**通信工程学院毕业设计（论文）周记 第 七 周**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **专业** | **信息安全** | **题目** | **基于Storm的实时推荐系统** | | |
| **学生姓名** | **冯小川** | **学号** | **01123111** | **指导**  **教师** | **张卫东** |
| **1、主要工作内容和进展** | | | | | |
| **完善系统方案与算法的设计，确定系统缓存数据库选择redis，以及安装其他的一些数据库Mysql和Hbase，这些数据库在系统中都会用到。确定编程语言为Java，IDE为IntelliJ IDEA 14，项目采用Maven进行管理，用git进行代码版本控制。编写爬虫进行系统所需数据的采集，并对采集的数据进行过滤，清洗，确保数据的正确。** | | | | | |
| **2、存在的主要问题和解决办法与思路** | | | | | |
| **存在的问题：抓取数据时需要模拟登陆且有目标站点有一定的反爬。**  **解决办法与思路：分析登陆过程，程序进行模拟登陆取得Cookies，在爬取得过程中进行速度限制和加上代理以应对目标站点的反爬。** | | | | | |
| **3、下周工作计划** | | | | | |
| **开始进行程序代码的编写，并将数据导入进行测试，改进系统的设计和算法。** | | | | | |
| **4、导师意见** | | | | | |
| **指导教师（签名）：**  **年 月 日** | | | | | |

说明：1. 本表每周由学生填写，导师签署意见。

2. 此表格作为答辩资格审查材料之一。

3. 文中字体采用宋体小四号，行间距为固定值20磅。