实时检测共享账户需求分析说明书

## 目 录

[1引言 1](#_Toc14295)

[1.1 背景 1](#_Toc3680)

[1.2用户特点 2](#_Toc23722)

[2 需求分析 3](#_Toc31884)

[2.1总体需求分析 3](#_Toc15190)

[2.2需求描述 3](#_Toc10135)

[2.2.1 概要分析 3](#_Toc7463)

[2.2.2 技术支持 3](#_Toc12529)

[2.2.3 参考资料 3](#_Toc28116)

# 1引言

## 1.1 背景

现在是一个数据大爆炸的时候，各种各样的数据漫天飞。开发商可以把这些数据拿来做生产数据，也可以拿这些数据进行离线分析操作，还可以去实时的抓取数据，然后对他们进行处理，最后得出想要的结果。

我们在运营商骨干网上采集现网流量流向信息，根据这些原始信息检测账号是否存在异常，如果多个终端使用同一个宽带账号，超过一定阈值则触发报警机制，例如阈值为5时，同一个账号同时连接的终端数量不能超过该值，如果超过则报警。

# 2需求分析

## 2.1总体需求分析

进行实时的数据采集以及利用大数据知识进行合理的数据分析主要是在学习大数据课程的基础上，在完成数据采集，数据分析，数据存储这三部分实验的基础上，对大数据开发整体进行一个模拟，通过实践加深对各个部分的管理功能的认识，还能进一步分析各个部分之间的联系，最后达到对完整开发流程的理解。同时，可以提高运用大数据知识解决实际问题的能力；锻炼实际的编程能力、创新能力及团队组织、协作开发软件的能力；还能提高调查研究、查阅技术文献、资料以及编写软件设计文档的能力。

## 2.2需求描述

### 2.2.1 概要分析

抓取采集到的现网流量流向信息，并根据信息运用相关技术分析异常的账户，每隔5分钟统计一次数据，并将数据写入MySQL中，运用大数据可视化工具将实时数据展示出来

### 2.2.2 技术支持

数据采集：Flume，kafka

数据分析：Storm

数据存储：MySql

据可视化：Echart

### 2.2.3 参考资料

《Hadoop权威指南第3版》

《Storm 实战：构建大数据实时计算》

《Kafka权威指南》