性和复杂性还要求对使用环境、使用目的等方面明确定义，以避免出现因上下游处理方式、网络、存储等因素引发的数据安全问题。文献[65-66]提出隐私政策语言（primelife policy language, PPL），引入数据处理的目的和下游使用约束声明，以及数据存储和分享时需要履行的责任定义。PPL在可扩展的访问控制标记语言（extensible access control markup language, XACML）基础上，预定义一组

统一资源标识符 (uniform resource identifier, URI) 作为处理方式、目的等取值, 用于实现机器可读的隐私策略, 支持在数据处理过程中自动执行对数据访问的权限控制。类似地, 文献[67]提出使用键值对格式维护数据使用时的存储、聚合和网络连接等条件定义。

尽管上述研究从多个维度丰富了使用权限模型对使用场景描述的能力, 但采用预定义的方式定义各个维度允许使用的取值范围, 导致权限模型相对固化、灵活性差, 难以应用在权限频繁变化和规模庞大的跨域数据交换中。针对该问题, 文献[68]提出了一种基于频谱聚类和风险的访问控制 (spectral clustering-risk-based access control, SC-RBAC) 模型, 基于用户历史访问数据, 采用改进的频谱聚类算法对角色进行聚类, 判断当前的权限使用场景。文献[69]则提出了RBAC角色的逆视图, 允许资源管理者根据自己的理解定义角色和角色层次结构, 而非采用预定义的有限角色类型。

综上所述, 现有研究通过对使用过程进行拆解, 采用函数标识和有限状态机等方式, 实现了使用过程和单次操作权限的可扩展定义。采用基于属性、关系、实体的权限模型中的一种或多种, 从数据使用者、数据资源、使用者和管理者的关系特征、对象属性等方面刻画使用场景, 引入上下游约束、存储、网络等使用环境描述, 进一步覆盖动态且复杂的使用场景。

2.2.2 管理权限模型

管理权限模型用于建模管理者对资源的管理能力, 除了对管理权限的明确定义, 还包括多个管理者之间的授权冲突解决机制等。

(1) 管理权限定义

传统授权服务中, 管理权限并未作明确定义。管理权限主要表现为管理者对数据资源权限委派的权利。不同管理者的权限具有同质化的特点, 由系统预定义或由系统管理员分发。大多数现有授权技术中, 授权由单一管理者决定, 即任意一个管理者可决定权属信息的产生, 不同管理者产生的权属信息之间具有覆盖关系。部分研究^[70-71]意识到协同授权管理的必要性, 要求数据的授权由 n 方中至少 m ($m \leq n$) 方共同决定。其中, 文献[70]提出资源所有者直接指派一个或多个授权体参与授权管理, 每个授权体为一个区块链节点, 享受平等的授权管理权限。文献[71]结合秘密共享技术, 允许多所有者共同管理数据, 每个所有者持有主密钥秘密的分片, 享有同等的秘密共享权和数据管理权。

为满足多样化的管理权限要求, 部分研究提出类似于RBAC的方式, 对授权管理角色进行定义。不同之处在于, 管理者允许执行的操作由系统预设, 授权服务通过识别不同角色来进行管理权限的控制。例如, 文献[72-73]根据角色与数据的关系以及参与数据分享的环节, 将数据管理者分为所有者、贡献者、利益相关者和传播者。类似地, 文献[74]引入原型的概念描述用户和对象之间的关系, 定义数据主体、数据提供者、数据持有者等角色作为管理者。文献[75]则对主体和数据、数据和数据之间的关系进行定义, 在使用、产生、控制、触发、衍生5种基本关系之上, 引入了所有、贡献、属于、不可用4种新的关系, 作为管理权限定义的基础。但上述工作均未明确地对不同管理者角色允许执行的操作进行定义和区分。

在多方授权管理中, 不同角色的授权意愿并不总是保持一致的, 存在冲突问题。

现有研究提出为不同角色分配不同的操作权重和优先级，以进一步综合多方授权意愿。文献[73]提出采用多级的敏感程度代替二元（私有/公开）敏感程度定义，并为不同所有者分配不同的管理权重。文献[29]则提出数据治理模型，按照层级结构对管理权限进行建模，将不同角色管理权限划分到不同优先级。另外，有研究提出不同类型的数据授权由不同角色参与。例如，不同敏感级别数据的授权由不同的管理者参与，敏感数据由监管者和所有者共同参与管理，非敏感数据仅需要监管者管理^[76]。**表2**总结了管理权限类型定义的不同维度。

（2）授权冲突解决机制

多方的授权意愿并不总是保持一致的。为解决多方授权冲突问题，一类方案采用投票机制，定义不同投票占比情况下的授权意愿合并规则，如基本规则“任一许可、大多数许可、全部许可”^[77]及细粒度规则“弱一致许可、强一致许可、弱大多数许可、强大多数许可、绝对多数许可”^[78]。另一类方案采用集合运算合并授权意愿。例如，文献[73]将数据的传播路径构建成树结构，由路径上的节点共同参与管理，取所有节点授权意愿的交集作为最终的授权结果。

多方管理存在相互依赖和制约，简单的投票机制或集合运算难以被应用于这类复杂的场景。例如，假设A是监管者，B、C、D是所有者，E是持有者。当监管者A允许或拒绝访问时，所有者和持有者的决策结果将被忽略；当监管者未参与决策时，

3个所有者投票产生共同授权意愿，若持有者E或3个所有者共同授权意愿为拒绝授权，则决策结果为拒绝授权。现有研究并未关注授权意愿冲突问题，但已考虑到与之类似的多策略组合问题。例如，传统访问控制策略语言XACML^[79]提出了“拒绝覆盖、允许覆盖、第一适用、单一适用、拒绝除非允许、允许除非拒绝”6种组合算法，用于对策略集中的多个策略进行合并。其中，“拒绝覆盖”表示当策略集中至少一个子策略评估结果为拒绝时，该策略评估结果为拒绝，“允许覆盖”是“拒绝覆盖”的对偶算法，“第一适用”则与策略结构相关，即按顺序评估策略，将第一个适用的子策略评估的结果作为最终的决策结果。文献[80]在XACML基础上，引入了（强）合取、析取、否定、弱化等操作，覆盖了更多的策略组合情况。

考虑到授权结果对数据流通、用户隐私和数据安全的影响，文献[74]提出采用“拒绝优先”或“允许优先”的策略得出最终的权属信息。为实现隐私和共享之间的平衡，文献[72]对数据共享涉及的隐私风险或敏感级别进行定义，进一步提出基于隐私考量的冲突解决机制，将多方授权意愿合并的问题转换为最优化问题。文献[73, 81]引入数据的敏感级别定义，允许通过敏感级别加权和阈值控制授权意愿合并结果。

除了上述预定义的规则和策略，文献[82]提出可以通过两轮决策、引入额外信道、引入新评估者、线上和线下交互等方式来

表2 管理权限类型定义

定义维度	文献	方法描述
系统管理员	[70-71]	管理者由管理员指派；管理者允许执行的操作由系统预定义
角色	[72-75]	根据用户与数据的关系定义角色
权重	[73]	不同管理者的权重不同
优先级	[29]	不同管理者的操作权限优先级不同，存在覆盖关系，采用分层模型

解决冲突。

部分研究发现在社交网络中，授权意愿并非固定不变。当授权意愿发生冲突时，多方之间存在潜在的让步策略，可通过协商后达成一致。该过程可建模为数学模型或博弈论。例如，文献[83]允许每个用户定义授权意愿的重要程度，采用克拉克税机制，经过博弈得到最大化整体效益的结果。文献[84]提出基于社交关系亲密度评估分享意愿，构建模糊规则集表达让步策略，完成多方之间的授权意愿协商。类似地，文献[85]通过构建社交网络图，结合博弈论知识研究多方决策间的相互影响。文献[86]则引入机器学习算法解决该问题，识别影响隐私决策的上下文、偏好和论点3类特征，采用多项逻辑回归模型预测最佳授权决策。**表3**总结了不同的授权冲突解决机制。

综上所述，现有方案允许从管理者的角色、操作、权重、优先级等方面定义管理权限，以进行异质化的管理权限定义。多方管理者的授权冲突解决可以采用意愿合并或协商两种方式。前者通过预定义规则或策略综合各方意愿，后者允许各方设置让步机制通过博弈最终达成一致。然而，不同于使用权限模型，管理权限模型尚未形成规范化的策略语言，以支持可扩展的管理权限规则自定义。

2.3 策略管理

策略管理作为数据管理者人工授权管理的替代方式，包括策略生成过程以及策略评估过程。根据面向互联网数据互操作的授权技术要求R4和R5，本节将分别从高效的策略生成过程和策略评估过程两个方面对现有研究展开讨论。

2.3.1 策略生成过程

在策略生成过程中，数据管理者根据自身的授权意愿和选用的使用权限模型，添加和设置授权规则，以根据授权请求自动化地生成权属信息。然而，人工配置的策略存在效率低下的弊端，机器学习和深度学习等算法被引入，根据过去的人工处理的访问记录日志或自然语言文本自动化地生成策略，提高策略生成的效率。例如，文献[87]采用支持向量机（support vector machine, SVM），使用实时监控和数据采集系统中的数据作为SVM的输入，以获得的用户-角色或角色-权限的分配关系作为输出。文献[88]则提出使用卷积神经网络来从自然语言中提取ABAC模型中的属性。文献[89]提出一种高效的ReBAC策略挖掘算法，采用进化算法搜索候选策略空间，在特征选择阶段利用神经

表3 授权冲突解决机制

思路	方法	文献	备注
意愿合并	投票机制	[77-78]	—
	集合运算	[73]	—
	合并策略	[79-80]	适用于不同优先级的管理权限合并
	最优化问题	[72]	平衡数据流通性和隐私风险
	多轮决策	[82]	—
授权协商	博弈论	[83, 85]	—
	让步策略	[84]	—
	算法推理	[86]	—

网络来识别有用的特征。文献[90]则采用决策树算法从ACL和有关实体信息中挖掘ReBAC策略。

在真实世界中,历史访问记录并不总覆盖所有的情况,特别是当数据授权复杂度高时,将影响算法生成策略的准确度。为了提高这种情况下的模型准确性,文献[91]提出方法Rhapsody,其建立在频繁项挖掘算法的基础上,从稀疏访问日志中挖掘规则。Rhapsody建立在频繁项挖掘算法APRIORI-SD的基础上,根据设计的可靠性方法,降低生成过于宽松的授权策略的可能性。该方法首先确定一组候选规则,这些规则覆盖日志中一定数量的请求;接着,根据可靠性阈值,过滤掉过于宽松的规则,替换为等效的较短规则。文献[92]则采用类似文献[90]的决策树方法,解决权限描述信息不完整或者某些属性值丢失或未知的问题。

综上所述,现有研究针对不同权限模型设计了不同的高效策略生成方法,但大多针对特定权限模型ABAC、RBAC、ReBAC等。因此,为适应互操作中可扩展的权限模型,策略生成算法还需要进行针对性的设计,以在稀疏的历史请求中,高效地挖掘出准确刻画授权意愿的策略。

2.3.2 策略评估过程

策略评估过程由策略评估引擎执行,以资源管理者制定的授权策略和资源使用者发起的授权请求为输入,匹配授权请求和授权策略中的规则,判断是否予以授权,输出相应的权属信息。

(1) 策略表示及存储

策略表示方法影响策略规则的查找速度,是策略评估效率的影响因素之一。决策图是策略表示的一类常见方法。最初,XEngine^[93]提出将策略规则中的条件属性

数值化,进一步将规则表示为决策图中的不同路径。但有研究^[94]发现,XEngine评估每个请求时,需要沿着路径依次匹配叶子节点,检查路径上所有的属性值,计算最终的决策结果,无法表示按照规则合并的策略。因此,文献[94]提出采用两棵树分别用于匹配和合并,前者采用二分法快速检索适用的规则,再采用后者对适用的规则进行合并和评估。但该方法采用字符串比较替代了数值比较,效率不及XEngine。文献[95]提出进一步的优化,将合并规则转换为求节点和边元素的并集和交集,实现仅采用一个决策图的策略表示方法。策略评估引擎也可以通过逻辑表达式实现^[96-97]。其中,文献[96]提出SBA-XACML(set-based algebra – extensible access control markup language),将策略表示为推理表达式,如集合交集、并集、包含、属于、蕴含等。文献[97]提出一种适用于资源受限的物联网场景的、轻量级基于语义的策略语言和方案,在SBA-XACML的基础上,引入上下文环境、用户行为等隐私约束评估,并结合深度学习模型根据不同的上下文环境和用户行为进行快速决策。相比基于决策图的XEngine,基于推理的实现方式性能更好。

为进一步提高策略评估的效率,部分研究针对策略存储结构进行优化。文献[98]使用自动机理论,将属性表示为位图,每个属性的取值对应于 n 位的比特串,其中第 i 位表示第 i 个规则包含该属性。策略评估时,对不同属性的位图串进行AND运算,以判断属性是否匹配。文献[99]采用哈希映射(hashmap)将属性取值映射为该属性的唯一索引,根据索引计算请求匹配的规则存储位置,每个规则占用2位来表示规则的结果。文献[100]则基于字典树构建规则,通过公共规则属性前缀加速评估过

程，同时借助 MB 树（Merkle-BTree）保证策略的完整性。

（2）策略匹配

策略评估要求找到所有策略中匹配的规则，以进一步评估和判断是否授权。因此，减少策略集中需要匹配的规则数量是提高策略评估效率的主要思路。文献[101]采用事件演算模型识别冗余策略，进一步减少策略集中的规则数目，加快策略评估速度。对策略集中的规则聚类，是有效减少匹配次数的思路之一。文献[102]提出一种基于隐含狄利克雷分布（latent Dirichlet allocation, LDA）主题模型聚类的策略评估引擎 XDLEngine，并对所有规则进行数字化和向量化，方便规则匹配，引入余弦相似度对每个主题下的规则进行分类，大大减少了规则匹配过程中的比较次数，提高了匹配效率。此外，考虑到不同主题的独立性，该方法在规则匹配过程中使用多线程并行搜索，显著缩短了 XDLEngine 的评估时间。类似地，文献[103]提出一种采用密度峰值聚类优化算法的策略评估引擎 DPEngine。神经网络也被引入加速策略匹配过程。文献[104]使用 K 均值（ K -means）聚类算法和异步随机梯度下降 K 均值（SGDK-means）聚类算法将策略集划分为不同的簇，根据各自的特征区分不同的规则，并对它们进行标记，

以用于神经网络的训练，利用神经网络来搜索适用的规则，并引入定量神经网络来降低服务器的计算成本。

综上所述，策略评估效率的提升依赖于决策图、逻辑表达式、位图等表示和存储结构，以及通过聚类、神经网络等算法缩小策略集，加速策略匹配。表 4 总结了策略评估的不同实现方法。

策略管理的本质是资源管理者的授权意愿管理。一种高效的授权管理模式是直接采用机器学习模型替代形式化的策略表达资源管理者的授权意愿，使用经过训练的机器模型自动化地决定授权请求是被授予还是被拒绝^[105-106]。文献[107]采用支持向量机对受时间约束的访问进行自动化授权。文献[108]提出了基于深度学习访问控制（deep learning based access control, DLBAC），以用户和资源的元数据为输入，以授权结果为输出，避免了属性和策略的制定过程。但这类方法的问题在于，管理者的授权意愿并不是一成不变的，非形式化的策略使管理者无法主动根据自身意愿的变化来调整规则。

3 未来展望

3.1 挑战

总结分析面向互联网数据互操作的授权技术在信任机制、权限模型和策略管理 3 个方面的研究进展，可以看到，当前授权技术仍面向有限范围或特定场景的数据流通展开研究，难以适应复杂且动态的互联网数据互操作场景，其未来面临的挑战如下。

（1）跨域授权成本高

当前，不同场景下的授权技术方案多种多样，不同机构和企业采用不同的授权技术实现。即便同样采用标准化协议

表 4 策略评估的不同实现方法

模块	方法	文献
策略表示	决策图	[93-95]
	逻辑表达式	[96-97]
	位图	[98]
策略存储	哈希映射	[99]
	字典树	[100]
	精简策略	[101]
策略匹配	聚类	[102-104]
	神经网络	[104]

OAuth2, 不同企业的实现也并不相同, 使数据流通过程中权属信息难以通用。因此, 为接入不同的数据, 应用程序需要适配不同的授权服务, 数据接入复杂度高。数据管理者也需要在不同的授权服务处分别管理数据, 数据管理成本高。

(2) 可信和效率难兼顾

授权服务从内部系统逐步发展为跨域系统, 通过将授权模型的各部分解耦, 使高复杂度的应用成为趋势, 依赖于区块链或采用完全分布式的授权服务是当前解决授权可信问题的主要方法。其中, 密码学技术和共识机制在保障机密性和可审计性的同时, 引入了额外的通信和计算成本, 降低了服务效率。然而, 互联网数据互操作要求数据访问时验证权属, 对授权服务的验证效率提出了高要求。信任机制和服务效率之间存在矛盾。

(3) 异质管理权限表达能力不足

现有研究提出的权限模型允许支持特定数据流通场景的权限保护, 如社交网络内容分享和传播链路中的多方权利。这类模型依据所服务的场景, 通过枚举的方式实现异质的管理权限类型定义, 存在控制粒度粗、规则单一等问题, 难以捕捉数据流通过程中数据派生、权限委派等产生的权限依赖关系。此外, 尽管部分研究也提出了多方授权管理冲突解决方案, 但尚未形成专用的授权管理策略语言, 难以对管理权限进行扩展和自定义, 无法适应数据流通过程中复杂、多方的协同管理模式。

(4) 复杂的使用权限难以被控制

在数据流通过程中, 数据使用者往往需要利用数据完成特定的计算任务。尽管已有研究的使用权限模型覆盖了引入计算权限的定义, 但仍然局限于特定的计算类型, 如匿名化、差分隐私算法等。这类特

定类型的计算, 允许静态程序分析检查是否在授予权限约束下使用数据。然而, 算法程序类型繁多, 静态程序尚无法识别所有操作的权限定义, 导致权属信息验证时难以检验使用者的操作是否满足被授予权限约束的要求。

3.2 研究方向

(1) 授权协议标准化

为降低开放数据授权的接入复杂度和管理成本, 亟须规范互联网数据互操作中的授权流程, 形成面向互联网数据互操作的通用授权协议。授权协议的核心是提供全网可达的数据授权服务, 实现数据权属信息的互联互通。当前, 互联网基础设施(如域名系统、PKI等)是实现主机和网站等互联网资源互联互通的关键, 用于维护互联网资源与其信息的映射关系。为实现数据全网范围的流通, 根据过去互联网实现主机和网站互联互通的经验, 以兼容现有互联网基础设施和协议的方式设计授权协议是未来授权协议标准化的主要路径。

(2) 加密权属信息存储优化

不加区分地加密权属信息是验证效率降低的根本原因, 导致当数据流通过程中出现长链路的权限委派时, 授权验证效率低。因此, 在保障必要信息安全的基础上, 设计高效的权属信息存储结构和检索方式是未来面向互联网数据的授权技术需要解决的问题。例如, 针对加密权属信息三元组的索引构建问题展开研究, 加速权属信息验证。

(3) 多方管理权限模型构建

未来面向数据互操作的授权技术需要基于现有访问控制策略扩展, 设计具有可扩展性的多方授权管理规则和策略语法,

形式化定义多方授权管理过程,确保多利益方协同管理的透明性和公平性。一方面,研究可扩展的管理权限定义方法,覆盖操作、权重、委派等特征,兼容现有的各类成熟的权限模型(适合单方授权管理);另一方面,研究可扩展的管理策略定义,例如,构建覆盖跨国、跨部门、跨机构的多层级多方协同授权管理方式的分层策略。进一步地,研究多方授权管理策略评估引擎,自动化执行多方授权决策。不同于传统策略模型评估以使用者请求和策略为输入,管理策略模型评估的输入还包括多方管理者的动态授权意愿,传统的基于决策图的策略表示方法难以适用,需要针对性地设计管理策略表示方法。

(4) 可约束的使用权限定义

未来,面向互联网互操作的授权技术仍需进一步研究可扩展、可度量的计算权限定义。一方面,探索类似信息熵等方式量化计算过程获取的信息量,以便于权属验证环节判断数据使用者权限的有效性,也可进一步用于数据流通过程中数据价值评估;另一方面,研究基于可信硬件的使用权限控制,确保数据使用者确实按照所声明的计算方式执行操作。

4 结束语

近年来,随着数字经济的快速发展,数据要素加速流通,数据互操作需求增加,面向数据互操作的授权技术成为当前授权技术研究的关键课题。然而,目前尚未有研究针对互联网数据互操作的授权技术进行系统性梳理。为填补国内外该领域的研究空白,本文对面向互联网数据互操作的授权技术要求展开了首次分析,从信任机制、权限模型、策略管理方面梳理和归纳

了当前授权技术的研究进展。研究发现,现在的研究尚未解决面向数据互操作的授权中的所有问题,各种技术方案大多专注于某一个领域的授权问题,因而在可用性、安全性、效率等方面有所取舍。结合当前的研究进展,本文分析了面向数据互操作的授权技术未来发展面临的挑战和研究方向,旨在为未来面向互联网数据互操作的授权研究提供参考和思路。

参考文献:

- [1] 李晓东. 价值互联网的“价值”所在[J]. 人民论坛, 2023(16): 29-33.
LI X D. The “value” of the Internet of value[J]. People’s Tribune, 2023(16): 29-33.
- [2] WEI J Q, LI Y, FU Y F, et al. Data interoperating architecture (DIA): decoupling data and applications to give back your data ownership[C]//Proceedings of the 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC). Piscataway: IEEE Press, 2023: 438-447.
- [3] FRASER B. RFC 2196: site security handbook[S]. United States: IETF, 1997.
- [4] 张纪林, 顾小卫, 张亦钊, 等. 跨域数据授权运营研究及应用[J]. 大数据, 2023, 9(4): 83-97.
ZHANG J L, GU X W, ZHANG Y Z, et al. Research and application of cross-domain data authorization and operation [J]. Big Data Research, 2023, 9(4): 83-97.
- [5] HARDT, DICK E. RFC 6749: the OAuth 2.0 authorization framework[S]. United States: IETF, 2012.
- [6] 汤奇峰, 邵志清, 叶雅珍. 数据交易中的权利

- 确认和授予体系[J]. 大数据, 2022, 8(3): 40-53.
- TANG Q F, SHAO Z Q, YE Y Z. Authenticating and licensing architecture of data rights in data trade[J]. Big Data Research, 2022, 8(3): 40-53.
- [7] KAHN R E, BLANCHI C, LANNOM L, et al. Digital object interface protocol specification[S]. Hong Kong: ATSD, 2018.
- [8] SAMBRA A V, MANSOUR E, HAWKE S, et al. Solid: a platform for decentralized social applications based on linked data[J]. MIT CSAIL & Qatar Computing Research Institute, 2016.
- [9] BRAUD A, FROMENTOUX G, RADIER B, et al. The road to European digital sovereignty with Gaia-X and IDSA[J]. IEEE Network, 2021, 35(2): 4-5.
- [10] TOURANI R, MISRA S, MICK T, et al. Security, privacy, and access control in information-centric networking: a survey[J]. IEEE Communications Surveys & Tutorials, 2018, 20(1): 566-600.
- [11] PACI F, SQUICCIARINI A, ZANNONE N. Survey on access control for community-centered collaborative systems[J]. ACM Computing Surveys, 2019, 51(1): 1-38.
- [12] Open Geospatial Consortium. OpenGIS Geography Markup Language (GML) encoding standard[R]. 2007.
- [13] WILKINSON M D, DUMONTIER M, AALBERSBERG I J, et al. The FAIR guiding principles for scientific data management and stewardship[J]. Scientific Data, 2016, 3: 160018.
- [14] FITZPATRICK G J, WOLLMAN D A. NIST interoperability framework and action plans[C]//Proceedings of IEEE PES General Meeting. Minneapolis: IEEE, 2010: 1-4.
- [15] HU V C, FERRAILOLO D, KUHN D R. Assessment of access control systems [M]. Gaithersburg: National Institute of Standards and Technology, 2006.
- [16] JØSANG A. A consistent definition of authorization[M]//Security and Trust Management. Cham: Springer International Publishing, 2017: 134-144.
- [17] DOWNS D D, RUB J R, KUNG K C, et al. Issues in discretionary access control [C]//Proceedings of the 1985 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 1985: 208.
- [18] BERTINO E, JAJODIAT S, SAMARATI P. Enforcing mandatory access control in object bases[M]//Security for Object-Oriented Systems. London: Springer London, 1994: 96-116.
- [19] FERRAILOLO D F, SANDHU R, GAVRILA S, et al. Proposed NIST standard for role-based access control[J]. ACM Transactions on Information and System Security, 2001, 4(3): 224-274.
- [20] BONATTI P A, SAMARATI P. A uniform framework for regulating service access and information release on the Web[J]. Journal of Computer Security, 2002, 10(3): 241-271.
- [21] RIVEST R L, LAMPSON B. SDSI—a simple distributed security infrastructure[R]. 1996.
- [22] ELLISON C, FRANTZ B, LAMPSON B, et al. RFC 2693: SPKI certificate theory [S]. United States: IETF, 1999.
- [23] MACHULAK M P, MALER E L, CATALANO D, et al. User-managed access to web resources[C]//Proceedings of the 6th ACM workshop on Digital identity management. New York: ACM, 2010: 35-44.

- [24] PACKET S. Institute of electrical and electronics engineers[J]. IEEE Standard Computer Dictionary a Compilation of IEEE Standard Computer Glossaries, 2009, 3: 128.
- [25] CHAJED T, GJENGSET J, VAN DEN HOOFF J, et al. Amber: decoupling user data from web applications[C]//Proceedings of the 15th Workshop on Hot Topics in Operating Systems. Berkeley: USENIX Association, 2015: 19.
- [26] WANG F, MICKENS J, ZELDOVICH N, et al. Sieve: cryptographically enforced access control for user data in untrusted clouds[C]//Proceedings of 13th USENIX Symposium on networked systems design and implementation. Berkeley: USENIX Association, 2016: 611–626.
- [27] ZHANG B, MOR N, KOLB J, et al. The cloud is not enough: Saving IoT from the cloud[C]// Proceedings of 7th USENIX Workshop on Hot Topics in Cloud Computing. Berkeley: USENIX Association, 2015.
- [28] VRIELYNCK P J, BENI E H, JANNES K, et al. DeFIREd: decentralized authorization with receiver-revocable and refutable delegations[C]//Proceedings of the 15th European Workshop on Systems Security. New York: ACM, 2022: 57–63.
- [29] MAHMUDLU R, DEN HARTOG J, ZANNONE N. Data governance and transparency for collaborative systems [M]//Data and Applications Security and Privacy XXX. Cham: Springer International Publishing, 2016: 199–216.
- [30] WANG L, KHAN U, NEAR J, et al. PrivGuard: privacy regulation compliance made easier[C]//Proceedings of 31st USENIX Security Symposium. Berkeley: USENIX Association, 2022: 3753–3770.
- [31] HUANG G, LUO C R, WU K D, et al. Software-defined infrastructure for decentralized data lifecycle governance: principled design and open challenges [C]//Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). Piscataway: IEEE Press, 2019: 1674–1683.
- [32] SACCO O, PASSANT A, DECKER S. An access control framework for the web of data[C]//Proceedings of the 2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications. Piscataway: IEEE Press, 2011: 456–463.
- [33] ESTEVES B, PANDIT H J, RODRÍGUEZ-DONCEL V. ODRL profile for expressing consent through granular access control policies in solid[C]//Proceedings of the 2021 IEEE European Symposium on Security and Privacy Workshops. Piscataway: IEEE Press, 2021: 298–306.
- [34] DEBACKERE L, COLPAERT P, TAELEMAN R, et al. A policy-oriented architecture for enforcing consent in solid[C]//Proceedings of the Companion Proceedings of the Web Conference 2022. New York: ACM, 2022: 516–524.
- [35] BRAUNCH J, PAPANCHEV V, KÄFER T. SISSI: an architecture for semantic interoperable self-sovereign identity-based access control on the web[C]//Proceedings of the Proceedings of the ACM Web Conference 2023. New York: ACM, 2023: 3011–3021.
- [36] GHAYVAT H, SHARMA M, GOPE P, et al. SHARIF: solid pod-based secured healthcare information storage and exchange solution in Internet of Things[J].

- IEEE Transactions on Industrial Informatics, 2022, 18(8): 5609–5618.
- [37] CAI T, HONG Z C, LIU S, et al. Social-Chain: decoupling social data and applications to return your data ownership[J]. IEEE Transactions on Services Computing, 2023, 16(1): 600–614.
- [38] TRUONG N B, SUN K, LEE G M, et al. GDPR-compliant personal data management: a blockchain-based solution[J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 1746–1761.
- [39] SHAFAGH H, BURKHALTER L, HITHNAWI A, et al. Droplet: decentralized authorization and access control for encrypted data streams[EB]. arXiv preprint, 2018, arXiv: 1806.02057.
- [40] ROUHANI S, BELCHIOR R, CRUZ R S, et al. Distributed attribute-based access control system using permissioned blockchain[J]. World Wide Web, 2021, 24(5): 1617–1644.
- [41] CHEN E, ZHU Y, ZHOU Z Y, et al. Policychain: a decentralized authorization service with script-driven policy on blockchain for Internet of Things[J]. IEEE Internet of Things Journal, 2022, 9(7): 5391–5409.
- [42] SAMBRA A, GUY A, CAPADISLI S, et al. Building decentralized applications for the social web[C]//Proceedings of 25th International Conference Companion on World Wide Web – WWW'16 Companion. New York: ACM, 2016: 1033–1034.
- [43] CAPADISLI S, GUY A, VERBORGH R, et al. Decentralised authoring, annotations and notifications for a read-write web with dokieli[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017: 469–481.
- [44] ZHANG Z Y, YU Y D, RAMANI S K, et al. NAC: automating access control via named data[C]//Proceedings of the MILCOM 2018 – 2018 IEEE Military Communications Conference (MILCOM). Piscataway: IEEE Press, 2018: 626–633.
- [45] WANG G J, LIU Q, WU J. Hierarchical attribute-based encryption for fine-grained access control in cloud storage services[C]//Proceedings of the 17th ACM conference on Computer and communications security. New York: ACM, 2010: 735–737.
- [46] WANG G J, LIU Q, WU J, et al. Hierarchical attribute-based encryption and scalable user revocation for sharing data in cloud servers[J]. Computers & Security, 2011, 30(5): 320–331.
- [47] CHEN N Y, LI J G, ZHANG Y C, et al. Efficient CP-ABE scheme with shared decryption in cloud storage[J]. IEEE Transactions on Computers, 2022, 71(1): 175–184.
- [48] DAS S, NAMASUDRA S. Multiauthority CP-ABE-based access control model for IoT-enabled healthcare infrastructure [J]. IEEE Transactions on Industrial Informatics, 2023, 19(1): 821–829.
- [49] KUMAR S, HU Y, ANDERSEN M P, et al. JEDI: many-to-many end-to-end encryption and key delegation for IoT [C]//Proceedings of 28th USENIX Security Symposium. Berkeley: USENIX Association, 2019: 1519–1536.
- [50] BIRGISSON A, POLITZ J G, ERLINGSSON Ú, et al. Macaroons: cookies with contextual caveats for decentralized authorization in the cloud[C]//Proceedings of 2014 Network and Distributed System Security Symposium.

- Reston: Internet Society, 2014.
- [51] ANDERSEN M P, KUMAR S, ABDELBAKY M, et al. WAVE: a decentralized authorization framework with transitive delegation[C]//Proceedings of 28th USENIX Security Symposium. Berkeley: USENIX Association, 2019: 1375–1392.
- [52] PARK J, SANDHU R. Towards usage control models: beyond traditional access control[C]//Proceedings of the seventh ACM symposium on Access control models and technologies – SACMAT '02. New York: ACM, 2002: 57–64.
- [53] PARK J, SANDHU R. The UCON ABC usage control model[J]. ACM Transactions on Information and System Security, 2004, 7(1): 128–174.
- [54] ZHANG X W, PARISI-PRESICCE F, SANDHU R, et al. Formal model and policy specification of usage control[J]. ACM Transactions on Information and System Security, 2005, 8(4): 351–387.
- [55] HARIRI A, IBRAHIM A, DIMITRAKOS T, et al. WiP: metamodel for continuous authorisation and usage control[C]//Proceedings of the 27th ACM Symposium on Access Control Models and Technologies. New York: ACM, 2022: 43–48.
- [56] CIRILLO F, CHENG B, PORCELLANA R, et al. IntentKeeper: intent-oriented data usage control for federated data analytics[C]//Proceedings of the 2020 IEEE 45th Conference on Local Computer Networks (LCN). Piscataway: IEEE Press, 2020: 204–215.
- [57] LI M, SUN X X, WANG H, et al. Privacy-aware access control with trust management in web service[J]. World Wide Web, 2011, 14(4): 407–430.
- [58] GERL A, BENNANI N, KOSCH H, et al. LPL, towards a GDPR-compliant privacy language: formal definition and usage[M]//Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018: 41–80.
- [59] FONG P W L. Relationship-based access control: protection model and policy language[C]//Proceedings of the first ACM conference on Data and application security and privacy. New York: ACM, 2011: 191–202.
- [60] CRAMPTON J, SELLWOOD J. Path conditions and principal matching: a new approach to access control[C]//Proceedings of the 19th ACM symposium on Access control models and technologies. New York: ACM, 2014: 187–198.
- [61] CHENG Y, PARK J, SANDHU R. Attribute-aware relationship-based access control for online social networks [M]//Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014: 292–306.
- [62] RIZVI S Z R, FONG P W L. Efficient authorization of graph-database queries in an attribute-supporting ReBAC model[J]. ACM Transactions on Privacy and Security, 2020, 23(4): 1–33.
- [63] BOGAERTS J, DECAT M, LAGAISSÉ B, et al. Entity-based access control: supporting more expressive access control policies[C]//Proceedings of the Proceedings of the 31st Annual Computer Security Applications Conference. New York: ACM, 2015: 291–300.
- [64] BUI T, STOLLER S D, LI J J. Greedy and evolutionary algorithms for mining relationship-based access control policies[J]. Computers & Security, 2019, 80: 317–333.
- [65] TRABELSI S, SENDOR J, REINICKE S. PPL: primeLife privacy policy engine

- [C]//Proceedings of the 2011 IEEE International Symposium on Policies for Distributed Systems and Networks. Piscataway: IEEE Press, 2011: 184–185.
- [66] AZRAOUI M, ELKHIYAOU K, ÖNEN M, et al. A-PPL: an accountability policy language[M]//Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance. Cham: Springer International Publishing, 2015: 319–326.
- [67] WANG F, KO R, MICKENS J. Riverbed: enforcing user-defined privacy constraints in distributed web services[C]//Proceedings of NSDI 19: 16th USENIX Symposium on Networked Systems Design and Implementation. Berkeley: USENIX Association, 2019: 615–630.
- [68] JIANG R, HAN S S, YU Y M, et al. An access control model for medical big data based on clustering and risk[J]. Information Sciences, 2023, 621: 691–707.
- [69] YEN I L, TIWARI A, BASTANI F. Access control in dynamic IoT scenarios [C]//Proceedings of the 2023 IEEE 15th International Symposium on Autonomous Decentralized System (ISADS). Piscataway: IEEE Press, 2023: 1–8.
- [70] BATTAH A A, MADINE M M, ALZAABI H, et al. Blockchain-based multi-party authorization for accessing IPFS encrypted data[J]. IEEE Access, 2020, 8: 196813–196825.
- [71] ILIA P, CARMINATI B, FERRARI E, et al. SAMPAC: socially-aware collaborative multi-party access control[C]//Proceedings of the Seventh ACM Conference on Data and Application Security and Privacy. New York: ACM, 2017: 71–82.
- [72] HU H X, AHN G J, JORGENSEN J. Detecting and resolving privacy conflicts for collaborative data sharing in online social networks[C]//Proceedings of the 27th Annual Computer Security Applications Conference. New York: ACM, 2011: 103–112.
- [73] HU H X, AHN G J, JORGENSEN J. Multiparty access control for online social networks: model and mechanisms[J]. IEEE Transactions on Knowledge and Data Engineering, 2013, 25(7): 1614–1627.
- [74] DAMEN S, DEN HARTOG J, ZANNONE N. CollAC: Collaborative access control [C]//Proceedings of the 2014 International Conference on Collaboration Technologies and Systems (CTS). Piscataway: IEEE Press, 2014: 142–149.
- [75] BERTOLISSI C, MARTINEZ ANTON A, ZANNONE N. Data sharing in social networks[C]//Proceedings of the 28th ACM Symposium on Access Control Models and Technologies. New York: ACM, 2023: 181–192.
- [76] ARULANANTH T S, BASKAR M, ANBARASU V, et al. Multi party secure data access management in cloud using user centric block chain data encryption [J]. Pattern Recognition Letters, 2021, 152: 295–301.
- [77] CARMINATI B, FERRARI E. Collaborative access control in online social networks[C]//Proceedings of the Proceedings of the 7th International Conference on Collaborative Computing: Networking, Applications and Worksharing. Piscataway: IEEE Press, 2011: 231–240.
- [78] LI N H, WANG Q H, QARDAJI W, et al. Access control policy combining: theory meets practice[C]//Proceedings of the 14th ACM symposium on Access control models and technologies. New York: ACM, 2009: 135–144.
- [79] Standard O. Extensible access control

- markup language (XACML) version 3.0 [S]. Boston: OASIS Open, 2003.
- [80] SHEIKHALISHAHI M, STORK I, ZANNONE N. Privacy-preserving policy evaluation in multi-party access control [J]. *Journal of Computer Security*, 2022, 29(6): 613–650.
- [81] HU H X, AHN G J. Multiparty authorization framework for data sharing in online social networks[M]//Lecture Notes in Computer Science. Berlin: Springer Berlin Heidelberg, 2011: 29–43.
- [82] KINKELIN H, NIEDERMAYER H, MÜLLER M, et al. Multi-party authorization and conflict mediation for decentralized configuration management processes[C]//Proceedings of 2019 IFIP/IEEE Symposium on Integrated Network and Service Management. Piscataway: IEEE Press, 2019: 5–8.
- [83] SQUICCIARINI A C, SHEHAB M, PACI F. Collective privacy management in social networks[C]//Proceedings of the 18th international conference on World wide web. New York: ACM, 2009: 521–530.
- [84] SUCH J M, CRIADO N. Resolving multi-party privacy conflicts in social media [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(7): 1851–1863.
- [85] DING K M, ZHANG J S. Multi-party privacy conflict management in online social networks: a network game perspective[J]. *IEEE/ACM Transactions on Networking*, 2020, 28(6): 2685–2698.
- [86] FOGUES R L, MURUKANNAIAH P K, SUCH J M, et al. Sharing policies in multiuser privacy scenarios[J]. *ACM Transactions on Computer-Human Interaction*, 2017, 24(1): 1–29.
- [87] ZHOU L, SU C H, LI Z, et al. Automatic fine-grained access control in SCADA by machine learning[J]. *Future Generation Computer Systems*, 2019, 93: 548–559.
- [88] ALOHALY M, TAKABI H, BLANCO E. A deep learning approach for extracting attributes of ABAC policies[C]//Proceedings of the 23rd ACM on Symposium on Access Control Models and Technologies. New York: ACM, 2018: 137–148.
- [89] BUI T, STOLLER S D, LE H. Efficient and extensible policy mining for relationship-based access control[C]//Proceedings of the 24th ACM Symposium on Access Control Models and Technologies. New York: ACM, 2019: 161–172.
- [90] BUI T, STOLLER S D. A decision tree learning approach for mining relationship-based access control policies[C]//Proceedings of the 25th ACM Symposium on Access Control Models and Technologies. New York: ACM, 2020: 167–178.
- [91] COTRINI C, WEGHORN T, BASIN D. Mining ABAC rules from sparse logs[C]//Proceedings of the 2018 IEEE European Symposium on Security and Privacy (EuroS&P). Piscataway: IEEE Press, 2018: 31–46.
- [92] BUI T, STOLLER S D. Learning attribute-based and relationship-based access control policies with unknown values[M]//Information Systems Security. Cham: Springer, 2020: 23–44.
- [93] LIU A X, CHEN F, HWANG J, et al. XEngine[J]. *ACM SIGMETRICS Performance Evaluation Review*, 2008, 36(1): 1–10.

- 265–276.
- [94] PINA ROS S, LISCHKA M, GÓMEZ MÁRMOL F. Graph-based XACML evaluation[C]//Proceedings of the 17th ACM symposium on Access Control Models and Technologies. New York: ACM, 2012: 83–92.
- [95] NGO C, DEMCHENKO Y, DE LAAT C. Decision diagrams for XACML policy evaluation and management[J]. Computers & Security, 2015, 49: 1–16.
- [96] MOURAD A, JEBBAOUI H. SBA-XACML: Set-based approach providing efficient policy decision process for accessing Web services[J]. Expert Systems with Applications, 2015, 42(1): 165–178.
- [97] CHEHAB M, MOURAD A. LP-SBA-XACML: lightweight semantics based scheme enabling intelligent behavior-aware privacy for IoT[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 19(1): 161–175.
- [98] DENG F, WANG S Y, ZHANG L Y, et al. Establishment of attribute bitmaps for efficient XACML policy evaluation[J]. Knowledge-Based Systems, 2018, 143: 93–101.
- [99] DENG F, ZHANG L Y, ZHANG C Y, et al. Establishment of rule dictionary for efficient XACML policy management[J]. Knowledge-Based Systems, 2019, 175: 26–35.
- [100] ZHENG K, TIAN X. Integrity-preserving and efficient policy evaluation for XACML[J]. International Journal of Network Security, 2022, 24(2): 262–272.
- [101] ZAHOOR E, CHAUDHARY M, AKHTAR S, et al. A formal approach for the identification of redundant authorization policies in Kubernetes[J]. Computers & Security, 2023, 135: 103473.
- [102] DENG F, YU Z H, LIU W J, et al. An efficient policy evaluation engine for XACML policy management[J]. Information Sciences, 2021, 547: 1105–1121.
- [103] YU Z H, YAN Y H, DENG F, et al. An efficient density peak cluster algorithm for improving policy evaluation performance[J]. Scientific Reports, 2022, 12: 5000.
- [104] DENG F, YU Z H, SONG H B, et al. Improvement on PDP evaluation performance based on neural networks and SGDK-means algorithm[J]. Soft Computing, 2022, 26(6): 3075–3089.
- [105] CAPPELLETTI L, VALTOLINA S, VALENTINI G, et al. On the quality of classification models for inferring ABAC policies from access logs[C]//Proceedings of the 2019 IEEE International Conference on Big Data (Big Data). Piscataway: IEEE Press, 2019: 4000–4007.
- [106] KARIMI L, ABDELHAKIM M, JOSHI J. Adaptive ABAC policy learning: a reinforcement learning approach[EB]. arxiv preprint, 2021, arxiv:2105.08587.
- [107] CHANG C, LIN I, LIAO C. An access control system with time-constraint using support vector machines[J]. International Journal of Network Security, 2006, 2: 150–159.
- [108] NOBI M N, KRISHNAN R, HUANG Y F, et al. Toward deep learning based access control[C]//Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy. New York: ACM, 2022: 143–154.

作者简介



李颖（1996–），女，中国科学院计算技术研究所博士生，主要研究方向为数据安全、数据互操作、域名系统。



李晓东（1976–），男，博士，中国科学院计算技术研究所研究员、互联网基础技术实验室主任，清华大学互联网治理中心主任、公共管理学院教授，伏羲智库创始人、主任，中国互联网协会副理事长，中国互联网络信息中心（CNNIC）原主任，互联网名称与数字地址分配机构（ICANN）原副总裁，主要研究方向为互联网基础技术、数据分析、网络安全和治理等方面的行业服务、技术创新和政策研究。曾负责国家域名全球服务平台规划建设和关键系统的研制、国际化域名和电子邮件国际标准制定、“中国”国家顶级域名入根、国家互联网大数据平台规划等工作。



费子郁（1999–），男，中国科学院计算技术研究所硕士生，主要研究方向为数据安全、数据治理。



彭博韬（1991–），男，博士，中国科学院计算技术研究所助理研究员，中国互联网协会青年专家，主要研究方向为向量数据相似性搜索算法、现代硬件加速技术。开创性地将现代硬件加速技术引入向量数据索引领域，在VLDB、IEEE ICDE、IEEE TKDE、VLDBJ等国际顶级学术会议与期刊中发表多篇相关高水平论文。

收稿日期: 2024-04-16
通信作者: 李晓东, XL@ict.ac.cn
基金项目: 国家自然科学基金项目 (No.62202450); 山东省中央引导地方科技发展基金 (No.YDZX20203700002564); 中国科学院网络安全和信息化专项咨询研究项目 (No.CAS-WX2023ZX02-04); 工业和信息化部工业互联网创新发展项目 (No.TC200A00S)
Foundation Items: The National Natural Science Foundation of China (No.62202450), Central Funds Guiding the Local Science and Technology Development of Shandong Province (No.YDZX20203700002564), Network Security and Informatization Research Project from Chinese Academy of Sciences (No.CAS-WX2023ZX02-04), Industrial Internet Innovation and Development Project from the Ministry of Industry and Information Technology of China (No.TC200A00S)