

张震

185-0043-0292

zbestz@gmail.com

系统研发工程师

经历

百度糯米搜索研发部，研发工程师	2016.4 至今
北京航空航天大学，计算机技术，硕士	2013.9 - 2016.3
东北大学，计算机科学与技术，本科	2009.9 - 2013.7

专业技能

编程语言：	PHP, Java, Shell
相关平台或技术：	MySQL, Solr, Redis, Lucene
其他相关知识：	Linux, BigPipe, Spark, MapReduce

项目经历

1. 店铺信息实时流（2016.6 - 2016.10）（PHP, BigPipe, Redis, MySQL）

- 参与设计和实现了糯米店铺信息中心到糯米检索的实时流系统，将店铺信息中心数据在糯米检索的生效时间从天级提升到秒级。
- 店铺信息实时流包括基础信息流和详情信息流，在特定情况下，两条数据流并发会造成推送给检索的数据有误，通过设计并发锁来解决该问题。
- 设计并完成了全网店铺信息实时流，糯米主站信息实时流与糯米检索开放平台的架构改造，使得全网店铺实时流与糯米检索开放平台实现互通，并使得糯米检索开放平台对特定的业务方支持从 200QPS 提升到 1000QPS 以上。

2. 科技搜索（2013.11 - 2015.3）（Java, JSP, MySQL, Solr）

- 基于 Servlet、JSP 和 Solr 完成了包括部分数据的爬取，索引建立，数据检索以及科技领域相关数据统计等功能的设计与开发。
- 在搜索功能实现中将通用逻辑进行封装，并将其应用在多个项目中。
- 在搜索功能实现过程中为 Ansj 分词器编写了 Solr 插件，使 Solr 能够使用 Ansj 进行分词。
- 使用 Solr 的内置的 Hash 算法进行网页全文去重，并编写了一个去重算法来处理网页摘要去重。
- 使用 Nutch 的 LinkRank 功能并结合自定义排序规则对科技相关网页结果进行排序，提升了科技类网页的权重。
- 设计算法计算出给定科技领域中的知名人物和知名研究机构，并达到了实时展示的要求。

3. 糯米 UGC 升级改造项目（2016.5 - 2016.8）（PHP, MySQL, Databus, Redis, NMQ）

- 由于历史原因糯米店铺配送流程比较复杂，重新梳理并简化糯米店铺评分配送流程，使得评分配送数据流程更加清晰。
- 解决了新 UGC 检索和储值卡功能使用 UGC 数据的兼容性问题。
- 对糯米评论数据实时到检索和离线文件到检索相关功能进行改进。
- 部分 UGC 离线任务运行不稳定，定位问题并修改了相关工具类，使得这些离线脚本能够稳定运行。

4. 词语相关性计算（2015.9 - 2016.1）（Java, Spark, Hadoop）

- 设计算法，根据两个词语在文档集合中的出现频次，共现频率，共现次数等指标对两个词语的相关性进行了计算，分别在 Spark 和 Hadoop 上对相关算法进行了实现，并采用爬取的 2000 万万方文献摘要数据进行了实验。