

程世文

硕士 ◇ 计算机科学与技术 ◇ 1991年1月

(+86) 15201523580 ◇ chengshiwen0103@gmail.com

北京市海淀区清华大学东主楼9-325室, 中国, 100084

<http://chengshiwen.com/> ◇ <https://github.com/chengshiwen>

兴趣领域

大数据: 大数据分析 & 处理, 如大数据平台架构设计、大数据算法研究、大数据挖掘
性能测量: 互联网服务性能测量, 如性能监测、异常检测、异常定位, 自动化、智能化监测
算法研发: 算法研究与设计, 如基于特定场景的算法研究, C/C++算法、大数据算法研究
后端开发: 后端设计与开发, 如服务器、缓存、数据库、负载均衡、分布式系统设计与开发

教育背景

清华大学, 硕士 计算机科学与技术系 指导老师: 裴丹 研究方向: 大数据分析 & 处理, 异常检测 & 定位	2014年8月 - 2017年7月
清华大学, 学士 计算机科学与技术系 排名: 前20%	2010年8月 - 2014年7月

主要经历

大数据分析 & 处理系统, 清华大学计算机系网络所 系统架构师 & 核心开发者	2016年3月 - 2016年8月 北京
<ul style="list-style-type: none">负责设计大数据分析 & 处理系统的架构, 该系统可以对实时流数据和离线数据进行处理, 其使用了当前流行的开源框架, 包括Kafka, Logstash, Spark, Hadoop, OpenTSDB, HBase, Redis, MangoDB等.该系统由数据采集、数据分发、数据处理、数据存储、数据查询、数据可视化组成.数据处理、数据存储、数据查询、数据可视化模块的核心开发者.数据请求峰值达到80万QPS.	
技术部研发, 北京水木快车科技有限公司 联合创始人 & CTO	2014年9月 - 2016年1月 北京
<ul style="list-style-type: none">北京水木快车科技有限公司的联合创始人和CTO, 主要产品是基于匿名社交的移动互联网应用: 枝丫.领导管理技术部, 并对所有的技术方案进行决策.后端的设计与开发者, 使用Python/Django、MySQL、Redis、负载均衡、分布式系统等技术, 并涉及到即时通讯, 消息推送, 短信通知, 云存储, CDN, 微信公众号, 微信支付, 统计分析等开发.开发了一套自动化、可视化的运营工具.巨大的成长: 专业技能, 沟通交流, 团队协作, 团队管理, 解决问题.	
运维部BFE, 百度公司 实习	2014年3月 - 2014年8月 北京
<ul style="list-style-type: none">研究网页搜索页面访问次数(PVs)的异常检测.研究并实现多种异常检测算法: 斜率、均值-阈值、时序分解(TSD)、滑动平均、ARIMA、CUSUM、三阶指数平滑(Holt-Winters)、聚类分析(K-Means)、主成分分析(PCA)、奇异值分解(SVD)等.	

- ◇ 采用了人工标注、自适应、自学习等机制，有效降低了误报率、漏报率，并对节假日等做了特定优化.
- ◇ 设计并实现内部的异常检测系统，使用了Python/Django、Memcached、并行化等.

edX, OPENedX

2013年10月- 2013年11月

贡献者

开源项目

- ◇ 参与开发国外edX开源项目 (<https://github.com/edx/edx-platform>, 2800+ github stars) 并贡献代码，其是OPENedX (<https://open.edx.org/>) 的子项目.

学堂在线, 清华大学

2013年5月- 2013年10月

核心开发者

北京

- ◇ 学堂在线 (<http://www.xuetangx.com/>) 是清华大学研发的中文MOOC (大规模开放在线课程, 简称慕课) 平台, 于2013年10月10日正式向全球发布, 面向全球提供在线课程.
- ◇ 独立负责视频模块的设计与开发, 涉及视频前端播放器开发, 后端视频上传、存储、同步、与第三方CDN厂商合作开发等工作.
- ◇ 内部多个代码分支的核心开发者和审核者.

个人技能

精通	C/C++, Python, Linux, MySQL, Redis, Web Full Stack, Algorithm, Data Structure
熟练	Java, PHP, Shell, Memcached, MongoDB, SQLite3, HTTP, TCP/IP, Big Data
熟悉	Nginx, Apache, Nagios, Kafka, Spark, HBase, OpenTSDB, Hadoop, Mesos, Docker
工具	Git, SVN, Vim, Markdown, Maven, yum, rpm, apt-get, pip, Fabric, Microsoft Office
语言	中文, 英语 (基本听说/良好读写)

奖项荣誉

新鸿基地产郭氏基金·清华大学励学金	2014年5月
清华之友—黄奕聪伉俪奖学金	2013年9月
新鸿基地产郭氏基金·清华大学励学金	2013年5月
清华大学学业单项优秀奖奖学金	2012年9月
国家励志奖学金	2012年9月
新鸿基地产郭氏基金·清华大学励学金	2012年5月
清华大学计算机系优秀学生干部	2011年9月
国家奖学金	2011年9月
第25届中国数学奥林匹克 (CMO) 银牌	2010年1月
第24届中国数学奥林匹克 (CMO) 银牌	2009年1月
第8届中国西部数学奥林匹克 (CWMO) 金牌	2008年11月

社会活动

计算机系党支部组织委员, 负责院系党支部各项活动的组织工作	2013年8月- 2014年7月
清华大学计算机系25班辅导员, 负责新生学习和生活引导及班级建设等工作	2012年8月- 2014年7月
清华大学团委实践组干部, 负责全校各院系的寒假和暑期实践的组织工作	2011年9月- 2012年8月
计算机系团委实践部干部, 负责计算机系寒假实践和暑期实践的组织工作	2010年9月- 2011年8月

个人评价

勤 奋	9 / 10	学习能力	10 / 10
沟通交流	8 / 10	团队协作	8 / 10
解决问题	9 / 10	责 任	9 / 10
专 注	9 / 10	创新能力	7 / 10