## 第七届全国计算机仿真大奖赛

# 文字解答

题目名称	共享单车一日流动停放仿真设计与分析
参赛人员.	董健
	张芳
_	
	袁和
辅导教师	陈彬
参赛单位.	国防科技大学
收题日期	2018-5-26
1/4 / <del>(-)</del>    //1	

二〇一八年五月二十六日

## 目录

摘要		5
第一章	引 言	6
1.1	题目重述	6
1.2	问题背景	6
	1.2.1 共享单车的发展与现状	6
	1.2.2 共享单车存在问题	7
1.3	题目分析	9
	1.3.1 乘客出行分析	9
	1.3.2 租借与停放特点分析	10
	1.3.3 单车引发问题分析	10
	1.3.4 城市选择	11
第二章	共享单车流动停放模型	13
2.1	模型整体架构	13
2.2	城市单车区域规划	14
2.3	共享单车租借模型	15
2.4	共享单车流动模型构建	16
	2.4.1 OD 矩阵	16
	2.4.2 共享单车行驶时间模型	17
2.5	共享单车停放模型构建	19
2.6	模型假设	19
第三章	数据获取及数据分析	21
3.1	数据获取及处理	21
	3.1.1 爬虫机制	21
	3.1.2 爬虫策略	22
	3.1.3 数据处理	24
3.2	数据驱动方法	25
3.3	数据分析及模型合理性验证	26
	3.3.1 聚类分析	26

	3.3.2 共享单车租借分析	. 26
	3.3.3 共享单车流动分析	. 28
	3.3.4 共享单车停放分析	. 29
第四章	仿真系统设计	.31
4.1	仿真系统设计	.31
4.2	仿真系统初始界面	.32
4.3	仿真实验界面	.32
4.4	软件使用说明	. 35
第五章	实验及分析	.36
5.1	实验设计	.36
5.2	实验结果分析	.36
第六章	结论	.41
参考文章	献	.42
附录:		.44

图	1 模型框架	14
图	2 网络到 OD 矩阵的转化示意图	17
图	3 百度地图骑车路径示意图	18
图	4 20 对起终点的实际距离、模型距离及其差值	19
图	5 爬虫机制	22
图	6 OFO 界面	23
图	7 爬虫程序流程	24
图	8 北京五环区域规划	26
图	9 北京五环内各个区域共享单车被租借骑出的数量	27
图	10 一天内各个时间段内共享单车被租借的数量	28
图	11 乘客出发频次最高区域(红星)以及乘客更倾向到达的区	域
	(蓝星)	29
图	12 区域内单车离散分布图	30
图	13 系统用例图	31
图	14 仿真系统初始界面	32
图	15 仿真界面总览图	33
图	16 仿真控制模块图	33
图	17 仿真参数显示模块图	34
图	18 仿真过程可视化模块图	34
图	19 软件使用说明流程图	35
图	20 累计出行次数分布图	37
图	21 骑行时间概率分布图	37
图	22 区域单车骑入-骑出对比图	38
图	23 仿真数据和真实数据骑入量与骑出量差值的空间分布热力	冬
		39
图	24 区域单车停放数量随时间变换热力图,从左至右时间依次	为
	2时、5时、8时、11时、14时、17时、20时、23时	40
图	25 共享单车损坏率对单车活跃量的影响	40

#### 摘要

共享单车在我国迅猛发展,其方便快捷、经济高效的特点受到人们推崇,但同时也带来了公共管理、资源浪费、分布不均等问题。针对此类问题,我们以北京市五环以内为研究区域,根据对共享单车出行特点进行分析,利用数据驱动的方法构建了共享单车一日内的停放流动模型。通过实证数据的研究,我们验证了模型构建的合理性。基于模型的仿真结果表明,共享单车的确存在区域分布不均,资源浪费问题,但也发现在一天内不会产生分布不均的问题,而是在长时间的积累下形成的。另外,通过对单车的损坏率对居民骑行出行的影响,我们发现当乘客出行意愿容易受到单车损坏度的影响而选择其它交通工具时时,单车的使用率会严重下降。

在共享单车一日流动停放仿真模型设计与分析过程中,主要解决了以下问题:

- (1)分析了共享单车租借流动特点,并利用实证数据进行了论证, 为模型的提出打下了基础:
- (2)提出了共享单车一日内的流动停放模型,租借一流动一停放模型。全面地分析了共享单车流动的全过程;
  - (3) 利用获取到的真实数据,驱动所构建的模型;
- (4)通过仿真实验分析了共享单车流动停放特点以及其所暴露出的问题;
  - (5) 通过真实数据与仿真数据的对比,验证了仿真的可靠性。

#### 第一章 引言

#### 1.1 题目重述

题目 C: 共享单车近年来在我国迅猛发展,由于其方便快捷、经济高效的特点受到了大众一致推崇,但由于管理上的滞后,其停放混乱、破坏严重、分布无序的问题也非常突出。请以自己所在城市为例,建立共享单车一日内流动停放的动态分布模型,并以一种共享单车(摩拜或OFO)为例编程,动态显示其 24 小时的流通情况。

#### 1.2 问题背景

共享单车是共享经济和物联网发展下的产物,它指企业在校园、地铁站点、公交站点、居民区、商业区、公共服务区等提供自行车单车共享服务,是一种分时租赁模式。而目前共享单车多指通过APP寻找车辆,利用扫码等智能方式解锁自行车,通过后台远程实时监控车辆健康和运营状态的单车智能出行新形式。

#### 1.2.1 共享单车的发展与现状

中国共享单车最初由国外兴起的公共单车模式引进,由政府主导分城市管理,多为有桩单车。从 2010 年开始,专门经营单车市场的企业开始出现,但公共单车仍以有桩单车为主。随着移动互联网的快速发展,以摩拜为首的互联网共享单车应运而生,更加便捷的无桩单车开始取代有桩单车。这种无桩式的设计,以及互联网和智能移动设备的运用,解决了以政府为主导的公共自行车的不便之处,使得自行车的寻找、使用、停放实现了随用随取、随停随放,真正体现了自行车的便利性。目前我们所说的共享单车常指的是这种无桩智能共享单车。

共享单车的发展经历了三大阶段。第一阶段是共享单车的萌芽期,这阶段共享单车行业中只有 ofo 和摩拜。在这一个阶段,ofo 完成了 A 轮融资,从破产的边缘走上了快速发展的道路,摩拜则完成了自己首款单车的研发工作,开始在上海投入运营;第二阶段是 ofo 和摩拜两家行业抢占用户市场,在这一个阶段也逐渐涌现了其他的共享单车品牌;第三阶段是共享单车的爆发期,市场上出现了大量的不同品牌的共享单车,一线城市已经是摩拜、ofo 的天下,但这并不代表两家公司完胜,广大二三线城市仍然有一定的创业机遇,二三线城市公共交通相对没有那么发达,汽车的普及同样造成了交通拥堵,人们平均出行距离并不一定比一线城市更短,共享单车的发展机会仍然广阔。众多的共享单车品牌表现出了一种"百花齐放、百家争鸣"的态势。

共享单车的出现,解决了政府交通拥堵和环境污染问题、市民出行 无序需求的问题以及为其运营商创造了商业价值。在环境方面,共享单 车的出行模式有利于节约资源、保护环境,顺应了国家大力提倡的低碳 环保出行方式,从而有效地改善了环境问题。在交通方面,共享单车作 为日趋完善的城市交通网络有效的补充,其自由灵活的租赁模式,能够 为 1~3 公里的短途出行提供服务,配合现存的公共交通系统,形成了 "点对点"的运输服务,提高现有公共交通系统的可达性,破解市民出 行"最后一公里"的难题,并且起到了缓解城市交通捅堵的作用。在产 业链方面,共享单车企业成功的连接了上游和下游的产业链。共享单车 真正实现了多角色、多行业互利互助,共同发展。

同时我们也应该看到共享单车在快速发展背后所暴露出来的一些问题,也只有通过不断地完善,才能够更好的发挥共享单车的优势作用,助力其实现更好、更快地发展。

#### 1.2.2 共享单车存在问题

#### (一)投放过量问题

2016年底以来,国内共享单车突然就火爆了起来,仿佛一夜之间,

共享单车已经到了"泛滥"的地步,各大城市路边排满各种颜色的共享单车,除了较早入局的摩拜单车、ofo 外,目前为止至少有 25 个新品牌汹涌入局,其中甚至还包括电动自行车共享品牌。各种品牌为了抢占市场,不断地向全国各地投放车辆,全国共享单车累计投放量超过 1600 万辆,北上广深就占了近 500 万辆。由于共享自行车过度投放、乱停乱放现象严重和超出城市非机动车可停放区域承载能力,多地开始了暂停共享单车的投放。共享单车暂停投放的背后暴露了其诸多问题。

#### (二)管理混乱问题

首先,其无桩化的设计,满足了随用随停的要求。但是在共享单车使用之后,一些人为了最大限度的方便自己,对城市管理规定视若无睹,把共享单车随意的违规停放,结果是方便了自己,给别人以及整个社会添了堵。其次,还存在有个别人将共享单车停放于自己的办公地方以及住宅小区内,试图将共享单车成为个人的出行工具。另外,由于共享单车数量众多,成为了恶意毁坏的对象,例如有些共享单车的二维码被涂画、刮花,导致不能使用,有些单车的轮胎被一些人扎破等,导致了共享单车不能够再次被别人使用。还有,对于拥有共享单车所有权的公司来说,单车的数量过多,分布又相对过于分散,可谓是遍布城市的各个角落,一旦单车出现损坏状况,需要维修师傅去维修,则及时、快速、准确的寻找到相关车辆是一个大的难题,这对单车公司来说成本是非常高的。

#### (三) 基础设施资源不完善

一方面,在我国众多城市的公共出行道路中,自行车道大多都要和机动车或人行道共享,随着共享单车数量进一步的提升,这将存在着极大的隐患。另一方面,共享单车的大量出现,将会占用大量的城市空间,而城市中专门的自行车停放点很少,对于其他出行工具的停放具有很大的挤出效应,占用大量的公共空间。同时,由于共享单车的所有权归属于单车公司,其属于商业性的运营模式。这两者之间的进一步融合还有待规则条例乃至法律的逐步完善。

#### 1.3 题目分析

根据题目所提问题,其目的旨在通过构建共享单车一日内流动停放的动态分布模型,分析共享单车一日内的流通情况以给出解决共享单车停放混乱、破坏严重、分布无序问题等的合理建议。针对此问题,下面将用户对用户出行特征、用户租借与停放特点进行分析。

#### 1.3.1 乘客出行分析

共享单车从 2014 年开始起步到如今,从校园走进城市、从一线城市到二三线城市,已经融入到了人们的日常生活中。短短几年时间共享单车能够如此快速发展,主要得益于它解决了人类出行最后一公里的问题。现在在现代城市出行活动中,人们往往需要换乘一种或多种交通工具才能到达目的地。这种组合的出行方式已经构成了现代日常生活中的一个重要组成部分,在一定程度上可以提高出行效率,因此越来越受到人们的青睐。

艾瑞咨询《2017年中国共享单车行业研究报告》显示,所在城市道路拥挤状况和空气质量是影响人们骑行意愿的主要因素,其中,路况拥挤程度对用户的骑行意愿影响最大,路况不佳时,用户更愿意使用共享单车以躲避交通拥堵,骑行意愿提升 10.1%,而当道路通畅时,用户的骑行意愿则会降低 11.4%;约七成的单车用户会将共享单车作为出行中的一环,其余百分之三十的用户倾向于不与公共交通结合,仅使用共享单车前往目的地,其中运动爱好者占多数,比例达到 37.4%,90%的用户在目的地与地铁站/公交站之间相距 3000 米内时选择使用共享单车,共享单车已经成为最有效的"最后一公里"解决方案;低频用户的使用场景较为单一,61.7%的用户出行场景集中在家/公司/商圈到地铁站/公交站间的往返,单次使用时间集中在 30 分钟以内,占比达到 84.4%;高频用户使用场景较为丰富,最主要的场景仍是家/公司/商圈到地铁站/公交站的往返,占比均在 70%左右,其次为商业圈内的代步或上下班通勤,占

比与主要场景差距不大; 骑行时长主要集中在 30 分钟内, 但 10%的高频用户骑行时长达到 1 小时以上, 是低频用户的 2.5 倍。

#### 1.3.2 租借与停放特点分析

当用户有用车需求的时候,刚好身边有车可用或距离用户不远处刚好有车,则用户会进行租借行为。如果说当用户有用车需求的时候,车辆距离用户太远找车需要花费太长时间,用户就更倾向于其他交通方式。比如找车需要十分钟骑车五分钟,即使走路需要二十几分钟,人们还是更愿意选择步行。

根据单车可以随时随地停放的特点,当用户完成骑车需求后就会就地停放。根据用户出行规律,单车一般的停放地为公共交通站点、商业街、公园和居民区等。有时候这些停放地之间是双向流动的,比如在用户上下班的过程中,单车在交通站点和居民区之间双向流动;但有时也存在单向流动,比如用户骑车去商业街、公园等地方,可能由于太累、太晚、天气原因不方便骑车,就选择打车回家。

## 1.3.3 单车引发问题分析

针对上述用户对共享单车的需求、租借和停放特点和出行特点分析,总结一下几个共享单车存在的问题:

- (1)单车分配不合理。考虑用户一日出行需求,可以了解到在上午上班期间,用户的租借单车的需要往往发生在居民区附近。但是由于单车资源分配不合理,导致在需求密集的时间段和地点的单车不能够满足用户需求。会产生早高峰期间许多用户无车可用的情况。
- (2)资源浪费。有的单车放置地点不合理,利用率较低。比如在一些交通线路的中央位置和以下人流量小的位置,放置的单车大部分都处于闲置状态,只有少部分车被利用。导致部分单车资源被浪费。
  - (3)人工调度不及时。由于在用户的使用下,单车存在单向流动,

有的地方的单车只进不出、有的地方的单车只出不进。这就需要投入人力资源进行调度。但是目前人工调度不够及时,往往会导致部分地点的车辆堆积。

(4) 损坏单车维修不及时。因为不是自己的私有财产,所以在使用的过程中更加容易损坏,同时还由于天气原因会对单车零部件进行损坏。当单车损坏之后往往公司的维修不及时,导致大量的坏车存在于街道上。用户在寻找单车时往往会出现,找了一辆是坏车再找一辆还是坏车的情况,严重影响用户体验。

#### 1.3.4 城市选择

为研究共享单车流动停放特点,分析其在监管上存在的问题,选择一个具有代表性的城市是有必要的(虽然题目做出了城市选择要求,之后会给出我们没有选择所在城市的原因)。

我们认为具有代表性的城市要具备以下特点

- 1)城市内有大量单车投放:有大量投放的共享单车,而且单车种类 多样是研究的基础,只有这样才能保证居民的选择性不受单车数量少、 种类单一的影响。
- 2) 骑行受城市地形影响较小: 共享单车很容易受到地形的影响,山地、丘陵地区都会使居民对单车的骑行意愿大大降低。
- 3)有完善的交通设施:共享单车主要是作为人类出行的补充性交通工具,旨在解决人类出行"最后一公里"的问题。地铁站、公交站、火车站、飞机场等区域会是居民出行最活跃位置。
- 4)城市经济发展突出:发达的城市,除了具有完善的交通设施外,还会有密集商业区、旅游地、公司、科技园等居民的日常活跃区域。
- 5)共享单车的监管存在问题:要分析共享单车存在的问题,选择问题突出城市会更容易发现问题所在,有助于更好地提供解决问题思路。

考虑以上几个方面,我们选择北京五环周边及以内(北京其它区域 共享单车数量少)作为研究对象。 一直处于经济发展前列的北京城区,在交通设施上已经相当完善,目前拥有 1200 条公交线、17 条地铁线以及 5 大铁路运输枢纽;繁华的商业区、大量旅游区,吸引了很多国内外游客;作为中国北方城市,地形平坦,便于骑行,而且气候较好,尤其是在四五月份更加适合骑车出行;共享单车遍布北京城,目前投入量已经达到 235 万辆,过度的投放、乱停乱放已经严重超出城市非机动车可停放区域承载能力,目前已经暂停投放。

没有选择长沙作为研究对象原因:

虽然属于国内一线城市,有着大量的共享单车投放,但是长沙作为一南方城市多雨季,而且目前天气炎热,影响了居民对共享单车的需求。另外一个更为重要的原因,相比北京,长沙交通设施较为落后,道路规划不合理,地铁线路少。目前长沙仅有两条通行地铁线,多条在建地铁线。正是由于市内地铁线在建原因,多条道路难以通行,导致了骑车意向的改变。这种外在的不确定因素给建模带来了很大困难。

#### 第二章 共享单车流动停放模型

#### 2.1 模型整体架构

根据对问题的整体分析,首先我们给出了模型的整体架构。由于共享单车都是按照区域进行投放,其空间分布往往会呈现出很多密集分布的区域。我们以城市内共享单车停放空间区域为基础,通过对区域内的租借、停放特点以及区域间的单车出行习惯建立模型,研究单车的流通情况。

通过对共享单车的流动分析,每一个单车正常完成一次流动可以分为三个阶段:租借一流动一停放。

租借: 当乘客有骑车需求时, 他会通过找车, 扫码, 解锁一系列操作完成对共享单车的租借;

流动:租借之后,乘客根据自己的骑行目的,到达目的地的过程;

停放:到达目的地后,乘客选择停放点,锁车,支付后即完成了单车的停放。

模型的整体架构如图 1, 共分为六部分, 城市单车区域规划、共享单车租借数量模型、共享单车流动动态 OD 矩阵、共享单车行驶时间模型、共享单车停放模型。城市单车区域规划是整体模型的基础, 共享单车的租借行为不仅具有时间的差异性, 还具有空间的差异性, 不同区域的单车租借情况是不同的, 同时单车的损坏率也会影响到租借行为, 这些都是模型中要考虑的因素。共享单车的流动模型包括对共享单车流动动态OD 矩阵和共享单车行驶时间模型, 这两个模型确定了乘客不同时间的目的倾向性以及单车的流动时间。共享单车停放模型是对单车在某一目的区域内停放的考虑, 停放的行为会影响到共享单车的租借。

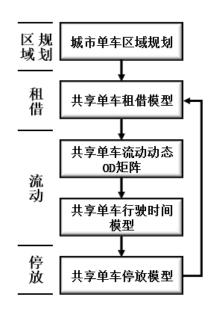


图 1 模型框架

#### 2.2 城市单车区域规划

城市内的共享单车一般按区域进行投放,这些区域是居民经常出入的地方,多集中在交通枢纽、商业区、公司等位置。但随着单车的流动,很多自行车会被无序停放,难以准确判断其所属区域。这里通过聚类方法构建单车停放区域。

聚类算法是一种典型的无监督学习算法,主要用于将相似的样本自动归到一个类别中。在聚类算法中,对于不同的相似度计算方法,会得到不同的聚类结果,对于共享单车的空间聚类,我们以欧式距离法作为相似度计算方法。其算法为:

- $\blacksquare$  初始化类别个数 K, 随机选取初始点为质心;
- ▲ 重复计算以下过程,直到质心不再改变;
  - ▶ 计算样本与每个质心之间的相似度,将样本归类到最相似的类中;
  - ▶ 重新计算质心;
- → 输出最终的质心以及每个类。

类别个数 K即区域个数由下式确定:

$$K = \frac{S_{city}}{d_{\min}^2}$$

其中 $s_{city}$ 为研究区域面积, $d_{min}$ 为正常情况下最短骑行距离。考虑到乘客在 200 米范围内的很少会选择骑车,所以这里 $d_{min}$ =300。

#### 2.3 共享单车租借模型

乘客租借共享单车的行为是有选择性的。不同时间、不同空间共享单车的租借情况是不同的。在早高峰和晚高峰时期,乘客对共享单车的需求量会很高,而在凌晨城市内流动的单车寥寥无几。大多乘客将共享单车作为其它交通工具的补充,在公交站、地铁站共享单车的流动频率会更高。时间和空间对单车租借情况的影响常常表现为时段性和空间区域性。共享单车的损坏是影响乘客租借的重要因素,单车的损坏率达到一定程度会严重影响乘客的满意度,从而导致其流通率的下降。以下租借模型考虑了在一定程度损坏率的情况下,不同时间段,不同空间区域的单车租借情况。

将一天时间 $t_{day}$ 划分成若干时间段 $t_1,t_2,...,t_r$ ,以单车的规划区域为基础,在每一区域每一时间段内共享单车在单位时间租借数量为:

$$r_i^t = \frac{r_{outi}^{t_r}}{t_{..}}$$

其中 $r_{outi}^{t_r}$ 为在 $t_r$ 时间段内从区域i出去的单车数量,每个区域内的共享单车是有损坏的,而且在停止投放的情况下,单车的损坏程度会逐渐加重。 考虑到共享单车的损坏情况,每个区域内可骑单车数量为

$$N_e = (1 - \lambda)N$$

λ为单车的损坏程度, N为区域内单车数量。单车的损坏率会影响到乘客对共享单车的依赖, 从而专向其它交通工具。假设租借量与区域内单车损坏程度正相关,则区域内实际租借量为:

$$r_{ir}^{t} = (1 - \lambda)r_{i}^{t}$$

#### 2.4 共享单车流动模型构建

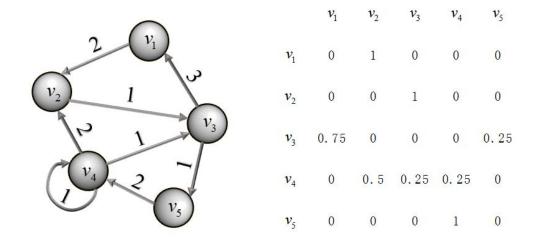
#### 2.4.1 OD 矩阵

OD(origin-destination)矩阵是记录交通网络中从起点到终点的出行概率的二维矩阵,反映了人们的出行习惯。动态 OD 矩阵则表示 OD 对在不同时段对交通的需求。通过建立共享单车的动态 OD 矩阵,可以研究居民的出行目的情况。

OD 矩阵以区域间的流动情况为基础,通过构建区域间流通网络G(V,E,W),确定共享单车从某一区域到另一区域的出行概率。在构建的网络图中 $V=\{v_1,v_2,...,v_N\}$ 表示网络中节点集合,每个节点代表每个划分区域,其中节点数为N;  $E=\{e_1,e_2,...,e_M\}$ 表示网络中连边集合,网络图中任意一条边 $e_i$ ,都会有其对应的一对节点 $(v_{i1},v_{i2})$ 存在,代表乘客从一区域到另一区域的出行路径,是有向的; $W=\{w_{e1},w_{e2},...,w_{eM}\}$ 表示网络连边的权重集合,每一条边有一权重,表示该边的连接次数,即乘客从一区域到另一区域的出行次数。图 2 (a) 为一流动网络示意图,其 OD矩阵表示为图 2 (b)。OD矩阵的每一行代表从某一区域到其它所有区域的出行概率,每一列代表某一区域吸引其它区域乘客到达的概率。OD矩阵的每个概率值在时间t内的计算方法为:

$$p_{ij} = \frac{w_{ij}}{\sum w_{i.}}$$

其中 $p_{ij}$ 表示从区域i移动到区域j的概率, $w_{ij}$ 为流通网络中从区域i到区域j的移动次数, $w_{i.}$ 为所有从区域i流出到其它区域的总次数。



(a) 流动网络示意图

(b) OD 矩阵

图 2 网络到 OD 矩阵的转化示意图

#### 2.4.2 共享单车行驶时间模型

在不考虑共享单车从解锁到开始骑行、从到达目的地到关锁的这两段时间,共享单车的行驶时间决定了其从租借到停放的这一完整流动时间。一般地,行驶时间按照下式计算:

$$t_{run} = \frac{d_{OD}}{v_{bike}}$$

其中, $d_{oD}$ 表示起始点到终点的行驶距离, $v_{bike}$ 表示共享单车的行驶速度。但是共享单车真实行驶路线的距离难以计算,考虑到北京城区交通完善,道路多呈南北走向、东西走向和环线,所以从起始点到终点的路线大多呈现为横纵的直线,假设所研究区域为一近似平面,通过经纬度,两点间的距离可近似为

$$d_{OD} = 2R \left( \left| \sin(\frac{\pi}{360} (long_D - long_O)) \right| + \left| \sin(\frac{\pi}{360} (lat_D - lat_O)) \right| \right)$$

其中  $long_D$ 、 $lat_D$ 分别为终点经纬度, $long_O$ 、 $lat_O$ 分别为起始点经纬度,R 为 地 球 半 径 , 这 里 取 R=6371km 图 3 为 在 百 度 地 图 (https://map.baidu.com/) 中选择的一对起终点,蓝色线表示利用百度地

图对起终点进行骑行路线规划路线,距离为25.0公里,红色线表示其利用公式计算距离为25.7公里。

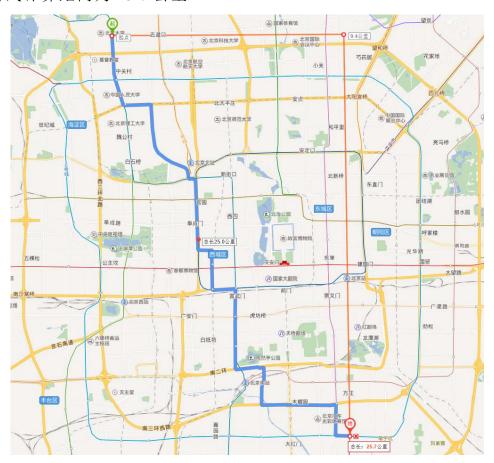


图 3 百度地图骑车路径示意图

图 4 为在百度地图上随机选择 20 对起终点,分别计算其骑车规划距离和模型中的距离及其差值,可以清楚地发现其距离相差很小,可以用模型中的距离近似为两点之间的骑车路线距离。

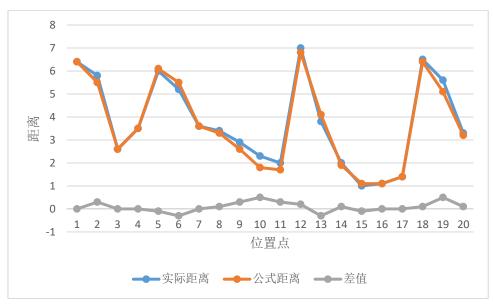


图 4 20 对起终点的实际距离、模型距离及其差值

#### 2.5 共享单车停放模型构建

当乘客到达目的地后,便会将共享单车停放在其目的地附近的某一位置。很多乘客选择停放点是有聚集效应的,更愿意选择共享单车聚集的地方,但也会有乘客为了方便,选择就近停车。正是由于乘客目的的多样性才导致了单车停放的无序性,但这种情况也是需要时间积累的。模型假设一天内,各个区域内单车停放位置的离散程度波动较小,趋于不变。这里定义区域内单车位置的离散程度为:

$$\varphi = \sum_{i=1}^{N} \left(d_{cp_i} - \overline{d_{cp}}\right)^2$$

其中,N为区域内单车的个数, $d_{cp_i}$ 为区域内的第i个小车到区域中心点的距离, $\overline{d_{cp}}$ 为所有小车到中心点的平均距离。假设的可行性会通过之后的真实数据加以验证。这样我们的模型可以以区域为研究对象,认为共享单车到达某一目的区域后,停放的位置在该区域上是均匀分布的。

## 2.6 模型假设

在一天内不存在人工对共享单车地异地拖运;

共享单车行驶中不存在因非自然因素,如修路,堵车,而导致的自 行车绕路形式情况;

任意两区域间的骑行路径不存在绕路情况;

当研究区域在城市级别,可近似其为一平面区域,而非球面;

乘客从将共享单车解锁到开始骑行和从到达目的地到关锁的时间忽略不计;

所研究的一天内天气良好,不是酷暑、严寒天气,也没有雨雪大风; 任意一区域内的共享单车数量能够满足乘客的使用需求;

#### 第三章 数据获取及数据分析

#### 3.1 数据获取及处理

网络爬虫是一种按照一定的规则,自动地抓取万维网信息的程序或者脚本。从 2017 年 1 月份微信上线了小程序功能,相应的很多知名平台纷纷推出了自己的微信小程序,OFO 共享单车也在其中,这意味着对于数据的获取又多了一个数据来源。我们根据微信小程序的 API 接口获取 OFO 位置数据。

#### 3.1.1 爬虫机制

从技术层面来说就是通过程序模拟浏览器请求站点的行为,把站点返回的 HTML 代码/JSON 数据/二进制数据爬到本地,进而提取自己需要的数据。其流程如图 5 所示,分为发送请求、获取相应内容、解析内容、保存数据四部分。发送请求就是用户将自己的信息通过浏览器发送给服务器,一般发送请求包括请求方式、请求 URL、请求头和请求体,其中请求方式包括 GET 方式和 POST 方式,请求 URL 用来定位互联网上一个唯一的资源,请求头和请求体包含了请求所需的属性和参数;发送请求后服务器接收请求,通过分析用户发来的请求信息,然后返回相应的状态码和内容;收到内容后通过本地解析得到所需数据,最后将数据以文本形式保存或者存储在数据库中。

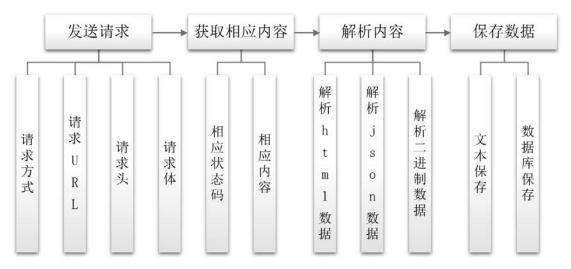


图 5 爬虫机制

#### 3.1.2 爬虫策略

我们通过爬取微信 OFO 小程序数据,获取一天内 OFO 静态位置。 打开微信 OFO 小程序可以看到如图 6 的界面,界面中可以找到某一位置 附近的所有可骑单车位置,爬虫的程序中就是利用这种方法。在北京(经 纬范围) 内等间隔选取(多少个)位置点,获取其周边静态小车。



图 6 0F0 界面

我们采用 Python 进行数据采集程序编写,采用了 Beautiful Soup 框架。首先确定一个初始位置,通过 URL 管理器获取一个待爬取 URL,根据 URL 定位网页位置,利用 urllib2 下载器下载请求内容,对下载内容解析成所需数据格式后进行存储。然后在所选区域内顺序地更换位置点,遍历完整个区域,直至爬取完一天的数据。

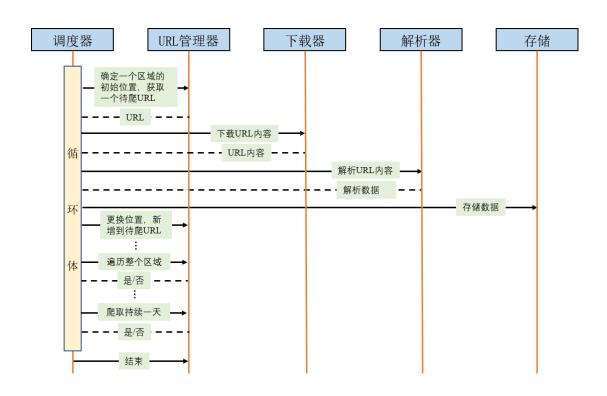


图 7 爬虫程序流程

#### 3.1.3 数据处理

通过爬取 OFO 小程序, 我们获取到北京(经度范围: 116.179981828~116.56741803, 纬度范围: 39.7569784355~40.034994274) 一天内的小车静态位置, 其数据格式为(部分数据见附录):

字段 (示例)	BPl08n	116.294719898	39.9665278571	2017-12-06 00:00:00
含义	单车 ID	単车所在经度	単车所在纬度	探测到单车在该位置
A X				的时间

表格 1 爬取数据字段及其含义

由于获取到的部分数据格式并不完全正确,存在缺少字段或者数据 不全的现象,而且数据要经过一定的处理才能作为分析的依据。

- 1) 去除缺少字段或者数据不全的记录: 缺少字段或者数据不全的记录缺少必要信息, 会干扰后续的分析。
  - 2) 去除由于 GPS 跳动或者人为短距离移动的数据: GPS 信号是不

稳定的,定位的位置会有跳动,几米到几十米不等;而且当小车处于锁定状态是会有人为的搬动,同样会导致单车的非骑行位置变化,人为在100米以内居民更愿意步行到达目的地,而非骑车。去除共享单车移动距离小于100米的数据。

3)去除不在首次全部覆盖探测到的单车:由于获取的数据是记录单车的静态信息,移动的单车是不会被记录的,所以第一次没有探测到的单车可能会出现在之后的探测。以第一次完整覆盖研究区域所探测到的单车为所有单车的集合,去除之后探测中不在这一集合中的单车。

#### 3.2 数据驱动方法

互联网时代使得数据获取变得容易,利用数据驱动模型的方法成为模型建立的重要依据。利用共享单车原始位置数据,通过重构其运动轨迹,提取共享单车的运动的时间、空间分布以及统计分布等,为模型提供数据支撑。在第二章介绍的模型中区域规划、OD矩阵、区域单车租借量将利用真实数据进行构建。

#### 区域规划:

共享单车并不是在任何位置都会有停放的,为了找出单车一般停放点,将处理后的一天数据中所有共享单车所在位置作为 k-means 聚类样本,找出其所有可能停放位置。所研究的北京五环区域面积约为 1021.982 平方千米,因此聚类个数为 10000。将每一类作为一个单车集中区域,聚类中心作为区域中心。

#### OD 矩阵构建:

虽然获取到的共享单车位置是静态的,但通过单车的标识 ID 可以提取每个单车在不同时间的位置变化。根据时间先后构建单车在区域间流动轨迹,通过轨迹的连通情况和连通次数提取区域间流通网络,依据网络构建 OD 矩阵。

#### 区域单车租借量:

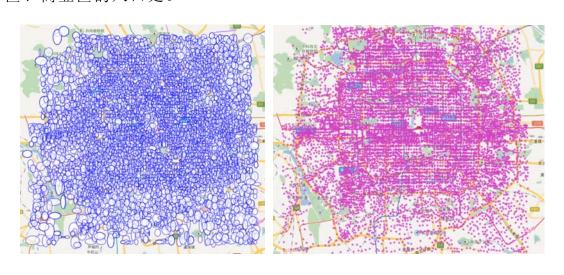
将上述构建的区域间流通网络按照时间进行切片,每一片的时间间

隔为一小时,每一片网络中每个区域节点的出度作为在该时间段该区域单车的租借数量。

#### 3.3 数据分析及模型合理性验证

#### 3.3.1 聚类分析

图为对获取到的北京五环一天的停放位置进行聚类,图 8 (a) 为聚类区域,图 8 (b) 为聚类区域的中心点。从图中可以看出,单车的停放区域在地图上表现近乎均匀,除了部分禁止停放单车的区域,如故宫,圆明园等景点区域,共享单车几乎遍布了整体区域。聚类区域越大,则说明该区域的单车分布较为分散,可以看出五环周边区域以及南部区域分布较为分散,其它区域则更为集中。从区域中心在地图上的分布可以看出,共享单车多集中在道路沿线,地铁站、公交站、校园以及一些景区、商业区的入口处。



(a) 聚类区域

(b) 聚类中心点

图 8 北京五环区域规划

## 3.3.2 共享单车租借分析

图 9 是一天内北京五环内各个区域共享单车被租借骑出的数量,其

中的小图是骑出数量大于 191 的区域,从图中可以看出不同的区域对共享单车的需求是不同的,有着明显的差异,其北部地区对共享单车的需求明显高于南部地区,尤其是在西城区和朝阳区表现明显。

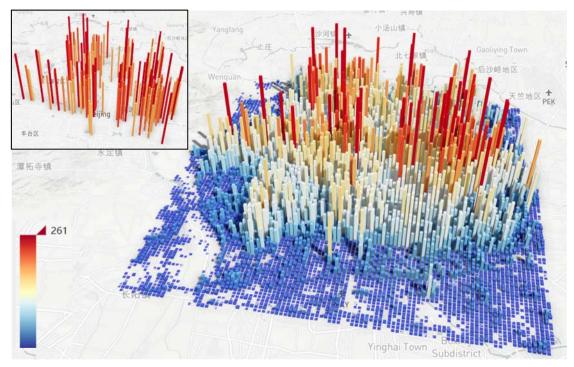


图 9 北京五环内各个区域共享单车被租借骑出的数量

图 10 为一天内各个时间段内共享单车被租借的数量,从图中可以看出,乘客的租借行为是有时间性的,在晚上的时间(00:00-05:00,21:00-00:00)人们出行较少,在白天会成倍增长,尤其在 7:00-10:00 和 18:00 左右,单车出行会出现高峰期,但这两个高峰期并不对称,乘客更倾向于早晨出去而后陆续返回。

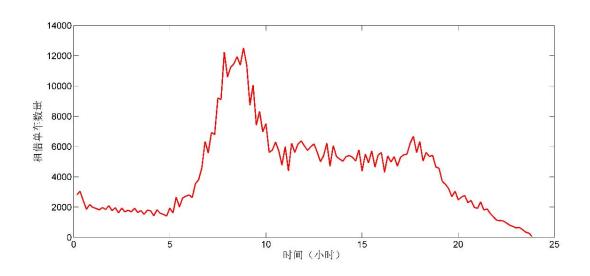


图 10 一天内各个时间段内共享单车被租借的数量

#### 3.3.3 共享单车流动分析

根据提取的共享单车出行 OD 矩阵,我们从中选出了乘客出发频次最高和乘客更倾向到达的区域,图 11 表明乘客出发位置和目的位置并不是随机的,而是倾向于某些热点区域,这些区域主要集中在学校、商业区、交通站点、景区等地方。而且这两部分区域在分布上具有相似性,一方面说明了共享单车是一种短途交通工具,乘客起终点距离较短。另一方面则说明热点区域流通的均衡性,保证了较小可能的某区域单车积累。

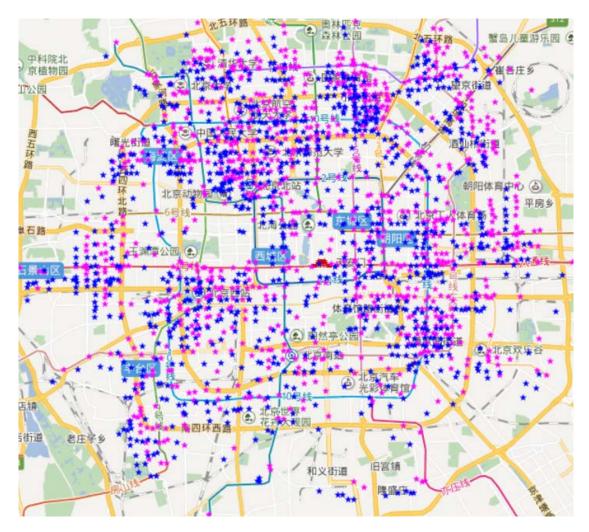


图 11 乘客出发频次最高区域(红星)以及乘客更倾向到达的区域(蓝星)

## 3.3.4 共享单车停放分析

为了验证停放模型,我们根据一天内探测的共享单车停放位置数据, 提取了一天初始时刻和结束时刻各个停放单车的位置。每一个单车有一 个所属区域,计算每一个区域内的单车空间分布离散程度。图 12 显示了 一天初始时刻和结束时刻各个区域内单车停放离散程度的差值分布,从 分布图可以看出,其大致为均值为零的正态分布,尖峰处的最大概率达 到了 0.55,可以认为,一天内各个区域内的单车分布是不变的。

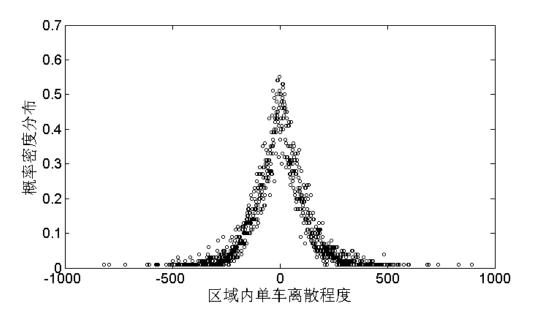


图 12 区域内单车离散分布图

#### 第四章 仿真系统设计

基于第二章的模型设计,完成了共享单车流通情况仿真分析系统,本章对系统功能及实现工作进行阐述。本系统在 Windows10 系统环境下,采用 Eclipse 通过 Java 编程语言进行开发,运行环境要求: 四核 CPU 2.5GHZ 以上,内存 8G,显示器分辨率 1920\*1080p, JRE (Java 运行环境) 1.8 及以上。

仿真引擎采用时间步进的方式推进仿真,每一仿真步长对应 1800 秒, 仿真总时长为 24 小时。

## 4.1 仿真系统设计

仿真系统主要分为以下两部分: 仿真系统初始界面和仿真实验界面。 其中仿真实验模块主要包括仿真过程可视化模块、仿真控制模块和仿真 参数显示模块。仿真系统采用数据驱动的方法,输入数据为区域规划结 果、OD矩阵、以及区域单车单位时间租借量,仿真结束后将结果输出至 文件中,以便后续分析。系统用例图如下 13 所示。

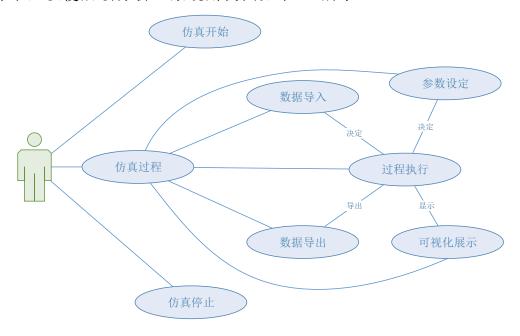


图 13 系统用例图

系统主要有以下功能:系统的开始与停止、数据的导入、仿真参数设定、仿真执行、可视化展示和结果数据导出。其中的功能相互作用相互影响:数据的导入和仿真参数的设定决定后续的仿真执行;将仿真执行的步骤进行显示做到可视化展示;仿真执行的结果将其进行数据导出以便进行下一步分析;数据导入、参数设定、仿真执行、可视化显示和数据导出共同组成了系统的仿真过程。

#### 4.2 仿真系统初始界面



图 14 仿真系统初始界面

初始界面主要是对仿真系统的初始展示,包括系统的名称、制作人和 系统控制按键。

## 4.3 仿真实验界面

仿真界面总览如下图 15 所示:

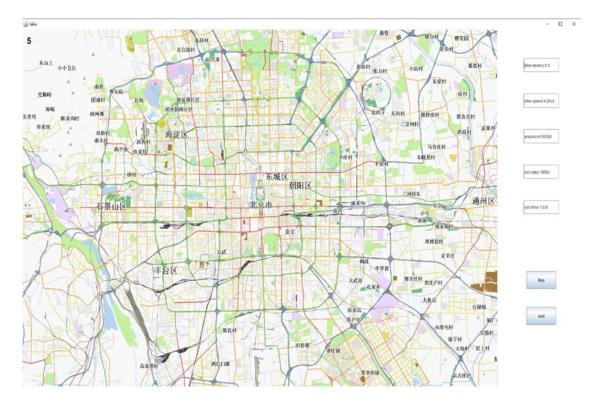


图 15 仿真界面总览图

#### 主要包括以下几个模块

(1)仿真控制模块。在控制模块中包含两个控件"Run"控件和"exit" 控件。"Run"控件可以控制仿真过程的开始,按下控件后启动仿真程序,并在仿真过程可视化模块中动态的显示仿真过程。"exit"控件可以控制仿真系统的退出,按下后退出仿真过程。

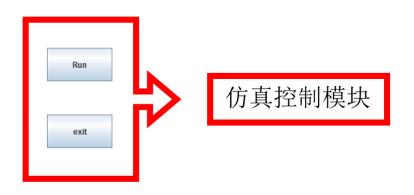


图 16 仿真控制模块图

(2) 仿真参数显示模块。在此模块中,主要显示了仿真系统的一些可调参数,主要包括共享单车的损坏率、单车的移动速度、仿真聚类数量、仿真步长和仿真时间。



图 17 仿真参数显示模块图

(3) 仿真过程可视化模块。在模块中导入了北京市的地图,并且可以通过鼠标滚轮控制地图比例尺。在仿真开始后在地图上现实单车的聚 类点和单车的流动情况,并且可以在不同的大小的地图上查看仿真情况。

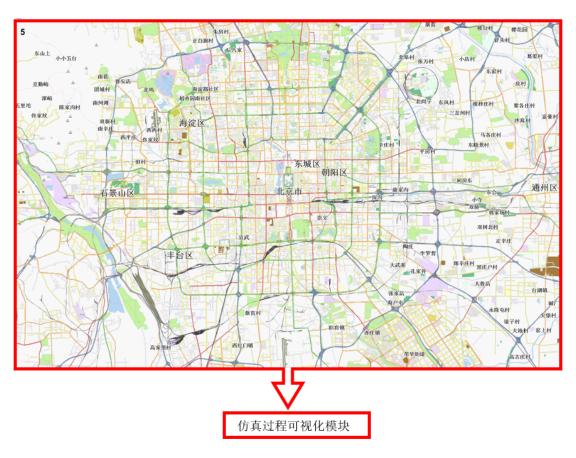


图 18 仿真过程可视化模块图

#### 4.4 软件使用说明

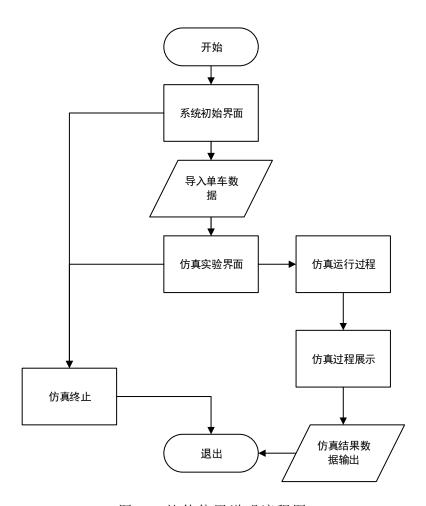


图 19 软件使用说明流程图

这里给出整个软件使用的简单说明。如上图 19 所示。打开系统后首先进入系统初始界面,其中包含"begin"和"exit"两个按钮。按下"exit"按钮后仿真终止,退出仿真系统。按下"begin"后进入仿真实验界面,在其中显示了此次仿真的一些默认参数包括:共享单车的损坏率、单车的移动速度、仿真聚类数量、仿真步长和仿真时间。控制模块中包含"run"和"exit"两个按钮,按下"exit"按钮后仿真终止,退出仿真系统。按下"run"后启动仿真,在仿真执行的过程中,在仿真仿过程可视化模块中可视化的显示仿真动态过程。由于数据量过大,仿真结束后自动将仿真的数据存入文件中以做分析。

#### 第五章 实验及分析

#### 5.1 实验设计

利用第四章开发的共享单车一日内流通情况仿真系统,我们进行了 两组设计实验,以验证模型的正确性和分析共享单车流通时存在的问 题。

实验一:从共享单车流通情况的时间性和空间性分析了共享单车骑行特点以及存在的问题,同时依据统计数据验证模型的正确性。在时间分析了随时间变化的累计出行次数和骑行时间的分布;在空间上对比分析了真实单车停放数据和仿真流通数据在各个区域骑出单车数量和骑入单车数量差值的分布。

实验二:利用共享单车一日内流通仿真系统,分析了单车损坏率对流通的影响。

#### 5.2 实验结果分析

累计出行次数:居民在 06:00 之前和 20:00 以后共享单车使用率很少,高峰期在白天,尤其是 07:00 到 10:00 最为明显,这一结果符合实证数据结果。

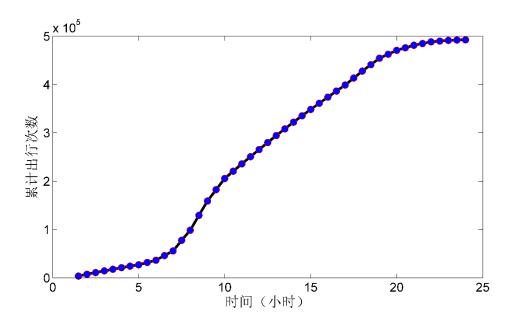


图 20 累计出行次数分布图

**骑行时间概率分布:** 乘客骑行时间主要集中在 1—10 分钟,大约是在几公里范围内,充分说明了共享单车在短途上发挥了重要作用,但同时也暴露出了在骑行时间上共享单车使用的资源浪费。骑行时间在十分钟之内的单车占了 85%,短时间骑行加大了共享单车的周转频率,但并没有使单车充分被利用。

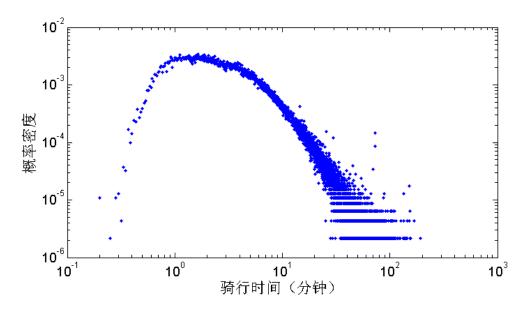


图 21 骑行时间概率分布图

各个区域骑出单车与骑入单车对比:图显示了在一天内每个区域内骑入量与骑出量之差。其中红线为真实数据,蓝线为仿真数据。图中表明区

域内骑入骑出相对均衡,只有部分区域会出现骑出量大于骑入量和骑入量大于骑出量,也说明这些区域是需要关注的区域,随着时间的推移,会导致区域单车分布不均,出现单车积累和匮乏的区域,这就需要管理者把握重点区域,进行合理的单车调度。

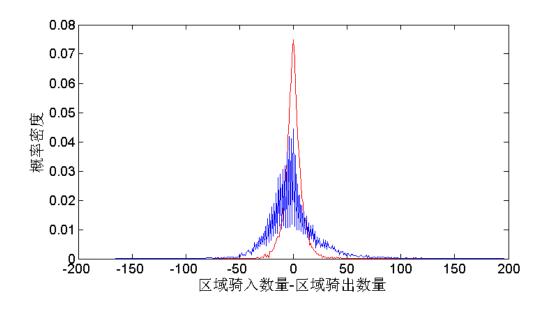


图 22 区域单车骑入-骑出对比图

图 23 (a) (b) 分别为仿真数据和真实数据骑入量减骑出量的空间分布热力图,图 23 (c) (d) 分别为仿真数据和真实数据骑出量减骑入量的空间分布热力图,这两幅图从空间全局上给出了会产生问题的重点区域。另一方面,通过真实数据与仿真数据的对比,说明了模型的合理性。



图 23 仿真数据和真实数据骑入量与骑出量差值的空间分布热力图

区域内停放单车数量随时间变化情况:图 24 分别为 2 时、5 时、8 时、11 时、14 时、17 时、20 时、23 时的各个区域内共享单车停放数量。从图中可以看出各个时刻区域内单车数量分布相似,只是在出行高峰时刻各个区域内单车数量明显减少,减少的区域多为北部地区,但最终分布又会和起始时刻相似。说明了在一天内共享单车的流通并不会出现由于出行而带来的明显的空间分布不均,但随着长时间的积累,会导致此类现象的发生。

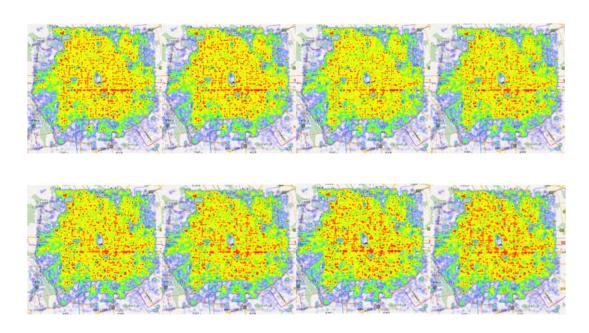


图 24 区域单车停放数量随时间变换热力图,从左至右时间依次为2时、5时、8时、11时、14时、17时、20时、23时

共享单车损坏率对出行的影响:这里从两方面考虑了共享单车的损坏率对单车的租借影响,如图 25。当乘客的骑车意愿不受单车损坏程度影响时,损坏率对活跃单车出行频次影响为图中红线。即使损坏程度很大,仍有很多可骑行的单车,出行频次仍然很高。一旦乘客的骑车意愿受单车损坏程度影响,那么很可能会找寻其它替代工具前往目的地,严重影响到单车的使用。损坏率对活跃单车出行频次影响为图中蓝线。对待损坏的即使维修,防止出现大面积报废单车对共享单车的运营尤其重要。

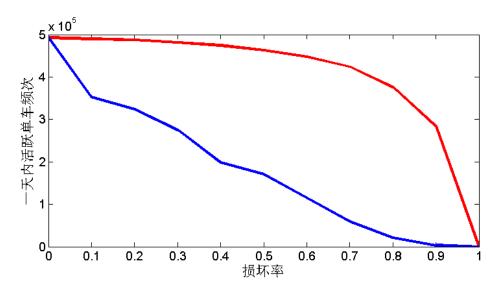


图 25 共享单车损坏率对单车活跃量的影响

#### 第六章 结论

本文以共享单车当前所暴露出来的诸多问题,如损坏严重、资源浪费、分布不均等进行研究分析。我们以北京市五环以内为研究区域,根据对共享单车出行特点进行分析,利用数据驱动的方法构建了共享单车一日内的停放流动模型,即租借一流动一停放模型,全面地分析了共享单车流动的全过程。通过对共享单车实证数据的研究,我们验证了模型构建的合理性。基于模型的仿真结果表明,共享单车的确存在区域分布不均,资源浪费问题,但也发现在一天内不会产生分布不均的问题,而是在长时间的积累下形成的。同时我们给出了在一日内共享单车出入不均的区域,这些区域会成为单车分布不均的主要因素,应对这些区域安排好单车调度。另外,通过对单车的损坏率对居民骑行出行的影响,我们发现当乘客出行意愿容易受到单车损坏度的影响而选择其它交通工具时时,单车的使用率会严重下降,定期对单车的维修同样对单车的运营有着不可忽视的作用。

#### 参考文献

- [1]. 李敏莲. 共享单车市场调研与分析[J]. 财经界(学术版), 2017(5).
- [2]. 薛强. 国人共享单车使用情况调查[J]. 金融博览(财富), 2017(1):24-26.
- [3]. 速途研究院. 2017 年第一季度国内共享单车市场调研报告[J]. 互联网天地, 2017(4):21-24.
- [4]. 朱萍, 朱亚成, 董雨薇,等. 2017-2018 中国共享单车发展报告[J]. 中国商论, 2017(31):143-147.
- [5]. 徐鑫垚. 共享单车 App 后台管理系统的优化[J]. 电子技术与软件工程, 2017(4):80-81.
- [6]. 李叶. 关于共享单车用户体验情况及现存问题的调查分析——以对某高校附近的调研为例[J]. 经济研究导刊, 2018(3):193-194.
- [7]. 闫东. 共享单车行业市场分析[J]. 环球市场信息导报, 2017(30):6-6.
- [8]. 孙迎坤,李永平,吴林芳. 互联网+分享经济模式下共享单车发展存在的问题及监管对策——以石家庄市 ofo 单车为例[J]. 现代商业,2017(31).
- [9]. 刘亚楠. 共享单车发展研究分析[J]. 时代金融, 2017(8).
- [10]. 余国磊. 浅析"共享单车"运营和管理中存在的问题与对策[J]. 知识经济, 2017(9):87-88.
- [11]. 李艳. 共享单车呼唤"共享诚信"[J]. 中国商界, 2017(3):90-91.
- [12]. 陈康, 景丽丽. 共享单车乱象原因分析及对策研究[J]. 经贸实践, 2017(3).
- [13]. Graf F, Kriegel H P, Renz M, et al. MARiO: Multi-Attribute Routing in Open Street Map[M]. Advances in Spatial and Temporal Databases. Springer Berlin Heidelberg, 2011:486-490.
- [14]. Wang M, Li Q, Hu Q, et al. Quality Analysis of Open Street Map Data[J]. ISPRS International Archives of the Photogrammetry, Remote

Sensing and Spatial Information Sciences, 2013, XL-2/W1:155-158.

- [15]. Anelli V W, Calì A, Noia T D, et al. Exposing Open Street Map in the Linked Data Cloud[C]. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Springer International Publishing, 2016:344-355.
- [16]. Fellows I. OpenStreetMap: Access to Open Street Map Raster Images[J]. 2016.
- [17]. 李伯虎, 王行仁. 综合仿真系统研究[J]. 系统仿真学报, 2000, 12(5):429-434.
- [18]. 黄柯棣. 系统仿真技术[M]. 国防科技大学出版社, 1998.

#### 附录:

3nxmNa,116.264140658,39.9714968889,2017-12-06 00:00:00 ozogyb,116.263891432,39.9716797446,2017-12-06 00:00:00 b6qXMK,116.256422016,39.9706081818,2017-12-06 00:00:00 dY4Z1B,116.255312212,39.9696231247,2017-12-06 00:00:00 zD7dx0,116.271180994,39.9687096439,2017-12-06 00:00:00  $qmk2qN, 116.254118492, 39.968729263, 2017\text{-}12\text{-}06\ 00\text{:}00\text{:}00$ GxXgVX,116.254785889,39.9703683102,2017-12-06 00:00:00 6L6nVP,116.274121745,39.9663655471,2017-12-06 00:00:00 DpZoAa,116.27505221,39.9653251666,2017-12-06 00:00:00 5w2JEm,116.274693698,39.9668772996,2017-12-06 00:00:00 wA8M3X,116.2746886,39.964935301,2017-12-06 00:00:00 Z5LomL,116.274881146,39.9673197605,2017-12-06 00:00:00 GV3GG6,116.279564747,39.9659618223,2017-12-06 00:00:00 A5zYa4,116.278160622,39.9663726337,2017-12-06 00:00:00 xv8wOE,116.278190654,39.9664356613,2017-12-06 00:00:00 YqpkxN,116.279936667,39.9659300312,2017-12-06 00:00:00 M0zR7d,116.277999524,39.9666285099,2017-12-06 00:00:00 7XGKDJ,116.28255861,39.9664336307,2017-12-06 00:00:00 Pekl5L,116.282589651,39.9665106642,2017-12-06 00:00:00 DM5amB,116.282733035,39.9665588853,2017-12-06 00:00:00 PMKyEB,116.282558628,39.9665826366,2017-12-06 00:00:00 3Kewvp,116.283307388,39.9665533793,2017-12-06 00:00:00 mpZxwJ,116.282971063,39.9667200495,2017-12-06 00:00:00 eMKrlO,116.283375467,39.9666214505,2017-12-06 00:00:00 Qedowd,116.283864942,39.9663129355,2017-12-06 00:00:00 A7wBVD,116.281922945,39.9665305416,2017-12-06 00:00:00 NwrbqG,116.281855462,39.9664741547,2017-12-06 00:00:00 wb7O3B,116.281954137,39.9665747355,2017-12-06 00:00:00 8anqZ2,116.281764816,39.9663418584,2017-12-06 00:00:00 lwypL2,116.281783281,39.9664933687,2017-12-06 00:00:00 wvNYqD,116.281748823,39.9665248504,2017-12-06 00:00:00 VdaYvg,116.281255019,39.9664009716,2017-12-06 00:00:00 mkJgDM,116.287232778,39.9665115704,2017-12-06 00:00:00 6Nw5qQ,116.286828312,39.9666441213,2017-12-06 00:00:00 PMmAQv,116.287760304,39.9655961345,2017-12-06 00:00:00 56pkb7,116.287441859,39.9667129101,2017-12-06 00:00:00 p3VLmo,116.287004365,39.9669407889,2017-12-06 00:00:00 Z4ALZB,116.288108999,39.9666365909,2017-12-06 00:00:00 X7NAw1,116.288026319,39.9667592056,2017-12-06 00:00:00 Rp6mAY,116.28806371,39.9667295989,2017-12-06 00:00:00 KKopPA,116.288251365,39.9667477656,2017-12-06 00:00:00 wvJwzD,116.288506352,39.9665910369,2017-12-06 00:00:00 M3qZNx,116.2885414,39.966621079,2017-12-06 00:00:00 yMx24b,116.288616491,39.9665931658,2017-12-06 00:00:00 zDVGZK,116.288736849,39.9665034799,2017-12-06 00:00:00 wbVpq1,116.28856459,39.9667816102,2017-12-06 00:00:00 GVVe0X,116.289829911,39.9654955664,2017-12-06 00:00:00 kwyzEo,116.290215404,39.9651209232,2017-12-06 00:00:00 674m2X,116.289834991,39.9667924884,2017-12-06 00:00:00 xRqkyz,116.290828663,39.9645050186,2017-12-06 00:00:00 bJOBV4,116.289939867,39.9673583701,2017-12-06 00:00:00 Np0y4n,116.288767742,39.9665037745,2017-12-06 00:00:00 VdzXLe,116.288782139,39.9665493719,2017-12-06 00:00:00 qBPr0Y,116.288837725,39.9666939621,2017-12-06 00:00:00 0BXO8D,116.288810755,39.9667464,2017-12-06 00:00:00