

Design and Research on the System of Public Safety Management Based on WIFI Probe

Zihui Ren¹, Yuqing Wang², Lin Wang²

¹China Electronics Technology Group Corporation No.38 Research Institute, Hefei Anhui

²Anhui Sun Create Electronics CO. LTD., Hefei Anhui
Email: renzihui2006@126.com, yuqingwang119@163.com

Received: Jun. 29th, 2017; accepted: Jul. 18th, 2017; published: Jul. 21st, 2017

Abstract

In recent years, with the rapid development and popularization of WIFI technology, the new wave of wireless city construction has been set up in China, and the AP devices have collected a large number of Probe data based on terminal MAC. How to effectively use the huge number of terminal MAC information providing strong information methods for urban public safety management causes close attention, and is also a hot area of research now. Through the distributed storage, calculation and analysis of the data from AP set up the system of public safety management based on WIFI probe in this paper. The system can achieve population counting, congestion warning, path analysis and so on.

Keywords

WIFI, AP, MAC, Public Safety

基于WIFI探针的公共安全管理系统设计与研究

任子晖¹, 王雨晴², 王 林²

¹中国电子科技集团公司第三十八研究所, 安徽 合肥

²安徽四创电子股份有限公司, 安徽 合肥
Email: renzihui2006@126.com, yuqingwang119@163.com

收稿日期: 2017年6月29日; 录用日期: 2017年7月18日; 发布日期: 2017年7月21日

摘 要

近年来, 随着WIFI技术的快速发展和普及, 国内掀起了无线城市建设的狂潮, 前端AP设备采集了大量以

终端MAC为主的探针数据。如何有效利用海量探针数据,为城市公共安全管理提供强有力的信息化手段引起了重点关注,是目前研究的热点问题。本文通过对前端AP采集的数据进行分布式存储、计算及分析,搭建基于WIFI探针的公共安全管理系统,可达到人群计数、拥堵预警、轨迹分析等功能。

关键词

WIFI, AP, 手机MAC, 公共安全

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

据2016年3月15日,360发布的《315 WIFI安全报告》显示,2015年,随着个人家庭网络的迅速普及,以及公共网络建设的快速发展,我国的WIFI迎来了迅猛发展。截止目前,全国活跃WIFI热点数量已超过1.5亿个,公共WIFI热点数量约800万个,个人家庭WIFI热点数量超过1.4亿,较比2014年增长40%。无论在家中,或者外出,民众对无线上网的需求越来越高。而手机上网已成为网民上网的首选,截止目前,手机网民已超过6.3亿。值得注意的是,其中92%的手机网民使用WIFI接入互联网,平均每天每人连接WIFI时长1.1小时。由此可见,WIFI已经成为民众日常生活中不可或缺的一部分。

目前,国内“无线城市”建设也在如火如荼地展开,国务院、地方省政府出台一系列关于建设新型智慧城市、加快宽带城市建设、推动互联网发展、促进信息消费、提升无线网络服务等文件,大力推进、高效指导城市公共场所无线局域网建设;台湾、香港、上海、北京、广东、湖北、江苏、山东、安徽等30余省的超300个城市的无线城市门户上线推广,已上线应用超过1.5万个。

在为民众提供便利的网络环境的同时,如何有效利用WIFI前端设备采集的探针数据,打造新型智慧城市,构建信息化环境下的安全社会,成为城市相关管理者亟待解决的问题[1]。

2. 系统整体设计

基于WIFI探针的公共安全管理系统整体框架如图1所示。

整体架构包含四层,各层作用如下:

前端采集:通过用户接入MAC采集设备(AP),获取移动终端的MAC地址、终端品牌等信息。

传输网络:通过网络将前端采集的数据上传到大数据处理平台。

大数据处理平台:对前端设备采集的数据进行存储和管理。

业务应用系统:利用大数据分析算法,对前端采集的海量数据进行挖掘分析,针对不同的应用要求,快速找到相关信息。

3. 技术路线

基于以上总体设计及业务功能需求,对本系统进行详细设计,在详细设计中,确定实现具体模块所涉及到的主要算法、数据结构、类的层次结构及调用关系。之后进入详细编码阶段,本系统采用Java + Linux为主要软件开发模式,系统的关键技术为大数据处理。根据大数据处理的几个核心阶段,本文分别采用不同的技术实现。本系统数据处理流程如图2所示。

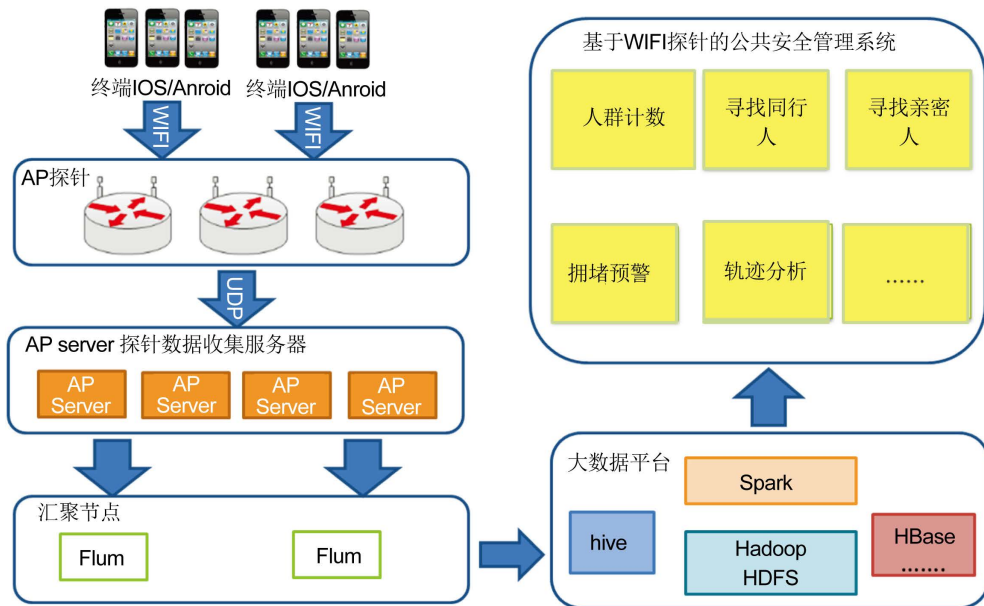


Figure 1. The framework of system

图 1. 系统整体框架



Figure 2. Data processing flow

图 2. 数据处理流程

3.1. 数据采集

本系统采集的数据为 WIFI 探针数据，WiFi 探针数据是 probe 帧[2]：终端设备在接入 AP 之前，会发送一些明文数据包(probe request 帧)给 AP，AP 然后反馈 probe response，以此来标识各自的存在，WIFI 探针就是 probe request 帧里面的数据，它以手机“MAC 地址”为重要组成，同时还包括了 AP 设备所在的经纬度，探测时间，终端类型及品牌等基础数据。

对于无线城市项目，往往会接入多个不同厂商的无线 AP 设备，因此需采用基于不同厂商 AP 探针模块的 UDP 包接收解析技术，实现对不同厂商 AP 设备探针模块发送数据的解析。另外，由于无线 AP 设备数量较大，接收的探针数据量巨大，因此本系统采用基于 UDP 的分布式集群，实现动态扩展，满足海量 AP 探针数据的接收及解析。

3.2. 数据传输

在本系统中，采用分布式架构，接收 AP 设备上报的报文数据，平台具备动态水平扩展能力，根据 AP 探针的数据量动态调整接收服务器数量。统一数据格式，将采集到的不同厂商的报文进行格式统一。通过自定义传输将收集处理后的数据从 AP server 中转至 HDFS、HBase、ElasticSearch 大数据平台。

3.3. 数据存储

基于功能业务需求及规划，本系统采用 HDFS、HBase、ElasticSearch 三种系统进行数据的统一存储。HDFS 是 Hadoop 体系中数据存储管理的基础。本文充分利用 HDFS 分布式优势，将 HDFS 作为基础的文

件存储平台，对于源源不断增长的大数据，HDFS 可进行很好的扩展与清理，并简化了文件的一致性模型，通过流式数据访问，提供高吞吐量应用程序数据访问功能。

面对海量数据的查询，对于系统的数据结构设计 with 性能优化都提出非常高的要求，传统数据库几乎无法完成多条件的查询。为了实现海量数据秒回查询，本系统对需要随机访问、实时读写的数据采用 HBase 数据库，并基于 Elasticsearch 建立分布式海量数据检索服务，作为 HBase 数据仓库的查询入口，实现在几秒内返回结果。HBase 是一个高可靠性、高性能、面向列、可伸缩、实时读写的分布式数据库。其利用 Hadoop HDFS 作为其基础的文件存储系统，利用 HadoopMapReduce 来处理 HBase 中的海量数据，利用 Zookeeper 作为其分布式协同服务。另外，HBase 技术在廉价 PC Server 上可搭建起大规模结构化存储集群，既具备 PB 级数据存储能力，又能节省系统资金投入。

在本系统中，考虑数据量巨大，对分布式服务器节点的性能和存储空间要求较大，通过采用 LZO 压缩技术，将原始数据压缩到 12% 左右，减少了 78% 的存储资源。

3.4. 数据计算及分析

3.4.1. 离线数据分析与统计

本系统基于 Hive 进行标准 SQL 化分布式数据计算与挖掘，它的底层是调用 Hadoop 的 Map Reduce，主要用于离线数据分析和统计，如计算及统计一天中每个 MAC 地址关联 AP 的个数，每个 AP 下关联的所有手机 MAC 个数等。Hive 是一个基于 Hadoop 的开源数据仓库工具，可以将结构化的数据文件映射为一张数据库表，并提供简单的 SQL 查询功能，可以将 SQL 语句转换为 MapReduce 任务进行运行。

3.4.2. 实时数据分析与统计

Map Reduce 框架适合离线计算，却无法满实时性要求较高的业务，如用户行为轨迹分析等。Spark 是一个类似于 MapReduce 的分布式计算框架，其核心是弹性分布式数据集，提供了比 MapReduce 更丰富的模型，可以快速在内存中对数据集进行多次迭代，以支持复杂的数据挖掘算法和图形计算算法，它的底层是调用 Hadoop 的 Yarn，其优点如下：

- 1) 提供分布式计算功能，将分布式存储的数据读入，同时将任务分发到各个节点进行计算；
- 2) 基于内存计算，将磁盘数据读入内存，将计算的中间结果保存在内存，这样可以很好的进行迭代运算；
- 3) 支持高容错；
- 4) 提供多计算范式。

本系统基于 Spark 进行实时数据流分析与计算，如亲密人和同行人的计算及分析。

4. 系统测试

在系统完成开发阶段之后，为了保证系统能达到预期的性能及指标，需对系统的界面、功能、性能及兼容性等进行测试。针对本系统，分别采用了黑盒和白盒的测试方法，经测试，系统操作性良好并且各项性能指标和业务功能都达到了预期目标。

5. 系统应用

5.1. 人群计数

人流量是智能化监控系统中的一个重要指标，在城市管理和决策中扮演着不可或缺的角色。在本系统中，通过统计每个 AP 设备下关联的手机终端数，可得到 WIFI 覆盖区域下的人流量信息，进而统计全市人群信息，根据统计结果描绘全市人群密度，实时监控城市人流密度情况。

另外，在交通枢纽、商场、机场、景区、公园、体育馆等公共场所，通过人群计数可有效辅助其运营组织工作，为人们的生活生产和研究等提供更好更优质的服务。

5.2. 拥堵预警

在本系统中，可以统计平时每个 AP 附近手机终端数量的均值，同时可以统计多个 AP 所代表的某个区域的手机终端数量的均值。通过均值设定合适的阈值 M [3]。

在系统中，对指定场所设置阈值 M ，当该场所中所有的 AP 采集到的手机终端数大于 M 时，系统立即发出拥堵预警，并可联动治安管理人员手机、邮箱等。

5.3. 轨迹分析

通过本系统，可以对任何进入 WIFI 覆盖范围内的手机 MAC 进行轨迹分析。对于指定的 MAC 地址，进入某个 WIFI 覆盖范围后，可以按时间顺序获得该手机接入过的 WIFI 采集设备 MAC 信息集合，通过该集合可以在地图上展示出人员行为轨迹。通过该功能，可以在大型毅行活动、马拉松比赛等现场，通过设定既定的路线、目标手机 MAC 地址，一旦有参赛人员逾越规定路线，系统立即发出报警，也支持赛后发现参赛人员未按规定路线到达终点立即取消其比赛成绩。另外，在景区内对于危险的未开发区域，通过该功能可以迅速发现进入危险区域的人员并加以阻止。

5.4. 寻找同行人和亲密人

寻找同行者：在本系统中，通过基于 AP 场所和时间两个维度的 MAC 地址搜索，计算及分析在相同的时间不同的 MAC 是否有相同的行为轨迹，可寻找同行者。

寻找亲密人：通过对同行次数的计算，得到不同 MAC 之间的亲密度，从而判断寻找亲密人。

6. 总结

利用无线城市发展契机，通过对前端 AP 设备采集的手机 MAC 数据的收集、传输、存储、计算分析等，搭建基于 WIFI 探针的公共安全管理系统，并提出该系统的四大应用场景：人群计数、拥堵预警、轨迹分析、寻找同行人和亲密人，为城市公共安全管理供强有力的信息化支撑和保障。

参考文献 (References)

- [1] 李许舫. 新型无线城市发展研究[J]. 移动通信, 2009, 21(1): 34-37.
- [2] Cunche, M., Kaafar, M.A. and Boreli, R. (2014) Linking Wireless Devices Using Information Contained in WIFI Probe Requests. *Pervasive & Mobile Computing*, **11**, 56-69. <https://doi.org/10.1016/j.pmcj.2013.04.001>
- [3] 高见, 袁得崙. 基于 WIFI 探针的预警系统设计与研究[J]. 中国人民公安大学学报, 2016, 3(1): 89-93.

期刊投稿者将享受如下服务：

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：hjdm@hanspub.org