

基于客流特征的北京地铁站点类型识别

尹 芹¹, 孟 斌², 张丽英³

(1. 首都师范大学资源环境与旅游学院, 北京 100048; 2. 北京联合大学应用文理学院, 北京 100191;
3. 中国石油大学地球物理与信息工程学院, 北京 102249)

摘 要:地铁站点是城市各种社会经济活动的关键节点,不同类型的地铁站点在城市中的区域条件、交通功能、土地利用类型等方面均存在差异,科学的站点分类有助于了解城市功能分区及评价轨道交通基础设施建设状况。本文基于2013年3月份14天工作日地铁刷卡客流量数据,运用引入客流特征的时间序列聚类方法,对北京市195个地铁站点进行分类。研究表明:①地铁站点客流量存在时空差异,也是城市功能分区时空差异的表现之一。②通过引入客流特征的时间序列方法,将地铁站点分为居住导向型、就业导向型、职住错位型、错位偏居住型、错位偏就业型、混合型、综合型及其他型8种不同类型。③利用地铁站点客流量数据,是将空间行为和实体空间进行关联比较的有效途径。

关键词:客流特征;时间序列聚类;地铁站点;北京

1 引言

地铁站点是城市轨道交通网中的关键节点,站点交通的可达性使其成为城市各种社会经济活动中的热点地区(傅搏峰等, 2008)。不同类型的站点在城市中的区域条件、交通功能、土地利用类型等方面均存在差异(段德罡等, 2013),科学的地铁站点分类,对了解城市功能分区及评价轨道交通基础设施建设情况具有重要意义。

国外有关地铁站点分类的研究大都以城市建成区站点为研究对象,郊区地铁站点作为其中子类(Korf et al, 1981)。一般依据站点交通节点特性或开放空间的场所特性进行分类指标选取,如车站标高形式、客流量大小、车站服务地区的主要功能、甚至于与某个指定参照系统的相关关系等(Bates Jr, 1978)。如日本东京、大阪等城市根据地铁站点所

处区位,首先将地铁站点划分为市区和郊区站点,再通过站点衔接方式种类、换乘比例等指标将其各划分为3个不同的等级(谢岫, 2011)。此外,国外学者除了对地铁站点分类的研究,还热衷于对站点客流量及站点环境等问题的探索。比如,有学者从站点尺度上分析了首尔地铁站点的客流量情况,发现地铁客流量随着起始站点及不同时间段的变化而变化,地铁客流量与周边建成环境有着重要关系(Choi et al, 2012);有学者研究了纽约市地铁站点的黑碳和PM_{2.5}的浓度(Vilcassim et al, 2014);还有学者分析了德黑兰地铁站点室内外空气传播真菌的浓度和分布特征(Hoseini et al, 2013)。

国内学者在地铁站点分类体系划分标准上尚未形成统一认识,主要的分类标准有以下几种:一是部分国内城市地铁建设采用的标准。如北京、上海、广州等大城市在相应的规划文件中,一般采用

收稿日期:2015-04;修订日期:2015-10。

基金项目:国家自然科学基金项目(41171136);北京市哲学社会科学基金项目(14CSA002);北京市属高等学校长城学者培育计划项目(IDHT20130322) [Foundation: National Natural Science Foundation of China, No.41171136; Beijing Philosophy and Social Science Foundation Grant, No.14CSA002; The Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions, No.IDHT20130322]。

作者简介:尹芹(1990-),女,山东青岛人,硕士,主要从事城市地理研究,E-mail: yinqin123@126.com。

通讯作者:孟斌(1971-),男,安徽肥东人,博士,教授,硕士生导师,主要从事城市地理和地理信息科学研究,
E-mail: mengbin@bnu.edu.cn。

引用格式:尹芹, 孟斌, 张丽英. 2016. 基于客流特征的北京地铁站点类型识别[J]. 地理科学进展, 35(1): 126-134. [Yin Q, Meng B, Zhang L Y. 2016. Classification of subway stations in Beijing based on passenger flow characteristics[J]. Progress in Geography, 35(1): 126-134.]. DOI: 10.18306/dlkxjz.2016.01.014

站点衔接的交通方式种类、轨道线路数及站点周边土地利用类型等作为分级指标,将地铁站点划分为3~4个等级(龚晓芳, 2009)。二是按站点周边土地利用方式分类,即按地铁站点的区位特征分类。吴娇蓉等(2007)结合实地调研数据,按照站点区位特征、站点周边土地利用性质、开发规模和强度等数据,将上海市郊区轨道交通站点划分为公共中心区站点、居住区站点、交通枢纽站点、工业园区站点、混合区站点、文教区站点和景区区站点七大类;段德罡等(2013)基于GIS研究确定轨道站点对周边地区的一般影响范围,划定轨道站点周边用地的合理影响范围,基于影响区域内用地功能特征,对西安地铁2号线的站点进行分类。三是通过轨道交通运营资料数据进行分类。马小毅(2013)从网络、线路、车站三个层面探讨轨道交通从单线向网络转变过程中的客流特征和成长规律,根据不同类型车站的客流特征差异,分为居住型、办公型、商业型和枢纽型;还有学者通过分析站点客流的波动性,发现北京轨道交通系统在不同时间维度和区域范围均呈现出明显的不均衡特征,针对车站特点进行聚类分析,归纳出周边不同用地类型的车站客流时空分布差异性特征规律(王静, 2013)。四是从站点的微观城市空间形态出发进行分类。可分为:以城市街区为主的站点、以站点为中心的节点、站点与城市空间融合的综合体、环绕城市中心发展的节点、城市内外各级交通枢纽等形式(张育南, 2011)。此外,有学者按站点敷设方式分类(石瑜, 2014),有学者按地铁站点的职能分类(谭啸, 2012),还有学者基于谱

聚类算法进行分类(余丽洁, 2014)。

从国内外研究现状看,国外研究中,区域、用地等差异性受到忽略,车站区位因素明显考虑不足,缺乏客观的定量标准(傅搏峰等, 2008)。国内学者大多通过实地调研数据或站点周边土地利用情况进行站点分类,基于其他数据和方法的分析相对较少。随着大数据挖掘及可视化技术日渐成熟,大数据逐渐应用于城市空间、城市等级体系及居民时空行为等研究领域(秦萧等, 2013),基于此,本文以北京市2013年3月份工作日的地铁刷卡客流量数据为数据源,采用引入客流特征的时间序列聚类方法,对195个地铁站点进行分类。

2 数据与方法

2.1 数据来源与处理

北京地铁刷卡基本信息包括进站刷卡时间、站点和出站刷卡时间、站点(表1)。考虑到数据完整性,以及2013年3月份北京市无重大节假日及活动,本文的数据来源为2013年3月1-20日的14个工作日地铁客流量数据(表1),由于个别站点信息缺失,所以进出站刷卡量存在差异(表2)。本文最终选择195个地铁站点为研究对象,不包括信息缺失站点及机场专线站点。

为充分揭示数据特征,本文将地铁客流量的原始数据按进出站处理成两个数据集(表3-4),每条记录代表该站点从4:00-23:00的客流量情况,客流量数值为14天工作日的平均客流量,在此基础上,还

表1 北京市部分地铁站点2013年3月1日刷卡数据

Tab.1 Examples of registered pass use at subway stations in Beijing on March 1st, 2013

ID	进站名称	进站刷卡时间	出站名称	出站刷卡时间
10007510*****6142	长椿街	2013/3/1 8:03:00	复兴门	2013/3/1 8:10:26
10007510*****5723	苹果园	2013/3/1 8:12:00	玉泉路	2013/3/1 8:29:08
10007510*****2821	通州北苑	2013/3/1 8:28:00	军事博物馆	2013/3/1 9:25:23
10007510*****6032	和平门	2013/3/1 9:36:00	五棵松	2013/3/1 10:07:08
10007510*****5779	古城路	2013/3/1 11:10:00	王府井	2013/3/1 11:58:26

表2 北京市部分地铁站点2013年3月客流量统计

Tab.2 Entering and exiting passenger flows of subway stations in Beijing, March 2013

	总客流量/(万/人次)	最大客流量/(万/人次)	最小客流量/(万/人次)	平均客流量/(万/人次)
进站	5383.38	467.11	357.71	384.53
出站	5503.56	483.29	365.45	393.11
合计	10886.94	-	-	777.64

注:仅统计3月1日-20日期间工作数据。

表3 部分站点进站客流量时间分布数据

Tab.3 Temporal distribution of entering passenger flow at some subway stations

站点名称	4:00 前		4:00-5:00		……	21:00-22:00		22:00-23:00	
	人数	标准化值	人数	标准化值		人数	标准化值	人数	标准化值
安华桥	0	-1.038	14	-0.992		58	-0.854	1	-1.035
八里桥	0	-0.712	75	-0.520		21	-0.657	2	-0.708
巴沟	3	-1.321	24	-1.258	……	97	-1.045	3	-1.320
北土城	1	-1.143	26	-1.088		106	-0.910	4	-1.137
草房	3	-0.753	268	-0.605		105	-0.696	3	-0.753

表4 部分站点出站客流量时间分布数据

Tab.4 Temporal distribution of exiting passenger flow at some subway stations

站点名称	4:00 前		4:00-5:00		……	21:00-22:00		22:00-23:00	
	人数	标准化值	人数	标准化值		人数	标准化值	人数	标准化值
安华桥	2	-0.874	48	-0.868		151	-0.495	1	-0.871
八里桥	1	-0.915	23	-0.911		238	-0.031	117	-0.482
巴沟	1	-1.269	88	-1.267	……	338	-0.508	109	-1.023
北土城	8	-1.070	117	-1.055		230	-0.615	60	-0.951
草房	0	-0.920	0	-0.920		1093	-0.170	357	-0.675

对客流量数值进行了标准化处理。数据标准化处理主要是无量纲化处理,解决数据的可比性问题;采用 z-score 标准化,这种方法是基于原始数据的均值和标准差进行处理,使得不同量纲的数据具有可比性,均值为 0,标准差为 1;大于 0 表明高于平均水平,小于 0 表明低于平均水平;本文标准化后的数值为正,表示该站点一小时内的客流量高于当天平均客流量,为负则表示低于平均客流量。

2.2 研究方法

时间序列聚类方法可分为基于原始测度数据的时间序列聚类法和基于特征的时间序列聚类法。基于特征的时间序列聚类即先对原始数据进行降维,抽取表征及其内在变化机制的特征作为相似性度量的基础,然后运用各种聚类方法对这些特征进行聚类(宋辞等, 2012)。本文采用引入客流特征的时间序列聚类方法进行分析,从不同的角度出发提取时间序列的特征,然后将时间序列聚类分析相应转化为用新特征值描述的时间序列之间的聚类分析(孙旭, 2009),再用层次聚类方法初步确定聚类数量,最后再用 K-均值聚类方法。

在层次聚类方法中,首先确定距离的基本定义和类间距离的计算方式,随后按照距离的远近,将距离较近的数据依次并入一类,依次类推,直到数据完全归为一类为止;或是首先将所有数据视为同一类别,然后按照距离远近,将距离远的数据依次

分离开来,直到所有的数据各自成为一类为止,整个分析过程都可以用树形图表示(张文彤等, 2013)。

K-均值聚类法也称快速聚类法,是为了将案例快速分成 K 类。一般而言,具体的类别个数需在分析前加以确定,整个分析过程使用迭代的方式,首先起步于一个初始的分类,然后通过不断的迭代将数据在不同类别间进行移动,直到最后达到一定的标准为止(张文彤等, 2013)。

2.3 时间序列特征值的选取与测算

时间序列具有多项特征,每项特征刻画了时间序列的一个方面,从对时间序列不同层次上的认知,可分为形态特征、结构特征以及模型特征。其中,形态特征主要指时间序列形状变化的特征,包括全局和局部特征,全局特征描述了时间序列的起伏变化,如上升、下降等;局部特征则表现为局部时间点上的异常观测值,如极值点、转折点等。结构特征是对时间序列全局构造或内在变化机制的描述,可很好地表现时间序列的全局特点,包括基本统计特征,如偏度和峰度、时域特征,如趋势和季节波动、频域特征,如周期解析强度和谱密度。模型特征描述了事物变化潜在的运动规律,这些模型包括高斯过程模型、ARMA 以及 ARIMA 模型等(宋辞等, 2012)。本文从形态特征和结构特征出发,提取时间序列的极大值点、偏度和峰度。

极值点是函数图像在某段子区间上最大值或

最小值的横纵坐标。根据地铁站点刷卡数据,可以得出各站点不同的客流小时分布曲线图,不同站点的客流波峰数量不同,由此,本文将波峰数量作为区分各站点的标志之一,选用极大值个数作为时间序列的特征值。

偏度是描述数据分布形态的统计量,描述的是某总体取值分布的对称性。该统计量需要与正态分布相比较,偏度为0表示其数据分布形态与正态分布的偏斜程度相同,偏度大于0表示为右偏,偏度小于0表示为左偏。

峰度是描述某总体取值分布形态陡缓程度的统计量,峰度为0表示该数据分布与正态分布的陡缓程度相同,峰度大于0表示该数据分布较为陡峭,为尖顶峰,峰度小于0表示较为平坦,为平顶峰。

本文除了提取时间序列的形态特征和结构特征,还引入客流特征进行聚类。地铁站点客流在全天不同时间上分布是不同的,通常用高峰小时系数和客流时段分布均衡系数分析客流的时间分布特征(王玉萍, 2011)。客流的时段(小时)系数 p 定义为:

$$p=Q_i/Q_d \tag{1}$$

式中, Q_i 表示第*i*小时的客流量(单位为人次); Q_d 表示全天的客流量; p 的最大值即为高峰小时系数。定义客流时段分布的均衡系数 U :

$$U=G/H \tag{2}$$

式中, G 表示高峰时间段的小时平均客运量; H 表示平峰时间段的小时平均客运量;当 $U>2$ 时,表示很不均衡; $1.6\leq U\leq 2$,不均衡; $U<1.6$,较均衡。

本文分别提取195个地铁站点进出站的极大值点、偏度、峰度、客流时段分布均衡系数和高峰小时系数,一个站点即一个时间序列,每个时间序列对应10项特征值(表5),最后,对用特征值描述的时间序列进行层次聚类 and K -均值聚类分析。在SPSS软

件19版本中,层次聚类时选用组间联接方法,以平方欧氏距离为度量标准,变量为10项特征值;在 K -均值聚类时,选用10项特征值为10个变量,最大迭代次数选择默认值10,统计量为ANOVA表及每个个案的聚类信息。

3 结果分析

3.1 层次聚类结果

层次聚类的结果为聚类表及树形图,通过聚类系数及树形图判断数据分成多少类别时较为合适,通过在不同距离下的层次聚类数(图1)可以明显看出,聚类数为14、10、8和5是转折点,较为合适,因此,本文将195个站点用 K -均值聚类方法分为8类。

3.2 K -均值聚类结果

用 K -均值聚类方法将195个地铁站点分为1类至8类站点,再根据各类站点进出站客流量小时分布曲线,8类站点可以总结为居住导向型、就业导向型、职住错位型、错位偏居住型、错位偏就业型、混合型、综合型及其他型(表6)。

表7给出的是方差分析结果,实际上是按照类别分组后对所有变量依次进行的单因素方差分析,从中根据 F 值的大小近似得到各变量在聚类分析中作用的大小。各个变量对聚类结果的重要程度排序为:出站高峰小时系数>进站高峰小时系数>出站峰度>出站偏度>进站偏度>出站客流时段分布均衡系数>进站峰度>进站客流时段分布均衡系数>出站极大值个数>进站极大值个数。

(1) 居住导向型站点

第1类站点的客流量小时分布曲线(图2)表现为单峰型,进站早高峰、出站晚高峰突出,是典型的居住导向型站点。此类站点进出站客流时段分布均衡系数的平均值较大,客流时段分布很不均衡;

表5 部分站点进出站特征值

Tab.5 Characteristic value of entering and exiting passenger flows at some subway stations

站定名称	形态特征		基本统计特征				客流的时间分布特征			
	极大值点		偏度		峰度		客流时段分布均衡系数		高峰小时系数/%	
	进站	出站	进站	出站	进站	出站	进站	出站	进站	出站
安华桥	3	3	1.34	2.46	0.87	7.38	3.523	4.419	16.69	24.94
八里桥	3	3	2.25	1.96	4.36	3.34	3.396	2.387	24.32	20.03
巴沟	4	4	0.70	1.09	-0.19	0.97	2.324	3.056	12.27	14.41
北土城	2	3	1.07	1.75	0.14	3.66	3.238	3.653	14.27	19.31
草房	3	2	2.25	1.93	4.48	3.22	3.294	2.396	23.02	20.75

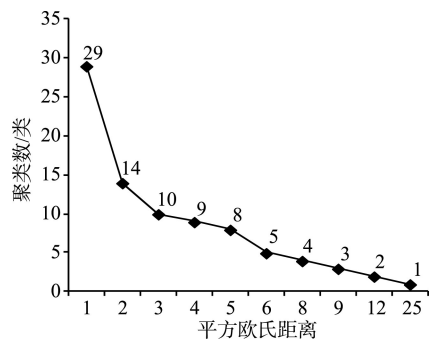


图1 不同平方欧氏距离下层次聚类数

Fig.1 The number of system clusters under different distances

高峰小时系数值大,进出站有明显的早晚高峰。例如,回龙观、回龙观东大街就是典型的居住导向型站点,这些站点周边土地利用类型较单一,站点的主要功能是为通勤人群提供交通便利。

(2) 就业导向型站点

这类站点的客流量小时分布曲线(图3)表现为单峰型,进站晚高峰、出站早高峰明显,是典型的就业导向型站点,如复兴门、西二旗等;此类站点进出站客流时段分布均衡系数的平均值较大,客流时段分布很不均衡;高峰小时系数值大,进出站有明显

的早晚高峰。

(3) 职住错位型站点

这类站点的客流量小时分布曲线表现为典型的双峰型(图4),同一时间段既有进站客流,又有出站客流,表明站点周边既有居住地又有就业地。在此居住的居民,工作地可能在其他地点,而在此类站点附近工作人群却多居住在别处,表现出较为明显的职住错位,本文将其归纳为职住错位型,典型站点如东四、望京西等。

(4) 错位偏居住型站点

第4类站点与第3类站点相比,虽然也表现为双峰型(图5),也是居住与就业的错位型,但两个波峰的高度存在显著差别,一波峰突出而另一波峰较平缓,进站高峰出现在早晨,而出站高峰出现在下午,站点类型更偏向于居住地,即错位偏居住型站点;此类站点客流时段分布均衡系数和高峰小时系数数值较大,例如,天通苑、亦庄桥、苹果园等站点。

(5) 错位偏就业型站点

第5类站点与4类站点特点相似,也表现为双峰型(图6),但进站高峰出现在下午,而出站高峰出现在早晨,站点类型更偏向于就业地,即错位偏就

表6 K-均值聚类结果
Tab.6 Result of K-means clustering

类别	名称	站点数量/个	站点名称
1类	居住导向型	13	常营、次渠南、稻田、巩华城、广阳城、回龙观、朱辛庄、回龙观东大街、篱笆房、良乡大学城西、生命科学园、沙河、沙河高教园
2类	就业导向型	4	复兴门、金台夕照、荣京东街、西二旗
3类	职住错位型	14	朝阳门、车公庄、东单、东四、国贸、海淀黄庄、呼家楼、知春路、雍和宫、立水桥、宋家庄、望京西、西直门、宣武门
4类	错位偏居住型	50	安河桥北、八宝山、八角游乐园、八里桥、北宫门、北苑、小红门、长阳、西红门、草房、慈寿寺、次渠、裕链坡、俸伯、高米店北、公益西桥、枣园、育新、天通苑南、古城路、管庄、果园、黄村火车站、黄渠、角门西、九棵树、玉泉路、永泰庄、苏庄、旧宫、梨园、立水桥南、良乡南关、林萃桥、临河里、龙泽、亦庄桥、新宫、天通苑北、马家堡、南部、苹果园、蒲黄榆、清源路、十里堡、石门、顺义、西苑、土桥、天通苑
5类	错位偏就业型	21	奥林匹克公园、白石桥南、车公庄西、东大桥、东四十条、知春里、西小口、万源街、阜成门、高碑店、国家图书馆、和平西桥、建国门、健德门、永安里、柳芳、苏州街、经海路、灵境胡同、荣昌东街、三元桥
6类	混合型	31	安定门、安华桥、安贞门、奥体中心、北土城、北苑路北、中关村、西土城、团结湖、磁器口、大葆台、大屯路东、大望路、大钟寺、灯市口、光熙门、五道口、望京、青年路、国展、花园桥、惠新西街北口、亮马桥、牡丹园、木樨地、魏公村、四惠、人民大学、南礼士路、农业展览馆、生物医药基地
7类	综合型	49	巴沟、北海北、北京大学东门、菜市口、传媒大学、崔各庄、长椿街、西四、五棵松、动物园、高米店南、郭公庄、海淀五路居、和平里北街、张自忠路、王府井、肖村、和平门、后沙峪、花梨坎、黄村西大街、惠新西街南口、亦庄文化园、万寿路、新街口、霍营、积水潭、金台路、劲松、军事博物馆、刘家窑、马泉营、圆明园、同济南路、孙河、南法信、南锣鼓巷、森林公园南门、上地、双井、双桥、四惠东、义和庄、通州北苑、太阳宫、陶然亭、天安门东、天安门西、天宮院
8类	其他型	13	北京南站、北京站、北新桥、崇文门、东直门、鼓楼大街、良乡大学城、西单、良乡大学城北、平安里、前门、芍药居、天坛东门

业型站点,如苏州街、三元桥、国家图书馆等;此类站点客流时段分布均衡系数和高峰小时系数值较大。

(6) 混合型站点

第6类站点具体表现为三峰型(图7),不仅早晚高峰均突出明显,在11:00-14:00也表现为一个小高

峰;和第3类站点类似,此类站点周边既有居住地又有就业地,但午间时段的小高峰可能意味着站点周边综合性服务设施较为完善,因此将此类站点称为混合型站点,如中关村、五道口、西土城、花园桥等最为典型。

(7) 综合型站点

第7类站点波形不规律(图8),客流量不但存在早晚高峰,而且在中午和晚高峰后会出现第三、第

表7 K-均值聚类ANOVA表
Tab.7 ANOVA table of K-means clustering

	聚类		误差		F	Sig.
	均方	df	均方	df		
进站峰度	113.559	7	1.830	187	62.045	0
进站偏度	8.387	7	0.101	187	82.931	0
进站客流时段分布均衡系数	2.637	7	0.415	187	6.346	0
进站高峰小时系数	896.166	7	4.834	187	185.404	0
进站极大值个数	1.124	7	0.311	187	3.612	0.001
出站偏度	10.459	7	0.105	187	99.916	0
出站峰度	175.055	7	1.481	187	118.220	0
出站高峰小时系数	756.462	7	3.66	187	206.707	0
出站客流时段分布均衡系数	15.388	7	0.202	187	76.131	0
出站极大值个数	1.194	7	0.327	187	3.650	0.001

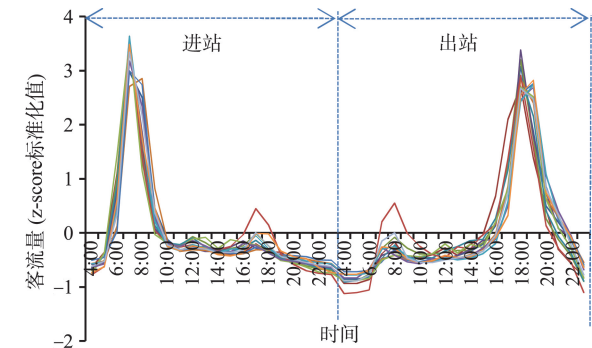


图2 居住导向型站点进出站客流量小时分布曲线
Fig.2 Temporal distribution of entering and exiting passenger flows of residential-oriented stations

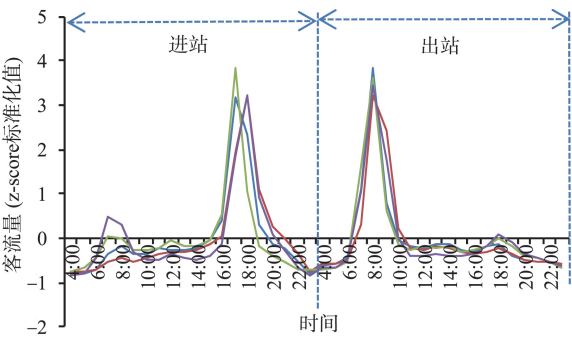


图3 就业导向型站点进出站客流量小时分布曲线
Fig.3 Temporal distribution of entering and exiting passenger flows of employment-oriented stations

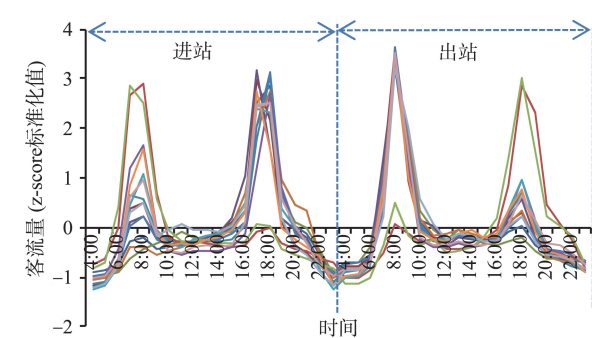


图4 职住错位型站点进出站客流量小时分布曲线
Fig.4 Temporal distribution of entering and exiting passenger flows of spatially mismatched stations

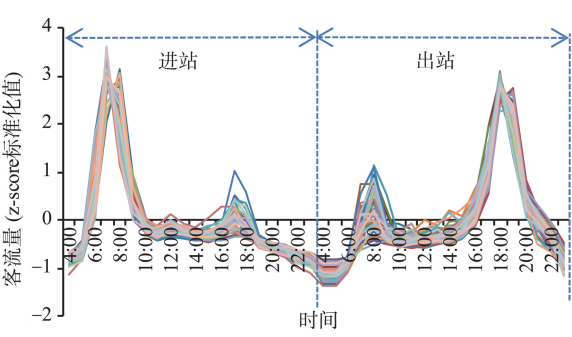


图5 错位偏居住型站点进出站客流量小时分布曲线
Fig.5 Temporal distribution of entering and exiting passenger flows of mixed mainly residential-oriented stations

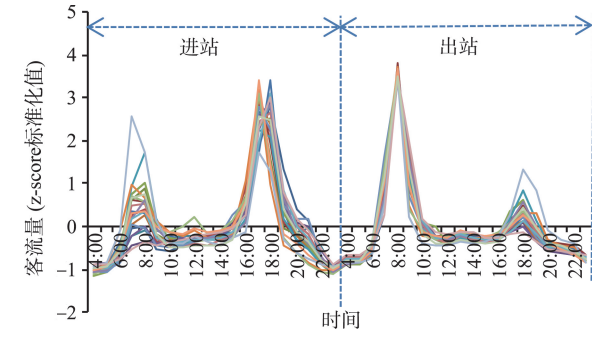


图6 错位偏就业型站点进出站客流量小时分布曲线
Fig.6 Temporal distribution of entering and exiting passenger flows of mixed mainly employment-oriented stations

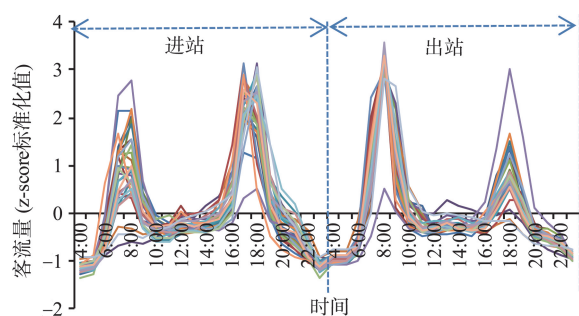


图7 混合型站点进出站客流量小时分布曲线
Fig.7 Temporal distribution of entering and exiting passenger flows of mixed type stations

四个小高峰,为综合型站点。站点周围土地利用类型多样,其中,景点型如南锣鼓巷、天安门东;商服型如王府井、动物园;科教型如北京大学东门、亦庄文化园等。部分站点客流量时段分布较均衡,高峰小时系数值较小。

(8) 其他型站点

第8类站点客流量时间分布曲线不规律(图9),主要表现为波峰数量多,原因是此类站点中包含大量交通枢纽型站点,如北京站、北京南站、东直门等;该类站点客流量时段分布均衡,高峰小时系数值小。

3.3 地铁站点类型空间分布

运用ArcGIS软件将分类结果进行专题图制作,可以直观看出其空间分布特点,结果如图10所示。

居住导向型站点集中分布于房山区和昌平区,即西部和北部近远郊城区分布较密集,在其周围分布有大型居住地。就业导向型站点数量少,分布分散,没有明显规律。错位型站点整体数量最多,空间分布广泛,反映出北京城市职住错位形势较为严峻,例如,典型职住错位型站点大部分分布于二环以内,如东四、西直门附近,且北部城区多于南部城区,少数分布于城市北部近郊,如立水桥附近;错位偏居住型站点内城数量少,集中于城市二环以外,分布范围广,形态上呈环状;错位偏就业型站点大部分集中于内城,如建国门、苏州街附近,且北部城区数量多于南部,少数站点分布于东南近郊,如高碑店附近。混合型站点从整体分布上看,南部城区数量少,分布零散,北部城区站点数量多,分布密集,以中关村、奥体中心和亮马桥等城市次中心附近最为集中。综合型站点数量多,分布广,在景区、商服中心、科教单位附近分布集中。其他型站点主

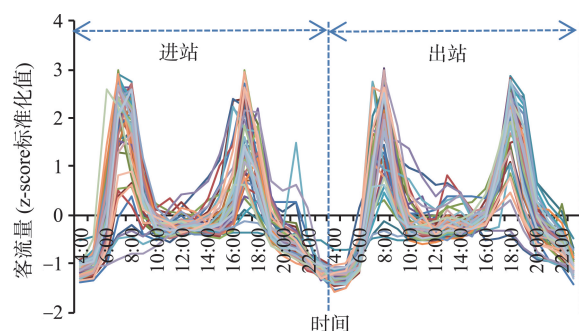


图8 综合型站点进出站客流量小时分布曲线
Fig.8 Temporal distribution of entering and exiting passenger flows of commerce- and attraction-oriented subway stations

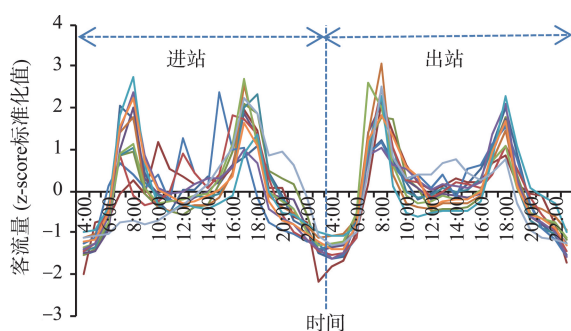


图9 其他型站点进出站客流量小时分布曲线
Fig.9 Temporal distribution of entering and exiting passenger flows of other stations

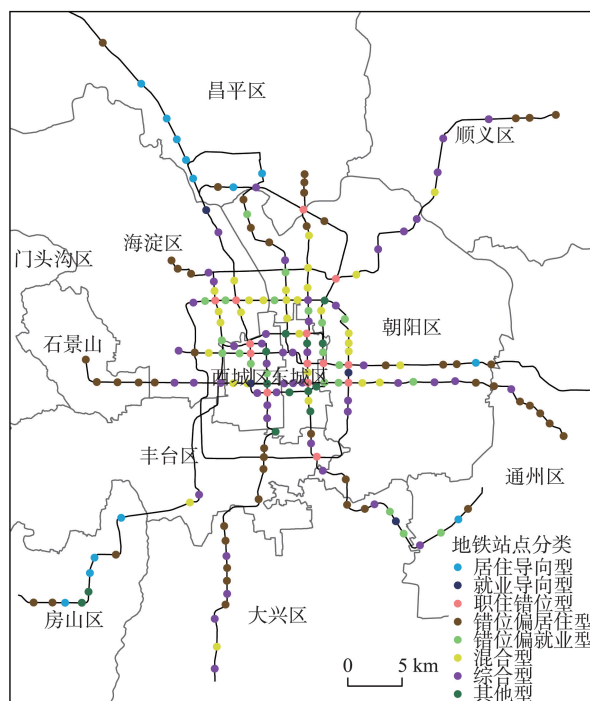


图10 地铁站点类型空间分布
Fig.10 Spatial distribution of different types of subway stations

要分布于内城区,在交通枢纽地区分布最为明显,如北京站和东直门站等。

4 结论与讨论

本文利用地铁刷卡客流量数据,采用引用客流特征的时间序列聚类方法,对北京市195个地铁站点进行分类。研究表明:

(1) 地铁站点客流量存在时空差异,表明了城市不同功能分区所对应的城市生活的时空的不平衡。通过地铁站点客流量时间分布的研究,可对站点进行分类,有助于进一步解读城市的不同功能分区及其所体现的社会行为特征,对理解城市空间格局及其演化具有重要意义。

(2) 根据聚类分析结果,本文将195个地铁站点分为居住导向型、就业导向型、职住错位型、错位偏居住型、错位偏就业型、混合型、综合型及其他型8种不同类型,证实了引入客流特征的时间序列聚类方法在处理此类大数据问题方面的有效性。

(3) 研究结果表明,地铁站点客流量数据,是将空间行为和实体空间进行关联比较的有效途径。相对于城市规划数据,地铁流量数据是城市居民对基础设施使用情况的直观反映,是从使用者角度对城市空间进行研究的新视角。

本文仅针对部分地铁站点的客流量数据进行分析,并且特征值的选取不够全面,在以后的研究中将进一步增加站点相关的辅助信息,如站点周边土地利用情况等信息,科学选取特征值,对地铁站点类型的差异作更全面的分析,以期深入挖掘城市功能分区的时空差异及城市空间格局的演变。

参考文献(References)

段德罡, 张凡. 2013. 土地利用优化视角下的城市轨道交通站点分类研究: 以西安地铁2号线为例[J]. 城市规划, 37(9): 39-45. [Duan D G, Zhang F. 2013. Study on classification of urban rail transit stations from the perspective of land use optimization: A case study on Xi'an subway line 2[J]. City Planning Review, 37(9): 39-45.]

傅博峰, 吴娇蓉, 陈小鸿. 2008. 郊区轨道站点分类方法研究[J]. 铁道学报, 30(6): 19-23. [Fu B F, Wu J R, Chen X H. 2008. Method of classification of suburban rail transit station sites[J]. Journal of the China Railway Society, 30(6): 19-23.]

龚晓芳. 2009. 现代城市轨道交通站点地区规划研究[D]. 西安: 长安大学. [Gong X F. 2009. The planning research

of metro rail transit station in modern city[D]. Xi'an, China: Chang'an University.]

马小毅, 金安, 刘明敏, 等. 2013. 广州市轨道交通客流特征分析[J]. 城市交通, 11(6): 35-42. [Ma X Y, Jin A, Liu M M, et al. 2013. Rail transit passenger flow characteristics in Guangzhou[J]. Urban Transport of China, 11(6): 35-42.]

秦萧, 甄峰, 熊丽芳, 等. 2013. 大数据时代城市时空行为研究方法[J]. 地理科学进展, 32(9): 1352-1361. [Qin X, Zhen F, Xiong L F, et al. 2013. Methods in urban temporal and spatial behavior research in the big data era[J]. Progress in Geography, 32(9): 1352-1361.]

石瑜. 2014. 基于GIS的城市地铁站点布设研究[D]. 石家庄: 石家庄铁道大学. [Shi Y. 2014. Research on the urban metro station layout based on GIS[D]. Shijiazhuang, China: Shijiazhuang Tiedao University.]

宋辞, 裴韬. 2012. 基于特征的时间序列聚类方法研究进展[J]. 地理科学进展, 31(10): 1307-1317. [Song C, Pei T. 2012. Research progress in time series clustering methods based on characteristics[J]. Progress in Geography, 31(10): 1307-1317.]

孙旭. 2009. 时间序列全局特征聚类分析方法及其应用[J]. 统计教育, (3): 55-59. [Sun X. 2009. Analysis and application of the global characteristics cluster[J]. Statistical Thinktank, (3): 55-59.]

谭啸. 2012. 天津城市轨道交通站点周边土地利用优化研究[D]. 天津: 天津大学. [Tan X. 2012. Study on optimization of land-use surrounding the urban rail station in Tianjin[D]. Tianjin, China: Tianjin University.]

王静, 刘剑锋, 马毅林, 等. 2013. 北京市轨道交通车站客流时空分布特征[J]. 城市交通, 11(6): 18-27. [Wang J, Liu J F, Ma Y L, et al. 2013. Temporal and spatial passenger flow distribution characteristics at rail transit stations in Beijing[J]. Urban Transport of China, 11(6): 18-27.]

王玉萍. 2011. 城市轨道交通客流预测与分析方法[D]. 西安: 长安大学. [Wang Y P. 2011. Study on urban rail transit passenger forecast and analysis[D]. Xi'an, China: Chang'an University.]

吴娇蓉, 毕艳祥, 傅博峰. 2007. 基于郊区轨道交通站点分类的客流特征和换乘系统优先级分析[J]. 城市轨道交通研究, (11): 23-28. [Wu J R, Bi Y X, Fu B F. 2007. Characteristics of in passenger different suburban rail stations and the priority of transfer system[J]. Urban Mass Transit, (11): 23-28.]

谢岫. 2011. 城市轨道交通站点周边空间形态整合浅析: 以南京为例[D]. 南京: 南京大学. [Xie S. 2011. Study on morphological integration of space surrounding urban rail transit stations: A case study of Nanjing[D]. Nanjing, Chi-

- na: Nanjing University.]
- 余丽洁, 李岩, 陈宽民. 2014. 基于谱聚类的城市轨道站点分类方法[J]. 交通信息与安全, 32(1): 122-125, 129. [Yu L J, Li Y, Chen K M. 2014. Using spectral clustering for urban rail station classification[J]. Journal of Transport Information and Safety, 32(1): 122-125, 129.]
- 张文彤, 董伟. 2013. SPSS 统计分析高级教程[M]. 2 版. 北京: 高等教育出版社. [Zhang W T, Dong W. 2013. SPSS tongji fenxi gaoji jiaocheng[M]. 2nd ed. Beijing, China: Higher Education Press.]
- 张育南. 2011. 城市轨道交通与城市空间整合发展问题初探[J]. 华中建筑, 29(10): 89-91. [Zhang Y N. 2011. Basic research on development of URT integrated with urban spaces[J]. Huazhong Architecture, 29(10): 89-91.]
- Bates Jr E G. 1978. A study of passenger transfer facilities[C]// Planning and design of rapid transit facilities. Washington, DC: Transportation Research Record.
- Choi J, Lee Y G, Kim T, et al. 2012. An analysis of metro ridership at the station-to-station level in Seoul[J]. Transportation, 39(3): 705-722.
- Hoseini M, Jabbari H, Naddafi K, et al. 2013. Concentration and distribution characteristics of airborne fungi in indoor and outdoor air of Tehran subway stations[J]. Aerobiologia, 29(3): 355-363.
- Korf J L, Demetsky M J. 1981. Analysis of rapid transit access mode Choice[C]//60th annual meeting of transportation research record, No. 817. Washington, DC: Transportation Research Record.
- Vilcassim M J R, Thurston G D, Peltier R E, et al. 2014. Black carbon and particulate matter (PM_{2.5}) concentrations in New York City's subway station[J]. Environmental Science & Technology, 48(24): 14738-14745.

Classification of subway stations in Beijing based on passenger flow characteristics

YIN Qin¹, MENG Bin², ZHANG Liying³

(1. College of Environment and Planning, Capital Normal University, Beijing 100048, China;

2. College of Arts and Science of Beijing Union University, Beijing 100191, China;

3. College of Geophysics and Information Engineering, China University of Petroleum, Beijing 102249, China)

Abstract: In cities there is a high concentration of social and economic activities at subway stations. Different types of subway stations have different functions, as reflected in their regional characteristics, traffic function, and land use. Meaningful station classification helps to understand urban functional partition and evaluate rail transportation infrastructure development. Based on the usage data of subway passes, this study classified the subway stations with time series clustering. The result shows that (1) subway station passenger flows have clear temporal and spatial differences. It reflects the temporal and spatial differences of urban functional partitions; (2) this study uses a time series clustering method. By considering passenger flow characteristics, subway stations can be divided into residential-oriented stations, employment-oriented stations, spatial mismatched stations, mixed mainly residential-oriented stations, mixed mainly employment-oriented stations, mixed type stations, commerce- and attraction-oriented stations, and other stations; (3) Using traffic data at subway stations is an effective way to compare spatial behavior and physical space.

Key words: passenger flow characteristics; time series clustering; subway stations; Beijing