

汪俊雄

TEL: 178-6310-9085 ✉: wud@wangjunx.top

Blog: wudly.cn Github: JX-Wang

教育背景

哈尔滨工业大学 - 本科 - 网络空间安全(Cyberspace Security)

2016.09 - 至今

在校期间拿过校级 ACM 奖, 通过 CET4, 大二上加入学校网络与信息安全技术研究中心并参与实验室工作, 主要研究方向为域名安全, 技术方向涉及后台开发, 数据处理等。

项目经历

域名基础信息获取 API

2018.10 - 至今

功能: 辅助其他实验室高效的获取 WHOIS 数据, 避免系统的多次部署。

负责部分及其特性:

WHOIS 信息获取, API 系统框架的设计与实现, 系统后台, 测试及运维。

- 分布式 API 以主节点转发, 从节点群探测的为模型, 多个从节点同时工作来高效的处理数据。
- 以 flask 为框架, 基于心跳的主从节点间状态的感知监控, 以便系统日常维护。
- 基于 uWSGI 和协程 gevent 以及 monkey_patch 补丁实现的高并发数据获取引擎。
- 基于日志(python logging 模块)的监测模块儿监测 API 运行状态, 了解 API 实时运行状况。

系统 API 单节点(单核)日探测量约 420w+, 数据正确率 70+%。

非法域名监测系统

2017.11 - 至今

功能: 对非法域名进行监测, 数据挖掘以及关联分析

(此系统为实验室与国家某中心合作项目)

负责部分以及特性:

域名 WHOIS 信息获取, 中间件(辅助开发), 系统后台以及 MySQL 数据库的日常维护。

- 基于 Tornado 的 RESTful 风格的 API 数据中间层。
- 基于 schedule 的定时数据的统计, 关键数据定时入库以及重要数据定时导出。
- 基于 token 的无状态身份认证。
- 优化的数据库操作以及数据缓存机制。

系统目前收录及监控了约 4500 万条域名相关的 WHOIS, WHOWAS 以及其他数据。

域名基础信息获取系统

2017.10 - 至今

功能: 通过与 WHOIS 服务器交互, 大规模获取域名 WHOIS 信息并解析。

负责部分及其特性:

基础数据获取, WHOIS 数据获取, 系统后台。

- 通过爬取 IANA 获取基础 WHOIS 服务器数据, 支持 WHOIS 探测过程。
- 底层采用 socks 代理, 核心数据轮询获取, 基于 RFC 协议等方式深度优化了 WHOIS 获取过程。
- 采用动态解析, 正则对原始数据进行提取, 将形式各异的 WHOIS 数据解析出统一特征项便于统计与分析。
- 额外的基于日志(python logging 模块)的监测模块儿监测系统运行状态, 了解系统实时运行状况。

系统目前获取了 1.2 亿的域名数据, 覆盖 1500 个顶级域, 以及 1800 个 WHOIS 服务器。

专业技能

熟练掌握 python, 有部分并发, 爬虫(Scrapy), 数据库(Mysql, MongoDB)以及 socket 的开发经验分布式, 有 Hadoop 集群生态系统内部 Zookeeper, HDFS, Yarn 的搭建和使用经验; 有基于 Tornado, Flask(Python Web 框架)的 RESTful 风格的 API 开发经验, 熟悉 uWSGI。有较丰富的 Linux 使用经验, 熟悉 MySQL 数据库, 有 Redis 和 Elasticsearch 的开发经验; 熟悉掌握 Git, Vim, Linux 命令, 了解 C++, 熟悉常用的算法, 数据结构; 除此以外有部分前端开发经验, 有 Tkinter(python UI 框架)的编程经验; 本科期间探究过伪基站(基于 OsmocomBB 和 OpenBTS 搭建), 有计算机视觉相关编程经验(基于 OpenCV3.4.2.16)。

自我描述

本科期间有较为丰富的项目经历, 主要为域名安全, 数据分析, Web 后台相关。对开源社区兴趣较浓, 个人博客更新较活跃。自学能力强, 善于独自解决问题, 责任心强, 技术尚可。踏实肯干, 有较好的数学, 英语基础, 兴趣较广, 爱好指弹和音乐。

期待岗位

期待实习岗位: 后台开发工程师 / Python 工程师 / 数据研发工程师 / 其他相关岗位都可, 预期实习时间为 2-3 个月, 可调节, 每周工作 4-5 天, 可调节。