

从全生命周期管理角度看大数据安全技术研究

李树栋^{1,2}, 贾焰², 吴晓波³, 李爱平², 杨小东⁴, 赵大伟⁵

1. 广州大学网络空间先进技术研究院, 广东 广州 510006; 2. 国防科技大学计算机学院, 湖南 长沙 410073;
3. 烟台职业学院信息工程系, 山东 烟台 264670; 4. 西北师范大学计算机科学与工程学院, 甘肃 兰州 730070;
5. 山东省计算中心(国家超级计算济南中心), 山东 济南 250014

摘要

数据信息在生产、存储、使用、传输、共享、销毁等环节暴露出的大数据安全问题突出, 成为了制约大数据应用发展的瓶颈。总结了世界各国的大数据安全发展战略; 从大数据的全生命周期管理角度出发, 重点聚焦于大数据的收集、存储、传输、使用、开放(访问控制、共享安全与隐私保护)、销毁和管理策略(体系和标准建设)等阶段, 对当前大数据安全防护技术的研究现状和趋势进行阐述。给出了未来大数据安全技术研究需要解决的几个新问题。

关键词

大数据安全; 存储安全; 传输安全; 共享安全; 隐私保护

中图分类号: TP309

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2017047

Techniques of big data security from the perspective of life cycle management

LI Shudong^{1,2}, JIA Yan², WU Xiaobo³, LI Aiping², YANG Xiaodong⁴, ZHAO Dawei⁵

1. Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China
2. School of Computer Science, National University of Defense Technology, Changsha 410073, China
3. School of Software Engineering, Yantai Vocational College, Yantai 264670, China
4. College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China
5. Shandong Computer Science Center (National Supercomputer Center in Ji'nan), Ji'nan 250014, China

Abstract

The big data security problems during the process of data production, storage, use, transmission, sharing and destruction have been explored. Firstly, the development strategies of big data security worldwide were summarized. Second, from the perspective of the whole life cycle management, the current technical research and the progress of big data security protection were introduced by focusing on its collection, storage, transport, use and open sharing, destruction and management strategy. Finally, some new problems that need to be solved in the future research of big data security were discussed.

Key words

big data security, storage security, transmission security, shared security, privacy protection

1 引言

互联网、云计算、大数据等信息技术正在深刻改变着人们的思维、生产、生活和学习方式,并延深进入人们的日常生活。伴随着社交媒体、电商、健康医疗、智能交通、电信运营、金融和智慧城市等各行业各领域的大数据的产生,大数据分析技术和应用研究使大数据呈现出不可限量的经济社会价值和科学研究意义,引起了国内外学术界和产业界的研究热潮,对此各国政府也高度重视并不断上升为国家战略高度。

自2008年《自然》推出专刊“Big Data”、2011年《科学》推出专刊“Dealing with Data”之后^[1],2012年欧洲信息学与数学研究协会出版专刊“Big Data”;2013年电气和电子工程师协会(Institute of Electrical and Electronics Engineers, IEEE)决定每年举办一次大数据科学国际会议;2012年、2013年中国科学院分别组织召开了关于大数据科学的香山会议;2013年国家自然科学基金委员会召开了双清论坛——大数据技术与应用中的挑战性科学问题^[2],并设立了重点项目群,以支持大数据技术基础研究。

然而,在大数据时代背景下,云计算技术的发展导致大数据在收集、存储、共享、使用等过程面临的安全威胁愈演愈烈,大数据泄露的企业个人隐私信息给用户带来了巨大的损失。因此,大数据的信息安全问题也成为了制约大数据应用发展的瓶颈,可以说没有安全防护的大数据,将是数据灾难。

为此,如何构建围绕大数据全生命周期的可管、可控、可信的数据安全体系,将大数据安全和数据应用深度融合,制定顶

层规划,保障大数据产业健康发展,成为了大数据应用和安全研究领域中的主要科学问题。

针对上述现状,本文将总结各国与大数据安全相关的战略规划,重点围绕大数据的全生命周期,阐述大数据的采集安全、存储安全、传输安全、共享安全、销毁和管理安全等阶段的研究现状及进展。

2 世界各国大数据安全发展战略

美国、英国、日本、澳大利亚、中国、欧盟等世界主要国家和地区纷纷出台制定了与大数据安全相关的战略和政策^[3],统筹规划大数据建设发展,加强对大数据的利用和安全防护,占领大数据技术研究国际竞争力的制高点。

美国是全球最早发布大数据发展战略的国家。早在2011年,美国总统科技顾问委员会提交报告指出:大数据蕴含着重要的战略价值,建议成立“大数据高级指导组织”。2012年3月发布了《大数据研究和发展战略》,将其上升到国家战略。在该战略框架下,美国国防部高级研究计划局在整理了先期执行的多尺度异常检测、洞察、灵眼及视频图片分析等多个大数据安全项目之后,启动了大数据与国家安全战略相结合的重要项目“XDATA项目”^[4],致力于大数据核心安全技术的解决方案,以突破大数据技术在军事领域的应用。2013年第13642号总统行政令提出了新的联邦大数据管理准则,指出在数据公开的同时,要保护好隐私安全性与机密性。2014年,总统办公室提交了一份报告《大数据:把握机遇,维护价值》,重点阐述了大数据为美国隐私保护、信息安全和社会发展带来的新挑战。同年,颁布了《2014年国家网络安全保护法案》与《网络安全信

息共享法案》，以积极应对日益凸显的大数据安全问题。

2012年6月，英国政府发布了《开放数据白皮书》，以促进公共服务数据的开放使用，专门针对个人隐私保护制定了《个人隐私影响评估手册》。2013年10月公布《把握数据带来的机遇：英国数据能力战略》，重点在提升大数据分析技术、推动产学研合作、确保数据安全存储使用等方面制定了一系列措施。

2013年6月，日本政府公布了2013—2020年的《新IT战略——创建最尖端IT国家宣言》，全面阐述了2013—2020年以发展开放公共数据和大数据为核心的国家战略。2014年7月，日本以引进的美国网络防御系统为基础，研究开发数据安全和保护技术。2015年4月，日本审议了《个人信息保护法》和《个人号码法》修正案，以推动并规范“大数据”的安全应用。

2012年7月，澳大利亚发布《信息安全管理指导方针：整合性信息的管理》，为大数据整合中涉及的安全风险提供了指导。同年11月，对《隐私法》中的隐私原则进行修订，规范了私人信息数据从采集、存储、安全、使用、发布到销毁的全生命周期管理。2013年8月发布《公共服务大数据战略》，提出发展大数据的6条原则和战略目标。

2015年，我国国务院印发《促进大数据发展行动纲要》（以下简称《纲要》），系统部署大数据的发展工作。《纲要》部署的三大主要任务之一是健全大数据安全保障体系，强化安全支撑。加强大数据环境下的网络安全问题研究和基于大数据的网络安全技术研究。2016年，国务院发布的《“十三五”国家信息化规划》（以下简称《规划》）中的第四部分“重大任务和重大工程”中明确指出，要“实施大数据安全保障工程，加强数据资源在采集、传输、存

储、使用和开放等环节的安全保护”。《规划》着重提出：要加快大数据安全的采集、存储、清洗、分析、发掘、可视化、安全与隐私保护等关键技术攻关。2017年召开了首届中国数据安全峰会，以“共建数据安全，共享安全数据”为宗旨，拟从网络安全、硬件安全、数据安全等全方位打造数据安全生态圈，建立“中国数据安全共同体”。

2010年11月，欧盟通信委员会向议会提交了《开放数据：创新、增长和透明治理的引擎》，围绕数据开放，制定大数据安全的发展战略。2012年1月，欧盟颁布了《通用数据保护条例》。2016年通过了新版《一般数据保护条例》，在数据保护监管与数据保护指令等方面进行了重要改革，该条例将于2018年5月25日全面实施。德国于2009年对《德国联邦数据保护法》进行修改并生效，旨在防止个人信息隐私泄露；2014年8月，通过了《数字议程（2014—2017）》，推动“网络安全”是其中的三大重要进程之一，突出了数据安全的保护的共识。法国于2013年2月发布《数字化路线图》，把大数据技术列为5项大力发展的战略技术之一。同年7月，发布了《法国政府大数据五项支持计划》，在数据科学家教育、设立基金等方面制定了措施，以吸引科学家开展研究。

3 大数据安全研究现状及进展

3.1 大数据采集安全

大数据的源头众多、数据多样、数据增长速度快，大数据采集的可信性是一个重要关注点。其面临的安全威胁之一是数据被伪造或刻意制造，如电商交易的虚假评论、互联网应用中的数据伪造或粉饰，

有可能诱导人们在分析数据时得出错误结论,影响用户的决策判断力。因此,如何对采集到的大数据进行评估、去伪存真,提高识别非法数据源的技术能力,确保数据来源安全可信,是大数据采集安全面临的一个重要挑战。

目前,基于大数据的数据真实性分析技术被广泛认为是有效的方法。传统信息安全技术无法完全保证数据的真实性,而利用大数据分析技术可以去除噪音,能够提高对伪造或虚假数据的鉴别能力。例如Yahoo公司和Thinkmail公司利用大数据分析技术过滤垃圾邮件;DataVisor公司提供恶意账户识别技术帮助Yelp网站减少虚假评论^①,此外还有基于内容、基于贝叶斯、基于数据挖掘等多种垃圾邮件过滤技术^[5,6],利用机器学习技术识别电商虚假评论^[7]、识别社交网络(新浪微博)中虚假身份^[8]、识别社交媒体中的水军^[9]等。

数据真实性分析技术类似于数据清洗和数据预处理,可以作为数据融合的基础。针对大数据的海量特征,利用大数据分析技术(包括机器学习和深度学习技术),通过构建数据分类模型,可以获得更高的识别准确率。因此,对于解决如何保证数据的可信度这一难题,除了需要考虑数据来源的真实性,还需要从数据传播途径、数据处理过程等多方面开展研究,这也是未来的一个研究热点。

3.2 大数据存储安全

目前,大数据主要是分布式地存储在大数据平台(Hadoop)中,采用云存储技术,以多副本、多节点、分布式的形式存储各类数据。数据的集中存储和滥用增加了被非法入侵和数据被泄露的风险。因此,如何保障大数据存储安全一直是重点研究的问题。

(1) 数据加密

数据加密是保障大数据存储安全的主流方法之一。当前,可以使用的国家商用密码局制定的应用标准包括SSF33、SM1、SM2、SM3、SM4、SM7、SM9等加密标准,对大数据进行加密处理后再存储。使用的技术包括基于属性的加密^[10]、同态加密^[11]等。然而对于海量数据来说,加解密操作不可避免会带来无法忽略的额外开销,这限制了数据加密技术在大数据存储安全中的应用范围。

(2) 磁盘存储安全

随着大数据的广泛应用,大数据平台或数据中心存储数据集中,其安全性异常重要。其中,硬盘是存储数据主要介质设备,针对传统磁盘存储数据的安全性一直是一个研究热点,主要工作集中在防磁盘数据篡改、防数据泄露失窃等方面^[12,13],还包括自存储安全解决方案^[14]、网络安全硬盘^[15]、安全云盘^[16]以及分布式存储系统安全^[17]等研究工作。固态硬盘因其延迟低、吞吐量大、能耗低等优点,正逐渐替代传统机械硬盘。新出现的可信固态硬盘技术依靠提供安全存储接口和协议,保证数据的机密性,并细粒度控制用户访问存储的数据,使得数据存储是可以信任的,从而保护了数据的存储安全以及机密性。因此,对于数据密集型应用的数据存储安全需求而言,可信固态硬盘有望成为保障大数据平台存储安全的新基础^[18]。

3.3 大数据的传输安全

目前大数据的分析、计算需要云计算平台强力支撑,在将大数据迁移到云环境中进行进一步存储、分析和计算时,可能会失去对大数据的安全控制,导致对安全边界之外的数据缺乏必要的控制,比如一些关键信息系统内外网的数据安全问题,

①

<https://www.jrzj.com/187908.html>

这将进一步导致大数据安全防护问题的复杂化。因此,在各个信息系统内的机密信息不断迁移到云平台的过程中,如何确保数据传输中的机密性和完整性^[19]是一个重要挑战。对此,被认同的普遍做法是利用密码学相关技术解决。

3.3.1 数据的机密性方面

加密算法能确保大数据的机密性,但传统的加密算法存在密钥管理复杂、计算开销过大等方面的不足,无法真正体现数据的高价值。一些新兴的加密体制能解决大数据面临的安全问题,但如何适应大数据应用场景的数据规模和数据增长速度仍需要进一步深入地研究。属性加密体制将用户的属性与数据属性相关联,只有满足解密数据属性的用户才能获得加密的数据,因此它能提供数据的保密性,提升数据的服务效率,但该加密算法的效率较低,阻碍了其在大数据中的广泛应用^[20,21]。代理重加密算法体制能实现数据解密权限的传递,提供消息的保密性和访问控制的灵活性,不会泄露加密数据的内容,但无法很好地细粒度控制第三方代理者的密文转换权限^[22-24]。全同态加密在不解密数据的前提下,能对加密数据进行检索、比较等操作,但对数据的处理效率依然很低^[25,26]。可搜索加密技术能实现密文数据的查询和相关排序,但存在查询效率低、支持的数据结构类型单一、扩展性差等问题^[27-29]。

3.3.2 数据的完整性方面

越来越多的用户选择云计算平台对海量数据进行存储管理,但同时用户失去了对数据的控制权,无法确保云端数据的完整性。云端存储数据的公开审计受到学术界和产业界的广泛关注,成为云计算安全

领域的研究热点^[30]。Wang等人^[31]基于环签名提出了具有身份隐私保护的群组数据公开审计方案Oruta,但该方案是一种无条件的隐私保护,同时审计的开销与群组用户数量线性相关。Wang等人^[32]基于群签名设计了群组数据完整性验证方案Knox,但该方案不支持公开审计。Yuan等人^[33]基于多项式认证标签和代理标签提出了一个支持群组用户撤销的方案,但无法抵抗恶意云服务提供商和撤销用户之间的合谋攻击^[34]。针对共享数据完整性验证过程中由于执行数据更新和共享用户成员撤销等操作而带来的安全问题,Wang等人^[35]提出了一种抗合谋攻击的共享数据动态完整性验证方案,但没有考虑隐私保护问题。然而,这些方案仅实现了支持隐私保护的公开验证、数据动态更新和群组用户动态更新、用户身份追踪等部分功能,目前没有一个审计方案能同时实现上述所有功能,并且具有较低的计算、存储和通信开销。为了保证大数据的可靠性和可用性,云服务提供商根据用户需求对数据进行多副本存储,以提高容灾抗毁和数据可恢复的能力,但对于支持动态群组和隐私保护的多用户多副本数据完整性验证的研究成果较少^[36,37]。

在云环境下,混合云融合了公有云和私有云的优点,被广泛使用,因此,如何确保混合云中大数据的完整性和机密性,也将是一个具有挑战且值得深入研究的问题^[38]。

3.4 大数据使用和开放安全

当前,各国均把数据看作一种战略资源,在数据开放和共享过程中会产生更大价值。然而,当前针对大数据的安全技术缺乏或安全管理能力不足带来了很多安全问题,比如地下数据的“黑灰”产业链、侵犯用户个人信息安全、数据滥用等问题,严

重阻碍了数据共享使用的发展。一方面,在大数据使用环节,在大数据的查询、访问过程中,不严格的权限访问将导致数据泄漏;另一方面,在数据共享开放环节,数据资源跨部门、跨域共享使用,不可避免地导致数据被各使用方存储使用,其中任何一个使用方措施不当,都可能导致数据泄漏。因此,如何保证大数据的共享安全成为了当前的研究热点,这方面研究主要集中在安全访问控制、共享安全和隐私保护等方面。

3.4.1 大数据安全访问控制

访问控制是实现数据安全共享的重要手段,大数据及大数据应用的诸多新特征使传统访问控制在授权管理、策略描述、细粒度控制、隐私保护、实施架构等方面都面临严峻挑战^[39]。大数据背景下的访问控制呈现出判定依据多元化、判定结果模糊化或不确定化、多种访问控制技术融合化的新特点^[40],进而也促进了大数据访问控制技术的新发展,角色访问控制^[41]、风险访问控制^[42]、半/非结构化数据的访问控制^[43]、针对隐私保护的访问控制^[44]、世系数据相关的访问控制^[45]、基于密码学的访问控制^[46]、基于大数据分析结果的内容访问控制^[47]、行为访问控制^[48]、协同访问控制^[49]和数据联系访问控制^[50]等在大数据场景中得以提出或改进创新,多元化且多技术融合的访问控制可以更有效地支持复杂的大数据访问控制需求。

未来,大数据服务于访问控制、大数据的访问控制框架等将成为重要研究课题。尤其是在云环境下,数据管理系统或存储平台内部管理人员存在潜在的数据窃取能力(如金融证券从业人员的数据泄露问题)^②,因此,如何解决云平台内部管理

员的特权行为的安全性问题^[51],以防止数据泄密,也是需要重点研究的方向之一。

3.4.2 大数据共享安全

政府信息系统和公共数据的互联共享、各行业领域大数据共享和整合使用是大数据应用的必然趋势,是推动大数据流动性、产业链完善的主要动力,也是大数据核心价值得以体现的关键环节。然而,在数据共享开放发布环节,由于数据资源跨部门、跨管理域、跨省共享使用,不可避免地导致数据被各使用方存储使用。如果其中任何一个使用方措施不当,都可能导致数据泄漏。

针对数据共享安全问题, Munier等人^[52]针对跨单位/公司的数据安全,提出了一种面向云存储的文件跨部门使用的自我保护体系结构,使得在提供数据外包存储服务时,确保文档的安全性和私密性。

在信息流的安全控制方面, Krohn等人^[53]提出了分散信息流安全控制方法,以控制应用程序和外部世界之间的数据流动。该方法在隐私方面,允许不受信任的软件使用私有数据,而受信任的安全代码控制数据发布。在完整性方面,允许受信任的代码保护不受信任的软件免遭恶意入侵。类似的工作还包括参考文献^[54]提出的方法。Pasquier等人^[55]使用信息流控制表达和执行云中数据存储位置的需求,以实现将数据流限制存储在特定的地理位置。在共享数据的安全属性标记方面,张大军等人^[56]提出了云服务中基于属性加密的数据共享机制,定义了云服务中每个数据拥有者的数据属性。中国通信标准化协会网络与信息安全技术工作委员会云计算安全工作组制定了《面向公有云服务的数据安全标记规范》^③,主要适用于公有云内和公有云间数据流动的审计和监管所需的标记的表示、生成、使用和管理。

②
<http://stock.qq.com/a/20161124/004201.htm>

③
http://www.ccsa.org.cn/ccsa_tx/tx_to_std.php?tbname=project&tx_num=C-06&orderby=&PageNo=1

在领域大数据共享安全方面,贾培宏等人^[57]针对地理空间数据,提出了基于数字水印技术的数据安全共享方法;张璐等人^[58]针对政府大数据,使用数据加密实现数据共享安全的管理。面向医疗大数据,Narayan等人^[59]提出云存储中电子健康记录的安全共享管理方案,使患者的健康记录动态和可伸缩地不同医疗保健提供者之间安全共享数据;雷婉^[60]基于属性加密技术提出了访问权和修改权分离的个人医疗健康数据安全共享方案。

当前,数据的安全共享问题成为了大数据应用的阻碍之一,如何使数据拥有者安全地保存数据,同时数据共享发布后被各方使用产生价值,成为数据安全共享要解决的困难问题。

3.4.3 大数据隐私保护

隐私保护是近几年大数据安全研究中的一个主要问题,大数据在商业化应用中如果数据处理不当将对用户隐私造成极大侵犯,而从法律上讲用户应当有权决定自己的信息如何被利用。因此,实现可控的隐私保护无疑是其中面临的一个重大挑战,这也是当前大数据应用面临的阻碍之一。当前,医疗信息隐私泄露、个人隐私信息泄露、位置隐私数据泄露^[61]等安全问题使得对大数据隐私保护的研究越来越多,这里主要对数据隐私保护框架、数据匿名处理等内容进行阐述。

(1) 大数据隐私保护模型

信息隐私保护模型最早是1999年由Culnan等人^[62]从管理学角度提出的,随后涌现了大量的面向静态数据的隐私保护模型^[63]、面向动态连续数据发布的隐私保护模型^[64-67]。最近,贾焰等人^[68]给出了大数据时代的隐私概念和生命周期保护模型。李凤华等人^[69]提出了隐私计算的概

念,包括隐私计算的定义、隐私计算研究范畴以及隐私计算框架,其中引入香农信息熵对隐私信息进行量化测度。

(2) 数据匿名处理技术

数据匿名化方法删除了敏感数据以保护用户隐私,传统的匿名化方法包括 k -Anonymity^[70]、 t -Closeness^[71,72]和 l -Diversity^[73]等方法。然而采用数据匿名时,若匿名化不够,攻击者可能通过多渠道获取关联信息,进而推测出用户隐私,这将导致用户隐私面临极大的泄露风险^[74]。如果删除信息过多,会使得大数据应用中面向数据挖掘的分析毫无价值,这会造成基于数据溯源的大数据分析和数据隐私保护之间的矛盾,因为数据溯源需要从数据源头获取大数据,而数据源头往往涉及敏感隐私信息^[75]。

(3) 社交网络匿名保护技术

发布社交网络数据时,一般将含有用户的标签、属性以及位置信息等个人隐私信息进行匿名化处理,然而攻击者往往可以通过社交网络数据中的社交关系推断出用户之间的关系及隐私信息,这方面的研究有基于图结构的抗结构攻击的网络发布的隐私保护方法^[76,77]、基于各类社交关系的用户关系预测^[78,79]等。然而,随着社交网络结构的不断演化,用户关系日益紧密,社交网络匿名保护研究将成为一个热门研究点^[80]。

当前,从大数据管理的全生命周期角度出发,进一步构建完善的大数据隐私保护体系,明确隐私保护内涵(包括隐私的知情权、遗忘权等)是隐私保护下一步研究的重要热点。其次,从信息安全角度来看,数据溯源和隐私保护之间的平衡仍然是一个值得关注的问题。

3.5 大数据安全销毁或删除

大数据的安全销毁或删除是近年大数

据安全的一个重要研究热点。用户对Web服务的依赖性越来越大,如果其存储在云端或云平台的数据删除不彻底,极有可能使其敏感数据被违规恢复,从而导致用户数据或隐私信息面临泄露的风险。传统的数据物理删除的方法是采用物理介质全覆盖的方法,然而针对云计算环境下的数据删除问题,这一手段并不可信。在云环境下,用户失去了对数据的物理存储介质的控制权,无法保证数据存储的副本同时也被删除,导致传统删除方法无法满足大数据安全的要求。因此,如何保证被删除的数据确实被删除,即保证数据可信删除,是一个重要挑战。

在这方面,2005年Perlman^[81,82]提出了一种支持两种类型的数据可信删除方法,该方法基于文件的创建时间,使过期文件无法被恢复访问,并支持按需删除个别文件,从而确保了文件被安全删除的可信性。2009年Geambasu等人^[83]提出了一个针对密文数据可信删除的方法,设计了一个验证的“消失原型”,能够确保在用户指定的时间期限之后,所有特定数据副本都变得不可读,而不需要对用户进行任何特定的操作,即使攻击者同时获得了该数据的缓存副本和用户的密码和密钥。2010年Yang等人^[84]针对保护被删除的数据,提出了基于数据加密标准的数据安全删除策略,使被删除的文件无法被任何人恢复,即使是云存储平台的管理人员。类似的还包括参考文献[85]提出的策略。2013年Cachin等人^[86]基于图论和密码学技术,提出了基于无环图删除策略的安全删除数据存储系统模型,该策略通过删除属性和保护类来说明数据销毁,当删除指定的数据属性时,同时也必须相应地删除该属性链接的保护类,从而最终达到数据安全删除的目标。2015年,Xiong等人^[87]提出了一种带有时间指定属

性的安全数据自毁方案,只有在允许的时间间隔内且与密文关联的属性满足密钥的访问结构时,才可以读取数据。在用户指定的过期时间之后,敏感数据将被安全地自毁。

在云计算环境下,个人数据被第三方(云数据存储平台)缓存、复制和存档,而这些数据往往没有真正被用户控制,在网络和云存储系统中,正确删除数据并清除所有痕迹的操作通常是无法预见的,因此,数据的可信删除技术仍然是未来大数据安全技术研究的热点。

3.6 大数据的管理安全

保障大数据安全除了技术研究之外,管理策略同样重要。大数据的建设是一项有序、动态、可持续发展的系统工程,因此,建设针对大数据的安全保障体系和标准规范至关重要。当前,这方面研究主要集中在两个核心要点:大数据安全保障体系和大数据安全标准化建设。

3.6.1 大数据安全保障体系

大数据应用具有极高的商业价值,也具有较高的社会价值,因此,针对包括政务、医疗、电信、智能交通、智慧城市等在内的不同领域行业大数据特点,研究提出有针对性的大数据安全保障和防护体系显得意义重大,且极具研究价值。

针对政府大数据,张璐等人^[58]从共享数据安全管理和数据授权访问两个安全问题角度出发,研究提出了面向政务大数据的安全保护模型。针对医疗领域大数据,2014年,Chen等人^[88]根据医疗数据的信息安全需求,构建了医疗数据面临的安全风险指数的计算方法,该计算方法考虑的因素包括数据来源、数据量、加密算法等。张

洪亮等人^[89]基于密码学中的身份认证、数字签名和加密等技术提出了医疗大数据安全保障体系,以解决当前健康医疗大数据面临的安全威胁与信任问题。针对电信运营商大数据,陶冶等人^[90]分析了电信行业大数据面临的安全现状,并对世界各大电信运营商的大数据安全策略进行了比较。张滨^[91]提出了针对电信运营商的大数据安全体系架构,该体系包括安全保障体系的基本执行框架、安全保障体系的能力支撑和效果评测等方面。裴金栋等人^[92]基于电信大数据的数据采集、数据标签、数据处理、功能模块和行业应用5个层次,构建了电信大数据全生命周期的安全管控体系。针对智能交通大数据面临的安全威胁和挑战,作为欧盟第七框架计划(FP7)资助的研究项目的一部分, Schreiner等人^[93]提出了用于管理智能交通系统的安全体系架构,该体系实现了对对象管理组的数据分发服务和高级访问控制。宋珊珊^[94]提出了应对智能交通大数据安全的管理策略,涵盖了存储安全、应用安全、管理安全等内容。针对智慧城市大数据,陈红松等人^[95]提出了安全管理策略和框架。针对大数据应用安全的防护体系,吕欣等人^[96]提出了大数据安全保障体系架构和评价指标体系,保障体系架构覆盖了战略保障、运行保障、技术保障、组织管理和过程管理5个层次,评价指标体系覆盖了建设情况、运行情况和安全态势3个方面。

3.6.2 大数据安全标准化

针对大数据安全合规方面的需求,大数据安全的标准化体系建设对规范和推动大数据产业的安全发展具有重要作用。当前,国内外标准化组织都认识到大数据的安全问题,并积极开展大数据安全的相关标准研究制定工作。

(1) 我国大数据安全标准化建设

我国开展大数据安全标准工作的组织主要是全国信息安全标准化技术委员会,2017年4月,该组织发布了《大数据安全标准化白皮书(2017)》^④,白皮书中提出从基础标准、平台和技术、数据安全、服务安全、行业应用5个方面构建大数据安全标准体系框架,规划设置了我国的大数据安全标准体系的工作路线。大数据安全标准体系框架如图1所示。

在该体系框架下,目前正在研究制定的国家标准有《大数据服务安全能力要求》《信息安全技术 大数据安全管理指南》《政府数据 数据分类分级指南》《政府数据 数据脱敏工作指南》等系列标准,并提出建议研制《大数据基础平台安全要求》《大数据交易服务安全要求》《大数据安全能力成熟度模型》等系列标准。

(2) 国际上大数据安全标准化建设

国际标准化组织/国际电工委员会标准化组织(ISO/IEC) JTC1正在开展与大数据安全直接相关的标准——ISO/IEC 20547《信息技术 大数据参考架构第4部分:安全与隐私》,涉及大数据安全内容。国际电信联盟电信标准化部门(ITU-T)已发布了涉及大数据安全的标准《大数据数

④

<http://www.cesi.cn/uploads/soft/170411/1-1F411162633.pdf>

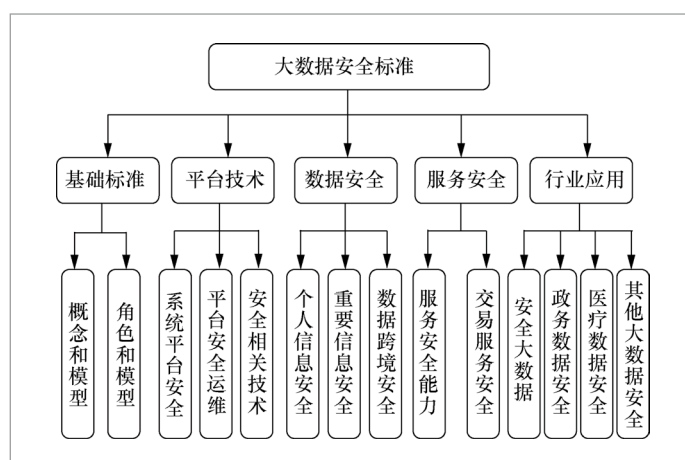


图1 大数据安全标准体系框架

据保全概述和要求》，目前正在推进研制的大数据安全标准有《移动互联网服务中的大数据分析安全要求和框架》《大数据即服务的安全指南》《电子商务业务数据生命周期管理的安全参考架构》等。美国国家标准与技术研究院(NIST)已经完成并发布了NIST SP 1500-4《第4册 安全和隐私保护》^[97]。

总之，在确保安全合规的情况下，如何规范指导各行业的大数据服务提供商，构建适合不同领域大数据的管理安全防护体系，以提供安全的大数据服务，仍然是急需解决的关键安全问题。

4 结束语

在大数据时代背景下，社交网络、健康医疗、智能交通、金融、电信等行业大数据蕴涵的巨大经济社会价值不断被学术界、产业界认同。然而，大数据安全问题成为了当前大数据产业资源开放共享程度低、大数据价值难以被有效挖掘利用的主要阻碍之一。

从科学研究角度，本文重点从大数据的全生命周期管理角度阐述了大数据在采集、存储、传输、共享、销毁和管理策略等方面的技术研究现状，分析了其发展趋势。未来大数据安全技术研究仍然面临着许多挑战，如何构建可控、可信、可管的大数据安全防护体系成为了大数据安全研究需要解决的重要科学问题。比如如何使企业既能安全保存大数据，又能科学挖掘利用数据价值，即数据共享与安全的平衡问题；如何使大数据既能被安全共享，又能保护个人隐私信息，即共享和隐私保护的平衡问题；如何不断完善大数据安全标准，推动国家“十三五”规划中关于大数据安全保障

工程的实施建设等，这些细分的问题都是未来需要研究解决的。

参考文献：

- [1] 李国杰, 程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域[J]. 中国科学院院刊, 2012, 27(6): 5-15.
LI G J, CHENG X Q. Big data research: a major strategic area for future science, technology and economic and social development [J]. Bulletin of Chinese Academy of Sciences, 2012, 27(6): 5-15.
- [2] 王成红, 陈伟能, 张军, 等. 大数据技术与应用中的挑战性科学问题[J]. 中国科学基金, 2014(2): 92-98.
WANG C H, CHEN W N, ZHANG J, et al. Challenging scientific problems for technologies and applications of big data [J]. Bulletin of National Natural Science Foundation of China, 2014(2): 92-98.
- [3] 周季礼, 李德斌. 国外大数据安全发展的主要经验及启示[J]. 信息安全与通信保密, 2015(6): 40-45.
ZHOU J L, LI D B. Experience and enlightenment of foreign big data security development[J]. Information Security and Communications Privacy, 2015(6): 40-45.
- [4] 张海粟, 戴剑伟, 刘玉超, 等. 美国国防部大数据研发核心项目XDATA分析[C]//第三届中国指挥控制大会, 2015年7月7日—9日, 北京, 中国. 北京: 中国指挥与控制学会, 2015: 20-24.
ZHANG H L, DAI J W, LIU Y C, et al. Survey of XDATA: a core big data research program of US DOD[C]// The 3rd China Conference on Command and Control, July 7-9, 2015, Beijing, China. Beijing: Chinese Institute of Command and Control, 2015: 20-24.
- [5] 张尼, 方滨兴. 垃圾邮件过滤技术综述[C]//2005全国网络与信息安全技术研讨会, 2005年8月30-31日, 北京, 中国. 北京: [出版者不详], 2005.

- ZHANG N, FANG B X. Review of spam filtering technology[C]// 2005 National Symposium on Network and Information Security Technology, August 30–31, 2005, Beijing, China. Beijing: [s.n.], 2005.
- [6] 欧德宁. 垃圾邮件过滤技术研究[D]. 济南: 山东大学, 2009.
- OU D N. Research on spam filtering technology[D]. Ji'nan: Shandong University, 2009.
- [7] LI S D, YAN Z, WU X B, et al. A method of emotional analysis of movie based on convolution neural network and bi-directional LSTM RNN[C]// 2017 IEEE Second International Conference on Data Science in Cyberspace, June 26–29, 2017, Shenzhen, China. New Jersey: IEEE Press, 2017.
- [8] NIE Y P, JIA Y, LI S D, et al. Identifying users across social networks based on dynamic core interests[J]. Neurocomputing, 2016(210): 107–115.
- [9] 莫倩, 杨珂. 网络水军识别研究[J]. 软件学报, 2014, 25(7): 1505–1526.
- MO Q, YANG K. Overview of web spammer detection[J]. Journal of Software, 2014, 25(7): 1505–1526.
- [10] BETHENCOURT J, SAHAI A, WATERS B. Ciphertext-policy attribute-based encryption[C]// 2007 IEEE Symposium on Security and Privacy, May 20–23, 2007, Berkeley, USA. New Jersey: IEEE Press, 2007: 321–334.
- [11] GENTRY C. Fully homomorphic encryption using ideal lattices[C]// The 41st Annual ACM Symposium on Theory of Computing, May 31–June 2, 2009, Bethesda, USA. New York: ACM Press, 2009: 169–178.
- [12] STANTON P. Securing data in storage: a review of current research[J]. Computer Science, 2004: 409034.
- [13] 高伟. 磁盘数据安全保护技术研究[D]. 上海: 上海交通大学, 2008.
- GAO W. Research on the security protection data stored on hard disk [D]. Shanghai: Shanghai Jiao Tong University, 2008.
- [14] GOODSON G R, WYLIE J J, GANGER G R, et al. The safety and liveness properties of a protocol family for versatile survivable storage infrastructures[R]. Pittsburgh: Carnegie Mellon University Parallel Data Laboratory, 2004.
- [15] MILLER E L, LONG D D E, FREEMAN W E, et al. Strong security for network-attached storage[C]// Conference on File and Storage Technologies, January 28–30, 2002, Monterey, USA. [S.l.:s.n.], 2002: 1–13.
- [16] WANG W, HUANG H, XIE C, et al. Secure cloud disc: internet certification service on network attached flash file system[J]. Wit Transactions on Information & Communication Technologies, 2014, 52: 1471–1478.
- [17] GANGER G R, KHOSLA P K, BAKKALOGLU M, et al. Survivable storage systems[M]. [S.l.:s.n.], 2001.
- [18] 田洪亮, 张勇, 许信辉, 等. 可信固态硬盘: 大数据安全的新基础[J]. 计算机学报, 2016(1): 154–168.
- TIAN H L, ZHANG Y, XU X H, et al. TrustedSSD: new foundation for big data security[J]. Chinese Journal of Computers, 2016(1): 154–168.
- [19] 魏凯敏, 翁健, 任奎. 大数据安全保护技术综述[J]. 网络与信息安全学报, 2016, 2(4): 1–11.
- WEI K M, WENG J, REN K. Data security and protection techniques in big data: a survey[J]. Chinese Journal of Network and Information Security, 2016, 2(4): 1–11.
- [20] CHASE M. Multi-authority attribute based encryption[M]// Theory of Cryptography. Heidelberg: Springer, 2007: 515–534.
- [21] ZHANG R, MA H, LU Y. Fine-grained access control system based on fully outsourced attribute-based encryption[J]. Journal of Systems and Software, 2017, 125: 344–353.

- [22] BLAZE M, BLEUMER G, STRAUSS M. Divertible protocols and atomic proxy cryptography[J]. Lecture Notes in Computer Science, 1998, 1403: 127–144.
- [23] GE C, SUSILO W, WANG J, et al. Identity-based conditional proxy re-encryption with fine grain policy[J]. Computer Standards & Interfaces, 2017, 52: 1–9.
- [24] SUN Y, LIU D. Cryptanalysis on a secret-sharing based conditional proxy re-encryption scheme[J]. Mobile Networks and Applications, 2017, 22(2): 209–215.
- [25] CANETTI R, RAGHURAMAN S, RICHELSON S, et al. Chosen-ciphertext secure fully homomorphic encryption[C]// IACR International Workshop on Public Key Cryptography, March 28–31, 2017, Amsterdam, Netherlands. Heidelberg: Springer, 2017: 213–240.
- [26] KOGOS K G, FILIPPOVA K S, EPISHKINA A V. Fully homomorphic encryption schemes: the state of the art[C]// 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), February 1–3, 2017, St.Petersburg, Russia. New Jersey: IEEE Press, 2017: 463–466.
- [27] CURTMOLA R, GARAY J, KAMARA S, et al. Searchable symmetric encryption: improved definitions and efficient constructions[C]// ACM Conference on Computer and Communications Security, October 30–November 3, 2006, Alexandria, USA. New York: ACM Press, 2006: 79–88.
- [28] DEMERTZIS I, PAPAMANTHOU C. Fast searchable encryption with tunable locality[C]// The 2017 ACM International Conference on Management of Data, May 14–19, 2017, Chicago, USA. New York: ACM Press, 2017: 1053–1067.
- [29] ROMPAY C V, MOLVA R, ÖNEN M. A leakage-abuse attack against multi-user searchable encryption[J]. Proceedings on Privacy Enhancing Technologies, 2017(3): 164–174.
- [30] ATENIESE G, BURNS R, CURTMOLA R, et al. Provable data possession at untrusted stores[C]// The 14th ACM Conference on Computer and Communications Security, October 29–November 2, Alexandria, USA. New York: ACM Press, 2007: 598–609.
- [31] WANG B, LI B, LI H. Oruta: privacy-preserving public auditing for shared data in the cloud[J]. IEEE Transactions on Cloud Computing, 2014, 2(1): 43–56.
- [32] WANG B, LI B, LI H. Knox: privacy-preserving auditing for shared data with large groups in the cloud[C]// The 10th International Conference on Applied Cryptography and Network Security, June 26–29, 2012, Singapore. [S.l.: s.n.], 2012: 507–525.
- [33] YUAN J, YU S. Efficient public integrity checking for cloud data sharing with multi-user modification[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(8): 1717–1726.
- [34] JIANG T, CHEN X, MA J. Public integrity auditing for shared dynamic cloud data with group user revocation[J]. IEEE Transactions on Computers, 2016, 65(8): 2363–2373.
- [35] WANG Z, HAN Z, LIU J. Public verifiability for shared data in cloud storage with a defense against collusion attacks[J]. Science China Information Sciences, 2016, 59(3): 1–3.
- [36] YUAN J, YU S. Efficient public integrity checking for cloud data sharing with multi-user modification[C]// The 33rd Annual IEEE International Conference on Computer Communications, April 27– May 2, 2014, Toronto, Canada. New Jersey: IEEE Press, 2014: 2121–2129.
- [37] CURTMOLA R, KHAN O, BURNS R. MR-PDP: multiple replica provable data possession[C]// The 28th IEEE International Conference on Distributed Computing Systems, June 17–20, 2008,

- Beijing, China. New Jersey: IEEE Press, 2008: 411-420.
- [38] ZHU Y, WANG H, HU Z, et al. Efficient provable data possession for hybrid clouds[C]// The 17th ACM Conference on Computer and Communications Security, October 4-8, Chicago, USA. New York: ACM Press, 2010: 756-758.
- [39] 冯登国, 张敏, 李昊. 大数据安全与隐私保护[J]. 计算机学报, 2014, 37(1): 246-258.
- FENG D G, ZHANG M, LI H. Big data security and privacy protection[J]. Chinese Journal of Computers, 2014, 37(1): 246-258.
- [40] 李昊, 张敏, 冯登国, 等. 大数据访问控制研究[J]. 计算机学报, 2017, 40(1): 72-91.
- LI H, ZHANG M, FENG D G, et al. Research on access control of big data[J]. Chinese Journal of Computers, 2017, 40(1): 72-91.
- [41] ZENG W, YANG Y, LUO B. Access control for big data using data content[C]// IEEE International Conference on Big Data, October 6-9, 2013, Santa Clara, USA. New Jersey: IEEE Press, 2013: 45-47.
- [42] JASON P O. Horizontal integration: broader access models for realizing information dominance[R]. [S.l.:s.n.], 2004.
- [43] FONG P W L. Relationship-based access control: protection model and policy language[C]// The 1st ACM Conference on Data and Application Security and Privacy, February 21-23, 2011, San Antonio, USA. New York: ACM Press, 2011: 191-202.
- [44] ZENG W, YANG Y, LUO B. Content-based access control: use data content to assist access control for large-scale content-centric databases[C]// 2014 IEEE International Conference on Big Data, October 27-30, 2014, Washington DC, USA. New Jersey: IEEE Press, 2014: 701-710.
- [45] BYUN J W, LI N. Purpose based access control for privacy protection in relational database systems[J]. The International Journal on Very Large Data Bases, 2008, 17(4): 603-619.
- [46] PARK J, NGUYEN D, SANDHU R. A provenance-based access control model[C]// 2012 Tenth Annual International Conference on Privacy, Security and Trust, July 16-18, 2012, Paris, France. New Jersey: IEEE Press, 2012: 137-144.
- [47] SANDHU R S, COYNE E J, FEINSTEIN H L, et al. Role-based access control models[J]. Computer, 1996, 29(2): 38-47.
- [48] WANG L, WIJESEKERA D, JAJODIA S. A logic-based framework for attribute based access control[C]//The 2004 ACM Workshop on Formal Methods in Security Engineering, October 29, 2004, Washington DC, USA. New York: ACM Press, 2004: 45-55.
- [49] HU V C, KUHN D R, FERRAILOLO D F. Attribute-based access control[J]. Computer, 2015, 48(2): 85-88.
- [50] LI J, SQUICCIARINI A, LIN D, et al. SecLoc: securing location-sensitive storage in the cloud[C]// The 20th ACM Symposium on Access Control Models and Technologies, June 1-3, 2015, Vienna, Austria. New York: ACM Press, 2015: 51-61.
- [51] 刘川意, 潘鹤中, 梁露露, 等. 基于小云审大云的云平台可信评测体系结构与技术研究[J]. 网络与信息安全学报, 2016, 2(10): 36-47.
- LIU C Y, PAN H Z, LIANG L L, et al. Cloud trustworthiness evaluation as a cloud service: architecture, key technologies and implementations[J]. Chinese Journal of Network and Information Security, 2016, 2(10): 36-47.
- [52] MUNIER M, LALANNE V, RICARDE M. Self-protecting documents for cloud storage security[C]// IEEE International Conference on Trust, Security and Privacy in Computing and Communications, June 25-27, 2012, Liverpool, UK. New Jersey: IEEE Press, 2012: 1231-1238.

- [53] KROHN M, YIP A, BRODSKY M, et al. Information flow control for standard OS abstractions[C]// The 21st ACM Symposium on Operating Systems Principles, October 14–17, 2007, Washington, USA. New York: ACM Press, 2007.
- [54] PASQUIER T, BACON J, EYERS D. FlowK: information flow control for the cloud[J]. Working Paper, 2014: 70–77.
- [55] PASQUIER J M, POWLES J E. Expressing and enforcing location requirements in the cloud using information flow control[C]// IEEE International Conference on Cloud Engineering, March 9–12, 2015, Tempe, USA. New Jersey: IEEE Press, 2015: 410–415.
- [56] 张大军, 李运发, 郑周. 云计算中数据资源的安全共享机制[J]. 信息网络安全, 2012(8): 79–82.
- ZHANG D J, LI Y F, ZHENG Z. A secure sharing mechanism for data resources in cloud computing[J]. Netinfo Security, 2012(8): 79–82.
- [57] JIA P H, JIN-SONG M A. Research on technical method of hiding watermark in GIS spatial data for sharing security[J]. Bulletin of Surveying & Mapping, 2007(2): 208.
- [58] 张璐, 李晓勇, 马威, 等. 政府大数据安全保护模型研究[J]. 信息网络安全, 2014(5): 63–67.
- ZHANG L, LI X Y, MA W, et al. Research on the security model on big data in government[J]. Netinfo Security, 2014(5): 63–67.
- [59] NARAYAN S, SAFAVI-NAINI R. Privacy preserving EHR system using attribute-based infrastructure[C]// The 2nd ACM Cloud Computing Security Workshop, October 8, 2010, Chicago, USA. New York: ACM Press, 2010: 47–52.
- [60] 雷婉. 权限分离的医疗数据安全共享机制研究与实现[D]. 西安: 西安电子科技大学, 2015.
- LEI W. Research and implementation of medical data security sharing mechanism based on privilege separation[D]. Xi'an: Xidian University, 2015.
- [61] 王璐, 孟小峰. 位置大数据隐私保护研究综述[J]. 软件学报, 2014, 25(4): 693–712.
- WANG L, MENG X F. Location privacy preservation in big data era: a survey[J]. Journal of Software, 2014, 25(4): 693–712.
- [62] CULNAN M J, ARMSTRONG P K. Information privacy concerns, procedural fairness, and impersonal trust: an empirical investigation[J]. Organization Science, 1999, 10(1): 104–115.
- [63] C A S A S - R O M A J, H E R R E R A - J O A N C O M A R T I J, T O R R A V. A summary of k-degree anonymous methods for privacy-preserving on networks[J]. Advanced Research in Data Privacy, 2015, 567: 231–250.
- [64] FUNG B C M, WANG K, FU W C, et al. Anonymity for continuous data publishing[C]// The 11th International Conference on Extending Database Technology, March 25–29, 2008, Nantes, France. New York: ACM Press, 2008: 264–275.
- [65] GORYCZKA S, LI X, FUNG B C M. m-Privacy for collaborative data publishing[J]. International Conference on Collaborative Computing: Networking, 2011, 26(10): 1–10.
- [66] BHATTACHARYYA D K. Decomposition+: improving l-diversity for multiple sensitive attributes[C]// CCSIT 2012, January 2–4, 2012, Bangalore, India. Heidelberg: Springer, 2012: 403–412.
- [67] YI T, SHI M. Privacy protection method for multiple sensitive attributes based on strong rule[J]. Mathematical Problems in Engineering, 2015(5): 464731.
- [68] 方滨兴, 贾焰, 李爱平, 等. 大数据隐私保护技术综述[J]. 大数据, 2016, 2(1): 1–18.
- FANG B X, JIA Y, LI A P, et al. Privacy preservation in big data: a survey[J]. Big Data Research, 2016, 2(1): 1–18.
- [69] 李凤华, 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.
- [70] LEFEVRE K, DEWITT D J, RAMAKRISHNAN R. Mondrian multidimensional k-anonymity[J]. International Conference on Data Engineering, 2006, 6(3): 25.
- [71] MACHANAVAJJHALA A, KIFER D, GEHRKE J. L-diversity: privacy beyond k-anonymity[C]// The 22nd International Conference on Data Engineering, April 3-7, 2006, Atlanta, USA. New Jersey: IEEE Press, 2006.
- [72] ZENG K. Publicly verifiable remote data integrity[J]. Lecture Notes in Computer Science, 2008, 5308: 419-434.
- [73] LIN H, LI T C, VENKATASUBRAMANIAN S. t-Closeness: privacy beyond k-anonymity and l-diversity[C]// The 23rd International Conference on Data Engineering, April 15-20, 2007, Istanbul, Turkey. New Jersey: IEEE Press, 2007: 106-115.
- [74] 张小松, 杨浩森, 汪小芬. 面向大数据安全的密码技术研究[J]. 信息安全研究, 2015, 1(3): 238-244.
- ZHANG X S, YANG H M, WANG X F. Research of cryptography technologies for big data security[J]. Journal of Information Security Research, 2015, 1(3): 238-244.
- [75] CHENG J, FU W C, LIU J. K-isomorphism: privacy preserving network publication against structural attacks[C]// 2010 ACM SIGMOD/PODS Conference, June 6-11, 2010, Indianapolis, USA. New York: ACM Press, 2010: 459-470.
- [76] LV L, ZHOU T. Link prediction in weighted networks: the role of weak ties[J]. Epl, 2010, 89(1):18001.
- [77] CLAUSET A, MOORE C, NEWMAN M E. Hierarchical structure and the prediction of missing links in networks[J]. Nature, 2008, 453(7191): 98.
- [78] WANG P, XU B W, WU Y R, et al. Link prediction in social networks: the state-of-the-art [J]. Science China Information Sciences, 2015, 58(1): 1-38.
- [79] 赵姝, 刘晓曼, 段震, 等. 社交关系挖掘研究综述[J]. 计算机学报, 2017, 40(3): 535-555.
- ZHAO S, LIU X M, DUAN Z, et al. A survey on social ties mining[J]. Chinese Journal of Computers, 2017, 40(3): 535-555.
- [80] 程燕. 大数据时代社交网络隐私保护问题研究[J]. 计算机科学, 2014, 41(10A): 74-78.
- CHENG Y. Research on digital forensic of cloud crime[J]. Computer Science, 2014, 41(10A): 74-78.
- [81] PERLMAN R. File system design with assured delete[C]//IEEE International Security in Storage Workshop, December 13-15, 2005, San Francisco, USA. New Jersey: IEEE Press, 2005: 83-88.
- [82] PERLMAN R. The ephemerizer: making data disappear[R]. Mountain View: Sun Microsystems, Inc., 2005.
- [83] GEAMBASU R, KOHNO T, LEVY A, et al. Vanish: increasing data privacy with self-destructing data[C]//The 18th USENIX Security Symposium, August 10-14, 2009, Montreal, Canada. [S.l.:s.n.], 2009: 299-316.
- [84] YANG T, LEE P P C, LUI J C S, et al. FADE: secure overlay cloud storage with file assured deletion[M]//International Conference on Security and Privacy in Communication Systems. Heidelberg: Springer, 2010.
- [85] TANG Y, LEE P P C, LUI J C S, et al. Secure overlay cloud storage with access control and assured deletion[J]. IEEE Transactions on Dependable & Secure Computer, 2012, 9(6): 903-916.
- [86] CACHIN C, HARALAMBIEV K, HSIAO H C, et al. Policy-based secure deletion[C]// ACM SIGSAC Conference on Computer & Communications Security, November 4-8, 2013, Berlin, Germany. New York: ACM Press, 2013: 259-270.
- [87] XIOMG J B, LIU X M, YAO Z Q, et al. A

- secure data self-destructing scheme in cloud computing[J]. IEEE Transactions on Cloud Computing, 2015, 2(4): 448-458.
- [88] CHEN H S, BHARGAVA B, FU Z C. Multilabels-based scalable access control for big data applications[J]. IEEE Cloud Computing, 2014, 1(3): 65-71.
- [89] 张静, 张洪亮. 基于密码技术的健康医疗大数据安全保障体系研究[J]. 信息安全研究, 2017, 3(7): 652-656.
- ZHANG J, ZHANG H L. Research on security system of healthcare big data based on cryptographic technique[J]. Journal of Information Security Research, 2017, 3(7): 652-656.
- [90] 陶冶, 张云勇, 张尼. 运营商大数据安全防护技术研究[J]. 邮电设计技术, 2014(7): 43-46.
- TAO Y, ZHANG Y Y, ZHANG N. Study on big data protection technology of telecom operators[J]. Designing Techniques of Posts and Telecommunications, 2014(7): 43-46.
- [91] 张滨. 运营商大数据安全保障体系研究[J]. 电信工程技术与标准化, 2016, 29(12): 1-7.
- ZHANG B. Study on big data security system operation[J]. Telecom Engineering Technics and Standardization, 2016, 29(12): 1-7.
- [92] 裴金栋, 赵旺飞. 运营商大数据安全管理策略研究[J]. 移动通信, 2016, 40(21): 5-10.
- PEI J D, ZHAO W F. Research on big data security management strategy for telecom operators[J]. Mobile Communications, 2016, 40(21): 5-10.
- [93] LANG U, SCHREINER R. Managing security in intelligent transport systems[C]// IEEE International Conference on Intelligent Transportation Systems, September 15-18, 2015, Gran Canaria, Canary Islands. New Jersey: IEEE Press, 2015: 48-53.
- [94] 宋珊珊. 智能交通大数据安全挑战[J]. 中国公共安全(学术版), 2015(4): 74-76.
- SONG S S. Security challenges of intelligent transportation big data[J]. China Public Security, 2015(4): 74-76.
- [95] 陈红松, 韩至, 邓淑宁. 智慧城市中大数据安全分析与研究[J]. 信息网络安全, 2015(7): 1-6.
- CHEN H S, HAN Z, DENG S N. Analysis and research on big data security in smart city[J]. Netinfo Security, 2015(7): 1-6.
- [96] 吕欣, 韩晓露, 毕钰, 等. 大数据安全保障框架与评价体系研究[J]. 信息安全研究, 2016, 2(10): 913-919.
- LV X, HAN X L, BI Y, et al. Research on the framework and evaluation system of big data security assurance[J]. Journal of Information Security Research, 2016, 2(10): 913-919.
- [97] 叶润国, 胡影, 韩晓露, 等. 大数据安全标准化研究进展[J]. 信息安全研究, 2016, 2(5): 404-411.
- YE R G, HU Y, HAN X L, et al. Survey and research on big data security standardization[J]. Journal of Information Security Research, 2016, 2(5): 404-411.

作者简介



李树栋 (1979-), 男, 博士, 广州大学副教授, 主要研究方向为网络与信息安全、社交网络分析、云计算安全等。

作者简介



贾焰 (1960-), 女, 博士, 国防科技大学教授, 主要研究方向为大数据、网络信息安全和社交网络。



吴晓波 (1979-), 女, 就职于烟台职业学院信息工程系, 主要研究方向为网络与信息安全。



李爱平 (1974-), 男, 博士, 国防科技大学研究员, 主要研究方向为大数据安全、网络攻防和社交网络分析。



杨小东 (1981-), 男, 博士, 西北师范大学副教授, 主要研究方向为密码学及云计算安全。



赵大伟 (1986-), 男, 博士, 就职于山东省计算中心 (国家超级计算济南中心), 主要研究方向为网络与信息安全、复杂网络。

收稿日期: 2017-08-14

基金项目: 国家自然科学基金资助项目 (No.61672020, No.61662069, No.61472433, No.61702309); 中国博士后基金资助项目 (No.2015T81129, No.2013M542560); 山东省高等学校科技计划资助项目 (No.J16LN61); 国家重点研发计划资助项目 (No.2017YFB0802204)

Foundation Items: The National Natural Science Foundation of China (No.61672020, No.61662069, No.61472433, No.61702309), China Postdoctoral Science Foundation (No.2015T81129, No.2013M542560), Shandong Province Higher Educational Science and Technology Program (No.J16LN61), National Key Research and Development Plan (No.2017YFB0802204)