

# 陈晓辉

男, 27

北京市海淀区科学院南路 2 号院 3 号楼搜狐媒体大厦 100190

+86 138 1179 8428

✉ [sdqxcxh@gmail.com](mailto:sdqxcxh@gmail.com)

🌐 [www.cxh.me](http://www.cxh.me)

👤 [chenxiaohui](#)

🌐 [chenxiaohui](#)

## 教育经历

- 2010.9–2013.6 清华大学软件学院, 软件工程硕士, 方向: 领域语义模型存储  
2006.9–2010.6 北京航空航天大学计算机学院, 计算机科学与技术本科, 保送清华软件学院读研

## 工作经历

- 2014.11–今 搜狐大数据中心推荐系统和精准广告团队, 高级研发工程师 (后台系统开发), L3.1, 绩效 AA  
2013.7–2014.10 支付宝北京基础数据部 OceanBase 团队, 研发工程师 (分布式数据库开发), P5, 绩效 3.5

## 外语能力

- 具备良好的英语听说读写能力, 熟练阅读英文文献。大学英语四级 (548)、六级 (549) 均已通过。

## 获奖情况

- 2010 北京航空航天大学优秀毕业生  
2010 北京航空航天大学计算机学院科技实践奖学金  
2009 北航第十九届“冯如杯”学生学术科技作品竞赛二等  
2008 北京市奥运会志愿者先进个人

## 个人陈述

- 专业兴趣浓厚, 热衷于钻研各种技术, 参与了很多项目, 实践能力和学习能力均比较强; 乐于分享, 关注开源, 长期维护技术博客; 崇尚高效率快节奏的生活方式, 喜欢对现有方法进行改进以提高效率。
- 受过规范的开发流程、做事风格和代码风格训练, 有较强的团队意识和一定的组织领导能力, 独立领导完成过中小型项目。扎实肯干, 易于相处。

## 专业素质

- 多年 C/C++ 经验, 熟悉主要分布式算法, 参与设计实现过分布式存储系统, 对关系数据库理论、NoSQL 有一定了解和经验。
- 对开源分布式存储系统有浓厚兴趣, 了解 redis、memcached、leveldb、zookeeper 等的实现。
- 熟悉推荐系统的基本算法和架构, 实现过推荐系统上下游服务。
- 熟悉精准广告系统的基本算法和架构, 了解搜索引擎的原理, 参与过全文索引和广告索引系统开发。
- 熟练 Python 语言, 使用 Django 等框架开发过 web 应用。用 Python 写过较多脚本, 部分开源在 Github 上。
- 熟悉 Java 语言, 独立开发过 Java 后台服务和分布式系统, 了解 Spring 等框架。
- 熟悉 JQuery/HTML/CSS 和 PHP, 独立做过网站前端、后端设计, 对网站开发和大型网站性能调优有一定经验。
- 熟悉基本 Linux 服务器运维, 熟悉 Linux Shell 编程。
- 熟练使用 SVN、GIT 等进行代码管理, 熟悉常见代码 review 工具、自动部署工具, 熟悉 VS、Eclipse、vim 等开发环境和工具。开发过多个 vim、sublime 插件。

## 项目经验-工作

- 2014.2–2014.11 阿里巴巴分布式数据库 Oceanbase1.0 版本设计实现  
参与 1.0 云版本事务模块的设计、分布式事务设计 (两阶段提交) 实现和日志同步设计 (基于 Raft)。设计主算法的测试框架。熟悉整个系统的实现原理。C++ 实现。

- 2013.7-2014.2 分布式数据库 **Oceanbase0.5** 版本重构  
Schema 管理（异步更新 Schema）、RootServer 重构、内部表重构等。C++ 实现。
- 2014.6-今 搜狐精准广告 **index** 系统重构  
维护并重构精准广告的索引系统，实现上类似于 Lucence，提供对满足某个定向条件的广告候选集的搜索，并持久化索引到磁盘。C++ 实现，分布式环境。
- 2014.4-2015.5 基于 **Zookeeper** 的配置管理客户端  
通过 Zookeeper 管理配置项，支持在线修改和事件通知，封装了 Zookeeper 的 Java 的客户端，目前已用在搜狐大数据中心多个业务。详见 [链接](#)。Java 实现。
- 2014.3-2014.4 搜狐微门户视频 **E&E** 策略 **index-server** 开发  
按照策略对刚进入系统的短视频进行试探性曝光 (Explore)，曝光到一定次数并获得相应点击的视频进入投放列表 (Exploit)，参照 Yahoo UCB 相关论文。目前已经上线并在上海地区运行。Java 实现。
- 2014.3-2015.4 搜狗输入法个性化推送平台  
通过聚类把用户分为多个类别，对不同类别用户编辑相关类别的热门新闻并推送。平台主要提供候选编辑，并对接搜狗的推送接口。人工介入以满足新闻质量要求。基于 Django。
- 2015.1-2015.3 推荐系统各业务线监控和人工标注系统  
通过 Thrift 接口从各个业务线取数据，Join 数据库对应新闻数据并展示以监控算法效果。提供为新闻手动打标签的界面为算法人工纠错。基于 Django。
- 2014.11-2015.1 手机搜狐个性频道实时 **UA(User Attributes)** 系统  
根据用户最近一段时间的浏览记录给用户打标签。参见 [链接](#)。Java 实现，通过 zookeeper 管理 server 列表并实现 HA。

## 项目经验-学校/创业

- 2013.3-2013.6 建筑领域语义 (**IFC**) 编译器和数据中心  
研究生毕设。实现了 IFC 语法的编译器，动态解析实体 Schema 并映射到 Python 语言，解析中间结果并提供存储、检索。基于 Python + MongoDB。参见 [链接](#)。
- 2011.3-2011.6 适合我旅游网 [链接](#)  
创业项目。负责了整个网站大部分的开发、运维工作，网站上线后运行良好，后因运营问题关停。项目包含了团购、支付和酒店机票爬取、搜索 (简单的垂直搜索) 等。基于 LAMP。代码见 [链接](#)。
- 2012.11-2013.3 矿山车联网监控系统  
创业项目。负责软件系统。包括服务端与硬件的接口和数据存储 (Python+TCP+MySQL)，前台的展示和统计 (LAMP)。
- 2009.11-2010.5 寻宝物语 SNS 游戏 (**TreasureStory**) [链接](#)  
创业项目。负责部分后台、后台管理系统开发以及数据平衡，2010 年 5 月游戏在 facebook 上线之后，在 Zynga 该类游戏中进入 Rank30，后登陆多种 SNS 平台。后台基于 LAMP。
- 2010.11-2011.3 金牌国旅企业管理系统  
外包项目。为金牌国际旅行社有限公司定制企业业务流程系统以及人事、财务管理软件。领导了整个项目开发并独立设计了系统架构，使用 .Net 平台，C/S 架构，数据库使用 SQLServer，基于异步非阻塞 Socket，目前仍在使用中。
- 2010.4-2010.6 **Admire** 系统 **MBus** 组件的改进和优化  
本科毕设项目。主要对 Admire 系统 (视频会议) MBus 消息总线 (跨进程、主机消息中间件) 进行改进和优化，使之支持 IPv6、变长消息格式，以及基于 CryptoAPI 实现 SSL 安全协议。VC 实现。

## 项目经验-其他

- vim、sublime plugin vim-config(vim 配置, [链接](#))、vim-cscope(关联 cscope, [链接](#))、ifdef-generator(生成 ifdef, [链接](#))、postreview.vim(生成 post-review 脚本, [链接](#))、file-list(文件管理器, [链接](#))、AllinOneCompile(编译选项管理, [链接](#))、BufItv2(缓冲区管理, [链接](#))、generate-ref(sublime 插件, 生成参考文献 [链接](#)) 等
- python 脚本 简历生成器 (通过文本文件生成简历 tex 的 generator) [链接](#)、category(分类电子书, [链接](#))、music-category(分类音乐, [链接](#))、proxy-pool(爬取免费代理并验证, [链接](#))、scrawl(url 规则爬虫, 自动切换代理防止封 ip, [链接](#)) 等
- 其他项目 C0 文法编译器 (课程设计, C, [链接](#))、正则表达式匹配器 (课程设计, C++, [链接](#))、中建地产企业产品线研发数据平台 (实验室项目, LAMP)、某卫星地面测试系统 (外包项目, MFC)、Admire 系统网站的 WebService 封装 (实验室项目, ASP.NET)、视频会议高清采集 (实验室项目, VC/DirectShow)