

大数据技术原理与应用课程建设经验分享

林子雨

厦门大学信息科学与技术学院, 福建 厦门 361005

摘要

大数据专业人才的培养是世界各国新一轮科技较量的基础,高等院校承担着大数据人才培养的重任。大数据专业作为典型的“新工科”专业,在课程体系建设方面还处于摸索阶段。首先剖析了大数据课程建设的难点,然后介绍了厦门大学建设的大数据课程体系,包括入门级课程、进阶级课程和实训课程,最后分享了大数据技术原理与应用课程建设的经验与方法,包括课程定位、培养目标、预备知识、大数据与云计算课程之间的知识切割、课程内容与学时安排、课程教材、实验环境搭建、配套资源建设、在线服务平台、线下培训与交流等。

关键词

大数据;课程体系;MOOC;公共服务平台

中图分类号:G423

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2018058

Course construction experience sharing of the principles and applications of big data technology

LIN Ziyu

School of Information Science and Engineering, Xiamen University, Xiamen 361005, China

Abstract

The training of big data professionals is the foundation of a new round of scientific and technological contest in the world. Colleges and universities assume the responsibility of training big data talents. As a typical “new engineering” major, the major of big data is still in the exploratory stage in the construction of curriculum system. Firstly, the difficulties in the construction of big data courses were analyzed, and then the big data course system built by Xiamen University was introduced, including introductory courses, advanced courses and training courses. Also the experience and methods of the course construction of the principles and applications of big data technology were introduced, which includes course orientation, training objectives, preparatory knowledge, knowledge partitioning between big data and cloud computing courses, course content and arrangement, teaching material, experimental environment construction, matching resources construction, online service platform, offline training and communication, and so on.

Key words

big data, course system, MOOC, public service platform

1 引言

大数据带来了信息技术的巨大变革,对社会生产和人们生活的各个领域都产生着深刻的影响^[1-4],所到之处,或是颠覆,或是提升,让人们深切感受到了大数据实实在在的威力。对于一个国家而言,能否紧紧抓住大数据发展机遇,快速形成核心技术和应用并参与新一轮的全球化竞争,将直接决定未来若干年世界范围内各国科技力量博弈的格局。大数据专业人才的培养是新一轮科技较量的基础,高等院校承担着大数据人才培养的重任,因此,各高等院校非常重视大数据课程的开设,大数据课程已经成为信息相关专业的重要核心课程。北京大学、厦门大学、中国人民大学等一批高校在国内率先开设大数据课程。2016年,北京大学、中南大学、对外经贸大学3所高校成为国内首批获得教育部批准设立“数据科学与大数据技术”专业的本科院校,此后,教育部又于2017年和2018年分别批准32所和248所本科院校设立数据科学与大数据技术专业。与此同时,根据教育部公布的“大数据技术与应用”专业备案和审批结果显示,截至目前,已经有累计208所职业院校获批“大数据技术与应用”专业。“数据科学与大数据技术”专业和“大数据技术与应用”专业一般被统称为“大数据专业”。随着大数据专业在国内众多高校中开设,大数据专业人才的培养迈入了全新的阶段。

大数据专业作为典型的“新工科”专业,在课程体系建设方面还处于摸索阶段^[5],没有太多可供借鉴的现成经验,需要一大批热爱教学的高校教师积极投身课程体系和教材的建设工作中,共同推动全国高校大数据教学工作不断向前发展。厦门大学

数据库实验室作为国内高校较早从事大数据教学资源建设的团队,从2013年开始,在大数据课程建设方面开展了很多有意义的尝试和探索,本文将分享笔者团队在这些方面的工作成果和经验做法。

2 大数据课程建设的难点

大数据专业课程涵盖范围较广,从学科角度而言,包括了数学(高等数学、线性代数、离散数学、数学建模等)、计算机(算法、数据结构、程序设计、数据库、操作系统、数据挖掘等)、统计(概率论与数理统计、多元统计分析等)等多学科知识。从数据分析流程角度而言,大数据专业课程包含了数据分析全流程的各种技术,包括数据采集、数据存储与管理、数据处理与分析、数据可视化等各个环节的技术。

本文探讨的大数据课程是指数据分析全流程涉及的大数据技术类课程。需要强调的是,在这些大数据技术类课程中,并非所有课程都是大数据时代新生的课程,比如,数据采集课程主要讲解网络爬虫技术,这些技术在大数据时代到来之前就已经存在很多年了,并非到了大数据时代才诞生。同理,数据可视化也是经历了多年发展的“老课程”,知识内容并没有因为大数据的出现而发生本质的变化。实际上,大数据技术之所以受到热捧,主要在于以Hadoop和Spark为代表的分布式框架解决了以较低的成本实现海量数据的存储和计算的问题。因此,本文探讨的大数据课程主要聚焦分布式存储和分布式处理这两个层面的大数据技术,这也是大数据课程区别于传统的计算机专业课程的一个重要方面。

经过上述聚焦以后,大数据课程建设的难点主要表现在以下几个方面。

(1) 缺少大数据课程体系

虽然本文把大数据课程聚焦到分布式存储和分布式处理这两个层面的大数据技术,但是,依然涵盖了大量新的知识内容。大数据技术体系非常庞杂,在分布式存储层面,涉及NoSQL数据库、分布式文件系统等内容;在分布式计算层面,涉及批处理、流计算和图计算等内容。如果要深入学习,每种技术都可以成为一门独立的课程。因此,如何系统归纳整理相关大数据技术,并且按照高校教学计划循序渐进地展开,是一个棘手的问题。此外,如何做好不同课程之间的知识布局,也具有较高的难度。因为在课程体系的建设上,需要做到课程与课程之间的“知识分割不重复”和“知识传承不断层”,让教师教课有纲可循,让学生学习脉络清晰。

(2) 缺少大数据课程教材

一门课程的顺利开设,离不开高质量的教材。作为新兴的专业,教材缺乏是一个正常的现象,也是必须面对和解决的问题。市场上已有的很多大数据书籍^[6-11]属于工具类数据,不合作为高校教材。

(3) 缺少大数据教学资源

教材只能满足开课的基本需求,如果要获得较好的教学效果,就必须建设大量配套的教学资源,比如PPT、MOOC、实验手册等,而在专业建设的初始时期,这些资源都是相对比较欠缺的。

(4) 缺少大数据课程师资

一门课程开设质量的高低取决于师资水平。若要教出合格的学生,就必须要求教师具备一定的大数据知识储备和编程实践能力。大数据课程需要教师具备大数据实验环境构建能力,掌握系统的大数据知识,具备开展编程实践的能力。而对于承担计算机类课程教学任务的很多教师而言,如果没有系统学习大数据知识,显然不具备这方面的能力。

(5) 大数据环境搭建复杂

大数据实验环境的搭建涉及Linux系统、虚拟机软件、JDK、Hadoop、Spark、HBase、Hive等多种开源软件的安装,安装过程涉及大量配置,需要具有一定的计算机操作功底,才能够顺利完成实验环境的搭建。而且有时候同样的操作方法,在某台计算机上可以成功,在另一台计算机上却会失败。这些给任课教师和学生带来了很大的挑战。

(6) 大数据学习门槛高

大数据知识体系的庞杂、大数据实验环境搭建的复杂性以及分布式编程的新知识都会成为大数据学习者很难跨越的“门槛”。如何有效降低大数据学习门槛,提高学习效率和效果,是开设大数据课程必须要解决的难题。

(7) 不同高校的差异化需求给教材建设带来了挑战

重点本科院校、普通本科院校和高职院校对人才培养的目标和方向都不尽相同,这种差异会在课程内容方面得到体现。因此,有些高校会更加侧重理论知识的讲解,而有些高校则会更注重学生技术应用能力的培养。这就要求课程教材和课程内容在知识点的布局和课时安排方面有不同的侧重点。如何根据现有的教材进行教学内容的个性化定制,是实际教学环节普遍遭遇的难题。

3 厦门大学建设的大数据课程体系

从2013年开始,经过为期5年的建设,厦门大学已经建成了涵盖入门课程、进阶课程和实训课程在内的大数据课程体系(如图1所示),已经可以提供包括教材、讲义PPT、习题、案例、视频、师资培训等全方位的大数据课程服务。

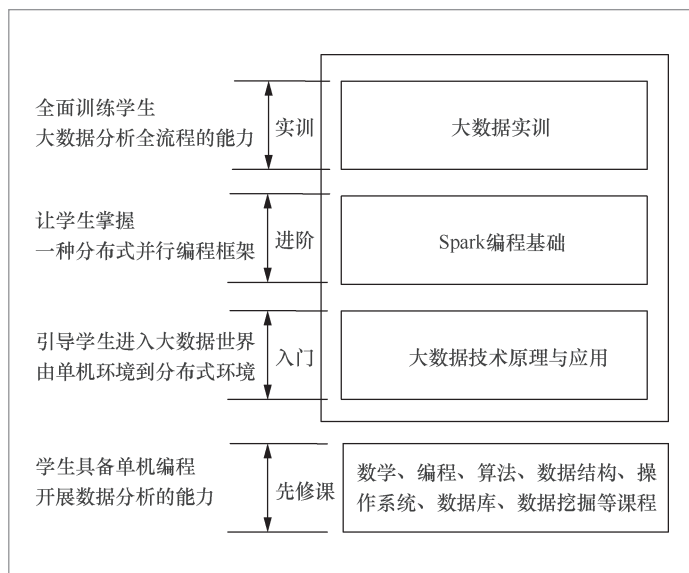


图1 厦门大学建设的大数据课程体系

在大数据专业课程体系中,数学、编程、算法、数据结构、操作系统、数据库、数据挖掘等课程属于历史已有课程,已经存在大量教材和配套资源,而分布式存储和计算框架则属于进入大数据时代以后的较新的课程,相关教学资源较为缺乏,因此,笔者单位选择大数据课程作为教学资源建设重点,投入大量人力、物力、财力,建设了全套的大数据课程资源。其中,“大数据技术原理与应用”属于大数据导论课程,旨在引导学生进入大数据世界,由单机环境进入分布式环境,面对和处理大数据问题^[12-13]。

“Spark编程基础”属于进阶级大数据课程,旨在帮助学生系统掌握当前流行的分布式并行编程框架Spark的编程方法^[14]。

“大数据实训”课程旨在满足高校开设综合实训课程的教学需求,案例涵盖数据预处理、存储、处理、分析、可视化等数据处理全流程涉及的各种典型操作和大数据软件,可以全面训练学生的大数据分析全流程的能力。

4 “大数据技术原理与应用”课程建设

“大数据技术原理与应用”课程是厦门大学数据库实验室建设的较为成熟的大数据课程,不仅在厦门大学受到学生的欢迎,而且已经推广到了国内很多高校。因此,本节内容将介绍这门课程的建设情况,包括课程定位、培养目标、预备知识、大数据与云计算课程之间的知识切割、课程内容与学时安排、课程教材、实验环境搭建、配套资源建设、在线服务平台、线下培训与交流等。

4.1 课程定位

该课程定位为入门级大数据技术课程,为学生搭建起通向“大数据知识空间”的桥梁和纽带。课程系统地梳理总结大数据相关技术,介绍大数据技术的基本原理和大数据主要应用,帮助学生形成对大数据知识体系及其应用领域的轮廓性认识,为学生在大数据领域“深耕细作”奠定基础、指明方向。在本课程学习的基础上,感兴趣的学生还可以通过其他工具书(如《Hadoop权威指南》等)继续深入学习和实践大数据相关技术。

4.2 培养目标

“大数据技术原理与应用”课程的预期目标如下。

- 能够建立对大数据知识体系的轮廓性认识,了解大数据发展历程、基本概念、主要影响、应用领域、关键技术、计算模式和产业发展,了解云计算、物联网的概念及其与大数据之间的紧密关系。

- 能够了解Hadoop的发展历史、重要特性、应用现状、Hadoop项目结构及其各个组件,并熟练掌握Hadoop平台的安装和使用方法。

- 能够了解分布式文件系统的基本概念、结构和设计需求,掌握Hadoop分布式文件系统(HDFS)的重要概念、体系结构、存储原理和读写过程,熟练掌握HDFS的使用方法。

- 能够了解分布式数据库HBase的访问接口、数据模型、实现原理和运行机制,并熟练掌握HBase的使用方法。

- 能够了解NoSQL数据库与传统的关系数据库的差异、NoSQL数据库的四大类型以及NoSQL数据库的三大基石;基本掌握Redis、MongoDB等NoSQL数据库的使用方法。

- 能够了解云数据库的概念、基本原理和代表性产品的使用方法。

- 能够熟练掌握分布式编程框架MapReduce的基本原理和编程方法。

- 能够了解大数据处理架构Spark的基本原理和编程方法。

- 能够了解大数据在互联网等领域的典型应用。

4.3 预备知识

“大数据技术原理与应用”属于大数据导论课程,引导学生进入大数据世界。因此,在学习该课程之前,需要学生已经学习过算法、数据结构、程序设计、数据库、操作系统等计算机类课程。

4.4 大数据课程与云计算课程之间的知识切割

目前,很多高校已经开设了大数据课程和云计算课程。由于大数据课程与云

计算课程有着千丝万缕的联系,因此,如果不能对这两门课程的知识进行合理切割,就会给实际教学环节造成很大的困扰,不仅会影响教师的教学,也会导致学生的重复学习。

云计算和大数据二者具有深厚的渊源。云计算的普及早于大数据,2006年亚马逊公司就推出了云计算产品Amazon AWS,而大数据概念被大众所熟知是在2010年左右。在大数据概念被广泛使用之前,以Hadoop为代表的大数据技术都被列为云计算技术,因此2010年之前出版的许多云计算教材^[15]通常有大量的篇幅介绍Hadoop等大数据技术,这类教材可以被称为复合型云计算教材。但是,2010年以后,以Hadoop为代表的分布式存储技术和分布式处理技术都被单独剥离出来,成为大数据技术。从这个角度来说,可以认为云计算为大数据提供了技术支撑。现在大家谈到云计算,主要是指为人类提供IT资源的一种全新商业模式及其背后涉及的相关技术。作为一种全新的商业模式,云计算彻底颠覆了人类获得IT资源的方式,云计算服务商可以通过网络、以服务的方式提供非常廉价的IT资源。

云计算和大数据存在如此紧密的关系,导致部分高校开设的云计算课程和大数据课程存在高度的知识重合,也就是说,两门课程都有大量的篇幅介绍Hadoop等大数据技术。为了解决这个尴尬的问题,就需要改变两门课程选择的教材,因为教材内容往往决定了课程的内容。云计算课程选择的教材应该尽量避免采用包含了大量大数据知识的复合型云计算教材,而应该采用单一型云计算教材。在单一型云计算教材中^[16],Hadoop等属于大数据课程的内容不会作为核心内容,只会用一个章节简要介绍Hadoop等大数据技术。

采用单一型云计算教材以后,云计算

和大数据两门课程的知识重点就有了明显的区分,不会重合,具体如下^[17]。

(1) 云计算课程的教学重点

云计算课程内容包括:云计算概念、云计算体系架构、数据中心、虚拟化技术(平台虚拟化、资源虚拟化、虚拟机的动态迁移、云操作系统)、SOA架构及开发技术、云数据中心设计与测试、云数据中心维护与管理、云安全架构、桌面云、PaaS应用开发平台、开源的云计算管理平台Openstack、Docker容器、大数据存储与管理(最多2个学时)。

(2) 大数据的教学重点

大数据课程内容包括:大数据的基本概念、Hadoop、HDFS、HBase、NoSQL数据库、云数据库、MapReduce、Spark、流计算、图计算、数据可视化以及大数据在互联网、生物医学和物流等各个领域的应用。

4.5 课程内容与学时安排

该课程在Hadoop、HDFS、HBase、MapReduce、Spark等重要章节安排了入门级的实践操作,让学生更好地学习和掌握大数据关键技术。具体学时安排见表1。

需要指出的是,在具体开展这门课程的教学时,各个高校可以根据自身的实际情况和特点,对课程内容进行适当调整。如果对理论层面内容比较注重,则可以适当地深入讲解理论内容,减少实验编程学时(比如可以把一些编程上机操作留作学生课后作业);如果更加注重实践操作层面,则可以减少理论内容的讲解,补充一些编程案例到课堂教学中。

4.6 课程教材

该课程将笔者编著的《大数据技术

原理与应用(第2版)》作为课程教材^[12]。该教材共分四大部分,包括大数据基础篇、大数据存储篇、大数据处理与分析篇和大数据应用篇。在大数据基础篇中,介绍了大数据的基本概念和应用领域,阐述了大数据、云计算和物联网的相互关系,并重点介绍了大数据处理架构Hadoop。由于Hadoop已经成为应用最为广泛的大数据技术,因此,该书的大数据相关技术主要围绕Hadoop展开,包括Hadoop MapReduce、HDFS和HBase。在大数据存储篇中,分别介绍了大数据存储相关技术的概念与原理,包括HDFS、HBase、NoSQL数据库和云数据库。在大数据处理与分析篇,介绍了大数据处理和分析的核心技术——MapReduce,然后,介绍了大数据时代两种新兴的数据分析技术——流计算和图计算,最后简单介绍了可视化技术。在大数据应用篇,介绍了大数据在互联网、生物医学和物流等各个领域的典型应用。

《大数据技术原理与应用(第2版)》属于入门级教材,它和进阶级大数据教材《Spark编程基础(Scala版)》^[14]以及大数据实训案例系列教材一起构成了可以满足高校教学需求的大数据教材体系。

4.7 实验环境搭建

在高校大数据教学过程中,实践是很重要的环节。因此,很多高校在开设大数据课程时,不仅要选择好的教材,同时也要选择好的大数据实验平台。总体而言,当前国内有以下几种大数据实验平台方案。

(1) 建设统一的大数据实验机房

条件较好的高校可以建设统一的大数据机房,为师生提供大数据上机环境。机房建设的方案也各有不同,部分高校采用“云桌面”解决方案,还有一部分高校采用

“小集群”解决方案。在“云桌面”解决方案中,需要在机房部署一个中央服务器,通过虚拟化技术可以构建出多个虚拟化主机,学生在终端机(相当于云桌面)上通过网络连接到中央服务器,中央服务器会为其分配一个虚拟机,学生在虚拟机上完成各种实验操作。在“小集群”解决方案中,可以为每个学生分配一台物理计算机,然后,由3~5个学生组成一个小组,让每个小组内部的多台计算机组建一个物理的集群,学生在集群中开展各种实验。

(2) 单机构建实验环境

单机构建实验环境主要是为了满足教师和学生课后学习实践大数据时的需求。因为机房统一上机的时间相对有限,大数据的学习还需要学生课后在自己的计算机上开展上机实践。单机构建大数据实验环境可以采用“虚拟机”和“双系统”两种方式。在“虚拟机”方式中,需要启动进入Windows系统,在系统中安装虚拟机软件,比如VMWare或者VirtualBox,然后,在虚拟机软件上安装Linux操作系统,再在Linux操作系统上安装Hadoop等大数据相关软件。若采用“双系统”的方式,在开机启动时,可以直接选择进入Linux系统,开展大数据实验。

(3) 实验室多机构建分布式环境

如果实验室有多台计算机可用,则可以利用这些机器构建一个集群,实现真正的分布式环境,从而完成一些科研数据的处理分析。通过多机的分布式计算,可以大大提高科研数据分析的效率。

4.8 配套资源建设

课程配套资源的建设是打造一门优质课程的重要环节。为了方便高校教师开设大数据课程,提高学生的课程学习效果,笔者团队制作了丰富的课程配套资源,包

表1 “大数据技术原理与应用”课程内容与学时安排

序号	主要内容	学时安排/课时
1	大数据的基本概念和应用领域,阐述大数据、云计算和物联网的相互关系	2
2	Hadoop	4
3	HDFS的基本原理和使用方法	4
4	HBase的基本原理和使用方法	4
5	NoSQL数据库的概念和基本原理	4
6	云数据库的概念和基本原理	2
7	MapReduce的原理和使用方法	4
8	Hadoop架构再探讨	2
9	Spark的原理与基础编程	4
10	大数据在互联网领域的典型应用:推荐系统	2

括讲义PPT、习题、实验手册、技术博客和课程视频等,并通过MOOC平台发布。

4.9 在线服务平台

丰富的课程配套资源必须通过方便快捷的方式传播给大数据学习者,只有这样才能充分发挥配套学习资源的价值。为此,笔者团队专门为课程配套建设了高校大数据课程公共服务平台,旨在促进国内高校大数据课程体系建设,提高大数据课程教学水平,降低大数据课程学习门槛,提升学生课程学习效果。该平台为全国高校教师和学生提供大数据教学资源一站式“免费”在线服务,包括课程教材、讲义PPT、课程习题、实验指南、学习指南、备课指南、授课视频和技术资料等。

4.10 线下培训和交流

具备合格的师资力量是各个高校开设大数据课程的前提条件。由于大数据知识体系较新,很多高校不具备大数据师

资储备。因此,笔者团队通过“公益”师资培训、大型研讨会、大数据公开课全国高校巡讲等活动,向全国高校传播大数据教学、普及大数据知识体系的同时,使不同高校教师互相分享教学理念和教学方法,共同推进中国高校大数据专业教学的发展。

5 结束语

大数据时代已经全面开启,高校大数据课程体系建设正在加速推进。目前,一些尚未开设大数据专业的高校已经在计算机和软件工程等专业中开设了大数据课程,以达到拓展学生知识结构、增强人才核心竞争力的目的。而对于已经开设了大数据专业的高校而言,更是把大数据课程作为核心课程进行重点建设,一方面通过参加师资培训锻炼教师队伍,另一方面建设高水平的大数据实训实验室,满足大数据课程对实验环境的要求。各所高校都在为如何做好大数据专业做出有益的尝试和探索。笔者相信,未来的3~5年内,在全国高校大数据教育工作者和企业界的共同努力下,高校大数据课程体系会逐渐走向完善,高校大数据教学水平会不断迈上新的台阶。

参考文献:

- [1] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶. 大数据时代: 生活、工作与思维的大变革[M]. 盛杨燕, 周涛, 译. 杭州: 浙江人民出版社, 2013.
MAYER-SCHONBERGER V, CUKIER K. Big data: a revolution that will transform how we live, work, and think[M]. Translated by SHENG Y Y, ZHOU T. Hangzhou: Zhejiang People's Publishing House, 2013.
- [2] 周春梅. 大数据在智能交通中的应用与发展[J]. 中国安防, 2014(6): 33-36.
ZHOU C M. Application and development of big data in intelligent transportation[J]. China Security & Protection, 2014(6): 33-36.
- [3] 秦萧, 甄峰. 大数据时代智慧城市空间规划方法探讨[J]. 现代城市研究, 2014(10): 18-24.
QIN X, ZHEN F. Discussion on the method of intelligent urban spatial planning in big data era[J]. Modern Urban Research, 2014(10): 18-24.
- [4] 车志宇, 段云峰. 电信客户离网分析方法[J]. 电信技术, 2004(10): 17-18.
CHE Z Y, DUAN Y F. Telecom customer off-line analysis method [J]. Telecommunications Technology, 2004(10): 17-18.
- [5] 朝乐门, 邢春晓, 王雨晴. 数据科学与大数据技术专业特色课程研究[J]. 计算机科学, 2018, 45(3): 3-10.
CHAO L M, XING C X, WANG Y Q. Unique curriculums for data science and big data technology[J]. Computer Science, 2018, 45(3): 3-10.
- [6] 陆嘉恒. Hadoop实战[M]. 北京: 机械工业出版社, 2011.
LU J H. Hadoop in action[M]. Beijing: China Machine Press, 2011.
- [7] 于俊, 向海, 代其锋, 等. Spark核心技术与高级应用[M]. 北京: 机械工业出版社, 2016.
YU J, XIANG H, DAI Q F, et al. Spark core technology and advanced applications[M]. Beijing: China Machine Press, 2016.
- [8] WHITE T. Hadoop权威指南[M]. 华东师范大学数据科学与工程学院, 译. 北京: 清华大学出版社, 2010.
WHITE T. Hadoop: the definitive guide[M]. Translated by School of Data Science and Engineering, East China Normal University. Beijing: Tsinghua University Press, 2010.
- [9] DIMIDUK N, KHURANA A. HBase实战中文版[M]. 谢磊, 译. 北京: 人民邮电出版社, 2013.
DIMIDUK N, KHURANA A. HBase in

- action[M]. Translated by XIE L. Beijing: Posts and Telecom Press, 2013.
- [10] 黄宜华. 深入理解大数据——大数据处理与编程实践[M]. 北京: 机械工业出版社, 2014.
HUANG Y H. Understanding big data: big data processing and programming[M]. Beijing: China Machine Press, 2014.
- [11] 蔡斌, 陈湘萍. Hadoop技术内幕——深入解析Hadoop Common和HDFS架构设计与实现原理[M]. 北京: 机械工业出版社, 2013.
CAI B, CHEN X P. Hadoop technology insider: deep analysis of hadoop common and the architecture design and implementation principles of HDFS[M]. Beijing: China Machine Press, 2013.
- [12] 林子雨. 大数据技术原理与应用(第2版)[M]. 北京: 人民邮电出版社, 2017.
LIN Z Y. Principles and applications of big data technology(2nd)[M]. Beijing: Posts and Telecom Press, 2017.
- [13] 林子雨. 大数据基础编程、实验和案例教程[M]. 北京: 清华大学出版社, 2017.
LIN Z Y. Tutorial on big data basic programming, experiments and cases [M]. Beijing: Tsinghua University Press, 2017.
- [14] 林子雨, 赖永炫, 陶继平. Spark编程基础(Scala版)[M]. 北京: 人民邮电出版社, 2018.
LIN Z Y, LAI Y X, TAO J P. Spark programming (Scala edition)[M]. Beijing: Posts and Telecom Press, 2018.
- [15] 刘鹏. 云计算(第三版)[M]. 北京: 电子工业出版社, 2015.
LIU P. Cloud computing (third edition)[M]. Beijing: Publishing House of Electronics Industry, 2015.
- [16] 顾炯炯. 云计算架构技术与实践(第2版)[M]. 北京: 清华大学出版社, 2016.
GU J J. Cloud computing architecture technologies and practice (second edition)[M]. Beijing: Tsinghua University Press, 2016.
- [17] 姚乐, 朱启明. 赋能大数据教育——全国高校大数据教育教学经验谈[M]. 北京: 电子工业出版社, 2018.
YAO L, ZHU Q M. Enabling big data education—talking about the teaching experience of big data education in national colleges and universities[M]. Beijing: Publishing House of Electronics Industry, 2018.

作者简介



林子雨(1978-), 男, 厦门大学信息科学与技术学院助理教授, 主要研究方向为数据库、数据挖掘和大数据等。

收稿日期: 2018-09-05