

# 陶友贤

安徽马鞍山 | 中共党员 | 24岁(1995.01) | 硕士三年级在读 | 资格认证: 软件设计师

☎: (+86)188-1137-1109

✉: taoyouxian@ruc.edu.cn



## 教育背景

- |          |                     |         |                 |
|----------|---------------------|---------|-----------------|
| - 中国人民大学 | 计算机应用技术 (3.83/4.00) | 硕士 (保送) | 2017.09-至今      |
| - 河海大学   | 计算机科学与技术 (2/85)     | 学士      | 2013.09-2017.07 |

## 学习概貌

实验室: 数据库与智能信息检索实验室 (DBIIR Lab)

研究方向: 大数据存储与分析系统, 自适应列存储布局优化

个人技能: Java | C++ | Python, 有 Hadoop/Presto/Hive 实践、区块链应用开发经验

语言水平: 英语六级 (522), 通过中级商务英语

## 科研/项目经历

- **Rainbow: 自适应存储布局优化系统** (<http://github.com/dbiir/rainbow>) 2017.07 - 2017.10  
描述: 针对不同查询负载自适应调整列存储的物理布局, 优化数据读取的跳读代价, 提高查询的 I/O 性能  
负责: (1) 动态收集查询负载, 根据 LRU 的缓存置换策略触发优化; 自动优化行组的大小  
(2) 提供交互式的 Web UI、CLI, 宽表分析 Benchmark
- **Pixels: 面向分析的自适应行列混合存储优化** 2018.03 - 至今  
描述: 研究宽表分析场景中负载驱动的行列混合存储, 考虑如何在水平方向上对不同查询动态选择合适的查询时水平数据分片大小  
负责: (1) 实现 Presto Connector 的设计, 动态选择数据分片大小  
(2) 实现 Etcd 分布式锁; 元数据管理; 数据装载工具 Pixels Load
- **广东省重大应用项目“ParaFlow 日志大数据实时分析系统”** 2017.10 - 2018.08  
描述: 基于实体 ID 的哈希数据分片, 提供结构化数据高时效、高吞吐的数据装载和高性能的查询  
负责: 实现元数据管理; 通过 Connector 访问 Paraflow Metadata 的元数据; 自适应列存储布局优化

## 论文/获奖情况

- Rainbow: Adaptive Layout Optimization for Wide Tables (A 类会议 ICDE2018 Demo 2 作)
- 基于微信公众平台的小区信息上报系统设计 (微处理机期刊 2017 第 4 期 1 作)
- 第一届 POLARDB 数据库性能大赛 (程序设计大赛) | 初赛 26/1653 复赛 46/200 2018.12  
负责: 设计存储方案, 以 Range 为核心, 将 Key/Value 数据分离, 划分多个 DB 分片
- 2018 阿里云全球区块链大赛 (创新应用大赛) | 初赛 1/605 复赛 12/140 2018.07
- 第七届中国大学生服务外包创新创业大赛国家级三等奖 (校园文化艺术直播平台) 2016
- 第四届“中国软件杯”全国大学生软件设计竞赛国家级二等奖 (高校教务微信公众平台) 2015
- 2014 年全国大学生英语竞赛 (NECCS) C 类二等奖 | 2015 年三等奖

## 实习经历

- 蚂蚁金服 | 区块链平台部 <可信计算与跨链小组> 实习开发工程师 2019.06 - 至今
  - 负责 mychain 中可信计算插件 mycloak 的发布迭代, kms 多方密钥管理开发工作
- 字节跳动网络有限公司 | Data 部门大数据研发工程师 2019.04 - 2019.05
  - 负责 ClickHouse 服务化、ETL 工具的更新和迭代功能开发, 满足 BI 查询需求

## 社会工作

- 中国人民大学 2017 级计算机硕士班 | 组织委员 2017.09 - 至今
- (本科) 河海大学物联网工程学院学生科协 | 软件设计部副部长 2014.09 - 2015.07