

# 基于大数据的智慧区域运营平台 设计与实现

## Design and Implementation of Intelligent Regional Operation Platform Based on Big Data

顾志峰(中国移动通信集团上海有限公司,上海 200060)  
Gu Zhifeng(China Mobile Group Shanghai Co.,Ltd.,Shanghai 200060,China)

### 摘要:

针对运营商现有生产经营系统在区域运营支撑等方面能力不足的问题,提出基于大数据技术建设智慧区域运营平台的方案,从系统功能设计、系统架构及开发实现等方面进行了研究。通过建设智慧区域运营平台,对运营商海量数据进行实时处理和挖掘分析,实现可视化运营、便捷化营销、网格化支撑能力,提升运营商市场运营效率。

### 关键词:

大数据;数据可视化;网格化运营;Geohash  
doi:10.12045/j.issn.1007-3043.2021.03.015  
文章编号:1007-3043(2021)03-0067-04  
中图分类号:TN919  
文献标识码:A  
开放科学(资源服务)标识码(OSID):



### Abstract:

Aiming at the problem that the existing production and operation system of operators is insufficient in regional operation support and other aspects, it puts forward a scheme of building an intelligent regional operation platform based on big data technology, and studies the system function design, system architecture and development implementation. Through the construction of intelligent regional operation platform, operators' massive data is processed and mined in real time, which realizes visual operation, convenient marketing and grid supporting ability, and improves the market operation efficiency of operators.

### Keywords:

Big data; Data visualization; Grid operation; Geohash

**引用格式:**顾志峰. 基于大数据的智慧区域运营平台设计与实现[J]. 邮电设计技术, 2021(3): 67-70.

## 0 引言

近年来,国内电信行业市场竞争日益激烈。受到流量红利快速释放、同业及跨界竞争不断加剧等经营环境和“提速降费”持续推进等政策性因素的影响,基础电信业务“量收剪刀差”进一步扩大。国内电信运营商正处于转型升级的关键时期,需要企业提升运营效率,实现高质量发展。

随着大数据技术的不断发展和企业大数据意识的不断提升,如何利用大数据推动企业管理模式的创新与运营能力的升级已经成为每个企业都需要思考的问题<sup>[1]</sup>。国内外电信运营商尝试利用大数据技术在市场营销、客户服务、经营决策和网络优化等方面开展应用,并取得积极成效。大数据技术及应用已经成为电信运营商破解经营难题的有效手段之一。

针对公司现有生产经营系统在区域运营支撑等方面能力不足的问题,本文基于大数据技术设计并开发实现智慧区域运营平台,通过数据赋能提升公司运

收稿日期:2021-02-16

营能力,发挥运营商大数据应用价值。

## 1 系统功能设计

### 1.1 总体设计思路

围绕可视化运营、便捷化营销、网格化支撑的建设目标,智慧区域运营平台基于大数据技术对运营商的网络域、业务域等海量数据进行实时处理和挖掘分析,将运营数据关联汇总纳入网格,利用GIS地图以可视化方式,实时全面洞察全市各类区域的特征情况;通过构建丰富应用场景及模型,实现精准营销、智慧选址、业网协同等网格化运营支撑能力,提升公司市场运营效率。

### 1.2 功能设计

#### 1.2.1 区域洞察

基于GIS地图圈划出全市各级行政区划、公司组织以及重点场景的区域,提供区域内客户、业务、网络、资源等多维度的可视化洞察服务。在常规区域方面,支持市、区、街道/镇等行政区划,以及市公司、分公司、区域运营部、网格等组织机构;在场景区域方面,支持5G基站、营业厅、高校、交通枢纽等重点关注区域。

a) 客流统计:基于可视化图表展示选定区域实时客流统计数据,并且按小时、日、月等颗粒度提供历史客流统计数据查询,以便业务人员实时、全面掌握区域客流情况。

b) 区域画像:基于可视化图表展示选定区域人口特征,包括居住人口、工作人口、流动人口、客源地、年龄段、性别等维度;展示区域业务经营指标,包括2G/4G/5G用户数、ARPU、DOU、MOU、网龄、套餐等维度;展示区域客群偏好特征,包括终端、APP使用及兴趣偏好等情况;展示区域内资源类别及数量,包括渠道网点、2G/4G/5G基站、家宽小区等;展示区域内基站流量、话务量等网络运行指标。

c) 客流热力图:基于GIS地图以可视化方式将选定区域的不同网格渲染成不同的颜色表示客流的密集程度,以便业务人员直观地掌握选定区域客流地理分布情况。

d) 客群洞察:针对重点发展业务,基于GIS地图将目标客群以热力图方式展示在地图上,方便业务人员直观地掌握用户地理分布情况。同时,可以通过区域画像交互式图表,点击人口特征、业务指标等细分类别(如居住人口、DOU大于8GB),展示该细分类别

客群的地理分布情况。

e) 资源视图:基于GIS地图以POI打点或多边形区域着色方式展示区域内渠道网点、2G/4G/5G基站、家宽小区等资源的地理位置。

f) 区域圈划:除默认常规区域、场景区域外,为业务人员提供所见即所得的自定义圈划区域功能,支持圆形、多边形等形状,实现快速、灵活支撑各类区域洞察需求。

#### 1.2.2 精准营销

根据营销需求,选定区域或自定义圈划区域,通过自助标签或营销模型筛选出目标客群后,利用系统自有短信、外呼等渠道触点,或其他营销系统部署执行营销活动,实现营业厅引流、地推活动预热以及其他各类场景的精准营销。

#### 1.2.3 智慧选址

面向渠道管理,可视化展示选定营业厅等渠道网点覆盖区域范围、周边客流热力分布、友商网点等相关情况,输出周边客流数量、人口特征、业务指标等量化数据,根据模型给出选址建议,为渠道网点优化与管理提供数据化、可视化支撑。

#### 1.2.4 业网协同

从业务角度,建立弱覆盖、异常核查、用户感知差、高负荷扩容、高倒流优化等模型,基于GIS地图提供选定时段的质差小区清单,精准定位故障片区,并持续跟踪指标改善情况,便于业务人员了解与业务发展相关的网络运行信息,支撑区域流量发展。

#### 1.2.5 区域及任务管理

通过可视化自助界面,业务人员能够查看、管理自建区域、营销任务等信息;具有系统权限的业务管理人员查看、管理全市各级组织的区域、营销任务等信息,对业务人员自建区域、营销任务等进行审批,分派工作任务等。

## 2 系统开发实现

智慧区域运营平台基于Hadoop和Spark框架的大数据平台开发。整个系统主要包括数据采集层、数据处理与分析层、数据服务层、应用层等,系统架构如图1所示<sup>[2]</sup>。

### 2.1 数据采集层

数据采集层主要负责完成各类数据的接入。系统使用的数据源主要包括网络DPI等实时数据,经分系统的用户信息表、渠道网点清单、家宽小区清单和

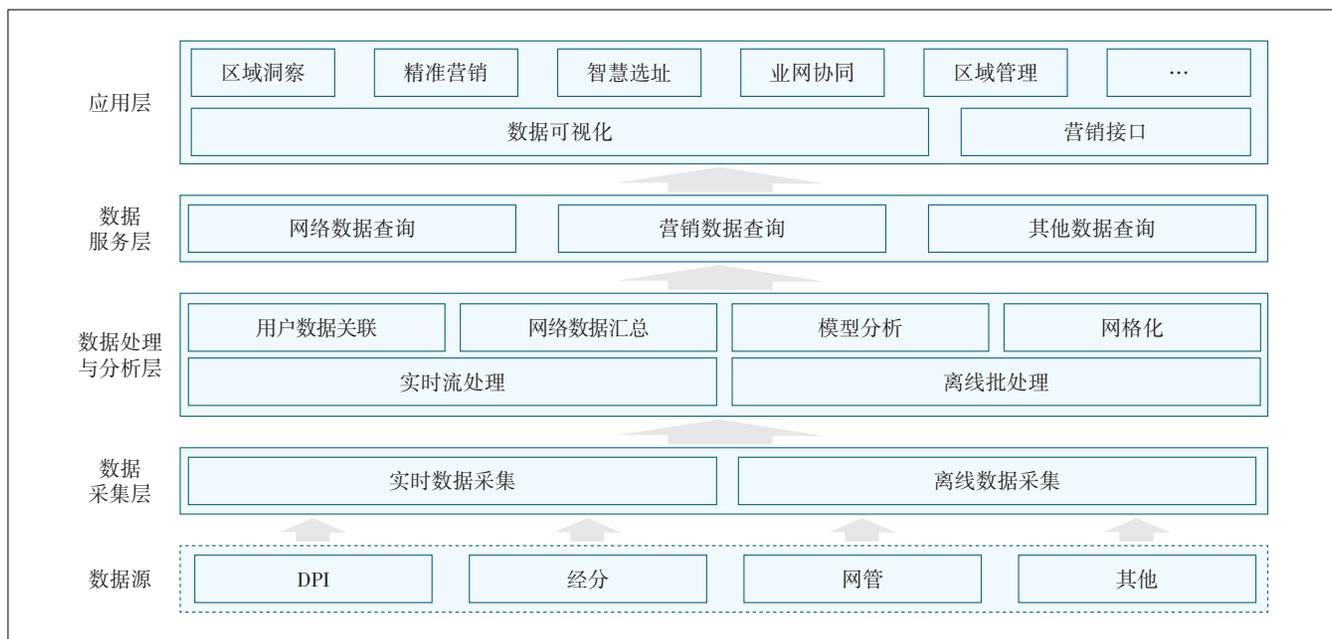


图1 系统架构

网管系统的网络性能报表、基站工参表等离线数据。系统采用Flume日志采集工具从公司的统一聚合平台消费实时数据,写入Kafka分布式消息队列中,供上层应用进行实时流处理或离线批处理。离线数据量相对较少,通过FTP接口机等方式接入,存储到Hadoop的HDFS分布式文件系统。

## 2.2 数据处理与分析层

数据处理与分析层主要负责网格化处理、实时数据关联汇总、离线数据挖掘分析。

在大数据处理前,网格化模块预先采用Geohash算法对基站、渠道网点、家宽小区等的经纬度数据进行编码。Geohash算法首先根据经度 $[-180, 180]$ 、纬度 $[-90, 90]$ 的区间范围,对经度和纬度分别不断进行二分(左区间取0,右区间取1)转换成二进制编码,然后按照偶数位放经度、奇数位放纬度的规则(从第0位开始),将经度和纬度二进制编码按位交替组合,最后根据Base32编码将组合后的二进制编码转换成字符串<sup>[3]</sup>。Geohash把二维的经纬度坐标编码成一维的字符串后,可以提升后续查询效率。Geohash编码字符串越长,精度就越高,对应的网格范围也越小。综合考虑精度和效率,本系统Geohash编码字符串长度取7位,相应网格尺寸为 $153 \times 153$  m。Geohash编码示例如表1所示。经过网格化,全市区域被划分为数万个网格,建立基站、渠道网点、家宽小区等与网格的关系,实现相关资源数据入格。

表1 Geohash编码示例

类型	名称	经纬度	Geohash编码
基站	东方明珠-xxxx	121.495 521, 31.241 662	wtw3sxq
渠道	长寿路旗舰店	121.436 702, 31.245 781	wtw3g25

实时处理部分采用Spark Streaming按设定时间间隔(5 min)对实时流数据进行处理,抽取网络信令XDR中IMSI、TAC、CELL ID、TIME等关键数据字段,与预先写入Redis内存数据库的用户标签(人口特征、位置、兴趣偏好等)、业务指标、基站网格编码等数据进行实时关联,形成中间数据并写入Kafka;其中根据用户所处基站,实现了用户相关数据入格。经过关联处理后,Spark Streaming读取中间数据,按设定的窗口时间长度(2 h)和滑动间隔(5 min),基于网格进行数据汇总统计,并将结果数据存入MySQL数据库。实时数据处理流程如图2所示。

离线处理部分由HDFS、Hive、Spark等组成,对DPI、经分等数据进行离线处理与分析。通过位置模型为用户打上居住地、工作地等位置类标签,根据精准营销、智慧选址、业网协同等应用需求,建立相关模型进行深度数据挖掘分析,其中用户标签数据写入HBase数据库。

## 2.3 数据服务层

数据服务层主要负责提供数据查询服务。根据系统前端的业务人员操作请求,网格数据查询模块按区域与网格关系从MySQL数据库中查询数据,计算出

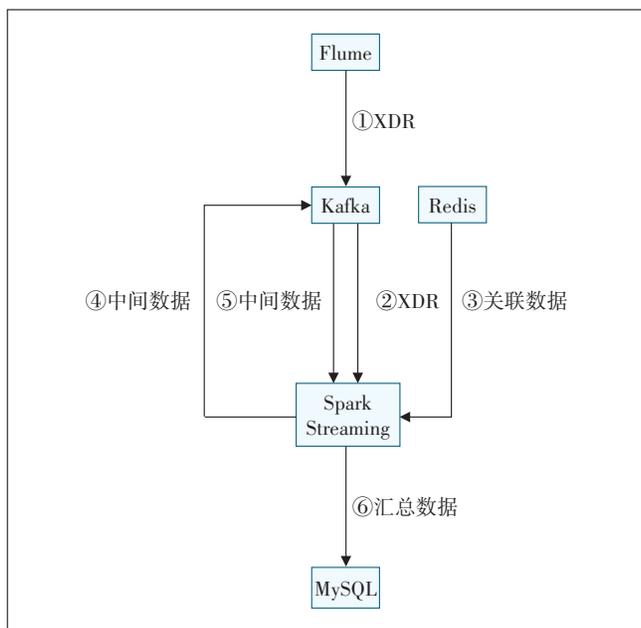


图2 实时数据处理流程

选定区域各类统计数据。营销数据查询模块首先基于 Solr 对 HBase 表中涉及筛选的字段和 RowKey 建立索引,然后通过 Solr 快速查询获得符合筛选条件的 RowKey,实时统计出目标客群人数,最后通过这些 RowKey 在 HBase 的全量用户数据中查询获取目标客群数据<sup>[4]</sup>。

## 2.4 应用层

应用层主要负责提供数据可视化展示及应用功能服务。系统通过 Web 门户为业务人员以可视化交互方式提供区域洞察、精准营销、智慧选址、业网协同、区域管理等应用功能。系统采用 B/S 架构,基于前后端分离的模式进行 Web 应用开发,前端采用 Vue.js 实现用户 UI 界面构建和数据展示,后端采用 Spring Boot 框架实现业务逻辑和数据提供;前端通过 AJAX 向后端接口请求数据,数据以 JSON 格式进行传输;此外,将客群、基站清单等数据按应用或接口要求生成文件。在数据可视化方面,为快速开发实现,GIS 部分采用百度地图 JavaScript API,实现地图底图绘制、热力图渲染、资源 POI 打点以及区域自定义圈划等功能<sup>[5]</sup>;图表部分采用 ECharts 开源可视化图表库,以折线图、柱状图、饼图等丰富形式,实现客流统计、区域画像等洞察数据可视化,系统 Web 界面如图 3 所示。营销接口模块根据接口规范打通短信网关、呼叫中心系统,或调用其他营销系统能力,通过 FTP 推送客群数据文件,执行精准营销活动。

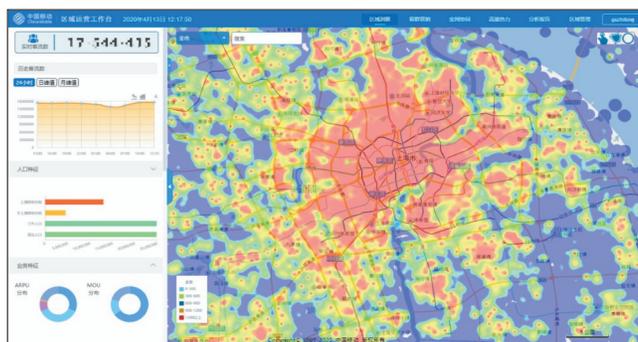


图3 系统 Web 界面

## 3 结束语

本文基于大数据技术研究开发智慧区域运营平台,创新性实现面向区域运营的数据网格化、实时处理、实时查询、数据可视化等关键技术,高效满足洞察分析、精准营销等应用需求。目前智慧区域运营平台已完成开发并在公司推广使用,面向各级运营人员提供多维度实时区域洞察,支撑营业厅引流赋能渠道业务发展、住宅小区位置营销拓展家庭宽带、智慧选址优化网点布局、业网协同助力流量提升等应用场景,达到了可视化运营、便捷化营销、网格化支撑的系统建设目标,取得了良好的经济效益。随着系统功能迭代优化、应用场景不断丰富,智慧区域运营平台将在网格化运营、渠道转型、5G 营销等方面发挥更大的价值。

## 参考文献:

- [1] 郭佳睿,魏进武,张云勇. 大数据助力运营商提升规模化运营核心策略[J]. 电信科学,2018(1).
- [2] 顾志峰. 基于大数据的运营商数据管理平台研究[J]. 电信快报,2020(5):16-19.
- [3] 周艺华,李广辉,杨宇光,等. 基于 GeoHash 的近邻查询位置隐私保护方法[J]. 计算机科学,2019(8).
- [4] 李聪颖,王瑞刚,于金良. 大数据分布式全文检索系统的设计与实现[J]. 计算机与数字工程,2016(12).
- [5] 吴志强,叶锺楠. 基于百度地图热力图的城市空间结构研究——以上海中心城区为例[J]. 城市规划,2016,40(4):33-40.

## 作者简介:

顾志峰,高级工程师,硕士,主要研究方向为大数据技术与应用。

