

## 魏翔

北北京市海淀区学院路 37 号北京航空航天大学 (地址)  
100083 (邮编)

18525425662 (电话)

weixiang@buaa.edu.cn (邮件)  
github.com/spaces-x (GitHub)



### 教育经历

- **北京航空航天大学** 计算机学院 · 计算机应用与技术 2018.09 – 至今 硕士在读  
– 矩阵理论 (88), 算法设计与分析 (91), 学术英语 (87), 信息系统集成技术 (93) 等
- **大连理工大学** 软件学院 · 网络工程 2014.09 – 2018.06 学士学位  
– 计算机网络, UNIX 网络编程, 数据结构与算法, C 语言入门, 操作系统, 离散数学等

### 技能专长

- **英语:** 六级(565)
- **语言:** Java, Python
- **工具:** Git, VS code, PostMan, IDEA 等常用编程辅助工具
- **操作系统:** 熟悉 Linux 下的基本命令, 具备在 Linux 系统下搭建小型分布式集群 (3 节点) 的经验

### 项目经历

- **高速交通大数据分析** 路网中心 2018.11 – 2019.01
  - 搭建 hadoop 分布式集群存储环境, 对路网中心上传的月数据量达 **500G** 全国收费站数据进行存储管理
  - 在交易数据记录中按属性进行聚类, 检测出离群点, 并将这部分离群点数据的信息完整性进行评估, 过滤出存在逃费的异常行为记录
  - 利用 spark 工具, 对交易数据进行分析, 生成不同省份的出行基本特征、区域运输特征以及通行费用基本特征的画像。根据分析结果, 给出合理的高速收费标准, 并针对不同的费率模拟推演计算盈亏数量
  - 依据路链拓扑信息, 采用 Dijkstra 计算不同高速收费站间最短路链集合、距离及行驶时间, 路段每日累计承担重量
- **联通信令数据挖掘** 联通/安科院 2019.03 – 2019.10
  - 基于联通的手机信令切片数据 (80G/天), 预处理对基站漂移和基站乒乓切换产生异常数据进行剔除
  - 针对每个用户 id, 对个人每日的停留点进行提取, 进而得到个人用户的出行链
  - 对出行链中的停留点属性判断 (职住), 并对特定区域 (大兴化工厂) 工作人员的夜间活动人数跟踪, 并与工业台账对比
  - 对固定出行模式的人员进行识别 (如: 外卖快递人员、出租车或网约车司机)
  - 基于联通实时数据 (5T/天) (Kafka 消费) 结合公交智能卡数据, 对密集人群区域监控 (南锣鼓巷、东直门枢纽)
- **广告数据平台告警服务** 暑期实习/字节跳动 2019.06 – 2019.08
  - 广告数据平台后台 ABTest 模块对实验组、对照组数据指标进行查询
  - 采用孤立森林的方法, 进行异常点检测并通过 Lark 消息对实验人进行报警
  - 分析报警后台任务时间较长的原因, 并将后台数据库由 Mysql 优化为 Redis 内存数据库
  - 发送消息后响应用户 (实验人) 的反馈, 将误判的实验 id 加入白名单
  - 将原来每 15 分钟执行的任务, 缩短到分钟级

### 自我评价

- 熟悉 Java 编程语言, 了解 JVM 的内存分区等, 了解并使用过 Spark Hadoop 等大数据分析工具
- 乐观积极, 热爱学习, 有较强的文档 (文献) 阅读能力, 做事认真负责, 有强烈团队意识和责任感