自动场景识别开发文档（居民区）

# 地图POI提取

从地图上的POI信息里面把包含居民区、住宅等等的所有楼宇都获取，然后进行分析，合并出居民区的目标列表。

*注：本文档以高德地图为例，百度地图及其他平台流程类似*

## 通过行政区域查询获取区县信息

申请完开放平台（http://lbs.amap.com/）Key后，在搜索服务API中可以通过city&citylimit参数指定希望搜索的城市或区县。而city参数能够接收citycode和adcode，citycode仅能精确到城市，而adcode却能够精确到区县。例如：

北京，citycode：010，adcode：110000

北京-海淀区，citycode：010，adcode：110108

故使用citycode仅能在北京范围内搜索，而adcode能够指定在海淀区搜索。

由于存在流量限制说明（附录1），故采取按照区县级切片爬取POI数据。

首先需要通过行政区域查询API服务地址（<http://restapi.amap.com/v3/config/district?parameters>）获取区县边界多边形坐标点集（parameters代表的参数包括必填参数和可选参数。所有参数均使用和号字符(&)进行分隔。下面的列表枚举了这些参数及其使用规则）

请求参数：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名 | 含义 | 规则说明 | 是否必须 | 缺省值 |
| key | 请求服务权限标识 | 用户在高德地图官网申请Web服务API类型KEY | 必填 | 无 |
| keywords | 查询关键字 | 规则：只支持单个关键词语搜索关键词支持：行政区名称、citycode、adcode | 可选 | 无 |
|  |  | 例如，在subdistrict=2，搜索省份（例如山东），能够显示市（例如济南），区（例如历下区） |  |  |
| subdistrict | 子级行政区 | 规则：设置显示下级行政区级数（行政区级别包括：国家、省/直辖市、市、区/县、乡镇/街道多级数据） | 可选 | 1 |
|  |  | 可选值：0、1、2、3等数字，并以此类推 |  |  |
|  |  | 0：不返回下级行政区； |  |  |
|  |  | 1：返回下一级行政区； |  |  |
|  |  | 2：返回下两级行政区； |  |  |
|  |  | 3：返回下三级行政区； |  |  |
|  |  |  |  |  |
|  |  | 需要在此特殊说明，目前部分城市和省直辖县是无法返回街道的，这块我们在进行优化，若您需要还请耐心等待一下。 |  |  |
|  |  | 例如：广东-东莞、海南-文昌市 |  |  |
| page | 需要第几页数据 | 最外层的districts最多会返回20个数据，若超过限制，请用page请求下一页数据。 | 可选 | 1 |
|  |  | 例如page=2；page=3。默认page=1 |  |  |
| offset | 最外层返回数据个数 |  | 可选 | 20 |
| extensions | 返回结果控制 | 此项控制行政区信息中返回行政区边界坐标点； 可选值：base、all; | 可选 | base |
|  |  | base:不返回行政区边界坐标点； |  |  |
|  |  | all:只返回当前查询district的边界值，不返回子节点的边界值； |  |  |
|  |  | 目前不能返回乡镇/街道级别的边界值 |  |  |
| filter | 根据区划过滤 | 按照指定行政区划进行过滤，填入后则只返回该省/直辖市信息 | 可选 |  |
|  |  | 需填入adcode，为了保证数据的正确，强烈建议填入此参数 |  |  |
| callback | 回调函数 | callback值是用户定义的函数名称，此参数只在output=JSON时有效 | 可选 |  |
| output | 返回数据格式类型 | 可选值：JSON，XML | 可选 | JSON |

### 然后根据坐标点集去MAX得到区县的外接矩形

返回结果参数：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 名称 | | 含义 | 规则说明 | |
| status | | 返回结果状态值 | 值为0或1，0表示失败；1表示成功 | |
| info | | 返回状态说明 | 返回状态说明，status为0时，info返回错误原因，否则返回“OK”。 | |
| infocode | | 状态码 | 返回状态说明，10000代表正确，详情参阅info状态表 | |
| suggestion | | 建议结果列表 |  | |
|  | keywords | 建议关键字列表 |  | |
|  | cites | 建议城市列表 |  | |
| districts | | 行政区列表 |  | |
|  | district | 行政区信息 |  | |
|  |  | citycode | 城市编码 |  |
|  |  | adcode | 区域编码 | 街道没有独有的adcode，均继承父类（区县）的adcode |
|  |  | name | 行政区名称 |  |
|  |  | polyline | 行政区边界坐标点 | 当一个行政区范围，由完全分隔两块或者多块的地块组成，每块地的 polyline 坐标串以 | 分隔 。如北京 的 朝阳区 |
|  |  | center | 城市中心点 | 必须说明：在区县级别，有28个区县不能返回中心点，在乡镇/街道界别，有9262个乡镇/街道不能返回中心点 |
|  |  | level | 行政区划级别 | country:国家 province:省份（直辖市会在province和city显示） city:市（直辖市会在province和city显示） district:区县 street:街道 |
|  |  | districts | 下级行政区列表，包含district元素 |  |

## 通过搜索服务API获取居民区POI信息

通过多边形搜索在选取区县进行矩形分区搜索

### 首先以选取区县外接矩形左上坐标点为起始点，以1000m为边长切割外接矩形。

### 判断该切割子矩形是否在选取的区县边界内，如果否，则舍弃该子矩形，如果是，则进行下一步。

### 对该有效子矩形进行POI搜索

多边形搜索API服务地址：

http://restapi.amap.com/v3/place/polygon?parameters

（parameters代表的参数包括必填参数和可选参数。所有参数均使用和号字符(&)进行分隔。下面的列表枚举了这些参数及其使用规则）

请求参数：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名 | 含义 | 规则说明 | 是否必须 | 缺省值 |
| key | 请求服务权限标识 | 用户在高德地图官网申请Web服务API类型KEY | 必填 | 无 |
| polygon | 经纬度坐标对 | 规则：经度和纬度用","分割，经度在前，纬度在后，坐标对用"|"分割。经纬度小数点后不得超过6位。         多边形为矩形时，可传入左上右下两顶点坐标对；其他情况下首尾坐标对需相同。 | 必填 | 无 |
| keywords | 查询关键字 | 规则： 多个关键字用“|”分割 | 可选 | 无 |
| types | 查询POI类型 | 多个类型用“|”分割，详见附录2 | 可选 |  |
| offset | 每页记录数据 | 强烈建议不超过25，若超过25可能造成访问报错 | 可选 | 20 |
| page | 当前页数 | 最大翻页数100 | 可选 | 1 |
| extensions | 返回结果控制 | 此项默认返回基本地址信息；取值为all返回地址信息、附近POI、道路以及道路交叉口信息。 | 可选 | base |
| sig | 数字签名 | 数字签名获取和使用方法 | 可选 | 无 |
| output | 返回数据格式类型 | 可选值：JSON，XML | 可选 | JSON |
| callback | 回调函数 | callback值是用户定义的函数名称，此参数只在output=JSON时有效 | 可选 | 无 |

返回结果参数：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 名称 | | | 含义 | | 规则说明 |
| status | | | 结果状态值，值为0或1 | | 0：请求失败；1：请求成功 |
| info | | | 返回状态说明 | | status为0时，info返回错误原因，否则返回“OK”。详情参阅info状态表 |
| count | | | 搜索方案数目(最大值为1000) | |  |
| suggestion | | | 城市建议列表 | | 当用户输入的词语为泛搜索词的时候，将显示城市列表 |
|  | keywords | | 唯一ID | |  |
|  | cities | | 城市列表 | |  |
|  |  | name | 名称 | |  |
|  |  | num | 该城市包含此关键字的个数 | |  |
|  |  | citycode | 该城市的citycode | |  |
|  |  | adcode | 该城市的adcode | |  |
| pois | | | 搜索POI信息列表 | |  |
|  | poi | | POI信息 | |  |
|  |  | id | 唯一ID | |  |
|  |  | name | 名称 | |  |
|  |  | type | 兴趣点类型 | | 顺序为大类、中类、小类 例如：餐饮服务;中餐厅;特色/地方风味餐厅 |
|  |  | typecode | 兴趣点类型编码 | | 例如：050118 |
|  |  | biz\_type | 行业类型 | |  |
|  |  | address | 地址 | | 东四环中路189号百盛北门 |
|  |  | location | 经纬度 | | 格式：X,Y |
|  |  | distance | 离中心点距离 | | 单位：米 必须说明， 此结果仅在周边搜索的时候有值 |
|  |  | tel | 该POI的电话 | |  |
|  |  | postcode | 邮编 | | extensions=all的时候显示 |
|  |  | website | 该POI的网址 | | extensions=all的时候显示 |
|  |  | email | 该POI的电子邮箱 | | extensions=all的时候显示 |
|  |  | pcode | poi所在省份编码 | | extensions=all的时候显示 |
|  |  | pname | poi所在省份名称 | | 若是直辖市的时候，此处直接显示市名，例如北京市 |
|  |  | citycode | 城市编码 | | extensions=all的时候显示 |
|  |  | cityname | 城市名 | | 若是直辖市的时候，此处直接显示市名，例如北京市 |
|  |  | adcode | 区域编码 | | extensions=all的时候显示 |
|  |  | adname | 区域名称 | | 区县级别的返回，例如朝阳区 |
|  |  | entr\_location | 入口经纬度 | | extensions=all的时候显示 |
|  |  | exit\_location | 出口经纬度 | | extensions=all的时候显示 |
|  |  | navi\_poiid | 地图编号 | | extensions=all的时候显示 |
|  |  | gridcode | 地理格ID | | extensions=all的时候显示 |
|  |  | alias | 别名 | | extensions=all的时候显示 |
|  |  | business\_area | 所在商圈 | | extensions=all的时候显示 |
|  |  | parking\_type | 停车场类型 | | 仅在停车场类型POI的时候显示该字段 展示停车场类型，包括：地下、地面、路边  extensions=all的时候显示 |
|  |  | tag |  | 该POI的特色内容 | 主要出现在POI为美食类的POI之中出现，此时代表特色菜 例如“烤鱼,麻辣香锅,老干妈回锅肉” 在其余POI类别的时候，会有小概率出现，此时代表此POI的特色内容 例如北京-第五季花艺 之中的“岁月静好,定情,如意花卡,圣诞之爱” 仅在extensions=all |
|  |  | indoor\_map |  | 是否有室内地图标志 | 1，表示有室内相关数据 0，代表没有室内相关数据  extensions=all的时候显示 |
|  |  | indoor\_data |  | 室内地图相关数据 | 当indoor\_map=0时，字段为空  extensions=all的时候显示 |
|  |  |  | cpid | 当前POI的父级POI | 如果当前POI为建筑物类POI，则cpid为自身POI ID；如果当前POI为商铺类POI，则cpid为其所在建筑物的POI ID |
|  |  |  | floor | 楼层索引 | 一般会用数字表示，例如8 |
|  |  |  | truefloor | 所在楼层 | 一般会带有字母，例如F8 |
|  |  | groupbuy\_num | 团购数据 | | 此字段逐渐废弃 |
|  |  | business\_area | 所属商圈 | | extensions=all的时候显示 |
|  |  | discount\_num | 优惠信息数目 | | 此字段逐渐废弃 |
|  |  | biz\_ext | 深度信息 | | extensions=all的时候显示 |
|  |  |  | rating | 评分 | 仅存在于餐饮、酒店、景点、影院类POI之下 |
|  |  |  | cost | 人均消费 | 仅存在于餐饮、酒店、景点、影院类POI之下 |
|  |  |  | meal\_ordering | 是否可订餐 | 仅存在于餐饮相关POI之下（此字段逐渐废弃） |
|  |  |  | seat\_ordering | 是否可选座 | 仅存在于影院相关POI之下（此字段逐渐废弃） |
|  |  |  | ticket\_ordering | 是否可订票 | 仅存在于景点相关POI之下（此字段逐渐废弃） |
|  |  |  | hotel\_ordering | 是否可以订房 | 仅存在于酒店相关POI之下（此字段逐渐废弃） |
|  |  | photos | 照片相关信息 | | extensions=all的时候显示 |
|  |  |  | titile | 图片介绍 |  |
|  |  |  | url | 具体链接 |  |

根据返回的types，floor等信息输出shape格式，csv格式及导入数据库操作。

### 为每一个居民区打重要程度标识，分为高级，中级和低级三个级别。

其中高德地图POI居民区分类和平台自动场景识别映射表如下：

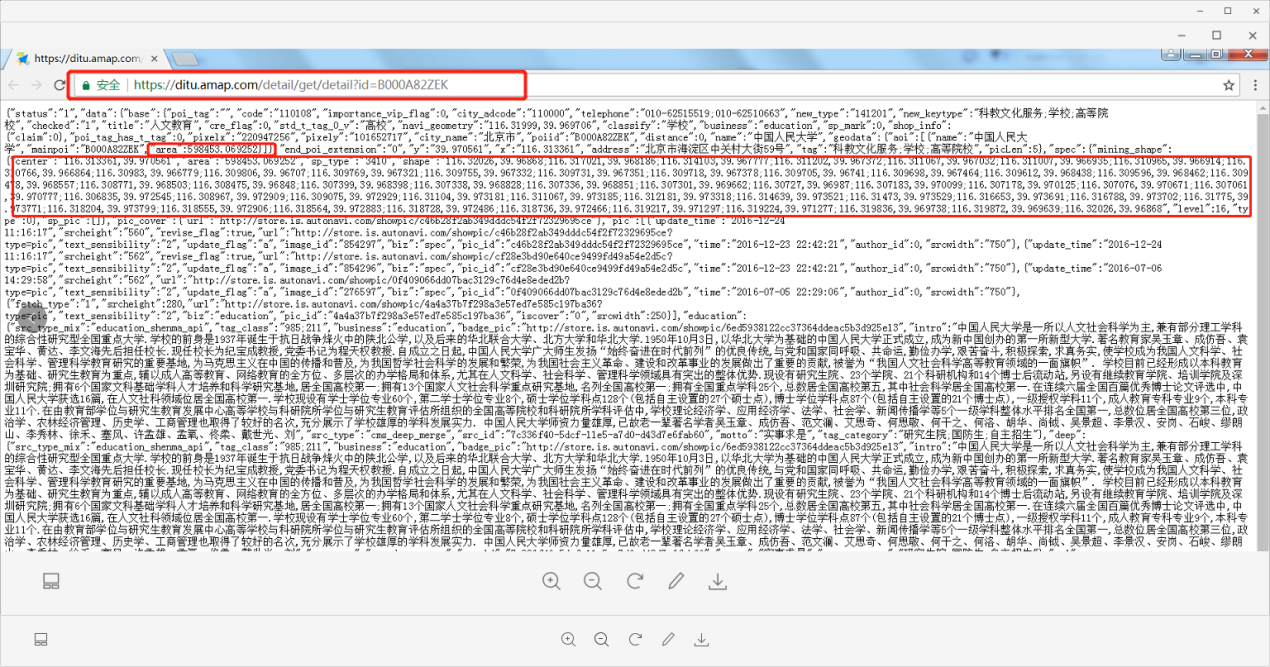
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **序号** | **NEW\_TYPE** | **大类** | **中类** | **小类** | **Big Category** | **Mid Category** | **Sub Category** | **重要程度标识** |
| 524 | 120100 | 商务住宅 | 产业园区 | 产业园区 | Commercial House | Industrial Park | Industrial Park | 高级 |
| 525 | 120200 | 商务住宅 | 楼宇 | 楼宇相关 | Commercial House | Building | Building Related | 高级 |
| 526 | 120201 | 商务住宅 | 楼宇 | 商务写字楼 | Commercial House | Building | Business Office Building | 高级 |
| 527 | 120202 | 商务住宅 | 楼宇 | 工业大厦建筑物 | Commercial House | Building | Industrial Building | 中级 |
| 528 | 120203 | 商务住宅 | 楼宇 | 商住两用楼宇 | Commercial House | Building | Commercial-residential Building | 低级 |
| 529 | 120300 | 商务住宅 | 住宅区 | 住宅区 | Commercial House | Residential Area | Residential Area | 高级 |
| 530 | 120301 | 商务住宅 | 住宅区 | 别墅 | Commercial House | Residential Area | Villa | 高级 |
| 531 | 120302 | 商务住宅 | 住宅区 | 住宅小区 | Commercial House | Residential Area | Residential Quarter | 高级 |
| 532 | 120303 | 商务住宅 | 住宅区 | 宿舍 | Commercial House | Residential Area | Dormitory | 中级 |
| 533 | 120304 | 商务住宅 | 住宅区 | 社区中心 | Commercial House | Residential Area | Community Center | 低级 |

## 通过查看地图请求服务爬取建筑物边界信息

由于高德提供的API接口POI信息没有建筑物边界信息，需要查看高德地图请求服务通过浏览器检测到其隐藏接口为：

https://ditu.amap.com/detail/get/detail? parameters

其中返回的面积为参数area，建筑物边界信息为坐标点集合参数shape



## 重点区域郊区区分

获取重点城区郊区地理图层，对居民区标签进行精细化处理，规则如下：

如判断该居民区坐标在城区范围，居民区标签不变；如判断该居民区坐标在郊区范围，则原高级居民区改为中级居民区标签，原中级居民区改为低级居民区标签，原低级居民区不变。

# 数据处理

## 经纬度偏移

通过坐标转换模块把坐标转换为GPS坐标

## 居民区相关无线小区

### 关联POI数据，工参数据和MR数据

通过POI中建筑物边界信息经纬度，判断工参中是否有小区在居民区形状内，是则关联该小区到POI居民区，否则执行下一步；

通过POI中居民区经纬度，取500m范围为半径判断周边小区是否在该小区覆盖范围内，是则关联该小区到POI居民区，否则执行下一步；

通过MR或MDT定位后的数据把主要覆盖范围在该居民区的小区关联到POI居民区小区列表中。

## 居民区覆盖维度指标统计

居民区目标展示的字段如下表：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 居民区名称 | 海兴大厦 |  |  |  |  |
| 城市 | 北京 |  |  |  |  |
| 区县 | 海淀区 |  |  |  |  |
| 经纬度 | 39.719345，116.043561 |  |  |  |  |
| 地址 | 北京市海淀区丹棱街16号 |  |  |  |  |
| 面积 | 100平米 |  |  |  |  |
| 关联小区信息 | 距离 | RSRP | SINR | 综合覆盖率 | RRC建立连接率 |
| 海兴大厦1 | 200m | -77dB | 0 | 99% | 99% |
| 海兴大厦2 | 201m | -78dB | 0 | 99% | 99% |

# 附录1

Web服务API用户分类及流量说明：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **服务** | **个人开发者** | | **企业开发者** | |
| 日调用超量封停 | QPS | 日调用超量报警 | QPS |
| （次/日） | （次/秒） | （次/日） | （次/秒） |
| 地理编码 | 6000 | 100 | 3000000 | 1000 |
| 逆地理编码 | 6000 | 100 | 3000000 | 1000 |
| 驾车路径规划 | 2000 | 20 | 300000 | 200 |
| 公交路径规划 | 2000 | 20 | 300000 | 200 |
| 步行路径规划 | 2000 | 20 | 300000 | 200 |
| 骑行路径规划 | 2000 | 20 | 300000 | 200 |
| 货车路径规划 | 100 | 1 | 100 | 1 |
| 距离测量 | 2000 | 20 | 300000 | 200 |
| 搜索 | 2000 | 20 | 300000 | 200 |
| 输入提示 | 2000 | 20 | 300000 | 200 |
| 行政区查询 | 2000 | 20 | 300000 | 200 |
| 交通态势查询 | 2000 | 20 | 300000 | 200 |
| 抓路服务 | 2000 | 20 | 300000 | 200 |
| 静态地图 | 1000000 | 400 | 6000000 | 1000 |
| IP定位 | 100000 | 100 | 3000000 | 1000 |
| 坐标转换 | 100000 | 100 | 3000000 | 1000 |
| 天气查询 | 100000 | 100 | 3000000 | 1000 |
| 云图存储 | 2000 | 20 | 300000 | 200 |
| 云图搜索 | 2000 | 20 | 300000 | 200 |

# 附录2



# 附录3



1.500POI选取合理性分析

2.1000m判断合理性

3.返回是小区名还是楼宇名（楼宇名汇聚算法）

4.重要程度打分规则及数据需求来源（多维度）

5.经纬度研究

6.实际覆盖居民区的小区列表排序（栅格）

数据库格式

Git

4台服务器

2个公网IP