运营商网络流量流向异常账号监测系统

目录

[1引言](#_Toc13744_WPSOffice_Level1) [3](#_Toc13744_WPSOffice_Level1)

[1.1背景](#_Toc16374_WPSOffice_Level1) [3](#_Toc16374_WPSOffice_Level1)

[1.2用户群体](#_Toc14523_WPSOffice_Level1) [3](#_Toc14523_WPSOffice_Level1)

[2需求分析](#_Toc13331_WPSOffice_Level1) [3](#_Toc13331_WPSOffice_Level1)

[2.1总体需求分析](#_Toc5888_WPSOffice_Level1) [3](#_Toc5888_WPSOffice_Level1)

[2.2需求描述](#_Toc7560_WPSOffice_Level1) [3](#_Toc7560_WPSOffice_Level1)

[2.2.1数据样本](#_Toc16374_WPSOffice_Level2) [3](#_Toc16374_WPSOffice_Level2)

[2.2.2检测方法](#_Toc14523_WPSOffice_Level2) [4](#_Toc14523_WPSOffice_Level2)

[2.2.3统计需求](#_Toc13331_WPSOffice_Level2) [4](#_Toc13331_WPSOffice_Level2)

## 1引言

### 1.1背景

以大数据处理方式来做流量异常检测这一问题出发，通过近五年来的科研文献调查研究大数据给流量异常检测带来的变化。包含基于大数据的网络异常流量检测架构，针对海量流量数据异常检测的有效算法，基于大数据网络流量流向异常检测仿真实验平台和数据集。

通过大数据来监测异常的运营商网络流量流向账号，从大量的数据中筛选出异常的账号对其限制，从而有效的保证运营商的利润。

### 1.2用户群体

运营商作为我们的客户，包括中国移动，中国联通，中国电信等。

## 2需求分析

### 2.1总体需求分析

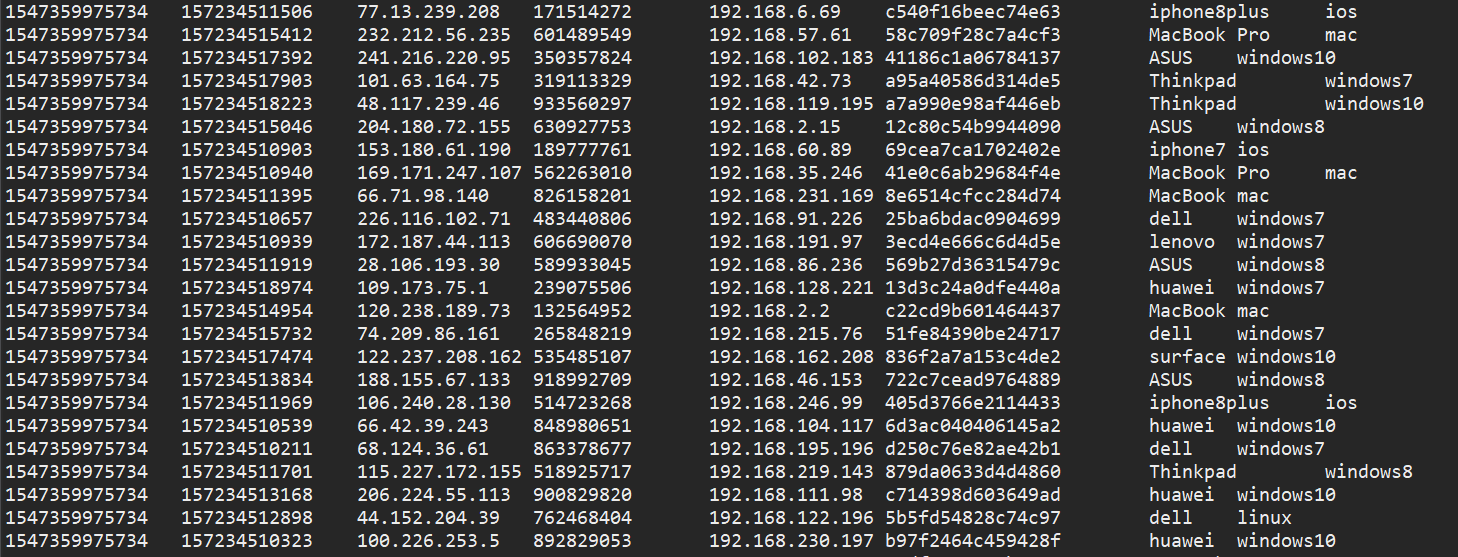
分局运营商骨干网上采集网络流量流向信息，根据这些原始信息检测账号是否存在异常。如果多个终端使用同一个宽带账号，超过一定阈值则触发报警机制，例如阈值为5时，同一个账号同时连接的终端数量不能超过该值，如果超过则报警。

数据量每天500万，每秒峰值60

5分钟是一个周期，每个周期生成一个结果文件，每个周期检测一次共享账号

### 2.2需求描述

#### 2.2.1数据样本



字段解释：

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 字段 | 类型 | 描述 |
| 1 | time | String | 数据统计时间 |
| 2 | userAccount | String | 宽带账户 |
| 3 | userIP | String | 用户以太网IP |
| 4 | qqid | String | QQ号 |
| 5 | natIP | String | 内网IP |
| 6 | cookieValue | String | Cookie值 |
| 7 | devName | String | 设备名称 |
| 8 | osName | String | 操作系统名称 |

#### 2.2.2检测方法

5分钟内，同一个账号（UserAccount）下，只要满足下面的任意一个条件，表示出现共享账户问题：

natIP去重求和数 > 5

qqid去重求和数 > 20

cookieValue + devName + osName去重求和数 > 5

#### 2.2.3统计需求

将1天中，每5分钟的异常账号总数绘制图表（只求异常账号总数），如下图所示：

