范忠瑞

中国科学院 计算技术研究所 应聘岗位:研发工程师 JAVA

E-Mail: fanzhongruis@163.com 电话: 13126666174

教育背景

2014.9-2017.7 中国科学院计算技术研究所 计算机应用技术 工学硕士(top 10%) 2010.9-2014.7 郑州大学 计算机科学与技术 工学学士(top 5%)

个人技能

- 熟悉 Java 编程, JVM 原理, Java 集合框架, 了解 C/C++, Python, Cloiure
- 熟悉数据结构和算法,了解常用设计模式,如策略模式,抽象工厂模式等
- 熟悉分布式数据库基本理论,熟悉 Redis, MySQL 等常用数据库
- 熟悉分布式计算框架 storm, Hadoop 等
- 伯乐在线专栏翻译 fzr, 目前在今日头条广告技术部门实习
- 阿里天池中间件大赛第一赛季第八名(单人组队)

■ 个人博客: http://blog.csdn.net/ymrfzr

项目经历

2016.6-2016.7 双十一实时计算模拟系统 天池阿里中间件性能挑战赛

核心研发人员

- ◆ *项目介绍:*模拟实时统计阿里双十一淘宝、天猫平台的交易数据。
- ◆ 负责工作:
 - **数据源与持久层构建:** 利用 JStorm 从 RocketMQ 拉取数据,按照要求将实时计算结果写入 Tair;
 - **计算平台每分钟的交易额:** 实时统计每一分钟内淘宝、天猫平台各自的交易值并将统计结果实时入库;
 - **计算整分交易金额比值:** 实时统计每整分时刻无线和 PC 端交易金额总比值并将统计结果实时入库。
- ◆ 工作成果:实测线上三千万数据,实时统计准确率 100%,实时计算结果平均耗时 115s, TPS11 万,排名第八。

2016.6-至今

京东数据云数千工坊

北京京东世纪贸易有限公司

软件开发工程师

- ◆ 项目介绍:通过拖拽、所见即所得的数据操作方式,对海量数据进行可视化分析的在线数据分析产品。
- ◆ 负责工作:
 - **数据抽取转换:** 完成用户中心数据由京东到数据计算平台 DCS 的迁移所需的 SDK 接口的开发;
 - ▼工作流构建:辅助实现向原云海任务调度系统中增加工作流构建功能的改造过程;
 - 即席查询:参与数千工坊 OLAP 即席查询功能的调研实现及测试上线;
- ◆ *工作成果:*数据转换程序达到 10M 每秒的速度,成功辅助改造工作流构建功能。

2015.8-2016.4

电子侦察信息处理系统

中国科学院计算技术研究所

核心研发人员

- ◆ 项目介绍:针对大量实时卫星数据和 AIS 数据,识别卫星目标,提供电子目标具体信息,挖掘航道和码头信息。
- ◆ 负责工作:
 - **系统架构设计:** 采用 redis 缓存热度数据,使用消息队列 ActiveMQ 实现多客户端的信息同步;
 - **实时位置数据接入及预处理**:解析数据包,将位置数据清洗入库,完成轨迹数据拼接及在线压缩船舶轨迹:
 - **目标预识别算法的设计与实现:** 选择目标特征,学习特征参数并进行电子目标的预识别处理;
 - 码头及航道挖掘:基于 DBSCAN 算法,面向不同密度分布类型的船只进行自动学习,实现码头挖掘和航道挖掘;
 - ◆ 工作成果:实际使用场景下,出情时间由 8min 降为 2min,可承受数据量提升 5 倍以上,目标识别准确率提高 5%,码头挖掘准确率达到 93%, 航道挖掘准确率为 90%。

2014.5-2014.9

搜索引擎

中国科学院计算技术研究所

核心研发人员

- ◆ *项目介绍:* 获取指定数据源,爬取数据并实现数据信息搜索,为科研项目提供实验平台和数据信息来源。
- ◆ 负责工作:
 - **系统关键技术的实现:** 为文档建立倒排索引,计算文档相关性,检索结果排序,查询自动补齐,搜索结果评价;
 - **网页聚类算法的设计与实现:** 使用 KNN 实现网页聚类, 相似网页的去重基于 sim-hash 算法:
 - 大规模短文本快速分类算法实现:设计基于网页标题的文本分类方法,使用 word2vec 及 SVM 算法实现分类;
 - 生成网页动态摘要: 针以用户查询词为中心,依据用户输入查询词提取摘要,旨在展现原文与查询的相关性。
- ◆ *工作成果:* 在 200 万网页数据,实现检索时间不超过 0.7s,短文本分类方法准确率比学术界 baseline 提高 3%。

获奖情况

- 2011.10 郑州大学优秀学生奖学金(**Top 10%**)
- 2012.05 河南省程序设计大赛铜奖

