山 东 科 技 大 学

本科毕业设计（论文）开题报告

题 目 基于共享单车骑行数据的上海市通勤模式分析

学院（系、部） 测绘与空间信息学院

专业班级 地理信息科学 20-2

姓 名 潘志清

学 号 202001020717

指 导 教 师 牟乃夏

填表时间： 2024 年 3 月 15 日

填表说明

1.开题报告作为毕业设计（论文）答辩委员会对学生答辩资格审查的依据材料之一。

2.此报告应在指导教师指导下，由学生在毕业设计（论文）工作前期完成，经指导教师签署意见、相关系主任审查后生效。

3.学生应按照学校统一设计的电子文档标准格式，用A4纸打印。

4.参考文献不少于10篇，其中应有适当的外文资料（一般不少于2篇）。

5.开题报告作为毕业设计（论文）资料，与毕业设计（论文）一同存档。

|  |  |
| --- | --- |
| 毕业设计（论文）题目 | 基于现代浏览器的GIS系统设计与实现 |
| 题目类型 | □毕业设计 ■毕业论文 |
| 题目来源 | □实验 □实习 □工程实践 □社会调查 ■科研选题等其他 |
| 1. 本课题的研究目的和意义 | |
| 1. 本课题的主要研究内容（提纲）   1. 数据搜集与数据预处理  1. 共享单车轨迹数据清洗：异常值剔除（范围异常、骑行时间异常、骑行速度异常）、重排序（按照骑行日期重新排序）、以天为单位切分为独立文件、某一天的独立文件再以 24 小时制排序。  2. 共享单车轨迹数据预处理：根据公式计算轨迹点 GeoHash 值作为新的属性保存在文件中、删除原先的用户及车辆ID属性以时间顺序重新编码记录索引号。  3. 乱序共享单车单车轨迹的重建与简化：根据 GeoHash 值重建骑行轨迹。  2. 基于共享单车轨迹的时空通勤模式分析  1. 统计用户通勤距离、骑行时间等参数随时间的变化规律。  2. 根据用户骑行数据绘制热力图，提取骑行热点区域，分析原因。  3. 使用 TraClus 算法提取用户骑行的公共轨迹。  4. 结合路网数据与骑行公共轨迹分析上海市道路建设对居民骑行通勤模式的影响。  3. 结论分析与制图  1. 结合上述分析步骤的结果，使用相关软件制作统计图表、热力图、频次空间分布图等图表用于辅助分析与表达。  2. 研究居民骑行通勤模式、结合上海市路网数据提出合理的道路建设优化建议。 | |
| 1. 文献综述（国内外研究情况及其发展） | |
| 1. 拟解决的关键问题   1. 骑行轨迹复原与简化：由于与服务器的通讯存在延迟等原因，原数据集记录的轨迹中存在个别乱序点，这导致整体轨迹无法直接使用。并且，匀速骑行的轨迹中含有大量值得剔除的冗余点，比如，某段沿平直道路前进的轨迹只需保存起点和终点，而中间点均可视为冗余点。  2. 基于轨迹数据的聚类分析与公共轨迹提取：一般研究会将骑行轨迹视作离散点云，再使用基于密度的聚类方法进行分析。这样做往往会忽略轨迹本身作为二维矢量数据的结构特征，并且，提取出的热点区域只是一个模糊的空间范围，对于后续空间分析的支撑作用有限。 | |
| 1. 研究思路和方法   1. 研究思路：轨迹模式的时空分析  1. 时间变化分析：数据集中记录了2016年8月及2020年8月这两个整月的上海市某品牌共享单车骑行数据，最起码可从年际、月度、逐小时这几个尺度来分析时间上的用户骑行数据变化。  2. 空间分布分析：同时对数据集使用基于密度聚类和公共轨迹提取，并结合上海市路网矢量数据进行空间分布的分析。重点考察，热点区域与公共轨迹是否重合，公共路径与哪些路段存在交点（经过了）等，为后续研究结论提供支撑。  2. 研究方法：  1. 骑行轨迹复原与简化：轨迹复原与简化可以合并到一个步骤中实现。首先，设计算法识别出位置快速变化的点，这些点一般是特征点或者是轨迹中的乱序错误点。然后，根据这些点将原来的轨迹截断，在对每一小段做简化后（例如使用道格拉斯-扑克法根据阈值简化），再根据不同小段的起止点就近连接。  2. 基于轨迹数据的聚类分析与公共轨迹提取：使用 TraClus 算法概括大量骑行轨迹的公共路径，精确寻找骑行热点路径。与传统的基于密度的空间统计方法不同，TraClus 算法是针对轨迹段进行空间聚类的算法，改算法首先将轨迹根据特征点切分为众多小段，然后再对这些小段进行相似度度量，最终归纳出众多轨迹的公共路径。 | |
| 1. 本课题的进度安排   第一阶段（2024.1.12 - 2024.3.15）: 确定研究方向，阅读国内外相关文献，总结论文思路、技术要点及待解决的实际问题。获取相关数据，研所可能会用到的技术框架，熟悉数据处理流程。  第二阶段（2024.3.16 - 2024.4.20）: 筛选、清洗、整合数据，数据处理算法实现与验证。  第三阶段（2024.4.21 - 2024.5.19）: 撰写论文初稿，提交初审，根据建议修改完善论文。  第四阶段（2024.5.20 - 2024.6.19）: 论文答辩。 | |
| 1. 参考文献   [1]李文翔,唐桂孔,刘博,等.基于摩拜骑行数据的上海市共享单车减排效益时空分析[J].环境科学学报,2021,41(11):4752-4759.DOI:10.13671/j.hjkxxb.2021.0213.  [2]王若萱,吴建平,奇格奇. 基于上海市数据的共享单车用户通勤模式研究（英文）[C]中国仿真学会.第三十三届中国仿真大会论文集.2021:16.DOI:10.26914/c.cnkihy.2021.025005.  [3]全雨霏. 南京市共享单车使用的时空特征及其骑行环境评估[D].东南大学,2024.DOI:10.27014/d.cnki.gdnau.2022.001097.  [4]常新. 基于共享单车轨迹数据的城市街道可骑行性研究[D].哈尔滨工业大学,2021.DOI:10.27061/d.cnki.ghgdu.2020.002483.  [5]Jie Bao, Tianfu He, Sijie Ruan , Yanhua Li, and Yu Zheng. 2017. Planning Bike Lanes based on Sharing-Bikes’ Trajectories[c]. In Proceedings of KDD’17, August 13–17, 2017, Halifax, NS, Canada., , 11 pages.  [6]刘泉宏,唐福星.基于K-means聚类算法与重心法的故障共享单车回收中心选址优化[J].运筹与管理,2023,32(07):85-91.  [7]谢光明. 基于改进时空图神经网络的共享单车流量预测[D].华东师范大学,2023.DOI:10.27149/d.cnki.ghdsu.2023.004430.  [8]刘冰,王舸洋,朱俊宇,等.基于共享单车大数据的骑行生活圈识别及其活动网络模式分析[J].城市规划学刊,2023(04):32-40.DOI:10.16361/j.upf.202304005.  [9]王俊,于爱荣.基于ConvLSTM的南京地区共享单车需求预测研究[J].软件工程,2024,27(02):55-59.DOI:10.19644/j.cnki.issn2096-1472.2024.002.011.  [10]谢国微. 天气及建成环境对共享单车出行需求的影响研究[D].南京林业大学,2024.DOI:10.27242/d.cnki.gnjlu.2022.000493.  [11]任丹. 基于TRACLUS算法的船舶轨迹分析系统的设计与实现[D].辽宁师范大学,2021.DOI:10.27212/d.cnki.glnsu.2020.001211. | |
| 指导教师意见  指导教师（签名）：  年 月 日 | |
| 所在系（教研室）意见  负责人（签章）：  年 月 日 | |