

## 2021 第六届“数维杯”大学生 数学建模竞赛论文

### 题 目 中小城市地铁运营与建设优化设计

#### 摘 要

地铁产业是中国城市基础设施中最有前景、最有市场的产业，但绝大多数地铁线路均存在亏损，诸如呼和浩特市这一类常住人口相对较少但交通道路较为拥堵的城市问题尤为显著，因此必须提出科学的运营方案和合理的地铁建设方案来有效降低运营成本并提高运行效率。

针对问题一，根据题目中所给出的附件 2 和当前呼和浩特市目前的发车方式及地铁基本信息，通过对乘车舒适度以及乘车满载率进行研究分析，发现目前发车方案不合理，因此根据舒适度与满载率我们建立了多目标发车间隔优化模型，采用基于 NSGA-III 的多目标遗传算法求解该问题的 Pare-to 解集，得到了最优参数。最后，综合考虑相邻时段的发车间隔的稳定性，确定了各时间段的发车间隔。具体结果见 5.1.3。

针对问题二，首先用 MATLAB 将站点在二维图上按其经纬度进行绘制并提出网络模型，将这些站点模拟为网络中的节点，利用节点的介数中心性计算出核心站点，并围绕这个站点提出两种选址方案，利用三角模糊数和集对分析理论描述不同方案的各指标特征，构建地铁选址方案指标体系，利用信息熵反映的效用值来计算各指标权重，建立综合评价模型，计算结果表明了该模型的正确性和可用性，可为地铁企业在建新路线时提供决策支持。其次，我们通过对地铁客运量及运营成本在不同行车间隔下的特性，分析出地铁运营企业成本特性、收入特性，最终预测出达到盈亏平衡时所需日均客运量。具体结果见 5.2.3。

针对问题三，结合瓶颈模型分析，以单个出行者的总效益最大化为目标，建立数学分析假设模型，并分析模型满足的平衡条件，探索于单一上班时间点出行者的出行行为和上班时间出行行为的优化方法，并以此为依据得出了最优的出行模式和上班时间点设置并实现最终平峰目标。具体结果见图 5-18。

针对问题四，根据附件 1 的站点经纬度信息，我们采用 DBSCAN 算法对这些站点进行密度聚类，得到密度相连的站点的最大集合，这些站点的重要性高，为枢纽站点，因此我们提出围绕集合里的站点，修建公交线路，与地铁相辅相成，以此提高出行效率。具体结果见图 5-21。

最后，讨论了模型的优缺点以及模型的推广，提出了科学的运营方案和合理的地铁建设设计以供参考。

**关键词：**多目标遗传算法；介数中心性；三角模糊数-集对分析；信息熵；盈亏平衡；瓶颈模型；DBSCAN

## 目 录

一、问题重述.....	3
二、问题分析.....	3
2.1 问题 1 的分析.....	3
2.2 问题 2 的分析.....	3
2.3 问题 3 的分析.....	4
2.4 问题 4 的分析.....	4
三、模型假设.....	4
四、定义与符号说明.....	4
五、模型的建立与求解.....	5
5.1 问题 1 的模型建立与求解.....	6
5.1.1 多目标发车间隔优化模型的建立.....	6
5.1.2 多目标发车间隔优化模型的求解.....	8
5.1.3 结论 .....	8
5.2 问题 2 的模型建立与求解.....	9
5.2.1 基于三角模糊数一集对分析的地铁车站选址模型的建立.....	9
5.2.2 基于三角模糊数一集对分析的地铁车站选址模型的求解.....	15
5.2.3 结论 .....	18
5.3 问题 3 的模型建立与求解.....	18
5.3.1 单一上班时间点下的出行行为分析及优化模型的建立.....	18
5.3.2 单一上班时间点下的出行行为分析及优化模型的求解.....	25
5.3.3 结论 .....	26
5.4 问题 4 的模型建立与求解.....	26
5.4.1 基于 DBSCAN 算法模型的建立.....	27
5.4.2 基于 DBSCAN 算法模型的求解.....	28
5.4.3 结论 .....	28
六、模型的评价及优化.....	29
6.1 模型的优点 .....	29
6.2 模型的缺点 .....	30
6.3 模型的推广 .....	30
参考文献.....	31
附录.....	32

## 一、问题重述

地铁有很多优点，鉴于地铁建成后为城市居民带来的诸多优点，中国大陆地区许多城市都将地铁建设纳入了城市长远发展规划中，然而因其高昂的建设成本、后期运营成本及便民的收费标准，众多地铁线路均存在一定规模的亏损。因此请以呼和浩特为研究对象，展开如下的分析建模：

问题 1：附件 2 中我们模拟了各站点在 2020 年 9 月 1 日至 2020 年 9 月 14 日进出站的数据，试分析目前发车方案的合理性，并提出一个最优的车厢数量及发车间隔确定模型，同时利用更多的仿真模拟数据对优化前后的方案展开对比分析。

问题 2：呼和浩特市未来仍在考虑新增地铁线路用以缓解交通压力和碳排放。附件 1 中我们给出了呼和浩特的站点经纬度，您能否提出一个合理的站点选址方案以确保更多的居民会选择地铁出行方式，并预测出每天总计乘坐地铁人数达到多少才能够实现地铁运营的盈利目标。

问题 3：考虑到疫情依旧持续较长的时间，您能否提出一种错峰的出行方案，这个方案通过对不同站点附近主要上学或上班等人员的合理上下班时间的制定最大限度的减少高峰时期出行人数，并实现最终平峰目标

问题 4：在充分考虑呼和浩特市城市快速路的基础之上，你能否提出一个地铁和公交互补的若干新增公交线路，以满足更多人员高峰出行时期的出行效率。

## 二、问题分析

### 2.1 问题 1 的分析

问题 1 要求我们分析目前呼和浩特发车方案的合理性，并提出一个最优的车厢数量及发车间隔确定模型，并对优化前后的方案展开对比分析。根据题目中所给出的附件 2 和当前呼和浩特目前的发车方式及地铁基本信息，我们建立多目标行车间隔时间优化模型；采用基于 *NSGA-II* 的多目标遗传算法求解该问题的 *Pare-to* 解集，得到了最优的参数；最后，综合考虑相邻时段的发车间隔的稳定性，确定了各时间段的发车间隔。

### 2.2 问题 2 的分析

问题 2 首先提出呼和浩特市未来仍在考虑新增地铁线路用以缓解交通压力和碳排放，要求我们结合附件 1 站点的经纬度，提出一个合理的站点选址方案以确保更多的居民会选择地铁出行方式，并预测出每天总计乘坐地铁人数达到多少才能够实现地铁运营的盈利目标。我们首先将这些站点在图上绘制，提出网络模型，将这些站点模拟为网络中的节点，利用节点的介数中心性计算出最核心的一个站点，并围绕这个站点提出 2 种选址方案，利用三角模糊数和集对分析理论描述不同方案的各指标特征，构建地铁车站选址方案指标体系，利用信息熵反映的效用值来计算各指标权重，建立综合评价模型，计算结果表明了该模型的正确性和可用性，可为地铁企业在站位选址时提供决策支持。其次，我们通过对地铁客运量及运营成本在不同行车间隔下的特性分析，得出了地铁运营企业成本和收入特性，预测出了盈亏平衡时所需日均客流量。

## 2.3 问题 3 的分析

问题 3 要求我们提出一种错峰的、通过对不同站点附近主要上学或上班等人员的合理上下班时间的制定最大限度的减少高峰时期出行人数，并实现最终平峰目标的出行方案。我们结合瓶颈模型，以单个出行者的总效益最大化为目标，建立数学模型，分析模型满足的平衡条件，研究在单一上班时间点下出行者的出行行为和上班时间点的优化，并以此为依据得出了最优的出行模式和最优的上班时间点设置。

## 2.4 问题 4 的分析

问题 4 以满足更多人员高峰出行时期的出行效率为目的，要求我们提出一个地铁和公交互补的若干新增公交线路。根据附件 1 的站点经纬度信息，我们采用 DBSCAN 算法对这些站点进行密度聚类，得到密度相连的站点的最大集合，这些站点的重要性高，为枢纽站点，因此我们提出围绕集合里的站点，修建公交线路，与地铁相辅相成，以此提高出行效率。

## 三、模型假设

- 1.同一时间段的发车间隔固定；
- 2.每一时间段内客流均匀到达和离开车站；
- 3.车辆选型和编组固定列车定员人数固定；
- 4.列车全程匀速运行，未发生安全事故；
- 5.同线路上各列一次运行的运营成本相同；
- 6.假定各周的乘车人数与附件 2 中的乘车人数不存在显著性差异；
- 7.地铁运行秩序良好，路上无阻塞情况也不会出现突然坏掉或燃料不足等情况；
- 8.进行客流量分析时，剔除观光和季节客流的影响；
- 9.假设所有企业都采用非弹性工作制，单位时间的工作效益是关于时间的给定函数，且工作者每天的工作时长固定，上班与下班是相互独立的；

## 四、定义与符号说明

表 4-1 符号定义说明表

符号定义	符号说明
$K = \{k   k = 1, 2, \dots, N\}$	城市轨道交通线路断面集合
$Q$	列车定员
$q$	线路的最大断面客流量
$h_1, h_2$	模拟期间的最小及最大发车间隔

$I$	模拟期间线路发车数
$m_1, m_2$	模拟期间的最小及最大车辆满载率
$d$	车内乘客站立人员密度
$U_w(t)$	在时刻 $t$ 从家出发的出行者的工作活动
$U(t)$	在时刻 $t$ 从家出发的出行者的总效用

\*其他未标明符号在文中注明

## 五、模型的建立与求解

### 1. 客流分析

#### (1) 时间分布特性分析

呼和浩特地铁客流在一天内随人们的生活习惯和工作需要而变化。如图 5-1 所示，工作日内某地铁 1 号线和 2 号线的进出站客流量分布图。从图中可知，该地铁客流在一天内形成 2 个客流高峰，在早高峰和晚高峰时，同时存在较高的进站和出站客流，其他工作日的客流量时间分布规律与此类似。

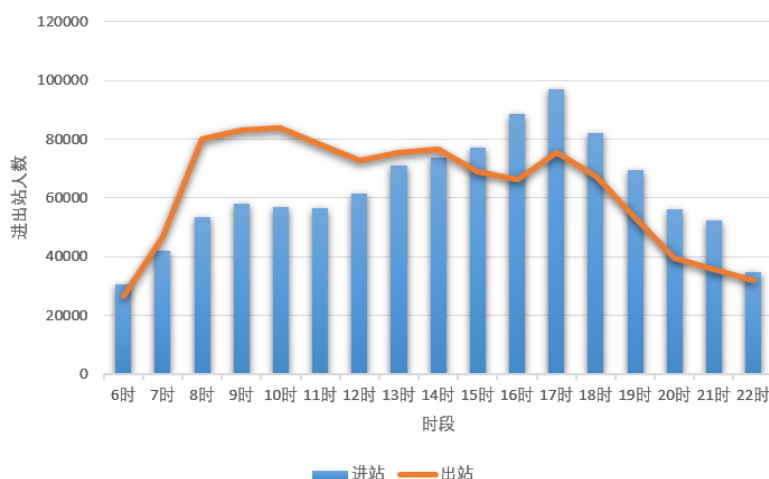


图 5-1 工作日进出站客流量图

由图 5-1 可见，晚高峰客流量较大，进站的早高峰客流峰度大于晚高峰，出站的晚高峰客流峰度大于早高峰，但早高峰的跨度大于晚高峰，说明晚高峰客流比较集中，早高峰峰值相对于晚高峰客流变化较缓和。由此可知，该地铁线路工作日客流时间分布具有通勤、通学特性。

图 5-2 为周末客流分布图，可知周六与周日的客流时间分布规律有相似之处，周末客流与工作日客流时间分布存在明显的差异，周末最高峰比工作日客流最高峰高，周末平峰时段的客流量也大于工作日的客流量，说明周末客流时间分布较分散。

## 5.1 问题 1 的模型建立与求解

### 5.1.1 多目标发车间隔优化模型的建立

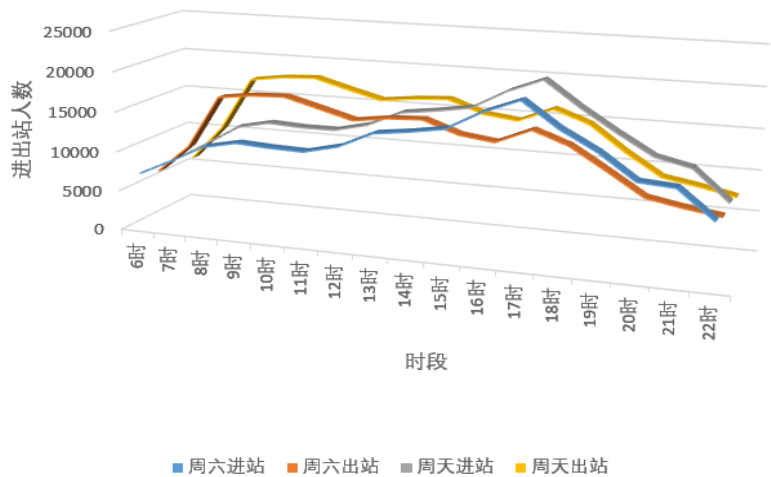


图 5-2 周末进出站客流对比图

#### 1. 模型分析及建立

##### (1) 乘客舒适度

地铁车厢拥挤程度直接影响乘客的乘车舒适度<sup>[1]</sup>，用车辆内乘客站立人员密度作为衡量舒适度的标准，也是定员标准<sup>[2]</sup>。《城市轨道交通工程项目建设标准》建议的车辆内乘客站立人员密度评价标准见表 5-1。

表 5-1 车内乘客站立人员密度评价标准表

站席密度/(人·m <sup>2</sup> )	乘客拥挤情况	评价标准
3	乘客可以自由流动，十分宽松	舒适
4	有较大宽松度，乘客可以看书报	良好
5	有一定宽松度，部分乘客可以看书报	良好
6	感到不宽松、不拥挤、稍可活动，是舒适度的临界状态	临界状态
7	感到有些拥挤，站席范围有些突破	有些拥挤
8	身体有接触，需错位排列，并突破站席范围，感到比较拥挤	比较拥挤
9	平均每人占有的空间非常拥挤，需突破站席范围，挤入座区，此情况偶有可能出现（车辆制造强度必须满足）	非常拥挤 (超员标准)

10	乘客突破站席范围，挤入座区，极为拥挤，难以忍受，影响上、下车行为和总时间，属极端情况	难以忍受
----	--	------

车内乘客站立人员密度 $d$ 可表示为

$$d = \left( \frac{q}{I} - S \right) / A \quad (5.1)$$

式中： $S$ 为列车坐席数量； $A$ 为列车立席区分配面积。

对于乘客而言，当舒适度空间恰好接触但不挤压乘客身体，乘客可小幅活动，此时乘客的舒适度良好，随着乘客数量的增加，舒适度逐渐降低，根据表 1-1 的标准，可构建乘客舒适度的分段效用函数如下：

$$U(d) = \begin{cases} 9, & d \leq 6 \\ 7, & 6 < d \leq 7 \\ 5, & 7 < d \leq 8 \\ 1, & 8 < d \end{cases} \quad (5.2)$$

## (2) 满载程度衡量

地铁运营管理者为了降低经营成本，往往会增大发车间隔，以增加列车的满载程度。满载率一般可用车内实际的乘客数与车辆定员的比值来表示，可构建以下效用函数对满载率进行衡量。

$$z_1 = 1 / \{1 + \alpha \exp[-\beta q / (QI)]\} \quad (5.3)$$

式中： $\alpha$ 为成本参数； $\beta$ 为强度系数。本文取 $\alpha=40$ ， $\beta=10$ 。

依据上述分析，构建地铁列车行车间隔优化模型如下：

$$\begin{aligned} \max z_1 &= U(d) \\ \max z_2 &= 1 / \{1 + \alpha \exp[-\beta q / (QI)]\} \\ h_1 &\leq \frac{T}{I} \leq h_2 \end{aligned} \quad (5.4)$$

s.t.

$$m_1 \leq \frac{q}{QI} \leq m_2, I \text{ 为正整数} \quad (5.5)$$

其中：式（5.4）为发车间隔限制约束；式（5.5）为车辆满载率约束。

## 2.求解算法

由于本模型由 2 个相互冲突的目标函数构成，因而需要寻求该模型的 Pare-to 最优集。遗传算法作为一种启发式的搜索算法，在多目标优化问题中得到了广泛的应用。其中，NSGA-II 采用简洁明晰的非优超排序和排挤机制，使算法具有逼近 Pare-to 最优前沿的能力，并采用排挤机制保证得到的 Pare-to 最优解具有良好的散布，表现出较好的综合性能<sup>[3]</sup>。本文基于 NSGA-II 设计模型的求解算法。

### (1) 染色体的构造

采用二进制编码方式表示模拟时间段内的发车数、列车坐席数量 $S$ 、列车厢数 $M$ 、列车立席区分配面积 $A$ ，发车数满足约束条件（5.4）（5.5）。

### (2) 交叉算子

按交叉概率 $p_c$ 从父代选择一些染色体，两两分组，并对每组染色体进行如下操作：随机产生两个交换位，将两条染色体中的基因进行交换，从而得到两条新的染色体。

### (3) 变异操作

对  $\text{popsize}$  个染色体以变异概率  $p_m$  进行变异：对被选择变异染色体的基因，随机产生一个变异位置，并重新产生其基因，从而得到一条新的染色体。

### (4) 拥挤度选择算子

快速非支配排序：设  $n_i$  为种群中支配个体  $i$  的个体数， $S_i$  为种群中被个体  $i$  支配的个体集合。快速非劣分层的具体步骤为<sup>[4]</sup>。

**步骤 1** 找出种群中所有  $n_i = 0$  的个体，并保存在当前集合  $F_1$  中。

**步骤 2** 对于当前集合  $F_1$  中每个个体  $i$ ，遍历  $S_i$  中每个个体  $l$ ，执行  $n_l = n_l + 1$ ，如果  $n_l = 0$ ，则将  $l$  保存在集合  $H$  中。

**步骤 3** 记  $F_1$  中得到的个体为第一个非支配层的个体，并以  $H$  为当前集合，重复以上操作，直到整个种群被分层。

确定拥挤度计算步骤为

**步骤 1** 对同层的个体初始化距离，令每个个体  $i$  的拥挤度初始值  $L[i]_d = 0$ 。

**步骤 2** 对不同的目标函数，重复步骤 3-1~3-3 操作。

**步骤 3-1** 对同层的个体按目标  $m$  的函数值进行优劣排序；

**步骤 3-2** 使得排序边界上的个体具有选择优势，给定一个大数  $L[0]_d = L[l]_d = \infty$ ，以确保进入下一代；

**步骤 3-3** 对边界点以外的个体，求拥挤度： $L[i]_d = L[i]_d + \frac{L[i+1]_m - L[i-1]_m}{f_m^{\max} - f_m^{\min}}$ ，其中： $L[i]_m$  为集合  $L$  中的第  $i$  个个体的第  $m$  个目标函数的值； $f_m^{\max}$ ， $f_m^{\min}$  为该目标函数的最大值及最小值。

经过快速非支配排序和拥挤度计算，群体中的每个个体  $i$  都得到两个属性：非支配序  $i_{rank}$  和拥挤度  $i_d$ ，进而可采用轮盘制选择算子进行选择，具体过程为：对于个体  $i, j$ ，当  $i_{rank} < j_{rank}$  或  $i_{rank} = j_{rank}$  且  $i_d > j_d$  时，选择  $i$ ，如果 2 个个体在同一级，取周围较不拥挤的个体。

## 5.1.2 多目标发车间隔优化模型的求解

根据当前呼和浩特地铁目前发车方案，计算以目前发车方案为条件的地铁满载率，结果见附件 1，结果显示在目前发车方案下，地铁满载率在高峰期反而很低，明显不合理，公共资源没有得到合理的运用。

对当前发车模式进行发车间隔优化，选取呼和浩特 1 号线和 2 号线为例，结合附件 2，设定参数为：优化后的车厢数  $M$ ，定员  $Q$  为  $400 \times M$  人，列车坐席数量  $S$ ，列车立席区分配面积  $A$ ，最大发车间隔为  $16 \text{ min}$ ，最小发车间隔为  $2 \text{ min}$ ，最大满载率为 1.1，最小满载率为 0.7。利用本文设计的模型及算法结合附件 2 的模拟客流对发车间隔进行优化。

经过算法优化后得到最优参数为： $M = 4$ ， $S = 240$ ， $A = 200 \text{ m}^2$ ，发车时间间隔见表 5-2。

## 5.1.3 结论

表 5-2 发车时间间隔

时间段	工作日发车间隔/min	时间段	工作日发车间隔/min
6:00-7:00	15	6:00-7:00	15
7:00-8:00	10	7:00-8:00	10



8:00-9:00	6	8:00-11:00	6
9:00-17:00	10	11:00-16:00	10
17:00-19:00	6	16:00-20:00	6
19:00-20:00	10	20:00 之后	15
20:00 之后	15		

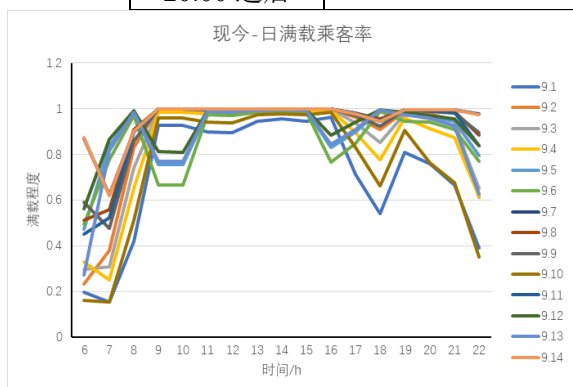


图 5-3 目前的地铁满载率图

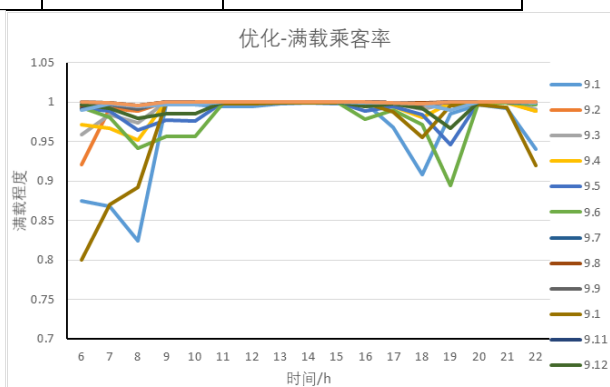


图 5-4 优化后的地铁满载率图

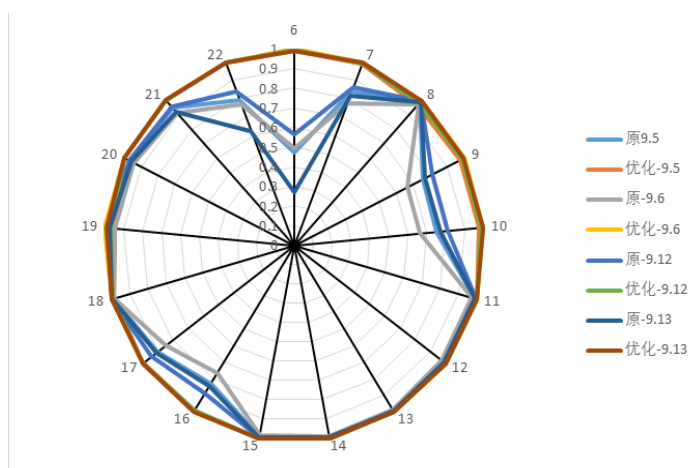


图 5-5 周末时地铁满载率对比图

综合考虑相邻时段发车间隔的稳定性，从优化结果来看，各时段满载率均在可接受范围，乘客的舒适度也会提高，说明本文设计的模型及算法具有较好的适用性。

## 5.2 问题 2 的模型建立与求解

### 5.2.1 基于三角模糊数—集对分析的地铁车站选址模型的建立

在城市地铁车站选址规划时，正确的选址方案将有助于缓解城市交通压力，减少工程投资，降低运营成本，提高旅客出行满意度。不合理的方案选择将有碍于城市发展，造成投资浪费<sup>[5]</sup>。因此，地铁车站选址方案的确定，是复杂的多属性决策问题，不同影响因素构成的指标体系具有定量与定性结合、模糊性高的特点。传统的选择方式通常采用专家评议、领导决策等主观性较强的定性方法，缺乏较全面系统的科学分析，易造成选择结果的失真<sup>[6]</sup>。

在进行地铁车站选址方案比选时，考虑各种影响因素的不确定性，集对分析理论与

三角模糊数结合,对模糊条件下定性的事物属性进行量化处理的方法,在多种评价模型中广泛应用<sup>[7]</sup>,本文在分析影响地铁车站选址方案主要因素的基础上,考虑各评价指标的确定性和不确定性,利用集对分析理论和三角模糊数理论<sup>[8]</sup>,构建影响地铁站位选址因素的指标体系.通过计算各影响因素和评价指标之间的差异度和联系度,对所有影响因素做出定量评价,利用信息熵计算指标矩阵,来确定各评价指标的权重,建立不同方案的综合评价模型。

以确保更多的居民会选择地铁出行方式,并以“盈亏平衡”为目的预测出每天总计乘坐地铁人数达到多少才能够实现地铁运营的盈利目标。

## 1.理论分析

### (1) 集对分析理论

集对分析理论的核心把系统中具有相互关联属性的两个集合构成集对,引入同、异、反联系度及其数学表达式来统一描述系统中的各种不确定性,从而将不确定性的辩证分析转化为数学运算<sup>[9]</sup>。

设集合  $A$  和集合  $B$  组成集对  $H(A, B)$ , 其中集合  $A = (a_1, a_2, \dots, a_n)$ , 集合  $B = (b_1, b_2, \dots, b_n)$ , 共得到  $n$  个特性, 其中  $S (S \leq n)$  个特性为集合  $A$ 、 $B$  所共有,  $F (F \leq n)$  个特性在关系上不确定,  $P (P \leq n)$  个特性相对立。则集合  $A$  和集合  $B$  的联系度  $v$  为

$$v = a + \sum_{d \geq 3}^n b_{d-2} i_{d-2} + cj \quad (5.6)$$

式中:  $a$  为集对  $H(A, B)$  同一度;  $d$  为评价指标数据划分等级;  $b_{d-2}$  为集对  $H(A, B)$  差异度;  $i_{d-2}$  为差异度系数, 取值范围为  $[-1, 1]$ , 当  $\sum_{d \geq 3}^n i_{d-2} = 1$  时, 差异度转化为同一度, 当  $\sum_{d \geq 3}^n i_{d-2} = -1$  时, 差异度转化为对立度, 当  $\sum_{d \geq 3}^n i_{d-2} \in (-1, 1)$  时, 差异度中同一与对立各占一定比例;  $c$  为集对  $H(A, B)$  的对立度;  $j$  为对立度系数, 取值为  $-1$ , 且  $a + \sum_{d \geq 3}^n b_{d-2} + c = 1$ 。

在式 (5.6) 中, 若  $d=3$  时,  $v = a + b_1 i_1 + cj$ , 为常用的 3 元联系度。为了使地铁站位设置方案选择结果更合理和符合实际, 本文将影响地铁站位设置的指标数据划分为 5 个等级<sup>[10]</sup>。当  $n=5$  时, 将式 (5.6) 的表达式修正为

$$v = a + b_1 i_1 + b_2 i_2 + b_3 i_3 + cj \quad (5.7)$$

式中:  $b_1$ 、 $b_2$  为差异度分量;  $i_1$ 、 $i_2$ 、 $i_3$  为差异度不确定分量系数<sup>[11]</sup>。

### (2) 三角模糊数理论

在对地铁选址方案比选时, 利用集对分析方法虽能表述地铁站位选址方案中的确定性特征, 但在处理不确定性因素时存在较大难度。因为联系度的差异度系数在表示相邻标准界别间的模糊性时求解较为困难, 可引入三角模糊数来描述差异度系数的模糊特性<sup>[12]</sup>。

定义 1: 设模糊数  $A$  属于实数域  $R$ , 可定义隶属度函数  $\mu_A(x): R \rightarrow [0, 1] (x \in R)$ 。若  $\mu_A(x)$  可表示为

$$\mu_A(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b \leq x < c \\ 0, & x \geq c \end{cases} \quad (5.8)$$

则称  $A$  为三角模糊数, 记作  $A = (a, b, c), (a \leq b \leq c)$ 。其分布如图 5-6 所示。

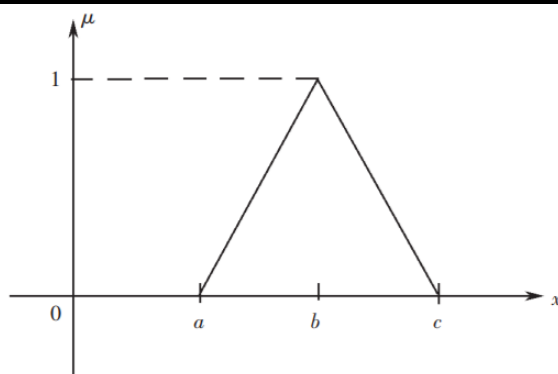


图 5-6 三角模糊数分布

如上图所示在集对分析理论中, 差异度系数 $i$ 的取值范围为 $[-1, 1]$ , 而三角模糊数的隶属度函数值域为 $[0, 1]$ , 为使这两个值域的范围一致, 可将其值域从 $[0, 1]$ 拓展到 $[-1, 1]$ , 这种处理方式建立综合评价模型提供理论基础。

## 2.综合评价模型构建

三角模糊数-集对分析模型在地铁站位选址方案比选中的应用步骤: 首先构造多方案指标矩阵, 选取各方案中各类指标的最优值(成本型指标以取值最小为优, 效益型指标以取值最大为优)作为参考标准, 各指标的实测值与参考标准值构成集对, 并将该集对做同、异、反决策分析, 利用三角模糊数构造其差异度系数 $i$ 和集对的联系度 $\mu$ , 运用信息熵评价实测数据的效用, 由效用值确定各评价指标的权重, 来综合评价地铁站位选址方案的优劣。

### (1) 指标矩阵的构造

设地铁站位选址有 $m(m = 1, 2, \dots, m)$ 个方案, 每个方案有 $n(n = 1, 2, \dots, n)$ 类指标, 可得多方案指标决策矩阵 $H$

$$H = \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1n} \\ I_{21} & I_{22} & \cdots & I_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ I_{m1} & I_{m2} & \cdots & I_{mn} \end{bmatrix} \quad (5.9)$$

构造同一度指标矩阵, 分别取给出 $m$ 的个方案中 $n$ 类指标的最优值 $f_j(j = 1, 2, \dots, n)$ , 成本型指标, 则 $f_j = \min \{I_{ij}\}$ , 该类指标的同一度值为 $b_{ij} = f_j / I_{ij}$ ; 如果第 $j$ 类指标为效益型指标, 则 $f_j = \max \{I_{ij}\}$ , 该类指标的同一度值为 $b_{ij} = I_{ij} / f_j$ 。这种处理方式可将整个指标矩阵统一转化为极大型指标体系。

### 1) 基于三角模糊数的差异度计算

文中将地铁站位选址评价方案中的各指标数据分为5个等级, 设等级界限值为 $s_1 = 1, s_2 = 0.75, s_3 = 0.5, s_4 = 0.25, s_5 = 0$ 。将 $s_1$ 定义为同一度 $a$ , 系数为1,  $s_5$ 为对立度 $c$ , 系数为-1;  $b_1$ 为与 $s_1$ 相差1级的偏同差异度,  $b_2$ 为与 $s_1$ 相差2级的中差异度,  $b_3$ 为与 $s_1$ 相差3级的偏反差异度;  $i_1, i_2, i_3$ 分别为偏同、中、偏反差异度系数<sup>[13]</sup>。各差异度系数的取值可采取分析取值法来确定, 设 $i_1 = 0.5, i_2 = 0, i_3 = -0.5$ 。考虑相邻等级之间的差异度系数存在着模糊性, 可用三角模糊数来表示。因此, 建立差异度系数的表达式为

$$i_1 = \begin{cases} \frac{s_1 - c}{2(s_1 - s_2)}, & s_1 > c \geq \frac{s_1 + s_2}{2} \\ \frac{s_1 + s_2 - 2c}{2(s_1 - s_3)}, & \frac{s_1 + s_2}{2} > c \geq \frac{s_3 + s_2}{2} \\ 0, & s_1 \leq c \text{ or } c < \frac{s_3 + s_2}{2} \end{cases} \quad (5.10)$$

$$i_2 = \begin{cases} \frac{s_1 + s_2 - c}{2(s_1 - s_3)}, & \frac{s_1 + s_2}{2} > c \geq \frac{s_3 + s_2}{2} \\ \frac{s_3 + s_2 - 2c}{2(s_2 - s_4)}, & \frac{s_3 + s_2}{2} > c > \frac{s_3 + s_4}{2} \\ 0, & \frac{s_1 + s_2}{2} \leq c \text{ or } c < \frac{s_3 + s_4}{2} \end{cases} \quad (5.11)$$

$$i_3 = \begin{cases} \frac{s_3 + s_2 - 2c}{2(s_4 - s_2)}, & \frac{s_3 + s_2}{2} > c \geq \frac{s_3 + s_4}{2} \\ \frac{s_3 + s_4 - 2c}{2(s_4 - s_3)}, & \frac{s_3 + s_4}{2} > c > s_4 \\ 0, & \frac{s_3 + s_2}{2} \leq c \text{ or } c < s_4 \end{cases} \quad (5.12)$$

其三角模糊数的分布如图 5-7 所示。

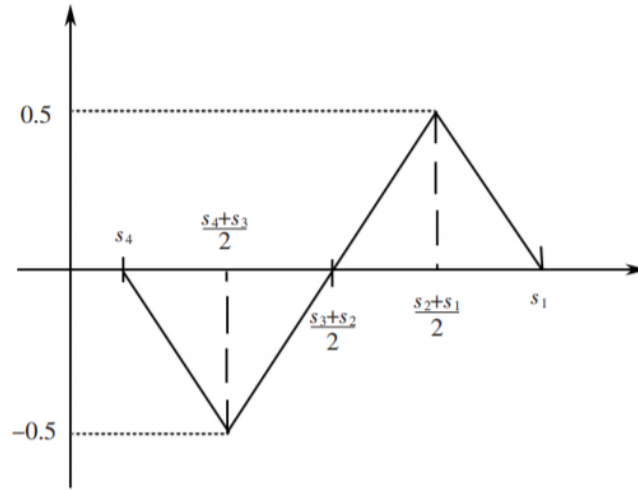


图 5-7 极大型三角模糊数分布

## 2) 基于三角模糊数的联系度计算

在地铁站位选址方案中，由于指标类型不同，在不同方案中各指标的实测值与标准等级的同、异、反的程度也存在差异。以 $s_1$ 级为标准，各指标实测值之间的联系度可表示为

$$\mu_{ij} = \begin{cases} 0, & c \geq s_1 \\ \frac{s_1 + s_2 - 2c}{s_2 - s_1} + \frac{2c - 2s_1}{s_2 - s_1} i_1 + 0i_2 + 0i_3 + 0j, & s_1 > c \geq \frac{s_1 + s_2}{2} \\ 0 + \frac{s_3 + s_2 - 2c}{s_3 - s_1} i_1 + \frac{2c - s_1 - s_2}{s_3 - s_1} i_2 + 0i_3 + 0j, & \frac{s_1 + s_2}{2} > c \geq \frac{s_2 + s_3}{2} \\ 0 + 0i_1 + \frac{s_3 + s_4 - 2c}{s_4 - s_2} i_2 + \frac{2c - s_3 - s_2}{s_4 - s_2} i_3 + 0j, & \frac{s_3 + s_2}{2} > c \geq \frac{s_3 + s_4}{2} \\ 0 + 0i_1 + 0i_2 + \frac{2s_4 - 2c}{s_4 - s_3} i_3 + \frac{2c - s_3 - s_4}{s_4 - s_3} j, & \frac{s_3 + s_4}{2} > c > s_4 \\ 0 + 0i_1 + 0i_2 + 0i_3 + j, & s_4 \geq c \end{cases} \quad (5.13)$$

式中： $c$ 为指标 $I_{ij}$ 的实际值； $\mu_{ij}$ 为实际值 $c$ 相对于 $s_1$ 的联系度值； $s_1$ 、 $s_2$ 、 $s_3$ 、 $s_4$ 分别表示各级指标的标准界定值。

由式 (5.10) ~ (5.13)，可计算基于三角模糊数的联系度值为

$$\mu_{ij} = \begin{cases} 0, & c > s_1 \\ \frac{(s_1-c)^2 + (s_2-c)^2}{(s_2-s_1)^2}, & s_1 > c \geq \frac{s_1+s_2}{2} \\ \frac{s_1^2 + s_1s_2 + 2cs_3 - 2cs_1 - s_1s_3 - s_2s_3}{2(s_3-s_1)^2}, & \frac{s_1+s_2}{2} > c \geq \frac{s_2+s_3}{2} \\ \frac{s_2^2 + s_2s_3 + 2cs_3 - 2cs_4 - s_2s_4 - s_3s_4}{2(s_4-s_2)^2}, & \frac{s_3+s_2}{2} > c \geq \frac{s_3+s_4}{2} \\ \frac{5cs_4 - 2s_4^2 - 2c^2 - 2cs_1 - cs_3 - s_3s_4}{2(s_4-s_2)^2}, & \frac{s_3+s_4}{2} > c > s_4 \\ -1, & s_4 \geq c \end{cases} \quad (5.14)$$

## (2) 方案指标权重的确定

对地铁站位选址方案的比选，需要考虑到不同指标的相对重要性，使得最终比选结果客观合理。在信息论中，由于熵值能反映信息的效用程度，可用来度量信息的效用。某项指标的效用值越高，则表示该项指标对决策的作用越大，其熵值越小。因此，建立信息熵评价指标矩阵，可用来评价各指标的权重。基于信息熵的权重计算步骤：

1) 对指标矩阵  $\mathbf{H}$  同一化处理，可得判断矩阵  $\mathbf{R}$

$$R = [r_{ij}], \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (5.15)$$

式中： $r_{ij}$  为方案  $i$  的第  $j$  个评价指标的值。

2) 定义判断矩阵  $\mathbf{R}$  中每个指标的投影  $g_{ij}$ ，计算表达式为

$$g_{ij} = \frac{r_{ij}}{\sum_{i=1}^m r_{ij}} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (5.16)$$

3) 计算各评价指标的熵值  $E_j$

$$E_j = -\left(\sum_{i=1}^m f_{ij} \ln(f_{ij})\right) / \ln(m), \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (5.17)$$

式中： $f_{ij}$  为  $g_{ij}$  的修正值，若  $g_{ij} > 0$ ，令  $f_{ij} = g_{ij}$ ；若  $g_{ij} = 0$ ， $f_{ij}$  的计算表达式为

$$f_{ij} = (1 + g_{ij}) / \sum_{i=1}^m (1 + g_{ij}) \quad (5.18)$$

4) 计算各项指标的熵权，得到指标权重向量  $\mathbf{W}$  为

$$\mathbf{W} = \Sigma[(w_j)_n] \mid w_j = (1 - E_j) / \left(n - \sum_{j=1}^n E_j\right), \quad \sum_{j=1}^n w_j = 1 \quad (5.19)$$

## (3) 综合评价模型

根据指标权重向量  $\mathbf{W}$  和联系度值  $\mu_{ij}$ ，可以得出三角模糊数-集对分析的综合评价模型：

$$b_i = \sum_{j=1}^n \omega_j \mu_{ij} \quad (5.20)$$

$$z_i = 0.5 + 0.5b_i \quad (5.21)$$

式中： $n$  为评价指标数量； $b_i$  和  $z_i$  分别为方案的联系度和综合计算结果，其中  $z_i$  的取值区间为  $[-1, 1]$ ， $z_i$  值越大，则表示该站位选址方案越好。

## 3. 评价指标

### (1) 评价指标体系的构建

城市地铁站位设置方案的比选结果，不仅要满足线路规划功能要求，还需满足技

术和经济上的要求。比选方案中各指标设定需要综合考虑其确定性和不确定性因素，构建可直接量化指标集合<sup>[14,15]</sup>。本文根据城市地铁站位方案比选特点，构建了3层结构的指标体系，方案层  $P$ 、准则层  $C$  和指标层  $I$ 。各层包含的内容如表 5-3 所列。

表 5-3 地铁车站站位选址方案评价指标

方案层	准则层	指标层
$P_i$	协调性 $C_1$	城市发展规划协调性地 $I_1$
		铁路网规划协调性 $I_2$
		其他交通方式协调性 $I_3$
	经济性 $C_2$	工程投资 $I_4$
		拆迁工程量 $I_5$
		可带动经济效益 $I_6$
		站点负荷强度 $I_7$
	交通功能性 $C_3$	平均出行时间 $I_8$
		运能匹配度 $I_9$
		客流断面均衡系数 $I_{10}$
	可实施性 $C_4$	工程建设难度 $I_{11}$
		对环境的影响 $I_{12}$

在表 5-3 列出的 12 个评价指标中， $I_1$ 、 $I_6$  为效益型指标，其他为成本型指标。

## (2) 评价指标量化

根据集对分析理论，对表 1 中的 12 个指标同一化处理，得到各指标的量化值。结合三角模糊数，将不确定的评价指标的转化为可计量的准确值。

### 1) 协调性准则 $C_1$

#### ① 与城市发展规划的协调性 $I_1$

$$I_1 = s_1 / s_2 \quad (5.22)$$

#### ② 与地铁路网规划的协调性 $I_2$

$$I_2 = 1 - \sqrt{\sum_{k=1}^K \left[ G_k (F_k - \lambda_k) / \left( \lambda_k \sum_{k=1}^K G_k \right) \right]} \quad (5.23)$$

#### ③ 与其他交通方式的协调性 $I_3$

$$I_3 = \sum_{m=1}^M \beta_m T_m \quad (5.24)$$

### 2) 经济性准则 $C_2$

#### ① 工程投资 $I_4$

#### ② 拆迁工程量 $I_5$

#### ③ 可带动经济效益 $I_6$

$$I_6 = 0.123 \ln(x_{k1}) - 0.55 \ln(x_{k2}) + 0.17 \quad (5.25)$$

### 3) 交通功能性准则 $C_3$

①客流吸引能力 $I_7$ 

$$I_7 = \sum_{j=1}^J \omega_j L_j, \omega_j = Q_j / \sum_{j=1}^J Q_j \quad (5.26)$$

②平均出行时间 $I_8$ 

$$I_8 = \sum_{n=1}^N \lambda_n T_n \quad (5.27)$$

③运能匹配度 $I_9$ 

$$I_9 = D / \left( \sum_{s=1}^S C_s \times P_s \right) \quad (5.28)$$

④客流断面不均衡系数 $I_{10}$ 

$$I_{10} = Q_k / q_k \quad (5.29)$$

4) 可实施性准则 $C_4$ ①客流吸引能力 $I_{11}$ ②平均出行时间 $I_{12}$ 

## 5.2.2 基于三角模糊数—集对分析的地铁车站选址模型的求解

1.根据呼和浩特市区域划分并结合附件 1 地铁站点的经纬度绘制出站点的位置图 5-8:

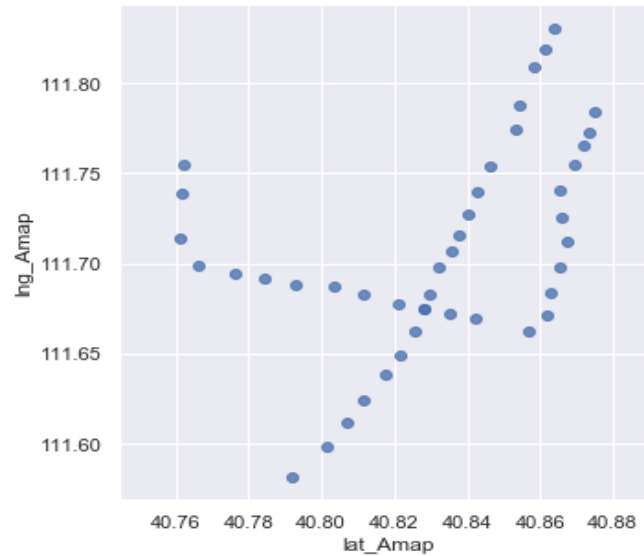


图 5-8 站点的位置

2.中心性定义了一个节点在网络中的重要程度,由与其相连的节点来确定,是判定网络中节点重要性的指标,是节点重要性的量化,介数中心性,以网络中所有节点对的最短路径中经过一个节点的最短路径数目来刻画节点重要性。经过一个节点的最短

路径越多，这个节点就越重要。介数中心性刻画了节点对网络中沿最短路径传输的网络流的控制力。 节点*i*的介数定义为 $BC(i)$ ，由

$$BC(i) = \sum_{s \neq i \neq t} \frac{n_{st}^i}{g_{st}} \tag{5.30}$$

式中： $g_{st}$ 为从节点*s*到节点*t*的所有最短路径的数目， $n_{st}^i$ 为从节点*s*到节点*t*的 $g_{st}$ 条最短路径中经过节点*i*的最短路径的数目符号说明。

计算图中每个站点的介数中心性，根据公式我们求得红色站点的介数中心性最高，记为核心站点。如图 5-9 所示。

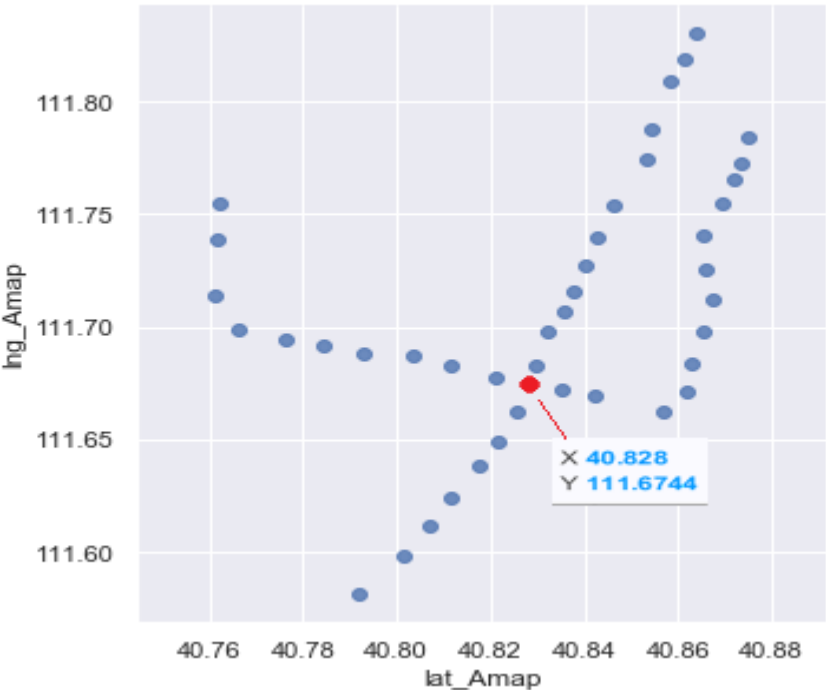


图 5-9 核心站点示意图

因此以该站点为标准，提出靠近该核心站点以北建立新路线为方案一，靠近该核心站点以西建立新路线为方案二，各方案互有优劣，在获取各指标的基础数据后，采用三角模糊数-集对分析模型对 2 个备选方案分别进行综合评价，通过计算结果的比较可得出选址方案的优劣。

通过实际的测量计算和论证分析，得到 2 个备选方案的评价指标数据如表 5-4 所列。

表 5-4 备选方案指标数据表

方案	与城市规划 协调性 $I_1$	与地铁路网 规划协调性 $I_2$	与其他交通 方式协调性 $I_3$	工程投资 $I_4$ / 亿元	拆迁工程量 $I_5/10^4m^2$	可带动经济 效益 $I_6$
P <sub>1</sub>	0.23	0.35	7.2	3.8	1.9	0.71
P <sub>2</sub>	0.47	0.22	9.4	2.6	1.6	0.53
方案	客流吸引能 力 $I_7/m$	平均出行时 间 $I_8/min$	运能匹配度 $I_9$	客流断面均 衡系数 $I_{10}$	工程建设难 度 $I_{11}$	对周边环境 影响 $I_{12}$
P <sub>1</sub>	279.2	13.7	0.54	1.2	0.78	0.54
P <sub>2</sub>	302.4	10.2	0.38	1.7	0.54	0.52

根据表 5-4 数据，构造指标矩阵并进行同一化处理，得矩阵*H*为



$$H=\begin{pmatrix} 0.49 & 0.63 & 1.00 & 0.68 & 0.84 & 1.00 & 1.00 & 0.74 & 1.00 & 0.71 & 0.69 & 0.96 \\ 1.00 & 1.00 & 0.77 & 1.00 & 1.00 & 0.75 & 0.76 & 1.00 & 0.70 & 1.00 & 1.00 & 1.00 \end{pmatrix}.$$

由式 (5.14) 可得不同方案中各评价指标的联系度值构成的矩阵 $\mu(\mu_{ij})$ 为

$$\mu=\begin{pmatrix} 0.27 & 0.49 & 1.00 & 0.39 & 0.07 & 1.00 & 1.00 & 0.27 & 1.00 & 0.33 & 0.37 & 0.73 \\ 1.00 & 1.00 & 0.21 & 1.00 & 1.00 & 0.19 & 0.23 & 1.00 & 0.35 & 1.00 & 1.00 & 1.00 \end{pmatrix}.$$

由式 (5.16) 可得各评价指标的投影 $g_{ij}$ 构成的矩阵 $g$ 为

$$g=\begin{pmatrix} 0.33 & 0.39 & 0.56 & 0.41 & 0.46 & 0.56 & 0.57 & 0.42 & 0.59 & 0.41 & 0.40 & 0.49 \\ 0.67 & 0.61 & 0.44 & 0.59 & 0.54 & 0.44 & 0.43 & 0.58 & 0.41 & 0.59 & 0.60 & 0.51 \end{pmatrix}.$$

根据式 (5.17) 和 (5.18) 可得第 $j$ 类指标的熵值 $E_j$ 为

$$E_j=(0.913 \quad 0.963 \quad 0.931 \quad 0.975 \quad 0.994 \quad 0.989 \quad 0.984 \quad 0.979 \quad 0.975 \quad 0.997 \quad 0.969 \quad 0.998).$$

利用式 (5.19), 得到第 $j$ 类指标的权重 $w_j$ 为

$$w_j=(0.261 \quad 0.141 \quad 0.207 \quad 0.075 \quad 0.018 \quad 0.033 \quad 0.048 \quad 0.063 \quad 0.075 \quad 0.009 \quad 0.093 \quad 0.006).$$

### 3. 预测每天总计乘坐地铁人数达到多少才能够实现地铁运营的盈利目标。

根据国内铁对运营企业的职能分工,运营企业的收入仅有票款收入,票款收入取决于平均票价和客流量两个因素。而平均票价的大小取决于票价的定价体系和开通站点的线网规模。站点开通后,在外界环境不发生明显变化的前提下,平均票价变化幅度很小,因此,在短时期内,运营企业收入取决于客流量。经过对呼和浩特地铁在基本行车间隔条件下客流量进行分析后,拟合出地铁客流趋势方程。

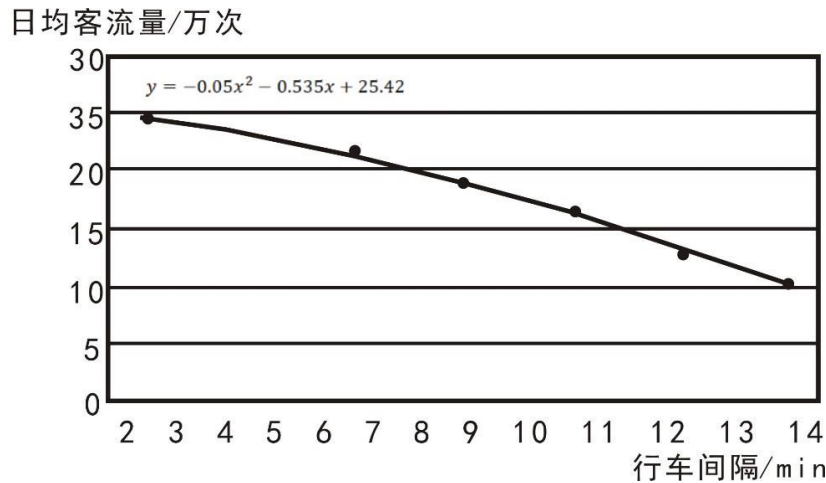


图 5-10 日均客流趋势拟合曲线

由该客流趋势方程,推算出客流量(表 5-5)。从图 5-5 不同行车间隔客流量的拟合曲线和表 1 数据可以看出,随着行车间隔的不断压缩,日均客流量趋势曲线的斜率趋于平缓,说明继续压缩行车时间间隔难以继续较大幅度提高日均客流量。在基本行车间隔压缩至 4-2min 后,要使客流量有较大幅度的增长,只有倚靠于开通线路的线网规模扩大或外界 社会环境发生明显变化。按照图 5-5 的日均客流量趋势拟合曲线来看,日均客流量最大值为 25.42 万人次/日。

表 5-5 不同行车间隔下的地铁客流量预测分析

行车间隔/min	12	10	8	6	5	4	3	2
日均客运量/万人次	13.38	16.87	18.91	20.4	21.5	22.48	23.37	24.15

地铁运营成本主要由工资及相关费用、运营能耗、生产维修费、营运费、管理费用、主营业务税金及附加、市场营销费 7 项组成。按成本是否随行车里程的变化,将地铁运营成本分为固定成本和可变成本。固定成本包括工资及相关费用、运营能耗中动力及照明用电、管理费用、维修费、营运费中不随行车里程变动部分,可变成本包括牵引用电、车辆维修成本、车辆保洁费、司机人工费等随行车里程变动部分。经过对呼和浩特地铁在行车间隔为 12-6 min 条件下的实际成本进行分析,及对行车间隔 5-2 min 条件下的成本进行预测后,得出不同行车间隔条件下地铁成本分析。如表 5-6 所示。

表 5-6 不同行车间隔条件下地铁成本分析表

行车间隔/min	12	10	8	6	5	4	3	2
日开行列车数/列	165	198	248	300	396	495	660	990
日开行列车增加数/列	——	16.5	25	26	96	99	165	330
年度总成本增加额/万元	——	36.76	52.09	31.12	73.84	569.56	615.3	1184.6
年度总成本增加率/%	——	1.018	1.044	1.06	1.097	1.38	1.688	2.278
固定成本占比/%	91.53	89.88	87.64	86.32	83.42	83.37	18.0	20.0
变动成本占比/%	8.47	10.12	12.36	13.68	16.58	16.63	82.0	80.0

表 5-7 不同行车间隔下的盈亏平衡日均客运量对比分析表

行车间隔/min	12	10	8	6	5	4	3	2
盈亏平衡日均客运量/万人次	23.07	23.49	24.09	24.46	25.4	32.18	38.02	49.14
预测可实现的日均客运量/万人次	13.38	16.87	18.91	20.4	21.5	22.48	23.37	24.15
日均客运量差额/万人次	9.69	6.62	5.18	4	3.9	9.7	14.65	24.99

### 5.2.3 结论

由式 (5.20) 和 (5.21), 可计算出各方案的综合评价价值, 其中方案 1 的值为 0.796, 方案 2 的值为 0.876。从综合评价结果可知, 方案 2 优于方案 1, 因此选择靠近核心站点以西建立新路线。

根据对客运支出和平均票价的测算, 可以得出呼和浩特地铁不同行车间隔下的盈亏平衡日均客运量, 结合表 5-4 客运量预测值可得出日均客流差额(表 5-6)。由表 5-6 可以看出, 当呼和浩特增加站点时, 要想实现地铁运营的盈利目标且日均客运量差额最小, 需要在增加站点的同时。缩短行车间隔为 5min, 需求日均客运量为 25.4 万人次。

## 5.3 问题 3 的模型建立与求解

### 5.3.1 单一上班时间点下的出行行为分析及优化模型的建立

出行行为理论作为交通管理和控制的重要理论之一, 对出行者的出行行为进行研究, 是合理制定交通错峰上班措施的前提, 是进行交通规划、建设和管理不可缺少的一项基础工作, 更是交通领域研究的重要课题。对个体出行行为的研究可以从整体上把握整个城市居民出行的基本情况, 从而为制定交通需求管理措施提供参考, 特别是随着经济水平的提高和和交通管理的科学化, 出行行为研究更是受到交通研究学者的重视。本文采

用“基于出行活动”的方法对出行者的出行行为进行分析。以单个出行者的总效益最大化为目标，建立数学模型，分析模型满足的平衡条件，研究单一上班时间点下出行者的出行行为和上班时间点的优化。

## 1.数学模型

### (1) 符号说明

出行者每天行驶于一条连接住宿地与工作地的地铁，早上从家去工作地，晚上从工作地回到住宿地，地铁上有一个通行能力有限的瓶颈(图 5-11)。

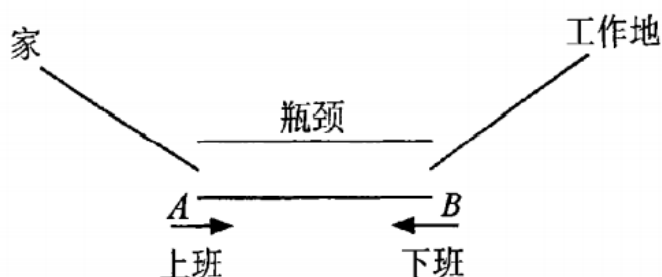


图 5-11 出行者一天活动的出行链

假设所有企业都采用非弹性工作制，单位时间的工作效益是关于时间的给定函数，且工作者每天的工作时长固定，令其为  $H$  小时。上班与下班是相互独立的，因此这里只对错峰上班下出行者的出发选择行为进行分析。并假设所有出行者具有相同的时间价值 (VOT),  $\alpha, \beta, \gamma$  以及  $t^*$ 。

由于模型中参数较多，为便于描述模型，下面对模型中用到的符号进行说明，

$t^0$ : 准时到达工作地的出行者的出行时间；

$SDE(t)$ : 早到工作地的延误；

$SDL(t)$ : 晚到工作地的延误；

$u_1(t)$ : 早上在家的边际效用函数；

$u_2(t)$ : 晚上在家的边际效用函数；

$T$ : 工作活动的边际效用随时间增加和减少的分界点；

$U(t)$ : 在时刻  $t$  从家出发的出行者的总效用；

$U_H(t)$ : 在时刻  $t$  从家出发的出行者早上在家的总效用；

$U_W(t)$ : 在时刻  $t$  从家出发的出行者的工作活动带来的总效用；

$C_T(t)$ : 在时刻  $t$  从家出发的出行者的负效用；

$TU$ : 每个出行者的出行效用；

$TTU$ : 系统总的出行效用；

### (2) 出行者的效用分析

由图 5-11 可看出，出行者一天的活动计划包含四种出行活动和两条出行链，四种出行活动分为：1)早上在家的活动；2)早上工作的活动 3)下午工作的活动；4)晚上在家的活动。

两条出行链为：1)早上从家到工作地的出行；2)晚上从工作地到家的出行。由于早高峰出行者出行行为的分析和晚高峰出行者的出行行为分析相类似，且早晚高峰出行者的出行行为是相互独立的，所以本文只研究早高峰期间出行者的出行行为。

出行者早上总的出行效用由以下三部分组成：1)早上在家活动产生的效用(早上在家为家人准备早餐)；2)工作活动产生的效用(工作赚钱)；3)早上从家去工作地出行产生的负效用(排队时间成本和计划延误成本)。

依据定义，我们有：

$$U(t) = U_H(t) + U_W(t) - C_T(t) \quad (5.31)$$

如图 5-12 所示，我们用  $t=0$  表示一天的开始时间， $t=24$  表示一天的结束时间，因此今天与昨天和明天都是相互独立的。假设出行者早上在家活动带来的效用、工作活动带来的效益和晚上在家活动带来的效用是关于时间  $t$  的线性函数，且早上在家的边际效用函数  $u_1(t)$  随时间  $t$  单调增加；晚上在家的边际效用函数  $u_3(t)$  随时间  $t$  单调减少；工作的边际效用函数  $u_2(t)$  随时间增加、 $u'_2(t)$  随时间减少。基于以上假设，早上在家和工作的边际效用函数可表示为：

$$u_1(t) = g_0 + g_1 t \quad (5.32)$$

$$u_2(t) = k_0 + k_1 t \quad (5.33)$$

其中， $g_0$ 、 $g_1$ 、 $k_0$  和  $k_1$  是参数，且  $g_1 < 0$ ， $k_1 > 0$

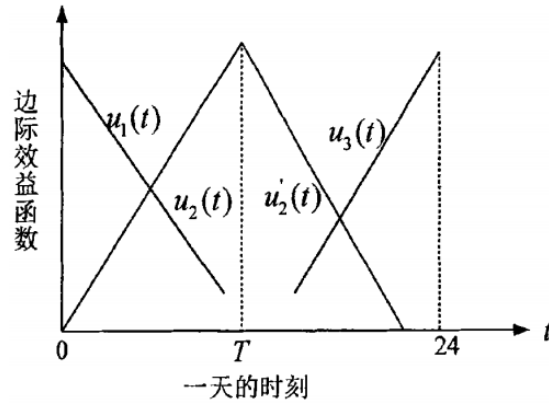


图 5-12 早上在家和工作的边际效用函数

由式(5.32)和(5.33)可得出在时刻  $t$  出发的出行者在家活动产生的总效用和工作活动产生的总效用分别为：

$$U_H(t) = \int_0^t u_1(t) dt \quad (5.34)$$

$$U_W(t) = \int_{t+T(t)}^T u_2(t) dt \quad (5.35)$$

$t$  时刻从家出发的出行者出行产生的负效用由排队时间费用和计划延误费用。(计划早到延误和计划晚到延误)两部分组成，如下：

$$C_T(t) = \alpha T(t) + \beta SDE(t) + \gamma SDL(t) \quad (5.36)$$

早到工作地和晚到工作地的延误分别为：

$$SDE(t) = \max\{0, t - T(t) - t\} \quad (5.35)$$

$$SDL(t) = \max\{0, T(t) + t - t\} \quad (5.37)$$

在  $t$  时刻瓶颈处的排队时间为：

$$T(t) = \frac{Q(t)}{s} \quad (5.38)$$

出行者排队长度的表达式满足下列微分方程：

$$\frac{dQ(t)}{dt} = \begin{cases} r(t) - s, & Q(t) > 0 \text{ 或 } r(t) \geq s, \\ 0, & \text{其他.} \end{cases} \quad (5.39)$$

把式(5.33)、(5.34)、(5.35)代入式(5.31)，得出单个出行者早上活动总的出行效益函数为：

$$U(t) = \int_0^t u_1(t)dt + \int_{t+T(t)}^T u_2(t)dt - \alpha T(t) - \beta SDE(t) - \gamma SDL(t) \quad (5.40)$$

## 2.单一上班时间点下的出行行为分析

### (1) 均衡条件

假设出行者的出发时间选择服从 UE 原则，即系统平衡时所有出行者的出行效用都相等，且没有人可以单方面通过改变自己的出行时间来使自己的总出行效用增加，满足以下平衡条件：

$$\frac{dQ(t)}{dt} = 0, \text{ 如果 } r(t) > 0 \quad (5.41)$$

### (2) 出行行为分析

单一上班时间点下出行者可能会遇到以下两种情形：(1)出行者总是早到工作地且排队；(2)出行者总是晚到工作地且排队，下面详细介绍这两种情况下出行者出发率的计算。在情形(1)下所有出行者都早到工作地且排队，即没出行者晚到工作地，相应的晚到惩罚为零，存在  $SDE(t) = t^* - T(t) - t$  和  $SDL(t) = 0$ ，把这两个式子和式(5.36)代入式(5.41)，得出单个出行者的总出行效用函数为：

$$U(t) = \int_0^t (g_0 + g_1 t)dt + \int_{t+T(t)}^T (k_0 + k_1 t)dt - \alpha T(t) - \beta [t^* - t - T(t)] \quad (5.42)$$

最先从家出发的出行者通过瓶颈时不需要排队，这种情形下，是排队从零开始形成的过程，相应的排队长度为：

$$Q(t) = R(t) - (t - t_s)ss \quad (5.43)$$

把式(5.39)、(5.43)和(5.44)代入平衡条件(5.42)，重新整理计算，可以得出在这种情形下出行者的出发率满足如下微分方程：

$$r(t) = \frac{\alpha + g_0 + g_1 t}{\alpha - \beta + k_0 + k_1 \left[ \frac{R(t)}{s} + t_s \right]}, t_s \leq t \leq t^0 \quad (5.44)$$

在情形(2)下所有出行者都晚到工作地，没有出行者早到工作地，且到达瓶颈所有出行者都需要排队，存在 $SDE(t)=0$ 和 $SDL(t)=-t^* + T(t) + t$ ，把这两个式子和式(5.36)代入式(5.41)，可得出这种情形下单个出行者的总出行效用函数为：

$$U(t) = \int_0^t (g_0 + g_1 t) dt + \int_{t+T(t)}^T (k_0 + k_1 t) dt - \alpha T(t) - \gamma [t + T(t) - t^*] \quad (5.45)$$

把式(5.39)、(5.44)和(5.46)代入平衡条件(5.42)，整理计算可得出在情形(2)下出行者的出发率满足如下微分方程：

$$r(t) = \frac{\alpha + g_0 + g_1 t}{\alpha + \gamma + k_0 + k_1 \left[ \frac{R(t)}{s} + t_s \right]}, t^0 \leq t \leq t_e s \quad (5.46)$$

最晚从家出发的出行者通过瓶颈不需要排队，即存在 $Q(t_e) = 0$ ，通过前面的定义可知 $R(t_e) = N$ ，把这两个式子代入(5.44)，整理变换可得出：

$$t_e - t_s = \frac{N}{s} \quad (5.47)$$

为了便于计算我们定义以下函数：

$$A(N) = (\beta + \gamma) - (g_1 - k_1) \frac{N}{s} \quad (5.48)$$

$$M(N) = \frac{(g_0 - k_0 - \gamma) \frac{N}{s}}{A(N)} + \frac{\frac{1}{2} (g_1 - k_1) \left( \frac{N}{s} \right)}{A(N)} \quad (5.49)$$

$$A_0(N) = 2(\beta + \gamma) - (g_1 - k_1) \frac{N}{s} \quad (5.50)$$

$$M_0(N) = \frac{(g_0 - k_0 - \gamma) \frac{N}{s}}{A_0(N)} + \frac{\frac{1}{2}(g_1 - k_1) \left(\frac{N}{s}\right)^2}{A_0(N)} \quad (5.51)$$

由平衡条件(5.42)知，平衡时所有人的出行效用都相等，最早出发和最晚出发的出行者效用也相等，存在 $U(t_e) = U(t_s) = U(t_0)$ 。把此式与式(5.48)相结合，可以得出最早出发出行者的出行时刻、最晚出发出行者的出行时刻以及准时到达工作地的出行者的出行时刻，如下：

$$t_s = \frac{\beta + \gamma}{A(N)} t^* + M(N) \quad (5.52)$$

$$t_e = \frac{\beta + \gamma}{A(N)} t^* + M(N) + \frac{N}{s} \quad (5.53)$$

$$t^0 = \frac{-(\alpha + g_0) + \sqrt{(\alpha + g_0)^2 - 2g_1 B}}{g_1} \quad (5.54)$$

式中：  $B = -(\alpha + k_0 - \beta)t^* - \frac{1}{2}k_1 t^{*2} - (g_0 - k_0 + \beta)t_s - \frac{1}{2}(g_1 - k_1)t_s^2$

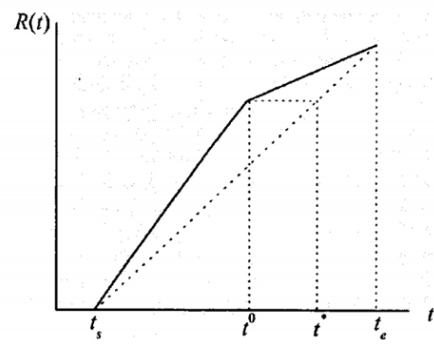
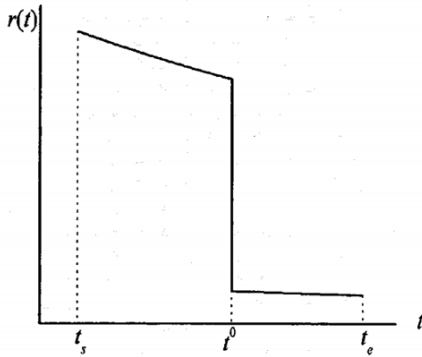


图 5-13 单个上班时间点下出发率随时间变化 图 5-14 单个上班时间点下累计出发人数随时间变化

图 5-13 和图 5-14 分别给出了 UE 平衡时单一上班时间点下的出行者出发率随时间变化和累计出发人数或累计到达人数随的时间变化。可看出与传统的基于出行的瓶颈模型相比，基于活动的瓶颈模型下出行者的出行行为虽然也分为早到和晚到两种情况，但出行者的出发率和出行时间都发生了变化。式(3.15)和(3.17)表明在基于出行活动的瓶颈模型下，出行者的出发率是关于出行时间 $t$ 的微分方程，而在基于出行的瓶颈模型下，出行者早到和晚到的出发率都为常数。

### 3.上班时间点的优化

由平衡条件可知系统总出行效用为系统总的出行人数与第一个出行者出行效益的乘积，由式(5.43)和(5.53)可得出：

$$\begin{aligned}
TTU = & N \left[ (g_0 - k_0 + \beta) \left( \frac{\beta + \gamma}{A(N)} t^* + M(N) \right) \right] + N \left[ k_0 T + \frac{1}{2} k_1 T^2 \right] \\
& + N \left[ \frac{1}{2} (g_1 - k_1) \left( \frac{\beta + \gamma}{A(N)} t^* + M(N) \right)^2 - \beta t^* \right]
\end{aligned} \quad (5.55)$$

式(5.56)中, 其他参数都是已知的, 只有一个未知参数 $t^*$ , 且系统总出行效用是关于上班时间点 $t$ 的二次函数, 系统总出行效用函数对上班时间点求一阶导数和二阶的偏导数可得:

$$\frac{\partial TTU}{\partial t^*} = \left( \frac{\beta + \gamma}{A(N)} \right)^2 (g_1 - k_1) t^* + \left( \frac{\beta + \gamma}{A(N)} \right) (g_1 - k_1) M(N) + \frac{\beta + \gamma}{A(N)} (g_0 - k_0 + \beta) - \beta \quad (5.56)$$

$$\frac{\partial^2 TTU}{\partial t^*} = \left( \frac{\beta + \gamma}{A(N)} \right) (g_1 - k_1) \quad (5.57)$$

由 $g_1 < 0$ ,  $k_1 > 0$  可知 $g_1 - k_1 < 0$ , 所以  $A(N) > 0$  始终成立, 由此可知式(5.58)小于零, 所以系统总效用函数是关于上班时间点 $t^*$ 的凹函数, 由 $\partial TTU / \partial t^* = 0$  可得出最优的上班时间点为:

$$t^* = \frac{A(N)^2 \beta}{(g_1 - k_1)(\beta + \gamma)^2} - \frac{A(N)(g_0 - k_0 + \beta)}{(g_1 - k_1)(\beta + \gamma)} - \frac{A(N)M(N)}{\beta + \gamma} \quad (5.58)$$

#### 4. 时变收费

拥挤收费作为交通需求管理的一种有效手段, 对城市整个交通系统的影响是综合性的。道路拥堵收费主要对城市道路基础设施投资政策、投资结构、城市路网布局以及交通供给的方式产生影响; 对交通需求而言, 拥堵收费会改变出行者的出行路线、出行时间和交通方式的选择。在一般确定型瓶颈模型中, 等待时间对出行者来说是一种纯浪费, 而拥挤收费不仅可以缓解道路交通拥挤, 减少出行者在上班途中的排队等待时间, 而且可以把拥堵收上来的费用作为交通管理部门新的财政来源, 把这部分费用用于交通基础设施的建设, 最终还是使广大的出行者受益。经典瓶颈模型中为了达到系统最优状态, 必须完全消除排队, 即出行者的排队出行费用为零。在系统最优状态下出行者的出发率和到达率都相等, 即为 $S$ 。由系统达到最优的定义可知, 最优的时变收费(收取的费用随时间连续变化)可以达到系统最优状态, 且不会影响总的出行交通量并完全消除排队。系统达到最优时高峰期的持续时间和总的计划延误费用并不发生变化, 但总的排队时间费用 为零。且在系统最优状态下每个出行者的出行效益不发生变化, 用某种上缴费用 代替瓶颈处的排队时间费用。在不收费状态下每个出行者的总的出行效用为:

$$\begin{aligned}
TU = & (g_0 - k_0 + \beta) \left( \frac{\beta + \gamma}{A(N)} t^* + M(N) \right) + \left( k_0 T + \frac{1}{2} k_1 T^2 \right) \\
& + \frac{1}{2} (g_1 - k_1) \left( \frac{\beta + \gamma}{A(N)} t^* + M(N) \right)^2 - \beta t^*
\end{aligned} \quad (5.59)$$



时变收费策略可以使出行者出行时间重新分布。时变收费下整个高峰期出行者的出发率等于瓶颈处通行能力，这样就可以消除出行者在瓶颈处的排队，时变收费水平为：

$$\tau(t) = \begin{cases} TU - \beta(t^* - t), & t \in [t_s, t^0] \\ TU - \gamma(t - t^*), & t \in [t^0, t_e] \end{cases} \quad (5.60)$$

式(5.60)表明，如果出行者不想排队就需缴纳一定费用；通过瓶颈不想缴费的出行者就要最早或最晚从家出发。在时变收费下出行者也可能会遇到早到工作地和晚到工作地两种情况。且两种情形下出行者到达瓶颈都不需要排队，所以早到和晚到的出发率都为  $s$ 。但在某一时刻  $t$  出发的出行者需要上缴一定的费用，所以在时变收费下，每个出行者的出行效用与不收费平衡相比并不发生变化。通过计算得出最早从家出发的出行者的出发时刻、最晚出发的出行者的出发时刻和准时到达工作地的出行者的出发时刻分别与式(5.53)、(5.54)和(5.55)相同。最优的时变收费可以完全消除瓶颈处排队，其通过改变出行者的出行分布来达到社会最优，并保持个人出行效用不变，减少了出行者的排队等待时间，增加了系统总的出行效用。

### 5.3.2 单一上班时间点下的出行行为分析及优化模型的求解

通过分析本文使用到的参数取值为  $\alpha = 10.0 \text{ \$/hr}$ 、 $\beta = 6.0 \text{ \$/hr}$ 、 $\gamma = 19.0 \text{ \$/hr}$ ，总出行人数  $N = 6000 \text{ veh}$ ，瓶颈处通行能力为  $s = 3000 \text{ veh/h}$ ， $T = 12.0$ ，边际效益函数的参数取值  $g_0 = 12.0$ ， $g_1 = -0.5$ ， $k_0 = 2.0$ ， $k_1 = 5.0/6.0$ 。

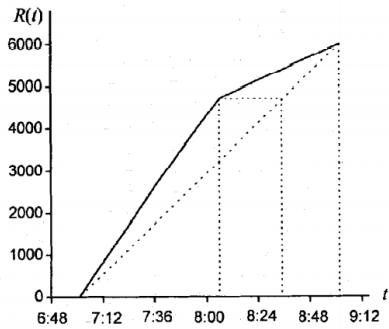


图 5-15 累计出发人数随上班时间变化

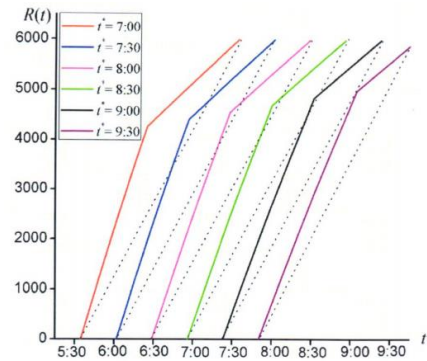


图 5-16 累计出发人数随上班时间的变化

图 5-15 为单一上班时间点下累计出发人数随时间的变化图。从图 5-15 可看出最优的上班时间点  $t^* = 8:35$ ，相应的出发时间为  $t^0 = 8:06$ 。最早和最晚从家出发的出行者的出行时间分别为  $t_s = 7:01$  和  $t_e = 9:01$ 。

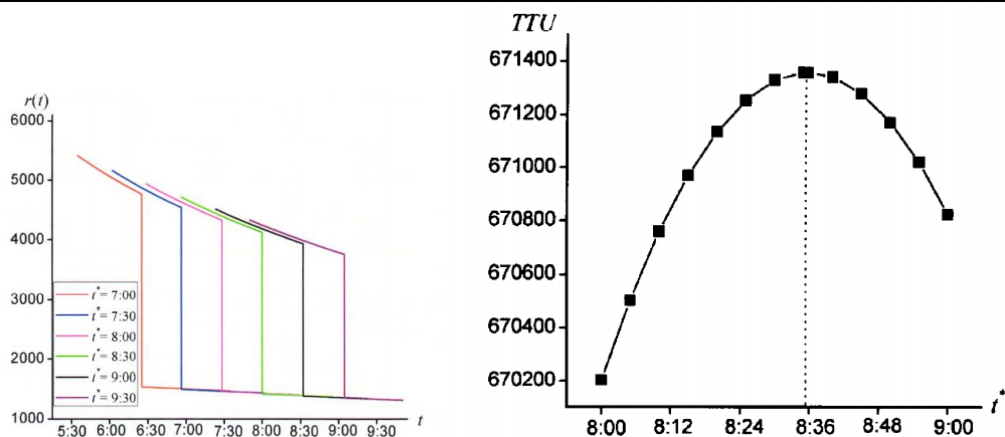


图 5-17 出发率随上班时间点的变化 图 5-18 系统总效用随上班时间点的变化

图 5-16、5-17 分别为上班时间点取 7:00、7:30、8:00、8:30、9:00 和 9:30 时累计出发人数和出发率随时间的变化。图 3.5 表明如果上班时间点较早，出行者在家的活动效益较高，因此早到工作地的出行者会比较少；上班时间点较晚，出行者在家的活动效益较低，则出行者会选择相对较早的时间点出发，因此晚到工作地的出行者较少。由图 3.6 表明上班时间点越大，出发率越小。

图 5-18 表示从 8:00 到 9:00 每隔 5 分钟取一个上班时间点，该上班时间点下系统总的出行效用。从图 3.7 看出随着上班时间点的增加，系统总效用先增加后减少，在上班时间点为 8:35 时，即最优的上班时间点下系统总出行效用达到最大，取值为 671361.60\$，与前面的理论分析结果相一致。

### 5.3.3 结论

本章主要研究了单一上班时间点情形下出行者的出行行为及上班时间点的优化。首先提出了满足条件的数学模型，得出模型满足的平衡条件，即平衡时所有出行者的出行效用都相等，出行者的出发时间的选择服从 UE 原则。接着分析了单一上班时间点下出行者出行行为并求得出行者出发率的解析解。以系统总效用最大为目标求得最优上班时间点的解析解。最后分析了在时变收费策略下出行者的出行时间为 8:35，能最大限度的减少高峰时期出行人数，实现了最终平峰目标。最优的时变收费可完全消除瓶颈处排队长度，与不收费平衡相比，每个出行者出行效用不发生变化，得益的是交通管理部门。

## 5.4 问题 4 的模型建立与求解

地铁与公交车换乘作为地铁站域集散行人的重要交通行为之一，在该换乘路径上存在着换乘距离模糊、相关功能业态布局杂乱、安全性与便捷性不佳等问题。换乘距离

过近，易使地铁站出入口与公交车站台之间产生人群拥挤，从而增加安全事故的发生几率；换乘距离过远，会影响行人寻路的舒适性和便捷性，有违地铁解决城市交通拥堵问题的初衷。为有效这个问题，提出基于 DBSCAN 来选择最优的位置修建公交线路。

### 5.4.1 基于 DBSCAN 算法模型的建立

#### 1.DBSCAN 算法原理

DBSCAN (Density-Based Spatial Clustering of Applications with Noise)，具有噪声的基于密度的聚类方法是一种基于密度的空间聚类算法。该算法将具有足够密度的区域划分为簇，并在具有噪声的空间数据库中发现任意形状的簇，它将簇定义为密度相连的点的最大集合。

在DBSCAN算法中将数据点分为三类：

- 核心点 (Core point)。若样本  $x_i$  的  $\varepsilon$  邻域内至少包含了  $MinPts$  个样本，即  $N_{\varepsilon}(X_i) \geq MinPts$ ，则称样本点  $x_i$  为核心点。
- 边界点 (Border point)。若样本  $x_i$  的  $\varepsilon$  邻域内包含的样本数目小于  $MinPts$ ，但是它在其他核心点的邻域内，则称样本点  $x_i$  为边界点。
- 噪音点 (Noise)。既不是核心点也不是边界点的点

在这里有两个量，一个是半径  $Eps(\varepsilon)$ ，另一个是指定的数目  $MinPts$ 。

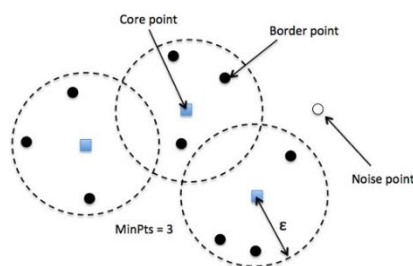


图 5-19 DBSCAN 算法原理

在DBSCAN算法中，还定义了如下一些概念：

**密度直达(directly density-reachable):** 我们称样本点  $p$  是由样本点  $q$  对于参数  $\{Eps, MinPts\}$  密度直达的，如果它们满足  $p \in NEps(q)$  且  $|NEps(q)| \geq MinPts$  (即样本点  $q$  是核心点)

**密度可达(density-reachable):** 我们称样本点  $p$  是由样本点  $q$  对于参数  $\{Eps, MinPts\}$  密度可达的，如果存在一系列的样本点  $p_1, p_2, p_2, \dots, p_n$  (其中  $p_1 = q, p_n = p$ ) 使得对于  $i = 1, \dots, n-1$ ，样本点  $p_{i+1}$  可由样本点  $p_i$  密度可达

**密度相连(density-connected):** 我们称样本点  $p$  与样本点  $q$  对于参数  $\{Eps, MinPts\}$  是密度相连的，如果存在一个样本点  $o$ ，使得  $p$  和  $q$  均由样本点  $o$  密度可达。

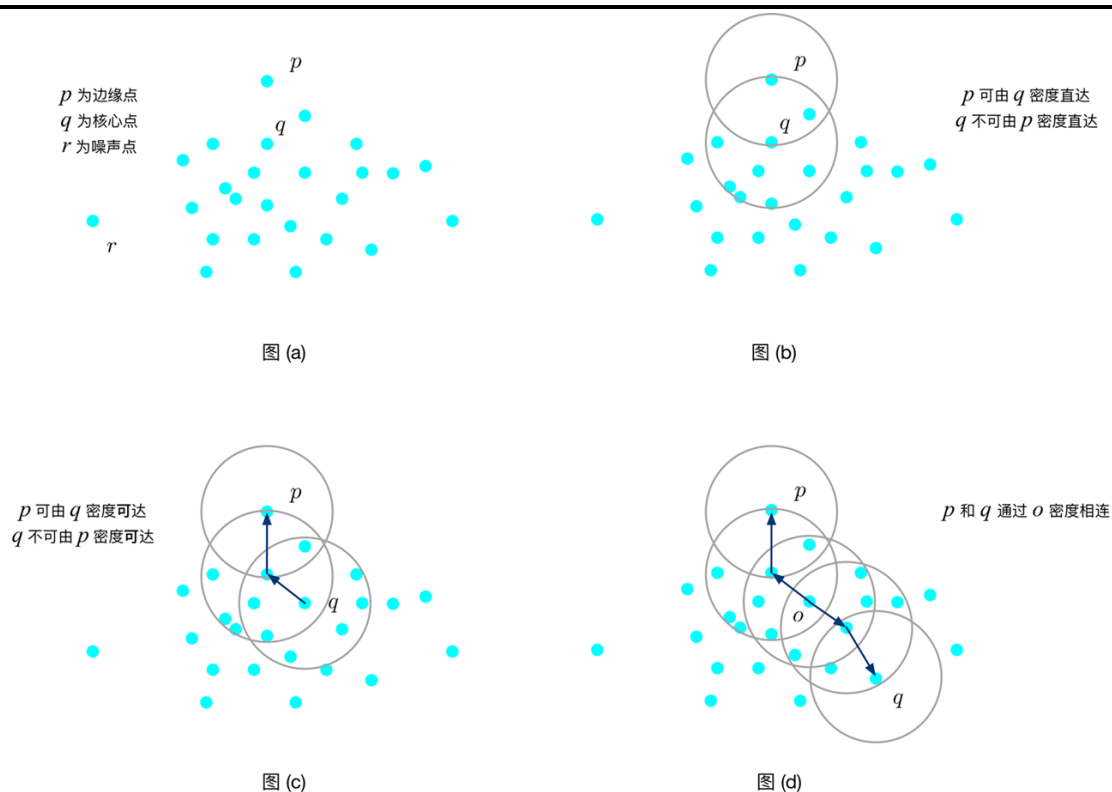


图 5-20 DBSCAN 算法原理

基于密度的聚类算法通过寻找被低密度区域分离的高密度区域，并将高密度区域作为一个聚类的“簇”。在DBSCAN算法中，聚类“簇”定义为：由密度可达关系导出的最大的密度连接样本的集合。

## 2. DBSCAN 算法流程

在DBSCAN算法中，有核心对象出发，找到与该核心对象密度可达的所有样本形成“簇”。DBSCAN算法的流程为：

- 根据给定的邻域参数 $Eps$ 和 $MinPts$ 确定所有的核心对象
- 对每一个核心对象
- 选择一个未处理过的核心对象，找到由其密度可达的样本生成聚类“簇”
- 重复以上过程

### 5.4.2 基于 DBSCAN 算法模型的求解

详情见图 5-21 结果

### 5.4.3 结论

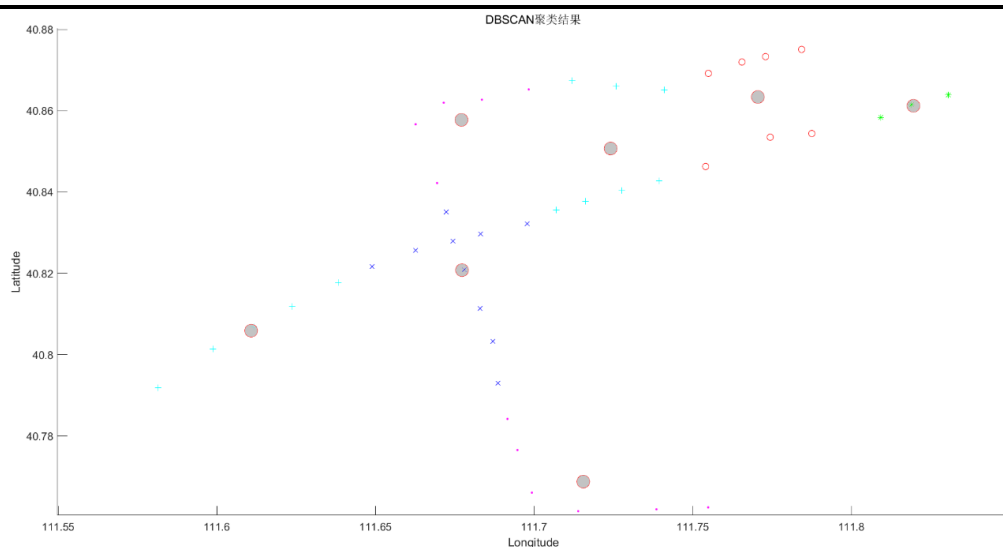


图 5-21 模型求解结果

经过 DBSCAN 聚类得到以下 7 个聚类点的纬度，经度如下：

(40.8634228571429,111.770498285714)

(40.8207847272727,111.677266181818)

(40.8612433333333,111.819531000000)

(40.7687320000000,111.715482166667)

(40.8058824000000,111.610802600000)

(40.8577890000000,111.677076400000)

(40.8507404285714,111.724102000000)

因此可以考虑在这七个密度聚类中心新增公交线路，以提高更多人员高峰出行时期的出行效率。

## 六、模型的评价及优化

### 6.1 模型的优点

1. NSGA-II 中提出了一种新的选择方法，这个方法不仅克服了第一代遗传算法的复杂的调参过程和高时间复杂度，还为后面的改进提出了可行的方向，也就是通过拥挤度来度量系统元素分布的情况，从而选择出那些分布均匀，获得信息最多的基因。随后这种方法被广泛采用于各个领域。

2. 本文将三角模糊数与集对分析理论结合，不仅提高了模型准确性，还为集对分析理论在描述事物属性的不确定性提供了一种有效的手段。

3. 相比其他聚类算法，DBSCAN 不需要预先声明聚类数量，可以对任意形状的稠密数据集进行聚类，相对的，其他聚类算法一般只适用于凸数据集。DBSCAN 可以在聚类同时发现异常点，对数据集中的异常点不敏感，DBSCAN 的聚类结果没有偏倚而其他聚类算法的聚类优劣依赖于初始值的选取。

## 6.2 模型的缺点

1. 基于 $NSGA-II$ 的多目标遗传算法没有考虑地铁线路间的乘客换乘影响，需要在今后进一步深入研究。
2. 专家给出的三角模糊数-集对分析的评价指标选择可能会有主观性。
3. 当空间聚类的密度不均匀、聚类间距差相差很大时，聚类质量较差，因为这种情况下参数 $MinPts$ 、 $Eps$ 选取困难。

## 6.3 模型的推广

不同行业一般通过设置不同的上班时间点来实现错峰上班。上班时间点的不同会影响出行者的出行时间的决策，在多个上班时间点下出行者的出行行为会变得更加复杂。对研究错峰上班下出行者的出行行为的研究一方面可以剖析错峰机理，为错峰上班制度的实施提供理论依据；另一方面可以帮助出行者制定更加合理的出行决策，减少出行的不确定性。在单个上班时间点下的出行行为分析的基础下，可以进一步研究多个上班时间点下的出行行为。

## 参考文献

- [1]刘兰芬,杨信丰.地铁客流分析及列车发车间隔优化研究[J].武汉理工大学学报(交通科学与工程版),2015,39(06):1119-1124.
- [2]中华人民共和国建设部.建标 104-2008, 城市轨道交通工程项目建设标准[S].
- [3]王林,陈璨.一种基于DE算法和NSGA-II的多目标混合进化算法[J].运筹与管理,2010,19(06):58-64..
- [4]KALYANMOY D,AMRIT P,SAMEER A, et al. A fast elitist multi-objective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation,2002,6(2):182-197
- [5]张天伟,聂磊,高桂凤.铁路客运站选址模型[J].交通运输工程学报,2011,11(05):83-87+92.
- [6]潘玲巧,焦永兰,林志安.基于集对分析的客运专线引入铁路枢纽客运站选址方案比选[J].中国铁道科学,2009,30(03):125-129.
- [7]吴红兵. 城市轨道交通线路与站点规划理论研究[D].重庆大学,2006.
- [8]刘兰芬,杨信丰.地铁客流分析及列车发车间隔优化研究[J].武汉理工大学学报(交通科学与工程版),2015,39(06):1119-1124.
- [9]谢海红,周浪雅,杨丽娟.铁路客运专线引入既有枢纽的方案选择[J].北京交通大学学报,2008(03):6-10.
- [10]MARIN A, NICKEL S, PUERTO J, et al. A flexible model and efficient solution strategies for discrete location problems[J]. Discrete Applied Mathematics, 2009, 157(5):1128-1145
- [11]魏丹.基于集对分析的道路交通拥堵判定[J].交通标准化,2013(03):56-58.
- [12]王纪鹏,欧阳治华,刘夏临.基于集对分析与三角模糊数的围岩稳定性综合评价模型[J].化工矿物与加工,2013,42(05):20-24.
- [13]赵克勤.SPA 的同异反系统理论在人工智能研究中的应用[J].智能系统学报,2007(05):20-35.
- [14]廖勇.基于三角模糊数的铁路客运站选址方案评价方法[J].中国铁道科学,2009,30(06):119-125.
- [15]夏臣芝,广晓平,王飞.基于集对分析的城市轨道交通线网规划方案比选研究[J].兰州交通大学学报,2010,29(04):139-143.

## 附录

代码:

```
clear;

close all;

clc;

k = 3;

Eps = 2;

%% 生成模拟数据

% n = 200;

% a = linspace(0,8*pi,n/2);

% u = [5*cos(a)+5 10*cos(a)+5]'+1*rand(n,1);

% v = [5*sin(a)+5 10*sin(a)+5]'+1*rand(n,1);

% mu1 = [20 20];

% S1 = [10 0;0 10];

% data1 = mvnrnd(mu1,S1,100);

load Lat.mat

load Lng.mat

data=[Lat,Lng];

% image = imread('data.png');

% image = image(:,:,1);

% [x,y]=find(image == 0);

% data=[x,y];

%% 准备变量，输出原始结果

[m,n] = size(data);

data=[(1:m)',data];

n = n + 1;
```



```
type = zeros(1,m);
cluster_No = 1;
visited = zeros(m,1);
class = zeros(1,m)-2;

figure(2);
plot(data(:,2),data(:,3),'k.');
grid on
daspect([1 1 1]);
xlabel('x');ylabel('y');
title('原始输入点');
hold on;

%% DBscan
Kdtree = KDTreeSearcher(data(:,2:3));

for i = 1:m
    % 抽取一个未访问点
    if visited(i)==0
        % 标为访问
        visited(i) = 1;
        point_now = data(i,:);
        Idx_range = rangesearch(Kdtree, point_now(2:3), Eps);
        index = Idx_range{1};
        if length(index) > k
            class(i) = cluster_No;
            while index
                if visited(index(1)) == 0
                    visited(index(1)) = 1;
                    if class(index(1)) <= 0
                        class(index(1)) = cluster_No;
                    end
                end
            end
        end
    end
end
```

```
        end
        point_now = data(index(1),:);
        Idx_range = rangesearch(Kdtree, point_now(2:3), Eps);
        index_temp = Idx_range{1};
        index(1) = [];
        if length(index_temp) > k
            index = [index, index_temp];
        end
    else
        index(1) = [];
    end
end
cluster_No = cluster_No + 1;
end
end
end

%% DBscan 聚类结果
figure;
for i = 1: cluster_No
    color = [rand(),rand(),rand()];
    data_class = data(find(class==i),:);
    plot(data_class(:,2),data_class(:,3),'.','Color',color,'MarkerFaceColor',color);
    hold on
end
data_class = data(find(class<=0),:);
plot(data_class(:,2),data_class(:,3),'k*');
hold on
grid on
daspect([1 1 1]);
xlabel('x');ylabel('y');
```

```
title('DBscan 聚类结果');
```