

大学生论文检测系统

文本复制检测报告单 (全文标明引文)

No: ADBD2023R_20230530172202472626297278

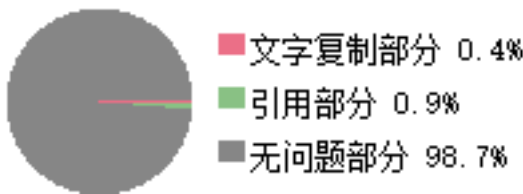
检测时间: 2023-05-30 17:22:02

篇名: 基于手机数据的城市空间结构分析和可视化感知
 作者: 林世豪
 指导教师: 宋轩
 检测机构: 南方科技大学
 文件名: 11810613_林世豪_毕业论文_中文.pdf
 检测系统: 大学生论文检测系统
 检测类型: 大学生论文
 检测范围: 中国学术期刊网络出版总库
 中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库
 中国重要会议论文全文数据库
 中国重要报纸全文数据库
 中国专利全文数据库
 图书资源
 优先出版文献库
 大学生论文联合比对库
 互联网资源(包含贴吧等论坛资源)
 英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)
 港澳台学术文献库
 互联网文档资源
 源代码库
 CNKI大成编客-原创作品库
 机构自建比对库
 时间范围: 1900-01-01至2023-05-30

检测结果

去除本人文献复制比: 1.3%
 去除引用文献复制比: 0.4%
 单篇最大文字复制比: 0.4% (广电网参与智慧城市建设的现状及发展空间分析)
 跨语言检测结果: 0%
 总文字复制比: 1.3%

重复字数: [176]
 总字数: [14028]
 单篇最大重复字数: [60]
 疑似段落最大重合字数: [120]
 疑似段落最小重合字数: [56]
 总段落数: [2]
 疑似段落数: [2]
 前部重合字数: [120]
 后部重合字数: [56]



指标: ☐ 疑似剽窃观点 ☒ 疑似剽窃文字表述 ☐ 疑似整体剽窃 ☐ 过度引用

相似表格: 0 相似公式: 没有公式 疑似文字的图片: 0

1.2% (120) 1.2% (120) 基于手机数据的城市空间结构分析和可视化感知_第1部分 (总9645字)
 1.3% (56) 1.3% (56) 基于手机数据的城市空间结构分析和可视化感知_第2部分 (总4383字)

(注释: 无问题部分 文字复制部分 引用部分)

指导教师审查结果
指导教师： 宋轩
审阅结果：
审阅意见： 指导老师未填写审阅意见

| | | |
|------------------------------------------------|------------------------------------------------------------------|----------------------|
| 1. 基于手机数据的城市空间结构分析和可视化感知_第1部分 | | 总字数：9645 |
| 相似文献列表 | | |
| 去除本人文献复制比：1.2%(120) 文字复制比：1.2%(120) 疑似剽窃观点：(0) | | |
| 1 | 广电网络参与智慧城市建设的现状及发展空间分析 张海耀;张益斌; - 《城市建设理论研究(电子版)》- 2022-12-15 | 0.6% (60) 是否引证： 是 |
| 2 | 基于广电网络5G技术的智慧城市建设分析 阎作铃; - 《电视技术》- 2022-02-15 | 0.6% (57) 是否引证： 否 |

原文内容

分类号编号
U D C 密级
本科生毕业设计（论文）
题目： 基于手机数据的城市空间结构分析和可视化感知
姓名： 林世豪
学号： 11810613
系别： 计算机科学与工程系
专业： 计算机科学与技术
指导教师： 宋轩副教授
2023 年 6 月 2 日
诚信承诺书
1. 本人郑重承诺所呈交的毕业设计（论文），是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料均真实可靠。
2. 除文中已经注明引用的内容外，本论文不包含任何其他人或集体已经发表或撰写过的作品或成果。对本论文的研究作出重要贡献的个人和集体，均已在文中以明确的方式标明。
3. 本人承诺在毕业论文（设计）选题和研究内容过程中没有抄袭他人研究成果和伪造相关数据等行为。
4. 在毕业论文（设计）中对侵犯任何方面知识产权的行为，由本人承担相应的法律责任。
作者签名：
2023 年 6 月 2 日基于手机数据的城市空间结构分析和可视化感知林世豪
（计算机科学与工程系指导教师：宋轩）
[摘要]：城市作为当前社会的主要活动场所，对空间结构的研究已经是一个不可忽视的课题。合理的城市空间结构，在城市健康发展中是不可忽略的重要条件，而城市空间结构分析应该与居民的活动有机结合。
本文使用手机信令数据，通过栅格化的方法，在保护居民个人隐私的同时，通过行为特征分析，分析出每个时间段有意义的活动，并且将这些活动设置为栅格的属性。通过相同属性栅格的聚合，分析空间地理位置的特征，得到城市空间的结构。手机信令数据记录了个人在每个时间点在哪里的大概信息，通过这些信息，使用栅格化算法，将个人的位置模糊到栅格当中，从而保护个人隐私的同时获取有意义的特征。
本文使用前端可视化的方法，将进行数据分析后得到的城市特征，通过地图可视化的形式呈现给使用方，极大的降低了使用方判断城市结构的成本。本论文将主要分析杭州市的城市结构，并且使用前端技术可视化地展示杭州市的城市空间结构。
[关键词]：栅格化；城市结构；可视化
I
[ABSTRACT]: As the main venue of current society, the study of spatial structure is already a topic that cannot be ignored. Reasonable urban space structures are important conditions that cannot be ignored in the healthy development of urban urban, and urban space structure analysis should be organically combined with residents' activities. Therefore, it is a scientific and reasonable analysis method through data analysis of the activity space attributes of residents and analyzing urban space structures, which can provide theoretical support for the optimization of the space structure.

This article uses mobile signaling data. Through the method of Rasterization, while protecting the personal privacy of residents, the experiment analyzes meaningful activities per time through behavioral characteristics, and sets these activities to the attributes of grid. Through the aggregation of the same attribute grid, analyze the characteristics of the geographical location of the space, and obtain the structure of the urban space. The mobile signaling data records the approximate information of the individual at each point in time. Through this information, using the grid algorithm to blur the personal location into the grid, thereby protecting personal privacy Essence.

This article uses the front-end visualization method, and the urban characteristics obtained after the data analysis are presented to the user through the form of the map visualization, which greatly reduces the cost of determining the urban structure of the user. This thesis will mainly analyze the urban structure of Hangzhou, and use the front -end technology to visually display the urban space structure of Hangzhou.

[Key words]: Rasterization, City Structure, Visualization

II

目录

| | |
|-------------------------------|----|
| 1. 绪论 | 1 |
| 1.1 背景 | 1 |
| 1.2 文献综述 | 1 |
| 1.3 研究方法和技术路线 | 2 |
| 1.4 研究的目的和意义 | 4 |
| 2. 数据介绍 | 4 |
| 2.1 数据获取与数据格式 | 4 |
| 2.2 手机信令数据的缺陷 | 6 |
| 2.3 poi 数据 | 7 |
| 3. 数据分析与处理 | 7 |
| 3.1 手机信令数据预处理和栅格化处理 | 8 |
| 3.2 停留活动的识别 | 9 |
| 3.3 移动识别 | 10 |
| 3.4 居住地点的识别 | 11 |
| 3.5 工作地点的识别 | 12 |
| 3.6 poi 地点统计 | 13 |
| 4. 可视化系统 | 13 |
| 4.1 设计的原则和目标 | 13 |
| 4.2 可视化系统介绍 | 14 |
| 4.3 视图和功能 | 15 |

III

| | |
|------------------------|----|
| 4.4 可视化的结果分析 | 20 |
| 5. 结论和展望 | 21 |
| 参考文献 | 22 |
| 致谢 | 24 |

IV

| | |
|-------------|--|
| 1. 绪论 | |
| 1.1 背景 | |
| 1. 国家战略的支持。 | |

随着智能化和信息技术的快速发展，城市数据和信息得到了快速积累与快速生成，就数据处理和利用提出了巨大的挑战。近年来，我国的信息化水平快速提升，已经成为了促进社会发展主要的因素之一，因此，人们对于信息化技术的运用的问题也越来越重视。

并且随着近年来大数据分析的火热，我国近年来也一直致力于推进大数据分析下的城市空间建设，特别是智慧交通[1]、安全等方面。并且虽然工程复杂，国家重要的网电机构也一直致力于在城市建设中做出贡献，[2]。并且我国发展改革委员会等部门于 2022 年 1 月联合发布的《“十四五”公共服务规划》明确，要加强智慧广电基础设施建设，推进城市的建设发展。在未来，大数据下的城市建设会是我国的重点方向。

2. 数据可视化发展

随着信息时代的到来，人们面临着海量的数据和信息。为了更好地理解和利用这些数据，人们需要有效的方式来展示数据。此时，数据可视化便成为了一个必要的手段。随着计算机图形学和数据可视化技术的发展，人们可以更加容易地展示和处理数据。从最初的简单的散点图到现在的复杂的网络图和三维图形，数据可视化技术不断地发展和壮大。随着信息化水平越来越高，数据积累和收集的能力越来越强，并且芯片的运算速度和渲染能力都达到了极高的程度，可视化工程越来越火热。同时阿里等各大公司都有推出相对应的可视化产品。将智慧城市与可视化结合，会是未来的一个大的趋势。

1.2 文献综述

城市空间结构的解析和建设,离不开居民活动空间的分析。居民活动空间定义的是个人进行大部分日常活动的空间,一般由个体日常活动过程中所经过的所有地点的集合表示[3]。在早期活动空间研究的数据来源一般为传统问卷调查,但是这种数据忽视了居民的个体差异性[4]。但是随着移动定位技术的不断发展与广泛应用,基于 1

GPS、社交网站签到、手机等数据已被广泛应用于居民日常活动与出行行为的研究中。这些数据拥有覆盖范围广、持续时间长,并且时间空间层面连续等优势,为居民活动分析提供了足够的技术支持。

在城市活动空间的影响方面,国内外已经有大量的研究。其中包括:人口与环境因素[5]、人口种族因素[6]、居民家庭经济属性[7]、不同城市属性[8]、工作日和休息日[9]、汽车拥有量[10]等。这些研究表明,居民的出行活动非常复杂,受到的影响因素众多。

其中手机通信数据分析目前是一个大方向,有国外研究者研究了十万个用户六个月的手机数据[11],通过一定的算法分析,获得一个结论:个人出行模式可由概率分布描述。这个结论表明个人的行为有一定的规律性在国际上有较大影响力。也有人利用手机数据研究了用户出行的可能性,最终实验获得的模型预测居民出行意愿准确度高达 93%[12]。

对于手机数据的真实性也是实验结果可靠的保证。有学者对手机数据描述真实信息的准确性进行验证,他们通过调查生产的模拟数据和手机数据进行对比,并且与真是情况进行比较,得到一个结果:手机数据所识别出的居住地分别有 70% 与 97% 在真实位置的 100 和 1000 米内[13]。因此手机数据的准确性是有保障的。

1.3 研究方法和技术路线

在判别城市结构这过程中,个人的信息就尤为关键。对一个个体的跟踪,可以得知个体在城市的各个地点联系。我们可以跟踪个体在各个时间点处于什么位置,个人通常有上班,回家睡觉,通勤等需求,每个需求都对应当前城市空间的作用,并且出行等活动还连接了城市不同空间,反映了两个城市空间存在一定的联系,当个体数量

足够大的时候,数据结果就可以代表整个城市的内部的关系。比如某两个地点经常会出现人口流动,那么对则两个地点的交通,可以适当的优化调整。

本文使用的手机信令数据本身携带个人的位置信息,通过一定的算法分析,可以得到个人的活动空间信息[14]。但是这有出现另外一个问题,那就是个人的隐私安全。

如果无条件的跟踪个人的活动,那么对个体的隐私就会造成侵犯,而且持有个体信息的政府机关也不会支持这种侵犯隐私的行为。通常需要和政府机构签订一些协议, 2

才能够拿到城市居民的数据,而且这种数据是个体抽样的数据,是随机的数据,对个体的隐私有一定的保护性。在智慧城市的建设过程中,模型安全问题[15]一直都是不容忽视的,在保证系统安全、个人信息安全等的前提下,进行智慧城市模型的有效建设,才是有正向价值的研究,并且模型的建设本身就有有助于安全风险解决[16]。

本文通过对手机信令数据进行分析,得到一种栅格化[17]的数据,这种数据因为是以栅格的形式将多个用户的信息统计到一起,因此对单个用户的隐私有保护性,这种栅格化数据虽然丢失了一定的精度,但是却可以更加放心的供私营单位使用,并且多个用户信息统计为栅格的数据,极大的减少了数据量,在使用算法处理数据时更加的顺畅。

通过栅格化数据,我们可以依据一定的算法,分析出每个栅格格子的属性,当拥

有相同属性的栅格聚集在一起的时候,这片栅格所在的地理位置就可以判断为拥有

一定的属性,比如较为集中的居住地和工作地。并且通过手机信令数据的路径轨迹跟踪[18],可以获取城市交通路线网图,进一步精确城市的空间模型。

最终通过前端部分可视化的展示,可以很清晰的看到地理空间的相对关系,对城市的结构判断、建设等等都有一定的益处。如图1表示本研究流程框架。

本研究主要分为数据处理和前端可视化系统搭建两个部分,在数据处理中,使用的数据为手机信令数据。本研究将会对手机信令数据进行去重、栅格化等操作,最终获得新的数据,在保证个人隐私的同时,分析出城市空间中的居住地和工作地,并且通过移动手机,展示空间中不同地方的联系程度。最终在前端可视化系统上,将城市空间结构展示出来。

图 1 技术流程图 3

1.4 研究的意义

如今城市作为人群活动的主题空间,有着不可替代的重要性,是人类当下最为重要的活动点。对城市有一个清晰的认知,有利于城市结构的合理规划和建设。但是在研究城市结构时,不可以单独的分析城市的静态结构,而应该结合城市居民的实际情况,分析出城市内部和城市之间的有机结合关系。目前城市结构的研究已经是一个成熟的论题。随着世界人口的不断增长以及经济社会的快速发展,城市的规模和数量都在持续扩大。城市结构研究旨在探究城市内部空间的分布结构和特征,以及城市与周边区域的联系和互动关系,对于城市规划、城市管理、社会经济等领域具有重要的理论意义和实践价值。

但是在研究城市结构中,个人的隐私问题也是一个难点。如何在保护个人隐私的同时,识别城市空间与城市居民的有机关系,也是一个备受争议的话题。本文采用数据信息栅格化的方法,对隐私进行模糊处理,可以有效的保护个人隐私。

本文通过栅格化的技术,可以极大程度上保护个人隐私,因此数据的可适用性非常的广,并且通过栅格化的方法,将手机信令数据进行一定的算法分析,识别空间地位位置的特征,将地理空间和居民行为有机的结合在一起。最终通过前端系统可视化的呈现出来,研究者可以可视化的识别空间中各个位置的特征。极大的节省了使用者对空间地理信息判断的成本。

2. 数据介绍

2.1 数据获取与数据格式

移动通信系统由移动终端(Handsets)与基站(Base Station)组成。基站有规律的分布在城市空间中,中心区密集而郊区稀疏,每个基站都装有更低功率的无线电发射器和接收器,通常认为每个基站覆盖区域的边界呈六变形,即蜂窝。一个通讯网络有很

多个蜂窝组成,每个接收一个固定信号发射器(称为基站收发信台,BaseTransceiver-Station, BTS)的服务。BTS 之间互相连通,因此当手机从一个蜂窝移动到另一个蜂窝时不会失去信号,如图2表示基站示意图。 4

图 2 基站示意图手机信令数据是本研究的基石,我所使用的数据是手机与移动基站进行交互的时候,产生的信令数据。除了常规的打电话,发短信,在开机、关机、连接基站变更等场景都会产生手机信令数据,能够很好的记录个人的移动轨迹。这

个数据会记录用户与基站的交互发生的时间，基站自己的地理位置，比如图 fig.3表示原始的手机信令数据。表格1为数据每个字段的含义。

表 1 手机信令数据字段

字段名意义

user_id 用户名

stime 时间

longitude 经度

latitude 纬度这种数据是个体手机与信号基站发现交互时就会记录下来。通常 1 左右分钟就会记录一次。userid 代表的是个体的 id，同一个 userid 代表相同的个体，stime 代表的是时间，longitude 和 latitude 代表的是基站的经度和纬度，因此这个数据可以记录一

个人在某个时间的大体位置信息。当信息连续，而且数量足够多的时候，改数据可以表示某个人的活动轨迹，本研究使用的是杭州市 2018 年 6 月份的信令数据。 5

图 3 手机信令原始数据示意图

2.2 手机信令数据的缺陷

对于这种手机信号数据，有一个缺陷，那就是手机有时候会同时接受到多个基站的信号，这就意味着多个基站在相近的时间内，记录了手机的信号，如图4所示。因此记录的数据可能会表示某个人在极端的时间内迅速移动的了位置，因为这个人的手机信令数据表示的经纬度坐标在几个基站之间来回切换。不过对于这个问题，使用的解决方法就是栅格化分析，来将用户的位置信息合并到一个栅格中，因为切换的基站大多数都会落在同一个栅格中，即使发送乒乓切换[19]的问题，也会被栅格化处理掉，视为没有发生变化。

图 4 基站同时覆盖居民 6

2.3 poi 数据

对于能耗的排放估算，需要使用一种具有地点建筑信息的数据，这种数据被称为

poi 数据，全称为 point of interest。这个数据中，记录了某个建筑所在的地理位置。比如某个公司的位置，或者某个图书馆的位置。同样的这个数据也要进行栅格化处理，然后统计出每个栅格上面的所有建筑信息。比如经过统计，可以得到在某个栅格空间里面，存在公司 10 个，政府机关 3 个等等。

3. 数据分析与处理

数据分析时通过数据自身携带的信息，通过一定的特征分析、算法分析，得到有意义的信息。在本研究中，我对数据信令数据进行一定的筛选、加工、分析，最终获得对应的数据，如图5表示对数据分析的流程。通过对数据信令数据进行栅格化，消除个人隐私信息后，对数据进行分析，得到停留活动数据和移动数据。再将停留活动数据进行二次分析，区分那些活动是工作地的活动，那些是居住地方活动。最终得到城市空间中居住地和在工作地的信息。将停留活动数据和 poi 数据联合起来，经过聚合，分类等操作，得到栅格能耗基础数据，这个数据包含了每个栅格包含的建筑数量信息和居民在这个栅格内活动的总时间。

图 5 数据分析流程图 7

3.1 手机信令数据预处理和栅格化处理

因为手机信号是不定时的被多个基站记录的，所以会出现没有意义的相同数据，这种一样的数据是因为同一个基站在同一分钟之内记录了两次数据，也可能是因为不同的基站在同一分钟之内记录了两次数据。对于这种数据，对相同的 userid 和 stime

进行剔除，只需要随机保留一条即可。因为在栅格化的过程中，地理空间位置会被整合到栅格中，最终表示为栅格中心点的位置，因此不必太关注单个时间点的地理空间的位置。

对于我们所得到的手机信号数据，我们并不能直接使用，因为这是一个涉及隐私的数据，通常运营商也不会将如此携带隐私的数据提供给研究机构使用，而且还有乒乓切换的问题需要解决，因此需要对数据进行栅格化处理。

我们会使用一种栅格化的思维来处理数据。栅格化是指选定一个空间坐标点，以这个点为中心，按照经纬度的方向画一个正方形的格子，这个格子的大小可以是

1000m，也可以是其他大小，比如 500 米。如果一个人的位置落在这个格之内，那么就重置这个人的位置，将这个人的位置移动到栅格的中心点。因此在栅格化后，个人的具体位置信息会被隐藏掉，取代的是栅格中心点的位置。因此栅格化是对个人隐私的一种保护。如图6为栅格化的示意图。因此城市空间可以本分割为多个栅格，个人在城市活动时，就可以将个人的位置信息都用栅格的中心点坐标代替。对于这种数据，隐私性较弱，更容易被接受用来做数据分析。

图 6 栅格化 8

3.2 停留活动的识别

在获得清洗后的数据之后，我们需要对个人的停留活动进行识别。

那么为什么要识别停留活动呢？因为停留活动证明了这个栅格所在的城市空间的价值。比如一个人在晚上的时间，在一个地点停留了很长时间，那么可以推断目前的空间是这个人的居住地，因为在晚上时间，大部分人都会在居住地中睡觉，那么这个栅格的特征可以标记为居住地，这个空间的结构便是居住地。当样本是随机分布，并且数量足够大的时候，可以识别出那个栅格是城市人群集中的分布地。同理我们也能识别出集中的工作地等等。

识别一个人是否处于停留活动的重要指标是一个人在同一个格子持续停留时间。

比如我们以 30 分钟为标准，如果一个人人在同一个栅格停留时间超过 30 分钟，那么就可以判定这个人正在这个栅格停留，并且进行一些有意义的活动，比如工作或者吃饭。

当然，这种识别方法也有一定的缺陷，那就是有的人经过一个格子时，并不想在这个格子中停留，但是因为个人某些原因，移动速度较为缓慢，导致持续在一个栅格中停留时间超过 30 分钟，这样我们就会误判，认为这个人正在进行有意义的活动。

但是在样本较大的时候，这种影响对实验结果就可以忽略。因为在栅格不是很大的情况下，一直移动但是却 30 分钟都没走出格子的情况非常少。而且我们可以动态的指

定栅格的大小，也可以指定停留判定的时间界限，这样可以对数据结果进行多样的分析。

最终数据会被转换为另外一种格式。如图7为停留活动数据的示例，表2为数据每个字段的含义。 9

图 7 停留点活动数据示例表 2 停留点活动数据字段

| 字段名 | 意义 |
|----------|---------|
| user_id | 观察者的标识 |
| stime | 活动开始时间 |
| etime | 活动结束时间 |
| LONCOL | 栅格经度编号 |
| LATCOL | 栅格纬度编号 |
| lon | 栅格中心点经度 |
| lat | 栅格中心点纬度 |
| duration | 活动持续时间 |

3.3 移动的认识

移动数据的识别最终获取到 OD 数据，在前端可视化系统上，出行视图使用的 geojson 数据便是移动数据转化而来的。

本研究将移动定义为在两个活动之间的移动，因此移动的特征是除了停留活动之外的时间段，都可以视为移动。因为停留活动必定在同一个栅格中，因此在两个停

留活动的地点肯定不是同一个栅格，对于两个活动位置的变更，可以理解观察对象从一个栅格移动到了另外一个栅格，即从一个地点移动到了另一个地点。

那么移动开始的时间，就是第一个停留活动结束的时间，而移动结束的时间，就是第二个停留活动开始的时间。就这样，我们可以确定移动的起始地点和移动的结束地点，同时也能确定移动的起始时间和移动的结束时间。

同样我们需要将个人移动的数据进行栅格化，将移动的起始点和终点都转换为格子的中心点。并且统计相同的起始点和终点有多少条数据，在数据较大的时候，可以推断这两个地点联系较为密切。 10

因此最终可以获得个人的移动数据，表3表示获取数据的字段信息。接下来对个人的移动数据进行进一步的处理，通过对移动的起始栅格和结束栅格进行分组，并且统计出在这两个栅格之间移动的次数，可以获得在两个栅格之间进行移动的总次数，这个总次数标为字段 count。因此通过这个处理，可以获得城市的两个地理空间位置移动联系的情况，当来回移动次数越大时，表示这两个位置存在的联系越大。

表 3 个人移动数据字段

| 字段名 | 意义 |
|----------|-----------|
| user_id | 观察者的标识 |
| stime | 移动开始时间 |
| etime | 移动结束时间 |
| SLONCOL | 开始栅格经度编号 |
| SLATCOL | 开始栅格纬度编号 |
| slon | 开始栅格中心点经度 |
| slat | 开始栅格中心点纬度 |
| ELONCOL | 结束栅格经度编号 |
| ELATCOL | 结束栅格纬度编号 |
| elon | 结束栅格中心点经度 |
| elat | 结束栅格中心点纬度 |
| duration | 移动持续时间 |

3.4 居住地点的识别

居住地的识别是在活动数据的基础上进行的。通过停留活动的开始时间、结束时间、持续时间等等信息，识别个人的居住地点。

居住地点有一个特征，那就是在晚上的时间，居民通常会睡在家里睡觉，而且这个时间通常都会非常长。我们可以根据这个特征，识别出那些栅格是居民的集中居住地点。

因此当获取到个人的活动数据后，首先需要筛选出某个活动在晚上的持续时间，因为有的活动可能会从白天一直持续到晚上，再从晚上持续到白天。

这里有个关键点参数，那就是晚上的定义。因为每个地方的生活习性不同，晚上的定义也会不同。有的地方下午 5 点后就下班了，上午 9 点才去上班，但是有的地方通常要下午 7 点才会下班，上午 8 点就要上班了，因此在数据分析时，可以动态地指定数据。这使得本研究有一定的灵活性和通用性。

对于时间的记录，可以给格子分配一个参数，这个参数用来记录某个人在晚上的 11

时间在这个格子中进行停留活动的总时间。那么当一个人在一个月内的活动都统计完成后，会有一个格子的参数的时间是最长的，那么这个格子就是这个人的居住地，因为居住地通常就是居民晚上停留时间最长的地点。

之后可以统计所有样本居民的居住地的栅格，将这些居住地的栅格进行聚合统计，统计出每个居住地格子的居住人口数量，较高的数量表面这个栅格是较为集中的居住地。如表4表示该数据的字段和意义。

表 4 居住地聚合数据

| 字段名 | 意义 |
|--------|--------|
| LONCOL | 栅格经度编号 |
| LATCOL | 栅格纬度编号 |

lon 栅格中心点经度
lat 栅格中心点经度
count 以此栅格为居住地的人数
3.5 工作地点的识别
工作地的识别也是在停留活动数据的基础上进行的。

| | | |
|----------------------------------------------|-----------------------------------------------------------------------------------------------|---------------------|
| 2. 基于手机数据的城市空间结构分析和可视化感知_第2部分 | | 总字数：4383 |
| 相似文献列表 | | |
| 去除本人文献复制比：1.3%(56) 文字复制比：1.3%(56) 疑似剽窃观点：(0) | | |
| 1 | 202005161541283499_无作者_高管团队特征对公司盈利能力的影响研究——基于中国上市公司数据 202082445 - 《大学生论文联合比对库》- 2020-05-16 | 0.8% (37) 是否引证：否 |
| 2 | 呼和浩特市2017-2021年中考化学试题深度研究 高梦静(导师：李玉珍) - 《内蒙古师范大学硕士学位论文》- 2022-06-30 | 0.7% (31) 是否引证：否 |
| 原文内容 | | |

方法和居住地的识别类似，但是使用的特征发生改变。居住地点有一个特征，那就是在工作日时，白天的时间段，观察对象会在工作地停留非常长的时间。因此与居住地的识别相似，需要将观察对象白天在每个栅格的活动时间统计出来，并且选择停留最长的栅格就行了。

但是对于白天的定义，在不同的城市、不同的时间，或许会有一定的差别。因此白天的定义可以提供多种选择，比如白天的开始时间可以设置为早上 7 点、8 点或者、9 点，而白天的结束时间可以定义为下午 5 点、6 点、7 点等等，这种可选择的设置能够在数据分析时，针对特定的人群，获得特定的数据。这增加了本研究的拓展性和功能性，在前端方面，白天定义的动态选择，也会有所体现。

同时需要处理一个问题，那就是有点居民是没有工作的。对于这样的人群，白天在一个地点停留的最大可能就是自己家里。因此在识别工作点的时候，应该将已经识别的居住地点排除。同时对于这个问题，还可以为工作点设置一个时间阈值，比如 5 个小时，在白天停留点最长的栅格内时间需要大于这个时间阈值，才能将当前的栅格定义为工作地点。然后同样对工作地点的栅格进行分组，再聚合统计出人数人数较多的栅格，这些 12 栅格就是较为集中的工作地点。如表5表示该数据的字段和意义。

表 5 工作地聚合数据

字段名意义

LONCOL 栅格经度编号

LATCOL 栅格纬度编号

lon 栅格中心点经度

lat 栅格中心点经度

count 以此栅格为工作地的人数

3.6 poi 地点统计

通过停留活动数据，可以得到在这个栅格当中，所有被观察对象在这个栅格中停留时间的总和。因此可以获取的数据字段如表6，其中 name 的值对应的建筑信息分别为公司、教育机构、娱乐场所、金融机构、政府机构、医院、酒店、自然保护区、居民住宅、旅游景点、餐馆、商店、体育馆、车站和交通要点。

表 6 poi 数据字段

字段名意义

name 各种建筑信息的个数

time 被观察者的停留活动总时间在前端部分，可以根据获取被观察者们在这个栅格的停留时间，并且对各种建筑信息进行加权计算，获得一个推理的能耗值。

但是目前的资料很难客观的表示每种建筑的能耗值为多少，因此在前端部分，我采用可以自定义输入的方式，将每种建筑的能耗比例使用动态指定的方式，进行计算。

4. 可视化系统

4.1 设计的原则和目标

前端可视化系统是为了将数据处理的结果通过图形的方式呈现出来。因此可视化系统应当充分发挥数据数据模型的所有优势，并且提供相对合理的功能，拓展可视化系统的能力。在使用上可视化系统提供相对统一的交互方式，并且采用分版面的方式将不同版面的交互隔离开来。如图8表示数据在和前端可视化系统交互过程中，各个数据和视图页面之间的相对关系。 13

图 8 视图页面与数据对照关系

4.2 可视化系统介绍

该可视化系统使用的前端框架为 vue 框架，采用的主要地图框架为由阿里提供的

antvL7。本研究的可视化系统采用标签页选择的方式展示不同的视图页面，在主要界面部分，采用左右结构的布局方式，左边是每个模块的选项和功能，右边是对应视图的地图信息，有点状图、线形图、柱状图等多种图形选择，并且可以自由选择图形的颜色。

本研究目前拥有出行视图、居住地视图、工作地视图和能耗热力图。在每个模块中，拥有相对应的功能选项。并且每个视图都是一个独立的模块，可以同时展示每个模块中的内容。每个视图也都提供了裁剪的功能，可以对特定区域的空间进行选择，在选择完成后，可以展示该空间的信息。如图9是目前的可视化系统的总体情况。 14

图 9 前端可视化系统

4.3 视图和功能

因为本研究拥有四种不同的视图界面，每种界面会有一些不同的功能，因此在这个模块，将一次简单介绍每个视图的信息和功能。

出行视图出行视图描述的是居民的移动情况，使用的是出行的 OD 数据。在 OD 数据中，主要体现的是出发点和终点，没有体现具体的活动轨迹。

在功能界面可以通过选择不同的图形类型，展示不同的图像，并且根据算法，可以动态的选择条个大小，活动时间等等。在可视化的界面中，可以清晰的观察到哪些地方的居民往来更加频繁，在视图上，线路越是密集的区域，居民往来就越是频繁。

而且可以通过选择量级限制，确定哪些地方的流量更加庞大。图10表示的是出行视图的一种情况。

表 7 出行视图功能选项能力图形种类选择展示的图层类型量级限制选择展示的路线图中，至少多少量级栅格大小/m 选择栅格化的大小活动时间限制表明多少分钟算为停留活动颜色选择选择视图的颜色 15

图 10 出行视图工作地视图

工作地视图使用的是工作地数据，目的是为了展示城市空间中那些区域为较为

密集的工作地。在视图功能版面中，除了提供了基本的能力选项，还提供了选择了选择工作时长的选择，可以动态指定至少平均每天工作多少小时，才能算这个人的工作地，这也是对工作地分析时，动态设置工作时间的特性的适配。

通过工作地视图，可以清除的看到那些地方是工作地密集区，对于这些地方，交通的需求优先级会更高，能耗的也会相对更高。图11表示的是图层类型为 3D 柱状图

的工作地视图情况，柱状图的高度越高，说明该栅格的工作地更密集。

表 8 工作地视图功能选项能力图形种类选择展示的图层类型量级限制选择展示的视图中，至少多少量级最小工作时长至少多少个小时，才能被认定为工作地栅格大小/m 选择栅格化的大小活动时间限制表明多少分钟算为停留活动颜色选择选择视图的颜色 16

图 11 工作地视图居住地视图居住地视图使用的是居住地数据，该视图展示了城市空间中的居住地情况。在居住地视图中，提供了白天开始时间和白天结束时间的选择功能，可以动态的指定白天时间段的定义，在本系统中，白天的开始时间可以选择 7 点、8 点或者 9 点，白天的结束时间可以选择下午 5 点一直到晚上 9 点。

在该视图中，可以清晰地观察到哪些区域是居住区密集区，对于这些地方，安全需求的优先级会一般会更高。如图12表示的是图层类型红色点状图的居住地情况。在点越密集的区域，居住地就越是集中。

表 9 居住地视图功能选项能力图形种类选择展示的图层类型量级限制选择展示的视图中，至少多少量级白天开始时间选择自定义的白天开始时间，也是夜晚结束时间白天结束时间选择自定义的白天结束时间，也是夜晚开始时间栅格大小/m 选择栅格化的大小活动时间限制表明多少分钟算为停留活动颜色选择选择视图的颜色 17

图 12 居住地视图能耗热力图

通过手机信令数据，可以计算出每个人在每个栅格的活动时间时间，并且通过 poi 数据，可以得到每个格子内部的建筑信息。根据建筑的能耗比例，可以计算出每个格子的相对能耗，因此可以比较出那些栅格的能搞更大，那些栅格的能耗更小。使用的公式为：

$$energy = times \times \sum (poiNum \times weight) \quad (1)$$

energy 表示能耗值，times 表示栅格中居民活动总时间，poiNum 表示某种建筑的数量，weight 表示该建筑的能耗比例。

如图13为经典热力图下，目前能耗情况。 18

图 13 能耗热力图对于建筑的能耗比例，因为通过查询相关的数据，并没有权威的指标，并且建筑的能耗比例因地制宜 [20]，随时时间、地区等因素变化[21]。因此前端通过动态选择的方式，让使用者可以动态的指定每种建筑的能耗比例，如图 14a。

同时，能耗热力图版面也提供了统计图表，可以分别统计数据中的总体建筑信息和剪切空间的建筑信息。并且提供饼状图，线形图和柱状图三种类型。图14b表示的是饼状图的情况。

a) 建筑能耗比例设置 b) 图表统计剪切功能在每个模块中，都有一个剪切功能，可以更具自己的需要，自定义的剪切地图空间，针对性的展示该空间对应的内容。同时在能耗热力图版面，图表绘制也会记录被 19

剪切空间的城市建筑信息。并且通过图表展示出来。如图15表示的是针对某个区域的剪切，并且使用网格热力图的图层类型展示该剪切区域的居住地情况。柱形的颜色越偏向蓝色，柱形越矮，该栅格的居住密度越低。反之柱形的颜色越红，高度越高，该栅格的居住密度就越高。

图 15 居住地视图剪切

4.4 可视化的结果分析

同时展示工作地，居住地，出行视图，如图16。1. 通过各种视图的对比可以发现，在杭州市中，人口的分布情况并不是均匀的，在部分建设较好的区域，吸纳了大部分的人口，因此政府应当重视基建薄弱地区的建设，加强杭州市的全面发展。2. 居民的活动也并不是毫无关联的，居民的工作区，居住区存在一定的联系，在工作区密集的地方，通常也是居住区密集地，并且这些人口密集点的交通流量也非常的大。在城市建设过程中，应当考虑到各种居民活动的有机结合，在合理的位置提供安全设

置。同时对于人员密集和交通密集的区域,加强交通的管理和建设。 20

图 16 视图合集

5. 结论和展望

对于数据处理:在本研究中,在手机信令数据的支持下,通过一定的算法,成功将手机信令数据转换为栅格化数据,在保护个人隐私的同时,分析出有实际意义的特征,并且通过前端可视化的方式,成功的展示出了城市的空间结构。但是目前的数据只是杭州市的数据,并且数据的潜力并没有完全挖掘,对于出行方式等特征,并没有识别出来。并且对于数据的覆盖范围、随机性等特征没有进行验证,可能存在数据本身有偏向性、偶然性等问题。

对于前端可视化:在前端可视化系统中,提供了部分参数可供选择的功能,强化了可视化系统的通用性,并且使用图表统计的方式,将部分重要数据呈现出来。可视化的方式确实极大地减少了使用者对城市空间结构的判断。但是前端方面其实也还有巨大的开发空间,对于界面的功能过于单一,和视图也不够美观等问题,都还有诸多的优化的空间。 21

参考文献

- [1] 由冰玉, 李佳钰, 谢雨梅. 我国智慧交通现状及存在问题[J]. 市政技术, 2022, 40(62-66). DOI: 10.19922/j.1009-7767.2022.05.062.
- [2] 张海耀, 张益斌. 广电网络参与智慧城市建设的现状及发展空间分析[J]. 城市建设理论研究(电子版), 2022, No. 425(164-166).
- [3] JAKLE J A, BRUNN S D, ROSEMAN C C. Human Spatial Behavior: A Social Geography[M]. Duxbury Press, 1976.
- [4] 颜亚宁. 快速城市化背景下北京郊区居民的日常生活活动空间[D]. 北京大学, 2009.
- [5] Zenk, N S, Schulz, et al. Activity space environment and dietary and physical activity behaviors: a pilot study[J]. Health & place, 2011, 17(5): 1150-1161.
- [6] Wong, WS D, Shaw, et al. Measuring segregation: an activity space approach[J]. Journal of Geographical Systems, 2011, 13(2): 127-145.
- [7] Ravensbergen, Lianne, Buliung, et al. Socioeconomic discrepancies in children's access to physical activity facilities: activity space analysis[J]. Transportation Research Record Journal of the Transportation Research Board, 2016, 2598: 11-18.
- [8] Xu, Yiping, Shaw, et al. Another Tale of Two Cities: Understanding Human Activity Space Using Actively Tracked Cellphone Location Data[J]. Annals of the American Association of Geographers, 2016, 106(2): 489-502.
- [9] 申悦, 柴彦威. 基于 GPS 数据的北京市郊区巨型社区居民日常活动空间[J]. 地理学报, 2013(04): 506-516.
- [10] 塔娜, 柴彦威, 关美宝. 北京郊区居民日常生活方式的行为测度与空间—行为互动[J]. 地理学报, 2015, 70(8): 1271-1280.
- [11] González, C M, HIDALGO, 等. Understanding individual human mobility patterns[J]. Nature, 2008, 453(7196): 779-782.
- [12] SONG C, QU Z, BLUMM N, et al. Limits of predictability in human mobility[J]. Science, 2010, 327(5968): 1018-1021.
- [13] CHEN C, BIAN L, MA J. From traces to trajectories: How well can we guess activity locations from mobile phone traces[J]. Transportation Research Part C: Emerging Technologies, 2014, 46: 326-337.
- [14] 段征宇, 赵浩然. 基于手机信令数据的居民活动空间及影响因素分析[J]. 交通与运输, 2023, 39(10-14).
- [15] 赵姗, 李任斯, 茹黄娜. 智慧城市中网络安全架构设计浅析[J]. 网络安全技术与应用, 2023, No. 269(100-101).
- [16] 冯世昌. 智慧城市中的行政处罚程序问题研究[J]. 河南科技大学学报(社会科学版), 2023, 41(73-77). DOI: 10.15926/j.cnki.hkdk.2023.03.011. 22
- [17] YU Q, YUAN J. TransBigData: A Python package for transportation spatio-temporal big data processing, analysis and visualization[J/OL]. Journal of Open Source Software, 2022, 7(71): 4021. <https://doi.org/10.21105/joss.04021>. DOI: 10.21105/joss.04021.
- [18] 郭煜东, 杨飞, 周涛, 等. 基于手机信令的城市机动化方式细分双层模型研究[J]. 交通运输系统工程与信息, [19] 余庆. 基于移动通信数据的居民活动空间检测与分析[J]., 2018.
- [20] 陈莉, 许娇. 基于 VAR 模型的建筑能耗宏观影响因素分析[J]. 徐州工程学院学报(自然科学版), 2018. DOI: 10.15873/j.cnki.jxit.000492.
- [21] 冯增喜, 杨芸芸, 赵锦彤, 等. 基于人工智能的建筑能耗预测研究综述[J]. 建筑节能(中英文), 2023, 51(22-29).

23

致谢

我在撰写本篇论文期间,得到了许多人的帮助和支持,我要在此向他们致以谢意。

首先要感谢的是指导我的导师宋轩教授和副指导老师余庆老师。他们在我选择研究课题、制定研究计划以及实践操作时都给了我很多指导。在我研究时期的每一

步,他都给予了我细致入微的指导和宝贵的建议,为我提供了持续的支持和动力。

其次,我要感谢实验室的赵奕丞师兄,他们对我这个新成员的加入十分友好,热心地帮助和指导我。在组内,他为我提供了必要的信息服务和时间安排方面的支持,为我顺利完成论文提供了先决条件。

最后我还要感谢我的同学们和朋友们。他们在我写论文的过程中提供了宝贵的建议和意见,以及实用的帮助。当我需要帮助时,他们时时刻刻都能给我支持和鼓励。 24

疑似剽窃文字表述

1. 顺利完成论文提供了先决条件。
最后我还要感谢我的同学们和朋友们。他们在我写论文的过程中提供了宝贵的建议和意见，

说明：1. 总文字复制比：被检测论文总重合字数在总字数中所占的比例

2. 去除引用文献复制比：去除系统识别为引用的文献后，计算出来的重合字数在总字数中所占的比例

3. 去除本人文献复制比：去除作者本人文献后，计算出来的重合字数在总字数中所占的比例

4. 单篇最大文字复制比：被检测文献与所有相似文献比对后，重合字数占总字数的比例最大的那一篇文献的文字复制比

5. 复制比：按照“四舍五入”规则，保留1位小数

6. 指标是由系统根据《学术论文不端行为的界定标准》自动生成的

7. 红色文字表示文字复制部分；绿色文字表示引用部分（包括系统自动识别为引用的部分）；棕灰色文字表示系统依据作者姓名识别的本人其他文献部分

8. 本报告单仅对您所选择的比对时间范围、资源范围内的检测结果负责



 amlc@cnki.net

 <https://check.cnki.net/>