**范忠瑞**

**中国科学院 计算技术研究所 专项技术研究中心**

E-Mail：[fanzhongruis@163.com](mailto:fanzhongrui@ict.ac.cn) 电话：13126666174

**求职意向**

数据挖掘研发工作；软件开发工作

**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

2010.9-2014.7 郑州大学 计算机科学与技术 工学学士

**项目经历**

**2015.10-至今 海运大数据分析平台 中国科学院计算技术研究所 核心研发人员**

* + *项目介绍：*建立供项目组内部使用，并用于处理分析海运数据的一整套平台系统。
  + *负责工作（独立完成）：*维护数据库集群，构建流式数据分析模块，在此基础上实现实时计算功能；
  + *主要方法：*
    - 基于**MySQL集群、Storm框架**，实现对应数据源的流式处理逻辑，完成流式分析功能；
    - 利用**内容缓冲、线程池、内存数据库**，提高数据处理效率；
  + *工作成果：*实时场景下集群反应时间为毫秒级。

**2015.4-2015.10 船舶水运信息平台 中国科学院计算技术研究所 核心研发人员**

* + *项目介绍：*综合多方数据，分析获取海上交通信息，预测海运与经济走势相关度。
  + *负责工作（独立完成）：*
    - 找出海上航运主要路线，实现航道发现；
    - 挖掘船舶和航道联系，用以识别船队及分析船舶异常。
  + *主要方法：*
    - 采用**决策树算法**从整合清理过的航运数据库中提取轨迹特征，从船舶数据库中抽取船舶特征；
    - 使用**层次聚类算法、DBSCAN算法**，对轨迹数据进行过滤、聚类；
    - 分析子航道属性，根据船舶特征和轨迹聚类结果，建立船队和子航道的对应关系。
  + *工作成果：*子航道识别准确率：96%，船舶航道对应准确率：90%。
    - 1. **电子数据处理系统 中国科学院计算技术研究所 核心开发人员**
  + *项目介绍：*接收卫星电子数据，通过数据分析手段获取海事信息。
  + *负责工作（独立完成）：*
    - 多数据源的数据规整化，实现各项服务接口，完成数据分析功能；
    - 挖掘舰艇船只之间的相似关系，为识别敏感船只目标提供核心依据。
  + *主要方法：*
    - 采用**决策树算法**对船舶数据进行船舶特征，主要依据候选船舶特征与现有敏感船舶特征参数库；
    - 利用船舶特征，采用**朴素贝叶斯算法**进行船只分类；
    - 提取观测船只与特征库船只的比较关系，完成观测船只的匹配结果。
  + *工作成果：*船只分类准确率：91%，召回率：90%；船舶相似匹配准确率：90%。

**获奖情况**

* 2011.10 优秀学生奖学金(**Top 10%**)

**个人技能**

■ 熟悉Java，C/C++， Python ■ 熟悉基本数据挖掘理论和方法 ■ 熟悉基本数据结构和算法

■ 熟练使用Hadoop，Storm ■ 英语：通过CET-6

**Zhongrui Fan**

**Institute of Computing Technology, Chinese Academy of Sciences**

Mobile: (+86)13126666174 E-Mail: [fanzhongruis@163.com](mailto:fanzhongrui@ict.ac.cn)

**Objective**

Data Mining Engineer, Software Engineer

**Education**

2014.9 - 2017.7 Master Institute of Computing Technology, Chinese Academy of Sciences Compute Application Technology

2010.9 - 2014.7 Bachelor Zhengzhou University Computer Science and Technology

**Project Experience**

**2015. 10**- **now Marine Big Data Analysis platform Institute of Computing Technology Designer & Developer**

* *Project Description:* Establish a platform to handle marine data for our project team.
* *Personal Responsibilities:* Maintain database cluster, construct data stream analysis module, and realize real-time calculation based on it.
* *Major methods:*
  + Based on **MySQL cluster and Storm framework**, finish the stream processing corresponding to data source;
  + Use **buffer, thread pool, cache database** to improve the efficiency of data processing;
* *Results:* The reaction time under the cluster real-time scene is in milliseconds.

**2015.4 - 10 Ship water transport information platform Institute of Computing Technology Designer & Developer**

* *Project Description:* Retrieve multi data sources, analysis and get maritime traffic information, forecast the correlation between maritime traffic and economic trend.
* *Personal Responsibilities:*
  + Find out main lines of shipping in the sea, and realize the discovery of the channel;
  + Mine links between ships and channels to identify the fleet and analysis of abnormal ships.
* *Major methods:*
  + By **Decision Tree algorithm** to extract ship feature of trajectory from the integrated and clean shipping database, and the vessel features from vessel data set;
  + Use **hierarchical clustering algorithm** and **DBSCAN algorithm** to filter and cluster the trajectory data;
  + Establish the corresponding relationship between fleet and sub channel, based on the clustering results of ship characteristics and trajectory.
* Results: Sub segment recognition accuracy: 96%, Accuracy rate of matching between ship and channel: 90%.

**2014.8 – 15.4 Electronic Data Target Recognition System Institute of Computing Technology Designer & Developer**

* *Project Description:* Receiving satellite electronic data, obtain maritime information through data analysis.
* *Personal Responsibilities:*
  + Data normalization of multi-source, and realization of service interface to complete data analysis;
  + Analysis of maritime intelligence, and mine the similar relationships between vessels to provide the core basis for the identification of sensitive vessels.
* *Major methods:*
  + Adopt **Decision Tree Classification Model** to extract vessel features, based on the features of the candidate ships and the characteristic parameter of existing ships;
  + Using **Naïve Bayesian algorithm** to classify the vessels, according to the characteristics of ships;
  + Extract the comparative relationship between the observed ships and ships in storage, and finish the matching result of observed vessels.
* *Results:* Vessel classification accuracy rate: 91%, Recall rate: 90%, Accuracy rate of ship similarity matching: 90%.

**Awards**

* Excellent Scholarship for Encouragement (**top 10%**).

**Personal skills**

■ Familiar with Java, C/C++, Python ■ Familiar with basic machine learning theory and practice

■ Familiar with data structure and algorithm ■ Familiar with Hadoop and Storm practice ■ English: CET-6