共享单车管理优化策略研究

**摘 要：**为应对共享单车带来的停车乱象、利用率低等问题，提出一种阶梯型分区收费的管理模式，拟为共享单车的优化整改提供参考。不失一般性，以浙江大学紫金港校区为研究区域，利用层次分析法对影响单车停放的适宜性因素进行分级和叠加，利用GIS软件得到适宜性分级图，与共享单车实际停靠站点的聚类结果进行比较，从而得到P/M/F分区。利用计算机模拟一段时间内一些用户可能使用单车的情况，以企业利润、单车使用率和单车停放混乱度为衡量指标，得到不同参数下的最优定价，并从多个角度验证了阶梯型分区收费的管理模式优于原先统一型的管理模式。

**关键词：**阶梯型分区收费；计算机模拟；层次分析；聚类；

# 引言

共享经济在近年的发展极大的丰富和方便了我们的个人生活，同时为社会经济发展提供了桥梁。然而共享经济所存在的不合理处将导致社会资源不能被物尽其用，并且对消费者和企业都造成困扰。共享单车作为一种新型的交通工具租赁业务，其主要依靠载体自行车，旨在充分利用城市因快速的经济发展而带来的自行车出行萎靡状况，最大化的利用公共道路通过率。然而，如今共享单车被随意停放已经成为扰乱交通、影响出行便利、加大管理者负担的重要因素。

针对上述问题，本文提供了一种解决思路——阶梯型分区收费的管理模式。其内在机制是将单车停放位置与用户个人利益结合起来，通过人“自利”心理的内在驱动，从根本上解决单车停放问题。具体来讲，我们将单车的停放区域分为三类，P区，F区和M区。P区为最佳停车区域，这是附近区域中单车使用率最高的地方，并且被政府划分为可停放非机动车的场所。F区为禁停区域，通常为草地、花圃、道路、水沟等场所，大多偏僻或阻碍交通。M区为次停车区域，即扣除P区F区之外的区域，大多人流较稀疏，单车使用率较低，或影响市容市貌，这片区域较大，政府和企业的管理维护成本较高，故不建议停车。根据单车所处区域的不同，来设置三种不同的收费标准，P区、M区、F区收费依次增高。

这种管理模式看似影响单车的自由流通，可能导致企业损失大量客户，实则不然。共享单车更多的被停放在人流密集而非偏僻无人的场所，提高了单车的使用率，不仅给企业增加利润，而且达到了资源共享的初衷。这种模式与“入栏结算”有本质上的区别，它并不禁止你停放，而是要求你为你的方便承担应有的费用。这种奖惩分明的方式能吸引更多的顾客，因为它把运营成本按用户的需求来划分，而非平均分摊到每个用户身上。

文献中针对于单车选址的研究较多，利用出行数据，分别应用了聚类算法、蚁群算法、机器学习等方法。然而，文献中对于综合考量选址和定价因素的研究少之又少，本文力图填补这一研究缺口。因此，本文力图给出一个通用的方法，在考虑共享单车管理优化的策略设计时，解决共享单车停放的层次型区域划分和阶梯型收费定价的相关问题。这种管理模式需要综合考虑用户的出行便捷性、企业的商业效益、政府的城市规划设计，这是由单车使用率、企业利润和单车停放混乱度来衡量的。

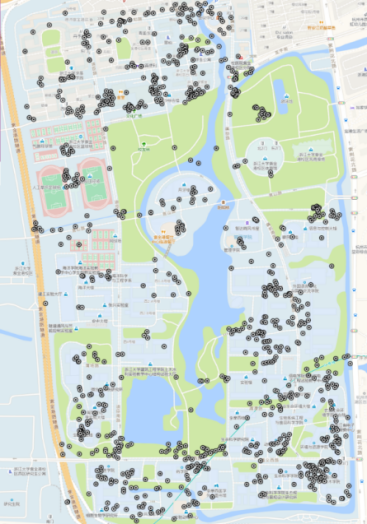
# 层次型区域划分模型

针对校园内共享单车杂乱停放的问题，本文提出了一个层次型区域划分模型，即根据不同区域的单车密度大小将校园区域依次划分为三部分：P区、M区、F区。P区，即最佳停车区，就是共享单车需求量最大的区域（也是人口流通量最大的区域）。M区，即次停车区，相对于P区而言M区的共享单车需求量并不大，但是却是校园单车停放混乱不堪现象的主力军。F区，即禁停区，它的来源主要有两个方面：一方面是不可能停放区域，例如建筑物、湖泊等；另一方面是学校出于管理和学园秩序的考量所设置的禁停区，例如蓝田、丹青等住宿区。根据消费者理论，当停放在P区的骑行费用比停放在M区的骑行费用更低时，部分原本停在M区的用户会选择转而停在P区。同样地，当F区的骑行费用十分昂贵时，大部分用户不再考虑将单车停放在禁停区，从而使得校园的单车停放位置更加整洁有序。接下来笔者将采用聚类算法和层次分析法生成层次型区域划分模型。

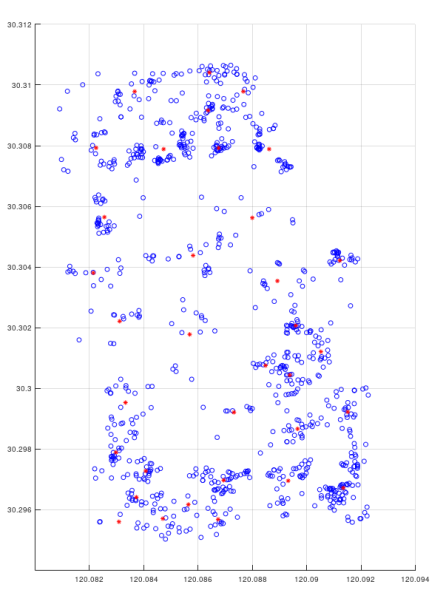
## 2.1 P区划分：聚类算法

一般而言，校园内共享单车可停放位置在空间上不可数并且会呈现出聚集的状态，例如教学楼、宿舍区等区域。这些区域往往是人口流通量较大的区域，也即共享单车需求量较大的区域。由此可见，最佳停车区域P区的规划与共享单车的空间分布情况有紧密联系[1]。因此，我们将借助哈罗单车的单车分布数据结合实地考察数据进行聚类分析，获得聚类中心位置作为P区域的中心，再选取合适的边长获得P区。

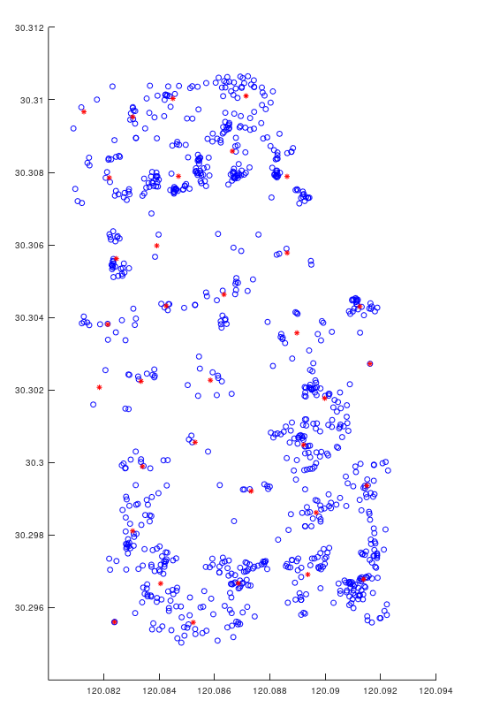
关于数据采集，我们借助紫金港校区共享单车市场合占比90%的哈啰单车和美团单车APP上的单车实况分布数据，对不同时间段、不同运营商的分布数据加和后得到紫金港校区的共享单车停放点集图（见下图）。因此，P区的选定有效地利用了之前单车停放的数据信息，增强了模型的说服力，为单车运营期进行管理整改提供参考。关于聚类中心标定，我们选择了较为典型的3种聚类算法。



1. Means聚类需要事先给定聚类簇数k和初始聚类中心{C\_1,..,.C\_k}，在一次循环中，遍历每个点，将其加入距离最近的类中，遍历结束后所有点被划分成k个类，接着计算每个类的均值向量作为下一次循环的初始聚类中心，循环直到均值向量保持不变为止。目的是为了得到最小化平方误差——这是对于每个类中的点与该类的均值向量的距离求和，再对类求和，这个量刻画了簇内样本围绕簇均值向量的紧密程度。它的优点是对噪声数据不敏感，缺点是对初始值敏感，且需要事先给定簇数k。针对上述缺点，我们采用ramdom库随机选取初始中心点，多次重复运行对比相同值保留，不同值参考实际单车停放需求；我们根据肘部法则确定簇数k，x轴是不同的k值，y轴是成本函数，即各个类畸变程度之和，每个类的畸变程度是该类重心与其内部成员位置距离的平方和。多次试验发现，实际效果最优的聚类簇数K与肘部法则计算得到的值相互印证。以下是改进后的效果图（见下图），可以看出聚类中心的标定与实际需求相符。



AGNES层次聚类需要实现给定簇数k，最初将每个数据点都自成一类，通过不断合并距离最近的两个聚类簇，直到合并到k个类为止。它的优点是对噪声数据不敏感，且不依赖于初始值的选取；缺点是需要事先给定簇数k，且对类别较少的数据集分类较慢，运行时间较长。针对上述缺点，簇数k的确定同样可以采用肘部法则；对于我们的数据集，程序运行了6分钟后可以得到结果，对于更庞大的数据集，可以通过设置阈值k，使得当聚类之间的距离大于k时就停止凝聚，这个k可以根据实际情况给定，如人们愿意行走寻找车辆的最大距离，或者是人们选择走路而不愿骑车的最小距离。以下是改进后的效果图（见图3），可以看出聚类中心的标定与实际需求相符。

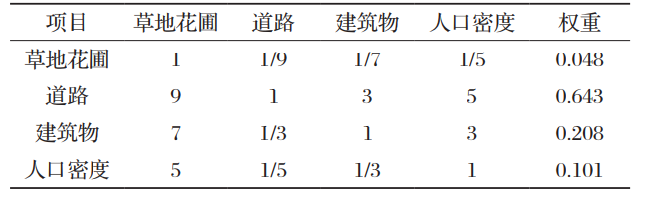


DBSCAN密度聚类需要实现给定邻域半径ε和最小包含的样本值MinPts，找到所有核心对象，核心对象就是所在ε-邻域内包含数据点个数大于MinPts的数据点，然后遍历每个核心对象，将其密度可达的点都归于这个类中，密度可达就是如果把在一个点的ε-邻域的数据点叫做它的朋友，那么密度可达的点就是核心对象的朋友的朋友的朋友，有几个核心对象最后就得到几个类。它的优点是可以完成各种形状的聚类分析；缺点是对噪声数据敏感，需要实现给定ε和MinPts，不适合数据集中密度差异较大的情况。将该算法应用到单车点集图时发现，无法找到合适的ε和MinPts，得到的结果簇数要么过少要么过多。这是因为，校内的单车停放部分过于分散，部分又过于密集，与期望的均匀分散不符。ε和MinPts取值偏小时会造成核心点过少而离散点过多，偏大时会导致核心点过多。因此，该算法不适用于本项目。

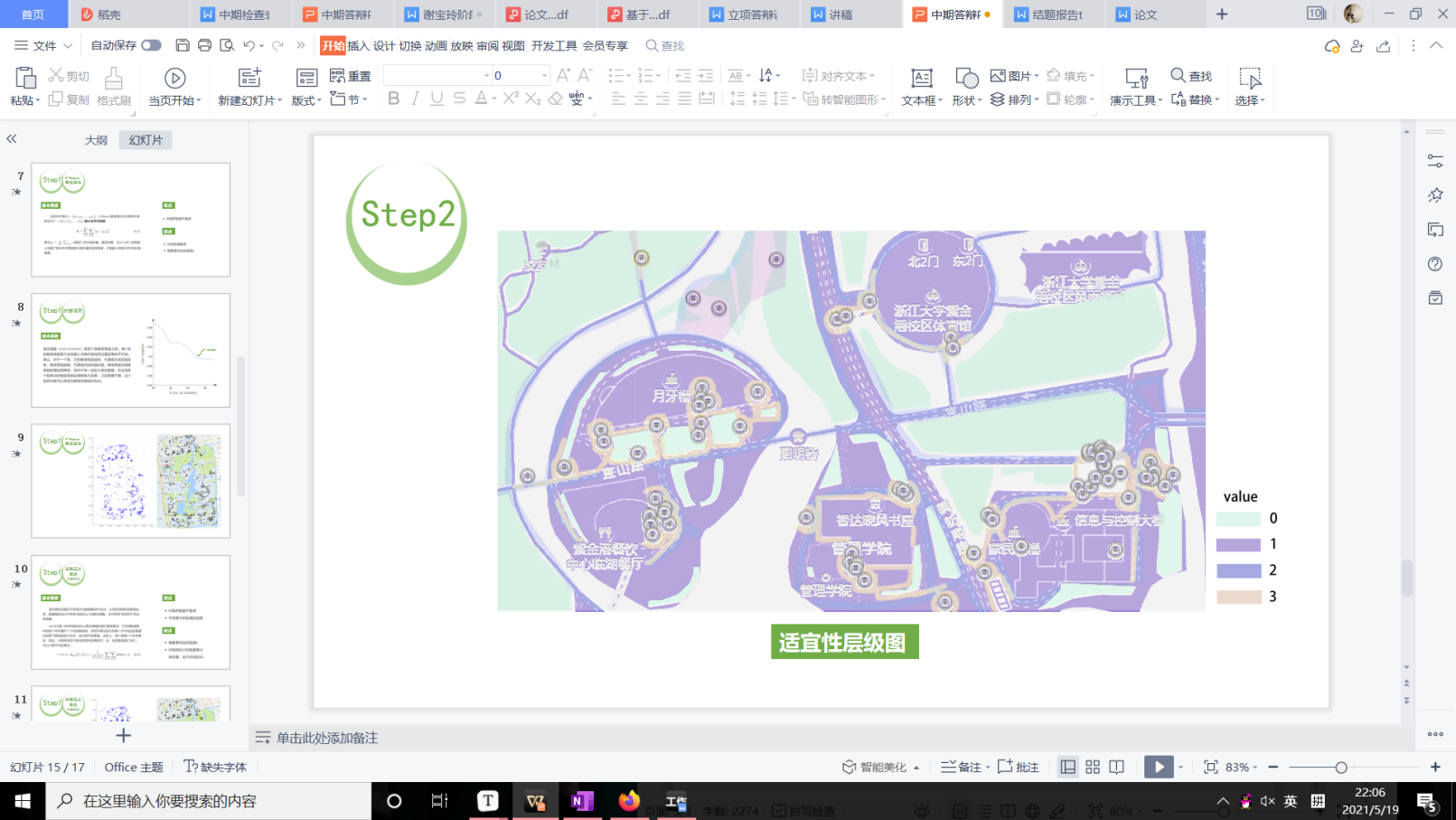
## 2.2 M区、F区划分：层次分析法

层次分析法是将决策问题各个因素按照不同属性自上而下地分解成若干层次结构，构造成对比较阵，然后用求解判断矩阵特征向量的办法，求得每一层次的各元素对上一层次某元素的优先权重，并做一致性检验，从而得到最优方案。适合于具有分层交错评价指标的目标系统，而且目标值又难于定量描述的决策问题。对于M、F区域的划分问题，我们采用的便是这一思路[2]。

首先，我们确定了影响单车停放的4个主要因素（草地花圃、道路、建筑物、人口密度），再对其进行两两相互打分，得出重要性对比程度，归一化后得到判断矩阵（见下图）。通过一致性检验后，得到各个影响因素的权重值即影响大小。



通过空间信息处理软件，分别对四个影响因素进行处理得到该因素分级图。最后按照以上权重对不同影响因素的分级图进行叠加，得到共享单车停放适宜性等级图（见下图），value值越大越适宜停放。在得到的适宜性层级图之后，我们需要对研究范围进行实地考察（如紫金港校区），来评估寻找策略的合理性。



至此，我们得到了层次型区域划分模型，它既可以在运营前为选址提供参考，又可以为运营后借助单车停放的历史数据为整改提供参考。

# 计算机模拟

在确定了P/M/F分区后，为了验证阶梯型定价决策在单车管理中的优越性，我们通过计算机模拟的手段模拟人们在阶梯型收费下的行为决策，并且观察人们决策与停车分区变化对于单车流动的影响，从企业利益、单车使用率和单车停放混乱程度三方面对于模型进行有效评估。

## 3.1 逻辑基础

对于在阶梯型定价下的单车使用客户，我们定义了以下三个基本逻辑，来模拟人们的选择。

①当目的地与起点距离处在合适范围内，且周围有单车存在时人们会选择骑车出行

②当目的地与最近的p区较近时，会选择将车停到该最近的p区点，以减少支出。

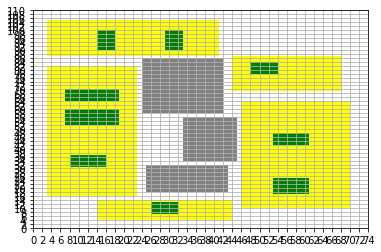
③目的地为f区点时，还会考虑周围是否有m区停车点，以便不支付过高车费。

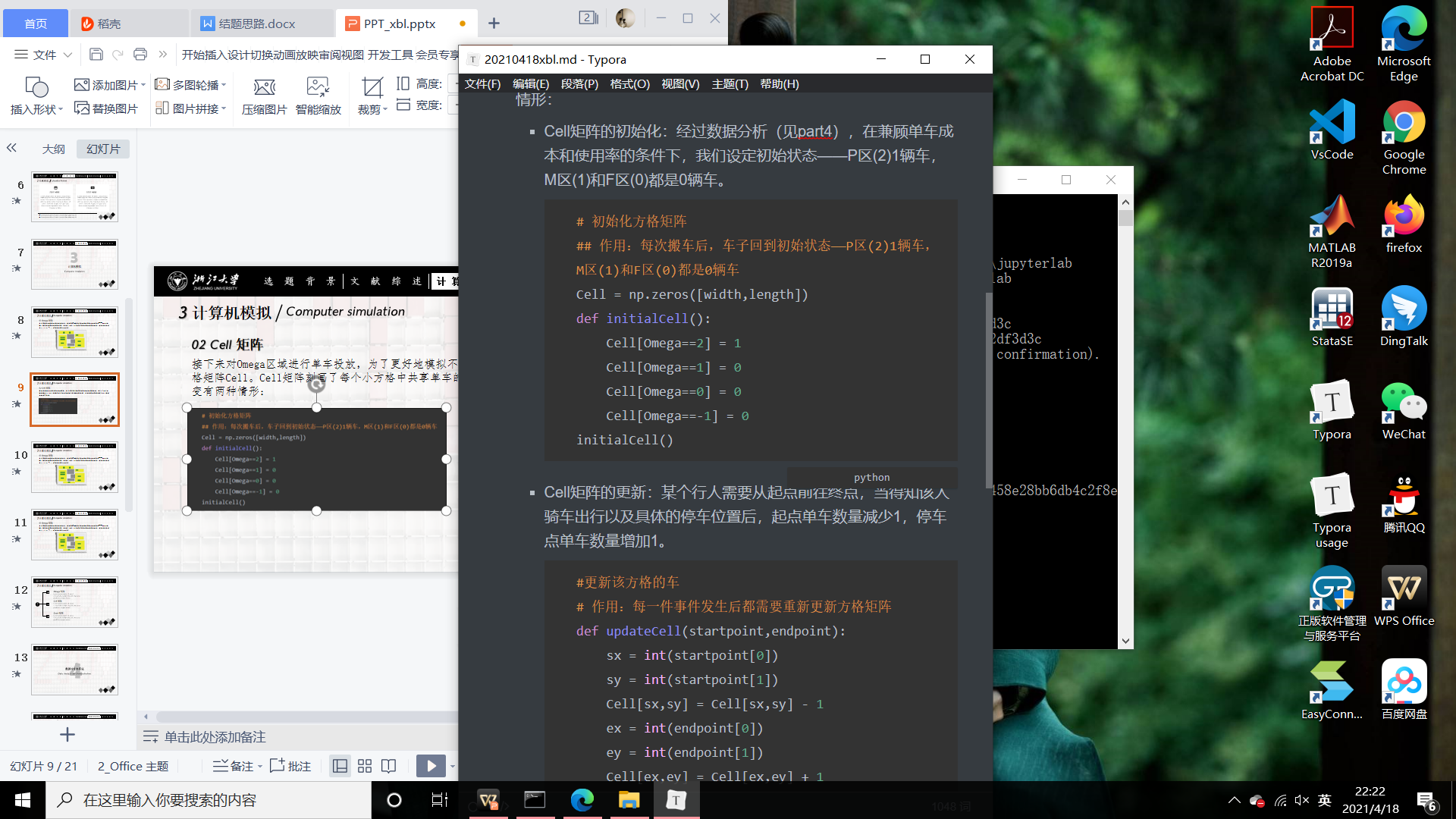
## 3.2 模拟实现过程

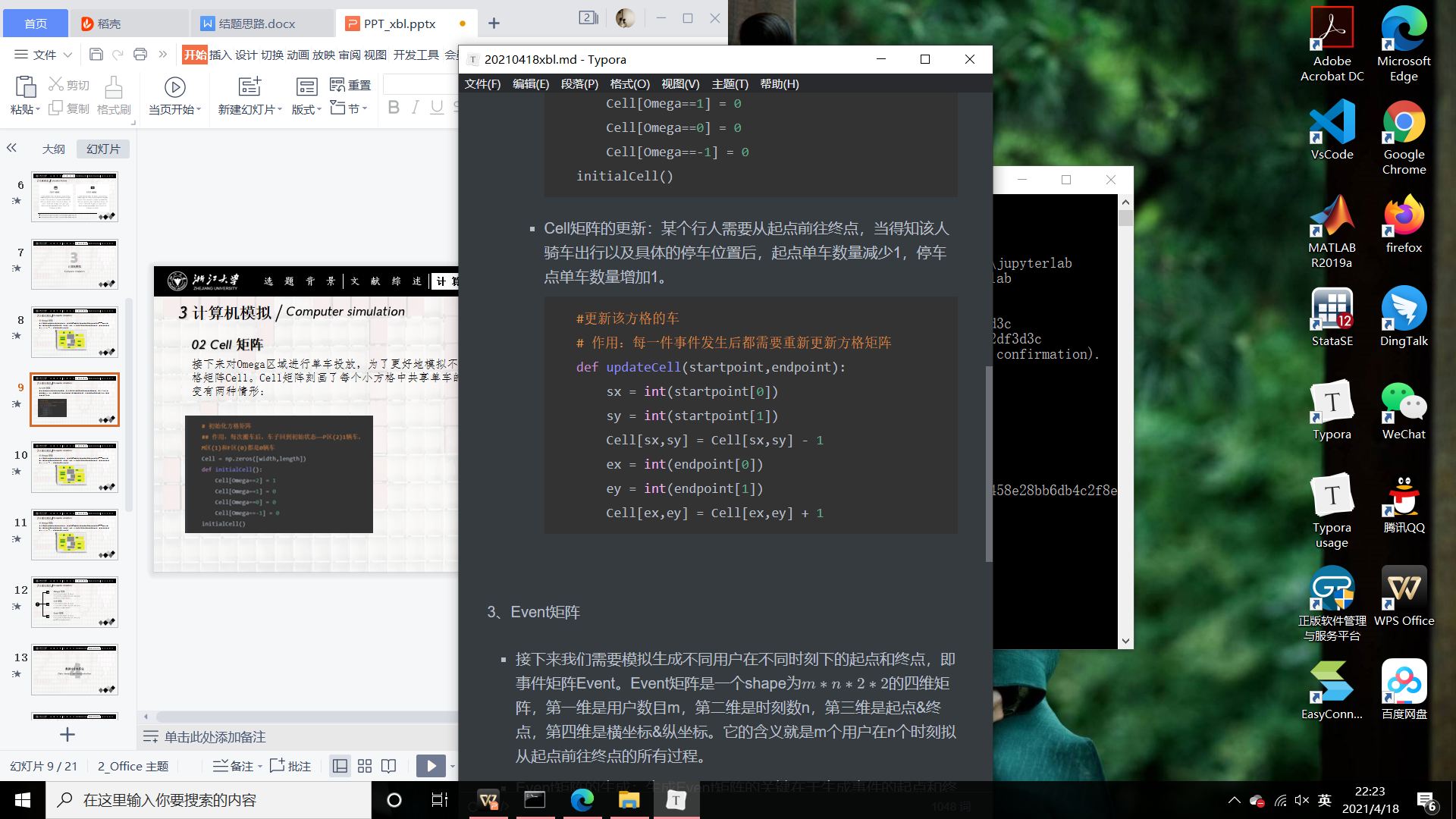
### 3.2.1 我们主要通过三个矩阵进行数据的计算与存储：Omega、Cell与Event

①Omega：用作存储地区为P/M/F区或是不可停放区的矩阵，根据前期聚类分析的结果，对P区、M区、F区和不可停放区域G区进行存储，对应元素分别为2、1、0、-1 。

可视化结果如下所示：

****

②Cell：用作存储不同范围内单车的数量的矩阵。Cell矩阵刻画了每个小方格中共享单车的数量，并且在整个模拟过程中Cell的改变有两种情形分别是：当有人在该区域用车或者停车时，改变其数值；当过一段时间会有工作人员搬运单车，改变Cell矩阵数值。

③Event：用作存储并随机不同时刻下不同人的出发与到达事件的矩阵。Event矩阵是一个shape为m\*n\*2\*2的四维矩阵，第一维是用户数目m，第二维是时刻数n，第三维是起点&终点，第四维是横坐标&纵坐标。Event矩阵的生成方式分为两步：第一步采用轮盘赌的方式确定行人所属P区域。第二步：产生以P区中心为期望、以 为协方差的二维正态随机投点，其中σ1，σ2 与P区的长、宽相关

### 3.2.2 通过之前1中所提及的逻辑基础进行模拟判断，调整参数，得到相应的各类数据。

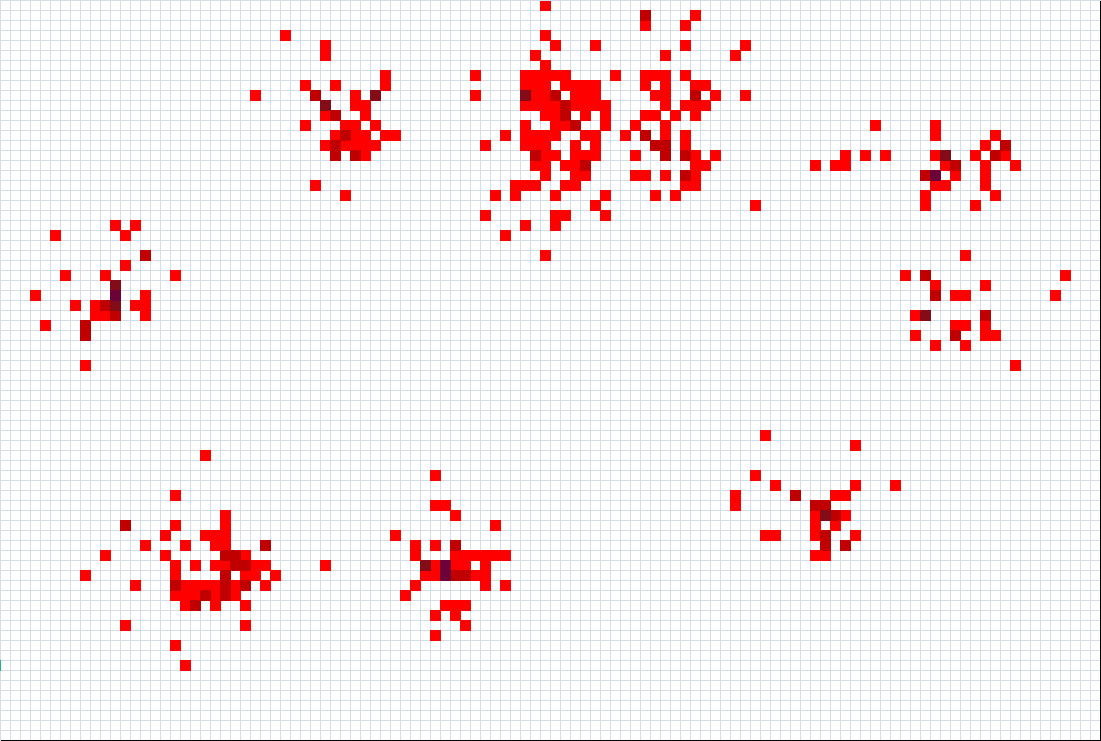
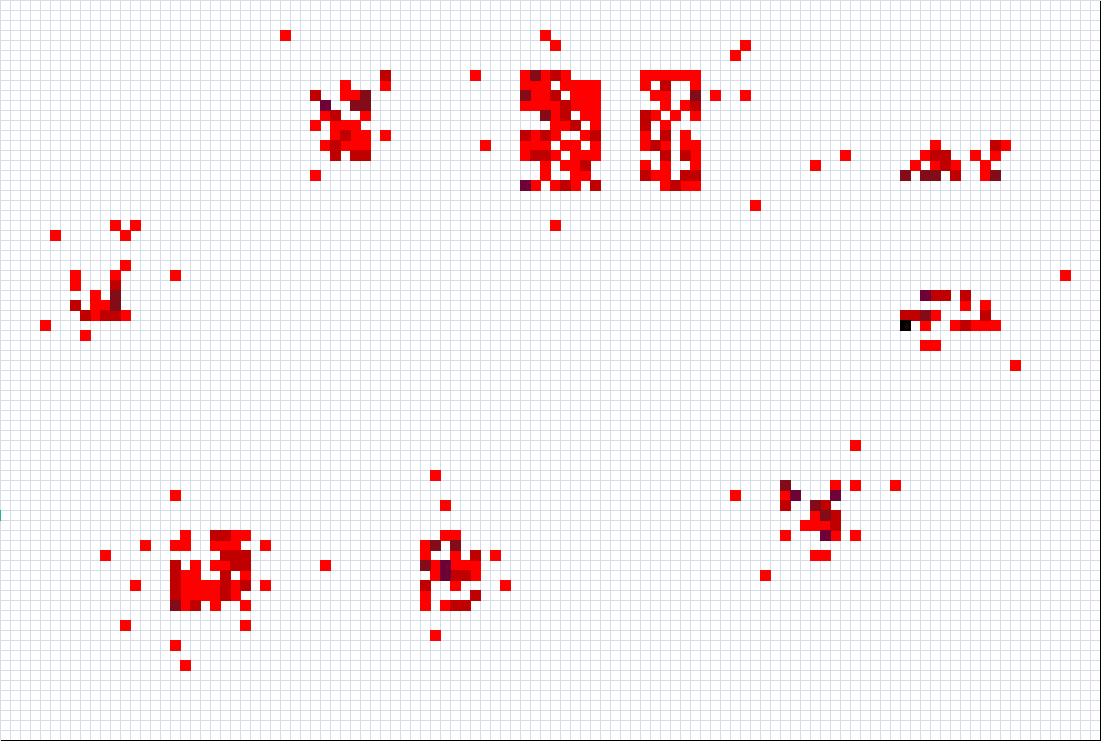
①骑行判定参数：也就是骑行距离、骑行价格、终点所属区域等因素对人是否骑车的决策参数，如图所示，其中a<Cp<C<Cm<Cf<Df<Dm<D<Dp<b

②停车点选择参数：这是阶梯型中特有的参数。它可以用来判断行人是否会因为不同区域的骑行收费不同而选择更改停车位置。以阶梯型的Cp、limitP为例：当P区骑行价格上升时，人们在相同距离下选择骑车的可能性会降低，即愿意选择骑车最小距离Cp增大；而当P区和M区的骑行价格差变大时，即使终点在M区，人们更愿意将车停在P区，即P区阈值limitP增大。再根据实际情况，给出不同价格下人们愿意行走的距离，用二次多项式拟合两者之间的关系，得到Cp、limitP。

### 3.2.3 计算机模拟部分流程图

|  |
| --- |
| 相关矩阵与系数准备：  1、Omega矩阵：生成一个m\*n的0矩阵，根据地图对于P/M/F/G区分别赋值2/1/0/-1。  2、Event矩阵：一个shape为m\*n\*2\*2的四维矩阵，第一维是用户数目m，第二维是时刻数n，第三维是起点&终点，第四维是横坐标&纵坐标。对于每个起点与终点进行轮盘赌随机到一个p区，再以该p区为中心∑为方差正态分布随机投点，将这2mn个随机点赋给Event。  3、Cell矩阵：存储不同范围内单车的数量，进行初始化，对于不同区域初始投入不同车辆数 |
| 代码流程：  1：for each i∈[0，m],j∈[0，n]，distance[i,j]=np.linalg.norm(Event[i,j,1]-Event[i,j,0])。将初始点与终点的距离计入矩阵distance。  2：if distance[i,j]∉[a,b],意味着出行太远或太近，不适合骑车，不会选择骑车。distance[i,j]清零。  3：vector1,flag = judgeCell( Event[i,j,0])，jungleCell函数判断了初始点附近是否有车，flag=0时无车，此时无法骑车出行。distance[i,j]清零。  4：if：Event[i,j,1]的Omega值为2，意味终点在P区，那么直接停在目的地，distance[i,j]不变。  5：elif：Event[i,j,1]的Omega值为1，意味终点在M区，用NearestPark函数寻找最近P区点，计算P区与目的地距离Kp，与参数Iimitp比较，较小则将停车点改为P区点。更新distance[i,j]的值。  6：elif：Event[i,j,1]的Omega值为0，意味终点在F区：用NearestPark函数寻找最近P/M区点，计算P/M区与目的地距离Kp/Km，与参数Iimitp、limitm比较，优先考虑P区，最后更新distance[i,j]的值。  7：distance[i,j]≠0时，即车被使用时，用函数updateCell对Cell矩阵实时更新，更新车辆流动状态。  8：i每增大x次，即过了x时刻，用函数initialCell再次初始化，模拟搬运工作。  9: 当走完Event矩阵，即m个人经过了n个时刻的用车情况，统计观察相应的Cell、distance等矩阵，进行数据分析。 |

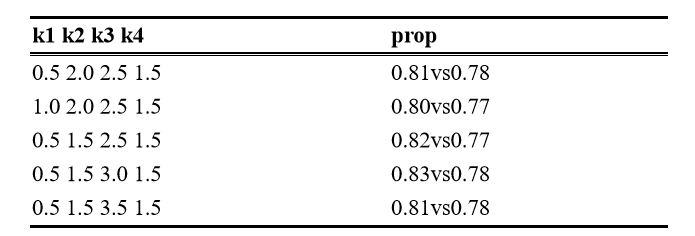
# 数据分析部分

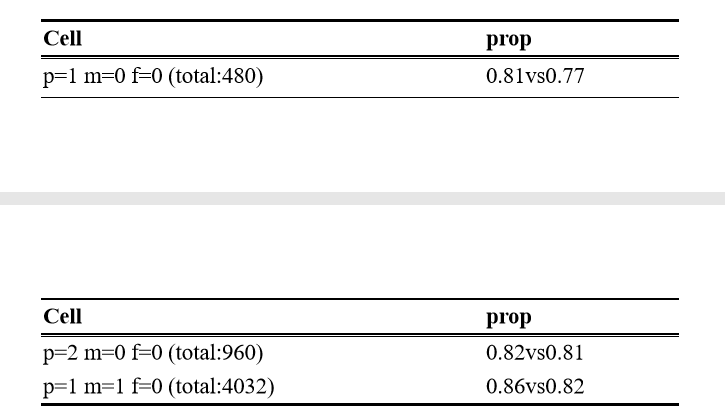
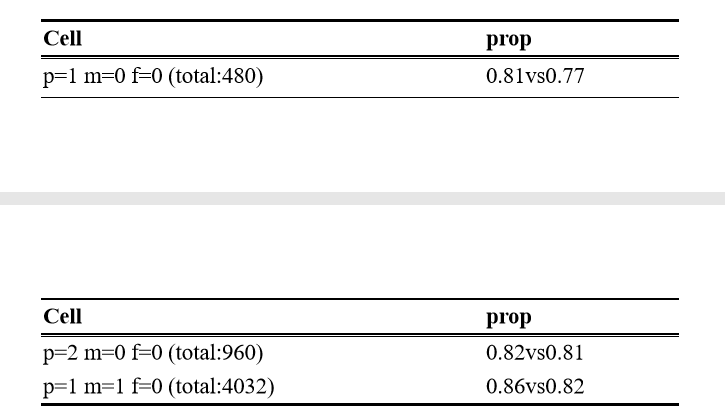
为了验证阶梯型收费的优势，我们需要从单车用户、单车企业以及社会三方面进行分析。因此我们对“单车散乱程度”、“企业利润”、“单车利用率”三种数据进行了分析，都能够发现阶梯型定价相比于统一型有着较好的模拟结果。

单车散乱程度：由于单车越散乱，人们越难以借到车辆，而且对于社会管理单车停放问题的难度会与散乱程度正相关。所以，这一指标与用户体验以及社会管理方面都有密切关系。我们每过一个时刻对Cell矩阵的数据进行记录，得到20个时刻不同的单车分布数。

据图，最终我们可以发现，随着时间推移，在统一型定价下单车散乱的程度更高速度更快，而阶梯型下，单车更趋向于分布在P区周围。如图是第20时刻两种定价的单车分布图，红色越深代表区域内单车数量越多，越聚集。（左图为统一型，右图为阶梯型）

对于企业利润，不同的定价方式主要改变了企业的“骑行收费”、“广告收费”与“管理搬运支出”三方面。在阶梯型定价下，由于p区价格低廉，会有更多的人选择骑行，这在一定程度上增加了骑行收费并且使广告效应更强。而统一型会导致单车更加分散、散乱，一定程度上会增加搬运费用的支出。经过模拟，我们也能发现事实上，阶梯型定价也做到了使企业骑行收费更高、广告收费更高且管理搬运费显著降低的效果。这证明了对于企业而言，阶梯型定价收费的优势所在。同时，通过整数规划或是枚举法的方式，在得到具体企业利润计算公式后，我们可以相应解出最佳的P/M/F分区定价值，以便于企业获得最高利润。

单车利用率prop=（出行使用单车人数/适宜范围内出行人数）。单车利用率高意味着会有更多的人选择骑车，而企业所投放的车辆能够被合适的运营起来。这可以使企业在投放相同数量车辆，投入相近资本的同时获取更大的利益，而用户也能更多地使用共享单车，对于共享单车的满意度更高。而经过验证，我们也能发现，在合适的参数下，阶梯型定价模式的单车利用率都高于统一型。（如下图所示）同时在不同定价下，不同的prop值使得我们可以通过整数规划与枚举法得到最佳的PMF区域定价值以及最佳的单车投放量。



# 结论

本文主要讨论了共享单车管理优化的设计问题，给出了一个通用的方法得到层次型区域划分模型，用计算机模拟得到最优阶梯型定价并对模型进行评估验证，未来的研究至少在以下几个方面是有益的。

首先，该管理优化方法具有极好的可重复性，对于任意一个新的单车运营区域，只需要利用该区域地图，并收集区域内的单车位置数据，我们便可以通过聚类算法+层次分析+地理信息处理的方式，得到相应范围内单车停放的P/M/F分区。之后通过相似的计算机模拟手段，可以直接对于该区域的单车管理运营进行观察和模拟。而我们也对于玉泉校区区域进行了简单的尝试，可以发现，这样的P/M/F分区手段与阶梯型收费方法是可以复刻的。其次，可延拓性也是我们方法的主要优点。在拥有企业盈利手段、社会管理单车方法等完整的信息与逻辑链后，我们可以在模拟模型中加入更多的影响参数，考虑更多的现实影响因素，使计算机模拟一步步逼近真实单车运营环境。此外我们也可以在此基础上模拟出更多的企业运营手段，比如单车会员带来的利润等等。此外，计算机模拟的手段对于数据的需求程度会更低，比较适用于一些缺乏数据甚至无法获得数据的情况。因此以它为企业带来模拟与运营建议，可以一定程度上降低单车管理所需的信息成本，降低一些试运行的人力物力财力。

当然我们的模型还具有一些有待改进的部分。目前的计算机模拟速度略有欠缺，在拥有过大的数据量时会出现运行时间较长的问题，可以继续研究如何在不影响模型效果的同时提高模拟效率。同时，我们对于共享单车公司的其他盈利手段例如现金流和广告收益方面没有过于准确的信息，在计算机模拟方面可以继续增加盈利的影响因素使得模型更加贴近实际情况。

**参考文献**

[1] 王霞,宋树华,汤军,刘远刚.基于混合聚类分析的共享单车停放点位置合理性研究[J].数字技术与应用,2019,37(07):58-61.

[2] 李林凤,李进强,耿莲.基于GIS的城市共享单车虚拟站点选址规划——以闽江学院校区为例[J].智能城市,2019,5(20):4-8.