

**本科毕业设计（论文）**

题目：基于共享单车骑行数据的上海市通勤模式分析

学 院：测绘与空间信息学院

专业班级：地理信息科学20级2班

姓 名：潘志清

学 号：202001020717

指导教师：牟乃夏

完成日期：（宋体小3）

教务处制

**BACHELOR'S DEGREE THESIS OF SHANDONG UNIVERSITY**

**OF SCIENCE AND TECHNOLOGY**

**Shanghai Commuting Patterns Unveiled through Shared Bicycle Data**

College：College of Geodesy and Geomatics

Subject：Geographical Information Science

Name：PanZhiqing

Directed by：MouNaixia Professor

**QINGDAO CHINA**

**郑 重 声 明**

本人呈交的毕业论文，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、图片资料真实可靠。尽我所知，除文中已经注明引用的内容外，本设计（论文）的研究成果不包含他人享有著作权的内容。对本设计（论文）所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确的方式标明。本设计（论文）的知识产权归属于培养单位。

本人签名： 签名内容 日期： \*\*\*\*\*\*

摘 要

由于

**关键词：**关键词1；关键词2；关键词3

**ABSTRACT**

Fluvial

**KEY WORDS:** sediment; rigid vegetation; settling velocity; turbulence characterize

**目 录**

**第1章 绪论 1**

1.1 问题的提出及研究意义 1

1.1.1 问题提出 1

1.1.2 研究意义 2

1.2 泥沙沉速的研究概述 2

1.2.1 泥沙沉速的影响因素 3

1.2.2 泥沙沉降阻力系数 3

1.2.3 泥沙沉速公式 4

1.2.4 动水中泥沙沉降的计算方法 10

1.3 植物对泥沙沉降的影响概述 11

1.3.1 植物对静水中泥沙沉速的影响 11

1.3.2 植物对明渠水流中泥沙沉降的影响 11

**第1章 绪论**

**1.1** **问题的提出及研究意义**

如广泛分布在黄河流域一带的黄土地质均匀，其粉砂含量占60%~70%，缺乏团粒结构，粒间的固结主要依靠硫酸钙质，这种硫酸钙质遇水极易溶解流失，加上黄土孔隙率极高，抗蚀能力很差[1]。

.......

.......

.......

**1.1.1** **问题提出**

近年来，随着环境的日益恶化，人们对生态日益重视，含有植物的水流问题也已经成为河流动力学研究中的热点之一[2]。直观的了解，河渠水流中的植物不仅减少了过水面积，加大了河渠地面的粗糙程度，降低了河渠的行洪能力，加大了两岸的洪灾威胁。

公式、图文示例：

（1）公式示例：

单颗粒球体在无限水体中等速下沉时，其沉速机理可看作对称绕流阻力与颗粒有效重力相平衡[13]，即

*（γs-γ）\*π = CD πγ*  （1.1）

（建议公式用微软office的公式编辑器输入）

Stokes曾以粘滞性流体的一般性的运动方程式作基础，忽略惯性项的条件下推导出滞留区的阻力系数为

CD=24/Red （1.2）

(公式按章编号，例如 第三章中的第二个公式 3.2)

（2）表示例：

表4.1 光滑明渠水流实验水力条件

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **实验**  **编码** | **H** | **Q** | **J** | **B** |  |
| cm | L/s | ‰ | cm | cm/s |
| w1 | 18 | 7.56 | 0.02 | 42 | 0.19 |
| w2 | 18 | 11.34 | 0.07 | 42 | 0.68 |
| w3 | 18 | 15.12 | 0.13 | 42 | 1.27 |
| w4 | 18 | 18.9 | 0.21 | 42 | 2.05 |
| w5 | 18 | 22.68 | 0.28 | 42 | 2.73 |

其中： 为摩阻流速， ，(其中R为水力半径)；J为水力坡降，B为水槽宽度，H为水深。

（3）图示例：



图4.1 清水明渠水流下w2工况下的三维时均流速图

**第2章 理论基础与相关技术**

**2.1** **相关理论基础**

**2.1.1** **行为地理学**

行为地理学是人文地理学的一个重要分支[12]，该学科在 20 世纪初期诞生，是地理学、心理学、行为科学的交叉学科，在后来的发展中逐步融合行为主义地理学与时间地理学两大理论基础，形成了一套相对完善的研究范式与理论体系。该学科主要探讨人与地理环境之间的互动，强调“以人为本”，从微观人地关系的研究出发探索宏观时空规律。受到行为经济学中将非理性的经济主体作为微观研究的对象启发，行为地理学所研究的“人”是一种有限理性的个体，这些个体的行为虽然受到地理环境等的客观制约而呈现出一定规律性，但也会出于主观做出一些随机行为。个体可以依据属性汇总为群体、可以依据社交网络与空间活动范围汇总为社群，这些不同类别不同层次的群体同样被归为“人”这一研究对象。行为地理学所研究的“地”指的是一种抽象的、主客观相结合的行为空间，比如城市意向、认知空间等。行为地理学强调人地关系的统一性，人与地相互依赖、相互影响、相互塑造，这二者是不可分割的。

与传统的将研究区域视作静态的“块”不同，行为地理学的研究强调动态性，这种动态性体现在：1）其研究图景是一个具有复杂时空流动模式的网络：个体与个体之间、个体与环境之间均存在复杂的互动关系，甚至这些关系本身也相互制约相互影响，这种有机的网络模型所具有的动态性是传统的静态区域块模型所无法比拟的；2）研究视野综合了宏观与微观并可动态调整：通过汇总个体行为，总结区域规律来不断扩大研究视野。通过不断缩小研究视野，深挖“人”这一复杂个体全生命周期内在认知的流变规律，更好地关怀个体生命质量；3）数据及对应处理方法层面的动态性：移动互联网、GPS 和移动位置服务(location-based services，LBS)等技术提供实动态数据支撑，人工智能、大数据技术等先进的数据处理手段提供快速动态的分析能力。

随着时代的发展，“人本主义”日益得到社会各界的认同，着重研究人的日常行为及其与环境的交互关系的行为地理学也越发受到重视。行为地理学以其独特的研究视角，在可持续发展、社会公平、国土空间规划与社会治理等方面发挥出越来越大的作用。

**2.1.2** **城市职住平衡与过剩通勤理论**

随着中国城镇化水平不断提高，城市交通拥堵问题日益凸显，其背后城市交通巨大供需失衡问题也逐渐被社会各界所认识。对于城市居民而言，住所与工作地点相距甚远往往会带来巨大的通勤压力，这也是城市交通拥堵的一个重要原因。研究并达成职住平衡可以有效缓解城市交通拥堵，职住平衡理论也可以为城市规划提供有用的指导。职住平衡理论（Jobs-Housing Balance）[13]最早由 Howard 在1902年出版的书籍《明日田园城市》中提出，他认为应当建设一种具有“自力性”（Self-contained）的社区，这种社区能够确保居民的就业需求与住房需求一并得到满足，最终由众多具有“自力性”的社区组成“田园城市”。这样的“田园城市”可以有效缓解交通拥堵及环境污染，减轻居民通勤压力，缩减中低收入家庭开支提升幸福感。

将多种城市功能理想化地集成在许多相互离散的小块土地上，这种理想化的城市模式在现实中往往难以实现。“职住分离”的现状是市场长期自主选择的结果，一方面企业追求产业聚集带来的规模效益，企业以零散的方式分布于小片地块会降低生产效率提高管理成本，另一方面，求职者更愿意为了高薪前往较远的工作地就职，随着城市公共交通系统的发展，通勤成本不断下降，这种职住距离还会被进一步拉远。另外，对于一个具体的家庭而言，求职并不是选择居住地唯一需求，医疗保健、子女求学及休闲放松等综合需求同样重要，尤其是优质的公共服务资源（重点学校、三甲医院等）往往成为一个家庭选择居住地的主要因素，单纯强调职住一体并不能很好地满足家庭的综合性需求。因此，需要客观看待城市职住分离的现状，着重研究城市通勤现状，过剩通勤可以作为衡量一个城市职住平衡的量化指标，较为客观地反映城市通勤现状。

过剩通勤（excess commuting）指的是实际通勤距离与理论通勤距离之间的差值，该理论最早在 1982 年由 Hamilton 首先提出，经过三十多年的发展逐步形成了一套较为完善的理论体系。在城市建模方面，该领域最早将城市建模为单中心模型，认为城市中心提供了所有的就业机会。后来随着计算机技术及地理信息系统的发展，有学者使用泰森多边形将城市划分为多个职住格网，更精细地建模城市职住空间分布情况。在理论通勤距离计算方面，主要计算方法包括：1）交通问题线性规划最优化函数（Transportation Problem in Linear Programming, TPLP）：采用线性规划的方法，以交通成本最小为优化目标，职住人数分布守恒为约束条件求解理论通勤距离；2）最大熵值法与蒙特卡罗仿真模型：最大熵值法是一种基于信息论的随机模拟方法，计算简单，不需要大量数据。蒙特卡罗模拟是一种基于随机抽样的统计方法，灵活性高，计算量大，结果受随机抽样影响；3）等比例匹配通勤：认为就业岗位分配与企业所占区域在区域中的比例相同。其中，TPLP 的其基本原理是将城市中的就业岗位和居住地视为节点，将通勤视为弧线，并以交通成本最小化为目标，求解所有居民的通勤路径和通勤距离。

其中， 为最小平均通勤距离， 为城市通勤总人口， 为城市居民总数， 为城市就业岗位总数， 为居民 与就业岗位 之间的通勤成本（一般为距离）， 为居民 与就业岗位 之间的通勤量（一般为人数）。

上述公式需要满足如下约束条件：

其中， 为居民 的出发地， 为就业岗位 的目的地。

居民实际通勤情况可以使用问卷调查、GPS 轨迹数据、公交刷卡数据等方式获取，通过计算实际通勤距离 结合理论通勤距离 可以计算出过剩通勤率 ：

**2.2** **相关研究技术**

**2.2.1** **聚类分析**

聚类分析指的是按照一定的分类规则（距离、相似性等）将一组数据划分为多个类簇的过程，是一种无监督的探索性数据分析（数据挖掘）方法，在大数据分析、计算机图形学、模式识别等专业领域应用广泛。聚类分析过程往往涉及许多算法与子任务迭代处理，且并非完全自动化，一般需要根据结果手动调整参数优化。

聚类分析的核心问题是聚类模型，典型的聚类模型可以分为以下几类：1）连接性模型：层次聚类是连接性模型的典型代表。它通过不断合并距离相近的数据对象来构建层次化的簇结构。该方法的优点是易于理解和实现，但对初始度量方法的选择比较敏感。2）质心模型：k均值算法是质心模型的代表，它通过迭代地移动簇的质心来使得簇内数据的距离平方和最小化，最终将数据划分到k个簇中。3）分布模型：使用统计分布对簇进行建模，例如期望最大化算法使用的多元正态分布。4）密度模型：DBSCAN和OPTICS算法是密度模型的代表，它们将数据空间中密度较高的区域定义为簇，并通过密度阈值或可达距离来识别簇中的数据点。5）基于图的模型：HCS聚类算法是基于图的模型的代表，它将数据点之间的相似性关系转换为图结构，并通过图的社区发现算法来识别簇中的数据点。6）神经模型：自组织映射（SOM）是神经模型的代表，它是一种无监督神经网络，将数据点映射到一个二维网格上，并通过相邻神经元的激活模式来识别簇中的数据点。

在上述聚类模型的基础上，可以进一步设计用于区分不同类别的规则，这些规则一般依赖一种相似性度量来划分，例如欧几里得距离、曼哈顿距离、闵可夫斯基距离、皮尔逊相似度、豪斯多夫距离等。根据分类规则对于分类算法的约束程度（是否允许重复分类、是否允许存在未分类目标等），又可以分为硬性聚类与模糊聚类（软聚类）。

**2.2.2** **时空轨迹聚类分析**

时空轨迹建模为某一对象按时间顺序记录的空间位置序列，对其的聚类分析就是根据轨迹间的时空相似性度量划分为不同类簇的过程。时空轨迹数据具有一定的特殊性，首先，它是以时间顺序记录的移动对象位置序列，内含时间因素代表其具有动态性，同一个对象在不同时刻具有不同位置属性，这与一般聚类分析的静态数据（如带有诸多静态属性数据的人或物）具有较大差别。其次，它有明确的线状空间结构，相较于一般的（多维空间中）点状数据模型，轨迹数据本质上为连续的同质顺序点列，轨迹间的相似性度量无法直接套用一般聚类的概念为团块的聚类模型已有的度量方法，而是需设计与之适应的特殊算法。

考虑到时空轨迹的特殊性，无法直接套用一般的聚类方法，而是需要根据轨迹的时空间结构设计特定的相似性度量算法，部分典型的时空轨迹相似性度量算法：1）轨迹间欧氏距离：将待分类轨迹视作随时间变动的点集，在每一个时间断面计算轨迹点间的欧氏距离并在全局时间轴方向上累积。对噪声敏感，需要采样频率一至，否则会有较大误差。2）最小外包络距离：首先使用轨迹特征点（方向快速变化的点）切分轨迹，然后提取这些子段的外包络矩形并组成顺序序列作为轨迹的概化表征，最后累计矩形间的距离以度量轨迹间的相似性。该方法对原始轨迹平滑去噪，但也会丢失一部分信息。3）子轨迹相似性度量[11]（TRACLUS）：该算法首先根据最小描述长度原则将轨迹划分为多个子轨迹，然后计算子轨迹间的三种距离（垂直距离、平行距离、角度距离）并加权计算的综合相似性量值。4）历史最近距离：在一段给定的时间范围内，检查各个时刻两个轨迹之间的距离（欧氏距离或其他），选取最小距离作为历史最近距离。

**2.2.3** **用户行为分析**

用户行为分析（User behavior analytics）是一种通过分析用户使用数据来了解用户行为模式的方法。企业可以收集用户点击次数、登录时间、浏览时长等数据，并通过数据挖掘、模式分析等手段进行分析。用户行为分析可以帮助企业在不打扰用户的情况下（相较于发放用户问卷）更好地了解用户群体，从而提升服务质量。此外，高效及时的用户行为分析还可以防范恶意攻击（例如DDoS攻击），过滤异常账号（刷单诈骗）。一些常用的用户数据挖掘方法包括统计用户使用频率并划分用户等级、计算并绘制用户贡献累计分布曲线（CDF曲线）、计算用户留存率等。其中，用户留存率（User Retention Rate）是一个用于衡量产品吸引力及用户粘性的重要指标，它指的是在某一时间段内开始使用产品或服务的用户，经过一段时间后仍然继续使用该产品或服务的用户的比例。

夏琼燕等人[14]下载并研究了2005年到2007年 Open Street Map 开源地理信息网站中志愿者用户数据（记录了用户的贡献行为），通过计算用户贡献累计分布曲线，他们发现该网站的贡献者符合“二八定律”，即该网站两成的贡献者贡献了约80%的数据，并根据贡献率将用户分为多个活跃等级。通过合理选取时间窗口，计算用户留存率，结合用户活跃度分类，他们发现OSM社区中低活跃度的用户留存率偏低，而高活跃度用户总体保持稳定，社区稳定存续主要依靠中高活跃度用户。

**2.3** **本章小结**

本章以行为地理学、城市职住平衡理论和过剩通勤理论为理论基础，并探讨聚类分析、用户行为分析和时空轨迹聚类分析等相关研究技术。行为地理学强调以人为本，探讨人与地理环境的互动关系，并在可持续发展和社会公平等领域发挥重要作用。城市职住平衡理论指出有效缓解城市交通拥堵的关键是合理规划就业与住房地点，而过剩通勤理论提供了衡量城市通勤状况的量化指标。在相关研究技术方面，聚类分析和用户行为分析为数据挖掘和模式分析提供了重要手段，特别是时空轨迹聚类分析针对轨迹数据的特殊性提出了一系列相似性度量算法。这些理论和技术的研究与应用有助于更好地理解和解决城市发展中的诸多问题，促进城市可持续发展和社会治理的进步。

**第3章 研究数据与方法**

**3.1** **研究范围**

**3.2** **数据来源与预处理**

**3.2.1** **数据来源与数据清洗**

**3.2.2** **轨迹重排序及长度量测**

**3.3** **研究方法**

**3.3.1** **轨迹终点最近邻匹配算法**

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.

[12] 塔娜,柴彦威.行为地理学的学科定位与前沿方向[J].地理科学进展,2022,41(01):1-15.

[13] 杨超,汪超.城市过剩通勤与职住平衡模型[J].同济大学学报(自然科学版),2013,41(11):1712-1716.

[14]夏琼燕,罗冠,张翔,等.OpenStreetMap志愿者贡献与留存分析[J].测绘与空间地理信息,2021,44(02):90-93+97.

致谢

本论文是在\*\*\*老师的悉心指导下完成的，感谢\*\*\*老师在论文的撰写中给予的细心的指导与帮助。老师那严谨的治学态度、精益求精的科研作风给我们树立了很好的榜样，这将时刻激励我在工作与学习中继续进步。

感谢朝夕相处的同学，、、、

感谢、、、、

感谢同小组的成员，感谢他们在毕业设计过程中给予的鼓励和帮助，他们对于学术上严谨的态度和不懈的追求令我难以忘怀，在此向他们表示衷心的感谢！