

郑健

手机: 18810300681 邮箱: zhengjian8972@126.com
学校: 北京邮电大学 地址: 北京市海淀区西土城路 10 号 (100876)



教育背景

- | | | | |
|---|---------------|----------------------|--------------|
| 2011.9—2014.3 | 北京邮电大学 | 计算机学院通信软件工程中心 | 工学硕士 |
| 专业方向: 下一代电信运营支撑系统与智能信息处理; 主修课程: 数据仓库与数据挖掘、现代软件工程、复杂网络; | | | |
| 2007.9—2011.7 | 厦门大学 | 软件学院 | 工学学士 |
| 专业: 软件工程; 主修课程: C/C++/Java 程序设计语言、数据结构与算法、软件工程、计算机网络、操作系统等; | | | |
| 2007.9—2011.7 | 厦门大学 | 经济学院 | 经济学学士 |
| 专业: 经济学双学位; 主修课程: 微观经济学、宏观经济学、统计学、会计学原理、国际经济学、金融学等; | | | |

项目经历

- 2011.10—2013.6 “大云”并行数据挖掘产品研发和应用项目 (数据挖掘模块组长)**
- 该项目是中国移动研究院“BigCloud-大云”平台的数据挖掘子系统, 基于 Hadoop 与 BSP 的海量数据并行 ETL 处理、分析与挖掘系统, 适用于经营决策、用户行为分析、精准营销、网络优化、移动互联网等领域的智能数据分析与挖掘应用。
- 主要负责数据挖掘算法的性能优化、数据 ETL 组件的开发;
 - 提出一套 Hadoop 参数优化方案, 并实现了一个 Hadoop 集群参数管理工具;
 - 实现了并行协同过滤推荐算法并基于该算法搭建了电影协同推荐系统;
 - 负责组织人员对“大云”平台在中国移动省分公司落地时使用的候选 Linux 服务器进行测试;
 - 组织人员使用 Hive 实现中国移动省分公司的电信日志处理流程在云计算环境下的并行化;
 - 负责整理项目周报和软件工程文档, 基于 SSH 实现了实验室内部使用的周报汇总系统。
- 2010.06—2010.9 保险理财规划系统 (项目成员)**
- 该项目根据用户输入的个人信息为用户计算出合理的保险规划列表, 并在用户确定保险组合后引导用户在网上完成投保操作。
- 主要负责调研保险产品的保费计算和投保过程, 为系统提供的可行性分析和需求分析;
 - 完成系统的数据库设计与实施。

奖励/活动

- 2013 北京邮电大学优秀研究生干部;
- 2011-2013 北京邮电大学研究生三等奖学金;
- 2011-2013 北京邮电大学计算机学院 2011 级研究生 4 班心理委员: 学习并在班级内部宣传心理健康知识;
- 2011-2012 北京邮电大学研究生会实践部干事: 负责北京邮电大学创业计划大赛宣传工作;
- 2011-2012 北京邮电大学计算机系研究生会宣传部干事: 负责计算机研究生会各项文体活动宣传工作;
- 2007-2008 厦门大学演讲协会宣传部部长: 负责协会日常活动和演讲比赛宣传工作;
- 2007-2008 厦门大学漳州校区艺术团合唱队成员: 担任男低音;
- 2009 厦门国际马拉松赛跑完全程 42.195km 路程, 耗时 6 小时零 6 秒;
- 2009 厦门大学“我爱我的祖国”主题暑期社会实践平潭县分队领队。

个人技能

- 熟练: Java、Hadoop 编程及参数配置、Hive、协同过滤推荐算法、常用数据结构和算法;
- 熟悉: C++、Linux 常用命令、MySQL 数据库设计与使用、SVN/Git 版本控制、设计模式、常用数据挖掘算法;
- 掌握: Python、Struts2、Hibernate、XML、HTML4 + CSS2 + Javascript、ActionScript + MXML;
- 英语 CET-6 级, 具有良好的听说读写能力;
- Github 代码库地址: <https://github.com/zhengjian8972>

工作实践

2013.6—2013.9 兰亭集势（北京）科技有限公司（分布式系统与计算研发实习生）

- 根据业务需求设计基于 Hadoop 的日志分析、数据挖掘、用户行为分析算法；
- 改进和维护公司内部的日志分析系统，包括维护原有功能、开发新功能、搭建新平台以快速响应需求。

2011.2—2011.5 厦门热推网络有限公司（手机客户端开发实习生）

- 参与“厦门小鱼”论坛的 Android 客户端的设计与开发；
- 调研 phonegap、sencha 等跨平台开发框架与原生 Android 程序的优劣。

发表论文

- BC-PDM: data mining, social network analysis and text mining system based on cloud computing. Le Yu, **Jian Zheng**, Wei Chong Shen, Bin Wu, Bai Wang, Long Qian, and Bo Ren Zhang. **KDD, page 1496-1499. ACM, (2012)**

本文介绍了基于 Hadoop 的 SaaS BI 平台 BC-PDM，并采用真实的电信数据集演示了业务推荐、客户群体划分和用户喜好分类功能。链接：<http://dl.acm.org/citation.cfm?id=2339764>。

- 基于资源与作业特征的 Hadoop 参数优化方法。**郑健**，徐六通。北京邮电大学第八届学术论坛（2013），优秀论文。

本文首先研究了 Hadoop 相关参数对资源消耗的影响，提出了一种基于集群资源配置和作业特征的 Hadoop 参数优化的方法，并实现了基于该方法的 Hadoop 参数调优工具。本文提出的方法在电信数据处理场景下效果显著。

自我评价

- 学习能力强，对知识充满渴求，善于思考，乐于钻研；
- 待人真诚热情，崇尚团队精神和奉献精神，工作积极主动，做事认真负责；
- 兴趣爱好：游泳、羽毛球、阅读、尝试新技术。