

# 基于时段客流特征聚类的地铁运营时段划分<sup>①</sup>



陈东洋<sup>1,2</sup>, 陈德旺<sup>1,2</sup>, 江世雄<sup>1,2</sup>, 徐 宁<sup>1,2</sup>

<sup>1</sup>(福州大学 数学与计算机科学学院, 福州 350108)

<sup>2</sup>(福州大学 智慧地铁福建省高校重点实验室, 福州 350108)

通讯作者: 陈东洋, E-mail: 2510249269@qq.com

**摘 要:** 准确合理的运营时段划分方案是制定地铁列车开行方案的前提和基础, 也是提高地铁运营效率的重要方式. 为了合理划分地铁运营时段, 本文构建时段客流特征向量以划分地铁运营时段. 以 10 min 为时间间隔对全日运营时段进行分段, 并根据时段内的客流变化特点构建各时段的特征向量. 并以此为基础采用 K-means 算法进行聚类, 同时以肘部法则、轮廓系数等聚类评估指标对结果进行评价, 以确定最优聚类数, 进而得到最优的运营时段划分方案. 最后以福州地铁一号线为例, 给出了该路线的运营时段划分方案, 验证了该方法的可行性.

**关键词:** 地铁; 客流; 运营时段划分; 聚类; K-means 算法

引用格式: 陈东洋, 陈德旺, 江世雄, 徐宁. 基于时段客流特征聚类的地铁运营时段划分. 计算机系统应用, 2021, 30(3): 256-261. <http://www.c-s-a.org.cn/1003-3254/7799.html>

## Division of Metro Operation Periods Based on Feature Clustering of Passenger Flow

CHEN Dong-Yang<sup>1,2</sup>, CHEN De-Wang<sup>1,2</sup>, JIANG Shi-Xiong<sup>1,2</sup>, XU Ning<sup>1,2</sup>

<sup>1</sup>(College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China)

<sup>2</sup>(Key Laboratory of Intelligent Metro of Universities in Fujian Province, Fuzhou University, Fuzhou 350108, China)

**Abstract:** An accurate and reasonable division scheme of operation periods is the premise and foundation for formulating metro train operation plans, and it is also an important way to improve metro operation efficiency. The feature vectors of passenger flow are constructed to divide metro operation periods. At an interval of 10 min, the operation period in a day is segmented, and the feature vectors of all the periods are constructed according to the characteristics of passenger flow changes in their corresponding periods. The K-means algorithm is used for clustering, and the results are evaluated by cluster evaluation indicators such as elbow method and silhouette coefficient to determine the optimal number of clusters, obtaining the optimal division scheme of operational periods. Finally, the division scheme of operation periods of Line 1 of Fuzhou Metro is given as an example, which verifies the feasibility of this method.

**Key words:** metro; passenger flow; division of operation periods; clustering; K-means algorithm

城市轨道交通的列车开行方案是运营公司日常组织运营调度的前提与基础<sup>[1]</sup>, 而合适的运营时段划分方案也是列车开行方案调整的基础. 对于运营公司而言, 分时段等间隔的发车方式同时具有灵活、易操作等特

点, 因此是运营公司广泛采用的一种发车模式<sup>[2]</sup>. 该模式通常根据客流规律将一天划分为高峰时期、正常时期、低谷时期等若干个运营时段, 并在各运营时段内采用合适的发车间隔. 但在实际的工作中, 地铁运营时

① 基金项目: 国家自然科学基金面上项目 (61976055); 智慧地铁福建省高校重点实验室建设项目 (53001703, 50013203)

Foundation item: General Program of National Natural Science Foundation of China (61976055); Construction Project of Key Laboratory of Intelligent Metro of Universities in Fujian Province (53001703, 50013203)

收稿时间: 2020-07-02; 修改时间: 2020-07-30; 采用时间: 2020-08-20; csa 在线出版时间: 2021-03-03

段的划分通常是技术人员根据经验等因素,主观地进行划分,难以符合实际的客流需求.因此国内外有许多学者针对此问题进行了研究.

对于时段划分问题的研究主要是针对公交、路口等问题,对于地铁的研究较少.在公交运营时段划分方面, Bulut 等<sup>[3]</sup>通过站间流量得到公交车使用率,并以此为特征变量采用聚类算法划分运营时段;文献[4-8]从车辆行程时间角度出发,以车辆单程时间、站间运行停留时间等作为特征变量,并采用有序样本聚类、K-means 等聚类算法划分运营时段;靳文舟等<sup>[9]</sup>以车队运营时间最小为优化目标构建模型,并借助遗传算法对模型进行求解,以寻找最优的运营时段划分方案.在交通路口时段划分方面,文献[10-13]以道路路口不同方向的流量等作为时段内道路的特征向量,并采用聚类算法划分交通时段;杜长海等<sup>[14]</sup>建立了一种结合混合蛙跳算法优点与模糊C均值算法优点的模型以解决交通运营时段的划分问题;杨立才等<sup>[15]</sup>借鉴生物免疫网络理论,建立了以人工免疫算法为基础的数据聚类分析算法模型,以解决城市交通运营时段的自动划分问题.在地铁时段划分方面,王文宪等<sup>[16]</sup>建立采用近邻传播聚类算法的地铁运营时段划分模型,模型中以目标线路各车站单位时间内的进站客流为描述变量;曾小旭等<sup>[17]</sup>根据实际客流数据建立了地铁单向OD概率矩阵的时序模型,并以此为基础采用最优分割法求解模型,给出时段划分方案.

以上研究为地铁运营时段划分提供了参考.但是可以发现,上述研究主要是将研究对象一定时间内各组成部分的客流量或行程时间作为特征变量来描述该时段的特点,如站点间的客流量、行程时间等.但是该方法容易受到一些特殊站点等因素的影响,且其直接采用客流量等作为描述变量,因此难以体现线路总体客流的变化特性.而且在求解划分方案时忽略了不同时段之间的相似性,降低了时段划分方案的灵活性.基于以上分析,本文以地铁线路总体客流的变化特点为基础,首先构建运营日的时段客流样本数据集.再根据各时段客流片段的客流变化特点,构建每个时段客流片的特征向量,并以此为基础采用K-means算法进行聚类,最后借助聚类评价指标确定最佳聚类数.本研究能根据实际客流需求,实现地铁运营时段的准确、高效划分,为地铁列车时刻表的优化设计提供了良好的基础.

## 1 时段客流特征向量构建

城市轨道交通自动售检票系统(AFC)中存储着大量的乘客出行交易信息,其中包括乘客的进出站点编号、进出站日期时间、交易卡号等.对这些信息进行挖掘和分析,可以进一步提取出地铁线路的客流变化特征等信息.本文以福州地铁一号线2019年5月13日-5月17日以及5月20日-5月24日共10个工作日的AFC客流数据为基础,研究福州地铁一号线工作日的各运营时段的划分方案.

### 1.1 单日客流序列构建

将目标线路在运营日 $k$ 中统计间隔 $i$ 内到达目标线路站点候车的乘客人数记作 $s_i^k$  ( $s_i^k \geq 0$ ),则运营日 $k$ 的客流序列可记作 $S_k = \{s_1^k, s_2^k, \dots, s_n^k\}$ ,且满足 $Q_k = \sum_{i=1}^n s_i^k$ ,其中 $Q_k$ 为目标线路在运营日 $k$ 的客流总量, $n$ 为运营日 $k$ 的统计间隔个数.本文中统计间隔选取1 min为单位,同时考虑福州地铁的实际运营情况,本文的运营日统计时段为06:00-23:59.因此本文中 $n$ 的取值为1080.

为后续运营时段划分方便,同时考虑到目标线路多个工作日的客流序列的相似性,因此采用均值处理方式,构建工作日单日客流序列 $S = \{s_1, s_2, \dots, s_i, \dots, s_n\}$ ,其中:

$$s_i = \frac{1}{K} \sum_{k=1}^K s_i^k \quad (1)$$

式(1)中, $K$ 为运营日的天数.

### 1.2 时段客流样本构建

运营日的客流序列反映了目标线路一天中乘客乘坐地铁的客流变化规律.选取合适的时间间隔 $d$ ,将运营日的客流序列划分为若干个长度相同的序列片段,则运营日的所有序列片段即构成了运营日的时段客流样本集 $D = \{D_1, D_2, \dots, D_m\}$ ,其中 $m$ 为样本集大小,且 $m = n/d$ .则有:

$$\begin{cases} D_1 = \{s_1, s_2, \dots, s_d\} \\ D_2 = \{s_{d+1}, s_{d+2}, \dots, s_{2d}\} \\ \vdots \\ D_m = \{s_{(m-1)d+1}, s_{(m-1)d+2}, \dots, s_{md}\} \end{cases} \quad (2)$$

本文选取10 min为时间间隔单位,即 $d=10$  min.因此样本集的大小 $m=108$ .

### 1.3 时段客流特征向量

样本集 $D$ 的每一个样本都是工作日单日客流序列

的某一个时段的客流变化序列, 每一个样本的客流序列都有其变化特点. 不同的时段客流样本可能具有相似的变化特点, 也可能具有不同的变化特点. 但是从客流数值上难以看出不同时段客流样本的异同, 因此本文构建时段客流特征向量来描述一定时间段内的客流变化特征.

本文采用时段客流的极大值、最小值、平均值、标准差、客流上升时间比、客流下降时间比共 6 个特征值来描述一定时间段内的客流变化特征.

对一段长度为  $N$  的时段客流序列  $c = \{c_1, c_2, \dots, c_N\}$ , 则上述特征值的计算公式如下:

时段客流序列  $c$  的极大值  $c_{\max}$  计算公式如式 (3):

$$c_{\max} = \max \{c_1, c_2, \dots, c_N\} \quad (3)$$

时段客流序列  $c$  的最小值  $c_{\min}$  计算公式如式 (4):

$$c_{\min} = \min \{c_1, c_2, \dots, c_N\} \quad (4)$$

时段客流序列  $c$  的平均值  $c_{\text{avg}}$  计算公式如式 (5):

$$c_{\text{avg}} = \frac{1}{N} \sum_{i=1}^N c_i \quad (5)$$

时段客流序列  $c$  的标准差  $c_{\text{std}}$  计算公式如式 (6):

$$c_{\text{std}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (c_i - c_{\text{avg}})^2} \quad (6)$$

时段客流序列  $c$  的客流上升时间比  $c_{\text{up}}$  计算公式如式 (7):

$$c_{\text{up}} = \frac{\text{片段中客流较前一个点上升的点的个数}}{N} \quad (7)$$

时段客流序列  $c$  的客流下降时间比  $c_{\text{down}}$  计算公式如式 (8):

$$c_{\text{down}} = \frac{\text{片段中客流较前一个点下降的点的个数}}{N} \quad (8)$$

这 6 个特征值中  $c_{\max}$  反映时段内每分钟进站客流的上限;  $c_{\min}$  反映时段内每分钟进站客流的下限;  $c_{\text{avg}}$  反映时段内每分钟进站客流的平均水平;  $c_{\text{std}}$  反映时段内按分钟统计的进站客流变化的波动幅度;  $c_{\text{up}}$  体现时段内客流的上升趋势;  $c_{\text{down}}$  体现时段内客流的下降趋势. 这 6 个特征值共同组成该客流片段的特征向量, 用以描述该客流片段的客流变化特征.

根据 1.1 节和 1.2 节生成的时段客流样本集, 本文采用上述方法构建时段客流样本集  $D$  对应的时段客流特征向量样本集  $C$ . 设样本  $D_i$  对应的特征向量为  $C_i$ , 且  $c_{\max}^i$ 、 $c_{\min}^i$ 、 $c_{\text{avg}}^i$ 、 $c_{\text{std}}^i$ 、 $c_{\text{up}}^i$ 、 $c_{\text{down}}^i$  分别对应样本  $D_i$  的

极大值、最小值、平均值、标准差、客流上升时间比、客流下降时间比, 则样本  $D_i$  的特征向量  $C_i$  表示如式 (9):

$$C_i = (c_{\max}^i, c_{\min}^i, c_{\text{avg}}^i, c_{\text{std}}^i, c_{\text{up}}^i, c_{\text{down}}^i) \quad (9)$$

## 2 聚类算法模型

数据挖掘是一项通过分析大量历史数据进而挖掘提取隐藏的、存在潜在应用价值的信息的技术手段. 而聚类分析作为数据挖掘的一种核心技术, 具有巨大的潜在应用价值, 是当下社会最具有应用前景的数据分析方式之一. 而 K-means 聚类算法是一种经典的解决数据聚类问题的算法, 其快速、简单、高效, 并且适合应用于大数据集上. 本文基于 K-means 算法进行地铁运营时段的划分.

### 2.1 K-means 聚类算法原理

K-means 是一种常用的聚类算法, 算法的核心思想是根据各样本之间的距离将最靠近的样本进行聚类, 并在多次迭代中不断更新各聚类中心的值, 直至达到算法设定的收敛条件, 得到最好的聚类结果.

算法首先从样本集中随机选取  $k$  个对象作为聚类的初始中心点, 然后计算样本集中所有剩余样本到各聚类中心点的距离, 并以此将剩余的各个样本划分到与其相距最近的类别中, 形成新的类别簇, 再重新计算各类别的新的中心值, 重复上述过程, 直至达到算法提前设定的停止条件. K-means 算法终止迭代的条件有迭代次数、各类簇的中心变化率、最小平方误差等. 本文采用最小平方误差作为 K-means 算法的迭代终止条件. 则算法的目标函数定义如下:

$$E = \sum_{j=1}^k \sum_{i=1}^{N_j} (x_i - \mu_j)^2 \quad (10)$$

式 (10) 中,  $E$  表示样本集中所有样本到各自类别中心的距离平方之和, 当某一次迭代之后  $E$  的值与上一次迭代后的值相比不发生显著变化时, 算法即收敛. 式中  $k$  为划分的类别数,  $N_j$  为类别  $j$  中包含的样本个数,  $x_i$  为类别  $j$  中的样本  $i$ ,  $\mu_j$  为类别  $j$  的中心.

本文采用欧式距离作为样本间距离的计算公式, 设两个样本向量分别为  $u = (u_1, u_2, \dots, u_n)$ ,  $v = (v_1, v_2, \dots, v_n)$ , 则  $u, v$  的欧式距离定义如下:

$$\text{dist}(u, v) = \sqrt{\sum_{i=1}^n (u_i - v_i)^2} \quad (11)$$

同时采用计算均值替代的方式更新每个样本类别的中心点 $\mu_j$ 的值,公式如式(12).

$$\mu_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i \quad (12)$$

基于上述选择,本文的K-means聚类算法的基本步骤如下:

步骤1. 在样本集 $C$ 中任意选取 $k$ 个样本对象 $\mu_1, \mu_2, \dots, \mu_k$ 作为每个类别的初始中心;

步骤2. 采用式(11)计算样本集中剩余的所有对象 $C_i$ 和 $\mu_1, \mu_2, \dots, \mu_k$ 的距离,并以此将其划分到与其距离最近的类别中心所属的类别中,即:

$$label_i = \arg \min_{1 \leq j \leq k} \|dist(C_i, \mu_j)\| \quad (13)$$

步骤3. 根据式(12)更新每个类别的中心 $\mu_1, \mu_2, \dots, \mu_k$ ;

步骤4. 根据式(10)计算该次迭代后的目标函数 $E$ ,如果该次迭代后的目标函数值与上次迭代后的目标函数值相差小于提前设定的阈值,则K-means算法终止,否则,重复步骤2~步骤4.

由算法流程可知,K-means算法有一个十分重要的参数,即聚类数目 $k$ 值的选取.不同的 $k$ 值所产生的聚类结果有很大差别,因此对于K-means算法,确定一个合适的 $k$ 值就十分重要.本文采用肘部法则以及轮廓系数两个评估指标确定最佳聚类数 $k$ .

## 2.2 肘部法则

肘部法则(elbow method)根据聚类的误差平方和(Sum of the Squared Errors, SSE)的变化程度来确定最佳的聚类数,SSE的计算公式如下:

$$SSE = \sum_{i=1}^k \sum_{p \in G_i} |p - m_i|^2 \quad (14)$$

式(14)中, $G_i$ 是第 $i$ 个簇, $p$ 是 $G_i$ 中的样本点, $m_i$ 是簇 $G_i$ 的中心,SSE是所有样本的聚类误差结果,代表了聚类结果的好坏程度.由SSE定义可知,SSE即为 $k$ 值确定情况下该 $k$ 值对应最终聚类结果的目标函数 $E$ .

采用肘部法则确定最佳聚类数的核心思想为:对于有一定的区分度的数据来说,伴随着K-means算法的聚类数 $k$ 值的增大,所有样本的划分会越来越精细,那么每一个类别的内部聚合程度会越来越高,则聚类的误差平方和SSE的值也就会变得越来越小.同时,当聚类数 $k$ 的值还没达到样本集最优的聚类数时,伴随

$k$ 值的增大,会极大增加样本中各类别的聚合程度,此阶段误差平方和SSE的值随 $k$ 值的增大有较大下降幅度.而当聚类数 $k$ 的值超过最优的聚类数时,伴随 $k$ 值的增大,所能够得到的各类别聚合程度的回报会快速的降低,此阶段的误差平方和SSE的降幅会相应的变小,并随着聚类数 $k$ 的值的增加而平缓增长.因此误差平方和SSE的值与聚类数 $k$ 的关系图变化曲线会形如手肘,而图中的这个肘部位置也即是曲线拐点的位置对应的聚类数 $k$ 就是所要找的最佳聚类数.

## 2.3 轮廓系数

对于一个聚类任务的结果,最希望的聚类情况是簇内样本尽量紧密,簇间的样本尽量远离.而轮廓系数(silhouette coefficient)同时考虑了类别内部的聚集度以及类别间的分离度,用以评价聚类结果的好坏.

轮廓系数的计算方法如下:

步骤1. 计算样本 $i$ 到同类别下其他样本的平均距离 $a(i)$ . $a(i)$ 越小,表明样本 $i$ 与该类簇的相似度越高,越应该被聚类到该类簇.

步骤2. 计算样本 $i$ 到与他相邻的最近的一个类别簇的所有样本的平均距离 $b(i)$ . $b(i)$ 越大,表明样本 $i$ 越不属于其他的类簇.

步骤3. 根据以上计算,定义样本 $i$ 的轮廓系数如下:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (15)$$

步骤4. 对所有的样本计算轮廓系数,再采用所有样本的轮廓系数的均值作为该聚类结果的轮廓系数.

由轮廓系数的定义公式可知,其取值范围为 $[-1, 1]$ .其中取值越大,就说明聚类得到的结果越加合理.当 $s(i)$ 越接近1时,表明对样本 $i$ 的聚类越合理; $s(i)$ 的值越接近-1,则表明样本 $i$ 更应该被分类到其他的类别簇中去.当 $s(i)$ 接近0时,表明样本 $i$ 在两个类别簇的边界上.使用轮廓系数来确定最佳聚类数时,应尽量选择使轮廓系数较大的聚类数 $k$ 值.

## 3 实验结果

基于本文第1节提取的样本特征向量集,采用第2节的K-means聚类算法模型进行聚类分析,其中聚类数 $k$ 的范围设定为 $[2, 9]$ .则各聚类数 $k$ 聚类结果如图1以及图2所示,图1为聚类数 $k$ 与误差平方和的关系图,也即肘部法则图,图2为聚类数 $k$ 与轮廓系数的关系图.

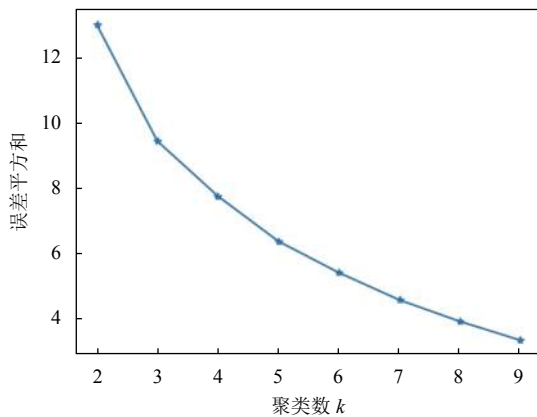


图1 肘部法则关系图

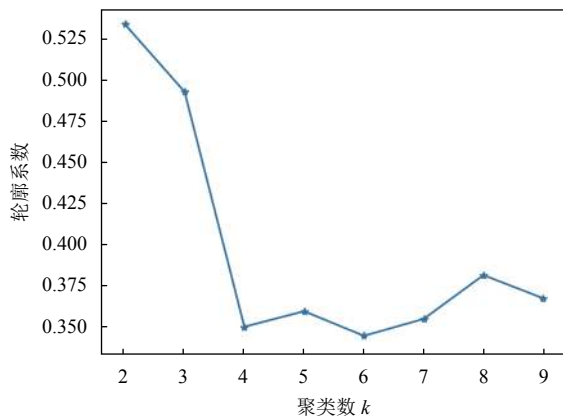


图2 轮廓系数与聚类数 k 关系图

由图1可知, 聚类数  $k$  与误差平方和的关系图的拐点出现在  $k=3$  处, 根据肘部法则的思想, 初步确定  $k=3$  为较好的聚类数。由图2可知, 当  $k=2$  时, 聚类结果的轮廓系数最大, 但结合图1可知, 此时的聚类结果的误差平方和还较大, 而  $k=3$  时的轮廓系数仅略低于  $k=2$  时的轮廓系数, 同时,  $k=3$  时的误差平方和也相对较低, 而且  $k=3$  也是肘部法则评判标准中的最佳聚类数。

因此, 结合肘部法则以及轮廓系数两个评价标准, 最终确定  $k=3$  为最佳聚类数, 即将样本聚为3类, 此时聚类结果的误差平方和的取值为 9.4758, 对应的轮廓系数取值为 0.4897。最终的聚类划分结果如表1所示。

由表1可知各样本所属类别。根据各样本所对应的时段, 将各相邻的样本对应的时段进行合并, 可以得到目标线路运营日中工作日一天的时段划分的初步结果如表2所示。初步的划分结果将一天划分为9个时段。

表1 样本聚类结果

类别编号	样本编号	对应运营时段
类别1	1-3	6:00-6:30
	99-108	22:20-24:00
类别2	4-7	6:30-7:10
	18	8:50-9:00
	20-68	9:10-17:20
	78-98	18:50-22:20
类别3	8-17	7:10-8:50
	19	9:00-9:10
	69-77	17:20-18:50

表2 初步的时段划分方案

时段编号	运营时段	时段编号	运营时段
1	6:00-6:30	6	9:10-17:20
2	6:30-7:10	7	17:20-18:50
3	7:10-8:50	8	18:50-22:20
4	8:50-9:00	9	22:20-24:00
5	9:00-9:10		

由表2可以得到初步的运营时段划分方案。但是在初步的划分方案中, 时段4以及时段5分别只有10分钟的长度, 这样的划分方式对于运营公司来说并不易于操作, 因此需要对该划分结果进行修正。考虑到时段5与时段3在聚类时属于同一类别, 因此将时段3、4、5合并为同一个运营时段。同时由于福州地铁一号线的首末班车发车时间分别为 6:30、23:00, 因此经过修正后的运营时段的划分方案如表3所示。

表3 修正后的时段划分方案

时段编号	运营时段	时段编号	运营时段
1	6:30-7:10	4	17:20-18:50
2	7:10-9:10	5	18:50-22:20
3	9:10-17:20	6	22:20-23:00

最终的运营时段划分方案将工作日一天划分为6个时段, 分别为 6:30-7:10、7:10-9:10、9:10-17:20、17:20-18:50、18:50-22:20、22:20-23:00。其划分结果示意图如图3所示。图中虚线即为各时段的起止时间点分割线。

在该运营时段的划分方案中, 每个时段内的客流都有着相似的变化特点, 运营公司可以据此时段划分方案, 根据各时段内的实际客流需求, 采取合适的发车间隔, 制定全天的列车运行方案。值得指出的是, 时段 6:30-7:10、9:10-17:20 以及 18:50-22:20 由于属于相同类别, 因此其客流变化特性具有相似性, 在制定发车间隔时可以考虑采用相同或者相近的发车间隔。同理,

时段 7:10–9:10 与 17:20–18:50 也是如此。这也为运营公司制定行车方案提供了更加丰富的思路。

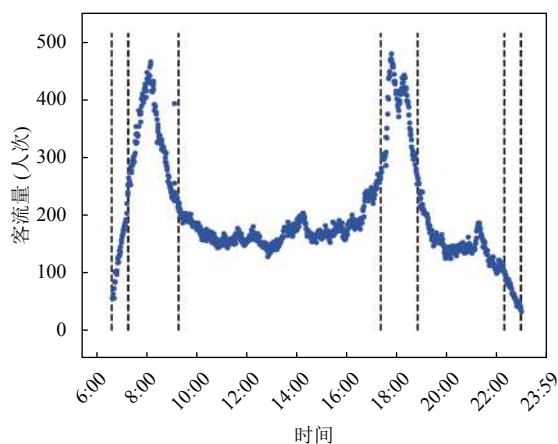


图3 工作日全天时段划分

#### 4 结论与展望

本文着眼于地铁运营时段划分问题进行了相应的研究。针对传统研究中采用客流量作为样本进行研究的局限等问题,提出一种构建时段客流特征向量的方法进行地铁运营时段的划分。首先构建线路目标运营日的单日客流序列,并采用合适的时间间隔将客流序列进行切片,再通过构建各时段客流样本特征向量的方式描述各时段的客流特征,最后借助于 K-means 聚类算法模型确定最佳聚类结果,进而得到最优的运营时段划分方案。

地铁运营时段的合理划分,能够帮助运营部门制定更加符合实际需求的发车方案。本文提出的基于时段客流特征聚类的运营时段划分方法,以大量实际客流为基础,考虑了地铁客流的实际变化特性,能够更加客观的给出划分方案,避免了人工划分的主观性等缺点。当然,为了得到更加合理的地铁列车运行时刻表方案,需要在本研究的基础上进行进一步的各时段发车间隔研究,这同时也是作者下一步的研究重点。

#### 参考文献

- 牛惠民, 陈明明, 张明辉. 城市轨道交通列车开行方案的优化理论及方法. 中国铁道科学, 2011, 32(4): 128–133.
- 毛保华. 城市轨道交通系统运营管理. 北京: 人民交通出版社, 2006.
- Bulut A, Yanik S. Partitioning the time of day for bus schedule optimization. In: Kahraman C, Cebi S, Onar SC, *et al.*, eds. Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making. Cham: Springer, 2020. 298–304.

- 张雪妍, 贺锋, 王玉刚, 等. 城市公交系统运营时段准确划分研究与仿真. 计算机仿真, 2019, 36(11): 126–129.
- 张瑾, 向一攀. 基于时间标签的公交运营时段划分方法. 物流科技, 2020, 43(5): 93–96.
- Jian W, Yang C. Operating time division for a bus route based on the recovery of GPS Data. Journal of Sensors, 2017, 2017: 1321237.
- Bie YM, Gong XL, Liu ZY, *et al.* Time of day intervals partition for bus schedule using GPS data. Transportation Research Part C: Emerging Technologies, 2015, 60: 443–456. [doi: 10.1016/j.trc.2015.09.016]
- 沈吟东, 张全辉, 徐甲. 基于 K-means 聚类算法的公交运营时段分析. 交通运输系统工程与信息, 2014, 14(2): 87–93. [doi: 10.3969/j.issn.1009-6744.2014.02.014]
- 靳文舟, 李鹏, 巫威眺. 基于多源公交数据和车时成本优化的公交运营时段划分方法. 中国公路学报, 2019, 32(2): 143–154. [doi: 10.3969/j.issn.1001-7372.2019.02.015]
- Wang XD, Cottrell W, Mu SC. Using k-means clustering to identify time-of-day break points for traffic signal timing plans. Proceedings of 2005 IEEE Intelligent Transportation Systems. Vienna, Austria. 2005. 586–591.
- Hauser TA, Scherer WT. Data mining tools for real-time traffic signal decision support & maintenance. Proceedings of 2001 IEEE International Conference on Systems, Man and Cybernetics. Tucson, AZ, USA. 2001. 1471–1477.
- Ratroun NT. Subtractive clustering-based k-means technique for determining optimum time-of-day breakpoints. Journal of Computing in Civil Engineering, 2011, 25(5): 380–387. [doi: 10.1061/(ASCE)CP.1943-5487.0000099]
- 姚佼, 徐洁琼, 韩印. 基于聚类分析的城市交通 TOD 优化控制方法. 交通运输工程学报, 2014, 14(6): 110–116. [doi: 10.3969/j.issn.1671-1637.2014.06.014]
- 杜长海, 黄席樾, 杨祖元, 等. 改进的 FCM 聚类在交通时段自动划分中的应用. 计算机工程与应用, 2009, 45(24): 190–193. [doi: 10.3778/j.issn.1002-8331.2009.24.057]
- 杨立才, 贾磊, 孔庆杰, 等. 基于人工免疫算法的交通时段自动划分方法. 控制理论与应用, 2006, 23(2): 193–198. [doi: 10.3969/j.issn.1000-8152.2006.02.006]
- 王文宪, 肖蒙, 成琳娜, 等. 基于近邻传播聚类的地铁运营时段划分. 运筹与管理, 2018, 27(12): 187–192.
- 曾小旭, 汪林, 罗贤迪, 等. 有序样本聚类方法在城市轨道交通运营时段划分中的应用. 都市快轨交通, 2017, 30(2): 108–112. [doi: 10.3969/j.issn.1672-6073.2017.02.022]