

中华人民共和国土地管理行业标准

TD/T 1085—2023

自然资源领域应用手机信令数据技术指南

Technical guidelines for the application of mobile signaling
data in the field of natural resources

2023-09-27 发布

2023-11-01 实施

中华人民共和国自然资源部 发布

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 总则	2
4.1 规范数据采集和处理	2
4.2 构建基础指标体系	3
4.3 提供应用场景参考	3
4.4 确保数据处理和应用中的用户隐私安全	3
4.5 按照统一的坐标参考	3
5 需考虑的因素	3
5.1 概要	3
5.2 数据采集和数据处理	3
5.3 基础指标体系	4
6 应用场景	5
6.1 概要	5
6.2 国土空间规划中的应用场景	6
6.3 自然资源监测监管中的应用场景	6
6.4 自然资源评价评估中的应用场景	6
附录 A (规范性) 数据采集和数据处理技术要点	7
附录 B (规范性) 基础指标体系组成和测算方法	10
附录 C (规范性) 国土空间规划中的应用场景示例	14
附录 D (规范性) 自然资源其他领域中的应用场景示例	18
参考文献	20

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》和 GB/T 20001.7—2017《标准编写规则 第7部分：指南标准》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中华人民共和国自然资源部提出。

本文件由全国地理信息标准化技术委员会(SAC/TC 230)提出并归口。

本文件主要起草单位：自然资源部信息中心、同济大学、吉林农业大学、深圳市规划和自然资源数据管理中心、武汉市规划编制研究和展示中心、智慧足迹数据科技有限公司、中移动信息技术有限公司。

本文件主要起草人：郭文华、钮心毅、陈红兵、刘文超、范延平、冯永恒、张帆、肖飞、郭晗、姜涛、张岩、康宁、邓颂平、王娜萍、李治君、胡玥、赵华、温长吉、吴多朋、李岩、曾建鹰、张玉韩、刘天羽、陈珍妮、聂可、林诗佳、刘思涵、施澄。

引 言

开展以人类活动为核心的总量统计、空间布局、流动情况等大数据标准制定,是行使自然资源两统一职能的重要支撑,是推进自然资源领域国家治理体系和治理能力现代化的必要条件。利用手机信令数据开展人口定量分析,弥补了传统统计数据的不足,提高了对人类活动的动态监测水平,具有良好的实用性与适用性。建立自然资源领域应用手机信令数据技术指南,有助于提升国土空间规划、自然资源监测监管、自然资源评价评估等方面的技术支撑能力。

为指导自然资源领域中手机信令数据产品的制作、构建科学实用的指标体系、拓展不同层级的应用场景、强化数据安全和隐私保护,制定本文件。

自然资源领域应用手机信令数据技术指南

1 范围

本文件提供了手机信令数据的采集和处理技术要点、基础指标体系,并给出了手机信令数据在国土空间规划、自然资源监测监管、自然资源评价评估等领域的典型应用场景。

本文件适用于国土空间规划、自然资源监测监管、自然资源评价评估以及自然资源领域其他需要应用手机信令数据的工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

- GB/T 18507—2014 城镇土地分等定级规程
- TD/T 1018 建设用地节约集约利用评价技术规程
- TD/T 1063 国土空间规划城市体检评估规程

3 术语和定义

下列术语和定义适用于本文件。

3.1

手机信令 mobile signaling

移动通信系统中,用来传输用户信息和保证正常通信所需要的设备与网络之间的协议控制信号。

[来源:JT/T 1182.1—2018,3.5]

3.2

手机信令数据 mobile signaling data

自然资源领域应用的手机信令数据应包含来自运营中的各代移动通信网络的信令数据。手机信令原始数据宜至少包含以下内容:

- 加密移动用户识别码;
- 信令时间戳;
- 信令事件类型代码:开关机、语音呼叫、短信、位置更新(包括用户移动触发的以及周期性触发两类)、切换,以及上网产生的附着(手机设备连接到运营商的网络)、去附着(手机设备断开运营商的网络连接)、数据事件等事件类型的信令记录;
- 经度:当前所在基站的地理位置的经度数值;
- 纬度:当前所在基站的地理位置的纬度数值。

3.3

基站 base station

在一定的无线电覆盖区中,通过移动通信交换中心,与移动电话终端之间进行信息传递的公用无线电收发信电台。

[来源:JT/T 1182.1—2018,3.1]

3.4

加密移动用户识别码 encrypted mobile subscriber identification code

由移动通信运营商对国际移动用户识别码进行加密而形成的,用于数据间关联处理的统一系统编码,与用户办理的号卡一一对应,是手机信令数据处理及服务中唯一识别信令所属移动用户的识别码。

[来源:JT/T 1182.1—2018,3.6,有修改]

注:一人多卡(同一用户办理多张号卡)的情况,会对应多个加密移动用户识别码;非人号卡(号卡的拥有者属于非人用户)的情况,会对应非人加密移动用户识别码。

3.5

原样数量 original sample quantity

直接使用移动通信用户的加密移动用户识别码测算得出的样本数量,也称为未扩样数量。

3.6

扩样数量 sample expanded quantity

原样数量通过某种扩样方法,推测得到的全体数量。

3.7

停留 dwell

移动通信用户位置连续位于某一区域内且超过最小停留时长,计为在该区域内的一次停留。

注:最小停留时长宜根据具体应用场景确定。

3.8

移动 journey

移动通信用户按照时间的先后顺序,从一个停留位置变动到另一个停留位置,计为一次移动。

3.9

重复率 repetition rate

在指定区域和时段内,满足某类规则的人群重复出现的天数占总时段天数的比例。

3.10

空间位置 location

用坐标数据表达的地理定位信息,一般采用经纬度方式表达。

4 总则

4.1 规范数据采集和处理

提出适用于自然资源领域的手机信令数据采集和处理技术要点,指导典型应用场景中数据特性和特

性值的选取。

4.2 构建基础指标体系

构建基础指标体系,实现手机信令数据记录的移动通信用户行为轨迹与自然资源领域应用场景的对应与衔接。

4.3 提供应用场景参考

为手机信令数据在自然资源领域不同应用场景中的应用条件和应用方法提供参考。

4.4 确保数据处理和应用中的用户隐私安全

数据处理中,采用匿名加工的手段是至关重要的。

数据应用中,宜采用聚合统计的方式,并严格执行《中华人民共和国个人信息保护法》等法律法规。

4.5 按照统一的坐标参考

空间数据的坐标转换是至关重要的,在数据应用前宜转换为 2000 国家大地坐标系(CGCS 2000),采用地理坐标表达。

空间信息处理严格执行《测绘地理信息管理工作国家秘密范围的规定》(自然资源发〔2020〕95号)。

5 需考虑的因素

5.1 概要

采集和处理手机信令数据时,可与移动通信运营商针对数据采集、数据处理、扩样、数据服务方式等方面确定技术要点和特征值选取规则,形成数据采集和数据处理及服务方案。数据采集和数据处理技术要点,见附录 A。

应用手机信令数据时,宜首先构建基础指标体系。基础指标体系由停留和移动指标、居住和就业指标、人口规模指标组成。基础指标体系组成和测算方法,见附录 B。

5.2 数据采集和数据处理

5.2.1 数据采集

可从移动通信数据源的选择、手机信令原始数据采集的基本内容、移动通信网络的地理空间和人群覆盖度、数据质量校验的方法和内容等方面确定数据采集阶段的基本需求,用以确保所采集手机信令数据内容和质量的科学性和合理性。

5.2.2 数据处理

可从剔除非人号卡、一人多卡去重、异常数据处理、坐标系统转换、数据安全性等方面确定数据处理阶段的基本需求,以确保数据处理过程的准确性和合法性。

数据处理过程中,应当严格遵守国家相关法律法规,确保用户隐私安全。

5.2.3 扩样

手机信令数据统计、分析工作需要数据扩样的,宜从应用场景和扩样步骤等方面确定扩样的必要性

以及具体的扩样方法,以确保扩样后的数据符合应用场景的需求。

5.2.4 数据服务方式

移动通信运营商通常可提供数据沙箱和按需统计两种数据服务方式。

自然资源领域应用手机信令数据时,宜根据应用场景的需求,与移动通信运营商确定适合的手机信令数据服务方式。

5.3 基础指标体系

5.3.1 基础指标体系组成

基础指标体系宜由停留和移动指标、居住和就业指标、人口规模指标组成。

- a) 停留和移动指标是直接由手机信令数据测算得到,用于表征移动通信用户单次行为轨迹特征的指标。
- b) 居住和就业指标是基于停留和移动指标,进一步结合时间、空间规律描述居住、就业活动的指标。
- c) 人口规模指标是基于居住和就业指标,进一步结合时间、空间规律描述自然资源领域应用场景的指标。

随着技术进步,指标内涵和获取方式会进行适当调整。

5.3.2 停留和移动指标

5.3.2.1 停留指标的测算

测算停留指标涉及停留位置、停留开始时间、停留结束时间,宜明确界定停留时段和最小停留时长,最小停留时长取决于应用场景。在某一位置发生小于最小停留时长的行为,不计为一次停留。最小停留时长的阈值不同,停留指标结果有差异。

注:停留指标包括停留位置、停留时长、停留人数,以及单日夜间停留地、单日夜间停留人数、单日日间停留地、单日日间停留人数等。

5.3.2.2 移动指标的测算

测算移动指标涉及移动的出发位置、出发时间及移动的到达位置、到达时间。移动的出发位置、到达位置分别是前后两次停留的位置。

注:移动指标包括移动距离、移动时长等。

5.3.3 居住和就业指标

5.3.3.1 居住指标的测算

手机信令数据测算的居住指标包括居住地位置、居住地人数等。

- a) 居住地位置可根据应用场景考虑连续数据周期、单日夜间停留地累计停留时长、单日夜间停留地重复率三个参数的取值。
- b) 居住地人数可根据某个区域范围内,连续 30 日以上的数据周期内,由手机信令数据测算出的具有固定居住地位置的用户数获得。

注:手机信令数据测算的居住活动,是指用户在夜间时段、在固定位置发生停留,达到单日一定累计停留时长并且多日重复的活动。

5.3.3.2 就业指标的测算

手机信令数据测算的就业指标包括就业地位置、就业人数、通勤距离、通勤时间等。

- a) 就业地位置可根据应用场景考虑连续数据周期、工作日单日日间停留地累计停留时长、工作日单日日间停留地重复率三个参数的取值。
- b) 就业人数可根据某个区域范围内,连续 30 日以上的数据周期内,由手机信令数据测算出的具有固定就业地位置的用户数获得。
- c) 通勤距离可根据居住地位置、就业地位置测算。
- d) 通勤时间可根据居住地、就业地的出发、到达时间测算。

注:手机信令数据测算的就业活动,是指用户在工作日的日间时段、在固定位置发生停留,达到单日一定累计停留时长并且多日重复的活动。手机信令数据测算的就业活动未包含不符合上述条件的就业活动,但可能包含符合上述条件的非就业活动。

5.3.4 人口规模指标

手机信令数据测算的人口规模指标包括常住人口、短期驻留人口、实际服务人口、城市对外人口流动量等。

- a) 常住人口的测算宜以连续 6 个月为测算时段,筛选居住天数最长的居住地位置所在区域,归属为该区域的常住人口。
- b) 短期驻留人口宜选取目标区域所在城市为基本空间单元,筛选单日连续停留时长 3 小时以上,但是按附录 B 中 B.4.1,尚不属于该城市常住人口的,作为该城市的当日短期驻留人口计算对象。
- c) 实际服务人口宜包括常态值和节假日极大值两个测算口径。实际服务人口常态值宜参考 TD/T 1063 中实际服务管理人口数量的规定测算,选取 11 月为特征月,提取单日在目标城市连续停留 3 小时以上的日均用户数作为常态值。实际服务人口节假日极大值宜选取包含节假日时段内在目标城市连续停留 3 小时以上用户数的单日最大值,测算过程中宜至少包括春节、国庆等重大节假日。
- d) 城市对外人口流动量的测算宜首先筛选有效停留城市,再依据用户停留发生的时间先后进行排序,得到从研究城市出发到外地的移动,以及从外地出发回到研究城市的移动,对前者以出发日期、对后者以到达日期进行统计,加和得到城市对外人口流动量。

测算人口规模指标宜根据应用场景需求设定适当的测算时段。

6 应用场景

6.1 概要

手机信令数据应用主要考虑数据应用方式、指标计算方式、数据隐私保护、应用模型构建、指标维度,还要能适应未来发展,应对技术和管理等因素变化情形。

a) 确定数据应用方式。手机信令在自然资源管理中的应用场景,首先考虑群体性人类活动的影响,使用聚类统计数据;其次在数据使用时也要考虑数据安全保密方面限制,尤其不能使用涉及个人隐私等信息。

b) 明确指标计算方式。手机信令数据可依据 5.3 进行数据处理计算,基础指标计算方法亦可根据需求的特点进行动态调整。

c) 指出指标维度。手机信令在自然资源中应用场景,不但要考虑具体位置信息、空间范围,还要考虑时间维度,并考虑由于技术进步、管理优化引起应用场景变化的情形。

d) 适应未来发展。当前技术条件和管理需求,手机信令可应用于国土空间规划、自然资源监测监管、土地利用评价评估、自然资源评价评估等领域,未来应用领域可逐步扩展。

6.2 国土空间规划中的应用场景

通过在允许人口活动的时空内测算人类活动的特征,手机信令数据可应用在国土空间规划编制及实施监督,特别是国土空间规划城市体检评估中的人口规模和分布测算、职住空间关系测算、公共活动中心和主要服务设施人数和范围分析、城市空间关联分析、城市内部出行分析等方面。

国土空间规划中的应用场景示例见附录 C。

6.3 自然资源监测监管中的应用场景

手机信令在自然资源监测监管中的应用分为确定监测监管单元、数据采集和处理以及数据场景应用三个步骤。

- a) 确定监测监管单元。针对自然资源业务特点确定监测监管范围,通过基站覆盖范围分析确定基站影响范围,通过空间数据分析等技术手段结合管理指标确定自然资源监测监管单元。
- b) 数据采集和处理。针对不同类型的监测监管单元,结合限制人类活动的范围、鼓励或允许人类活动的范围以及特定区域等空间约束条件,数据运营商生产不同应用场景下基于手机信令的人类活动监测数据集。
- c) 数据场景应用。利用基于手机信令的人类活动监测数据集,结合自然资源业务监测监管相关信息以及自然资源“一张图”信息,在不同的业务应用场景利用约定指标实现分析应用和监测预警。

自然资源监测监管中的应用场景,见附录 D。

6.4 自然资源评价评估中的应用场景

手机信令数据可应用在建设用地节约集约利用评价和自然资源分等定级估价场景。

- a) 应用于建设用地节约集约利用评价场景。手机信令数据可应用于建设用地(含各类开发区)节约集约利用评价业务中常住人口规模测算、疑似土地闲置情况监测、就业情况监测等,辅助开展开发区节约集约利用评价。
- b) 应用于自然资源分等定级估价场景。手机信令数据亦可应用于自然资源分等定级、价格评估中指标量化和数据校对等,用作自然资源评价评估的补充信息来源。当前也可应用于城镇土地分等定级估价中住宅用地人口密度、商业用地人口密度以及产业聚集度等测算等。

自然资源评价评估中的应用场景,见附录 D。

附 录 A
(规范性)
数据采集和数据处理技术要点

A.1 数据采集环节技术要点

A.1.1 移动通信数据源选择

自然资源领域中应用的手机信令数据可以采集自一家移动通信运营商,也可以采集自多家移动通信运营商。如果同时采集多家移动通信运营商的数据,宜优先用于相互比较、相互校核以提高结果可靠性。因存在跨运营商的一人多卡现象,多家移动通信运营商数据统计结果不宜直接累加。若使用累加数据,需要标注清楚。

A.1.2 地理空间和人群覆盖

手机信令数据宜按照行政区划单元采集,常用行政区划单元是地级市。移动通信网络宜完全覆盖该行政区划单元内的空间范围。

移动通信运营商记录的扩样前日均活跃用户数宜高于该区域常住人口的10%。

注:活跃用户指当日留下至少一条信令记录的用户。

A.1.3 数据质量校验

数据质量校验包括对信令数据检查和对基站坐标检查。

- a) 对信令数据检查。包括对信令记录数量、记录到用户数的检查。小时变化检查是以小时为单位,日数量变化检查是以日为单位,如信令日变或小时变化检查发现异常,均宜排除该日期数据。
- b) 对基站坐标检查。基站坐标检查是检查基站坐标经度、纬度的完整性以及是否在空间边界内,并确保基站均能获得正确的空间定位坐标。对于确实存在错误的,需通过原始轨迹对其正确位置进行估算校准。

A.2 数据处理环节技术要点

A.2.1 剔除非人号卡

移动通信运营商的号卡不仅包括真实用户,也包括物联网设备等非人号卡。自然资源领域使用的手机信令数据,需要识别并剔除物联网卡等非人号卡产生的信令记录。识别物联网卡等非人号卡要依据物联网特殊号段以及从移动用户识别码获取的设备类型等进行。利用号卡与加密移动用户标识码的一一对应关系,将非人号卡对应的信令记录剔除。

A.2.2 一人多卡去重

由于存在一人随身携带并使用多个手机号卡的情况,需要进行一人多卡去重。一人多卡去重包括网内一人多卡去重和网间一人多卡去重两种情况。

- a) 网内一人多卡去重。同一用户随身携带并使用同一运营商的多个手机号卡,可通过测算多个加

密移动用户识别码的轨迹重合度进行网内一人多卡去重。

- b) 网间一人多卡去重。由于各移动通信运营商带加密移动用户识别码的数据都仅能在各自机房指定安全区域内完成数据处理,同一用户随身携带两个及以上运营商号卡情况的去重,即网间一人多卡去重当前尚难以实现。

A.2.3 “乒乓效应”处理

移动通信系统中,在多个邻近基站交叉覆盖区域,由于两者信号强度相似,手机会在两者之间频繁进行切换,形成所谓的“乒乓效应”。具体处理过程中,需要对每日的信令记录进行统计,以基站对的形式,考察各对之间的切换数量,同时考虑时间跨度,如果在短时间内切换数量过多(一般以超过3次/分钟为准),即认为存在“乒乓效应”。宜以用户在“乒乓效应”发生期间连接次数最多的基站为主基站,剔除“乒乓效应”期间产生的信令数据中主基站外的记录,完成对数据的清洗。

A.2.4 “基站漂移”处理

“基站漂移”是指移动通信用户突然从邻近基站切换至远处基站,一段时间后又切回邻近基站的情况,导致用户的出行异常。具体处理过程中:一方面需在信令数据采集中保证基站的地理位置信息为最新数据,另一方面可以根据每条信令记录的停留时间、停留位置及其下一条记录的停留时间、停留位置,测算每次移动的速度,宜通过设定速度阈值,剔除短时间内发生远距离位移的信令记录。

A.2.5 数据安全

从加密移动用户识别码无法识别特定个人并无法复原对数据安全是至关重要的,在一般情况下只许输出聚合统计后的指标结果。

同一个移动用户识别码在不同区域加密后得到的加密移动用户识别码应相同。

A.3 扩样环节技术要点

A.3.1 扩样的定义

以手机信令数据获取的原样数量为统计样本,通过一定的算法推算全量数据,这个过程简称“扩样”。

A.3.2 扩样的使用场景

扩样是手机信令数据统计、分析阶段的数据处理工作,不是必要环节,可根据具体应用场景的实际需要决定是否进行扩样。如果进行扩样,宜说明扩样方法。

- a) 当应用场景中只需要描述人数占比相对值时,不需要进行扩样,可直接使用统计得出的原样数量。
- b) 当应用场景中需要描述人数总量的绝对值时,需要进行扩样,使用扩样数量。
- c) 使用多家移动通信运营商的数据不能取代扩样环节。全部运营商的数据总和虽能实现手机设备的全样本统计,但是手机设备的全样本并不等同于人数的全样本。

A.3.3 扩样的方法

使用一家移动通信运营商的手机信令数据,扩样方法宜按照以下两个步骤进行。

- a) 从“一家移动通信运营商网内不计重总数”结合该移动通信运营商在当地的市场占有率(粒度需至少达到地市级尺度)得到“全部移动通信运营商网内不计重总数”,市场占有率的测定宜参照活跃用户的规模以保证该结果的可靠性,或者通过长期观测的话单数据估算得到。

- b) 从“全部移动通信运营商网内不计重总数”结合目标区域手机普及率得到“实际人数的扩样数量”,手机普及率的测定可通过问卷调查,或者考虑当地老龄化程度、少儿人口比重,以及社会发展水平估算得到。

A.4 数据服务方式

移动通信运营商可以提供两种手机信令数据服务方式。

- a) 数据沙箱:第一种数据服务的数据处理结果是在符合 A.2 数据处理环节技术要点的条件下,保留信令记录,对 A.1.2 信令数据字段中的信令时间戳、信令事件类型代码、经度、纬度进行必要模糊处理。第一种数据服务的的结果必须保存在移动通信运营商机房指定安全区域内,以数据沙箱形式提供服务。
- b) 按需统计:第二种数据服务的的结果是在符合 A.2 数据处理环节技术要点的条件下,按指定空间位置、指定时间段汇总用户人数,不保留加密移动用户识别码、不保留信令记录。第二种数据服务的的结果可以提供文本格式文件。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/497035102122006131>