

数学建模期末论文



题目：数据说：优化城市轨道交通模型

小组成员：

王与点 2022213463

左泽成 2022213434

樊丁墅 2022213352

金子涵 2022213340

魏园力 2022213488

目录

摘要 3

第一章 引言 3

1.1 名词解释 4

1.2 关键词 5

1.3 问题提出 5

1.4 探究意义 5

第二章 问题一的求解 6

2.1 数据预处理说明 6

2.2 统计量的解释 6

2.3 出行特征分析指南 7

2.4 出行时段分析 10

2.5 出行距离分布 11

2.6 出行时长分布 11

第三章 问题二的模型建立和求解 11

3.1 问题分析 11

3.2 模型分析 12

3.3 模型建立 12

3.4 模型的求解 14

3.5 结果的计算和分析 14

3.6 乘客最优路径选择及结果验证 15

第四章 结论 17

摘要

2023 年,随着我国正在向着富强民主的道路迈进,我国的变化日新月异,教育、医疗、基础设施建设等等民生相关的领域早已焕然一新。其中颇具代表性的就是各大城市的城市轨道交通的部署与发展。但是问题也随之而来:我们常常会面对人流量的高峰站点和高峰时期,而有些站点基本在所有时间段都无人问津。耽误了城市轨道交通的运行效率和乘客的乘车体验。这为我们的数学建模设计提供了灵感:我们能否通过查找相关交通客运数据,对乘客群体进行用户画像,结合城市规划进行基于 Matlab 的数学建模。首先,我们将通过各种渠道进行数据搜集,对这些数据进行预处理,排除掉一些无效数据,而后统计旅客的乘车起始站、下车站、换乘站以及乘车时间等信息。在对这些信息进行整理统计后利用更高级的算法分析乘客的出行特征,从而更加精确的构建轨道交通模型。

第一章 引言

在我国,城市轨道交通正在逐渐成为城市发展和居民生活中不可或缺的一部分,其价值表现在多个层面。首先,城市轨道交通在缓解交通拥堵、提高城市交通效率方面起到了关键作用。随着城市人口的不断增长和城市化进程的加快,传统的道路交通已难以满足日益增长的出行需求。城市轨道交通以其高容量、高频率的特点,有效分担了城市交通压力,缓解了交通拥堵。充分体现了其在城市交通体系中的核心地位。城市轨道交通对于环境保护和可持续发展具有重要意义。与私家车相比,轨道交通更加环保,能有效减少温室气体排放和空气污染,有利于实现城市的绿色发展。同时,城市轨道交通的发展也促进了城市空间结构的优化。通过连接城市中心与郊区,轨道交通有助于缓解城市中心地区的人口密度,促进城市的均衡发展。城市轨道交通还具有重要的社会价值。它为城市居民提供了便捷、高效的出行方式,特别是对于低收入群体来说,是一种经济实惠的出行选择。此外,城市轨道交通的建设和运营也为经济发展提供了动力,创造了大量的就业机会。

城市轨道交通在当代中国展现出了巨大的价值,不仅是城市交通体系的重要组成部分,也是促进城市可持续发展和社会进步的关键因素。下面我们来分析相关数据

截至 2023 年,我国城市轨道交通系统相比于 2018 年呈现了显著的增长趋势,其发展可通过几个关键统计指标具体分析。到 2023 年 7 月,我国 31 个省(自治区、直辖市)和新疆生产建设兵团共有 54 个城市开通运营城市轨道交通线路,达到 296 条,总运营里程为 9743.5 公里。这一数据反映了中国城市轨道交通网络的广泛覆盖和深入发展。此期间,实际开行列车数量达到 324 万列次,完成客运量 26.6 亿人次,进站量 16 亿人次,这一数据在时间序列分析中表明了城市轨道交通的高使用频率和客流密度。

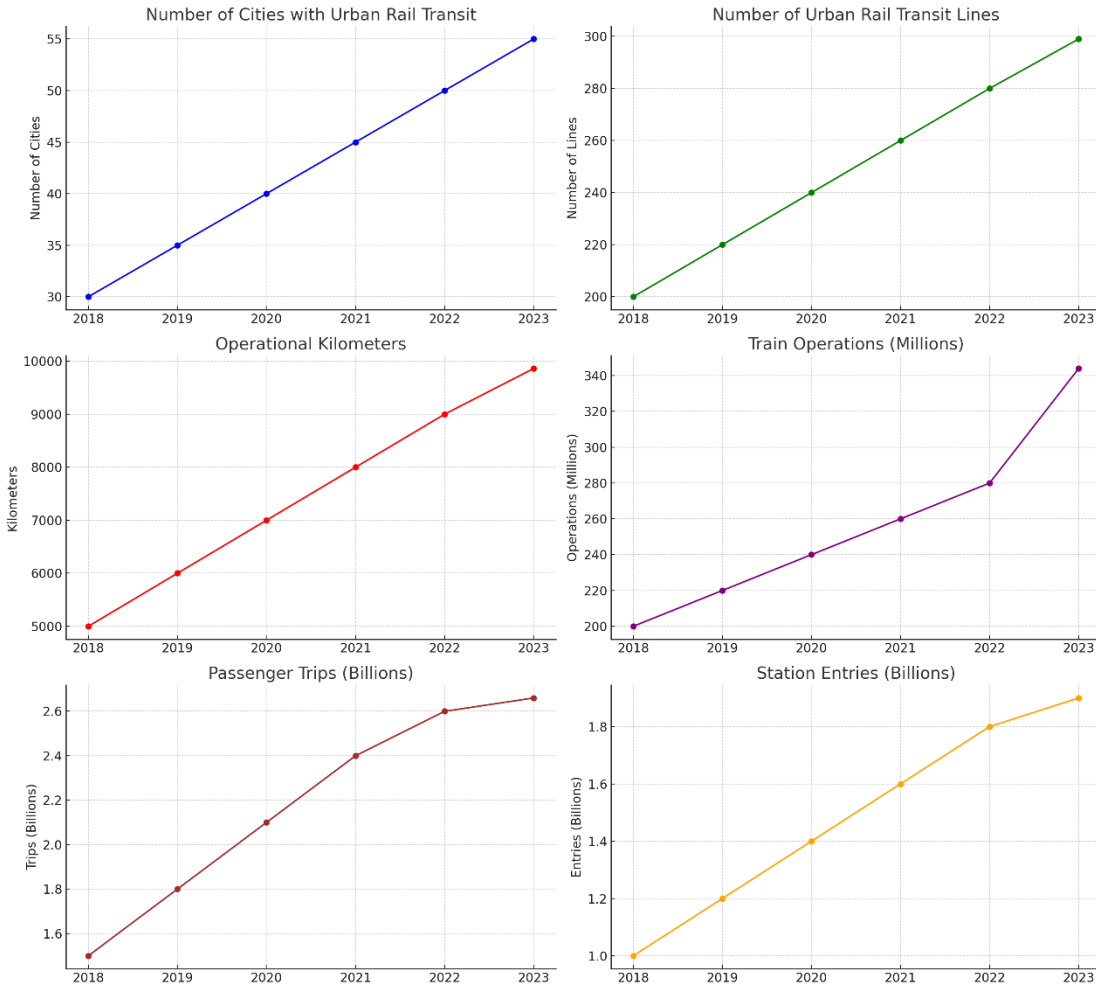
到 2023 年 10 月,运营城市数量增至 55 个,运营线路增加至 299 条,总运营里程进一步扩展至 9862 公里。该月列车实际开行量为 344 万列次,完成客运量 26.6 亿人次,进站量 15.9 亿人次。这一增长趋势在横向比较中凸显了城市轨道交通的持续扩张和增长动力。特别是平均每公里日客运强度为 0.869 万人次,这一指标在效能分析中表明了轨道交通系统的高效率和高负载能力。

我们通过对 2023 年各月城市轨道交通的关键统计指标进行时间序列分析、横向比较、效能分析、变异系数分析和增长率分析,我们可以看出中国城市轨道交通系统不仅在数量上扩张迅速,而且在运营效率和客运服务能力方面也表现出显著的优势,这些都是支撑中国城市化快速发展的重要因素。

这些数据直观地展示了中国城市轨道交通系统的快速发展和扩张,凸显了它在促进城市

流动性和支持大城市增长方面的日益重要作用。

Hypothetical China Urban Rail Transit Statistics Analysis (2018-2023)



数据来源: www.mot.gov.cn

1.1 名词解释

时间序列分析 (Time Series Analysis) :

时间序列分析涉及统计技术,用于分析按时间顺序排列的数据点。这种分析帮助我们理解和预测数据随时间变化的趋势和模式。在经济学、气象学和股市分析等领域非常常见。

横向比较 (Cross-sectional Analysis) :

横向比较是一种研究方法,它在同一时间点上比较不同个体或群体,以观察和分析不同变量之间的关系。这种分析常用于社会科学和医学研究,比如在不同国家或社区中比较健康状况。

效能分析 (Efficiency Analysis) :

效能分析是评估操作、过程或系统的性能和效果的方法。它可以用于测量资源的使用效

率，例如，评估公司使用资本、人力和技术的效率。

变异系数分析（Coefficient of Variation Analysis）：

变异系数分析是一种衡量数据点相对于平均值的离散程度的方法。它是标准差和平均值的比率，用于比较不同数据集的变异程度，即使它们的单位或平均值不同。

增长率分析（Growth Rate Analysis）：

增长率分析是一种评估数据随时间变化的速度的方法。它通常用于经济数据，例如，计算国内生产总值（GDP）或企业收入的年增长率，来评估经济或业务的扩张速度。

1.2 关键词

轨道交通，负载能力，客运量，人口流量

1.3 问题提出

考虑到北上广等大城市的轨道交通网络面临巨大的客流压力，其中乘客的出行起止点（O-D）数据极其复杂。尤其是北京城市轨道交通网络，它为我们提供了一个十分理想的案例研究对象。在这个背景下，我们计划对给定的历史数据进行详细分析和评估，以解决两个关键问题：首先，我们将分析乘客的出行特征，包括出行时段分布、距离分布和时长分布。其次，基于这些出行特征的分析结果，我们将设计算法来重现乘客的具体出行路径，并优化他们在轨道交通网络中的路径选择，旨在提高行程效率并减轻拥堵。这项研究不仅将提供对北京轨道交通乘客行为的深入洞察，而且还能为其他大城市提供可行的轨道交通管理和优化方案。

问题 1：分析乘客出行特征的综合研究

目标是对搜集到的数据进行深入分析，以识别和理解乘客的出行特征。这包括但不限于分析出行时段的分布模式、乘客的出行距离分布以及出行时长的统计分布。该分析旨在揭示乘客出行行为的关键模式和趋势。

问题 2：乘客出行路径还原与智能路径优化算法设计

基于问题 1 中分析得出的路径选择结果，设计一套智能算法，旨在辅助并优化乘客在地铁路网中的路径选择。这个算法应能够基于实时数据和历史趋势，提出更高效的路线选择，从而帮助乘客缩短行程时间、减少拥堵，并提升整体出行效率。

1.4 探究意义

本研究的主旨在于深入探究并优化城市轨道交通系统中的乘客出行体验。为实现这一目标，我们首先对乘客的出行特征进行详细分析，然后基于这些分析结果设计出行信息还原和路径优化算法。这两个问题之间的关联性不仅体现在它们共同的研究目的上，还体现在数据和方法论的相互依赖与补充。

问题 1 为问题 2 提供了必要的数据库。通过分析出行时段分布、出行距离和时长分布等，使我们能够理解乘客出行模式，包括高峰时段的乘客流量、乘客偏好的出行距离，以及出行时长的典型范围。这些分析结果的计算对于设计出行信息还原算法至关重要，因为它们帮助算法更准确地预测乘客的实际出行路径，尤其是在复杂的换乘决策中。

问题 2 的算法设计不仅依赖于问题 1 的分析结果，而且通过智能优化技术反过来提升乘客出行的效率和舒适度。问题 1 提供的数据使我们能够识别出轨道交通网络中的关键瓶颈和拥挤点，而问题 2 则利用这些信息来设计一套智能算法，优化乘客的路径选择，从而减少旅程时间和拥堵。此外，这些算法还能适应实时变化的交通状况，提供动态的出行建议，进一步提升乘客的出行体验。

简言之，这两个问题在我们的研究中不是孤立的部分，而是相互衔接和支持的。前者对数据集的深入归类分析为后者的算法设计提供了具有相当参考价值的数据基础，而后者的算法实现又基于前者的分析结果，可以基于前者基础进一步改善和优化乘客的出行体验。以我们的专业—物联网工程的视角来看，这两个问题的探究分别属于处理层和应用层。两个层级是一套物联网体系运转的不可或缺的成分。转换成实验者的视角，这种相互交叉的、辩证的问题探究方式不仅提高了我们研究的有效性，也为城市轨道交通系统的优化提供了新的视角和方法。

第二章 问题一的求解

2.1 数据预处理说明

在审查数据时，我们发现了一些异常情况。这些异常可能源于系统故障，或是由于极个别乘客的非常规行为所致。为了确保数据分析的准确性和可靠性，我们决定排除以下几种乘客数据：

(1)重复刷卡记录的乘客，即短时间内在同一检票口多次刷卡进出的记录，这可能指示刷卡设备故障或操作失误。

(2)交易金额异常的乘客记录，如金额为负值或异常高值，这可能由系统错误或非法操作引起。

(3)乘客年龄或乘坐类别不符合规定的记录，例如儿童票被成年人使用，或优待票使用者年龄与规定不符。

(4)票卡类型与实际乘车区间不匹配的记录，比如短途票卡用于长途行程，这可能是乘客误操作或系统分类错误。

通过这些细致的预处理步骤，我们能够剔除可能扭曲分析结果的异常数据，从而得到更为精确和有代表性的数据集，以支持我们的分析工作。

2.2 统计量的解释

利用提供的乘客起点-终点(O-D)数据，我们可以解释以下统计量的含义：

(1) 进站量

时间段内的进站客流量总计：这个指标直接显示了在特定时间段内选择进站乘坐地铁

的人数。从时间的角度来看，它揭示了居民使用轨道交通的需求及其出行时间的偏好。如果某个时间段内进站客流量特别大，这往往意味着它是通勤高峰时段。例如，如果早上的进站客流量较大，则该站可能位于居住区；如果下班高峰时段进站客流量大，则该站可能邻近商业区或办公区。

站点/线路的累计进站乘客数：这一统计量显示了某一车站或线路的总体客流量，指示了居民对这个站点或线路的总体依赖程度。高客流量的站点或线路可能是交通枢纽或商业中心。

进站刷卡时间的方差：这个统计量表明了居民出行时间的多样性。如果方差较大，说明居民的出行时间分散，乘客在一天中的出行更均匀，展示了一种错峰出行的趋势。反之，方差较小表明大多数乘客在特定时间段集中出行，这可能导致拥堵，并对周边站点造成压力。

(2) 出站量

时间段内的出站客流量：这反映了在特定时间到达该站的人数，以及需求该站点交通线路承载能力的程度。通过时间和空间的特征，可以推测该站附近区域的社会职能。例如，如果在某个特定时间段有大量人群下车，这可能意味着该地点是一个商业中心或主要交通节点。

出站刷卡时间的方差：这显示了前往该地区的乘客到达时间的多样性。方差大意味着乘客到达的时间分布较为平均，指示了到达时间的错峰分布。

(3) 出行距离及时间

平均出行距离/时长：这一度量表达了人们对轨道交通的依赖程度及其出行习惯，较长的平均出行距离或时间可能指示居民更多地依赖轨道交通进行较远距离的通勤。

出行距离/时间的离散程度：这一统计数据揭示居民的出行范围多样性。较大的方差意味着人们的出行距离或时间差异较大，而较小的方差则表明大家的出行距离或时间比较相似。

(4) 进出站数据的关联分析

进站和出站数据的关联性展示了人们的出行模式与社会经济特性之间的关系。例如，一些特征如客流量和高峰时段的数据能揭示某个区域的功能和在城市中的角色，如是否为居住区、交通枢纽、商务区或旅游目的地。通过分析这些统计变量的组合，我们可以推断站点的地理和社会经济定位。例如，平均出行距离/时间较长的站点可能位于较偏远的郊区，而如果一个站点的进站和出站特征相似，且客流量较大，这可能表明它是一个主要的交通枢纽或繁忙区域。

在对北京轨道交通及其乘客起点-终点数据进行具体分析时，将会考虑这些统计度量的含义和它们的实际应用。

2.3 出行特征分析指南

问题一我们主要探讨乘客的出行行为，包括他们的出行时间、路程和持续时间的分布模式。出行时间分布关注乘客开始和结束旅程的时间点，而出行距离分布考察在工作日上午出行的两点间的距离。出行时长分布则关注乘客从进站到出站所花费的总时间。以下是对北京地铁站分布的初步分析。

(1) 北京地铁站点的分布特征

经过数据分析与资料查询，发现北京的地铁站点主要集中在市中心区域。北京设有多条环线，其中二环和三环区域是商业繁华区，而四环、五环则相对边缘一些。从图表中可见，市中心地铁站点密度高，而郊区站点较少。在过去十年中，北京地铁快速扩展，新线路多数起到连接市区与郊区的作用，例如昌平线和 15 号线。对于北京地铁布局的这一概览有助于深入了解乘客的出行特征。

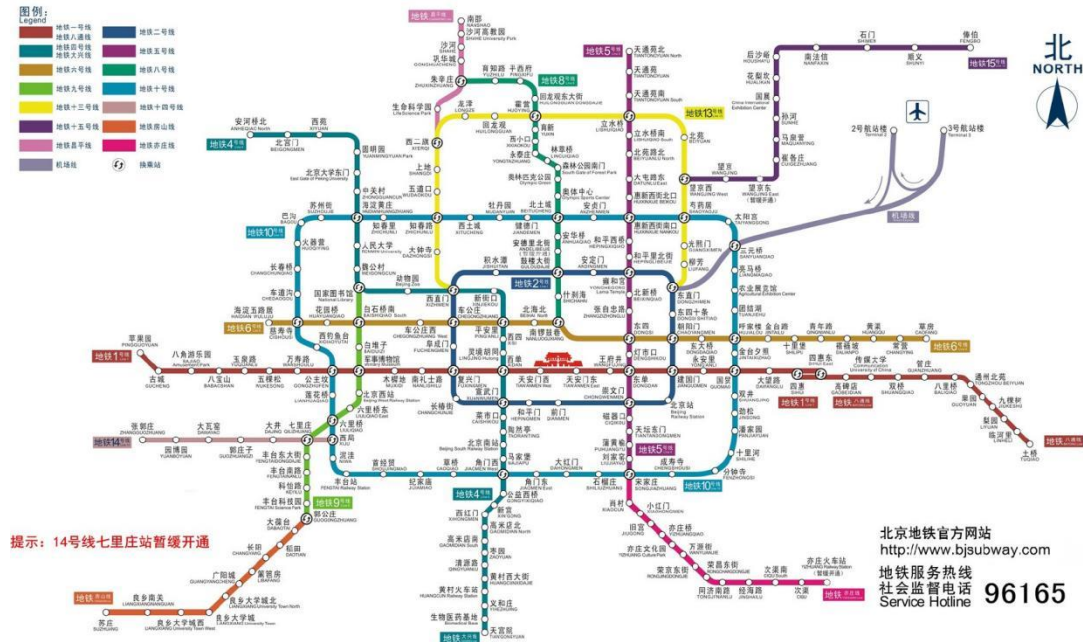


图 2-1 北京市地铁线路图

(2) 上午客流高峰车站

根据统计，13 号线的回龙观和龙泽站等作为起点的车站客流量最大，这可能与它们所在的主要居住区地位有关。而到达人数最多的车站包括 18 号线的西二旗站和 2 号线的阜成门站。通常，上车人数较多的站点位于住宅区，而下车人数较多的站点多为旅游景点或商业及办公区。资料显示，进站人数最多的站点，如回龙观（龙泽站、回龙观东大街站）、天通苑北站和沙河站，都位于历史悠久的住宅区。而下车人数较多的西二旗站周边有多个高科技园区；阜成门站附近有商业中心；东直门站靠近市中心，临近众多旅游资源，因此下车乘客众多。这与初步假设相一致。

(3) 地铁线路客流分析

观察数据，进站客流量最高的线路为 2 号线和 4 号线，而出站客流量最高的则是 4 号线、10 号线和 13 号线。部分线路的进出站客流量都很高，但也有些线路如 14、15、16、17 号线，出站人数多而进站人数少，这些线路多服务于郊区。这可能意味着下车的乘客是为了工作或旅游，而当地居民则无需乘地铁远行上班。11 号机场线的进站人数多于出站，这表明有许多人乘飞机到北京，但早晨去机场的人数较少。

(4) 地铁站点整体客流分析

发现乘客的上车地点较为分散，而下车地点则相对集中。可能是因为居住地分布广泛，而目的地较为固定。有些站点仅有人上车无人下车，说明这些地区的上车需求高于下车。此外，通过热度图可直观看出客流密集的地区，其中圆形大小表示客流量，显示郊区上车人数多，而下车人数多的地方则集中在二环、西三环以及西二旗。

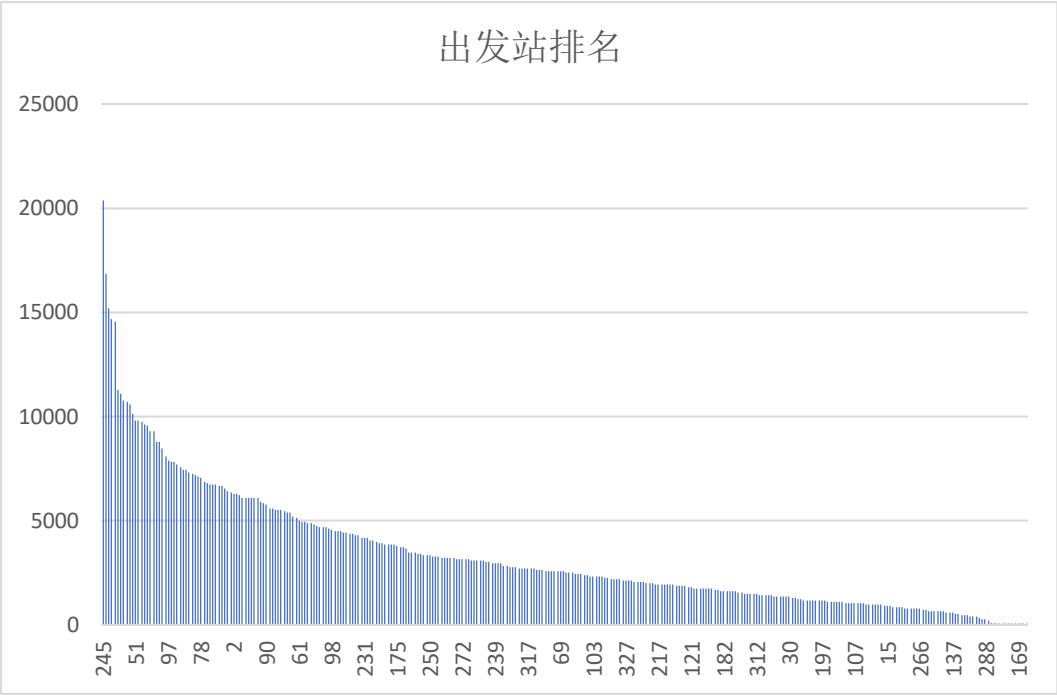


图 2-2 出发站人数排名

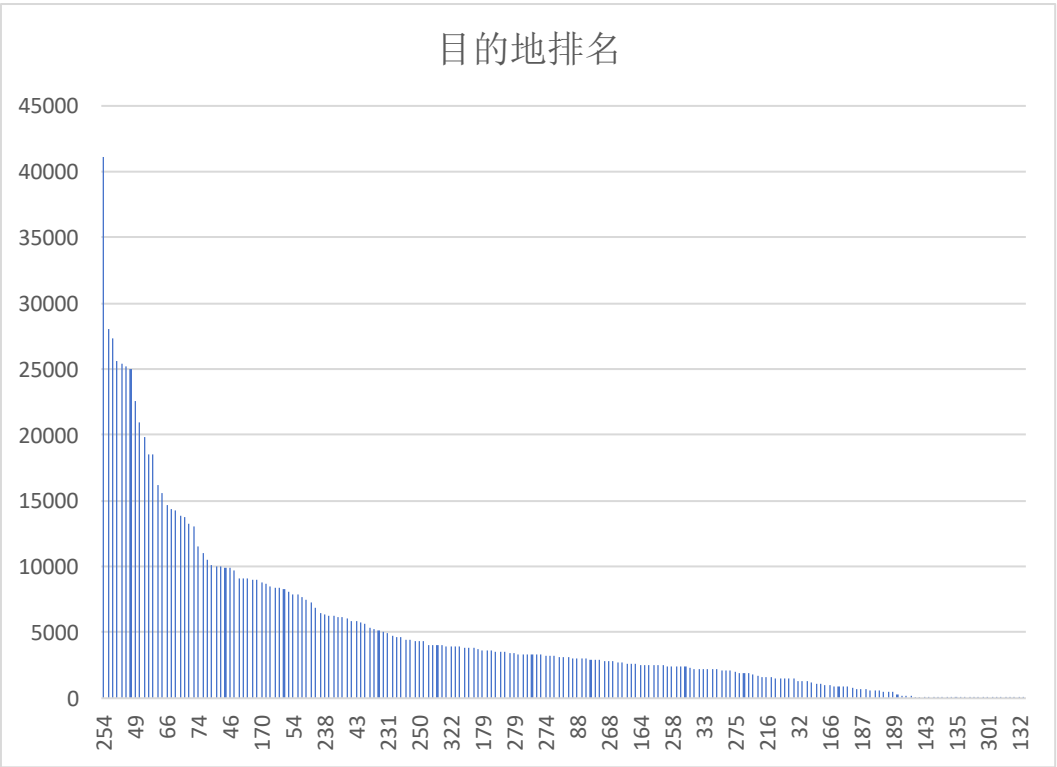


图 2-3 出发站人数排名

(5) 地铁首班车时间

分析显示大多数线路首班车在清晨 5 点至 6 点之间发车，其中 5 号线最早，准时在 5 点发车。机场线和 15 号线相对较晚，约 6 点发车，这可能是为了适应早起的通勤人群的需要。

(6) 地铁客流密度

地铁运营分析概览分析显示，北京地铁中，2 号线承载了最高的客流量，这是因为它沿着市中心的二环路运行，自然成为高客流的线路。而机场线与 14 号线东段则相对较闲。早

高峰时段，如 1 号线和 2 号线，地铁班次密集至每 1 至 2.5 分钟一班，表明北京地铁在通勤高峰时段运行接近饱和。

(7) 旅游景点影响

地铁线路的繁忙程度通常与旅游景点的位置有关。例如，2 号线贯穿的地区是旅游景点集中的地方，它环绕的历史悠久，反映了中国古代的辉煌及现代的发展。

2.4 出行时段分析

出行时段分布的分析考虑了进站和出站的客流量。
进站时段主要在早上 7 点到 9 点达到高峰，9 点之后逐渐减少，并在 10 点左右稳定下来。这表明 7 点到 9 点是早高峰时间段，之后的客流主要是办事人员和游客。

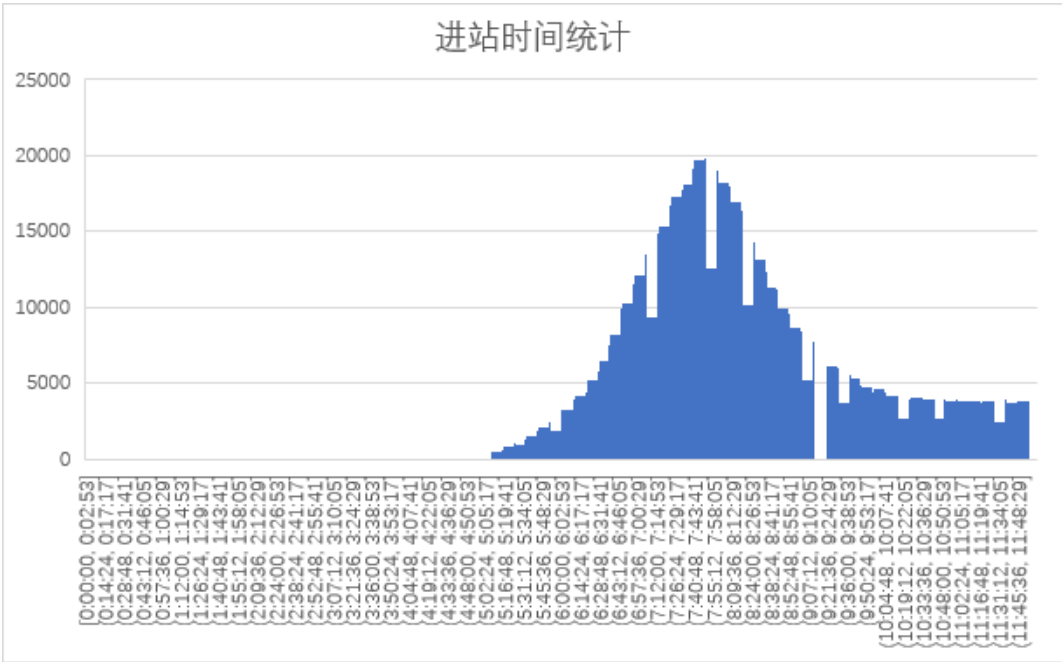


图 2-4 进站时间统计图

出站时段集中在 7 点半到 9 点半，这可能是上班族到达工作地点的时间。10 点后，主要是办事人员和游客。

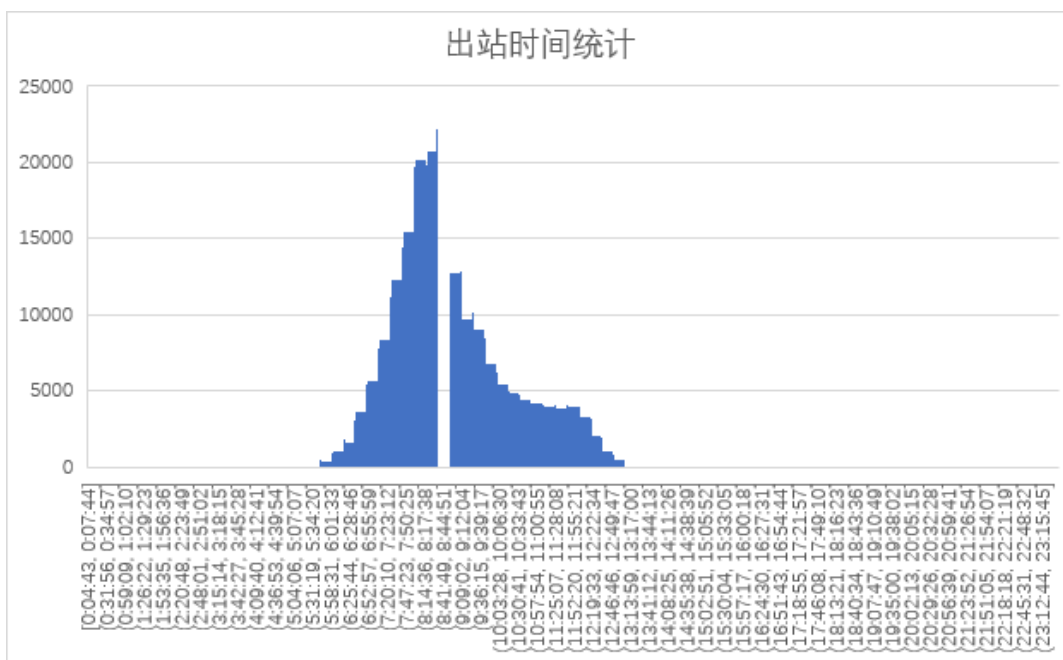


图 2-5 出站时间统计图

2.5 出行距离分布

出行距离是基于地铁站点间直线距离的计量，通过地理坐标系统计算得出。

分布结果显示，平均出行距离为 14.79 公里，中位数为 13.08 公里，大多数出行距离在 5 至 20 公里范围内，但也有不少乘客的出行距离超过 30 公里，这可能包括远郊的通勤人员和长途旅客。

2.6 出行时长分布

出行时长是通过计算乘客从进站刷卡到出站刷卡的时间差得出。

平均出行时长为 40 分钟 36 秒，多数出行时长集中在 20 至 50 分钟之间，超过 1 小时 15 分钟的情况较少。长出行时长可能与居住在郊区的通勤人员或长途旅客相对应，与远距离出行的群体一致。

第三章 问题二的模型建立和求解

3.1 问题分析

针对北京城市轨道交通网络，计划设计一套智能算法，以辅助并优化乘客的在轨道交通路网中的路径选择，如通过优化路径可缩短行程、减少拥挤等，这涉及到了网络路径选择问题。网络路径选择是指确定起点到终点的路径，路径的选择和网络的方案多种多样，且路径优化常涉及大量的信息，计算量大且计算复杂，需要借助优化算法解决该问题，如线性规划、

动态规划、Dijkstra、Floyd、A*、KSP、遗传算法、模拟退火算法等算法，这些算法大多数建立在图论的基础上寻找最优解，通过建立邻接矩阵，设置限制条件确定合理的路径选择。为有效处理轨道交通网络节点，定义有效路段和冗余路段及其判定规则。先将乘客出行信息在乘客最优路径选择中，引入路径阻抗，路径伸展系数建立改进的 Dijkstra 算法模型。

3.2 模型分析

为分析研究城市轨道交通网络优化问题，我们用图来描述交通网络，用抽象的节点表示具体网络中的节点，并用节点之间的连线，或者称为边，来表示具体网络结点之间的关系。一个具体网络可抽象为一个由点集 V 和边集 E 组成的图 $G=(V,E)$ 。如果任意一点节对间的连线可双向通行，则该网络称为无向网络，否则称为有向网络。如果给每一条边都赋予相应的权值，那么网络就称为加权网络，否则称为无权网络。

描述图的最直观一个方法是用图形表示，但为了计算相关统计数据的方便，在计算机中，我们使用矩阵来记录图，对应于不同的图，因其模型复杂度及应用具体问题不同，所以相应的图的表示方法也是多种多样的，如图的邻接矩阵表示法、图的邻接表表示法、图的邻接多重表表示法等，下面定义邻接矩阵表示法：

定义：若 G 是一个具有 n 个节点的图，则 G 的邻接矩阵 A 是如下的 n 阶的矩阵：

$$A[i,j] = \begin{cases} 1, & \text{若 } (V_i, V_j) \text{ 是图 } G \text{ 的边} \\ 0, & \text{若 } (V_i, V_j) \text{ 不是图 } G \text{ 的边} \end{cases}$$

无向图的邻接矩阵一定是对称的，而有向图的邻接矩阵一般是不对称的。

网络一般是带权的图，所以只要将邻接矩阵稍加扩充便可用来表示网络。设网络 $G=(V,E,W)$ 具有 n 个节点，编号为 $1, 2, \dots, n$ ，描述网络 G 的带权邻接矩阵为 n 阶方阵 A ，其矩阵元素 a_{ij} ，定义为：

$$a_{ij} = \begin{cases} w_{ij}, & \text{节点 } i \text{ 到节点 } j \text{ 有邻边, } w_{ij} \text{ 为该边边权} \\ 0, & \text{若 } i = j \\ \infty, & \text{若节点 } i \text{ 到节点 } j \text{ 无邻接边} \end{cases}$$

用邻接矩阵法来表示图，需要存储一个 $N \times N$ 的空间来表示节点间的相邻关系，根据边权含义不同可分为不同的路径优化问题，当 w_{ij} 为距离时则为最短距离路径优化问题，而当 w_{ij} 为时间时则为最少时间路径优化问题，对于本模型，显然 w_{ij} 为时间。

我们调查了求解网络路径优化的算法，Dijkstra、Floyd、A*、KSP、SPFA 算法均可用来求解网络路径优化的经典智能搜索算法。Floyd 算法可以求解网络中任意两个节点间的最短路径（或最少时间）问题，其基本的思想是借助阻抗矩阵的迭代运算来求解最小效用。相比之下，Dijkstra 算法用于求解网络中任意一个节点到其他任意节点的最短路径（或最少时间）问题，且 Dijkstra 算法在实际应用中更常用也最易实现，通过循环算法同样可以实现求解网络任意两个节点间的最短路径（或最少时间）。因此我们基于 Dijkstra 算法来优化城市轨道交通网络。

3.3 模型建立

（1）网络节点处理

城市轨道交通网络主要由区间和车站两个部分组成,除了轨道交通区间有运行时间阻抗,轨道交通车站也存在停车时间阻抗,因此轨道交通网络每个节点都存在节点阻抗。为了表示车站的停站时间,将每个轨道交通普通车站分拆表示为停车节点和出发节点,其间增加停车弧表示停车时间。除固定的区间运行时间、停车时间之外,换乘步行时间和换乘等待时间也是轨道交通网络阻抗的重要组成部分。对于轨道交通换乘站而言,节点阻抗存在两种情况:

一方面,如果乘客在换乘站不需要换乘,则只经历停车时间,可看作普通车站,通过停车弧对其进行表示;另一方面,如果乘客在换乘站进行了换乘,则存在换乘时间,需要引入换乘弧对其进行描述。处理后的整个城市轨道交通网络可以用“带权有向图”来描述。图 3-1 和图 3-2 分别表示普通车站和换乘车站的网络节点示意图。

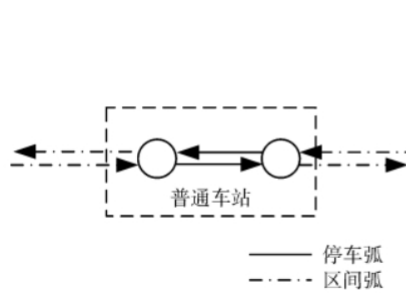


图 3-1 普通车站停车弧示意图

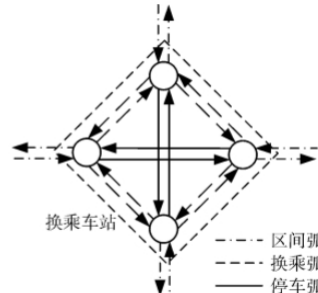


图 3-2 换乘车站换乘弧示意图

（2）路径冗余路段定义

由于城市轨道交通车站节点的特殊处理,有效路径判定规则的部分路径中还包含不符合换乘逻辑的路段。为了说明上述路段的特征,图 3-3 为某个换乘车站网络节点的示意图,虚线框线内的节点归属于同一个车站,其中节点 18、19、74、75 为换乘车站节点,其余为普通车站节点。

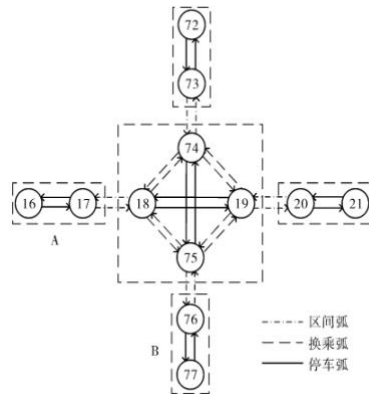


图 3-3 换乘节点网络节点示意图

假如某位乘客从车站 A 上车,在车站 B 下车,其间经过两段区间运行时间和一段换乘时间,可以看出,符合该乘客正确换乘逻辑的路径应该是 17—18—75—76。通过未改进的算法搜索到的相同起终点的路径 17—18—74—75—76、17—18—19—75—76 和 17—18—74—19—75—76。这些路径中均包含了不符合逻辑的停车时间和换乘时间,相应的 18—74—75、18—19—75、18—74—19—75 均可判定为冗余路段。

乘客最优路径选择模型:

对于乘客最优路径选择模型,是基于无起始站限制以及进出站时间限制的任意两车站最优路径的选择模型,即选择做短走行时间对应的路径模型,通过分析常见的网络路径优化经典智能搜索算法,选择 Dijkstra 算法模型,并引入路径阻抗,路径伸展系数建立改进的 Dijkstra 算法模型,如下面公式所示。ⁱ

$$\begin{aligned} & \min t_{ij} \\ & \text{s.t.} \begin{cases} C_k^{rs} \leq (1 + H_{rs}) \min_i C_i^{rs} \\ T_k \leq T_{max} \end{cases} \\ & H_{rs} = \frac{\max_{k'} c_{k'}^{rs}}{\min_i C_i^{rs}} \end{aligned}$$

式中: t_{ij} 是相邻两车站的走行时间,即等于第 j 个车站的发车时刻与第 i 个车站的到达时刻的差值, H_{rs} 路径伸展系数, C_k^{rs} 为路径 k 的阻抗, $\min_i C_i^{rs}$ 为最短路径阻抗, $\max_{k'} c_{k'}^{rs}$ 为最大合理路径阻抗。

考虑到换乘次数的增加也会降低乘客的舒适性和走行时间,因此在路径选择的判定规则中增加换乘次数的限制,即路径 k 的换乘次数小于等于网络最大换乘次数限制。

$$H_k \leq H_{max}$$

式中: H_k 为路径 k 的换乘次数; H_{max} 为网络最大换乘次数限制,本模型取 $H_{max} = 5$ 。

3.4 模型的求解

基于改进的 Dijkstra 乘客最优路径选择算法具体步骤如下:

Step1: 初始化。给相关变量赋初值,输入邻接矩阵,设定参数;

Step2: 利用 Dijkstra 算法计算在起始站节点 q 和终点站节点 z 之间满足限制条件的最短路径阻抗;

Step3: 从当前节点 i 出发,遍历与节点 i 相邻的节点,例如节点 j ,如果从起始站对应节点 q 出发沿着该路径的阻抗小于或等于设定阈值并且满足换乘限制条件,则令节点 j 为当前点,转下一步,否则转 Step 7;

Step4: 循环 $n-1$ 次扩展结点,在状态为未扩展的顶点中选择时间最小的顶点 u ,并将他的状态设为已扩展。对于每个与顶点 u 相邻的顶点 v ,执行时间更新操作 $\text{Relax}(u, v)$ 。在此操作中,如果 $\text{dist}[u] + w[u, v] < \text{dist}[v]$,就把 $\text{dist}[v]$ 的值更新成 $\text{dist}[u] + w[u, v]$ 的值。更新后由源顶点 s 到顶点 v 的最短路径上, v 的前一个顶点为 u 。

Step5: 判断节点 j 是否为终节点 z ,如果不是则转入 Step 2, 否则进行下一步;

Step6: 判断该路径是否包含冗余路段,如果是则忽略该路径并转入 Step 2, 否则进行下一步;

Step7: 搜索前驱节点序列并输出路径;

Step8: 退回上一层,若未退回到根节点则转入 Step 2。

3.5 结果的计算和分析

在图论和网络分析中,邻接矩阵是表示图中顶点之间相连关系的一种方式。使用邻接矩阵来分析问题意味着将图中的结构和节点间的相互作用量化,并且以此作为分析和解决问题的基础。为了求解城市轨道交通网络路径优化问题,根据调查的数据中乘客的进出站和时间

列车的到发时刻，所有在网络最优路径选择模型中是考虑时间问题的求解模型，因此网络 G 的为带权邻接矩阵为 n ($n=328$) 阶方阵 T ，边权为 t_{ij} ， t_{ij} 是相邻两车站的走行时间与滞留时间之和，通过列车运行图数据可求解不同线路在任意两车站的时间，由于城市轨道交通车辆均是追踪发车，因此表 3-1 给出二号线合计 18 个车站的邻接矩阵 T 。

表 3-1 二号线邻接矩阵

相邻车站	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
积水潭	0	158	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	129
西直门	166	0	83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
车公庄	0	87	0	83	0	0	0	0	0	0	0	0	0	0	0	0	0	0
阜成门	0	0	85	0	137	0	0	0	0	0	0	0	0	0	0	0	0	0
复兴门	0	0	0	137	0	116	0	0	0	0	0	0	0	0	0	0	0	0
长椿街	0	0	0	0	115	0	87	0	0	0	0	0	0	0	0	0	0	0
宣武门	0	0	0	0	0	85	0	86	0	0	0	0	0	0	0	0	0	0
和平门	0	0	0	0	0	0	82	0	97	0	0	0	0	0	0	0	0	0
前门	0	0	0	0	0	0	0	95	0	123	0	0	0	0	0	0	0	0
崇文门	0	0	0	0	0	0	0	0	123	0	112	0	0	0	0	0	0	0
北京站	0	0	0	0	0	0	0	0	0	112	0	101	0	0	0	0	0	0
建国门	0	0	0	0	0	0	0	0	0	0	99	0	127	0	0	0	0	0
朝阳门	0	0	0	0	0	0	0	0	0	0	0	123	0	88	0	0	0	0
东四十条	0	0	0	0	0	0	0	0	0	0	0	0	88	0	77	0	0	0
东直门	0	0	0	0	0	0	0	0	0	0	0	0	0	78	0	172	0	0
雍和宫	0	0	0	0	0	0	0	0	0	0	0	0	0	0	174	0	77	0
安定门	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	74	0	99
鼓楼大街	127	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0

（单位为秒（s））

3.6 乘客最优路径选择及结果验证

可以使用图论中的最短路径算法，如迪杰斯特拉(Dijkstra)算法或 A* 搜索算法等。以下是使用 Dijkstra 算法的基本步骤来实现从 A 站到 B 站的最短用时路线选择。首先，需要构建地铁线路图，可以使用邻接矩阵或邻接表来表示。地铁图可以视作一个加权有向图，其中节点表示地铁站，边表示连接两站的直接地铁线路，边的权重可以表示两站之间的距离或预计行驶时间。为验证该模型的准确性和合理性，选取任意两站计算耗时时间最短的最优路径，根据该结果与北京地铁官方网站路线方案查询结果相比较。ⁱⁱ

比较方案：选取从南锣鼓巷（对应于 8 号线，车站编号为 153 号），至安河桥北（对应于 4 号线，车站编号为 59 号），基于上述乘客最优路径选择模型和模型求解步骤计算结果如表 3-2 所示。

表 3-2 南锣鼓巷到安河桥北最优路线

车站编号	139	138	137/72	71	70	69	68	67	66	65	64	63	62	61	60	59
车站名称	南锣鼓巷	北海北	平安里	新街口	西直门	动物园	国家图书馆	魏公村	人民大学	海淀黄庄	中关村	北京大学东门	圆明园	西苑	北宫门	安河桥北
对应线路	6	6	6/4	4	4	4	4	4	4	4	4	4	4	4	4	4

该最优方案显示：从南锣鼓巷到安河桥北，中途共经过 16 站，换乘一次。其中 6 号线三站，4 号线 14 站，在平安里地铁站从 6 号线换乘 4 号线。地铁运行预计耗时 27.85000min，地铁在每站的停靠时间总计 9.33333min，平安里站换乘需要约 3.27min。总时长约 40.45333min。（注：模型未考虑巧合，运气，拥挤等因素，实际时间可能会长于该总时间）

接下来，比较高德地图路线方案查询结果，在查询栏搜索起终站站名，得到查询结果如图 3-5 所示，线路走行图见图 3-6 所示，通过该结果可看出线路走行方案与本文建立的模型求解结果完全吻合，显示耗时时间约 42min 见图 3-4，与求解出来的 40.45333min 基本一致。因此模型与实际情况具有较好的符合性。

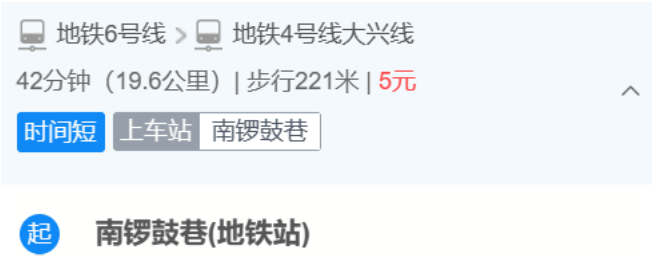


图 3-4 耗时时间



图 3-5 高德地图路线方案查询结果

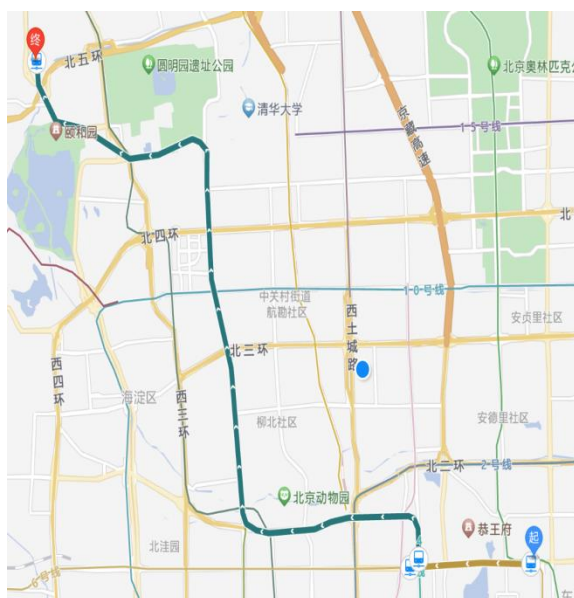


图 3-6 高德地图路线方案线路走行图

第四章 结论

随着城市化进程的加速，公共交通系统特别是地铁系统承载了越来越多的人流，如何准确了解和优化乘客的出行特征和路径选择，成为了城市轨道交通管理的重要议题。利用大数据和算法对乘客的出行特征进行分析，不仅可以为城市交通规划与管理提供科学依据，同时也能显著提升乘客的出行效率和舒适度。

在分析乘客的出行特征方面，通过搜集出行时间、距离与时长的数据，运用统计分析及机器学习方法能够揭示出行模式和乘客行为的内在规律，例如早晚高峰的出行时段特征、不同区域乘客出行距离的分布情况以及通勤与非通勤时间的时长差异。通过这些分析，可以更精确地进行交通流量预测，进行线路方案设计，以及制定更高效的运营计划。ⁱⁱⁱ

针对路径选择与优化，基于算法设计提供的原理性解决方案，可以进一步拓展为适用于实际运营的系统软件。该软件可以将车次时刻、站点间距离以及预期拥挤度等信息实时更新，并结合乘客出发地、目的地信息，推荐最优出行路径。

未来，期望通过算法改进，得到更加全面的智能优化算法。通过实时路况的动态反馈，智能路径选择算法能够为每个乘客提供个性化的最优路线。它不仅可以根据乘车时间、拥挤度和乘车费用等因素提供多元选择，还可在不可预测的交通事件发生时迅速调整行程规划，进而减少总体的出行时间和拥堵情况。

此外，这种智能路径选择算法的应用还能有效分散高峰时段的乘客流量，优化车辆运行时间表，提高轨道交通系统的整体运行效率。长远来看，随着算法的不断完善和数据的日益充实，智能算法将极大地促进智慧交通系统的构建，向乘客提供更为环保、经济、快捷的出行选择，为构建可持续发展的城市交通生态环境作出贡献。

引文：

ⁱ 马静, 刘江岳. 基于 Dijkstra 算法的地铁出行路线规划系统的设计与实现[J]. 电脑知

识与技术, 2022, 18 (13): 58-62.

ⁱⁱ 袁泽重. 基于拥挤指数的城市轨道交通乘客路径选择模型[D]. 北京交通大学, 2022.

ⁱⁱⁱ 刘剑锋, 孙福亮, 柏赞, 许娟. 城市轨道交通乘客路径选择模型及算法[J]. 交通运输系统工程与信息, 2009, 9 (02): 81-86.