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
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　　项目负责人

- *项目介绍：*针对海量船舶数据和卫星电子数据，建立从数据存储到数据分析，从离线分析到实时处理的一整套数据分析平台系统。
- *负责工作：*
  - **作为项目负责人：**承担平台系统架构的方案设计，项目管理和沟通，安排资源分配，核心技术研究；
  - **实时处理方案设计与实现：**结合 Storm 框架与 Esper 系统，实现船舶数据的流式处理，完成流式分析功能，在此基础上利用内容缓冲和线程池减小数据处理延时；
  - **设计实现船舶数据的在线分析：**针对船舶实时位置数据，使用卡尔曼滤波实现在线过滤噪音数据，并实现在线压缩船舶轨迹与轨迹分段。
- *工作成果：*实现在线船舶目标跟踪，跟踪反馈延时<=10s。

2015.4-2015.10　　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：*
  - **位置数据接入及预处理：**解析 27 种格式数据包，数据清洗入库，并根据 AIS 动态位置数据特征拼接船舶轨迹数据；
  - **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，设计实现了能对网格信息发送数目、信息发送间隔、船舶经过次数、船舶停留时间等多维交通密度特征进行统计的算法；
  - **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，使用层次聚类算法及 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>=70%；完成异常船舶检测时间<=1min。

2014.7-2015.4　　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：*海军某部门出情系统，综合处理大量实时性要求高的卫星数据和船舶 AIS 数据，为多个情报部门提供军事情报，提供有效的战场态势分析。
- *负责工作：*
  - **系统架构设计：**采用传统数据库与内存数据库相结合的二级缓存机制，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%。

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

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，能够使用数据挖掘算法解决实际问题
- 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

Zhongrui Fan

Institute of Computing Technology, Chinese Academy of Sciences
Mobile: (+86)13126666174 E-Mail: fanzhongruis@163.com

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 **Storm framework**, finish the stream processing corresponding to data source;
  - Use **buffer, thread pool, memory 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 select 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: 90%, 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-sources, 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 select 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%**).
- Bronze Medal at Henan Province Programming Contest.

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