

范忠瑞

中国科学院 计算技术研究所

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 15%)
2010.9-2014.7	郑州大学	计算机科学与技术	工学学士(top 10%)

项目经历

2015.8-至今	船舶水运信息处理系统	中国科学院计算技术研究所	核心研发人员
◆	项目介绍: 综合海量卫星数据和船舶 AIS 动静态数据, 进行数据分析和实体挖掘, 实现船舶异常行为建模并基于船舶轨迹数据挖掘船只异常行为。		
◆	负责工作:		
	● 实时位置数据接入及在线预处理: 解析 27 种格式数据包, 过滤噪音数据, 将位置数据清洗入库, 并进行轨迹数据拼接及在线压缩船舶轨迹;		
	● 海上交通密度计算与分析: 基于船舶轨迹数据, 将全球划分为细粒度网格, 设计实现能对多维交通密度特征进行统计分析的算法;		
	● 码头及航道挖掘算法的设计与实现: 利用海量船舶位置及轨迹数据, 使用 DBSCAN 算法, 面向不同类型不同密度分布的船只进行自动学习优化参数, 实现码头挖掘和航道挖掘;		
	● 构建船舶异常行为模型: 针对船舶异常行为侦察监视的需求, 实现船舶异常行为模型的构建及异常评价算法的设计实现。		
◆	工作成果: 利用两年内中国滚装船的实际 AIS 数据和国际滚装船实际数据, 码头挖掘准确率达到 93%, 航道挖掘准确率为 90%。支持发现船舶异常行为并告警, 异常船舶检测正确率 $\geq 70\%$; 完成异常船舶检测时间 $\leq 1\text{min}$ 。		

2014.10-2015.7	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: 针对大量且实时性要求较高的卫星数据和 AIS 数据, 识别卫星目标, 为多个情报部门提供电子目标信息, 以及有效的战场态势分析。		
◆	负责工作:		
	● 系统架构设计: 采用基于内存数据库的热度数据缓存机制, 使用消息队列实现多客户端的信息同步;		
	● 数据预处理: 解析接收的电子数据及 AIS 数据压缩包, 数据清洗入库;		
	● 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;		
	● 数据存储与缓存优化方案的设计与实现: 针对系统数据量增长较快的特点, 设计实现数据库分表与常用数据提取方案。		
◆	工作成果: 在实际使用场景下, 系统出情时间由前系统的 8min 降为 2min, 系统可承受数据量提升 5 倍以上, 系统目标识别准确率提高 5%。		

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, 爬取数据并实现搜索引擎, 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: 从数据源爬取数据, 建立倒排索引, 计算文档相关性, 检索结果的排序, 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;		
	● 大规模短文本分类算法实现: 为实现对大规模文本快速有效分类, 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: 针对用户查询词为中心, 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%。		

个人技能

- 熟悉 Java, C/C++, Python, SQL 等常用语言
- 熟悉数据挖掘理论和方法, 能够使用基本聚类、回归算法解决实际问题
- 熟悉基本数据结构和算法
- 熟悉常用数据库, 了解 Hadoop, Storm 分布式计算框架

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