

大数据技术前瞻

梅宏¹, 杜小勇², 金海³, 程学旗^{4,5}, 柴云鹏², 石宣化³, 靳小龙^{4,5}, 王亚沙¹, 刘驰⁶

1. 高可信软件技术教育部重点实验室(北京大学), 北京 100871;
2. 中国人民大学信息学院, 北京 100872;
3. 华中科技大学计算机科学与技术学院, 湖北 武汉 430074;
4. 中国科学院计算技术研究所, 北京 100086;
5. 中国科学院大学计算机科学与技术学院, 北京 100049;
6. 北京理工大学计算机学院, 北京 100081

摘要

世界主要国家高度重视大数据发展, 我国也将发展大数据作为国家战略, 发展大数据技术具有重要意义。大数据技术涉及从采集、传输到管理、处理、分析、应用的全生命周期以及生命周期各阶段的数据治理。选取数据生命周期中的管理、处理和分析技术以及大数据治理技术来梳理国内外技术发展现状, 特别是研判我国大数据技术发展与国际先进技术之间的差距。另外, 在大数据应用需求的驱动下, 计算技术体系正面临重构, 从“以计算为中心”向“以数据为中心”转型, 在新的计算技术体系下, 一系列基础理论和核心技术问题亟待破解, 新型大数据系统技术成为重要发展方向。在计算体系重构的背景下, 提出大数据技术发展的四大技术挑战和十大发展趋势。

关键词

大数据技术; 大数据管理; 大数据处理; 大数据分析; 大数据治理

中图分类号: TP399

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2023009

Big data technologies forward-looking

MEI Hong¹, DU Xiaoyong², JIN Hai³, CHENG Xueqi^{4,5}, CHAI Yunpeng², SHI Xuanhua³,
JIN Xiaolong^{4,5}, WANG Yasha¹, LIU Chi⁶

1. Key Lab of High Confidence Software Technologies (Peking University), Ministry of Education, Beijing 100871, China
2. School of Information, Renmin University of China, Beijing 100872, China
3. School of Computer Science & Technology, Huazhong University of Science and Technology, Wuhan 430074, China
4. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100086, China
5. School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China
6. School of Computer Science & Technology, Beijing Institute of Technology, Beijing 100081, China

Abstract

Major countries in the world attach great importance to the development of big data technology. China also puts big data as a national strategy, of great significance to develop in the long run. Big data technologies include data collection, transmission, management, processing, analysis, and application, forming a data life cycle as well as the data governance

2023009-1

related to each procedure. Big data management, processing, analysis, and governance in four areas were selected, to identify the gap between China and the world. On the other hand, driven by diverse successful big data applications, the system architecture of computing technology is being restructured. From “computation-centric” to “data-centric”, fundamental computing theories and core technologies need to be redesigned, therefore a new type of big data system technology is becoming an important research direction. Against this background, four technical challenges and ten future development trends of big data technologies were aimed at identifying.

Key words

big data technology, big data management, big data processing, big data analytics, big data governance

0 引言

世界主要国家高度重视大数据发展,我国也将发展大数据作为国家战略。2015年8月,国务院印发了《促进大数据发展行动纲要》(国发〔2015〕50号),明确提出“数据已成为国家基础性战略资源”。2016年3月,《中华人民共和国国民经济和社会发展第十三个五年规划纲要》正式提出“实施国家大数据战略”。2017年10月,党的十九大报告中提出加快推进大数据与实体经济深度融合发展。2017年12月8日,在中共中央政治局第二次集体学习时,习近平总书记发表讲话,做出了“大数据是信息化发展的新阶段”这一重要论断,要求“审时度势精心谋划超前布局力争主动,实施国家大数据战略,加快建设数字中国”。党的二十大报告指出数字技术与实体经济深度融合具有重大意义。

大数据技术链条长,从采集、传输到管理、处理、分析、应用,形成一个完整的数据生命周期,同时在多个阶段涉及数据治理。由于数据采集用到的传感器技术和数据传输依赖的通信技术分别属于基础器件和通信学科领域,本文不涉及这两类技术。同时,大数据应用面比较广,应用发展状态不是技术问题,故本文也不涉及。另外,虽然大数据的发展已经经历了10多个

年头,在一些应用领域(特别是互联网领域)取得了较好的成效,但是总体而言,大数据基础理论和核心技术仍不成熟,大数据治理体系远未建立,大数据发展仍然处于初级阶段。因此,为了更好地发展大数据技术,本文聚焦数据生命周期中大数据管理、处理、分析和治理4类技术,梳理国内外技术发展现状,特别是亟须清楚研判我国大数据技术发展与国际先进技术之间的差距。此外,信息技术发展的主流一直是以计算为中心的,数据仅作为输入和输出围绕着计算任务组织,信息系统设计和优化的核心目标是计算效能的提升。随着数据体量的快速增长,以计算为中心的技术体系开始显现出弊端,算力增长难以跟上数据规模的增长,形成“剪刀差”。近年来,一种以数据为中心的新技术体系日益受到关注,有望成为缩小“剪刀差”的突破点。一些具有潜力的创新技术已经出现了,包括:在算法理论体系上,从由多项式精确算法主导演变为由亚线性概率近似算法主导;在大数据方法体系上,从经验沉淀驱动演变为数据模型驱动,从单纯依赖机器发展到“人在回路”的协作计算模式;在大数据系统结构上,通过以数据为中心的泛在操作系统,围绕着数据间的互操作,高效组织广域计算资源等。因此,在大数据应用需求的驱动下,计算技术体系正面临重构,从“以计算为中心”向“以数据为中心”转型。在新的计算技术体系下,一些基

础理论和核心技术问题亟待破解,新型大数据系统技术成为重要发展方向,同时面临以下四大挑战。

挑战一:如何构建以数据为中心的计算体系。全球大数据规模增长快速。2020年全球新增数据规模为64 ZB,是2016年的400%,预计2035年新增数据将高达2 140 ZB,数据量呈现指数级增长。随着数字经济的发展和数字化转型的深入,愈来愈多的数据资源正以数据要素的形态独立存在,并参与数字经济活动的全过程。因此需要构建以数据为中心的新型计算体系,以适应新的应用环境。如何组织和管理超大规模的数据要素已经成为一项难题,如大数据管理面临着数据跨境访问带来的各种问题、系统规模持续增大带来的可用性下降、维护大规模数据带来的成本和能耗持续增高等严峻挑战。

挑战二:如何满足大数据高效处理的需求。数据规模呈指数级增长,数据动态倾斜、稀疏关联、应用复杂,传统大数据处理架构处理成本高、时效性差,如何满足规模海量、格式复杂、需求多变的大数据高效处理需求是大数据处理面临的重要挑战。

挑战三:如何实现多源异构大数据的可解释性分析。随着数据量持续地爆炸式增长和各类应用的不断拓展与深化,基于深度学习的主流方法因其仅关注单源单模态数据且模型只知其然不知其所以然的特性,已无法满足发展需求。如何打破数据多源异构造成的隔阂,融合多域甚至全域数据中蕴含的知识,实现分析结果的可解释,从而提升其可用性,是当前大数据分析面临的主要挑战。

挑战四:如何形成系统化大数据治理框架与关键技术。针对大数据应用过程中的对数据汇聚融合、质量保障、开放流通、

标准化和生态系统建设的需求,大数据治理技术逐渐成为发展热点。然而当前系统化的大数据治理框架尚未形成,开放共享、质量评估、价值预测等关键技术远未成熟,这成为制约大数据发展的主要瓶颈。

本文在计算体系重构的背景下,对大数据管理、处理、分析和治理4个方面的主要技术挑战、国内外研究现状与差距做出研判,并提出大数据技术发展的十大趋势。

1 大数据管理技术

1.1 主要技术挑战

大数据管理是大数据生命周期中的重要环节。大数据管理技术主要包括大数据的组织、存储、计算与访问等相关技术。尽管数据管理技术与系统已经经历了近60年的发展,但由于大数据具有超大规模、高速增长、类型多样、高复杂性等特点,大数据管理仍然面临诸多严峻挑战,特别表现在以下3个方面。

(1) 高性能挑战

一直以来,追求数据访问的高性能是数据库系统的核心问题。随着摩尔定律接近终结,传统硬件性能提升乏力。在数据规模高速增长的情况下,大数据管理遇到的高性能挑战更加突出,迫切需要寻找性能提升的新驱动力,尤其是要充分利用近年来发展迅速的各种新硬件和人工智能技术,最大限度发挥异构架构带来的机会。

(2) 可用性挑战

数据管理系统是大数据产业的核心基础设施,因此其可用性至关重要。大数据规模呈现指数级高速增长,大数据业务环境也非常复杂,大数据管理系统往往规模极大且具有很高的复杂度,因此各类错

误发生的概率显著提升,可能导致服务中断、无法保证正确性,甚至可能存在恶意节点引发虚假消息或数据篡改等严重问题。这给大数据管理系统的可用性带来更大的挑战。

(3) 能效挑战

目前维护和管理超大规模数据的能耗占比已经很高。现有技术往往以性能提升为导向,能效不高,因此面向快速增长的大数据,迫切需要发展和应用高效大数据管理技术。

1.2 国内外研究现状与差距比较

新一代大数据管理技术的创新发展主要围绕高性能、高可用、高效能3个方面展开。

在高性能大数据管理领域,新的性能提升动力主要包括新硬件和人工智能驱动两个途径。在新硬件方面,美国在内存数据库并发控制、非易失存储器(non-volatile memory, NVM)数据管理、GPU数据库、异构计算数据库等方面的技术和产品均处于领先地位^[11],德国的SAP HANA是高性能数据库的代表之一^[12]。我国在NVM、远程直接内存访问(remote direct memory access, RDMA)、超文本标记语言(hyper text mark-up language, HTML)等新硬件数据管理技术方面也提出一系列先进的创新成果,如上海交通大学的RMDA和HTML结合方案^[13],以及中国科学院计算技术研究所的NVM日志与恢复方法^[14]等。在人工智能驱动的高性能大数据管理方面,美国率先提出学习索引^[15]、Self-Design^[16]等技术,微软^[17]、甲骨文^[18]、亚马逊等企业已在数据库产品中使用智能驱动技术。我国在学习索引、自动调优、智能查询优化、智能数据分区等领域的创新也非常活跃,如阿里巴巴^[9]、华为^[10]、

腾讯等企业的数据库产品加入了数据自动调优、自动运维等自治技术,清华大学等高校提出一系列AI4DB创新技术^[11-12]。

在高可用大数据管理领域,数据冗余是应对软硬件错误的主要手段,核心技术是分布式共识和分布式事务处理,难点在于同时实现高可用和高性能^[13]。美国是该方向的引领者,核心算法和最早的系统都来自美国^[14-15]。国内阿里巴巴、清华大学、上海交通大学、中国人民大学等在核心的分布式事务和共识协同优化技术等方面取得了系列创新成果^[16-17]。国产数据库PolarDB、TiDB等可用性水平已达到国际一流水准^[16,18]。另一类可用性问题是网络中部分节点存在恶意,如发送虚假消息或篡改数据等。目前国际主流应对技术是区块链,主要发展趋势是通过公链和联盟链相互融合、适度降低安全性要求等方式提升区块链的性能^[19-20]。我国的区块链大数据管理技术发展和应用已接近国际一流水平,国产联盟链系统(蚂蚁链、腾讯区块链和百度超级链等)每秒交易量均超过1万,并在区块链发票、区块链司法存证、区块链物流溯源等领域广泛应用。

在高效能大数据管理领域,云数据管理技术通过资源共享、资源调度来显著减少硬件资源和能源的消耗,并使用低功耗硬件和数据压缩等软件方法进一步降低能耗。美国的亚马逊和Snowflake分别是联机事务处理(on-line transaction processing, OLTP)和联机分析处理(on-line analysis processing, OLAP)领域云数据库的全球引领者;美国亚马逊、谷歌、微软、Meta等公司率先探索基于异构硬件和压缩的低功耗数据管理技术^[21]。而国内的高能效云数据库技术发展非常快,已与美国技术水平接近,如阿里云的PolarDB率先提出分离内存技术^[22],中国人民大学等高校也通过异构计算和压缩数据直接

计算方式构建了新型高能效数据管理引擎^[23]。近年来,国内外在云数据管理的基础上探索国家范围内的一体化高能效数据管理,例如美国苹果、谷歌等公司通过智能编排,将数据智能地在多个数据中心进行分布和计算,降低总体能耗^[24]。我国提出了算力网络的概念和国际标准,正式启动“东数西算”工程,充分利用中西部地区的气候、能源等优势,在全国布局算力网络国家枢纽节点,并逐步在人工智能、图像渲染、金融和政企业务等领域应用^[25]。我国在高能效一体化数据管理技术方面走在世界前列。

综上所述,近年来我国大数据管理技术和产品的发展都非常快,如在联机事务处理基准测试TPC-C排行榜中,OceanBase以707 351 007 tpmC的性能排名世界第一,打破了甲骨文等国外公司长期垄断的局面。我国大数据管理技术与国外顶尖水平(绝大部分情况指美国)虽然存在一定差距,但大部分领域的差距并不大,具备赶超的机会;在高能效一体化大数据管理等领域,甚至有局部领先。

1.3 小结

美国在新一代大数据管理技术方面仍处于全球领先和主导地位;欧盟在部分领域的产品和创新技术上有一定优势;日本、俄罗斯等其他国家相对而言在技术、产品和生态上均缺乏优势。而我国大数据管理技术进步非常快,在我国大规模市场的培育下,头部企业的产品能力和顶尖高校的创新能力已经达到或接近国际一流水准,明显超越日本、俄罗斯等国家一些领域的技术创新也领先于欧洲,与美国的差距在迅速减小;尤其是在高能效一体化大数据管理等领域,我国已经在国际上抢先一步发展。

2 大数据处理技术

2.1 主要技术挑战

过去10年,数据规模呈指数级增长,数据处理的时效性问题成为大数据处理系统面临的核心问题。同时数据应用蓬勃发展,数据深度价值挖掘、数据实时处理等新型处理需求进一步提高了数据处理复杂度,大规模数据处理系统中数据动态倾斜、稀疏关联、超大容量等特征给系统带来资源效率低、时空开销大、扩展困难等严重问题。作为大数据领域典型关联关系的图数据,由于其不规则数据访问、计算-访存比小、依赖关系复杂等特点,给现有大数据处理架构带来了并行流水执行效率低、访存局部性低、内外存通道利用率低和锁同步开销大等技术挑战。

2.2 国内外研究现状与差距比较

大数据处理通常有两种方式,一种为离线处理模式,另一种为在线处理模式,与之对应的系统为批处理系统和流处理系统。同时作为互联网以及科学计算领域重要的数据形态,图数据因其结构特殊,多采用专用系统处理。本节将从批处理、流处理以及图计算3个方面进行比较。

(1) 批处理

作为最早的大数据处理模式,从谷歌公司提出MapReduce^[26]模型开始,美国一直引领该领域的发展。近10年,从开源Apache Hadoop到美国加州大学伯克利分校的AMPLab研发的Spark系统,在技术

生态上完全诠释了大数据处理从“扩展性优先”设计到“性能优先”设计的过渡,大幅提升了大数据的处理性能,同时也为更多类型的大数据处理业务(如数据挖掘、机器学习等)提供了支持。为了更好地支撑分布式AI应用,同样由美国加州大学伯克利分校领导的AnyScale公司提出了Ray框架,通过动态定制计算依赖,取得了比Spark更优异的计算性能。另外,异构计算成为大数据处理系统时效性提升的重要手段。美国NVIDIA公司牵头针对不同应用领域,开源了RAPIDS GPU数据科学库^[27-28],存内计算^[29]之类的近数据处理方法在体系结构领域获得广泛关注。

我国过去10年在大数据处理系统领域也取得了巨大进展,尤其以大型互联网企业为代表,如阿里巴巴自研的MaxCompute引擎可提供高效的数据处理功能,在大型互联网企业的数据仓库和BI分析、网站的日志分析、电子商务网站的交易分析、用户特征和兴趣挖掘等领域获得较好应用。此外,阿里云基于Flink开发了Blink系统,实现了流处理和批处理的统一,在批流融合相关技术指标方面达到了国际领先水平。在学术界,国内在大数据处理的单点技术突破上也取得了系列进展,如上海交通大学的Espresso^[30]、华中科技大学的Deca系统^[31],性能大幅领先国际同类系统。目前我国的大多数大数据批处理系统还是建立在国际开源的基础之上的,自研系统国际占有率非常有限,整体技术水平仍有较大差距,生态差距更大。

(2) 流处理

国际上现有流处理系统按体系结构划分主要可分为并行流处理系统^[32]、分布式流处理系统^[33]。采用分布式集群架构的分布式流处理系统成为当前国际上流处理系统的主流,近年来发展出围绕Storm^[33]、Flink、Spark Streaming^[34]等开源系统的

软件生态。为了提高分布式流处理系统的性能,现有工作主要采用数据并行^[35-36]、流水线并行^[37]等技术提升系统性能。数据并行主要充分利用单个计算节点上的多核并行资源,对部署到某个计算节点的具体计算阶段进行多实例并行化,从而充分发挥多核资源的并行性,提高系统的吞吐率。数据并行优化方面的代表性工作是Nasir M A U等人^[35-36]提出的PKG (partial key grouping)数据划分策略,该策略解决了系统由动态倾斜性带来的负载不均问题,提升了系统的资源效率,并被Apache的Storm开源项目采纳集成。流水线并行的主要思想是将流处理的计算逻辑分解为多个阶段,将不同的计算阶段部署到分布式集群中的不同计算节点上,通过流水线技术提高系统资源并行效率,从而提升系统的整体性能。流水线并行方面的代表性工作是Abdelhamid A S等人^[37]提出的Prompt系统。

国内流处理系统方面主要基于现有国际开源的相关工作进行优化改进。阿里巴巴通过收购德国创业公司Data Artisans,购入了Apache Flink,并进行了若干优化改进,目前绝大部分阿里巴巴的业务跑在流处理平台上。数据并行方面的典型系统工作是PStream^[38],该系统基于Apache Storm平台对高频的键值使用轮询划分的方式平衡负载,同时对低频的键值使用哈希划分的方式避免额外的数据聚合开销。相比于国际上最新的PKG系统,PStream将系统吞吐率提升了2.3倍,处理时延缩短了64%。PStream在华为公司有所应用,但在开源社区的推广方面不及PKG系统。在流水线并行方面,目前国内研究较少。同时,由于现有系统多基于国际开源软件架构,国内的流处理系统研究也承袭了现有国际开源框架的弊端,如难以针对多语义查询进行灵活部署等。

(3) 图计算

图计算技术是由谷歌公司首次提出的。为了支持分布式图计算,谷歌公司设计了国际上首个分布式图计算系统Pregel^[39]。该系统将迭代图算法表示为简单编程的多次迭代,以简化分布式图计算应用的开发和在底层分布式平台上的高效执行。后续国外研究团队也提出了诸多软件和硬件优化技术^[40-44]来提高图计算性能。例如,为了支持PC上的图计算,国外研究团队研发了首个单机图计算系统GraphChi^[45],其采用并行滑动窗口技术,大幅降低外存的乱序访问次数。普林斯顿大学的研究人员研发了首个面向图计算的专用加速器Graphicionado^[46],其将以顶点为中心的编程模型中的功能模块抽象为相应的流水线阶段,在保证通用性的同时,显著提升了图计算指令效率。

相比而言,国内研究团队在图计算领域起步略晚,但随着国内互联网行业和技术迅猛发展,国内图计算市场需求日益高涨。“十三五”期间,我国科技部启动了“面向图计算的通用计算机技术与系统”项目。国内一批研究人员在图计算领域进行了深入研究和大量攻关工作,提出了多种高性能图计算系统^[47-49]和图计算硬件加速器^[50-52]。清华大学研究团队开发的基于神威·太湖之光的超大规模图计算系统“神图”系统入围国际超算大会戈登贝尔奖^[53]。华中科技大学研发的DepGraph^[54]斩获2021年11月国际Green Graph 500和Graph 500榜单两项全球第一,基于鹏城云脑II系统研发的图计算系统再次获得2022年11月Graph 500 SSSP性能第一。随着国内图计算技术的迅猛发展,国内图计算研发能力在图计算的理论研究、系统研发和体系结构设计方面均已达到世界前列。

2.3 小结

作为大数据技术生态的重要一环,大数据处理技术在过去10年飞速发展,尤其因为其巨大的产业需求,国际、国内大型企业均投入了大量人力、物力参与研发,技术生态上依旧蓬勃。我国在该领域发展迅速,但后发特征明显,单点技术突破较好,整体生态与欧美差距较大,缺乏生态引领系统出现。在批处理系统、流处理系统上,我国基本以国外开源跟踪改进为主,图处理技术由于与国际发展时间基本同步,国内有部分系统与国际同类系统处于同一技术水平。

3 大数据分析技术

3.1 主要技术挑战

大数据分析技术旨在实现从数据到信息再到知识甚至到决策的价值转换。近年来,大数据分析技术迅猛发展,性能和效率均取得了显著的提升,并促进了相关行业或产业的智能化发展。例如,基于深度学习的蛋白质折叠分析技术帮助生物学取得了重大进展;金融数据分析技术帮助银行业大幅降低了金融欺诈的风险。然而,广泛深入的应用在给大数据分析技术的发展带来更高的需求的同时,也使其面临着更深层的挑战。首先,大数据分析的对象逐渐从相对小规模、单模态的数据转变为大体量、多模态的数据,如何对大规模异构多模态数据进行融合分析?其次,单一来源的数据往往体量较小或统计分布的代表性不够,相应分析模型的性能受到制约,如何在保护数据安全与隐私的前提下进行多

方数据的联合学习与分析从而实现数据价值的最大化?最后,诸多应用不再满足于对大数据相关性的简单建模与分析,更期待能够挖掘现象背后的因果规律,如何通过因果推断满足分析技术在可解释性、稳定性、公平性以及可回溯性等方面的更高需求?

3.2 国内外研究现状与差距比较

针对前述3个方面的挑战,本节对国内外相应的研究现状进行了梳理,并比较了差距。

随着信息技术的飞速发展,多模态数据已成为数据资源的主要形式,国内外研究者近些年对多模态数据融合分析给予了高度关注,也取得了显著研究成效。当前,国际上多模态技术的顶尖研究团队有卡内基梅隆大学的MultiComp Lab和麻省理工学院的Synthetic Intelligence Lab等。这些团队在AAAI、ICLR等相关领域的国际顶会上发表了诸多突破性成果^[55]。除此之外,微软2022年提出通用多模态基础模型BEiT-3^[56],它在视觉和视觉-语言任务上都实现了当时最先进的迁移性能。同年,谷歌提出了多模态大模型PaLI^[57],在多语言图文数据上进行训练,效果超过了BEiT-3。目前,多模态技术已在军事、交通等领域得到了越来越深入的应用。例如,美国桑迪亚国家实验室提出了多模态的军事概念装备。紧跟国外研究的步伐,国内研究团队也加快了对多模态数据分析技术的攻关,并取得了一些突破性成果。例如,2021年,中国科学院自动化研究所研发了全球首个三模态大模型“紫东太初”^[58],并在2022年的世界人工智能大会上获得最高奖项。在应用方面,国内将多模态技术与商业、医疗等领域进行了有效结合。例如,地平线的多模语音算法团队在2020年长安UNI-T车型上实现大规模量产,是全球首个能够在端上实时预测并实现大规模量产的团队。综合来看,国内的多模

态技术在基础理论和核心技术方面与国际先进水平还存在着一定的差距,但是与具体应用领域结合较好,有效实现了科技成果的落地转化。

随着信息化进程的发展,各个企业或同一企业的不同部门生产并存储了大量应用数据,跨部门、跨机构的数据流通与共享能够更大程度地挖掘数据的潜在价值。然而,数据的流通共享受到数据安全、商业机密、个人隐私等多方面的严格约束。为了应对上述问题,2016年谷歌提出联邦学习(federated learning)的概念^[59],用于解决多方数据联合学习与分析的挑战,并推出TensorFlow Federated联邦学习开源框架,引起学术界和工业界的强烈关注。除此之外,Open Mind推出的Pysyft联邦学习框架^[60]能够与主流深度学习框架兼容,热度居高不下。目前,联邦学习已经得到一定程度的领域应用。苹果公司将联邦学习应用在IOS 13跨设备QuickType键盘“Hey Siri”的人声分类上;NVIDIA公司推出了NVIDIA Clara医疗学习平台^[61],将联邦学习技术应用在医疗领域。在国内,微众银行AI团队2018年系统性地阐述了联邦学习理论,其能够保证各企业在自有数据不出本地、不违规的前提下进行联合建模,大幅提升了机器学习建模的效果,之后微众银行AI团队又于2019年开源了首个工业级联邦学习技术框架FATE^[62],受到广泛关注。此外,近几年,百度推出了开源联邦学习框架PaddleFL,阿里巴巴利用联邦学习等隐私计算技术推出了DataTrust平台,二者均在自然语言处理和推荐算法等领域进行了落地实践。综合来看,我国相关团队与企业已成为联邦学习技术的重要贡献者,积极参与到联邦学习技术的研发与标准的制定中。

现有大数据分析技术在稳定性、可解释性、公平性、可回溯性等方面存在着天

然不足,主要原因是现有技术往往只关注变量统计意义上的相关性分析与建模,而这其中势必存在不符合变量因果规律的伪相关。因果推断技术旨在发现变量之间的因果规律。现有最主流的因果模型为潜在结果框架^[63]和结构因果模型^[64],两个模型分别由美国科学家Rubin和美国以色列裔科学家Pearl J等人提出。他们都认为因果关系指的是两个事物之间改变一个是否能够影响另一个的关系。Pearl J等人^[65]提出了“因果之梯”的概念,自下而上将问题划分为关联、干预和反事实3个层次,分别对应于观察、行动和想象3类活动。而回答因果问题需要借助干预或者反事实。基于这两个主流的因果模型,微软^[66]等外国公司在积极探索因果相关的研究。国内因果相关研究起步较晚,但发展较快。清华大学Cui P等人^[67]提出了稳定学习的概念,在传统机器学习框架下通过引入因果推断技术,寻求数据中的“不变性机制”。中国科学技术大学Zhang Y等人^[68]关注因果推荐系统相关的研究,尝试挖掘用户与商品之间的因果关系。南京大学Zhu Z M等人^[69]则关注因果强化学习的相关研究。综合来看,因果推断技术国外起步较早,并开创了奠基性的框架。国内研究虽然起步较晚,但相关研究的发展呈良好态势。

3.3 小结

从上述分析可以看出,面向大数据分析前沿需求,以美国为代表的西方国家起步较早,在多模态数据融合技术、联邦学习技术以及因果推断技术的基础理论和核心技术上已经形成了比较完善的分析和应用生态,且在诸多领域得到了广泛应用。国内大数据分析技术在基础理论和核心技术等方面不断靠近国际先进水平,保持着快速发展的良好势头。

4 大数据治理技术

4.1 主要技术挑战

在应用蓬勃发展的过程中,大数据面临数据产权、交易流通、收益分配等一系列问题,从而使大数据治理逐渐成为各国关注的热点,相关技术也发展迅速。大数据治理是以数据为对象,以最大限度释放数据价值为目的,在确保数据安全的前提下,贯穿于数据全生命周期的由多方主体参与的共建共享共治的数据价值释放的过程。大数据治理技术面临的主要挑战包括5个方面:一是汇聚融合,如何将海量、多源、分散、异构的数据汇聚融合成便于分析利用的数据资源;二是质量保障,如何评估大数据的质量、检测及修正其中存在的错漏,提升数据质量;三是开放流通,如何打通数据孤岛,使数据可以突破原生信息系统的边界,通过开放流通形成更大范围、更高价值的数据库资源;四是数据的安全与隐私保护,如何在释放数据价值的同时保证数据不被破坏、泄露和滥用;五是标准化与生态系统建设,如何协调各利益方的诉求,建立共同遵循的数据标准体系,并促进形成大数据软硬件生态系统等。

4.2 国内外研究现状与差距比较

在大数据汇聚融合方面,图灵奖获得者Michael Stonebraker教授的研究最具代表性,其将大数据融合处理相关研究分为3代^[70],分别解决传统数据仓库场景(十余个数据源)、跨组织数据共享场景(数十个数据源)和数据湖场景(数百个

数据源)的数据融合问题。相应的技术路线也存在显著的差异,从第一代基于规则的ETL系统,到第二代基于机器学习的Data Curation系统,再发展到第三代机器驱动、人在回路的智能融合系统。清华大学围绕物联网场景,主导研发了Apache IoTDB,在国内外产生了一定的影响力;阿里巴巴、华为、京东等企业提出了建设“数据中台”^[71],大数据汇聚融合作为数据中台的核心技术能力,在实践中得到了创新发展。

在大数据质量保障方面,ACM/IEEE Fellow加拿大滑铁卢大学教授Ihab Ilyas研发了HoloClean系统^[72],该系统主要解决关系数据的错误检测和修复问题,在城市、医疗等领域的真实数据集上将数据质量提升了1~2倍。谷歌公司构建了Knowledge Vault系统,提出了知识融合方法,解决了万维网事实抽取过程中的数据质量问题,构建了规模远大于现有开源知识图谱的结构化知识库^[73]。清华大学的相关学者提出了人在回路的数据质量提升方法,提出在算法的回路中优化地引入人的识别与推理能力,在提升数据质量的同时,有效地控制人工参与的成本^[74]。哈尔滨工业大学的相关学者提出了一系列的数据清洗方法,解决了工业时序数据场景下的错误检测与修复问题,显著地提升了数据质量^[75]。在工业界,阿里巴巴公司推出了DataWorks全链路数据治理工具,华为公司推出数智融合系统,这些系统主要针对企业数据仓库和数据湖的真实场景,通过数据建模、数据集成、数据血缘等技术手段,支持面向数据全生命周期的质量保障能力。

在大数据开放流通方面,数联网成为互联网之上实现可信可管可控的数据互联互通和大数据应用的核心技术。数联网是基于软件定义的,将各种异构数据平台和系统连接起来,在“物理/机器”互联网

之上形成的“虚拟/数据”网络,它正在世界范围内发展成为大数据时代的一种新型信息基础设施。当前最有影响力的数联网技术路线是互联网发明人、图灵奖得主Robert Kahn提出的数字对象架构(digital object architecture, DOA)^[76],美国、英国、德国、俄罗斯、中国等国是DOA标识解析系统Handle全球根节点的参与者。北京大学自主研发的“黑盒式”互操作技术及燕云DaaS系统^[77],提出颠覆式的数据互操作技术途径——“黑盒”思路,消除了系统源码、数据库表、后台权限、原开发团队等“白盒”依赖,信息孤岛开放效率得到大幅提升。在此基础上,2018—2021年,北京大学、清华大学、中国科学院等,在科技部“云计算与大数据”国家重点研发计划专项的支持下提出了数联网中国云方案。中国信息通信研究院于2018年开始建设兼容DOA的国家工业互联网标识解析系统。我国以DOA为代表的数联网系统软件方面取得了国际先进的技术成果^[78]。

在大数据安全与隐私保护方面,同态加密允许数据分析处理过程中,直接在密文数据上进行计算而无须事先解密,很好地解决了大数据价值利用与数据安全和隐私保护的矛盾,近年来受到各界高度重视,从学术研究逐渐进入实际应用。2017年IBM、微软、Intel、NIST的研究者主导创立了同态加密标注化委员会,致力于研制同态加密安全性、API和应用的标准。安全多方计算在保持各方数据隐私的同时,使多方可以合作完成某个共同的计算目标。此项技术是密码学的分支领域,在大数据应用的推动下逐渐成为热点,特别是针对机器学习场景,谷歌率先提出联邦学习概念,近年来得到快速发展。2020年美国Meta、日本NTT、中国阿里巴巴等众多企业共同创建了安全多方计算联盟MPC Alliance,旨在推进安全多方计算的认知、接受和采用。

当前谷歌、OpenMind、LatticeX基金会,以及我国的百度、字节跳动等公司推出了各自的开源联邦学习框架,该领域呈现出百家争鸣的态势。

在标准化与生态系统建设方面,ISO/IEC JTC 1、ITU-T、IEEE等国际各大标准组织积极制定大数据相关标准,已形成包括参考模型、关键技术、安全隐私、领域应用等在内的比较全面的标准体系。我国成立了大数据标准工作组与大数据安全标准特别工作组,现已发布35项国家标准,并积极参与国际标准的制定。同时,围绕大数据的开源软硬件社区蓬勃发展,全球最大的代码托管平台GitHub截至2020年采用社交化方式汇聚了全球约1亿代码仓,以及近6 000万名开发者,成为科技创新的强大引擎。开源硬件方面,OpenCores已经成为全世界最大的免费开源硬件IP核线上社区。在大数据开源软件领域,截至2020年木兰社区托管的代码仓库量超过1 500万。在开源硬件方面,2007年美国赛灵思公司的大学计划资助创建了中国首个开源硬件社区OpenHW。中国科学院的科学数据银行、上海交通大学推动白玉兰开放数据集社区做了很好的尝试,已经产生影响力。

4.3 小结

大数据治理技术整体上尚不成熟,很多技术仍然处于探索阶段。在大数据汇聚融合方面,对比美欧日俄,我国的优势在于数据资源丰富、政府推动有力、行业应用广泛,这些为技术的创新发展奠定了良好的基础。然而,我国也显示出单点研究居多、整体性的系统创新不足、研究比较分散、缺乏有影响力的主流系统、缺乏应用示范效应等不足。在大数据质量保障方面,我国处于追赶状态,差距主要体现为基础理论

薄弱、缺乏面向大数据全生命周期的统一的质量治理模型;缺乏可实现数据质量、治理成本、治理时延的统一优化的通用数据质量保障系统;缺乏尚无突破多组织跨辖域环境下的全链路数据质量追踪与治理体系。在大数据开放流通方面,我国的“黑盒”互操作技术和燕云DaaS系统已成为打破数据孤岛的“撒手锏”技术,达到国际领先水平;在此我国已提出了数联网中国云方案,制定了DOIP新版标准,总体与国外处于技术并跑的阶段。一批行业级和区域级数联网基础设施正在开展建设,使得我国在数联网应用方面取得国际领先地位。在大数据安全与隐私保护方面,我国应用较多,基础性、原创性成果不足,还有待大力发展;在标准化与生态系统建设方面,我国紧跟国际发展前沿,大数据标准体系设计已基本形成,并在稳步推进中。我国开源生态建设方面近年来虽然有所进展,但总体处于跟跑阶段,部分我国企业主导的开源项目(如华为CarbonData)崭露头角,但是我国主导的生态系统尚未建立。整体上我国数据治理技术发展较晚,体系远未成形,技术产品生态仍由外国主导,同时在以数联网为代表的数联网开放流通技术方面与国际先进水平相当。

5 大数据技术的未来发展趋势

在大数据应用需求的驱动下,计算技术体系正在重构,从“以计算为中心”向“以数据为中心”转型,在新的计算技术体系下,一些基础理论和核心技术问题亟待破解。本文提出新型大数据系统技术发展的十大趋势。

趋势一:数据与应用进一步分离,实现数据要素化。数据一开始是依附于具体应用的。数据库技术的出现使得数据与应用

实现了第一次分离。数据存储于数据库中，不再依赖具体的应用而存在。数据要素化的需求将推动数据与应用进一步分离，数据不再依赖于具体的业务场景，数据以独立的形态存在于数据库中，并通过数据服务为不同的业务场景提供服务。例如，人口数据库可以为全部的涉及人口信息的业务场景提供服务。

趋势二：数联网成为数字化时代的新型信息基础设施。将形成一套完整的数联网基础软件理论、系统软件架构、关键技术体系，包括：针对数联网软件以数据为中心的特点，需要从复杂网络和复杂系统等复杂性理论出发，研究数联网软件的结构组成、行为模式和外在性质；针对数联网软件的数据传存算一体化需求，需要采用数据互操作技术和软件定义思想，研究数联网软件运行机理、体系结构与关键机制；针对数联网软件跨层级、跨地域、跨系统运行带来的可靠性、可用性、安全性等质量挑战，需要以数据驱动为手段，研究数联网环境下保障服务质量与保护质量的原理、机制与方法。

趋势三：从单域到跨域数据管理，促进数据要素的共享与协同。以数据为中心的计算机的核心目标是数据价值的最大化，关键要打破“数据孤岛”，实现数据要素的高效共享与协同。传统数据管理局限于单一企业、业务、数据中心等内部，未来大数据管理将从传统的单域模式发展到跨域模式，跨越空间域、管辖域和信任域。但跨空间域会造成网络时延较高且不稳定；跨管辖域会造成数据与应用异构，数据管理复杂度大大提升；跨信任域则要求具备容忍各类恶意错误的能力。跨域带来的这些变化将为大数据技术带来新的机遇和挑战。

趋势四：大数据管理与处理系统体系结构异构化日趋明显。体系结构创新进入“黄金十年”，围绕不同数据处理特征

的新型加速器（GPU、TPU、APU等各种xPU）层出不穷，存储器件快速发展，高速固态硬盘（solid state disk, SSD）、新型非易失内存、新型计算网络等成为大数据处理系统的重要硬件配置，计算与存储的融合趋势明显。为了最大限度地发挥数据管理能力，大数据管理系统在存储、网络、计算等硬件上最大化挖掘新型硬件的处理能力。在处理上针对不同数据处理需求，配置不同计算与存储硬件成为大数据处理系统的主流架构。数据驱动的计算架构快速发展，以数据流为中心的系统结构成为重要的系统设计理念。

趋势五：扩展性优先设计到性能优先设计。数据规模急剧增长，大数据处理需求越来越走向深度价值挖掘，数据处理计算愈发密集，数据管理与处理的成本成为大数据管理与处理系统的重要考量因素，传统“扩展性优先”的大数据处理系统设计将会被“以性能优先”的系统设计代替。Spark、Flink等系统在大数据处理生态系统中的占有率明显体现了这一趋势，图计算（图加速器、图计算框架等）、深度学习框架（Tensorflow、PyTorch等）等领域专用大数据处理系统的崛起也是这一系统设计理念在技术生态上的表现。智能化数据管理、近似计算等新兴管理和处理方法成为性能优先设计的重要技术手段。

趋势六：近数处理成为突破大数据处理系统性能瓶颈的重要途径。存算一体类体系结构技术快速发展，新型SSD等新型存储赢家功能愈发丰富，分布式计算系统边缘能力迅速发展。以上3种体系结构技术发展为大数据近数处理提供了良好的发展契机。近数处理体现在“存储上移”（如在GPU、现场可编程门阵列（field-programmable gate array, FPGA）等计算设备上集成HBM）、“算力下沉”（如在DRAM内存或者SSD存储设备上集成处理

能力)、“分布扩展”(如在云、边、端分布式处理数据,降低数据处理中心压力)3个方面。

趋势七:从单域单模态分析到多域多模态融合,实现广谱关联计算。传统大数据分析技术大多仅聚焦于单一来源、单一模态的数据,而实际应用中往往要对来自不同来源、不同模态(如文本、图像、音视频等)的数据进行联合分析,从而实现不同来源与不同模态数据之间的信息互补。此外,诸多领域的大数据具有重要的时空属性,当前研究对这类信息的利用还不够充分。因此,探究能够跨模态关联、跨时空关联的广谱关联技术是大数据分析处理的一个重要趋势。

趋势八:从聚焦关联到探究因果,实现分析结果可解释。如何让大数据分析模型更加稳定且具有可解释性,从而使其分析结果对于用户而言变得更加可信、更加可用最好还能具备一定的可回溯性,是大数据分析面临的巨大挑战。虽然因果推断与可解释性分析技术取得了一定进展,但总体来说尚处于起步阶段,离实际应用还有很长一段距离。因此,从关联到因果也是未来大数据分析技术的重要研究方向。

趋势九:高效能大数据技术是可持续发展的关键。全球大数据量的持续高速增长,以及“碳达峰、碳中和”目标的提出,要求大数据技术栈必须走低碳高效、可持续发展的路线。例如云数据管理系统以资源共享、节能高效为主要特点,将是未来大数据管理的主要基础形态;在云数据管理基础上的全国一体化高效能大数据管理,由于算力和数据要素的大规模调度与流通,可以进一步成为未来大数据管理的主要方向,形成低碳发展新格局。

趋势十:大数据标准规范和以开源社区为核心的软硬件生态系统将成为发展的重点。随着大数据在各个领域应用的迅速

普及,标准化需求将不断增长,与大数据流动融合、质量评估,以及与行业、领域应用密切相关的大数据标准将成为发展重点。开源社区在大数据软硬件生态建设中的地位不断加强,对开源社区的主导权争夺将成为各国技术、产品和市场竞争的重点。

6 结束语

回顾国内外大数据技术在管理、处理、分析与治理4个方面近10年的发展,可以看出:数据规模高速增长,现有处理计算能力已经成为瓶颈;数据成为生产要素,但数据价值释放不充分;从产业生态重点的变迁看,呈现出“应用先于理论技术,市场先于标准法规”的现象。虽然大数据已经在一些应用领域(特别是互联网领域)取得了较好的成效,但是大数据基础理论和应用技术不成熟,大数据治理体系远未建立。总体上,大数据发展仍然处于初级阶段。

我国在大数据发展方面取得了积极进展,但总体上较国际先进水平,仍存在差距。具体地,大数据管理技术大部分领域与国外顶尖水平基本相同或接近,其中高效能一体化大数据管理领域处于国际领先水平;大数据处理技术多数领域与国外顶尖水平尚存在技术差距,在资源和互联网应用领域大数据处理技术应用较好,与国际最高水平基本持平;大数据分析的基础理论与核心技术方面与国际先进水平还存在着一定差距,在商业等领域应用方面已超越国外;大数据治理技术整体上发展较晚,体系远未成形,技术产品生态仍由国外主导,同时在以数联网为代表的数据开放流通技术方面与国际先进水平相当。

面向未来,在大数据应用需求的驱动下,计算技术体系有必要进行重构,以数据

为中心的新型大数据系统技术成为重要方向,信息技术体系将从“以计算为中心”向“以数据为中心”转型,新的基础理论和核心技术问题仍有待探索和破解。以大数据管理、处理、分析和治理为核心的大数据技术在原有通用计算体系上的持续优化仍有发展空间,以数据为中心的新技术体系将成为缩小大数据规模指数级增长与大数据计算需求和能力之间“剪刀差”的突破点。

参考文献:

- [1] 裴威, 李战怀, 潘巍. GPU数据库核心技术综述[J]. 软件学报, 2021, 32(3): 859-885.
PEI W, LI Z H, PAN W. Survey of key technologies in GPU database system[J]. Journal of Software, 2021, 32(3): 859-885.
- [2] SHERKAT R, FLORENDO C, ANDREI M, et al. Native store extension for SAP HANA[J]. Proceedings of the VLDB Endowment, 2019, 12(12): 2047-2058.
- [3] SHEN S J, CHEN R, CHEN H B, et al. Retrofitting High availability mechanism to tame hybrid transaction/analytical processing[C]//Proceedings of 2021 Operating Systems Design and Implementation. [S.l.:s.n.], 2021: 219-238.
- [4] LIU G, CHEN L Y, CHEN S M. Zen: a high-throughput log-free OLTP engine for non-volatile main memory[J]. Proceedings of the VLDB Endowment, 2021, 14(5): 835-848.
- [5] KRASKA T, BEUTEL A, CHIEH, et al. The case for learned index structures[C]//Proceedings of 2018 International Conference on Management of Data. New York: ACM Press, 2018: 489-504.
- [6] CHATTERJEE S, JAGADEESAN M, QIN W, et al. Cosine[J]. Proceedings of the VLDB Endowment, 2021, 15(1): 112-126.
- [7] DAS S, GRBIC M, ILIC I, et al. Automatically indexing millions of databases in microsoft azure SQL database[C]//Proceedings of 2019 International Conference on Management of Data. New York: ACM Press, 2019: 666-679.
- [8] AHMED R, BELLO R, WITKOWSKI A, et al. Automated generation of materialized views in Oracle[J]. Proceedings of the VLDB Endowment, 2020, 13(12): 3046-3058.
- [9] LIU X Z, YIN Z, ZHAO C, et al. PinSQL: pinpoint root cause SQLs to resolve performance issues in cloud databases[C]//Proceedings of 2022 IEEE 38th International Conference on Data Engineering. Piscataway: IEEE Press, 2022: 2549-2561.
- [10] LI G L, ZHOU X H, SUN J, et al. OpenGauss: an autonomous database system[C]//Proceedings of the International Conference on Very Large Databases. [S.l.:s.n.], 2021, 14(12): 3028-3041.
- [11] ZHOU X H, LI G L, CHAI C L, et al. A learned query rewrite system using Monte Carlo tree search[J]. Proceedings of the VLDB Endowment, 2021, 15(1): 46-58.
- [12] WANG J Y, CHAI C L, LIU J B, et al. FACE: a normalizing flowbased cardinality estimator[C]//Proceedings of the International Conference on Very Large Databases. [S.l.:s.n.], 2022, 15(1): 72-84.
- [13] DEPOUTOVITCH A, CHEN C, CHEN J, et al. Taurus database: how to be fast, available, and frugal in the cloud[C]//Proceedings of 2020 ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2020: 1463-1478.
- [14] CAO W, LIU Z J, WANG P, et al. PolarFS: an ultra-low latency and failure resilient distributed file system for shared storage cloud database[J]. Proceedings of the VLDB Endowment, 2018, 11(12): 1849-1862.
- [15] TAFT R, SHARIF I, MATEI A, et al.

- CockroachDB: the resilient geo-distributed SQL database[C]//Proceedings of 2020 ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2020: 1493-1509.
- [16] CAO W, LIU Z J, WANG P, et al. PolarFS: an ultra-low latency and failure resilient distributed file system for shared storage cloud database[J]. Proceedings of the VLDB Endowment, 2018, 11(12): 1849-1862.
- [17] WANG Y Y, WANG Z K, CHAI Y P, et al. Rethink the linearizability constraints of raft for distributed key-value stores[C]//Proceedings of 2021 IEEE 37th International Conference on Data Engineering. Piscataway: IEEE Press, 2021: 1877-1882.
- [18] HUANG D X, LIU Q, CUI Q, et al. TiDB[J]. Proceedings of the VLDB Endowment, 2020, 13(12): 3072-3084.
- [19] WANG H X, XU C, ZHANG C, et al. A blockchain system ensuring query integrity[C]//Proceedings of the ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2020: 2693-2696.
- [20] DANG H, DINH T T A, LOGHIN D, et al. Towards scaling blockchain systems via sharding[C]//Proceedings of 2019 International Conference on Management of Data. New York: ACM Press, 2019: 123-140.
- [21] ALAKUIJALA J, FARRUGGIA A, FERRAGINA P, et al. Brotli: a general-purpose data compressor[J]. ACM Transactions on Information Systems, 2019, 37(1): 1-30.
- [22] CAO W, ZHANG Y Q, YANG X J, et al. PolarDB serverless: a cloud native database for disaggregated data centers[C]//Proceedings of 2021 International Conference on Management of Data. New York: ACM Press, 2021: 2477-2489.
- [23] ZHANG F, WAN W T, ZHANG C Y, et al. CompressDB: enabling efficient compressed data direct processing for various databases[C]//Proceedings of 2022 International Conference on Management of Data. [S.l.:s.n.], 2022: 1655-1669.
- [24] WOJTOWICZ D T, YIN S Y, MORVAN F, et al. Cost-effective dynamic optimisation for multi-cloud queries[C]//Proceedings of 2021 IEEE 14th International Conference on Cloud Computing. Piscataway: IEEE Press, 2021: 387-397.
- [25] 王建冬, 于施洋, 窦悦. 东数西算: 我国数据跨境流通的总体框架和实施路径研究[J]. 电子政务, 2020(3): 13-21.
- WANG J D, YU S Y, DOU Y. East-west computing transfer: research on the overall framework and implementation path of cross-domain data circulation in China[J]. E-Government, 2020(3): 13-21.
- [26] DEAN J, GHEMAWAT S. MapReduce: simplified data processing on large clusters[J]. Communications of the ACM, 2008, 51(1): 137-150.
- [27] FEY M, LENSSEN J E. Fast graph representation learning with PyTorch geometric[J]. arXiv preprint, 2019, arXiv:1903.02428v2.
- [28] RASCHKA S, PATTERSON J, NOLET C. Machine learning in python: main developments and technology trends in data science, machine learning, and artificial intelligence[J]. Information, 2020, 11(4): 193.
- [29] AHN J, YOO S, MUTLU O, et al. PIM-enabled instructions: a low-overhead, locality-aware processing-in-memory architecture[J]. Computer Architecture News, 2015, 43(3): 336-348.
- [30] WU M Y, ZHAO Z M, LI H Y, et al. Espresso: brewing Java for more non-volatility with non-volatile memory[C]//Proceedings of the 23rd International Conference on Architectural Support for Programming Languages and Operating Systems. New York: ACM Press, 2018: 70-83.
- [31] SHI X H, KE Z X, ZHOU Y L, et al.

- Deca: a garbage collection optimizer for in-memory data processing[J]. *ACM Transactions on Computer Systems*, 2018, 36(1): 1–47.
- [32] ZEUCH S, MONTE B D, KARIMOV J, et al. Analyzing efficient stream processing on modern hardware[J]. *Proceedings of the VLDB Endowment*, 2019, 12(5): 516–530.
- [33] TOSHNIWAL A, TANEJA S, SHUKLA A, et al. Storm@twitter[C]//*Proceedings of 2014 ACM SIGMOD International Conference on Management of Data*. New York: ACM Press, 2014.
- [34] ZAHARIA M, DAS T, LI H Y, et al. Discretized streams: fault-tolerant streaming computation at scale[C]//*Proceedings of the 24th ACM Symposium on Operating Systems Principles*. New York: ACM Press, 2013.
- [35] NASIR M A U, MORALES G D F, GARCÍA-SORIANO D, et al. The power of both choices: practical load balancing for distributed stream processing engines[C]//*Proceedings of 2015 IEEE 31st International Conference on Data Engineering*. Piscataway: IEEE Press, 2015: 137–148.
- [36] NASIR M A U, MORALES G D F, KOURTELLIS N, et al. When two choices are not enough: balancing at scale in distributed stream processing[C]//*Proceedings of 2016 IEEE 32nd International Conference on Data Engineering*. Piscataway: IEEE Press, 2016: 589–600.
- [37] ABDELHAMID A S, MAHMOOD A R, DAGHISTANI A, et al. Prompt: dynamic data-partitioning for distributed micro-batch stream processing systems[C]//*Proceedings of 2020 ACM SIGMOD International Conference on Management of Data*. New York: ACM Press, 2020: 2455–2469.
- [38] CHEN H H, ZHANG F, JIN H. PStream: a popularity-aware differentiated distributed stream processing system[J]. *IEEE Transactions on Computers*, 2021, 70(10): 1582–1597.
- [39] MALEWICZ G, AUSTERN M H, BIK A J C, et al. Pregel: a system for large-scale graph processing[C]//*Proceedings of 2010 ACM SIGMOD International Conference on Management of Data*. New York: ACM Press, 2010: 135–146.
- [40] WANG Y, DAVIDSON A, PAN Y C, et al. Gunrock: a high-performance graph processing library on the GPU[C]//*Proceedings of the 21st ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. New York: ACM Press, 2016.
- [41] ZHOU S J, KANNAN R, PRASANNA V K, et al. HitGraph: high-throughput graph processing framework on FPGA[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2019, 30(10): 2249–2264.
- [42] RAHMAN S, ABU-GHAZALEH N, GUPTA R. GraphPulse: an event-driven hardware accelerator for asynchronous graph processing[C]//*Proceedings of 2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture*. Piscataway: IEEE Press, 2020: 908–921.
- [43] SHUN J L, BLELLOCH G E. Ligma: a lightweight graph processing framework for shared memory[C]//*Proceedings of the 18th ACM SIGPLAN symposium on Principles and practice of parallel programming*. New York: ACM Press, 2013: 135–146.
- [44] GONZALEZ J E, LOW Y C, GU H J, et al. PowerGraph: distributed graph-parallel computation on natural graphs[C]//*Proceedings of the 10th USENIX Conference on Operating Systems Design and Implementation*. New York: ACM Press, 2012:17–30.
- [45] KYROLA A, BLELLOCH G, GUESTRIN C.

- GraphChi: large-scale graph computation on just a PC[C]//Proceedings of the 10th USENIX Symposium on Operating Systems Design and Implementation. New York: ACM Press, 2012: 31-46.
- [46] HAM T J, WU L S, SUNDARAM N, et al. Graphicionado: a high-performance and energy-efficient accelerator for graph analytics[C]//Proceedings of 2016 49th Annual IEEE/ACM International Symposium on Microarchitecture. Piscataway: IEEE Press, 2016: 1-13.
- [47] ZHANG Y, LIAO X F, JIN H, et al. HotGraph: efficient asynchronous processing for real-world graphs[J]. IEEE Transactions on Computers, 2017, 66(5): 799-809.
- [48] ZHANG K Y, CHEN R, CHEN H B. NUMA-aware graph-structured analytics[C]//Proceedings of the 20th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming. New York: ACM Press, 2015: 183-193.
- [49] ZHU X W, CHEN W G, ZHENG W M, et al. Gemini: a computation-centric distributed graph processing system[C]//Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation. New York: ACM Press, 2016: 301-316.
- [50] ZHANG Y, LIAO X F, GU L, et al. Asyngraph: maximizing data parallelism for efficient iterative graph processing on gpus[J]. ACM Transactions on Architecture and Code Optimization, 2020, 17(4): 1-21.
- [51] DAI G H, HUANG T H, CHI Y Z, et al. ForeGraph: exploring large-scale graph processing on multi-FPGA architecture[C]//Proceedings of 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays. New York: ACM Press, 2017: 217-226.
- [52] ZHAO J, YANG Y, ZHANG Y, et al. TDGraph: a topology-driven accelerator for high-performance streaming graph processing[C]//Proceedings of the 49th Annual International Symposium on Computer Architecture. New York: ACM Press, 2022: 116-129.
- [53] LIN H, ZHU X W, YU B W, et al. ShenTu: processing multi-trillion edge graphs on millions of cores in seconds[C]//Proceedings of International Conference for High Performance Computing, Networking, Storage and Analysis. Piscataway: IEEE Press, 2018.
- [54] ZHANG Y, LIAO X F, JIN H, et al. DiGraph: an efficient path-based iterative directed graph processing system on multiple GPUs[C]//Proceedings of the 24th International Conference on Architectural Support for Programming Languages and Operating Systems. New York: ACM Press, 2019: 601-614.
- [55] PHAM H, LIANG P P, MANZINI T, et al. Found in translation: learning robust joint representations by cyclic translations between modalities[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 6892-6899.
- [56] WANG W H, BAO H B, DONG L, et al. Image as a foreign language: BEIT pretraining for all vision and vision-language tasks[J]. arXiv preprint, 2022, arXiv:2208.10442.
- [57] CHEN X, WANG X, CHANGPINYO S, et al. Pali: a jointly-scaled multilingual language-image model[J]. arXiv preprint, 2022, arXiv:2209.06794.
- [58] LIU J, ZHU X X, LIU F, et al. OPT: omni-perception pre-trainer for cross-modal understanding and generation[J]. arXiv preprint, 2021, arXiv:2107.00249.
- [59] MAMMEN P M. Federated learning: opportunities and challenges[J]. arXiv preprint, 2021, arXiv:2101.05428.
- [60] ZILLER A, TRASK A, LOPARDO A, et al. PySyft: a library for easy federated

- learning[M]//Federated learning systems. Cham: Springer, 2021: 111–139.
- [61] WELTEN S, MOU Y L, NEUMANN L, et al. A privacy-preserving distributed analytics platform for health care data[J]. *Methods of Information in Medicine*, 2022, 61(S 01): e1–e11.
- [62] LI Q B, WEN Z Y, WU Z M, et al. A survey on federated learning systems: vision, hype and reality for data privacy and protection[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021: 10.1109/TKDE.2021.3124599.
- [63] RUBIN D B. Estimating causal effects of treatments in randomized and nonrandomized studies[J]. *Journal of Educational Psychology*, 1974, 66(5): 688–701.
- [64] PEARL J. *Causality: models, reasoning and inference*[M]. Cambridge: Cambridge University Press, 2009.
- [65] PEARL J, MACKENZIE D. The book of why: the new science of cause and effect[J]. *Journal of MultiDisciplinary Evaluation*, 2018, 14(31): 47–54.
- [66] SUN X W, WU B T, ZHENG X Y, et al. Recovering latent causal factor for generalization to distributional shifts[C]// *Advances in Neural Information Processing Systems*. [S.l.:s.n.], 2021: 16846–16859.
- [67] CUI P, ATHEY S. Stable learning establishes some common ground between causal inference and machine learning[J]. *Nature Machine Intelligence*, 2022, 4(2): 110–115.
- [68] ZHANG Y, FENG F L, HE X N, et al. Causal intervention for leveraging popularity bias in recommendation[J]. *arXiv preprint*, 2021, arXiv:2105.06067.
- [69] ZHU Z M, CHEN X H, TIAN H L, et al. Offline reinforcement learning with causal structured world models[J]. *arXiv preprint*, 2022, arXiv:2206.01474.
- [70] STONEBRAKER M. The solution: data curation at scale[M]//Getting data right. [S.l.]: O'Reilly, 2016.
- [71] 华为公司数据管理部. 华为数据之道[M]. 北京: 机械工业出版社, 2020.
- Data Management Department of Huawei. *Enterprise data at Huawei*[M]. Beijing: China Machine Press, 2020.
- [72] REKATSINAS T, CHU X, ILYAS I F, et al. HoloClean: holistic data repairs with probabilistic inference[J]. *arXiv preprint*, 2017, arXiv:1702.00820.
- [73] DONG X, GABRILOVICH E, HEITZ G, et al. Knowledge vault: a web-scale approach to probabilistic knowledge fusion[J]. *SIGKDD Explorations*, 2014(CD/ROM): 597–606.
- [74] 郝爽, 李国良, 冯建华, 等. 结构化数据清洗技术综述[J]. *清华大学学报(自然科学版)*, 2018, 58(12): 1037–1050.
- HAO S, LI G L, FENG J H, et al. Survey of structured data cleaning methods[J]. *Journal of Tsinghua University (Science and Technology)*, 2018, 58(12): 1037–1050.
- [75] 丁小欧, 王宏志, 于晟健. 工业时序大数据质量管理[J]. *大数据*, 2019, 5(6): 1–11.
- DING X O, WANG H Z, YU S J. Data quality management of industrial temporal big data[J]. *Big Data Research*, 2019, 5(6): 1–11.
- [76] KAHN R, WILENSKY R. A framework for distributed digital object services[J]. *International Journal on Digital Libraries*, 2006, 6(2): 115–123.
- [77] 梅宏, 黄罡, 刘譞哲, 等. 网构软件研究: 回顾与展望[J]. *科学通报*, 2022, 67(32): 3782–3792.
- MEI H, HUANG G, LIU X Z, et al. Research on internetware: review and prospect[J]. *Chinese Science Bulletin*, 2022, 67(32): 3782–3792.
- [78] 黄罡. 数联网: 数字空间基础设施[J]. *中国计算机学会通讯*, 2021, 17(12): 58–60.
- HUANG G. Internet of Data: infrastructure of digital space[J]. *Communications of the CCF*, 2021, 17(12): 58–60.

作者简介



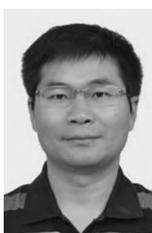
梅宏(1963-),男,博士,北京大学教授、高可信软件技术教育部重点实验室(北京大学)主任,中国科学院院士,发展中国家科学院院士,欧洲科学院外籍院士,中国计算机学会理事长。主要研究方向为软件工程与系统软件。



杜小勇(1963-),男,博士,中国人民大学教授、校长助理,中国计算机学会大数据专家委员会主任,主要研究方向为数据库与大数据。



金海(1966-),男,博士,华中科技大学计算机科学与技术学院教授,中国计算机学会副理事长,主要研究方向为计算机系统结构、并行与分布式计算。



程学旗(1971-),男,博士,中国科学院计算技术研究所研究员、副所长,主要研究方向为大数据分析系统、Web信息检索与数据挖掘等。



柴云鹏(1983-),男,博士,中国人民大学信息学院教授、计算机系主任,主要研究方向为数据库系统、云计算、存储系统等。



石宣化(1978-),男,博士,华中科技大学计算机科学与技术学院教授,主要研究方向为并行与分布式计算、异构计算。



靳小龙 (1976-), 男, 博士, 中国科学院计算技术研究所研究员, 主要研究方向为知识图谱、知识工程、社会计算、社交网络等。



王亚沙 (1975-), 男, 博士, 高可信软件技术教育部重点实验室(北京大学)教授, 主要研究方向为大数据分析、普适计算、城市计算。



刘驰 (1984-), 男, 博士, 北京理工大学计算机学院教授、副院长, 主要研究方向为大数据分析、智能物联网。

收稿日期: 2022-12-22