

YANSONG LI

Peking University, Beijing

✉ liyansong@pku.edu.cn

☎ 185-1466-2277

🌐 <https://deepcooking.cn>

EDUCATION

M.S. in Computer Science

Peking University 📅 2018.9 – 2021.7

B.S. in Computer Science

Jilin University 📅 2013.9 – 2017.7

INTERSHIP

Software Development Engineer

ByteDance (Data-System-Network)

📅 2019.6 – 2019.10 📍 Beijing, China

Traffic Monitoring

- In order to **monitor** whether the **network traffic in the data center** is abnormal, we have deployed **sflow** on the switch. The software **samples data-gram** on the switch and sends it to the designated server.
- Implement **trafficmonitor** to collect UDP packets, and write to redis and Kafka after analysis.
- In order to **efficiently use memory** to store data and improve the **accuracy** of hot data statistics. I implemented the **heavyguardian** data structure to save hot data.

NAT Traffic scheduling

- Establish the **NAT** servers of CT, CU and CMCC respectively, then we can send route to server.
- The server establishes a BGP connection with the router through **ExaBGP software**.

HONORS

- 2015-2016 The First Prize Scholarship
- 2016-2017 The Third Prize Scholarship
- 2016-2017 The University Excellent Student
- AceTradeer V1.0 software copyright

PROJECT

Acetrader futures trading software

- This project is mainly based on the CTP interface of futures provided by the Shanghai futures exchange. Through rewriting the interface, it realizes the function of subscribing to futures market, and sending trading instructions (open/close) to futures companies.
- The main advantage is that it can improve the transaction speed.
- The main framework is written in C++ and QT.

Chinese named entity recognition

- Based on tensorflow, use BiLSTM + CRF model for training.
- Set batchsize = 64, use adam optimizer and randomly initialize embedding.
- The F1 value is 88%.

Full text retrieval system based on Lucene

- Using pdfbox to extract and save the content of English PDF file to mongodb database.
- Using Lucene to index files.

SKILLS

C++

Python

Golang

CET-6

Operating System

Data Structure

Computer Network

李延松

北京大学信息科学与技术学院

✉ liyansong@pku.edu.cn

☎ 185-1466-2277

🌐 <https://deepcooking.cn>

教育背景

北京大学 信息科学技术学院 硕士

计算机系统结构 2018.9 - 2021.7

吉林大学 计算机科学与技术学院 学士

计算机科学与技术 2013.9 - 2017.7

实习经历

软件开发工程师

字节跳动 (data-系统-网络)

2019.6 - 2019.10 北京, 中国

Traffic Monitoring

- 为了监控数据中心的网络流量是否异常，我们在数据中心的交换机上部署了 Sflow，该软件采样交换机上的数据报发送给指定的服务器
- 服务器上面部署 TrafficMonitor 收集指定 socket 上面的 UDP 数据包，分析后写入 Redis 和 Kafka
- 为了高效的使用内存存储数据，以及提高热点数据统计的正确率，实现了 HeavyGuardian 数据结构来保存热点数据

NAT 流量调度

- 分别建立电信、联通、移动的 NAT 服务器，通过 ipip 库得到每个 ip 对应的运营商，然后将路由发送到对应服务器
- 服务器通过 ExaBGP 与路由器建立 BGP 连接，发送和接收路由更新消息

奖励荣誉

- 吉林大学 2015-2016 学年三等奖学金
- 吉林大学 2016-2017 学年一等奖学金
- 吉林大学 2016-2017 学年校优秀学生
- “AceTradeer 程序化交易系统 V1.0” 软件著作权

项目经历

AceTrader 期货交易软件

- 该项目主要是基于上期技术提供的期货 CTP 接口实现的，通过对接口的重写，实现了期货行情订阅的功能，返回实盘实时的一档行情；并实现了基于给定的价格向期货公司发送交易指令（开/平仓）
- 与以往的期货交易软件相比，主要的优势是可以在开仓成功的瞬间，基于设定的价格，发送出平仓单来提高平仓的成功率。
- 主要框架使用 C++ 来编写，图形化界面部分使用 Qt 来编写

中文命名实体识别

- 基于 TensorFlow，使用了 BiLSTM+CRF 的模型来进行训练
- 设定 batchsize=64，使用 Adam 优化器，随机初始化 embedding
- F1 值为 88%

基于 lucene 的全文检索系统

- 使用 pdfBOX 将英文 PDF 文件中的内容提取存到 mongodb 数据库中
- 使用 lucene 对 PDF 文件建立索引，实现对 PDF 文件的全文检索

技能

C++

Python

Golang

CET-6

操作系统

数据结构

计算机网络