

基于天地图的朔黄铁路通信运维数据分析系统设计与实现

胡跃华

(朔黄铁路肃宁分公司,河北 肃宁 062350)

摘要:朔黄铁路公司为保障本公司 LTE-R 网络的顺利运行,需要按计划对 LTE 网络进行维护,并对维护过程中收集的数据进行分析,以排查网络中可能出现的问题。目前,数据分析大多通过人工的方式来进行,效率低下且很难获取故障发生位置的地理及基站分布情况。同时,由于数据总量巨大,传统的数据存储及处理方法很难满足数据的存储及分析需求。针对这种情况,研究并实现了一个基于天地图的通讯运维大数据分析系统。实际应用表明,该系统能够直观、形象地展示出铁路沿线通信网络的整体情况,并能对常见的通信故障进行排查,大大提升了朔黄铁路公司通信运维的效率,具有较高的实用价值。

关键词:天地图;Cloudera;impala;LTE-R;大数据;通信故障

中图分类号:TP391 **文献标志码:**A **文章编号:**2095-0373(2019)01-0098-07

0 引言

朔黄铁路公司开创了将 LTE 宽带移动通信业务应用到铁路中的先河,不仅在朔黄铁路公司尚属首次,在世界范围内也属于首例。该技术在铁路通信网络中的应用,可为朔黄铁路的运输及各部门的业务提供更大的带宽、更多的信息、更安全的管理、更便捷的服务、更高速化的信息管理^[1]。在这样的背景下,加快 LTE-R 网络故障的排查速度,提高网络运行维护效率已经成为保障朔黄铁路运输能够顺利进行的趋势。传统的以人工排查为主的网络运行维护方式已经不能够满足这种需求。总体来说,提高朔黄铁路通信运维效率应当解决以下两个主要问题:一是传统的通信运维数据,以及分析结果大多以表格、图片的形式存在,难以获取铁路沿线相关具体信息(如:地理位置信息、基站分布状况、地形地貌等),导致在分析数据时,不能对铁路通信网络的整体状况进行全面的考虑,提高了发现故障,以及排除故障的难度;二是通信运维数据的数据量巨大,如果采用传统的数据存储及处理方法,存在着存储空间不足且处理速度较慢、处理效率低下及用户体验较差等问题。目前,已有许多专家和学者针对通信网络运行维护的问题进行了研究,大多集中在民用通信领域^[2-4];也有一些研究人员将 GIS(Geographic Information System)技术引入网络运行维护中,以提高数据分析的全面性,提升可视化效果^[5-7];还有学者将大数据技术应用到网络运行维护中,以提高系统的数据存储及处理能力^[8-10]。

根据朔黄铁路公司通信运维的实际需求,将 GIS 技术以及大数据技术综合起来,设计并实现了一个基于 GIS 的通讯运维大数据分析系统。

1 相关技术

1.1 天地图

天地图即国家地理信息公共服务平台,是由国家测绘地理信息局监制的地理信息综合服务平台,是

收稿日期:2017-04-21 网络出版日期:2019-06-26 责任编辑:车轩玉 DOI:10.13319/j.cnki.sjztdxxbzb.20170089

网络出版地址: <http://kns.cnki.net/kcms/detail/13.1042.n.20170626.1608.002.html>

作者简介:胡跃华(1963—),男,高级工程师,主要从事轨道交通信息化的研究。E-mail:16112084@bjtu.edu.cn

胡跃华.基于天地图的朔黄铁路通信运维数据分析系统设计与实现[J].石家庄铁道大学学报:自然科学版,2019,32(1):98-103.

数字中国的重要组成部分,是运行于互联网环境的国家地理信息公众服务平台,通过门户网站向公众提供权威、可信、统一的公益性地图信息服务,通过二次开发接口提供地理信息服务、资源的增值应用开发环境^[11]。目前市面上开放的地理信息系统主要有:Google Map、Bing map、百度地图等。与这些平台相比,天地图除了能够提供相对简洁、易懂的二次开发接口之外,还具有一个重要的优势:天地图采用的坐标系统为 CGCS2000,同朔黄通信运维数据中所采用的 WGS84 坐标系统十分相似,两个坐标系之间的偏差几乎可以忽略,利用天地图进行开发能够提升系统与数据的兼容性,以及展示结果的准确性。

1.2 Cloudera 大数据平台

Cloudera 是一个企业级的综合数据管理平台,用户可以通过 Cloudera 中的工具快速、便捷地实现对 Hadoop 以及相关大数据组件的安装和部署,同时,Cloudera 也具有较高的数据安全性^[12]。其中,Cloudera Impala 提供了高效、便捷的大数据查询服务^[13]。不同于 Apache Hive,Impala 完全抛弃了 MapReduce,省掉了 MapReduce 作业启动的开销,也省去了把中间结果写入磁盘的步骤,在实时交互式查询方面更加快捷,极大地提升了数据 I/O 的速度。Impala 与 Hive 等的关系如图 1 所示。

2 系统总体设计

2.1 总体架构

根据朔黄铁路通信运维的需求,本文提出了一个基于 GIS 的朔黄铁路通信运维大数据分析系统。为方便通信运维人员使用,将系统设计为 B/S 架构,不仅能够省去安装部署的工作,同时也能支持通信运维人员在多种终端进行使用。其中,服务端主要由 Cloudera 搭建的大数据集群提供支持,主要负责数据的存储及处理;浏览器端主要利用天地图及 HTML5 等进行可视化展示。系统的总体架构如图 2 所示。

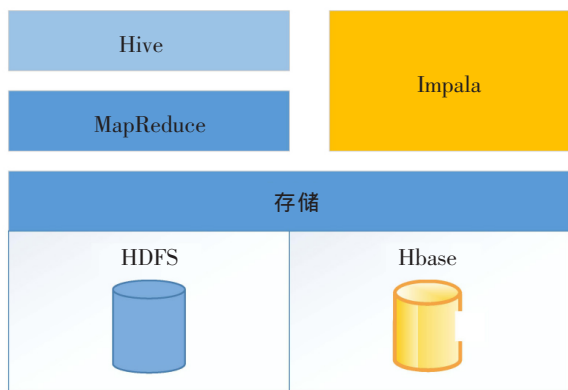


图 1 Impala 与 Hive 关系图

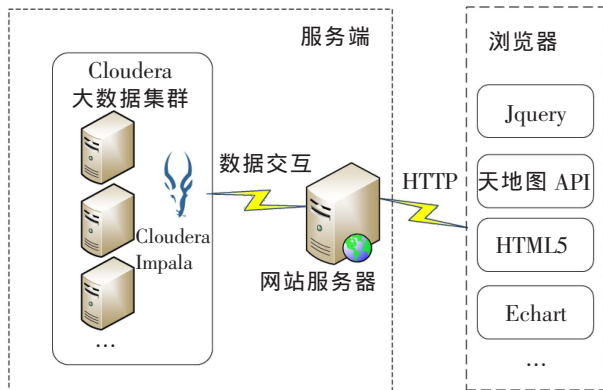


图 2 系统的总体架构图

2.2 系统功能设计

系统应当包括以下几个主要功能:①数据 ETL(Extraction-Transformation-Loading)。主要功能是对原始数据进行预处理,包括删除无效值(如 NULL 值、空值等)、提取通信事件(如 PING 事件、Attach 事件、Switch 事件等);将数据文件类型转化为 CSV 文件,上传到 HDFS,并通过 Impala 把这些数据链接到数据表中。②数据分析。数据分析指的是能够对常见的通信故障(如越区覆盖、弱覆盖、乒乓切换等)进行建模,并根据模型从数据中筛选出常见的通信故障。③可视化展示。利用天地图将分析结果和地理位置信息以及铁路沿线基站信息进行集中展示,将存在问题的区域以直观、形象的方式展示出来,为通信故障的排查提供辅助决策。

3 关键功能实现

3.1 基于 Cloudera 的数据 ETL

数据 ETL 指的是数据的提取、转化和加载。朔黄铁路公司通信运维数据的原始数据大多以 XLS 和

CSV 的形式存在,且其中存在着 NULL 形式的无效值,通信事件如:Ping 事件、Attach 事件、Switch 事件等则以原始信令的形式存在,需要将通信事件发生的时间,持续时间,事件结果等内容从原始数据中提取出来。原始数据如图 3 所示。

	A	B	C	D	E	F	G	H	I	J
1	Computer	Handset	Longitude	Latitude	Event	Event Detail	Direction	Message	Message Detail	LTE Work L
2	35:14.1	35:14.1						NULL		
3	35:14.1	35:14.1	117.0581	38.33036				NULL		NULL
4	35:14.4	35:14.4	117.058	38.33034						NULL
5	35:14.7	35:14.7	117.0579	38.33032						
6	35:15.1	35:15.1	117.0579	38.33031						
7	35:15.1	35:15.1	117.0579	38.33031						
									Message	
									L3Message	
									dir : UPLINK	
									message	
									ATTACH_REQUEST	
									EPS_Attach_Type : (1)EPS Attach	
									NAS_key_set_identifier	
									TSC : (0)native security context	
									(for KSI_ASME)	
									NAS_Key_Set_Identifier : (0)possible	
									values for the NAS key set identifier	
									Old_GUTI_Or_IMSI	
									odd_even_indic : (0)even number of	
									identity digits and also when the GUTI is	
									used	
									Type_Of_Indentity : (6)GUTI	
									MCC : 460	
									MNC : 30	

图 3 原始数据

基于 Cloudera 的数据 ETL 主要分为以下几个步骤:步骤一,将数据中存在的无效值“NULL”替换为空字符串;步骤二,从原始数据中提取关键事件,主要记录事件是否成功以及事件的持续时间,以 Attach 事件的提取过程为例,通信事件提取过程的伪码如下:

```

startTime=0;//事件开始时间
endTime=0;//事件结束时间
isSuccess=false;//事件是否成功
isStart=false;//事件是否开始
for(var item in rawData){//遍历原始数据
    if(item.event.contains("LTE Msg Attach Request,LTE DataService Connect Request") && is-
Start=false)//当开始标志位 false,同时看到 Attach 事件开始标志
        isStart=true;//更改开始标志位 true
        startTime= item.currentTime;//记录当前时间为开始时间
        if(item.event.contains("LTE Msg Attach Request,LTE DataService Connect Request") && is-
Start=true)//当开始标志位为 true,同时看到 Attach 事件开始标志
            isSuccess=false;//标记上次事件结果为失败
            duringTime= item.currentTime- startTime;//记录持续时间
            startTime= item.currentTime;
            if(item.event.contains("LTE Msg Attach Complete,LTE DataService Connect Success") && is-
Start=true)//当开始标志位为 true,同时看到 Attach 事件结束标志
                isSuccess=true;//标记上次事件结果为成功
                duringTime= item.currentTime- startTime//记录持续时间
            .....}

```

其它事件如:Ping 事件、Switch 事件提取方法和 Attach 事件相同;步骤三,将原始数据统一转化为 CSV 格式,上传到 HDFS,并通过 Impala 的 load 函数,将数据添加到数据表中。

3.2 常见通信故障分析与建模

本功能模块的主要功能是,针对朔黄铁路通信运维过程中经常发现的:“越区覆盖”、“弱覆盖”、“乒乓切换”,这 3 种常见故障进行建模,并根据模型对原始数据进行分析,从而快速准确地从数据中排查出此类故障,3 种故障模型分别建立如下。

(1)弱覆盖。弱覆盖的判别方法较为简单,对于每一条数据记录,设定门限为 Threshold,该记录的 RSRP 值为 Valuer,当 $Valuer \leq Threshold$ 时,则判断该记录出现了弱覆盖现象。

(2)越区覆盖。越区覆盖是指基站小区信号越过了本小区的覆盖范围,越区覆盖到临区,给临区带来严重的干扰,引起掉话、频繁切换等^[14]。

从数据中判断越区干扰需要通过两个参数进行判断:① eNodeBID,即 Evolved Node B,LTE 中基站的名称。② PCI(Physical Cell Identity),物理小区标识。具体判断步骤如下:步骤一,将当前数据点的两个参数与基站信息表进行关联,获取当前连接的基站名称、经纬度地址等信息;步骤二,根据当前数据点的经纬度信息,从基站信息表中获取相邻的两个基站的名称;步骤三,如果从相邻两个基站中不包含当前所连接的基站,则判定在当前数据点出现了越区覆盖。

(3)乒乓切换。乒乓切换指的是在进行通信时,切换到目的小区后驻留时间过短又立即切回了源小区^[15]。乒乓切换会产生不必要的开销,增加了切换失败的概率,同时也会对吞吐量产生影响。判断乒乓切换是否发生的具体方法如下:筛选数据中所有的切换事件(handover),记录其中切换成功的事件,假设 handover 事件列表为 listh,单个 handover 事件为一个三元组 $recordh = (currentTime, PCIpre, PCInext)$,设定判定乒乓切换发生的最小时间 $minTime = 1min$,乒乓切换发生的模式为“ABA”,筛选乒乓切换事件的伪码如下:

```
For(i=0;i<listh.length;i++){
    If(listh[i+1].currentTime-listh[i].currentTime<60)//两次 handover 事件相差 1 分钟以内
    If(listh[i+1].PCInext==listh[i].PCIpre){
        //判定乒乓切换事件,记录事件发生的时间、经纬度以及 PCI 信息
    }
}
```

3.3 基于天地图的可视化展示

本功能主要利用天地图 API,将分析出的通信故障以及能够反映通信状况的关键指标如 RSRP、SINR 以及 RSRQ 等,在天地图上展示出来,帮助运维人员快速定位故障发生位置,以及确定故障发生地点的基站和地理情况。为了实现以上功能需求,基于天地图的可视化展示主要应当解决两个问题:①如何利用天地图绘制铁路沿线通信状况以及基站信息。②由于通信运维数据量巨大,而浏览器的渲染能力又十分有限,造成在展示时页面响应速度缓慢,需要考虑一种方法在保证展示效果及准确性的同时提升页面的响应速度。

针对问题一,研究了天地图接口 API,利用天地图 API 提供的 TCircle 类来绘制圆点,不同颜色的圆点表示数据值的大小,利用天地图 API 绘制不同颜色的点的代码如下:

```
var config = {
    strokeColor : "red", // 圆边线颜色(红)
    fillColor : "red", // 填充颜色(根据该点表示的信号值的大小不同)。
    strokeStyle : "solid" // 圆边线线的样式,solid 或 dashed
};
var circle = new TCircle(new TLatLng(lon, lat), radius, config);//创建一个点对象,并指定点的位置、半径和配置信息
map.addOverlay(circle);//在地图上添加点对象
```

针对问题二,采取的解决方案如下:首先在服务端利用 Cloudera impala 读取 HDFS 中的数据,大大提升了数据读取的速度。其次,在浏览器端采用“分层展示”的策略,当地图所属的层级较高时,无需过

多展示数据的细节,所以,当需要绘制大量的数据时,根据层级的不同,对数据进行插值,层级越高,所需展示的数据则越少。同时,层级不同时,为保证所绘制图形的大小不变,需要实时对绘制图形的大小进行改变,当绘制的图形为圆形时,半径 $radius$ 与层级 $zoom$ 满足如下关系

$$radius = 3 \times 2^{18-zoom} \quad (1)$$

最后,由于查询数据库的速度较慢,在实现策略时应当尽可能减少查询数据库的次数。考虑到在地图 API 中,鼠标滚轮可以触发地图层级的变化,数据库的查询操作应当在鼠标滚轮完全停止时进行。本文的策略是记录最后一次鼠标滚动事件与当前时间的时间差,当时间差大于一定程度时,判断当前地图层级,与上一次发生数据更新时的地图层级是否相同。如果是,更新数据、重绘地图并记录当前层级;反之,则不作任何操作。数据更新策略的算法流程如图 4 所示。

在解决以上问题的基础上,KPI 关键指标和通信故障便能够在地图上展示出来,具体思路如下:当展示 KPI 指标及通信事件时,根据 KPI 指标值的大小或者事件时延的大小来绘制不同颜色的点,根据前述方法将这些点绘制在地图上,即可

对铁路沿线通信情况进行展示。当展示通信故障时,首先根据所展示通信故障的不同,利用 3.2 所建立故障模型对原始数据进行筛选,然后根据故障种类的不同,选择有关的关键指标进行展示。例如,当展示弱覆盖问题时,选择 RSRP 指数进行展示,将 RSRP 值小于阈值的数据点挑选出来,根据 RSRP 值的大小,在地图上绘制不同颜色的点进行展示;当展示乒乓切换的点时,则主要关注 PCI 值,需要根据 PCI 值的不同绘制不同颜色的点;当展示越区覆盖时,主要关注的则是数据点与基站之间的连接关系,首先绘制出所有沿线的基站,然后将发生越区覆盖的数据点同实际连接的基站用连线表示出来,便可直观地展示出越区覆盖问题。

4 应用效果

4.1 应用环境

设计的系统在朔黄铁路公司进行了应用,部署环境为 7 台浪潮英信 NX5840 的刀片服务器,CPU 型号 Xeon E5-2620 v2,内存容量为 8 GB,操作系统为 Centos6.5。其中一台作为网站服务器,网站发布环境为 Tomcat8.0 和 jdk1.8.0_91。

4.2 实现效果

为测试系统的实现效果,导入朔黄铁路实际网络运维数据进行测试,数据反映的是肃宁北到黄骅港区段的通信状况。实现效果如图 5 所示。

图 5(a)是对铁路沿线通信 RSRP 指数的分析和展示,通过不同的颜色来指示不同的 RSRP 值;图 5(b)通过对原始数据的分析,展示出现乒乓切换问题的地区,可以看到对问题发生的地点和基站信息都能够从图中方便地获取;图 5(c)展示的是出现越区覆盖问题的地点和基站信息,可以看到图中基站的覆盖区域显然超过了正常范围;图 5(d)是对原始数据中 Ping 事件发生情况进行分析,并利用不同颜色的点展示出在铁路沿线进行 Ping 事件时的时延情况。通过实现效果可以看出,本系统能够直观、形象地展现出铁路沿线的通信状况,并迅速地从数据中排查出存在的通信故障。

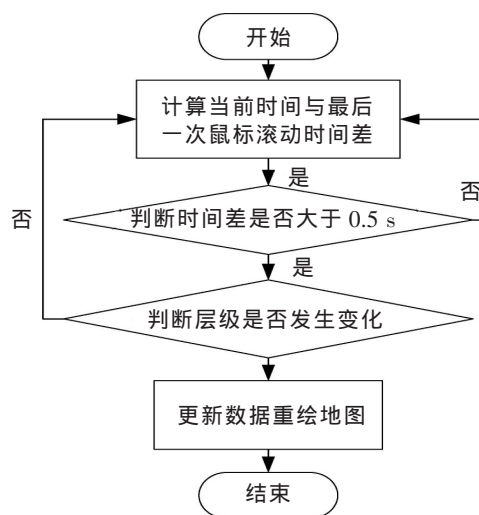


图 4 数据更新策略流程图

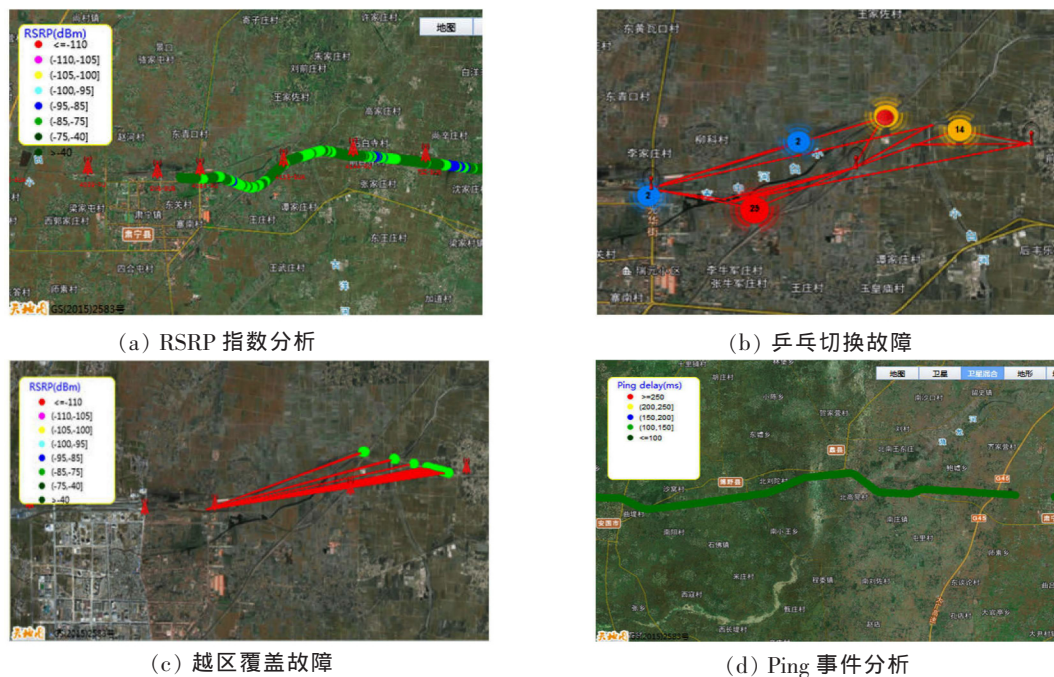


图 5 系统实现效果

5 结论

针对朔黄铁路通信运维的具体需求,设计并实现了一个基于 GIS 以及大数据技术的朔黄铁路通信运维大数据分析系统。利用 Cloudera 大数据平台,解决了大规模通信运维数据存储及处理的问题;通过数据 ETL 过程将原始通信运维数据进行清洗,并将通信事件如 Ping 事件、Attach 事件、Switch 事件等提取出来;对通信故障进行建模,使得本系统能够将常见的通信故障如:越区覆盖、乒乓切换、弱覆盖等从数据中筛选出来,大大节约了人工的开支。天地图 API 的使用,不仅能够将铁路沿线的通信状况直观、形象地展示出来,并能够将周边地理信息以及基站信息等一并展示给用户,为故障排查和检修提供了便利,大大提升了朔黄铁路公司通信运维的效率,具有较高的实用价值。今后的工作将主要集中在丰富故障模型,提供常见问题的维修建议,真正实现整个通讯网络的信息化管理。

参 考 文 献

- [1]李鹏. LTE-R 技术应用研究[J]. 自动化与仪器仪表,2014(3):11-12.
- [2]王艺旋,邢友松,刘娟. 基于 SG-TMS 的市县一体化通信运维管理模式[J]. 信息通信,2016(11):217-218.
- [3]周筠. 电力通信移动运维系统研究及应用[D]. 镇江:江苏大学,2016.
- [4]黄振. 电力通信运维管理系统设计与实现[D]. 成都:电子科技大学,2013.
- [5]张士鹏. 基于 GIS 的 GSM 无线网络优化系统的设计和实现[D]. 北京:北京邮电大学,2009.
- [6]王兴举,刘佳,周杨,等. 基于 GIS 铁路网络运营管理和分析系统[J]. 石家庄铁道大学学报:自然科学版,2015,28(3):99-104.
- [7]金龙,王越,商霖,等. 基于 SuperMap IS.NET 的河北省铁路地理信息数据管理系统设计与实现[J]. 石家庄铁道大学学报:自然科学版,2016,29(2):56-60.
- [8]余海波. 大数据在电信移动通信网络优化中的应用[J]. 广西通信技术,2014(4):8-11.
- [9]张国栋. 基于 Hadoop 技术的电信大数据分析平台的设计和实现[D]. 上海:上海交通大学,2014.
- [10]徐峰. 一种基于大数据分析的网络优化方案[J]. 江苏通信,2014(4):38-39.
- [11]魏金明,邵飞,仲伟政. 基于天地图的地图服务方法初探[J]. 测绘通报,2014(S2):265-268.
- [12]NetApp 携 Cloudera 提供卓越的 Hadoop 解决方案[J]. 硅谷,2011(22):73.

(下转第 120 页)

Preparation of Boron Nitride Nanosheets / g-C₃N₄ Composite Photocatalysts for Rhodamine B Degradation

Wang Yijin^{1,2}, Zhang Chenxu^{1,2}, Zhang Han^{1,2}, Cao Zuolin^{1,2}, Wu Xiangfeng^{1,2}

(1. School of Materials Science and Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050043, China;

2. Hebei Provincial Key Laboratory of Traffic Engineering Materials, Shijiazhuang 050043, China)

Abstract: The boron nitride nanosheets BNNS/g-C₃N₄ composite photocatalysts was fabricated through a calcination method in this work. The effects of BNNS content on the structure, morphology, uv/visible optical property and electrochemical impedance properties of the as-prepared samples were discussed by using XRD, TEM, UV-vis and EIS, respectively. Rhodamine B solution was used to evaluate the degradation efficiency of the as-prepared samples under the visible light irradiation. Experimental results showed that the similar heterojunction, which was in favor of improving the separation efficiency of raw carriers, had been obtained by loading BNNS on g-C₃N₄. Moreover, with the increase of BNNS content, the degradation efficiency was firstly increased and then decreased. When the content of BNNS was 8%, it showed the highest degradation efficiency of 73.34%, which was increased by 32.05% compared to pure g-C₃N₄, and 40.47% compared to P25 (titanium dioxide, as a contrast sample), respectively. In addition, a possible synergistic degradation mechanism of the as-prepared composite photocatalysts was also proposed.

Key words: g-C₃N₄; boron nitride nanosheets; photocatalysis; Rhodamine B

(上接第 103 页)

[13] 郭超, 刘波, 林伟伟. 基于 Impala 的大数据查询分析计算性能研究[J]. 计算机应用研究, 2015(5): 1330-1334.

[14] 张钢. GSM 网络中越区覆盖的分析[J]. 中国无线电, 2005(12): 17-18.

[15] 王敏. 网络优化中乒乓切换的解决方案的研究[D]. 长春: 吉林大学, 2014.

Design and Implementation of Data Analysis System for Communication Operation and Maintenance of Shuohuang Railway Based on Tianditu

Hu Yuehua

(Shuohuang Railway Company Suning Branch, Suning 062350, China)

Abstract: To ensure the trouble-free operation of the company's LTE-R network, the Shuohuang Railway Company needs to maintain the communication situation of LTE network as planned, and analyzes the data collected in the maintenance process to find communication problems. At present, most of the data analysis is carried out manually, which is inefficient and difficult to obtain the geographical and base station distribution of the location where the fault occurred. Moreover, due to the huge amount of data, traditional data storage and processing methods are difficult to meet the demand of data storage and analysis. In view of this situation, this paper puts forward an analysis system of Tianditu communication operation based on big data. The practical application shows that the system can intuitively and vividly demonstrate the overall situation of the railway communication network, and carry out the investigation of communication failures. The analysis system greatly enhances the efficiency and communication operation of Shuohuang Railway Company, and has high practical value.

Key words: Tianditu; cloudera; impala; LTE-R; big data; communication failure