

# 数据密集型科学与工程:需求和挑战

宫学庆<sup>1)</sup> 金澈清<sup>1)</sup> 王晓玲<sup>1),2)</sup> 张 蓉<sup>1)</sup> 周傲英<sup>1),2)</sup>

<sup>1)</sup>(华东师范大学软件学院 上海市高可信计算重点实验室 上海 200062)

<sup>2)</sup>(复旦大学 上海市智能信息处理实验室 上海 200433)

**摘 要** 科学研究在经历了实验科学、理论科学、计算科学阶段后,进入了数据密集型科学阶段,与之相伴的是大数据时代的到来.大数据泛指规模达到几百TB,甚至PB级的数据<sup>①</sup>,其典型的特征是分布、异构、低质量等.尽管传统数据库管理技术(特别是商业关系型数据库)在过去40年间取得了巨大成功,但是这些技术和系统无法有效管理支持数据密集型科学与工程(Data-Intensive Science and Engineering, DISE)的大数据.文中探讨数据密集型科学与工程的具体需求和现实挑战.它涵盖的内容表现在4个层面,包括数据存储与组织、计算方法、数据分析以及用户接口技术等.同时,数据质量、数据安全、数据监护等内容也需要在各层面得到重视.文中尝试梳理了数据密集型科学与工程的整体架构,回顾了相关领域的新近发展,分析了面临的挑战,探讨了未来的研究方向.

**关键词** 大数据;数据密集型科学与工程;需求;挑战

中图法分类号 TP311 DOI号: 10.3724/SP.J.1016.2012.01563

## Data-Intensive Science and Engineering: Requirements and Challenges

GONG Xue-Qing<sup>1)</sup> JIN Che-Qing<sup>1)</sup> WANG Xiao-Ling<sup>1),2)</sup> ZHANG Rong<sup>1)</sup> ZHOU Ao-Ying<sup>1),2)</sup>

<sup>1)</sup>(Shanghai Key Laboratory of Trustworthy Computing, Software Engineering Institute, East China Normal University, Shanghai 200062)

<sup>2)</sup>(Shanghai Key Laboratory of Intelligent Information Processing, Fudan University, Shanghai 200433)

**Abstract** Scientific exploration after experimental science, theoretical science and computational science phases, into data-intensive science phase, are accompanied by the arrival of the big data era. Generally, big data refers to a data set with a size of hundreds of TB, or several PB or even above, and it is often distributed, heterogeneous and in low-quality. It is critical to devise novel methods to manage big data since traditional database management techniques are unfeasible to manage big data efficiently and effectively, though such techniques, especially the commercial relational DBMSs, have achieved great success in the past decades. This paper discusses concrete requirements and realistic challenges of Data-Intensive Science and Engineering (DISE), ranging from data storage and organization, computational method, data analysis, to user interfaces. Meanwhile, data quality, data security and data curation should be paid more attentions. In this paper, we attempt to describe the architecture of DISE, review the recent progress, and discuss the challenges and future work briefly.

**Keywords** big data; Data Intensive Science and Engineering (DISE); requirement; challenge

收稿日期:2011-06-13;最终修改稿收到日期:2012-07-06. 本课题得到国家核高基项目(2010ZX01042-002-003-004)、国家自然科学基金项目(60925008,61070052)资助. 宫学庆,男,1974年生,副教授,主要研究方向为数据库技术、Web数据管理等.金澈清,男,1977年生,博士,教授,主要研究领域为数据流管理、不确定性数据管理技术、基于位置的服务等.王晓玲,女,1975年生,教授,博士生导师,主要研究兴趣为XML数据管理、Web服务等.张蓉,女,1978年生,博士,副教授,主要研究兴趣为Web服务、云计算等.周傲英(通信作者),男,1965年生,教授,博士生导师,主要研究兴趣为数据管理与信息系统,包括Web数据管理、中文Web基础设施、Web搜索与挖掘、数据流与数据挖掘、复杂事件处理与实时商务智能、不确定数据管理及其应用、数据密集的计算、分布存储与计算、对等计算及其数据管理、Web服务计算等. E-mail: ayzhou@sei.ecnu.edu.cn.

① 1TB相当于 $10^{12}$ 字节,1PB相当于 $10^{15}$ 字节.

## 1 引 言

大数据管理技术业已成为当前和未来一段时间内计算机领域的重要研究课题之一. 美国的总统科学顾问委员会于 2010 年底提出一份报告, 认为“数据规模呈指数增长, ……各联邦机构都需要制定应对‘大数据’的策略”<sup>①</sup>. 所谓大数据, 泛指规模达到几百 TB, 甚至 PB 级的数据, 广泛出现在科学研究、Web 集成、多媒体等诸多领域中. Jim Gray(图灵奖获得者)认为, 当今时代已是数据密集型科学研究时代(即 eScience 时代), 科研人员的工作重心多已转向如何分析海量数据, 并获取知识<sup>[1]</sup>. 例如, 天文望远镜是重要的天文观测工具, 正在建设中的 LSST 天文望远镜(Large Synoptic Survey Telescope)在服役后每天将会产生 30TB 数据<sup>②</sup>. Web 数据的规模增长也非常迅猛, 根据中国互联网络信息中心(CNNIC)的调查报告, 截至 2010 年底, 我国网民规模为 4.57 亿人, 网页数目约 60G, 总内容超过 1.8PB<sup>③</sup>, 而国内网页的内容量仅占整个互联网的一小部分. 多媒体数据(包括视频、音频、图片等)也非常庞大. 据 2008 年数据统计, Facebook 公司存储着 10 亿张照片, 约 1PB 规模<sup>④</sup>. 其它领域例如医疗、金融、电子商务、物流等也产生并使用大数据.

大数据除了规模庞大之外, 还兼具分布式、异构、不准确、不一致等特性, 因而管理难度加大. 以海洋观测科学研究为例, 为了监测较大海域的洋流变化情况, 需在海水中、海平面、海床分别部署大量传感设备和工业机器人, 这些设备(或者机器人)通过无线网络将数据传回服务器, 分析人员结合现有资料(例如文档、图片、视频等)分析数据. 在此过程中, 数据由分布在不同地理位置上的多个节点产生; 受传感设备精度、海洋环境、人工操作等因素的影响, 所采集数据可能并不准确; 数据来源多样, 因而可能存在不一致性, 需要采用多种手段提升质量.

大数据管理已经得到了国内外学者的极大关注. 《自然》杂志的 2008 年 9 月专刊提出的观点是, 人类已经进入拍字节时代(petabyte era), 大数据在各领域广泛出现. 该杂志还预测下一个 IT 巨头的主营业务将会是大数据管理<sup>[2]</sup>. 微软研究院出版的《The fourth paradigm》一书以科学研究为切入点, 阐述了如何在 eScience 时代做数据密集型的科学研究<sup>[1]</sup>. 《科学》杂志的 2011 年 2 月份专刊讨论如何管理大数据<sup>[3]</sup>. 我国学者也密切关注大数据领域.

《中国计算机学会通讯》2011 年 7 月刊的专题《数据密集型计算》从不同角度阐述了数据密集型计算(Data-Intensive Computing)的相关技术, 也介绍了几位专家正在开展的工作<sup>[4]</sup>. 以上论著侧重于从一些具体方面介绍、分析大数据对传统数据管理技术的强烈冲击. 事实上, 大数据的出现已经在数据管理的各个层面产生了深刻的影响, 数据密集型科学与工程(Data-Intensive Science and Enigneering, DISE)已经成为目前国内外的一个研究热点. 本文尝试梳理了 DISE 的整体架构, 回顾了相关领域的新近发展, 探讨了 DISE 的研究方向和所面临的挑战. DISE 的基本框架如图 1 所示, 包括数据存储和组织、计算方法、数据分析、用户接口等 4 个方面, 以及在整个数据管理过程中需重视的数据质量、数据安全、数据监护等问题.

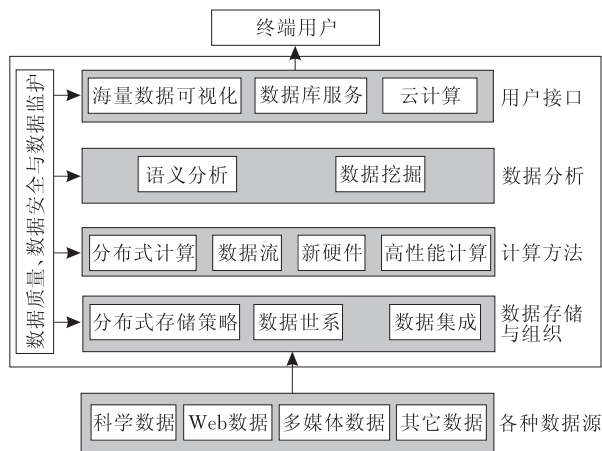


图 1 DISE 的基本框架

首先, 需要合理地存储与组织各种大数据, 如科学数据、Web 数据、多媒体数据等. 鉴于集中式存储策略无法提供良好的 I/O 访问效率, 需要设计分布式存储策略. 在数据存储和组织过程中还需要考虑数据世系(data provenance)和数据集成(data integration)因素. 数据世系是指数据产生及演化的整个过程, 在数据质量评价、数据核查、数据恢复和数据引用等方面都具有重要意义<sup>[5]</sup>. 数据集成将不同来源的数据有效地集成在一起, 提供统一的数据访问模式.

其次, 需要设计高效的计算方法来管理海量数

- ① President's Council of Advisors on Science and Technology. Designing a digital future: federally funded research and development in networking and information technology. <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrtd-report-2010.pdf>
- ② <http://www.lsst.org/lsst/faq#q1>
- ③ <http://research.cnnic.cn/html/1295338825d2556.html>
- ④ <http://mashable.com/2008/10/15/facebook-10-billion-photos/>

据,包括分布式计算、数据流技术、新硬件技术、高性能计算等。分布式计算方法将整个计算任务分摊到不同计算机上执行,大大降低了查询与分析的时间,其挑战性在于如何执行任务的分解和归约操作。目前用的比较多的任务分解方式是数据分解。与分布式计算方法不同,数据流管理技术强调如何提升单台计算机的计算效率,它一般仅仅使用少量内存空间,以单遍扫描方式访问数据集,并获得较高质量的近似结果;若给予更多的内存空间,查询结果的质量可以进一步提升。另外,开发新型计算不可忽视硬件技术的革新。近几年来,嵌入了多个“核”的多核处理器得到长足发展,每个“核”都具有独立的计算能力,因而其计算能力大大超过单核处理器。但是传统的串行算法无法发挥多核处理器的计算能力,需要重新开发多线程并行算法。图形处理单元(GPU)和现场可编程逻辑器件(FPGA)等硬件设备也得到了越来越多的关注,这些器件都具备并行处理能力,适合解决特定问题。还有,超级计算机具有不可比拟的计算能力,因而高性能计算也是大数据管理的重要内容。

再者,数据分析层需要挖掘隐藏在海量数据背后的知识和规律。语义分析和数据挖掘是两种较为常用的技术手段。语义分析技术从海量数据中找出对象的具体特征,并以这些特征来描述对象。数据挖掘技术也是典型的知识抽取技术。用户接口层是用户与计算机系统之间的接口,主要包括海量数据可视化、数据库服务(DaaS)和云计算(Cloud)等。海量数据可视化以图形化界面形式呈现数据,使用户更深刻地理解数据本质,该技术可视为是计算机技术、设计美学、领域知识的完美结合。DaaS是指企业(数据所有者或数据库所有者)将自身的数据库创建、访问、维护、升级、管理等任务委托给专门的第三方(服务提供者或数据库服务提供者)进行管理和维护。云计算是近期兴起的一项新技术,以云的方式向用户提供服务。大数据的内容丰富,包含高价值的信息和知识,经常是在云计算平台上通过服务的方式提供给用户。

最后,数据质量、数据安全和数据监护贯穿 DISE 的各个层面。数据质量高低会严重影响数据管理的成败。产生低质量数据的原因很多,包括设备精度、外部环境、人工操作等;低质量数据的表征也有多种形式,包括冗余、不一致、不准确、缺失等。在大数据应用中,数据来源更广,因而数据质量问题也更为严重。数据清洗技术可以提升数据质量,但无法完全解决数据质量问题,因此仍需在数据质量较低的

情况下进行查询分析,并获得查询分析结果。由于大数据应用往往涉及诸多用户和机构,数据安全的重要性不言而喻,典型应用包括军事、金融、电子商务、保险等。在数据存储与组织层,需要考虑以密文方式存储;在计算方法层和数据分析层,需要考虑设计数据存取策略,分权限访问数据;在用户接口层,需要制定服务执行的访问策略。数据监护是指对数据资产的选择、维持、维护、收集和打包的整个过程。

本文第 2~6 节分别介绍 DISE 的 4 个层面的基本内容(即:数据存储与组织、计算方法、数据分析、用户接口)以及贯穿各个层面的数据质量、数据安全和数据监护相关技术,并分析所面临的挑战;第 7 节列举 2 个典型的大数据应用,即社会计算和信息-物理融合网络;最后,第 8 节总结全文。

## 2 数据存储与组织

### 2.1 分布式数据存储策略

为了提高数据吞吐效率、降低故障率,一般采用分布式策略存储海量数据。基本策略如下:所有数据被放置在多个存储设备之中;存储设备通过高速网络相互连接;为各数据准备多个副本,分别放置在不同存储设备之中;仅当数据块的所有副本均无法被访问时,该数据块才无法被读取;建立分布式索引,当系统接收到数据访问请求时,能够快速决定从哪些存储设备中读取数据。

谷歌公司的 GFS(也被称为 GoogleFS)<sup>[6]</sup>和 Hadoop 项目的 HDFS<sup>[7]</sup>是两个最知名的分布式文件系统。GFS 将集群内的节点划分成两类,包括一个主服务器(master)和多个块服务器(chunk server)。主服务器维护文件系统的元数据信息,但不存储文件内容;块服务器存储文件内容。块服务器的最小读写单位是块(chunk),缺省为 64 MB。每个文件首先被划分成若干个块,分别放置在不同的块服务器中。主服务器记录各块的存储细节。当用户试图获取某个文件时,需向主服务器发送请求;主服务器搜寻该文件的元数据信息,并反馈给用户;用户再与相应块服务器通讯以获取数据。HDFS 也采用类似方式存储数据:namenode 节点(和辅助 namenode 节点)存放文件的元数据信息,datanode 节点放置文件信息<sup>[7]</sup>。GFS 和 HDFS 系统能有效管理无结构数据,也能管理简单结构的数据,如码-值对(key-value pairs)结构的数据。此后,为了管理复杂结构数据,科研人员又纷纷基于这两种文件系统推出新的数据

存储策略. 谷歌公司提出了基于 GFS 的 BigTable 技术<sup>[6]</sup>, 它维护多张表 (table), 每个表是一个多维的稀疏图; 表由行和列组成; 各行和列指定的存储单元均包含值和时间戳信息. 当表规模太大时, 可将表根据行分割成若干个子表, 即 Tablet. 基于 HDFS 文件系统的有结构存储方法很多, 包括 Hive、HBase、HadoopDB 等<sup>[7]</sup>. 这些系统增加了数据格式, 可提供类似 SQL 语句的查询表达能力.

行存储 (row-store) 和列存储 (column-store) 是两种典型的数据库物理存储策略. 行存储方式较为传统, 它在磁盘中依次保存每条记录, 比较适合事务操作; 列存储方式垂直划分关系表, 以列为单位存储数据, 列存储还具有数据压缩 (compression)、延期物化 (late materialization)、块循环 (block iteration) 等特性<sup>[9]</sup>. 由于数据分析任务往往仅使用较少字段, 因此列存储方式的效率更高. 数据分析任务在大数据应用中更为常见, 因此许多系统尽管无法完全实现列存储的所有特性, 但也或多或少地借鉴了相关概念, 包括 BigTable、HBase 等. BigTable 在逻辑上提供了列族 (column family) 的概念, 在物理存储上仍旧是按照行方式存储<sup>[8]</sup>; HBase 将每一个列存储在独立的文件之中<sup>①</sup>.

分布式数据存储策略对于大数据应用来说非常关键. 但是现有研究成果仍然具有一些局限性: (1) 未能充分满足应用需求, 特别是针对复杂数据结构的要求. 倪明选等人<sup>[10]</sup>认为, 现有工作并未延伸到时空数据, 特别是移动对象数据领域, 没有考虑到时空数据的复杂结构和查询处理的复杂性; (2) 正如廖小飞等人<sup>[11]</sup>所认为, 应该重视数据的放置和调度, 不合理的数据放置策略会增加网络和存储 I/O 开销, 降低系统性能; (3) 数据的存储与组织还应该关注列存储技术, 设计优化措施. 正如文献<sup>[9]</sup>提到的那样, 仅仅将数据按列存储并不充分, 还需要引入其它一些优化措施.

## 2.2 数据世系

数据世系一般出现在包含多数据集的应用中, 用于描述数据的产生并随着时间推移而演化的整个过程<sup>[5]</sup>, 其价值在于可评价数据质量并在发生故障时恢复数据. 若缺乏外部辅助信息, 评估数据集的质量比较困难, 但是当充分了解数据进化过程时, 就较易做出判断. 以食品安全监督为例, 菜品从田间到超市经历多道环节, 仅凭人工观察无法正确判断菜品的质量, 而当菜品从生产、物流、批发、零售等各环节均可溯源时, 则可以更全面地反映出菜品的安全

状况. 数据世系也可用于数据恢复, 当某一个中间环节被检验出错误时, 可以利用前序过程恢复数据.

数据世系一般有两类基本方法: 非注解的方法 (non annotation-based approach) 和基于注解的方法 (annotation-based approach). 前者采用模式映射方式, 使用数据处理函数和其相对应的反向函数. 例如, 令  $R$  表示关系  $S_1$  和  $S_2$  的连接结果, 则  $R$  中的每条记录均可在  $S_1$  和  $S_2$  中找到对应记录. 在更复杂的例子中, 可能并不存在集合之间的可逆函数, 必须使用注解描述世系. 事实上, 基于注解的方法的应用范围要远远高于非注解的方法. 数据世系可针对多种数据类型, 包括关系型数据、XML 数据、不确定数据等. 加州大学圣克鲁兹分校的 DBNotes 项目主要研究关系型数据库的世系, 采用的主要手段是增加注释信息<sup>②</sup>; Buneman 则尝试研究针对 XML 数据的世系工作<sup>[12]</sup>. 斯坦福大学的 Widom 将不确定性和世系结合起来, 研究针对不确定数据的世系管理<sup>③</sup>.

在大数据时代, 数据世系的研究工作需关注以下几个方面: (1) 如何有效整合工作流世系 (workflow provenance)<sup>[13]</sup> 和数据世系. 工作流世系和数据世系的粒度、侧重点均不相同. 工作流世系强调数据集的演化过程, 数据世系则记录各个元组的演化过程. 在大数据时代, 仅强调数据世系意味着产生海量注解信息; 仅强调工作流世系则无法得到数据的具体演化信息, 因而有必要探寻二者的结合点; (2) 如何解决异构世系标准的融合问题. 大数据应用将涵盖更多的原本可能相互隔离的数据集合. 例如, 大型医疗保健系统需要结合医疗机构、卫生部门、民政部门的数据, 如何将适用不同标准的数据世系信息整合在一起就很关键; (3) 还需要关注数据世系的可视化问题. 图形化界面是最好的描述数据演化过程的方式, 探讨如何将数据世系以友好的方式呈现给用户也是亟待解决的问题.

## 2.3 数据集成

数据集成的目标是把不同数据源整合在一起. 大数据应用往往涉及诸多数据源, 并且它们的数据模式也可能不一致. 造成这种状况的原因是多方面的. 首先, 这些数据源在设计或收集阶段相对独立, 并未预计到在将来可能会整合在一起; 其次, 各数据源的应用目标不尽相同, 客观上无法采用相同的数据模式; 最后, 各数据源的相关说明文档可能并不全

① <http://hbase.apache.org/book.html>

② <http://users.soe.ucsc.edu/~wctan/Projects/dbnotes/>

③ <http://infolab.stanford.edu/trio/>

面或者已经丢失. 图 2 描述了数据集成的过程, 包含 3 个阶段. 在第 1 阶段, 将不同数据源的数据模式映射为一个中间模式, 以该中间模式描述所有数据源; 在第 2 阶段, 通过重复数据检测技术检测出位于不同数据源但描述真实世界的同一对象的数据(这些数据可能并不一致); 最后, 数据融合(data fusion)技术综合各数据源的数据, 找到恰当的形式来描述真实世界的各个对象<sup>[14]</sup>.

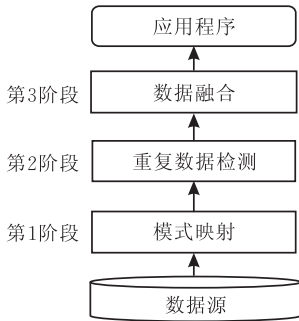


图 2 数据集成的过程<sup>[14]</sup>

Deep Web 数据集成就是一个典型案例. 所有 Web 数据按照所包含的信息的深度大致划分为两大类, 即 Surface Web(也称为 Visible Web)和 Deep Web, 其中 Surface Web 包含可以被搜索引擎索引的页面, 而 Deep Web 则是指传统搜索引擎无法索引的 Web 内容, 例如私有 Web 页面、动态内容、未链接的内容、脚本产生的内容等. Deep Web 的规模非常大, 有调查显示, 仅在 2000 年, Deep Web 的规模就已经达到 7.5 PB<sup>①</sup>. 文献[15]综述了在 Deep Web 数据集成方面的工作.

在大数据应用中, 数据集成仍有许多问题值得探讨: (1) 如何在一个系统中集成不同类型的数据源. 例如, Halevy 认为, 在 Surface Web 中包含很多表格, 将这些表格与 Deep Web 相融合将能更好地管理 Web<sup>②</sup>; (2) 如何设计自动化方法实现从数据源到中间模式的映射. 当数据源数目众多且数据模式差异较大时, 仅靠人工手段无法高效完成这项任务, 需要制订一些规则自动生成中间模式; (3) 如何获得完整、简明、一致的数据融合方案, 即方案不仅可以包含数据源中所有相关对象和属性, 也可以包含相关的描述, 这些描述既不赘述也不相悖<sup>[14]</sup>.

## 3 高效的计算方法

### 3.1 分布式计算架构

对于计算量大的任务来说, 一个很自然的方案

就是采用分布式计算架构, 分解整个任务, 并在不同硬件设备上运行子任务, 以减少总计算时间. 分布式数据库系统就是如此. 然而, 如前所述, 分布式数据库系统保留了许多集中式数据库管理系统的特性, 管理开销较大, 当数据规模持续增大时, 过大的管理开销最终导致无法有效地管理数据. 因此, 研究人员转而尝试设计更简洁的分布式计算架构来管理大数据, 例如 MapReduce 技术(特性比较见表 1).

表 1 关系型数据库和 MapReduce 的比较<sup>[7]</sup>

	传统关系型数据库	MapReduce
数据大小	GB	PB
访问	交互式 and 批处理	批处理
更新	多次读写	一次写入、多次读取
结构	静态模式	动态模式
完整性	高	低
横向扩展	非线性	线性

MapReduce 是谷歌公司提出的分布式计算框架, 比分布式数据库系统简洁<sup>[16]</sup>. 该架构包含两个核心概念: Map 和 Reduce. Map 函数将整个任务划分成若干子任务, 并在不同硬件节点执行子任务; Reduce 函数整合 Map 函数产生的中间结果, 生成最终查询结果. Map 任务之间相对独立, 当某一个 Map 任务出现故障时, 可重新启动运行该任务而不会影响其它正在运行的 Map 任务. 用户可以根据任务量大小调整执行 Map 和 Reduce 函数的主机数量. Hadoop 项目也实现了 MapReduce 架构. 表 1 比较了关系型数据库和 MapReduce 的主要区别, 可以看出, MapReduce 所管理的数据规模更大, 可扩展性更佳<sup>[7]</sup>. Dryad 是微软公司提出的分布式计算架构, 它使用无环图(DAG)描述任务的分解, 支持数据并行的应用<sup>③</sup>.

尽管 MapReduce 和 Dryad 框架已经得到了广泛研究与应用, 但是研发新分布式计算架构的脚步并未停滞. 从某种意义上讲, 与分布式数据库系统相比, MapReduce、Dryad 只是从一个极端(至繁)走向了另外一个极端(至简). 分布式数据库系统倾向保留数据库系统的更多特性, 而 MapReduce 和 Dryad 则希望尽量简化数据管理步骤. 正如文献[1]指出的那样, 尽管这些简化版的分布式计算架构已经显示出管理大数据的巨大潜力, 但这并不意味着它们已经是最优选择了, 适当引入某些新特性(例如, 合适

① [http://en.wikipedia.org/wiki/Invisible\\_Web#Size](http://en.wikipedia.org/wiki/Invisible_Web#Size)

② <http://www.cs.washington.edu/homes/alon/files/NSFII2010.pdf>

③ <http://research.microsoft.com/en-us/projects/dryad/>

的索引)之后,系统性能还可能显著攀升.此外,过敏意等人<sup>[17]</sup>认为应该将 GPU 引入到框架中,以满足计算密集和数据密集混合的需求.

### 3.2 数据流技术

如前所述,分布式计算架构利用更多硬件设备来降低计算时间,而数据流技术则侧重提升单台主机的处理能力.令  $N$  表示数据流大小, $\epsilon$  表示误差参数,数据流算法在内存中维护一个大小为  $\text{poly}(\log N, \epsilon)$  的概要数据结构,在必要时也可以根据该数据结构计算查询结果.数据流算法一般仅能够返回近似查询结果,但是当内存资源增加时,查询结果的质量会相应上升.重要的概要数据结构包括 AMS sketch<sup>[18]</sup>、Bloom filter<sup>[19]</sup>、FM-sketch<sup>[20]</sup>等,其中,AMS sketch 能估算频数矩阵;Bloom filter 可判断指定元素是否存在于一个集合之中;FM-sketch 可估算相异元素的个数.典型的模型有界标模型、滑动窗口模型和衰减窗口模型,其中,界标模型考虑所有的数据元素;滑动窗口模型仅考虑最新的  $W$  个元素( $W$  是指窗口的大小);在衰减窗口模型中元组的重要性会随着时间推移而降低.数据流技术仅需单遍扫描数据,从而可以显著降低磁盘 I/O 开销.

分布式数据流管理技术可在分布式网络拓扑环境下管理多数据流.典型的网络拓扑结构有星形结构和树形结构.在星形结构中,所有客户端节点均与一个服务器通过(无线)网络相连接;在树形结构中,所有节点组成一棵树,非叶节点既作为数据中转站,也可以采集、处理数据.在分布式环境下,内存和 CPU 资源则相对丰富,而网络传输带宽资源可能成为瓶颈.分布式数据流技术的目标在于尽量降低网络传输开销,并保证查询结果的质量.

但是,现有数据流研究成果仍旧无法有效管理大数据,需要开发新技术.原因之一在于现有的单数据流解决方案仅利用有限硬件资源(一般指单台计算机),能够在较短时间内处理完“较大”规模数据,却无法快速处理“大规模”数据.原因之二在于现有的分布式数据流技术以降低网络通信开销为主要目的,而网络带宽并不总是系统的资源瓶颈.目前,一些学者已经在探索将数据流技术与 MapReduce 框架相结合,以降低查询处理时间<sup>[21]</sup>;还有一些学者则尝试使用多核处理芯片来加速处理数据流问题<sup>[22]</sup>.总体来说,这些研究均刚刚起步.

### 3.3 新硬件技术

#### 3.3.1 多核处理器

在过去几十年间,业界主要通过提升时钟频率、

优化指令执行、使用缓存等方式来提升处理器的性能.但是当处理器的主频已经达到一个高点且指令执行策略已经比较复杂时,再大幅度提升单个处理芯片的性能就会非常困难.因此,工程师们将研发重点转向多核处理器,即在一个处理器内部嵌入多个“核”,每个“核”都有计算能力.一般来说,每个核的运算能力可能弱于最先进的单核处理器,但是多核的总运算能力则会远高于单核处理器.目前已经出现了 8 核芯片,在未来几年间,单个处理器的核芯片数目还会进一步增加.例如,龙芯 3B 处理器已经集成了 8 个 64 位超标量处理核<sup>①</sup>.从某种意义上讲,一个配备了多核处理器的计算机可被视为共享内存的并行处理系统.

多核芯片的出现对软件开发产生了深远的影响.传统的串行算法无法发挥多核芯片的潜能,必须设计并行算法,将计算任务均匀分摊到各个核之中.多核编程一般采用多线程开发模式.目前主要有两种分解模式,即任务分解和数据分解.任务分解是指依据任务执行的先后次序制定工作流程,再将整个工作流程划分成若干个子任务,相互独立的子任务可以划归不同核来执行,整个任务的完成时间等价于工作流程的最长路径长度.数据分解将待处理的数据集合划分成若干个小数据集合,由不同线程进行管理,最后将所产生的中间结果聚集起来,生成最终结果.加速比(speed-up)是衡量并行算法优劣的重要度量,它描述指定并行算法相对于最佳串行算法的性能提升程度.近期也有一些工作针对多核计算进行了优化.文献[22]提出了一种在多核芯片上挖掘频繁元素的方法,这种基于合作机制的设计模式优于传统的基于竞争机制的设计模式.文献[23]提出了两种利用多核架构来计算 skyline 的方法,即 pskyline 方法和并行 BBS 方法.实验结果表明,pskyline 方法优于并行 BBS 方法.

在海量数据情况下,多核技术仍然面对一些挑战:(1)如何分解任务和整合中间结果.不同数据管理任务的处理难度不尽相同,例如,整体型聚集函数(holistic aggregate function)的结果依赖于全部数据;(2)如何解决负载均衡问题.不同线程所处理的任务量可能并不均匀,从而导致一部分线程空闲而另一部分线程满负荷运转.出现这种情况时,需要动态调整负载,使各核的使用情况相当;(3)需要考虑

① <http://www.loongson.cn/>

超多核情况。在未来,单个处理器的核数目可能会远远高于现阶段,各个核之间的资源竞争更为剧烈,如何设计更优的编程模式成为一个重要问题。

### 3.3.2 GPU 和 FPGA

除了通用处理器 CPU 外, GPU (图形处理单元)、FPGA (现场可编程逻辑器件) 等也频繁用在科学计算上。GPU 可以被看作是一类遵循单指令流多数据流 (Single Instruction stream, Multiple Data stream, SIMD) 架构的并行处理器。当需要在大批量数据上执行相同指令时,能够显著提高执行效率。目前有多种 GPU 编程模型,包括 NVIDIA 的 CUDA、AMD 的流 SDK 和 OpenCL 等。文献[24]提出了一种利用 GPU 集群对海量文档进行聚类的算法,该算法也可以计算文档集合的 TF-IDF 向量。实验结果表明,与单独采用 CPU 的技术相比,同时使用 CPU 和 GPU 之后能将性能提升十倍以上。GPU 也常被用于执行海量数据挖掘任务,例如聚类<sup>[25]</sup>。

现场可编程逻辑器件 (FPGA) 芯片包含大量逻辑门 (比如 AND、OR、XOR、NOT 等),逻辑门之间的配线可以通过软件编程实现。合理设置逻辑门之间的配线可以实现并行计算。文献[26]演示如何基于 FPGA 实现任务级、数据级和流水线级的并行计算,并综述了新近的研究成果。

现有研究成果表明,合理使用 GPU 和 FPGA 等非传统型处理器能够有效提升数据管理效率。但是,熟练使用这些设备并不容易,需要深入理解相关的硬件知识。GPU 面向 SIMD 并行架构,需要从待解决的问题中提取出数据并行的步骤,并进行优化。

### 3.4 高性能计算

高性能计算使用超级计算机和计算机集群来解决复杂计算问题,尤其是针对科学与工程应用中出现的数据处理问题。当前,高性能计算机每秒能进行超过千万亿次的浮点运算,它集成了网络与安全系统管理、并行处理技术,综合了数字电子、计算机系统结构、系统软件、程序语言、算法和计算技术。高性能计算是高技术科技和基础科研领域的工作顺利开展的基石。这些领域通常要对处理对象进行数值模拟、动态显示和结果预测及验证。在过去的几十年中,模拟计算是科学研究的重要内容,用于验证一些未知或者很难分析的理论、模拟一些难以实际操作或代价昂贵或太危险的实验。传统上,计算性能提升的一大途径是商业计算硬件的快速发展,比如处理器、存储器和数据存储系统。

如前所述,大数据不仅仅来源于科学与工程应用,还包括网络数据以及电子商务数据。为适应网络时代的数据存储、分析工作,近几年,部分高性能计算系统的研究工作已经转向计算机集群和网格计算——通过增加系统规模 (如增加组件) 来扩充性能。例如,采用集群计算机来支持商业数据处理,包括数据仓库、在线交易处理等。性能指标不仅仅指 FLOPS (每秒浮点计算能力),还包括图操作处理速度、处理响应速度、高带宽和低延迟等。

数据密集型计算和集群密不可分,而数据分析工作往往不可能完全分割成独立的模块,因此,在集群上的高性能计算与独立的超级计算机相比要面对更多的挑战性任务,其中一个最主要的问题是传输瓶颈:大规模数据在不同处理器间的传输以及在节点间的传输;还有并行处理中的错误检测和处理。为了支持高效的数据密集型计算,还需要继续在硬件、软件系统和算法上做基础研究,而大规模并行算法是必需的技术之一。硬件研究主要包括设计新型的 IC (Integrated Circuit) 用于集成大规模处理器、新的芯片通信框架、系统级的宽带互联网络等。综合考虑软硬件的设计有助于实现可靠的并行计算系统、面向特定应用 (如特别数据结构、特别算法) 的机器设计。软硬件设计的新方法必须满足数据密集型计算的要求,包括针对非数值型应用,如图形操作。特别需要提及的是,这些算法、软件系统和框架必须适应超大规模数据高性能处理的需求,如 PB 数量级<sup>[27-28]</sup>。

## 4 数据分析

### 4.1 语义实体

语义实体是利用机器学习方法从数据中获取的实体,该类实体在一段时间内具有稳定的语义。语义实体可以是人、物、时间日期、事件、产品等,例如“iPad2”、“上海世博会”、“三鹿”等<sup>[29]</sup>。从概念上来讲,语义实体不同于命名实体 (Named Entity)。命名实体识别是基于已知实体集合 (包括人名、地名、机构名、时间、数值、货币、百分比等) 进行实体抽取,而在语义实体研究中却没有预先定义的实体,需要发现和挖掘未知的新实体。

语义实体的识别与检测对于海量数据分析非常重要。例如,在管理 Web 数据时,传统挖掘技术返回一些相关的网页链接,但是需要用户去甄别这些网页链接的意义;而语义实体挖掘则可以直接返回待

挖掘实体的一些相关信息,节省了用户参与的工作量,提高了用户体验程度。微软在 Web 挖掘和 Libra 学术搜索<sup>①</sup>中已经采用了语义实体概念。语义实体识别主要需要解决以下几个问题:(1)待处理文档的规模庞大且数据类型丰富,包括图像、视频、音频、文本、博客数据和论坛数据等新型应用产生的数据;(2)语义实体内部结构比较复杂,无明显特征;(3)大量语义实体属于新出现的对象,比如“三鹿”和各种突发事件等。因而传统的基于统计学习的命名实体识别算法(例如最大熵模型、决策树、支持向量机、隐马尔可夫模型、条件随机场)用于语义实体识别时遇到困难,或者无法使用,或者准确度不高。

在海量数据应用中,语义实体技术面临如下挑战:(1)海量数据的来源多种多样、数据结构复杂化、数据类型也是多样化的,目前的语义实体识别技术仅限于在某个领域进行数据分析与实体识别,如何实现跨领域、跨类型的、自适应的语义实体的研究,是一个崭新的课题;(2)如何结合多种语义实体识别方法处理一项任务,实现多种语义实体识别方法的相互补充,也是值得探讨的一个研究方向;(3)如何设计高效的语义实体识别技术,以实时方式识别出有意义的语义实体或者事件,对于某些舆情分析、论坛监测等应用比较关键;(4)如何结合数据挖掘、统计学习方法和语义 Web(Semantic Web)来识别和管理语义实体,也是一个值得探讨的研究方向。

## 4.2 数据挖掘

数据挖掘技术是从纷繁芜杂的数据集合中抽取有用的知识,以辅助决策支持。典型的数据挖掘算法包括关联规则、分类、聚类、离群点检测等。随着大数据的广泛出现,有些学者开始关注针对大数据的挖掘技术。Han 等人<sup>[30]</sup>认为,必须开发数据密集型挖掘技术,以有效处理大数据(特别是 TB 级别或者 PB 级别规模的数据)。数据密集型挖掘技术的应用背景很广阔,包括信息网络分析、移动对象数据分析等诸多领域。

自然界中互相关联的数据会构成信息网络,包括因特网、交通运输网、无线通信网络、蛋白质网络和社交网络等。这些网络的规模已经非常庞大,而且仍然在不断进化、扩大之中。信息网络中各个节点或者链路包含一些多维信息,节点重要性(依据入度与出度)的差异可能很大,链路权重的差异性也可能非常大。以因特网为例,各主页被视为节点,超链接

被视为链路,从而构成一个信息网络。该信息网络随着时间推移不断进化,节点的重要性可用 PageRank<sup>[31]</sup>等多种方法计算得到。信息网络的主要研究内容包括基于链路的对象排序分析(Link-Based object Ranking, LBR)和图挖掘等。前者关注对象排序,以返回最关键的若干个对象,广泛应用在搜索引擎等领域;后者注重挖掘图中隐藏的知识,例如挖掘频繁子图模式等。

随着基于位置的服务(Location-Based Services, LBS)的普及,亟需发展海量移动对象数据分析和轨迹分析技术。现有定位技术大致可以分为两类,即室外定位技术和室内定位技术,其中室外技术主要是指 GPS 技术,而室内定位技术则比较多样化,包括蓝牙、红外、RFID、Wi-Fi 等。近年来,越来越多的设备可以通过上述定位技术实时采集移动对象的位置,再经由无线网络发送给服务器进行分析。移动对象数据具有很强的时空特性,通常要求获得实时应答。移动对象分析的内容很多,包括移动模式挖掘、周期性模式挖掘、聚类、预测、分类和离群点检测等<sup>[32]</sup>。为了应付超大规模数据, Yang 等人<sup>[33]</sup>采用 MapReduce 框架管理轨迹数据,获得了较好的实验结果。

数据密集型挖掘技术需要关注如何提升计算效率。与一般查询处理问题相比,数据挖掘问题的复杂度往往更高,其计算复杂度往往并不与数据规模成线性增长,而是增长更快。可以充分利用第 3 节中所提到的分布式计算方法、数据流方法、新硬件以及超级计算等方式提升计算效率。Apache Mahout 是一个基于 Apache Hadoop 的项目,它使用了 MapReduce 模式,集成了一批典型的大数据挖掘和机器学习算法,包括聚类、分类、频繁模式挖掘等<sup>②</sup>。此外,知识的表示方式并非一成不变,需要关注在新领域中出现的新的挖掘问题。

## 5 用户接口

### 5.1 海量数据可视化

数据可视化利用图形化手段分析和呈现数据,其呈现方式可以是诸如直方图、饼图、柱状图、折线图统计图表,也可以是许多复杂度更高的图形。数

① <http://academic.research.microsoft.com/>

② <http://mahout.apache.org/>



据可视化技术需要综合利用计算机技术、专业领域知识和设计美学,其中计算机技术是实现基础,专业领域知识能够深入理解数据,而设计美学有助于美化图形显示界面.数据可视化技术并不过分追求绚丽的界面,而要在功能性与美观之间达到一个平衡点.数据可视化的最终目的是呈现数据的内涵,并挖掘出隐藏在大数据内部的规律.换句话说,数据可视化是手段,挖掘知识或规律才是目的.

海量数据可视化是对海量数据的可视化.随着海量数据的应用越来越多,海量数据可视化也得到了越来越多的关注.这里仅列举几个项目.犹他大学的 Vistrails 项目面向工作流和世系管理,集成了良好的可视化界面,可以显示数据演化步骤信息<sup>①</sup>. Visus 项目开发面向海量数据流的可视化技术,以处理美国 LLNL 实验室产生的海量数据<sup>②</sup>.“我们感觉很好(We feel fine)”项目采集互联网的博客信息,从中抽取出表达情感的关键词,并维护一个包含数百万人情感的数据库,在图形界面上以不同颜色和大小圆形表达用户的情感<sup>③</sup>.

目前,海量数据可视化仍然存在诸多挑战:(1)海量数据可视化需要丰富的想象力,这既包括美学想象力,也包括计算机实现技术想象力.前者设计出具有美感的界面,后者能高效地呈现该界面;(2)当数据规模逐步增加时,无法在显示器(或电视墙)上显示所有数据以及所有信息,因而需要深入分析和理解数据内涵,合理划分数据的重要程度,在界面显示时更具有层次感;(3)时间维度往往是数据的重要特征,可以依据该维度动态地淡入和淡出显示内容;(4)海量数据可视化需要把握“全局”与“局部”的关系,不仅可以显示数据的全貌,还可以显示数据的局部特征.

## 5.2 DaaS

数据库服务(Database as a Service, DaaS)<sup>[34]</sup>是数据库外包(Database Outsourcing)在云计算环境中的一种崭新的数据库应用形式<sup>[35]</sup>.DaaS 是指企业(数据所有者或数据库所有者)将自身的数据库创建、访问、维护、升级、管理等任务委托给专门的第三方(服务提供者或数据库服务提供者)进行管理和维护.目前已经存在很多数据库服务,如 MySQL-Hosting<sup>④</sup>、IBM Data Center Outsourcing Services<sup>⑤</sup>等.云计算中的多租户数据库也是一种典型的对多用户提供 DaaS 服务的应用.

图 3 用面向服务(SOA)的观点刻画了数据库

服务的架构.该架构包含三个角色(分别是数据拥有者、数据库服务提供者和数据请求者)和数据流(分别是数据源、查询与结果等).体系结构可以是统一客户模式(Unified Client Model)、多查询者模式(Multi-Querier Model)和多数据拥有者模式(Multi-Owner Model)等<sup>[36]</sup>.DaaS 的主要研究内容包括委托数据库的可用性、可靠性、数据加密、分布式存储、备份机制、多版本管理、数据更新机制、基于隐私保护的查询、访问控制等.

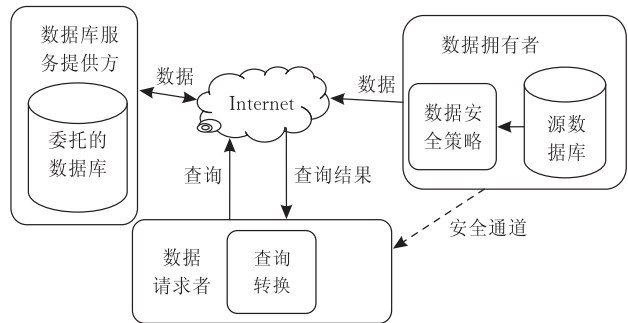


图 3 DaaS 的框架结构<sup>[34]</sup>

在面向数据密集型的应用中,DaaS 存在以下问题和挑战:(1)查询的高效性和有效性.从查询性能上看,传统的数据外包中的查询主要是对数值型数据的检索,数据量比较小.而在数据库服务提供者的数据中心的数据量很大,需要考虑如何降低数据网络传输量.从查询有效性方面,目前只有在基于数据分布的数值型数据上实现了精确匹配查询,而数据中心所存储的数据类型多种多样,需要对多种类型数据进行有效查询;(2)隐私数据管理.目前大多数研究工作主要关注如何保护委托数据库中用户的敏感信息,较少考虑数据被存放在数据库服务提供方,用户可能需要查询私有信息,因而需要考虑隐私数据查询技术,保护查询者的身份和隐私数据;(3)访问控制技术.由于数据库服务提供方不可信,因而 DaaS 的数据访问控制非常重要,它是控制用户执行合法授权的第一步.需要探索灵活、有效的访问控制机制,将访问控制和保证数据安全和隐私的方式结合起来,以增强 DaaS 的可用性;(4)多版本数据的备份和更新机制.多版本技术能保证查询效率和数据的高可用性,也是目前的一个研究热点.

① <http://vistrails.sci.utah.edu>

② <http://www.pascucci.org/visus>

③ <http://www.wefeelfine.org/>

④ <http://adhost.com/>

⑤ <http://www-1.ibm.com/services>

### 5.3 云计算

云计算是一种典型的网络计算方式,是以数据为中心的一种超级计算,具有在虚拟计算环境下的可扩展性和可用性.目前,亚马逊、微软、谷歌、IBM 等公司都提出了“云计划”,如亚马逊的 AWS(Amazon Web Services)、IBM 和谷歌联合进行的“蓝云”计划等.学术界也纷纷对云计算展开了深层次的研究,如谷歌、华盛顿大学以及清华大学开展的云计算学术合作计划,卡内基梅隆大学等对数据密集型的超级计算的研究,本质上也都是对云计算的相关技术展开研究.实践证明,全球财政在云计算的各种服务上的投入以每年约 30% 以上的速度递增.本节主要讨论云计算环境给软件开发和演化带来的影响和挑战.

云计算从 3 个方面融合了现有技术,包括使用方式、计算方式以及基础设施的部署方式等.使用方式采用服务的形式,即 SaaS、IaaS、PaaS 和 DaaS;计算方式是分布式的,即网格计算和 SOA;基础设施的部署方式即虚拟化,如集群计算.互联网平台为软件行业的发展带了新的机遇,如高生产率、快速反馈、便于在线升级等<sup>[37]</sup>,但是网络软件组织使用模式(如任务协作性、模块并行性、业务装配化等)和网络数据生成的特性(如数据的分散性和多样性、资源的异构性、规模庞大性等)使得人们把目光从对单机的软件开发与演化系统的研究转向了与社会网络以及分布式大规模并行系统紧密关联的相关问题的探索和研究,如社会软件工程<sup>[38]</sup>、软件自适应演化<sup>[39]</sup>、软件可信评估<sup>[40]</sup>与选择以及超大规模系统的演化<sup>[41]</sup>等.

在云计算环境中,软件开发和演化的挑战课题主要有以下几个方面:(1)软件系统的可扩展性.软件使用环境的变化,包括软件系统的网络化,(异构的、非数值型的)数据量的激增,社会协同的分布式作业的产生(网络社区)等等,使得各种软件组件有必要实施同时作业,而这样一种工作模式极大地增加了软件创建、测试和维护的复杂度.异构设备和操作环境加剧了解决这类问题的难度.目前解决这类问题的主要手段是人工作业,需要花费大量精力并且通过不断地下载和安装补丁来保证系统的可信性;(2)提高软件生产力.网络发展推动了协同工作,加大了软件任务的复杂度,因而增加了软件开发的难度.在软件行业的发展过程中,由于协作性工作任务增多,以人为中心的软件设计和开发模式变得非常重要,即软件开发将更侧重与人密切相关的

上层抽象设计而不是那些底层的系统软件.从而使得软件开发可以快速理解和掌握开发以及维护过程;(3)由于新的分布式应用的出现,如物联网等,这些新的应用要求我们针对这些异构的系统设计出新颖的软件系统,要求在编程语言、算法、分析工具和系统软件上下功夫从而简化软件开发人员的工作增加工作效率.

## 6 数据质量与隐私保护

### 6.1 数据质量

数据质量极大地影响到海量数据的管理过程.低质量的数据集合可能无法使用,或者产生质量低下下的查询分析结果.影响数据质量的因素很多,包括仪器精度、外部环境、人工误操作、历史因素等.例如,在传感网络中产生的数据的精度往往受到传感器本身精度的制约,周围环境的复杂度也会影响数据传输质量.数据质量的具体表征有:数据不准确、数据缺失、数据不一致等.据报道,大型机构所产生的数据中 60% 都是冗余的<sup>[42]</sup>.此外,RFID 技术现已广泛应用到物流等多个行业之中,而 RFID 数据的准确度仅达到 60%~70%<sup>[43]</sup>.

提升原始数据质量的措施很多,包括提高物理仪器精度、简化运行环境、加强人员培训、提升设计质量等.这些措施能够提升数据质量,但却无法消除数据质量问题.待处理的数据质量可能仍然不高,有待进一步的提升.在此过程中,自动检测与修复技术就显得非常关键,即以常识和专业基础知识为前提,通过分析 and 理解数据内容以及数据之间的联系来实现问题数据的自动检测和修复.强调自动化具有两层含义,首先,对海量数据进行人工干预的成本非常高昂,其次,自动检测和修复能够提高时效性.一般有两类方法可以进行自动检验.一种是与应用无关的检验技术(普适的技术),包括缺失值处理、去除重复等.另一种是与特定应用相关的检测技术,需要借助相关领域的具体知识才能够做到<sup>[42]</sup>.数据世系也是一种提升数据质量的手段.数据世系保存了数据从产生开始不断演变的整个过程.追溯这个过程即可显示数据质量变化的历程,从而帮助提升数据质量.

敖莉等人<sup>[42]</sup>回顾了重复数据删除技术,认为主要有两类技术手段:相同数据的检测技术和相似数据的检测与编码技术.而相同数据的检测又分为相同文件以及相同数据块两个层次.这方面的处理比

较简单,可以通过散列或者滑动块技术进行管理,而相似数据的查找则稍显困难一点,需要定义相似度距离,以编码技术进行求解。

事实上,在大多数情况下无法获取完美质量的数据集合,更何况这个数据集合本身是海量规模。即使在应用了高级的数据质量提升技术之后,也是如此。因此,一个比较重要的任务即在于如何从弱质量数据上执行数据管理任务。不确定数据管理技术是一种在弱质量数据上执行查询分析任务的手段。数据质量不高时,以概率密度函数的形式表征数据,而非某一个特征值。

## 6.2 数据安全

数据的安全和数据所涉及到的隐私保护是数据管理需要解决的重要问题之一。数据安全和隐私保护技术涉及各种应用领域,例如教育、医疗、能源、交通、国土安全、数字化民主、经济等。这些数据(如销售数据、工资信息、医疗记录、身份信息、证券交易信息)在一定程度上涉及到个人的私密信息,例如,销售记录反映出消费者的个人购买兴趣,医疗记录反映了患者的身体状况,银行记录透露了储户的财产信息,这些信息的安全保护措施非常重要。如 2005 年 6 月 4 亿多 Master 信用卡持有者的姓名和信用卡号被泄露<sup>①</sup>,2006 年 5 月存放 2 亿多老兵姓名、社会安全号码和出生日期的信息的硬盘被盗,都造成了大量隐私信息泄露<sup>②</sup>,对个人以及社会造成巨大损失。因此,数据安全和敏感数据的隐私保护是数据管理中贯穿数据获取、数据存储、数据维护、数据查询、数据分析等整个数据生命周期的一个重要研究内容。

可信系统和信息安全的研究是基础研究。传统数据库系统所提供的各项技术手段能较好地解决封闭应用环境下数据的安全和隐私保护问题,但随着互联网技术和应用的普及,给在开放的、分布的互联网上面向信息共享的敏感数据隐私保护技术提出了新的挑战,也成为多租户数据库、数据服务外包等应用的重要问题。面向信息共享的敏感数据隐私保护技术主要包括以下几个方面:

- (1) 完善的数据安全机制,保证数据的内容不会泄露(数据的机密性<sup>[44]</sup>);
- (2) 数据内容不会被破坏(数据的完整性<sup>[45]</sup>);
- (3) 数据是正确的,返回客户的结果是完备的(数据的完备性<sup>[46]</sup>);
- (4) 服务提供者无法察觉客户查询数据的目的(查询隐私保护<sup>[47]</sup>);
- (5) 数据的可用性。

现有很多数据安全技术大多属于被动防守技术,例如加密、防火墙、锁等。随着新型应用(例如社交网络)的广泛出现,需要设计更多主动和被动技术,以防范内、外攻击者的攻击。内部攻击包括即使某些不被信任的用户可能共享了一部分社交网络空间,也要能保护整个应用的数据安全。这方面的崭新研究方向包括:(1) 分析海量数据,利用群体智慧来发现系统攻击、过滤垃圾信息、检测社会工程攻击;(2) 建立信息安全基础,包括安全的硬件、软件和网络。目前应用工业实践的方法仍不充分,针对基础架构和过程控制系统的信息安全需要完备的理论基础,有必要建立在传统和量子理论基础上的认证、授权和可信管理的基础机制;(3) 隐私保护需要行为学和认知科学的支持;(4) 安全和隐私政策的形式化定义和应用<sup>③</sup>。

## 6.3 数据监护

数据监护(data curation)<sup>[48]</sup>来源于图书馆领域的“digital curation”<sup>④</sup>,是指对数据资产的选择、维持、维护、收集和打包的整个过程,主要目的是解决数据的保管和使用问题。据 OCLC 统计<sup>⑤</sup>,1998 年、1999 年和 2000 年存在的网站 IP 到 2002 年时分别仅剩 13%、19% 和 33%,因此,Web 数据的保存问题很严峻。对于科学家、研究者和学者花费了大量时间和精力所产生的大量科学数据,保管和使用非常重要。一旦数据遗失,将使科研成果不完整,并且严重影响其使用价值和创新价值,然而目前的现状是往往一个项目产生的数据在项目结束后就很难再继续使用。

学术界开始研究数据监护,尝试解决数据的有效存储和使用问题。英国 JISC (Joint Information Systems Committee) 对数据监护给出的定义是:为了确保数据的当前使用目的,并能用于未来再发现及再利用,从数据产生开始就对其进行管理和完善的活动。数据监护意味着需要进行持续性补充和更新,以使数据符合用户需求,发挥其使用价值。目前,数据监护逐渐成为国外图书馆界科研和实践的新热点,2007 年美国 NSF 启动了 DataNet 计划<sup>⑥</sup>,以图

① <http://www.nytimes.com/2005/06/20/technology/20credit.html>

② <http://www.nytimes.com/2006/05/25/washington/25identity.html>

③ <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd-report-2010.pdf>

④ [http://en.wikipedia.org/wiki/Digital\\_curation](http://en.wikipedia.org/wiki/Digital_curation)

⑤ <http://www.oclc.org/research/activities/past/orprojects/wcp/stats/misc.htm>

⑥ <http://www.nsf.gov/pubs/2007/nsf07601/nsf07601.htm>

书馆为主体,用 1 亿美元、5 年时间资助数据监护的研究.伊利诺伊大学的 DCEP(Data Curation Education Program)<sup>①</sup>计划把数据监护的教学包含的课程定义为:数据采集与管理、知识表达、数字资源保护与存档、数据标准、数据政策等,课程包括信息组织、信息建模、本体论和元数据理论和实践等.

在数据库领域,国际知名的数据库系统和数据库理论专家 Peter Buneman 于 2004 年最早在爱丁堡大学成立了英国数据监护中心(UK Digital Curation Center)<sup>②</sup>,是数据监护技术从数据库领域进行探索的第一个科研机构. Peter 教授在数据监护的基础上,提出了监护数据库(Curated Database)概念<sup>[49]</sup>.从数据管理系统的角度看,监护数据库的挑战包括:数据的标注、数据模式的演化、数据内容的正确性、数据世袭的管理、数据的灵活访问、数据的发布和引用等.在最新的工作中<sup>[50]</sup>,他们实现了 Database+Wiki=Curated Database 的原型系统,并在 SIGMOD2011 上演示,这个系统通过 XML 模型来管理演化的数据库模式、利用 Wiki 实现了简洁的版本管理和世系管理、提供了对结构化数据的 SQL 的关系查询.

目前监护数据库还刚刚起步,如何监护海量数据库,是一个挑战性的问题.由于数据量大、数据内部的结构多元化、数据之间的关联错综复杂,由此带来的数据标注的工作量巨大,而且数据的一致性、完整性很难维护,数据世系的分析也成为一难点.从数据模式的演化来看,虽然 XML 可以建模数据模式的演化问题,但是,针对海量数据的 XML 数据管理以及高效的数据检索和查询,都是目前需要解决的问题.

## 7 新型应用分析

本节举例分析两个新型大数据应用,即社交计算和信息-物理融合网络.

### 7.1 社交计算

社交计算(social computing)<sup>③</sup>是当前网络应用中的一个热点技术,其主要研究对象是社交网络.社交网络的主要特点是拥有个性数据(包括爱好和经历)的用户参与到合作式的信息交互中(人际关系网络化),从而使互联网从门户网络发展到社交网络.社交计算的相关概念包括社交网络(social network)、社交媒体(social media)、社交网络服务(social network service)等.从结构上看,可以将每个用户视作

一个节点,且附加一些个人信息,如住址、邮件等;可以将用户之间的各种社会关系视作边,这类边往往可以加上权重,从而构成了一个复杂的社交网络图.在社交网络图中,用户直接或者间接地产生网络信息.提供社交计算功能的网站称为社交媒体.相对于以往的 Web 应用,这些社交媒体除了传统的文本和多媒体外,还带有标记、注释、评分、评论等信息,可以向用户提供更加个性化和社会性的信息服务——社交网络服务.典型的社交媒体包括 Facebook、Twitter、Wikipedia 等.

社交计算是典型的大数据应用.社交计算涉及的用户量非常巨大,数据量也很大.据最新统计<sup>④</sup>,全球互联网用户数已经达到 20 亿,Facebook 的用户数在 2010 年末已经达到 6 亿,互联网上博客的数量约为 1.52 亿,2010 年的 Twitter 帖子数达到 250 亿.因此,根据用户发布的消息、所作的评论、评分等数据抽取出知识(例如用户的兴趣等)就显得极为重要,需要综合采用数据管理、自然语言处理、机器学习和信息检索等技术.社交计算的挑战性在于每个用户作出的决定既受到他人的影响,也影响到他人,从而在整个网络内产生“聚集”效果;此外,社会计算通常含有“博弈论”思想,因此相关的社会学问题都属于社交计算的范畴,例如圈子的动态进化、信息与利益的形式化描述等.

在社交计算中,互联网的大量用户、海量数据会带来如下挑战:(1)在复杂系统和互联网环境下产生的“collective intelligence”<sup>[51]</sup>,需要对用户协作行为和解决问题的方法进行深入探讨.在给定目标、任务的情况下,如何更好地组织分布的大量用户、海量数据形成的复杂网络,实现不同的应用需求?从经济模型来看,针对一个合作的任务,需要研究参与该任务的用户兴趣动机、激励机制和经济模型等.从计算技术来讲,针对一个复杂任务,如何实现利用分布的人力来共同完成一个子任务间的协调?(2)传统搜索引擎是基于信息网络(Information Network),而在社交网络的框架下,如何做知识抽取和信息挖掘,以及用户隐私的保护和信任关系的衡量.事实上,除搜索引擎和问答站点外,已经有越来越多的人将 Twitter 作为信息获取的一种方式<sup>[52]</sup>,除了实时、广泛的特点外,这种模式是在与真实的人交流,而不

① <http://www.lis.illinois.edu/events/2011/02/18/data-curation-education-program-tei-workshop>

② <http://www.dcc.ac.uk>

③ [http://en.wikipedia.org/wiki/Social\\_computing](http://en.wikipedia.org/wiki/Social_computing)

④ <http://www.birduser.net/>

是提交问题给某个算法/机器系统。

综上所述,探索面向大数据的数据管理技术和观点挖掘技术,利用复杂网络技术解决社会学问题、实现计算机的自动识别、理解、管理大量分布的自由的人和设备,成为目前非常值得探讨的研究课题。

## 7.2 信息-物理融合网络

因特网的出现使得人们可以在计算机之间共享信息,促进了整个信息产业的发展。时至今日,更多物理设备(不仅仅是计算机)具备了数据传输功能(有线传输或者无线传输),融入到因特网之中,从而扩大了因特网的内涵和范畴,典型案例包括智能家电、车载系统、手持设备、野外观测设备等。这个融入了诸多物理设备的网络系统被称为信息-物理融合系统(CPS),其宗旨在于通过3C(Computation、Communication、Control)技术来有效融合整个系统之中的物理设备,实现实时感知、动态控制和信息服务,如图4所示<sup>[53]</sup>。

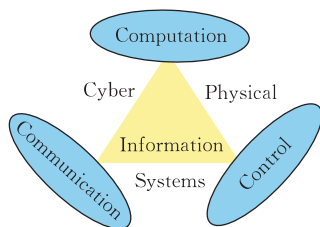


图4 CPS的3C概念图

CPS系统管理大量数据,它们来源不同、结构各异且并不一致,既包括各种传感器实时采集的环境数据,也包括其它渠道产生的动态数据集合,例如Web信息、生产日志等。CPS系统中的数据也需要考虑时间-空间属性,分析历史时间,预测未来走势。例如,许多车辆内置了GPS通信模块,可以实时采集位置信息。用户可据此获得基于位置的服务,即:当用户的位置改变时,所得到的服务是与他的位置信息紧密相关的。

传感器是CPS的重要组件之一,它采集周围环境的各类信息,从而实现将物理世界和信息世界融为一体。典型的传感器包括温度传感器、湿度传感器、红外传感器等。在未来,人们将研发功能更强、精度更高、体积更小、能耗更低、成本低廉的传感器。首先,不同应用需要使用具备特定感知能力的传感器,但是有些需求由于应用面比较狭窄,并未被研制出来,例如在化学和生物等领域之中。其次,缩小传感器的体积会充分激发人类的想象力,制造出更新奇的设备。微机电系统(Microelectromechanical system, MEMS)就是一个较好的例子。与传统机械相比,

MEMS的尺寸更小,小到几个微米,大到不足1个厘米,而在狭小的空间内却集成了微型机构、微型传感器、微型执行器和辅助电路等几个部分,在医疗、工业、测试仪器等领域具有广泛的用途。此外,降低能耗、减少成本均可促使传感器应用更加普及。

CPS数据管理面临较大的挑战:(1)异构数据融合。要克服不同数据来源的不一致性,获取一致的结果;(2)数据以流的形式持续到达,并期待实时的查询结果。传统的数据管理方法主要从一个静态数据集中执行快照查询,并返回查询处理结果,这并不符合CPS系统的场景。需要采用连续查询方式,首先在系统中注册查询,当输入数据到达时,即可处理查询,并实时刷新查询结果;(3)需要提高CPS系统的自动化程度,以构建决策支持系统。高级传感器为CPS自动化提供了基础设施;高效的数据分析能力可以帮CPS系统解决特定问题;而CPS系统的自动化技术(或半自动化技术)则从系统层面解决问题。CPS系统自动化的应用领域非常广泛,包括智能交通系统、无人驾驶飞机、工业机器人等,不仅可以提高工作效率,还可以在极端条件下工作。但是实现这一点并非那么容易,我们不仅需要开发高级对象识别技术,提高对象鉴别能力,还需要新的学习算法来适应周围的环境。

## 8 总 结

数据密集型科学与工程(DISE)诞生于当今的数据爆炸时代,传统的数据管理技术无法有效应对大数据。本文从数据存储与组织、计算方法、数据分析、用户接口、数据质量、数据安全和数据监护等方面介绍了DISE的研究内容和面临的挑战:(1)分布式存储策略是主流的数据存储与组织形式,另外在许多应用中还涉及数据世系和数据集成等内容;(2)可以采用多种方式提升计算能力,包括分布式计算架构、数据流、新硬件、高性能计算等;(3)需要发展数据分析技术以获取知识,典型的内容包括语义实体分析和数据挖掘;(4)需要提供友好的用户接口,例如海量数据可视化、数据库服务和云计算等;(5)DISE还关注数据质量和数据安全问题,并且执行数据监护。总之,与传统数据管理技术相比,DISE在各个层面的差异都非常显著。尽管目前已经有一些相关的研究工作,但是总体上来说,DISE仍然非常年轻,诸多问题尚待研究。

## 参 考 文 献

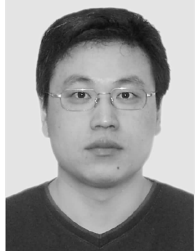
- [1] Hay T, Tansley S, Tolle K. The Fourth Paradigm: Data-Intensive Scientific Discovery. Redmond, Washington: Microsoft Research, 2009
- [2] Big data: Science in the petabyte era. *Nature*, 2008, 455: 1-136
- [3] Dealing with data. *Science*, 2011, 331(6018): 639-806
- [4] Xiao Nong. Data-intensive computing. *Communications of the CCF*, 2011, 7(7): 6-7(in Chinese)  
(肖依. 数据密集型计算. 中国计算机学会通讯, 2011, 7(7): 6-7)
- [5] Gao Ming, Jin Che-Qing, Wang Xiao-Ling, Tian Xiu-Xia, Zhou Ao-Ying. A survey on management of data provenance. *Chinese Journal of Computers*, 2010, 33(3): 373-389 (in Chinese)  
(高明, 金澈清, 王晓玲, 田秀霞, 周傲英. 数据世系管理技术研究综述. 计算机学报, 2010, 33(3): 373-389)
- [6] Ghemawat S, Gobioff H, Leung S T. The Google file system//Proceedings of the 19th ACM Symposium on Operating Systems Principles. Bolton Landing, NY, 2003: 29-43
- [7] White Tom. Translated by Zhou Min-Qi, Wang Xiao-Ling, Jin Che-Qing, Qian Wei-Ning. Hadoop: The Definitive Guide. 2nd Edition. Beijing: Tsinghua University Press, 2011(in Chinese)  
(White Tom 著. 周敏奇, 王晓玲, 金澈清, 钱卫宁译. Hadoop 权威指南. 第 2 版. 北京: 清华大学出版社, 2011)
- [8] Chang F, Dean J, Ghemawat S, Hsieh W C, Wallach D A, Burrows M, Chandra T, Fikes A, Gruber R. Bigtable: A distributed storage system for structured data//Proceedings of the 7th Symposium on Operating Systems Design and Implementation. Seattle, WA, 2006: 205-218
- [9] Abadi D, Madden S, Hachem N. Column-Stores vs. Row-Stores: How different are they really?//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data. Vancouver, BC, Canada, 2008: 967-980
- [10] Ni Lionel, Luo Wu-Man. The technology transform in the era of data explosion. *Communications of the CCF*, 2011, 7(7): 12-20(in Chinese)  
(倪明选, 罗吴蔓. 数据爆炸时代的技术变革. 中国计算机学会通讯, 2011, 7(7): 12-20)
- [11] Liao Xiao-Fei, Fan Xue-Peng, Xu Fei, Li He, Jin Hai. Large-scale data-intensive computing system. *Communications of the CCF*, 2011, 7(7): 33-40(in Chinese)  
(廖小飞, 范学鹏, 徐飞, 李鹤, 金海. 数据密集型大规模计算系统. 中国计算机学会通讯, 2011, 7(7): 33-40)
- [12] Buneman P, Khanna S, Tan W C. Why and where: A characterization of data provenance//Proceedings of the 17th International Conference on Data Engineering. London, UK, 2001: 316-330
- [13] Davidson S B, Freire J. Provenance and scientific workflows: Challenges and opportunities//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data. Vancouver, BC, Canada, 2008: 1345-1350
- [14] Bleiholder J, Naumann F. Data fusion. *ACM Computing Surveys*, 2008, 41(1): 1-41
- [15] Liu Wei, Meng Xiao-Feng, Meng Wei-Yi. A survey of Deep Web data integration. *Chinese Journal of Computers*, 2007, 30(9): 1475-1489(in Chinese)  
(刘伟, 孟小峰, 孟卫一. Deep Web 数据集成研究综述. 计算机学报, 2007, 30(9): 1475-1489)
- [16] Dean J, Ghemawat S. MapReduce: Simplified data processing on large clusters//Proceedings of the 6th Symposium on Operating System Design and Implementation. San Francisco, CA, 2004: 137-150
- [17] Guo Min-Yi, Xu Cheng-Zhong. Programming model and self-adaptive resource management for cloud computing. *Communications of the CCF*, 2011, 7(7): 26-33(in Chinese)  
(过敏义, 须成忠. 云计算的编程模型与自适应资源管理. 中国计算机学会通讯, 2011, 7(7): 26-33)
- [18] Alon N, Matias Y, Szegedy M. The space complexity of approximating the frequency moments//Proceedings of the 28th Annual ACM Symposium on Theory of Computing. Philadelphia, Pennsylvania, USA, 1996: 20-29
- [19] Bloom B. Space/time tradeoffs in hash coding with allowable errors. *Communications of the ACM*, 1970, 13(7): 422-426
- [20] Flajolet P, Martin G N. Probabilistic counting algorithms for data base applications. *Journal of Computer and Systems Sciences*, 1985, 31(2): 182-209
- [21] Li B, Mazur E, Diao Y, McGregor A, Shenoy P. A platform for scalable one-pass analytics using MapReduce//Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data. Athens, Greece, 2011: 985-996
- [22] Das S, Antony S, Agrawal D, Abbadi A E. Thread cooperation in multicore architectures for frequency counting over multiple data streams. *Proceedings of the VLDB Endowment*, 2009, 2(1): 217-228
- [23] Park S, Kim T, Park J, Kim J, Im H. Parallel skyline computation on multicore architectures//Proceedings of the 25th International Conference on Data Engineering. Shanghai, China, 2009: 760-771
- [24] Zhang Y, Mueller F, Cui X, Potok T. Data-intensive document clustering on graphics processing unit (GPU) clusters. *Journal of Parallel and Distributed Computing*, 2011, 71(2): 211-224
- [25] Che S, Boyer M, Meng J, Tarjan D, Sheaffer J W, Skadron K. A performance study of general-purpose applications on graphics processors using CUDA. *Journal Parallel Distributed Computing*, 2008, 68(10): 1370-1380
- [26] Müller R, Teubner J. FPGAs: A new point in the database design space//Proceedings of the 13th International Conference on Extending Database Technology. Lausanne, Switzerland, 2010: 721-723

- [27] Kindratenko V, Enos J, Shi G, Showerman M, Arnold G, Stone J, Phillips J, Hwu W. GPU clusters for high-performance computing//Proceedings of the 2009 IEEE International Conference on Cluster Computing. New Orleans, Louisiana, USA, 2009: 1-8
- [28] Fuller S, Millett L. Computing performance: Game over or next level? IEEE Computer, 2011, 44(1): 31-38
- [29] Cai P, Luo H, Zhou A. Semantic entity detection by integrating CRF and SVM//Proceedings of the 11th International Conference on Web-Age Information Management. Jiuzhaigou, China, 2010: 483-494
- [30] Han J, Gao J. Research challenges for data mining in science and engineering//Kargupta H ed. Next Generation of Data Mining, Chapman and Hall/CRC, 2008
- [31] Page L, Brin S, Motwani R, Winograd T. The pagerank citation ranking: Bringing order to the web. Computer Science Department, Stanford University; Technical Report 422, 1998
- [32] Han J, Li Z, Tang L A. Mining moving object, trajectory and traffic data (tutorial)//Proceedings of the 15th International Conference on Database Systems for Advanced Applications. Tsukuba, Japan, 2010: 485-486
- [33] Yang B, Ma Q, Qian W, Zhou A. TRUSTER: Trajectory data processing on clusters//Proceedings of the 14th International Conference on Database Systems for Advanced Applications. Brisbane, 2009: 768-771
- [34] Tian Xiu-Xia, Wang Xiao-Ling, Gao Ming, Zhou Ao-Ying. Database as a service — Security and privacy preserving. Journal of Software, 2010, 21(5): 991-1006(in Chinese)  
(田秀霞, 王晓玲, 高明, 周傲英. 数据库服务——安全与隐私保护. 软件学报, 2010, 21(5): 991-1006)
- [35] Hacigumus H, Mehrotra S, Iyer B. Providing database as a service//Proceedings of the 18th International Conference on Data Engineering. San Jose, CA, 2002: 29-40
- [36] Mykletun E, Narasimha M, Tsudik G. Authentication and integrity in outsourced databases. ACM Transactions on Storage, 2006, 2(2): 107-138
- [37] Yin Gang, Wang Huai-Min, Yuan Lin, Zhu Yan-Xu, Shi Dian-Xi, Mi Hai-Bo. Construction of Internet-based trustworthy software Production service system. Journal of Frontiers of Computer Science and Technology, 2011, 5(10): 880-890(in Chinese)  
(尹刚, 王怀民, 袁霖, 朱沿旭, 史殿习, 米海波. 构造基于互联网的可靠软件生产服务系统. 计算机科学与探索, 2011, 5(10): 880-890)
- [38] Sarma A. A survey of collaborative tools in software development. Institute for Software Research, Donald Bren School of Information and Computer Science, University of California, Irvine: UCI-ISR-05-3, 2005
- [39] Yang Fu-Qing, Lv Jian, Mei Hong. Network configuration software technology system: An architecture-centric approach. Scientia Sinica E: Information Sciences, 2008, 38(6): 818-828(in Chinese)  
(杨芙清, 吕建, 梅宏. 网构软件技术体系: 一种以体系结构为中心的途径. 中国科学 E 辑: 信息科学, 2008, 38(6): 818-828)
- [40] Davide Taibi, Luigi Lavazza. Sandro morasca; OpenBQR: A framework for the assessment of OSS. IFIP Working Group 2.13 on Open Source Software, Limerick, Ireland, 2007: 173-186
- [41] Northrop L, Feiler P, Gabriel R, Goodenough J, Linger R, Longstaff T, Kazman R, Klein M, Schmidt D, Sullivan K, Wallnau K. Ultra large-scale systems: The software challenge of the future. Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 2006
- [42] Ao Li, Shu Ji-Wu, Li Ming-Qiang. Data duplication techniques. Journal of Software, 2010, 21(5): 916-929 (in Chinese)  
(敖莉, 舒继武, 李明强. 重复数据删除技术. 软件学报, 2010, 21(5): 916-929)
- [43] Jeffery Shawn R, Garofalakis Minos, Franklin Michael J. Adaptive cleaning for RFID data streams//Proceedings of the 32nd International Conference on Very Large Data Bases. Seoul, Korea, 2006: 163-174
- [44] Hacigümüs H, Iyer B, Li C, Mehrotra S. Executing SQL over encrypted data in the database service provider model//Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data. Madison, Wisconsin, 2002: 216-227
- [45] Wang H, Yin J, Perng C S, Yu P S. Dual encryption for query integrity assurance//Proceedings of the 17th ACM Conference on Information and Knowledge Management. Napa Valley, California, 2008: 863-872
- [46] Mouratidis K, Sacharidis D, Pang H. Partially materialized digest scheme: An efficient verification method for outsourced databases. The VLDB Journal, 2009, 18(1): 363-381
- [47] Ge T, Zdonik S B. Answering aggregation queries in a secure system model//Proceedings of the 33rd International Conference on Very Large Data Bases. Vienna, Austria, 2007: 519-530
- [48] Yang He-Lin. Data curation: A new development of university libraries in the U. S. Journal of Academic Libraries, 2011, 29(2): 18-21(in Chinese)  
(杨鹤林. 数据监护: 美国高校图书馆的新探索. 大学图书馆学报, 2011, 29(2): 18-21)
- [49] Buneman P, Cheney J, Tan W C, Vansummeren S. Curated databases//Proceedings of the 27th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems. Vancouver, BC, Canada, 2008: 1-12
- [50] Buneman P, Cheney J, Lindley S, Müller H. DBWiki: A structured wiki for curated data and collaborative data management//Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data. Athens, Greece, 2011: 1335-1338

- [51] Siegel David A. Social networks and collective action. *American Journal of Political Science*, 2009, 53(1): 122-138
- [52] Kwak H, Lee C, Park H, Moon S. What is twitter, a social network or a news media? // *Proceedings of the 19th International Conference on World Wide Web*. Raleigh, North

Carolina, USA, 2010: 591-600

- [53] He Ji-Feng. Cyber-physical systems. *Communications of the CCF*, 2010, 6(1): 25-29(in Chinese)  
(何积丰. Cyber-physical systems. *中国计算机学会通讯*, 2010, 6(1): 25-29)



**GONG Xue-Qing**, born in 1974, associate professor. His research interests include database technology, Web data management etc.

**JIN Che-Qing**, born in 1977, Ph. D., professor. His research interests include data stream management, uncertain data management, location-based services.

**WANG Xiao-Ling**, born in 1975, professor, Ph. D. supervisor. Her research interests include XML data management, Web service.

**ZHANG Rong**, born in 1978, Ph. D., associate professor. Her research interests include Web services and cloud computing.

**ZHOU Ao-Ying**, born in 1965, professor, Ph. D. supervisor. His research interests focus on data management and information system, inclusive of Web data management, Chinese Web infrastructure, Web searching and mining, data streaming and mining, complex event processing and real-time business intelligence, uncertain data management and applications, data-intensive computing, distributed storage and computing, peer to peer computing and management, Web services.

## Background

This paper introduces Data Intensive Science and Engineering (DISE) that belongs to the database category. DISE is the discipline to manage big data, which now widely exists in science researches and engineering practices. Although traditional data management techniques have achieved great success in past decades, such techniques cannot manage PB-scale data straightforwardly, so that it is necessary to develop novel solutions to handle the issues. More and more researchers all over the world has recognized the importance of the big data. In September, 2008, *Nature's* special issue on big data says that we are now in the petabyte era. In February, 2011, *Science's* special issue on dealing with data focu-

ses on the challenges and the opportunities on scientific data deluge. In July, 2011, CCF communication's special issue discusses Data Intensive Computing (DIC) from several different aspects.

This paper aims at providing the whole picture for DISE, from data storage and organization, computational method, data analysis, to the middlewares to communicate with users. Moreover, data quality and privacy preservation are two key factors that cannot be ignored. Besides introducing the present progress, we also try to point out the main challenges for the future work. DISE is critical in many areas and there still have rich topics to be conquered.