

基于交通大数据的智能信息服务平台^①

彭晨伟¹, 巴继东²

¹(武汉邮电科学研究院, 武汉 430074)

²(武汉长江通信智联技术有限公司, 武汉 430010)

摘要: 随着智能交通领域的迅速发展, 日益增长的交通数据量已经达到TB甚至PB级别, 智能交通领域也开始运用大数据技术对海量的行车数据进行深入的挖掘分析, 向着构建一个综合性智能交通信息服务平台方向发展. 本文提出了一种基于交通大数据的智能信息服务平台的总体设计方案, 重点研究了系统的总体架构, 应用架构和数据中心的设计等三个方面. 通过测试, 该平台可以很好的满足用户的前期需求, 平台基于分层和分模块的设计思想可以很好的应对用户后期需求变更.

关键词: 大数据; 智能交通; 电子车牌; 物联网; 智慧城市

引用格式: 彭晨伟, 巴继东. 基于交通大数据的智能信息服务平台. 计算机系统应用, 2017, 26(7): 97-103. <http://www.c-s-a.org.cn/1003-3254/5851.html>

Intelligent Information Service Platform Based on Traffic Big Data

PENG Chen-Wei¹, BA Ji-Dong²

¹(Wuhan Research Institute of Posts and Telecommunications, Wuhan 430074, China)

²(Wuhan YCIG iLink Technology Co. Ltd., Wuhan 430010, China)

Abstract: With the development of intelligent transportation technology, the increasing traffic data volume has reached TB or even PB level. The intelligent transportation field has begun to use big data technology to analyze deeply the massive traffic data, to build a comprehensive intelligent traffic information Service platform. In this paper, an intelligent information service platform based on traffic big data is proposed. The paper mainly analyzes this platform from the overall architecture, application architecture and data center design of the system. Through the test, the platform can meet the early needs of the users. The platform-based hierarchy and sub-module design can be a good response to the user needs change later on.

Key words: big data; intelligent transportation; ERI; IOT; smart city

近几年来, 随着信息技术与互联网技术的飞速发展, 大数据成为最近最为火爆的技术概念, 尤其是在贵阳召开的“中国大数据产业峰会”之后, 大数据已经成为国家和各个城市在IT基础领域下一步的重点发展方向. 继大数据分析 & 挖掘技术在淘宝、京东、亚马逊^[1-3] 等电商行业取得巨大的商业成功后, 大数据开始融入人们生活的方方面面.

智能交通系统是交通领域的一项重大工程, 也是

各个国家的战略方向. 根据我国《国家中长期科学和技术发展规划纲要(2006-2020年)》^[4], 智能交通将成为国内未来交通运输业优先发展的主题. 伴随着智能交通战略的发展, 海量的交通数据涌现, 大数据处理技术为智能交通发展带来新的机遇和挑战. 传统的交通出行信息主要通过广播来向用户传达, 同时信息服务平台在获取实时交通情况也存在一定困难, 这就不可避免的存在: 1) 交通信息发布实时性不高; 2) 可提供的信

① 收稿时间: 2016-11-07; 收到修改稿时间: 2016-12-05

息服务种类单一,多样性缺失;3)路段交通情况共享准确性不高;4)系统平台的升级转型困难等缺点.基于大数据的智能信息服务平台结合了云计算、机器学习、数据融合与数据挖掘等先进技术^[5,6],与交管、公安、环保等部门合作,从而能够更加精准的为用户提供各项专业服务,智能交通信息服务平台具备:1)为交管部门提供具有各种时效性的实时道路交通导流、阶段性交通拥堵情况预测、年度交通管网数据分析等服务;2)多维交通信息数据源的融合,为更加准确的分析用户行为提供数据支撑^[7],从而为各部门协调管理提供技术支持;3)现如今物联网技术促使智能交通行业发展迅速^[8],再加上数据量增长的势态迅猛,信息服务平台需要有一定的宽容度为以后的平台升级节约成本等.本文从大数据的角度出发,提出了基于交通大数据的信息服务平台的开发与研究.

1 交通数据源分析

交通数据源是指智能交通信息服务平台中涉及到的基础数据和业务数据的来源地.其中业务数据来源是指的电子车牌采集点、卡口设备以及视频设备.

电子车牌(electronic registration identification of the motor vehicle)^[9]是一种基于RFID(Radio Frequency Identification)技术的电子信息化标签,其具有高精度识别、高准确采集、高灵敏度的特点,电子车牌包含该车辆基本信息,可以帮助高频识别读写器,也即道路网络的采集点来采集各机动车的基本数据,再通过采集点本身的地理位置信息与车辆的基本信息进行关联就可以产生丰富的车辆行车数据.

电子车牌端采集的交通数据在数据量、准确度、效率、复杂程度等天然具备车联网大数据端的属性,这些数据不仅在数量上庞大,在信息的维度上也丰富,这对后期的数据挖掘、数据变现产生直接影响.

除采集点采集的车辆行车数据之外,还包含与本系统有密切业务往来的其他平台的数据.它们包括交警执法记录,环保部门黄标车基本数据,路桥收费管理处收费记录,电子停车场数据,高速路收费记录以及市政银行等机构的数据.

2 系统平台总体设计

2.1 总体设计思路

城市交通行业数据既有静态数据,又有实时动态

数据;数据类型既包括结构化数据、半结构化数据,也包括大量非结构化数据;数据管理在组织上存在多部门之间的数据整合与交换等实际业务情况.考虑到以上因素,本系统采用云计算+大数据处理平台+数据服务总线的设计框架,可以满足高可靠、高可用建设成本低等要求.

本平台的设计思路如下,按功能职责分类主要分为通用服务组件和数据整理与挖掘两个大的部分,其中通用服务组件又包含了四个子功能:1)交通流量常规统计与模糊分析;2)车辆分类识别;3)实时交通信息发布与诱导;4)区域交通流量分析与告警.数据整理与挖掘需要长期的数据积累和不断的算法优化,将在平台的后期逐步完善.无论是通用服务还是数据挖掘功能,都依托于对数据的高效处理,结合具体的业务数据特点和业务需求来制定相应的处理方案将直接影响平台的运行效率.本平台的特点是数据量庞大,所以需要对一些特定的实时性要求较高的业务进行优化处理,譬如建立索引等措施实现海量数据的快速查询.而对于海量数据的存储与处理需要建立分布式、多节点的集群服务器,充分利用服务器资源和充分使用业务模型可以进行趋势预测和决策分析.计算平台和存储平台可以根据实际的需求实现线性的弹性扩展是系统设计对宽容度的考虑,可保证平台长期满足业务需求和升级换代的低成本体现.

2.2 系统总体框架

智能交通信息服务平台的总体设计参照物联网层次模型,采用分层设计思想,每一层为上一层提供面向服务的接口调用,同时结合物联网、云计算、大数据、服务计算等技术特点,平台总体设计总共分为5层,由低到高分别为感知设备层、传输层、平台层、服务层和展示层^[10].采用分层架构不但可以降低系统各个模块之间的耦合性,同时提高模块之间的功能内聚性,而且有利于在系统实际开发过程中实现代码的复用,大大提高了工作效率并降低开发成本.系统总体架构图如图1所示.

2.3 系统各层功能分析

1)感知设备层:感知设备层是指能够对标识对象,即机动车的行为进行感知,并执行采集操作的所有方法的集合.设备主要包括RFID设备、卡口设备、视频设备,完成主要完成车辆标识、位置、时间、状态等信息的协议适配、信息采集和上传,采集的信息主要

包括电子车牌信息、抓拍图片、标清视频以及位置等其他信息。

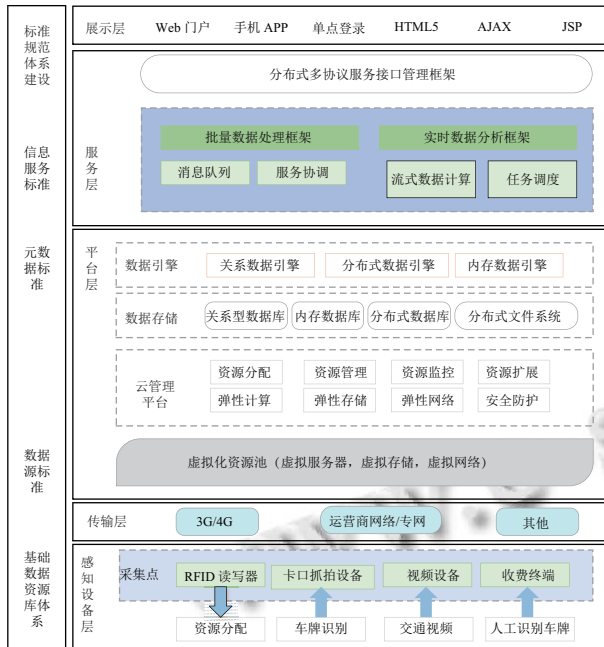


图1 系统总体框架图

2) 传输层: 传输层是将感知设备层感知(采集)到的行车数据传输到数据中心。提供城域网、Wi-Fi、2G/3G/4G等多种通信方式, 实现采集端与后台双向数据传输。

3) 平台层: 平台层包含数据处理引擎, 数据存储和云平台管理三个子功能。其中数据处理引擎和数据存储两者是协作统一的关系, 数据处理引擎将传输层传递过来的基本数据进行剔重、清洗和补漏等操作, 把原始数据转换为系统的统一标准数据格式, 数据存储是将各类数据按照规则存入关系型数据库或分布式数据库; 云管理平台完成对整个支撑服务平台相关IT资源的管理和调度, 为平台其他部分提供资源服务。平台层是整个系统的核心, 采用云计算和大数据技术, 支持对底层的物理资源及各类数据进行有效管理, 为上层的业务应用提供部署环境和业务构件, 在大数据处理框架Hadoop下提供通用服务组件。利用虚拟化技术实现在线迁移、高可用等资源调度技术, 利用storm、map-reduce技术实现实时、批量的大规模数据处理, 建立服务SOA架构的企业服务中心, 建立与外部系统之间的数据交换系统。

4) 服务层: 服务层访问基于平台层提供的服务组

件, 开发业务应用, 并封装为Dubbo架构的服务接口(可暴露为多种协议, 如dubbo://, rmi://, hessian://, http://, webservice://, thrift://, memcached://, redis://等), 供展示层调用。

5) 展示层: 利用HTML5、手机APP、统一门户、单点登录等技术完成和用户的交换, 支持通过不同终端使用平台的各类服务。

2.4 标准规范体系建设

针对交通数据以及相关业务的复杂多样性的特点, 需要建立一套标准规范体系来指导开发过程, 保证项目的先进行、规范性和持续性, 标准规范体系建设包含以下四个方面: 信息服务标准, 元数据标注, 数据源标准和基础数据资源库体系。模型每一层均按照一套标准规范体系来实施建设, 指导整个开发过程顺利进行。

1) 基础数据资源库体系: 用于划分数据来源, 将通过设备采集到的数据或与其它平台交换得到的数据进行归类、划分, 为建立不同的数据传输、交换协议提供基础。

2) 数据源标准: 用于数据引擎处理不同数据源, 将数据源传输过来的数据进行标准化处理, 根据从各数据源得到的数据对象之间的关联关系, 在本系统中将各类数据加以关联整理, 为后续的数据库设计, 数据存储和数据挖掘提供支持。

3) 元数据标准: 将得到所有数据按照人、机、物、法、环的概念分类^[11], 统一制定一组描述数据的元数据标准, 保证数据的规范性和持续性。

4) 信息服务标准: 主要为对外数据服务提供支持, 按客户的角色不同、等级权限等标准为客户提供不同等级的数据服务, 保证数据的安全性。

3 系统应用架构

3.1 系统应用架构分析

本信息服务平台满足城市智能交通信息采集、传输、交换、清洗、存储以及大数据分析处理、资源管理与调度、平台监控与运维、安全保障等方面的要求, 系统应用架构包括以下几个部分: 通用组件服务/数据服务, 数据处理, 云管理平台, 运维管理, 安全管理以及其他应用。系统应用架构如图2所示。

3.2 应用架构模块分析

1) 通用组件服务/数据服务: 提供智能交通应用通用的软件组件服务, 包括交通流量常规统计与模糊分

析、车辆分类识别与告警、实时交通信息发布与诱导、区域交通流量分析与告警.在通用服务组件基础之上,可以根据服务对象的等级、权限不同,为用户提供不同的服务内容、服务质量和信息安全等级的数据服务,数据服务将分析结果按照信息服务标准封装成一定形式提供给外部客户使用.这些客户按照等级权限可以区分为公权和私权用户,包括前述的与本平台有密切业务合作的公安,环保,市政,银行以及商家等客户.

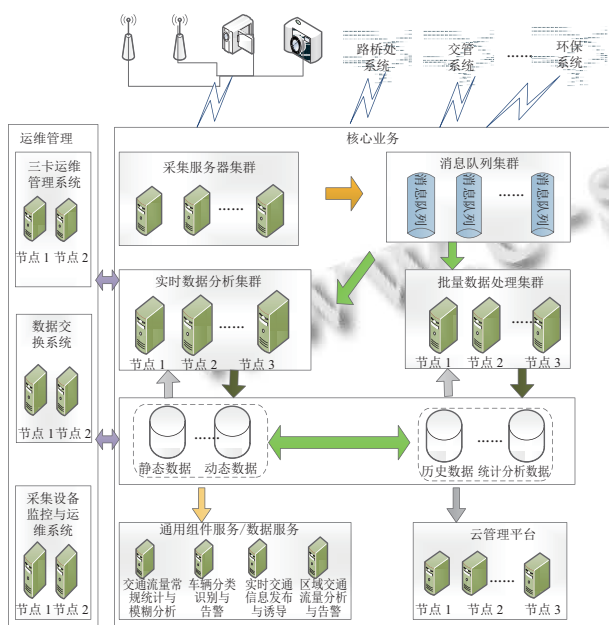


图2 系统总体框架图

2) 数据处理: 包括了数据采集系统、消息队列、实时处理系统、批量处理系统和数据存储与管理系统等几个部分.

① 数据采集系统完成对数据的接收、重整、清洗等操作, 收集所有采集点的采集信息传递给消息队列.

② 消息队列负责接收采集系统传来的数据, 根据处理需求分发给不同的系统, 支持发布订阅模式, 支持集群的弹性扩展.

③ 实时处理系统主要面向实时性要求较高的业务(T+0), 对采集到的数据进行即时的处理, 满足实时交通应用需求.

④ 批量处理系统主要面向非实时/准实时的分析型业务(T+1), 对采集到的数据进行分析挖掘, 满足分析型交通应用需求.

⑤ 数据存储和管理系统主要对采集的数据和分析挖掘的数据进行存储和管理, 完成数据进行持久化和

访问管理等功能.

3) 云管理平台: 对整个支撑服务平台相关IT资源的管理和调度, 为平台其他部分提供资源服务. 主要完成服务器、存储和网络设备的整合与服务提供, 通过服务的形式把资源提供给不同业务系统使用.

4) 运维管理: 主要完成对IT设备的监控与运维、采集设备的监控与运维、电子牌运行状态统计与分析等子系统, 结合相关的人员和运维制度, 建成统一的、多维度的、涵盖业务的运维保障系统. 提高平台的可用性和运维效率.

5) 安全管理: 安全管理要建立一套符合等保三的安全体系, 涵盖安全的各个方面, 当前阶段要重点考虑云平台的安全和数据的安全管理.

6) 其他: 数据交换系统主要完成平台与外部数据源之间的数据同步与交换, 保证各方在数据方面保持一致, 便于系统协作和信息管理.

4 数据中心网络设计

4.1 网络设计图

数据中心网络设计遵循以下五点基本原则: 1) 实用性; 2) 安全性; 3) 先进性; 4) 适用性; 5) 可靠性^[12]. 在数据中心网络结构设计中, 采用了区域化的结构设计. 区域化的结构设计所建设的网络具有良好的扩充性并且便于区域性管理, 而且后期若要变更子网模块设计时, 在不破坏现有的网络情况下, 新的子网模块和新的网络技术能被更容易集成进整个系统中. 该设计将整个系统按应用层次的不同分成不同的区域, 包括核心交换区、服务器区、运行维护区、广域网接入区等. 各个区域均通过接入交换机与核心交换机冗余连接, 实现清晰的系统边界. 数据中心网络设计图, 如图3所示.

4.2 网络设计分析

网络结构设计上分为三个层次来设计: 核心层、汇聚层和接入层.

1) 核心层设计: 核心层交换模块承担着互联互通各个组件的任务, 所有的内部模块间的通信都通过该模块进行, 所以核心交换模块需要功能足够强大的网络交换机来支撑, 且具备很高的背板交换能力、端口密度、设备可靠性. 采用的网络拓扑为双核心双归属拓扑结构, 用两台核心交换机互为备份, 其它所有业务区域模块都通过双归的方式接入双核心. 作为核心层采用万兆光纤模块链接各功能区域的汇聚交换机, 核心层之间采用冗余链路设计, 通过万兆捆绑.

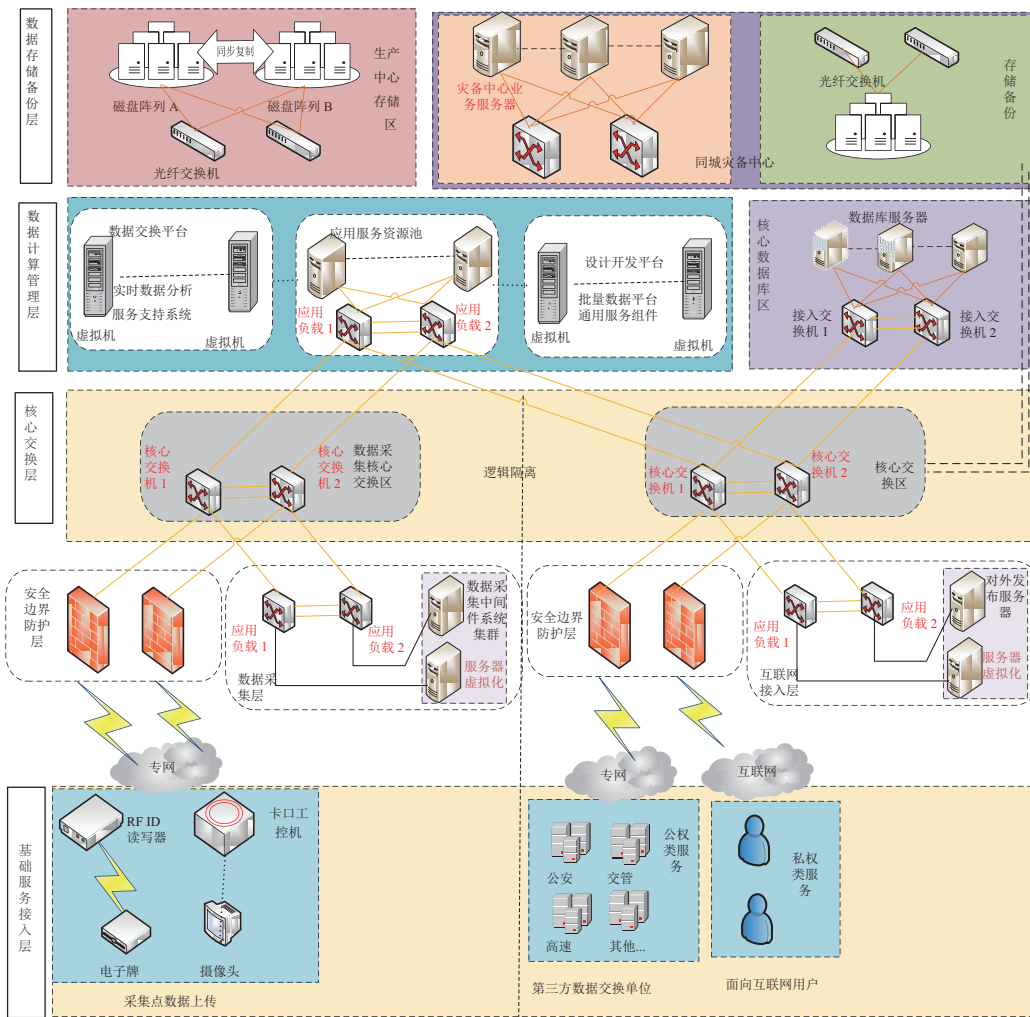


图3 数据中心网络图

双核心双归属拓扑结构具有以下优势^[13]:

① 网络拓扑结构完全对称,核心交换机与汇聚交换机之间存在两条等值负载分担的链路,既保证了系统的高带宽,同时也提高了系统的可靠性.

② 网络设备少,占用机房空间,耗电小,管理维护简单,成本低.

③ 对称结构避免了网络的单点故障,任何一个核心节点失效,都不会影响系统的正常运行,系统的可靠性高.

2) 汇聚层设计:每个功能区启用两台汇聚交换机分别与核心层连接.接入交换机采用堆叠方式,通过千兆链路上行到汇聚交换机,汇聚交换机通过万兆光纤链接核心层,每个功能区的汇聚交换机之间采用捆绑冗余链路设计.

3) 接入层设计:接入采用支持堆叠、可扩展万兆

的独立式三层交换机,与汇聚交换机采用千兆连接,实现信息点的最终接入.通过交换机堆叠技术,可以多台千兆交换机构成一个堆叠组并可配置万兆上行口、千兆备份接口实现与核心交换机之间的双归属连接,完全对称链路连接状况,通过启用三层协议,方便实现基于链路开销(cost)的负载均衡.

5 平台测试结果与分析

5.1 测试环境说明

本平台所有的测试结果均在满足下测试环境的条件下所得,测试环境如表1所示.

5.1 测试环境说明

下面对几个主要功能点的实现情况作如下说明:通过指定采集方向和车流量统计时间起始点进行

采集方向车流量统计查询, 查询结果如图4所示。

表1 测试环境

项目	类型
测试工具	ETC阅读器和ETC设备
起止时间	2015.11.10~2015.11.25, 总计历时15天
硬件环境	ETC阅读器、ETC设备、Linux服务器
软件环境	JDK、Oracle数据库、HBase数据等
测试约束	1. 为保证接近系统真实使用环境, 测试数据均为真实数据抹去个人信息 2. 全部测试都基于相同的测试环境(包括硬件配置、操作系统、软件版本); 3. 测试数据全部由相同ETC设备测试程序生成; 4. 日

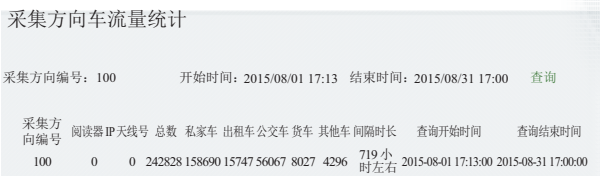


图4 采集方向车流量统计

通过指定大桥编号和查询起始时间, 以小时为单位为展现统计结果, 查询结果如图5所示。

按车辆档案号轨迹分析主要按照车辆档案号进行模糊轨迹查询, 如图6所示。

黄标车违章告警主要按照业务规则, 对于需要进

行告警的黄标车辆进行分类识别, 对于符合业务规则的轨迹数据进行告警记录并转发至指定目标地址, 查询结果如图7所示。

通过以上对主要的功能项的测试可以看出, 本平台已经满足用户前期提出的基本需求, 并且能支持50个拥堵路段车流及路段平均车速的并发实时分析, 响应时间不超过5秒, 且能在6:00-23:00期间持续运行; 支持150个并发车辆的历史跟踪查询, 单车轨迹查询响应时间不能查过15秒; RFID历史数据查询: 一周内1-2秒, 一个月内5-10秒, 3个月内10-30秒内; 肇事逃逸等特定车辆追溯查证: 一周内5-8秒, 一个月内10-20秒, 3个月内20-40秒内。

6 结论

本文给出了在交通大数据背景下, 智能信息服务平台的总体方案, 重点介绍了系统的总体架构, 系统的应用架构和系统数据中心的设计, 从不同角度描述了系统所采用的设计思想. 该信息服务平台已在某地区智能交通项目中投入测试, 测试表明, 该平台满足用户前期提出的基本需求, 在技术方面充分证明了其可行性. 在后期的平台建设方面还要结合用户的具体需求做进一步的改进和完善, 该平台的设计对于发掘交通大数据在智能交通信息服务领域的研发具有一定的参考价值 and 借鉴意义。

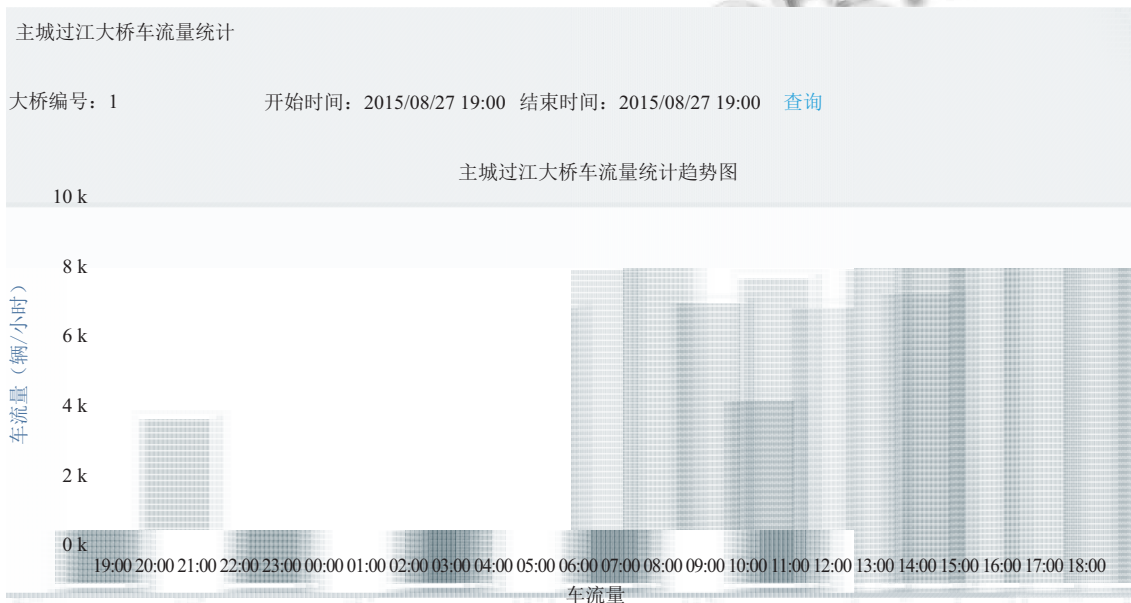


图5 主城过江大桥车流量统计



图6 按照车辆档案号进行模糊轨迹查询

黄标车违章告警			
开始时间: 2015-08-25 20:01:35 结束时间: 2015-08-31 20:01:46 查询			
车牌号码	车辆类型	报警信息	报警产生时间
渝 Y28887	黄标车	经过采集方向.采集编号: 606	2015-08-31 09:58:32
渝 T25539	黄标车	经过采集方向.采集编号: 389	2015-08-31 09:59:23
渝 Q47471	黄标车	经过采集方向.采集编号: 987	2015-08-31 09:58:47
渝 M95334	黄标车	经过采集方向.采集编号: 191	2015-08-31 09:52:54
渝 L27341	黄标车	经过采集方向.采集编号: 568	2015-08-31 09:55:00
渝 K11734	黄标车	经过采集方向.采集编号: 137	2015-08-31 09:53:45
渝 E56676	黄标车	经过采集方向.采集编号: 542	2015-08-31 09:53:30

图7 黄标车违章告警

参考文献

- 杨春华. 大数据在中国B2C电子商务精准营销中的应用研究——以京东商城为例. 中小企业管理与科技(下旬刊), 2016, (3): 104-107.
- 罗洁, 张彩霞. 中国B2C电子商务快速发展原因浅探——以天猫商城为例. 市场论坛, 2016, (4): 47-51.
- Cao S, Manrai AK. Big data in marketing & retailing. Journal of International & Interdisciplinary Business Research, 2014, (1): 23-42.
- 中华人民共和国国务院. 国家中长期科学和技术发展规划纲要(2006-2020). 北京: 中华人民共和国国务院, 2006.
- Xia YJ, Chen JL, Lu XD, et al. Big traffic data processing framework for intelligent monitoring and recording systems. Neurocomputing, 2016, 181: 139-146. [doi: 10.1016/j.neucom.2015.07.140]
- Shi Q, Abdel-Aty M. Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. Transportation Research Part C: Emerging Technologies, 2015, (58): 380-394. [doi: 10.1016/j.trc.2015.02.022]
- Tao S, Corcoran J, Mateo-Babiano I, et al. Exploring bus rapid transit passenger travel behaviour using big data. Applied Geography, 2014, (53): 90-104. [doi: 10.1016/j.apgeog.2014.06.008]
- Liu CL. Intelligent transportation based on the internet of things. 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet). Yichang, China. 2012. 360-362.
- 赵文婧, 王刚. 基于RFID的电子车牌管理系统. 微计算机信息, 2010, 26(29): 150-152, 193. [doi: 10.3969/j.issn.2095-6835.2010.29.059]
- 王兴文, 黄础章. RFID技术在智能交通中的大规模应用模式分析. 中国电子商情(RFID技术与应用), 2009, (1): 20-24.
- 易光旺. 智能建筑安全防范系统的评价指标体系研究. 中国安全生产科学技术, 2010, 6(3): 147-151.
- 张天祖. 基于RFID技术的物联网智能交通系统开发研究. 兰州交通大学学报, 2012, 31(4): 112-116.
- 解书凯. 基于NGN的企业通信网络平台设计[硕士学位论文]. 成都: 电子科技大学, 2012.