

侯建民,郭凯,崔满丰,等. 2023. 基于移动大数据的地震人口热力实时获取系统实现与应用——以泸定  $M_s 6.8$  地震为例. 中国地震, 39(3):663~670.

# 基于移动大数据的地震人口热力实时 获取系统实现与应用 ——以泸定 $M_s 6.8$ 地震为例

侯建民<sup>1)</sup> 郭凯<sup>1)</sup> 崔满丰<sup>1)</sup> 方毅<sup>2)</sup>  
董霖<sup>2)</sup> 翟颖<sup>1)</sup> 马秀丹<sup>1)</sup>

1) 中国地震台网中心, 北京 100045

2) 每日互动科技股份有限公司, 杭州 030024

**摘要** 震后第一时间获取震中的人口情况对于开展应急处置决策非常关键。获取海量移动互联网人口分布数据,采用分布式集群进行海量数据的实时处理和网格化,为确保震后系统能够达到秒级产出的效率,设计了多级的多策略抽稀计算并基于 HeatMap 实现可视化展示,震后实现秒级震中周边指定范围的热力人口数据计算和分布图的自动绘制。以泸定  $M_s 6.8$  地震为例进行应用,体现出基于海量移动互联网用户数据获取震中人口分布的优势。

**关键词:** 移动大数据 人口热力图 分布式 泸定地震

[文章编号] 1001-4683(2023)03-0663-08 [中图分类号] P315 [文献标识码] A

## 0 引言

随着地震信息服务技术不断发展革新,面对我国地震多、强度大、分布广、灾害重的基本国情,具备特定区域人口热力洞察监测的能力是智慧应急发展的必然要求,也是应急管理体系现代化的重要组成部分(赵深等,2022)。大数据、云计算等信息技术可为震后应急救援和公众关注热点提供地震相关背景信息。随着大数据分析技术和地理信息技术的发展,目前热力图的发展形成比较成熟的应用产品,如 Google 公司提供输入 shp 数据自动生成热图 API 工具,OpenSignal 网站可以提供全球手机基站信号覆盖热图,微软公司的统计数据访问量热图网站(杨微等,2012)。在国内,人口热力图技术被广泛应用于旅游、交通、市政管理等行业,如腾讯区域热力图、百度景区热力图、基于 LBS 人口数据可视化监测系统(姜保庆等,2016)。

基于手机信令大数据的应急人口热力图系统充分结合手机信令大数据的连续性、无感知性、真实性的特点与优势,将个体信息汇集、建模、计算、处理形成实时动态的人口时空分

[收稿日期] 2023-02-08 [修定日期] 2023-05-08

[项目类别] 地震科技星火计划项目(XH223706YB)、山西省重点研发计划(202202020101009)共同资助

[作者简介] 侯建民,男,1978年生,研究员,主要研究方向为地震信息服务技术。E-mail:houjm@seis.ac.cn

布。在城市应急指挥救援过程中,实时动态的人口时空分布是政府应急救援态势分析、指挥研判的重要依据之一,从而有效提升城市应急指挥救援的效率和科学管理水平(赵深等,2022)。热力图以不同的颜色及亮度动态反映人口活动空间的聚集程度,尽管热力图并非准确的人口分布数据,但其相较于手机信令和公共交通刷卡等运营商数据更易于获取,更能精准表示一定时空范围的人口分布,并在规划设计领域得到了大量的应用探索(张海林,2021)。随着互联网 2.0 的出现,大数据平台使城市动态人口数量及空间分布的获取成为可能(周芳检等,2018)。Murakamia 等(2012)研究了 2011 年日本大地震相关的推特(Twitter)信息,发现在灾害的不同阶段,物资需求的数量会随着受灾人口的变化而存在较大差异。由此可见,地震后震区人口的实时情况对于地震应急救援和政府决策指挥具有重要的参考意义。利用大数据技术绘制人口热力图,为地震震后应急处置提供数据支撑。目前,通过地震信息、大数据技术应用可快速自动实现地震影响范围内的人口数量、分布情况等基本信息,经可视化技术处理后可直观了解震区人口概况(侯建民等,2022)。

人口热力图在地震应急领域得到了进一步应用,基于“互联网+地震”模式,地震信息播报机器人系统具备强大的数据储备能力、空间分析能力和基于海量大数据的网络检索功能,极大地丰富了地震速报服务内容,有效地支撑了地震应急工作(侯建民,2018),能够在震后数秒内自动产出与当前地震相关的数百字及十多张图片,涵盖速报参数、震中位置、周边信息、历史地震等二十余项内容。通过获取地震机器人的地震参数数据,秒级即可实现自动绘制地震人口热力图并展示效果,同时在震后不同时间段,通过人口热力图对比展示震区范围内人口动态情况。在 2017 年 8 月 8 日四川九寨沟 7.0 级地震等多次地震中,地震新媒体机器人自动产出十多项内容,如震区海拔、热力人口图,周边的历史地震、历史地震分布图、村庄、乡镇、县城、天气等信息,5s 便可以实现信息输出,及时提供了震区范围内的基本情况,为地震救援决策提供了有价值的地震信息参考。

强烈地震发生后第一时间获取灾区人口分布情况非常重要,而传统统计普查数据往往具有滞后性,本文研究通过获取海量实时互联网人口位置数据,并采用分布式处理系统,在震后快速获取震中人口分布数据,基于 heatmap 完成人口热力图计算和可视化展示,可以快速评估受灾人群分布,为地震应急救援提供最及时的决策依据。

## 1 数据获取与加工

### 1.1 移动互联数据获取和处理

据工信部权威数据,截至 2022 年 9 月底,中国移动、中国电信和中国联通三家基础电信企业移动电话用户总数达 16.82 亿户,庞大的移动用户群体,使得基于移动互联网的链接数更加准确,用互联网人口代替实际人口不仅准确性高,而且时效性也强,从而解决了如何在地震应急处置工作中的第一时间判定准确受灾人数这一难题。中国地震台网中心通过与百度地图和每日互动两家互联网公司合作,得到了最庞大的移动互联网用户链接大数据。

海量用户数据通过大数据平台进行分布式实时分析处理,采用 Kafka 消息队列进行计算结果的高速传输,采用 Spark 分布式计算技术进行数据清洗、去重,并将全国划分成约 150m×150m 的网格,构建了百米级的全国 3700 多万个热力人口网格数据,以“像素”的方式进行网格区域人口的计算,并将结果存储到分布式数据库 Hbase 中。通常情况下,一个地理

位置由经度和纬度共同表示,这种二维展示方式便于人们理解,但并不适合计算与存储,因为其数据量巨大的计算工作,会对服务器造成很大压力。为解决这一问题,系统通过 GeoHash 编码方式,将二维的空间经纬度数据编码成一维的字符串数据,GeoHash 编码字符串的长度越长,则经纬度细分的区间越小,GeoHash 所表示的定位区域也越小,定位精度也越精确,且对于同一 GeoHash 编码字符串,不同长度的字符串之间存在着包含关系,以编码 wx4g0ec1 为例,wx4g0e 的定位区域包含着 wx4g0ec1,而 wx4g0 所表示的定位区域则包含着 wx4g0e,以此类推,便会极大压缩计算量,为整个技术系统实现秒级产出奠定了基础(图 1)。



图 1 基于 GeoHash 编码实现热力人口的高效管理

## 1.2 基于 HeatMap 可视化展示

人口热力图是一种常用的基本数据可视化技术,通常采用颜色编码数值大小,并以矩阵或方格形式整齐排列,在二维平面或者地图上呈现数据空间分布,被广泛应用于众多领域(姚笛,2017)。HeatMap 是一种常用的热力可视化技术,为确保震后系统能达到秒级产出的效率,本文设计了多级的通过多策略抽稀计算并基于 HeatMap 实现可视化展示,在震后实现秒级震中周边指定范围的热力人口数据计算和分布图的自动绘制(图 2)。

对海量用户数据的网格化进行“像素”级处理后,在震后第一时间获取震中区域的人口分布情况,如果进行直接计算,对系统的计算性能要求非常高,很难在秒级完成计算,因此,按照东、西部不同区域的人口分布密度即“西疏东密”的原则,对西部区域按照约 500m 范围进行数据抽取合并,东部区域按照约 1km×1km 范围进行数据抽取合并。

考虑到实时人口分布数据(小时级更新)可能并不能更稳定表示一定区域范围内的真实人口,实时变化较大,在一定程度上会影响评估震中受灾人口的准确性。我们在实时人口分布接口上,设计了月均人口热力数据接口(图 2),通过实际检验更适用于评估震中区域的人

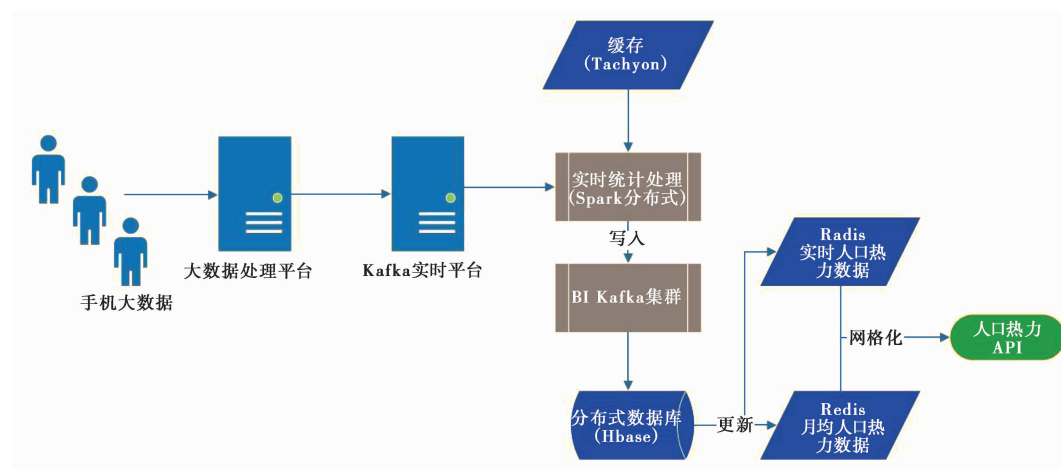


图 2 人口热力数据处理及接口设计

口分布情况。计算公式为

$$Y_i = p_n \times a + Y_{i-1} \times b \tag{1}$$

其中,  $Y_i$  为第  $i$  个网格的月均人口热力数据;  $p_n$  为当前最新的小时级区域人口分布数量;  $Y_{i-1}$  为前一小时的月均人口分布数量;  $a$  和  $b$  为权重系数, 用于准实时更新区域月均人口数据。

## 2 系统实现与部署

### 2.1 系统实现

通过大数据技术, 部署分布式处理集群实现海量用户数据的实时处理, 网格化后通过信息加密实时推送到中国地震台网中心, 按照式(1)完成数据的计算并更新到本地数据库。在地震信息实时获取上, 系统实时连接地震信息服务平台 (EQIM) 获取实时地震速报信息, 在地震发生后第一时间按照震中位置从数据库中获取人口分布数据。基于 Heatmap 完成人口热力可视化计算和不同震中距离的人口分布计算后, 调用地震 API 接口生成可视化结果, 如图 3 所示。

### 2.2 基于云技术的轻量级部署和服务

基于云计算技术实现地震人口热力图系统接口服务 API 的部署和运行。整个应用系统基于新浪 SAE 和阿里云平台设计开发, 自动化程度高, 稳定性强; 采用云计算的 SaaS 技术实现海量数据的高效处理, 全国任一位置均可实现全自动秒级产出与多场景应用服务推送, 这不仅为大震应急工作赢得了时间, 还提供了重要的数据支撑。基于上述技术, 系统成功实现地震人口热力图的全自动计算与产出, 并通过地震信息播报机器人、第三方平台对接转发等多种途径开展应用。

## 3 四川泸定 6.8 级地震震例应用

2022 年 9 月 5 日四川泸定发生 6.8 级强震, 全自动产出的地震人口热力图自动推送到中国地震局应急指挥中心, 发挥了重要的数据支撑作用。在历史地震数据中, 选取了 2022 年以来



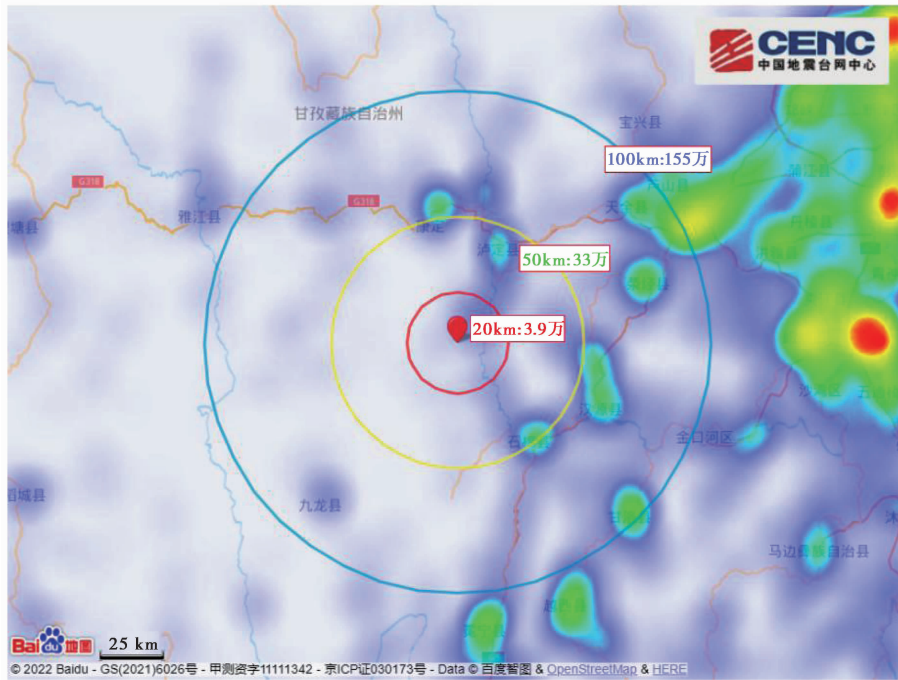


图3 面向移动端的人口热力数据可视化

中国大陆发生的 6.0 级以上地震数据参数,根据地震人口热力大数据,震后第一时间获取震中周边人口分布数据,在地震信息服务平台上实时显示所发生的地震震中 20km、50km 和 100km 范围内的周边人口数量,同时显示震中周边较近的村镇和县城。地震震中周边人口数量分布数据的可视化展示为地震应急指挥提供了参考和依据(表 1)。通过对比可以发现整体人口人力分布和普查统计基本在同一数量级,在人口密集区域如雅安市,可以看到人口热力数量明显高于普查统计,说明人口热力数据相对普查数据更能反应真实的人口分布情况。

表 1 2020 年以来中国大陆 6.0 级以上地震人口数据对比

震级	时间 (年-月-日 T时:分)	纬度 /(°N)	经度 /(°E)	深度 /km	位置	震中 20km 内人数 /万	震中 50km 内人数 /万	震中 100km 内人数 /万	普查统计 人数情况 (约 100km 范围)/万
6.1	2023-01-30T07:49	40.01	82.29	50	新疆阿克苏地区沙雅县	0	0	0.6	0
6.0	2022-06-10T01:28	32.25	101.82	13	四川阿坝藏族自治州 马尔康市	0.1	1.4	18.8	18.3
6.1	2022-06-01T17:00	30.37	102.94	17	四川雅安市芦山县	6.9	116.5	788.3	252
6.0	2022-03-26T00:21	38.50	97.33	10	青海海西蒙古族藏族 自治州德令哈市	0	0	0.1	0.15
6.9	2022-01-08T01:45	37.77	101.26	10	青海海北藏族自治州 门源县	0.2	3.1	71.2	91.4

2017 年 8 月 8 日九寨沟 7.0 级地震发生后,基于“地震人口热力大数据”形成了震中周边人口热力图。大数据准确地估算出震中附近的人口总数:震中 20km 范围内人口数约 2.1 万,50km 范围内约 6.3 万,100km 范围内约 30 万。人口热力图利用获取的手机数据,定位该

区域的用户数量,通过用户数量渲染地图颜色,实时展示该地区人口密度。由此可见,人口热力图便于观察一个地区的人口密集数量及实时的人流量情况。

同时,系统通过 API 接口将地震人口热力大数据实时推送至已经部署好的地震信息服务平台,震后第一时间获取震中人口分布可视化数据(图4),并在震后的媒体现场采访和报道中得到应用。从4图中我们可以看到震中区域 20km 范围内有 3.8 万人口分布,震中 50km 范围内有 32.6 万人分布,100km 范围内有 154.1 万人口分布,准确的信息数据以及基于 Heatmap 的可视化人口分布情况(图4),为第一时间评估受灾人口情况、开展应急救援方案制定提供数据支持,提升了应急处置效率。

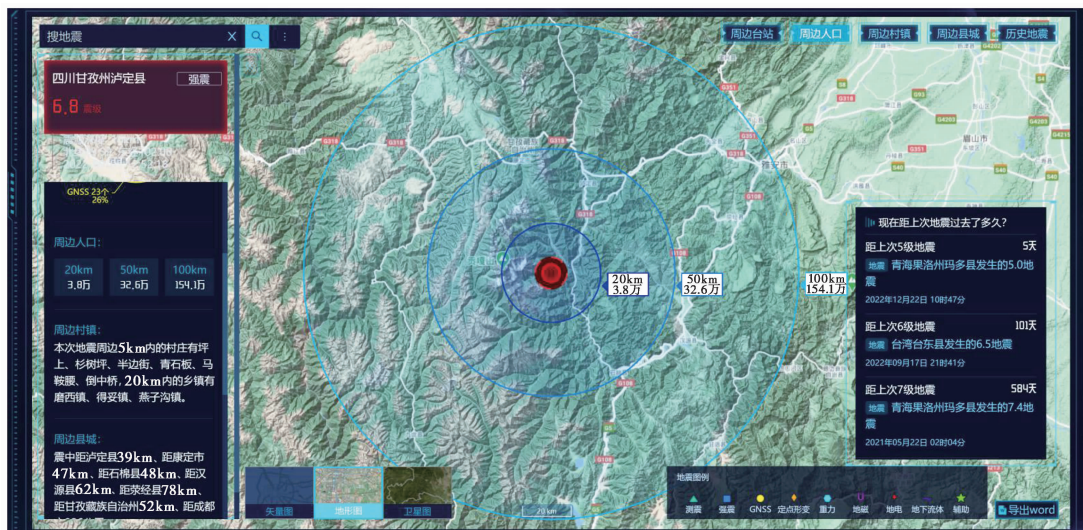


图4 面向应急决策的人口人力大屏可视化

四川泸定6.8级地震震中距泸定县39km、距康定市47km、距石棉县48km、距汉源县62km、距荥经县78km、距成都市226km。据2021年5月11日第七次全国人口普查数据结果显示,震中周边范围内的人口分布情况如下:泸定县常住人口数量为8.4204万,康定市常住人口数量为12.68万,石棉县常住人口数量为11.4116万,汉源县常住人口数量为28.5558万,荥经县人口数量为13.15万,震中50km范围内常住人口数量共计约32.51万。该地震震中20km内的乡镇有磨西镇、得妥镇、燕子沟镇,据国家统计局2020年数据显示,磨西镇户籍人口数量为7243人,得妥镇户籍人口数量为8114人,燕子沟镇户籍人口数量为5400人,震中区域20km范围内户籍人口数量共计约2.08万人。

通过对四川泸定地震计算的人口数据和实际人口数据的对比分析,震中区域50km范围内地震人口热力大数据约为32.6万人,相比于普查统计人口数据约32.51万人多0.09万人。震中区域20km范围内地震人口热力大数据约3.8万人,相比于普查人口数据2.08万人多1.72万人。人口热力大数据多于普查人口数量初步分析原因为四川泸定地区作为热门旅游景点,每年往来游客络绎不绝,地震人口热力数据量高于普查人口统计数据显示了人口数量变化的动态数据。从时间对比来看,计算的人口大数据更彰显了地震发生时当地的人口实时流动性的数据特征,普查的人口数据则具有时限性。从数据空间角度对比来看,地

震人口热力数据量能够直观反映出地震发生时不同区域单位标准范围的人口位置流动量信息,相比普查统计数据而言,可以显示相同单位标准人口流动数据密集度。

## 4 结语

移动人口热力大数据在震后的深度挖掘与多途径应用服务,使得地震应急工作的开展更加精准和高效,对地震灾害的评估和判断也更为准确,可有效助力地震应急处置工作,服务于公共安全治理水平提升。通过数次实际震例的检验,地震人口热力大数据产出结果与震中实际人口分布情况高度吻合,在震后应急中起到了支撑作用。

对比人口热力数据与传统人口数据,在人口数据处理效率方面,利用大数据和云计算等信息技术手段,前者提高了人口信息数据处理效率,对人口数据统计分析更精准细化,能够通过整体人口数据进行精细化划分和行为分析,增加对人口数量和行为的认知。人口热力数据能够体现最新的人口数量分布,通过对人口位置发生位移数据的实时捕捉和更新处理,能够快速获取和反映出最新的区域人口密度和数量,为大震应急指挥提供参考依据。

继续深度挖掘人口热力数据的价值,将地震灾害发生后的实时震感分布与震区实时热力人口数据进行深度业务融合,根据地震、余震等次生灾害的影响评估震区中人口热力变化,动态绘制地震人口热力变化趋势图,准确评判震中的影响范围,从而对地震灾害的应急指挥处置工作预判更加精准。

## 参考文献

- 侯建民. 2018. 大数据时代的地震信息播报. 新闻与写作, (3): 106~108.
- 侯建民, 郭凯, 崔满丰, 等. 2022. 基于可视化技术的地震信息服务系统设计与实现. 中国地震, 38(3): 574~584.
- 姜保庆, 郝锐朋. 2016. 基于 LBS 人口数据可视化监测系统. 计算机与现代化, (7): 107~110, 114.
- 杨微, 刘纪平, 王勇. 2012. 基于 Heatmap 的地理对象空间分布热度计算方法. 见: 第四届“测绘科学前沿技术论坛”论文精选. 北京: 测绘出版社, 398~400, 405.
- 姚迪. 2017. “大数据”应用于地震应急的新进展——手机热力图机器人全国地震值班系统. 中国应急救援, (6): 9~11.
- 张海林. 2021. 基于百度热力图的人口活动数量提取与规划应用. 城市交通, 19(3): 103~111.
- 赵深, 李靖, 朱文广. 2022. 基于手机信令大数据的应急人口热力图系统研究. 湖南邮电职业技术学院学报, 21(1): 16~19.
- 周芳检, 何振. 2018. 大数据时代城市公共安全应急管理面临的挑战与应对. 云南民族大学学报(哲学社会科学版), 35(1): 117~123.
- Murakami A, Nasukawa T. 2012. Tweeting about the tsunami?: mining twitter for information on the Tohoku earthquake and tsunami. In: Proceedings of the 21st International Conference on World Wide Web. Lyon, France: ACM.

## Implementation and Application of A Real-time Acquisition System for Earthquake Population with Heat Based on Mobile Big Data——A Case Study of Luding $M_s6.8$ Earthquake

Hou Jianmin<sup>1)</sup>, Guo Kai<sup>1)</sup>, Cui Manfeng<sup>1)</sup>, Fang Yi<sup>2)</sup>, Dong Lin<sup>2)</sup>, Zhai Ying<sup>1)</sup>, Ma Xiudan<sup>1)</sup>

1) China Earthquake Networks Center, Beijing 100045, China

2) Daily Interactive Technology Co., Ltd., Hangzhou 030024, China

**Abstract** Obtaining the population data around the epicenter immediately after an earthquake is crucial for making emergency response decisions. In this paper we use distributed clusters for real-time processing and gridding of massive data to obtain massive mobile internet population distribution data,. In order to ensure that the system can achieve efficiency in second level output after the earthquake, a multi-level multi-strategy thinning calculation and HeatMap based visualization are designed. After the earthquake, the thermal population data calculation and distribution map of the designated range around earthquake epicenter in the time scale of second are automatically drawn, and the Luding  $M_s6.8$  earthquake is applied as an example case. Our results demonstrate the advantage of obtaining epicenter population distribution based on massive mobile internet user data.

**Keywords:** Mobile big data; Population heat map; Distributed; The Luding earthquake