

# 基于 APC 数据的公交客流分析

## 一、数据预处理

APC 数据集来自于对俄亥俄州立大学校园公交车的统计结果，数据集包含一天内同一条线路上 5 辆公交车（共 87 个 trip）的 ID、始发时间、每辆车在各个站点的（共 18 个站点）上下车人数以及每个站点的累计距离。经检查，无异常数据。

## 二、问题一的分析与求解——客流量时序图

### 2.1 问题假设

整条线路总长大约为 8 千米，该大学的公交车的平均时速（计算时含停靠点时间）经过综合考虑大概在  $14\text{mph}$ ，公交车在整条路线的行驶时间大约在  $21\text{min}$ 。由于本题以  $30\text{min}$  为单元绘制客流量时间序列图，因此假设公交车的客流量均全部属于各自发车时间所对应的  $30\text{min}$  之内。

### 2.2 客流量时序图及指标计算

将一天划分为  $m$  个时长为  $T$  的时间段，合并相邻时段的客流量  $n_i$ ，得到每个周期时间段的总客流量  $N_i$ ，公式如下：

$$N_i = n_i + n_{i+1} + \cdots + n_{i+h-1}, i \leq m - h + 1, h = \frac{60}{T}。 \quad (1)$$

若有  $N_{\max} = N_k = \max\{N_1, N_2, \cdots, N_{m-h+1}\}$ ，则高峰小时为： $[kT, (k+h-1)T]$ 。

对于本题而言，参数有： $T = 30\text{min}, h = 2, m = 48$ 。于是，对原数据进行整理后（整理结果见 q1.xlsx），绘制的客流量时间序图如下（代码见 q1.m）。

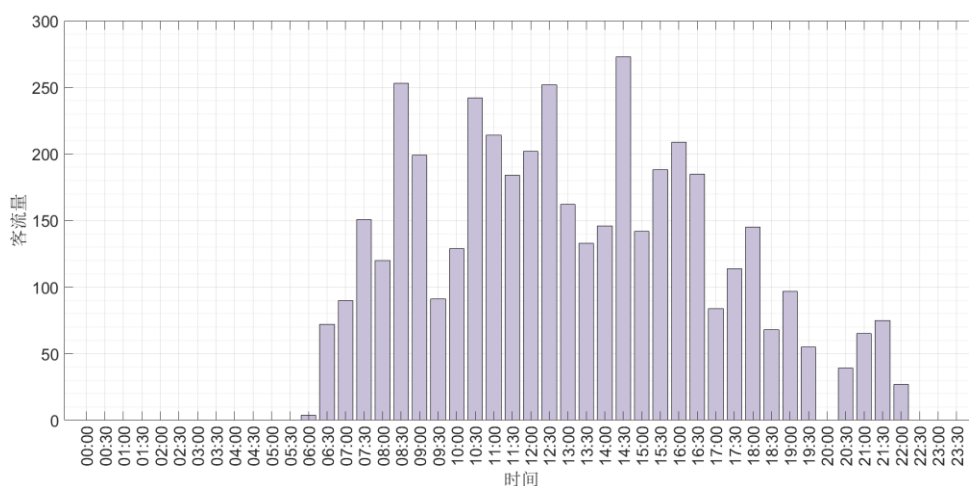


图 1.客流量时间序列图

- 1、经过合并相邻两个时间段的客流量可知： $N_{\max} = N_{22} = 456$  人/h，所以其中一个高峰小时大概在 [10:30, 11:30] 之间。
- 2、假设学生乘坐公交全为刷卡模式，扩样系数  $\eta=1$ 。则高峰小时客流量  $PFV_{ph} = N_{\max}/\eta = 456$  人/h。
- 3、24 小时客流量  $PFV_{24h} = \sum_{i=1}^{48} n_i = 4410$  人，高峰小时乘车率  $R_{ph} = \frac{PFV_{ph}}{PFV_{24h}} \approx 0.103$ 。

### 三、问题二的分析与求解——高峰小时断面客流

#### 3.1 基于客流 OD 矩阵计算断面客流

##### 3.1.1 整条线路客流 OD 矩阵

由问题一可知，取高峰小时段为 [10:30, 11:30]。合并位于高峰时间段的每一站点的客流，得到每个站点的上下车人数如下表所示：

表 1.各站点上下车人数

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
上车(人)	63	0	9	13	25	7	0	11	36	52	12	95	46	53	8	17	9	0
下车(人)	0	2	5	7	15	13	3	14	16	15	4	22	88	58	17	21	85	71

利用迭代比例拟合算法，求出该线路客流 OD 矩阵如下（代码见 q2\_ODmat.m， $\varepsilon = 0.01$ ）：

表 2. 高峰时段线路客流 OD 矩阵

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	2	5	6	11	6	1	6	6	3	1	3	6	3	1	1	2	2
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	1	2	1	0	1	1	1	0	0	1	0	0	0	0	0
4	0	0	0	0	3	2	0	2	2	1	0	1	1	1	0	0	1	0
5	0	0	0	0	0	4	1	4	4	2	0	2	4	2	0	0	1	1
6	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	3	1	0	1	2	1	0	0	1	1
9	0	0	0	0	0	0	0	0	0	6	1	5	10	5	1	1	4	3
10	0	0	0	0	0	0	0	0	0	0	2	8	18	8	2	2	7	5
11	0	0	0	0	0	0	0	0	0	0	0	2	4	2	0	0	2	1

12	0	0	0	0	0	0	0	0	0	0	0	0	0	40	19	4	4	16	12
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	3	4	13	10
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	7	23	17
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	3
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	7
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

由表可知，在高峰客流小时期间内，2号、7号车站均无人上车（不含终点站）。

### 3.1.2 断面客流计算及分析

通过线路客流 OD 矩阵，可以计算两个相邻站点之间的断面客流。令相邻两站点为 $n,n+1$ ，计算该对站点间的断面客流公式如下：

$$P(n\sim n+1)=\sum_{j=n+1}^N OD_{(n,j)}+\sum_{i=0}^{n-1}\sum_{j=n+1}^N OD_{(i,j)},\quad OD_{(0,j)}=0。$$

(2)

其中： $N$ ——车站总数； $OD_{(i,j)}$ ——线路客流OD矩阵中第*i*行第*j*列的值。

通过 MATLAB 计算的相邻两车站间的段面客流四舍五入结果如下（代码见附件 q2\_pf.m）：

表 3.高峰小时断面客流

	1~2	2~3	3~4	4~5	5~6	6~7	7~8	8~9	9~10
断面客流（人/h）	63	61	65	71	81	75	72	69	89

表 4.高峰小时断面客流（续）

	10~11	11~12	12~13	13~14	14~15	15~16	16~17	17~18
断面客流（人/h）	126	134	207	165	160	151	147	71

可视化结果如下：

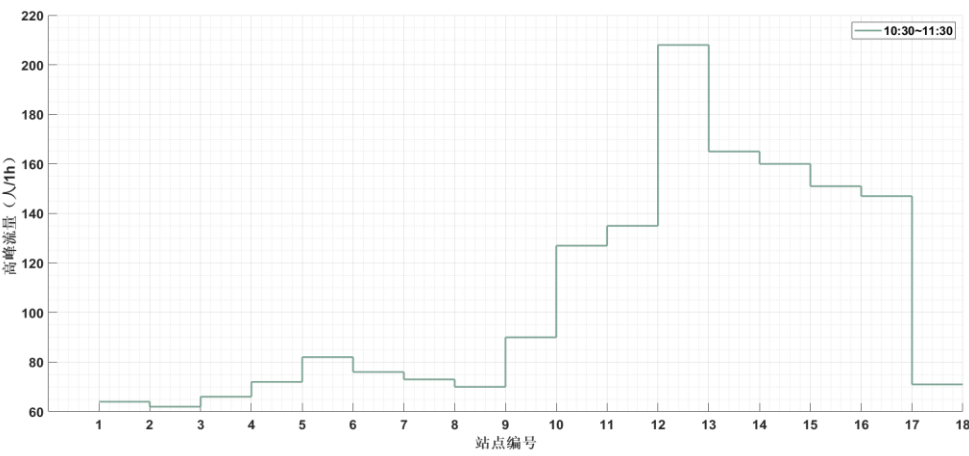


图 2.高峰小时断面客流

上图可知,高峰小时最大断面客流的位置在12~13 号站点之间,其断面客流量大小为208人。

### 3.2 基于各站点上下车人数计算断面客流

断面客流可用断面载荷量来表示,断面载荷量是指在单位时间内,通过线路某一地点(区间)的客流量。其计算公式可表示为:

$$P_i = \sum_{k=1}^i (P_k^B - P_k^A) \quad (3)$$

其中:  $P_i$ ——断面*i*车内载客量;  $P_k^B$ ——断面*i*上游站点*k*上客人数;  $P_k^A$ ——断面*i*上游站点*k*下客人数。

该计算公式可以简化为<sup>[1]</sup>:

$$P_{i+1} = P_i - P_{id} + P_{iu}, i = 0, 1, 2, \dots, N-1。 \quad (4)$$

其中:  $P_i$ ——某单一线路在第*i*个断面的客流量。

$P_{id}, P_{iu}$ ——某单一线路在第*i*个站的下车人数。

通过该方法计算出的结果与 3.1 节所述内容计算出的结果相同(见 q2\_xlsx), 此处不再赘述。

### 3.3 一些指标的计算及实际分析

#### 3.3.1 断面空间不均衡系数

断面空间不均衡系数表示最大客流断面客流量 ( $\overline{SPFV}$ ) 与断面客流量的平均值 ( $\overline{SPFV}$ ) 之比。上述可知最大断面客流量  $SPFV_{\max} = SPFV_{12,13} = 207$  人/h。平均断面客流量计算如下:

$$\overline{SPFV} = \frac{\sum_{i=1}^{N-1} SPFV_{i,(i+1)}}{N-1} \quad (5)$$

计算可知  $\overline{SPFV} = \frac{1807}{17} \approx 106$  人/h。

因此断面空间不均衡系数  $f_s = \frac{SPFV_{\max}}{\overline{SPFV}} = \frac{207}{106} \approx 2.0$ 。

#### 3.3.2 断面时间不均衡系数

断面时间不均衡系数用于衡量该公交线路的流量在特定时间段内的分布情况, 进而评估线路的负载情况。其计算公式如下:

$$\alpha_1 = \frac{p_{\max}}{\sum_{t=1}^H p_t / H} \quad (6)$$

其中:  $p_{\max}$ ——单向高峰小时最大断面客流量。  $p_t$ ——单向分时最大断面客流量。  $H$ ——全

日营业小时数。

以一个小时为时间单元，计算整点时间段的最大断面客流如图：（代码见附件 q2\_fig.m）：

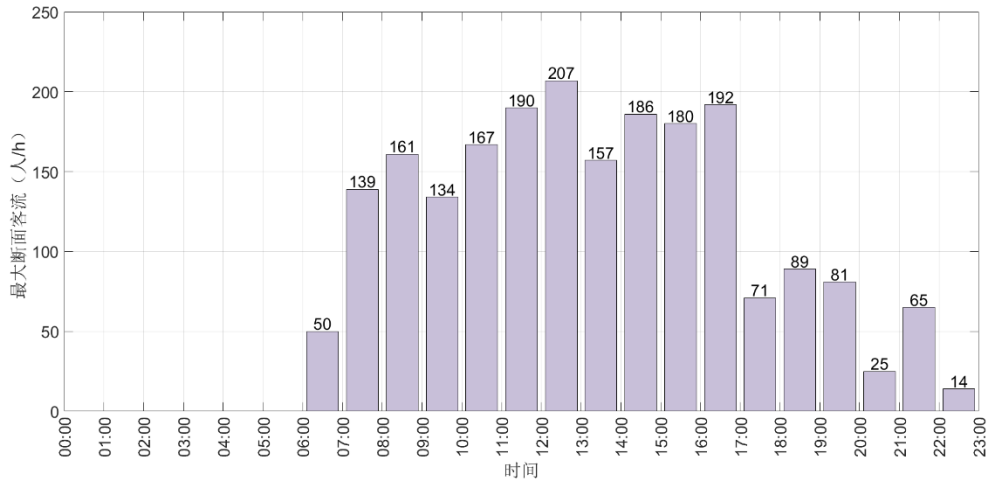


图 3.断面客流时间序列柱状图

由图可得知，式（6）中的一些参数的值为： $p_{\max} = 207 \text{ 人/h}$ ， $H = 17$ 。且通过计算可知

$$\sum_{t=1}^H p_t / H = 124 \text{ 人/h}。因此断面时间不均衡系数 \alpha_1 \approx 1.7。$$

### 3.3.3 分析

由断面时空不均衡系数可知：在该校园内的某些时间段（如上学、下课、下班等高峰时段），客流量较大，公交车辆可能会面临较大的客流压力。在非高峰时段，客流量相对较少，公交车的利用率可能较低。

## 四、问题三的分析与求解——行车间隔

为确定一个高峰小时 [10:30, 11:30] 期间内合理的行车间隔，本文首先考虑到大学校园内学生或教职工的乘车体验。即在高峰时期，每辆车不存在满载（60 人）的情况。本文简单定义车辆饱和度  $D$  来描述这种关系：

$$D = \frac{PF_{\max}}{VO \cdot n} \quad (7)$$

其中： $PF_{\max}$ ——断面客流量。 $VO$ ——公交车最大容量。 $n$ ——车次数量。

假设经调研分析，该高峰时期该校园公交的饱和度  $D = 0.85$ 。

由第一问可知，高峰小时断面客流最大值  $PF_{\max} = 207 \text{ 人/h}$ ，所以可计算车次数量，

$$n = \frac{PF_{\max}}{D \cdot VO} \approx 4.05，因此合理的车辆安排应为 5 辆。$$

定义行车间隔  $IVL$ ，所以粗略的行车间隔可以是  $IVL = \frac{60}{n} \approx 14.8\text{min}$ 。

为了保证更好的服务水平，并考虑到实际中的随机波动（例如乘客上下车时间、交通状况等），本文对行车间隔稍微增加一点余量（假设为10%），以保证乘客在高峰期的乘车舒适性和可靠性。

因此，最合理的行车间隔  $IVL_{best} = 1.1 \times IVL = 15.3\text{min}$ 。

## 五、问题四的分析与求解——最大断面客流时序图

### 5.1 最大断面客流即运输能力时序图绘制

接（3.3.2）讨论的内容，把图 3 转化为折线图。叠加数据集中每个小时段的发车数量，计算每个小时段的计划运输能力（假设每辆公交最大承载 60 人）。得到下图（代码见附件 q4.m）：

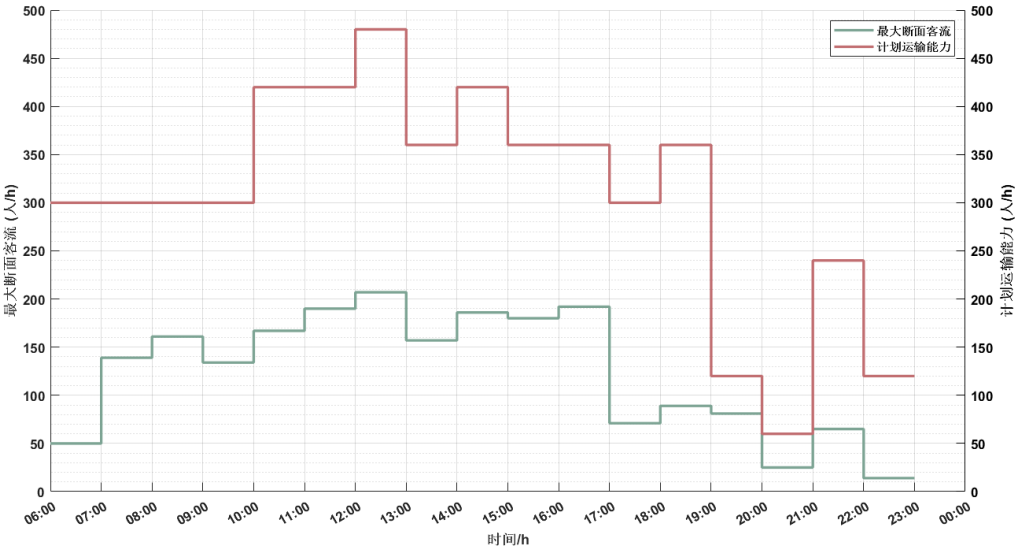


图 4. 最大断面客流即运输能力时序图

### 5.2 整体趋势分析

计划运输能力（红色线）与最大断面客流需求（绿色线）整体高差较大，每个时段都呈现出“供大于求”的关系。6:00~19:00，21:00~23:00 的计划运输能力均远大于最大断面客流。19:00~21:00 的运输能力与客流需求匹配情况基本吻合。

### 5.3 各时段分析

#### 5.3.1 时段 7:00-9:00

绿色线显示早高峰期间客流逐步上升，但整体处于适中水平，约在120~170人/h之间。

红色线的计划运输能力稳定保持在300人/h，明显高于实际客流，说明早高峰期间有充足的运力应对乘客需求，可能存在一定的运力浪费。

5.3.2 时段 9:00-17:00

在此时间段，最大断面客流在一定小范围内波动，最大客流上升至约200人/h，但是计划运输能力也有相应上升，一直远大于最大断面客流，约在380人/h。也存在运力浪费的情况。

5.3.3 时段 17:00-19:00

最大断面客流急剧下降到80人/h左右，并基本稳定在一个较低的值附近，虽然有一定波动，但仍远低于计划运输能力。也存在运力浪费的情况。

5.3.4 时段 19:00-21:00

该时段内，最大断面客流在100人/h以下，而计划运输能力也几乎保持在同一水平，略高于最大断面客流。运力与客流匹配良好。

5.3.5 时段 21:00-23:00

在21:00 之后，客流迅速下降至50人/h附近，而计划运输能力却上升至一个较高水平。这一时段的运力明显超出实际需求，可能导致运力浪费。

六、 问题五的分析与求解——指标计算

6.1 客运周转量

客运周转量表示乘客乘坐里程的总和，其计算公式如下：

$$PPK = \sum l_{ij} \cdot q_{ij}, i,j \in [1, 18]. \tag{8}$$

其中： $l_{ij}$ ——站点*i*和站点*j*之间的线路长度。 $q_{ij}$ ——站点*i*和站点*j*之间的客流量。

已知每个站点的站间累计距离如下表所示：

表 5.站点累计距离

	1	2	3	4	5	6	7	8	9
累计距离 (km)	0.000	0.513	1.117	1.625	2.358	2.793	3.304	3.550	3.918

表 6.站点累计距离 (续)

	10	11	12	13	14	15	16	17	18
累计距离 (km)	4.123	4.447	4.654	5.228	5.566	5.983	6.286	6.797	7.836

通过 MATLAB 编程计算（代码见 q5.m），得到 $PPK = 8433 \text{ 人} \cdot \text{km}$ 。

6.2 平均运距

平均运距是指乘客每次乘车的平均距离，它的计算公式如下：

$$ADC = \frac{PPK}{Q} \tag{9}$$

其中： $Q$ ——客运总量。

该数据集中的客运总量即为所有站点*i*和*j*的客流量之和，如下公式：

$$Q = \sum_{i,j}^{18} q_{ij}, i, j \in [1, 18] \text{ 且 } i < j. \quad (10)$$

所以计算得 $Q = 4410$ 人。进而有平均运距 $ADC = 1.91km$ 。

### 6.3 运力利用率

运力利用率是指客运周转量与客运里程之比，其计算公式如下：

$$CCUR = \frac{PPK}{SK} \times 100\% \quad (11)$$

其中： $SK$ ——客运里程，即运营车的客位数与载客里程的乘积。

假设每辆公交车的最大载客容量为 $n = 60$ 人，一天之内共有的 trip 数 $m = 87$ ，每一个 trip 的运行载客里程 $L = 7.836km$  因此对于  $SK$  的计算公式如下：

$$SK = m \cdot n \cdot L \quad (12)$$

计算得 $SK = 40903.92$ 人· $km$ ，进而有运力利用率 $CCUR = 20.6\%$ 。

## 七、问题六的分析与求解——IPF 求 OD 矩阵

高峰小时段[10:30, 11:30]，由 3.1.1 算出来的结果可知，在高峰时段线路客流 OD 矩阵请见

3.1.1（表 2. 高峰时段线路客流 OD 矩阵）。由表 2 可知，客流量前 5 的 OD 对如下：

表 7.客流量前 5 的 OD 对

	OD 值	OD 对
1	40	12~13
2	23	14~17
3	19	12~14
4	18	10~13
5	17	14~18

## 八、问题七的分析与求解——客流特征及建议

### 8.1 客流特征

#### 8.1.1 高峰客流特征与时间分布

通过客流量时序图（图 1），发现在7:00—9:00和17:00—19:00这两个时段，公交客流量



迅速上升，说明这些时间段与校园内的上下课及上下班时间高度吻合。这意味着早高峰时：学生和教职工集中前往学校和办公室，反映出典型的单向客流特征；晚高峰时：可能与下午的课程结束或下班时间重叠，出现大量人员从学校或办公区域返回居住地的需求。这种高峰时段的客流呈现双向分布（部分乘客前往晚间活动场所，另一部分乘客则返回家）。

#### 8.1.2 断面客流的空间分布不均衡

由 3.1.2（图 2.高峰小时断面客流）可知站点之间的断面客流存在明显的差异。具体表现为：

- 1、热闹区域：12~13 号站点集中了大量上下车乘客，断面客流量尤其大。因为这些站点靠近公园、艺术中心、博物馆、图书馆、实验室、学院楼及学生会办事处等工作学习地点。
- 2、空闲区间：部分区间的客流量较少，尤其是远离上述关键站点的区域，公交车在这些区段的运力利用率较低，造成了资源浪费。

#### 8.1.3 运力过剩与浪费现象

由 5.1（图 4.最大断面客流即运输能力时序图）可知，无论是高峰时段还是非高峰时段的运力都远超过需求，只是在19:00~21:00时段，这种极度不均衡的关系才得以改善。高峰时段之外的乘客数量相对较少，导致大量公交车辆在非高峰时段处于低负荷或空载状态，未能有效利用运力资源。尤其是在夜间时段，客运需求下降的趋势更加明显，进一步拉低了全天的运力利用率。

这种情况显示，公交系统的运行效率存在问题，运力分配未能很好地与客流需求相匹配，造成运力利用率较低的情况，问题五具体的指标计算结果也能很好地反映这个问题。

说明公交系统的运力配置在全天范围内有待优化，运力利用率可以通过灵活调整班次和车次来进一步提高。

#### 8.1.4 客流 OD 分布特征

问题六通过 OD 矩阵估计，揭示了高峰时段主要的客流对（OD 对），发现前五的 OD 对占据了较大的客流比例。这表明：主要的乘客流动集中在少数关键的站点之间，客流流动呈现出强烈的集中趋势。大多数乘客的出行模式较为固定，更多的是往返于特定的起点和终点之间（例如宿舍区与学习区、办公区、休闲区之间）。

### 8.2 相关建议

#### 8.2.1 动态调整车间隔和班次

根据不同时段的客流量变化，灵活调整公交车的发车间隔和班次。例如：在早晚高峰时段，适当减少发车频率，适当扩大行车间隔，以应对高峰时期运力空闲的情况（视情况而定，也有可能增加车次）；在非高峰时段，更应该减少发车频率，尤其是在夜间时段，根据实际需求适当减少运力，以避免不必要的资源浪费。

### 8.2.2 针对重点站更新运力

在高峰时段或某些特定站点（如 12 号、13 号站点）的客流量特别集中，建议在这些站点适当增加运力调度：适当减少该区段的公交车次，或在该区段内安排短途公交，以分散客流压力；也可以考虑采用“大站快车”等策略，绕过部分乘客较少的站点，快速运送乘客到达集中客流站点，提高运送效率。

### 8.2.3 提升运力利用率

为了减少非高峰时段的运力浪费，可以采取以下措施：采用小型公交车或者新能源公交车进行运营，以减少空载时的能源消耗和运力浪费；在一些需求不稳定的时间段引入“按需公交”服务，即乘客提前预约公交车，使车辆的调度更加灵活、精准，降低空驶率。

### 8.2.4 改善关键 OD 对服务

针对问题六中确定的主要 OD 对，建议重点优化这条线路上的服务。例如：在这些 OD 对的区间内，增加公交车次，或者提供直达快线服务，以缩短乘客的出行时间，提升整体的服务效率；考虑引入快速公交系统（BRT），在关键 OD 对上提供高速、便捷的运输服务，进一步提升运力利用率和乘客满意度。

## 8.3 分时出行宣传和激励机制

通过校园内的宣传或激励机制，鼓励乘客在非高峰时段出行，减缓高峰期的客流压力。例如：通过优惠券、积分等方式，鼓励学生或教职工在高峰期以外的时段搭乘公交车，平衡各时段的客流量分布；与校园管理部门协作，推行灵活上下课或上下班的制度，以减少高峰期的集中出行需求。

## 九、参考文献

- [1]. 黄琛. 基于公交运营数据的广州公交实时断面客流研究与分析[D]. 广东:广东工业大学,2023.