

文章编号: 1673-9469(2017)04-0094-05

doi:10.3969/j.issn.1673-9469.2017.04.021

基于互联网开放数据的城市实时交通圈提取及分析

张龙深, 葛星, 刘迪

(河海大学地理信息科学与工程研究所, 南京 211100)

摘要: 以北京市为研究区, 结合 Python 语言获取在样本点位置百度地图 API 所提供的实时交通数据, 通过空间插值的方法将样本点转化为可达性面, 并将面数据进行分级处理, 最终获得交通圈数据。通过北京市公交刷卡数据来对获取的时空圈进行置信度评价, 以此为依据, 划分驾车和公共交通两种模式下的高峰和非高峰时空圈。结果表明: (1) 获取的交通数据与刷卡数据有较强的相关性, 说明通过百度地图 API 获取实时交通数据方法可行; (2) 比起传统的交通圈获取方法, 利用百度地图 API 获取的数据能够更准确更实时地反映交通圈信息; (3) 相比驾车模式、北京市公共交通模式下的交通圈在偏向式延伸方面支撑力较差; (4) 公共交通模式获取的时空圈考虑到步行的距离和时间, 这对未来研究城市内部公共交通分布的均衡性提供依据。

关键词: 百度地图 API; 城市交通圈; 范围; 结构

中图分类号: U12

文献标志码: A

Analysis and extraction of urban real - time traffic circle based on internet open data

ZHANG Longsheng, GE Xing, LIU Di

(Institute of Geographic Information Science and Engineering research, Hohai university, Nanjing 211100, China)

Abstract: In this paper, taking Beijing as the research area, the Python language is used to obtain the real-time traffic data provided by the Baidu map API at the sample point location, then, the sample points are transformed into reachability surface by spatial interpolation, and the surface data is graded. Finally, we get traffic data. Through the Beijing bus credit card data to obtain the space-time circle for the evaluation of confidence, as a basis the two modes of driving and public transport under the peak and non-peak space-time circle are divided. The results show that: ① The obtained traffic data has a great correlation with the credit card data, which shows that the real-time traffic data method can be obtained through the Baidu map API. ② Compared with the traditional traffic circle acquisition method, the data obtained by Baidu map API is more accurate and more real-time traffic circle information; ③ Compared to the driving mode, the Beijing public transport mode traffic circle is in the bias-type extension of the poor support; ④ Public transport mode takes into account the distance and time, which will provide a basis for the future study on the balance of the distribution of public transport in the city.

Key words: Baidu map API; city traffic circle; scope; structure

等时间交通圈集成时间与空间两个维度, 一般指从中心地出发, 在一定时间阈值内能够到达的空间范围, 是交通基础设施对城市与区域发展的引导、支撑与保障能力的直观反映^[1]。国内学者曾经通过

通勤^[1-4]以及经济距离^[5-6]这两种方面来对研究区域交通圈进行综合评价: 黄翌等采用基于路网的网络分析分别对上海主城区、京津冀城市群、厦漳泉城市群进行交通圈提取分析^[1-4]。但基于实时拥堵度的

收稿日期: 2017-09-25

基金项目: 国家自然科学基金资助项目(41201394)

作者简介: 张龙深(1993-), 男, 江苏盐城人, 硕士研究生, 主要研究方向是地图学与地理信息系统。

交通圈分析，目前国内学术界尚未有更好的方法进行量化及提取方法^[7-8]，为此，本文以北京市为例，结合百度地图 API 所提供的实时路况数据，通过 Python 语言和 ArcGIS 软件对实时交通圈进行提取和分析。

1 研究区概况

本次研究区为北京市市域，总面积约 16 410 m²，常住人口 2 170.5 万人，GDP 22 968.6 亿元。北京是全国交通最发达的城市之一，全市公交线路共有约 700 余条、地铁 19 条线路，运营里程约 600 km，六环内公交站约 4 000 个，在运营出租车约 6.6 万余辆。

2 方法与数据

2.1 百度地图 API

本次用到的互联网开放数据是百度开放平台提供的接口：Direction API，它是一套以 HTTP/HTTPS 形式提供的公交、驾车、骑行、步行查询检索接口，根据请求接口参数，可返回 xml 或 json 格式的检索数据。其主要参数如表 1 所示。

通过输入参数后返回的 json 格式数据格式如下：其中，distance 代表从起点到终点经过路径规划

```
" result: " {
  " routes: " [
    {
      " distance " :5817,
      " duration " :975,
      " steps " :[ " ..."],
      " toll " :0,
      " originLocation " :{ " ..."},
      " destinationLocation " :{ " ..."},
    }
  ],
}
```

后的最佳路径长度，其单位为 m。duration 代表经历的时间，单位为 s。

由于 Direction API 这个接口是基于百度地图所监测的实时路况进行计算规划，因此返回的最佳路径以及最短时间在不同的时段是考虑拥挤度后计算的，可以借此方法来对研究区不同时段的时空圈进行计算。

2.2 样本点选取及数据获取

本次研究以天安门为中心，0.005 经度和纬度间距构建一万个样本点阵。由于样本点数过于庞大，使用人工处理较为繁琐，因此通过 Python 语言的 requests 包向 Direction API 传递参数，将返回的 json 数据进行读取存储，最终获得每个样本点实时到达中心点的路径距离和时间。采用 Python 并行运算的处理方式，获取一万个样本点数据的时间约为 15 min，可以认为是实时的交通数据。

2.3 数据插值

由于获取的数据是点数据，为了使其更能完整地表达出研究区的交通圈特征，需要对点进行插值处理从而获得面数据。考虑到样本点的数目以及其密集程度，本次采用普通克里金方法对其进行插值处理。其主要原理是考虑样本点的属性和样本点的空间位置的变异分布，通过分析样本点之间的空间相关性，选定一个对未知点具有影响的范围，范围内的所有样本点共同确定其属性值，其最基本的表达式如下：

$$Z(X_i) = \sum_{j=1}^n \lambda_j Z(X_j)$$

其中：Z(X_i) 代表 X_i 处插值的结果，Z(X_j) 代表 X_j 处的观测值，λ_j 为克里金系数，n 为样本点的个数。

2.4 公交刷卡数据

为证明所获取数据的可靠性，本次研究通过北

表 1 Direction API 参数及含义
Tab.1 Parameters and meaning fo direction API

参数	参数含义
origin	起点名称或经纬度
destination	终点名称或经纬度
mode	出行模式，包括：driving(驾车)、walking(步行)、transit(公交)、riding(骑行)
output	表示输出类型，可设置为 xml 或 json，默认为 xml
coord_type	坐标类型，可选参数，默认为 bd09ll。允许的值为：bd09ll(百度经纬度坐标)、gcj02(国测局加密坐标)、wgs84(gps 设备获取的坐标)

京市六环内公交车刷卡数据来进行回归分析验证,北京市内公交站点分布如图1所示。北京公交实行上下刷卡制度,通过刷卡数据可以得到从上车站点到下车站点期间经历的时间,本次同样以天安门公交站点为起点,计算从该站点出发到其他站点所经历的时间,若出现站点不能一次性到达的情况时,则计算到该站点换乘的最短距离并将换乘期间的的时间相加,即可取得以天安门公交站为起点到达其他各个公交站点所经历的时间。

3 结果与分析

3.1 置信度分析

通过 ArcGIS 的提取分析,提取出公交站点所在位置的非高峰期交通圈数值,与公交可达性时间大小作为顺序与该数据进行对比,结果如图2所示:

将二者进行回归分析得:

从图2中可以看出,在30min内,公交刷卡时

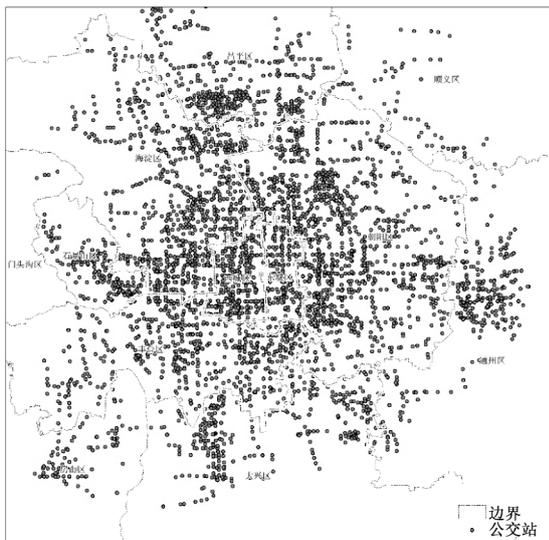


图1 北京市内公交站点分布
Fig.1 Distribution of Beijing bus station

表2 回归分析结果
Tab.2 Regression analysis results

回归结果参数	值
Multiple R	0.733 214
R Square	0.537 603
Adjusted R Square	0.537 494
标准误差	17.872 71
P-value	6.3E-45

间曲线低于基于百度地图计算的可达时间拟合曲线,这可能是由于刷卡数据仅考虑上车时间,未考虑步行时间,而百度地图规划可达时间是加上计算的步行时间的。当时间超过50min后,基于百度地图计算的可达时间拟合曲线低于公交刷卡时间曲线,这可能是由于百度地图API的公共交通方式采用公交和地铁混合计算,越远距离,地铁的优势度升高,到达时间比单纯公交线路快。

由回归结果可以看出(表2),测定系数(R Square)约为0.54,代表二者拟合较好,置信度(P-value)约为0,代表二者关系具有很高的相关性。标准误差较高的原因是由于百度地图数据计算时添加了到各个位置的步行时间以及混合了地铁计算,以及二者数据虽然是在相同时间段,但日期不同,可能道路通达性不同等。

3.2 交通圈分析

本次获取和处理的可达性数据时间段分为高峰期(18:00—18:30)和非高峰期(20:00—20:30)两种情况,出行模式选择驾车出行和公共交通出行。通过对4种情况点数据进行普通克里金插值,采用30min为分级界限分为五个等级,其结果如图3、图4所示:

以点阵的范围做边界(5331km²),统计驾车和公共交通模式在30min以及60min两种情况交通圈面积占比,如表3所示:

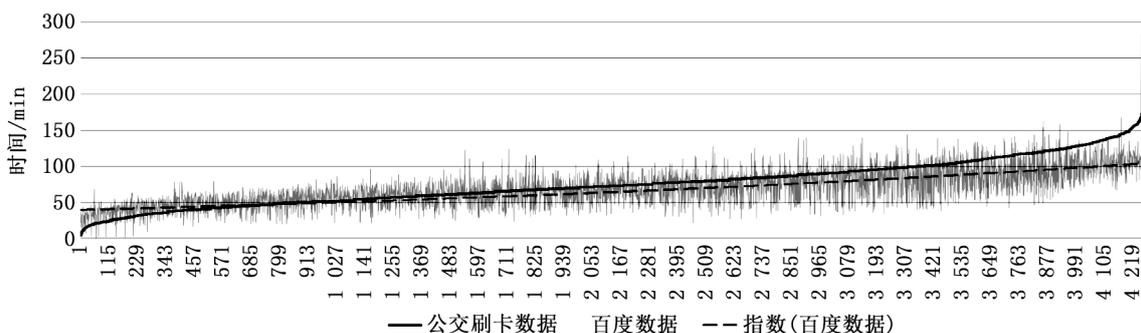


图2 刷卡数据与交通圈数据对比(非高峰)
Fig.2 Comparison of credit card data and traffic circle data(non-peak)

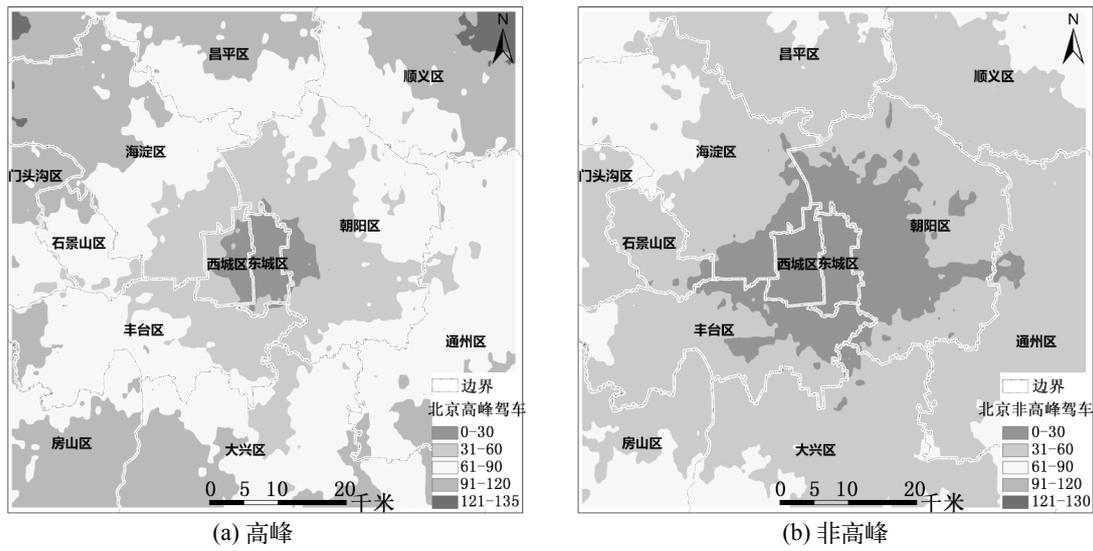


图 3 北京市驾车交通圈
Fig.3 Beijing driving traffic circle

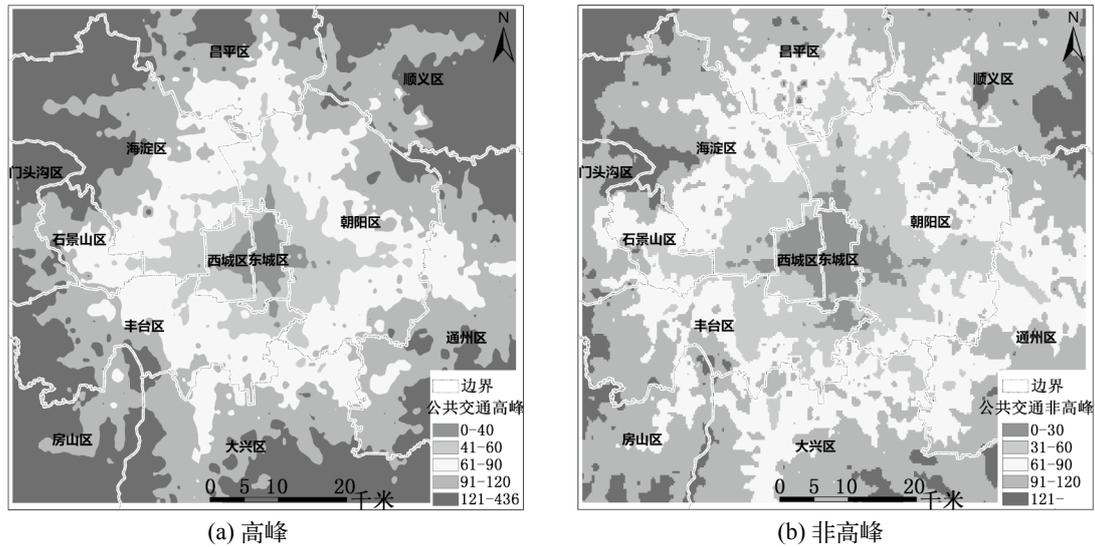


图 4 北京市公共交通交通圈
Fig.4 Beijing public transit traffic circle

结合图表我们可以看出：

(1) 基于驾车模式插值生成的可达性交通圈，其两种形态都偏向于圆形，说明了北京市在道路方面规划较为均衡，车辆在 30 min 内到任何方向的距离差距不大。但 30 min 内高峰和非高峰交通圈面积差距有约 6 倍左右，表明高峰时期对于驾车来讲影响较大。

(2) 基于公共交通模式插值生成的可达性交通圈，其交通圈在高峰和非高峰期形态偏向于菱形，这可能是由于公共交通模式混合有地铁方式，北京市地铁大多数是横纵分布，对于横向和纵向位置可以较为便捷到达，公交到达斜向较为方便，但相比地铁来说较慢。高峰与非高峰时期的交通圈面积变化相对于驾车模式不大，非高峰面积约为高峰时期

表 3 不同模式 / 时间下交通圈面积对比
Tab.3 Comparison of traffic circle in different modes / times

类型	30 min 内可达面积 /km ²	占比 /%	60 min 内可达面积 /km ²	占比 /%
高峰驾车	148.071	2.79	1 018.954	19.114
非高峰驾车	767.496	14.39	4 508.203	84.566
高峰公共交通	89.908	1.69	485.560	9.108
非高峰公共交通	188.307	3.53	935.738	17.553

面积的2倍。

(3) 驾车模式和公共交通模式二者对比可以看出,在30 min的高峰期内,驾车可达的范围与公共交通可达范围其实差距不大,即当高峰期时,在30 min内的路程可以考虑公共交通代替驾车。但在60 min内,驾车模式时空圈相比于公共交通模式存在优势,其二者面积相差约1倍,即在30 min以上的路程,驾车比乘坐公共交通工具具有更大的优势。

4 结论

1) 通过公共交通模式时空圈提取出的数据与公共刷卡数据进行对比,发现其相关性较大,说明基于百度地图API获取的实时数据方法可行。

2) 与网络分析法等传统方法相比,利用百度地图API获取的可达性数据可以更准确地描述该区域的交通圈形态、范围和结构。并且能在不同的时段获取实时的交通可达数据,使得数据更精细,更具有时效性,为未来城市内部交通研究提供了新的思路。

3) 北京市驾车模式下的交通圈形态偏向圆形、公共交通模式下偏向于菱形,表明北京市公共交通方面对于斜向的交通支撑力较弱。在高峰时期,驾车模式相比较公共交通模式在30 min内交通圈相差并不大,但随着时间增长,驾车模式交通圈扩张更快。公共交通模式下交通圈较为稳定,受高峰拥挤度影响较小。

4) 获取的公共交通实时数据考虑到了步行的距离和时间,这将有利于分析城市内部的公共交通分布的均衡性。

本研究还存在一些不足:由于API每天配额有限,在0.005经纬度间隔密度下只取了一万个采样

点进行的研究,使得研究区域并不能完全覆盖,并且需要15 min获取全部数据,未来应增加采样点的数量并优化算法;数据样本点研究只采用一天的数据,这可能会受某些其他因素造成结果有偏差,或考虑对多天进行关联分析;未对样本点数据进行清洗,这可能导致部分样本点因位于交通不便处使得数值异常,造成插值结果不准确。

参考文献:

- [1] 陈卓,金凤君.北京市等时间交通圈的范围、形态与结构特征[J].地理科学进展,2016,35(3):389-398.
- [2] 黄翌,李陈,欧向军,等.城际“1小时交通圈”地理学定量研究:以上海主城区为例[J].地理科学,2013,33(2):157-166.
- [3] 王德,郭玖玖.北京市一日交流圈的空间特征及其动态变化研究[J].现代城市研究,2008(5):68-75.
- [4] 曾月娥,伍世代.厦漳泉城市群小时交通圈研究[J].地理科学进展,2016,35(8):975-982.
- [5] 吴新文,罗阳辉.哈尔滨一小时经济圈空间范围的界定研究[J].经济论坛,2012(11):75-78.
- [6] 刘迪.基于路网可达性的中国新城新区公共服务均衡性水平的度量——以医疗保障为例[J].河北工程大学学报:自然科学版,2017,34(2):91-94.
- [7] BENENSON I, MARTENS K, Rof Y, et al. Public transport versus private car GIS-based estimation of accessibility applied to the Tel Aviv metropolitan area[J]. The Annals of Regional Science, 2011, 47(3): 499-515.
- [8] CAO X, CHEN H, LINNA L I, et al. Private Car Travel Characteristics and Influencing Factors in Chinese Cities——A Case Study of Guangzhou in Guangdong, China[J]. Chinese Geographical Science, 2009, 19(4): 325-332.