大数据安全隐私保护研究与认识

摘要

大数据是经济发展的新动能, 社会发展的新引擎, 塑造国家竞争力的战略制高点, 对人民生活 具有重大影响. 然而随着社会对数据价值认知的提升和大数据平台建设的蓬勃发展, 大数据安全问题 日益成为阻碍大数据应用推广的瓶颈. 同时, 由于大数据技术、框架仍在不断演变当中, 研究人员对大数据安全内涵的核心认知和关键特征理解还存在差异, 尚未形成相对统一的大数据安全框架. 当前亟需对大数据安全技术发展现状进行梳理, 为大数据安全重点问题的研究和突破提供参考. 本文结合阐述大数据系统技术框架与安全框架, 结合相关法律法规, 最后总结了大数据安全技术有待解决的核心问题和发展趋势.

**关键词：大数据; 大数据安全; 隐私保护; 信息安全;**

# 1. 引言

当今, 社会信息化和网络化的发展导致数据爆炸式增长. 据统计, 平均每秒有２００万用户在使用谷歌搜索, Facebook用户每天共享的东西超 过４０亿, Twitter每天处理的推特数量超过3.4亿. 同时, 科学计算、医疗卫生、金融、零售业等各行业也有大量产生.

目前大数据的发展仍然面临着许多问题, 安全与隐私问题是人们公认的关键问题之一. 当前, 人们在互联网上的一言一行都掌握在互联网商家手中, 包括购物习惯、好友联络情况、阅读习惯、检索习惯等等, 多项实际案例说明, 即使无害的数据被大量收集后, 也会暴露个人隐私[1] .

图1 大数据下的信息安全

事实上, 大数据安全含义更为广泛, 人们面临的威胁并不仅限于个人隐私泄漏. 与其它信息一样, 大数据在存储、处理、传输 等过程中面临诸多安全风险, 具有数据安全与隐私保护需求. 而实现大数据安全与隐私保护, 较以往其 它安全问题（如云计算中的数据安全等）更为棘手. 这是因为在云计算中, 虽然服务提供商控制了数据的存储与运行环境, 但是用户仍然有些办法保护自己的数据, 例如通过密码学的技术手段实现数据安 全存储与安全计算, 或者通过可信计算方式实现运行环境安全等. 而在大数据的背景下, Facebook等商家既是数据的生产者, 又是数据的存储、管理者和使用者[2] , 因此, 单纯通过技术手段限制商家对用户信息的使用, 实现用户隐私保护是极其困难的事

# 2. 大数据安全技术框架

## 2.1 大数据架构

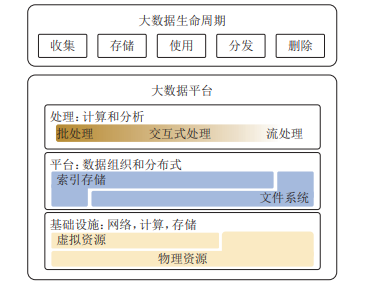
大数据生命周期包含收集、存储、使用、分发和 删除各环节. 数据通过收集进入大数据平台并进行 存储, 通过使用发掘其潜在价值, 通过分发传递和共 享数据或分析结果, 最后删除不再需要的数据[3] . 因此 可以说大数据生命周期是数据转换为价值的过程.

图2 大数据架构

## 2.2 大数据安全体系

2015年, NIST提出一种大数据参考架构, 将大数据系统参与者划分为数据提供者、数据消费者、大数据应用提供者和大数据框架提供者 4 种角色[2] . 其中, 应用提供者执行数据的采集、预处理、分 析、可视化和访问, 框架提供者提供数据的处理、存储框架和基础设施[4] .

NIST 大数据参考架构在国际国内都有较大的影响力, 国际标准化组织/国际电工委员会下的大数据工作组、我国信息技术标准化技术委员会在建立大数据参考架构时都参考了 NIST 所提架构[3] ., 结合大数据业务流程和大数据系统技术框架组成特点, 兼顾数据安全与平 台安全、安全防御与安全管理, 提出符合大数据业务特点的安全技术框架[5] ,查阅资料得知技术框架如下图：

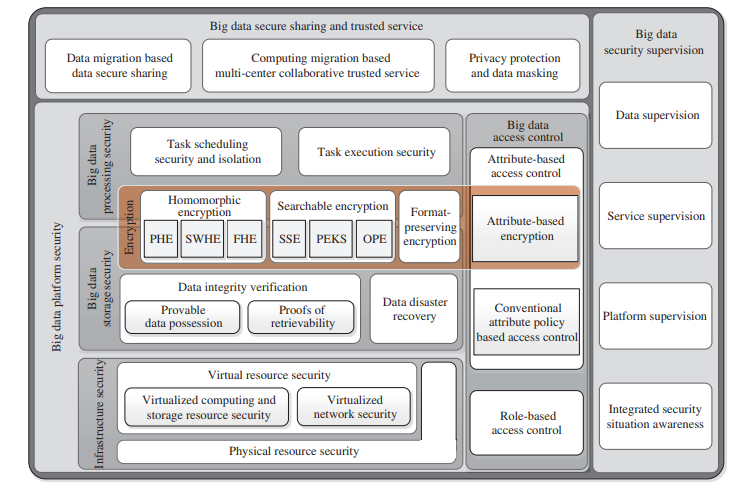


图2 大数据安全技术框架

大数据平台为大数据共享与服务提供数据存储与处理的基础支撑, 负责数据存储、处理和访问等 的实际执行[6] .

大数据平台安全是大数据安全的基础, 主要包括大数据处理安全、大数据存储安全、基 础设施安全和大数据访问控制. 其中, 大数据存储安全是平台安全的重中之重, 并以密码技术作为核心. 同态加密、可搜索加密、保留格式加密和属性加密等密码机制不仅能够提供数据机密性保护, 还能为密态数据的统计、分析、搜索和访问控制等提供支持[7] [8] .

大数据安全监管是大数据安全的保证, 主要解决数据自身、大数据服务和大数据平台安全的监控 与评估等问题, 具体包括数据监管、平台监管、服务监管和综合安全态势感知.

# 3. 相关法律法规

随着大数据的安全问题越来越引起人们的重视, 包括美国、英国、欧盟和中国在内的很多国家和 组织都制定了大数据安全相关的法律法规和政策以 推动大数据应用和数据保护. 2009年, 美国发布《开放政府的指令》要求政府通过网站发布数据等方式公开政府信息[9] ；2012年 5月, 出台《数字政府：构建一个21世纪平台以更好 地服务美国人民》支撑美国电子政府发展[5] . 欧盟早 在1995年就发布了《保护个人享有的与个人数据处 理有关的权利以及个人数据自由流动的指令》（简 称《数据保护指令》）, 为欧盟成员国保护个人数据 设立了最低标准. 2015年, 欧盟通过《通用数据保护 条例》（GDPR）, 该条例在《数据保护指令》的基础 上进行了大刀阔斧的改革, 对欧盟居民的个人信息提出更严的保护标准和更高的保护水平. 巴西、韩国和日本等国家也都发布了《个人信息保护法》[10] , 对 个人信息保护提出明确要求.

# 4. 结论

大数据作为国家基础性战略资源, 已广泛应用 于各重大行业, 其安全问题得到学术界和产业界的 高度重视和积极研究. 本文介绍了大数据安全相关的法律法规和标准现状, 分别从大数据生命周期安全和大数据平台安全两个角度分析目前大数据面临的安全问题, 阐述大数据安全关键技术研究现状及 其开源项目, 最后提出了大数据安全在标准缺口、关 键技术难点和大数据安全分析3个方面的开放问题. 大数据安全技术的发展, 不仅是大数据产业发 展所驱动的结果, 还是国家部署的重要战略[11] . 加强研 究大数据安全保护技术, 可推动大数据的开放共享, 有力支撑大数据产业的持续发展, 更加增强国家网 络空间安全的防御能力.

# **参考文献**

* + 1. 丁红发,孟秋晴,王祥,蒋合领.面向数据生命周期的政府数据开放的数据安全与隐私保护对策分析[J].情报杂志,2019,38(07):151-159.
    2. 张佳乐,赵彦超,陈兵,胡峰,朱琨.边缘计算数据安全与隐私保护研究综述[J].通信学报,2018,39(03):1-21.
    3. 陈兴蜀,杨露,罗永刚.大数据安全保护技术[J].工程科学与技术,2017,49(05):1-12.
    4. 黄永刚.基于区块链技术的电子健康档案安全建设[J].中华医学图书情报杂志,2016,25(10):38-40+46.
    5. 何晓琳,钱庆,吴思竹,张泽,孙小康.健康医疗可穿戴设备数据安全与隐私研究进展[J].中华医学图书情报杂志,2016,25(10):32-37.
    6. 曹珍富,董晓蕾,周俊,沈佳辰,宁建廷,巩俊卿.大数据安全与隐私保护研究进展[J].计算机研究与发展,2016,53(10):2137-2151.
    7. 李昊,张敏,冯登国,惠榛.大数据访问控制研究[J].计算机学报,2017,40(01):72-91.
    8. 吕欣,韩晓露.大数据安全和隐私保护技术架构研究[J].信息安全研究,2016,2(03):244-250.
    9. 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(01):246-258.
    10. 钟瑛,张恒山.大数据的缘起、冲击及其应对[J].现代传播(中国传媒大学学报),2013,35(07):104-109.
    11. 胡艳.云计算数据安全与隐私保护[J].科技通报,2013,29(02):212-214.