首都师范大学学报(自然科学版)
第41卷 第6期 Journal of Capital Normal University No.6
2020年12月 (Natural Science Edition) Dec.,2020

大数据视角下的京津冀地区城市体系现状

——以POI数据和手机数据为例

王崑阳[1] 谯博文[2] 张 晶[2*]

(1. 广岛大学综合科学研究科,广岛 7398521; 2. 首都师范大学三维重点实验室,北京 100048)

摘要:京津冀一体化发展,是面向未来打造新型首都经济圈,实现京津冀优势互补,促进环渤海经济区发展的重大国家战略.为此,充分把握当前京津冀都市圈的城市体系现状是非常必要的.本文选取京津冀都市圈内6类城市基础设施的信息点数据和手机位置服务交通交互数据,从资源和交互2个方面对当前京津冀的城市体系现状展开研究;以土地利用-交通相互作用模型为理论背景,采用均衡性评价、聚类和社区发现等多种方法,深度分析了京津冀地区城市体系结构.研究结果表明:当前京津冀地区的城市基础设施资源高度集中在北京和天津,以邢台和邯郸为代表的南部城市的基础设施资源严重不足;人口交互上呈现出北京、天津双中心的结构,且北京具有更高的交互强度;石家庄具有较丰富的城市基础设施资源,却在交通层面表现不足,在网络结构中的核心地位与层次性不明显.京津冀南部地区在城市体系中处于薄弱状态,加强石家庄对于南部地区的影响力,可以有效地为京津冀一体化提供助力.

关键词:POI数据;位置服务;京津冀;城市体系;交通流

中图分类号:K901 **DOI**:10.19789/j.1004-9398.2020.06.011

0 引 言

城市体系是由一系列不同等级规模、不同职能分工和相互密切联系的城市组成的系统[1],其结构反映了城市群内部的空间特征[2].城市体系将一定范围内的城市视为一个具有层次等级,且不断改变的整体[3].区域一体化是城市体系的一种发展形式,是指区域内的2个或多个城市突破行政区划制约,使发展要素和资源在彼此之间自由流动和优化配置,促进社会经济紧密融合,形成优势互补、共同繁荣的整体效应[4],如京津冀一体化.

城市基础设施是城市发展的支柱[5],基础设施的数量反映了一个地区的发展规模.城市基础设施的建设程度在一定程度上代表了某一区域的城市发展等级.分析京津冀地区城市基础设施的分布,对全面了解京津冀地区现状与未来发展趋势具有重要的意义.各类基础设施在区域内的分布特征反映了其空间分布上的公平性与异质性[6-8],可以体现城市之间的功能差异[9-11].同时,城市基础设施的数量必须考虑该地区的人口数量,城市基础设施的建设要以人口为基础,人口的数量反映了该城市对于基础设施的需求量.单独对人口或城市基础设施进行讨论是片面的,并不能反映城市的实际情况,将城市基础设施数量与人口数量结合讨论是必要的.通过对比城市基础设施与人口的关系,可以确定该区域是否存在基础设施供不应求的情况,而基础设施的供需不平均是造成人口迁徙的原因之一[12].

另一方面,交通网络是制约区域内各城市之间各类供需关系的重要因素[13].随着技术的发展,手机的位置移动数据逐渐被应用到人口与城市的研究中[14].通过手机内置的基于位置服务(location based services,LBS)功能,定位设备的所在位置,再对比不同时刻设备所在位置的变化,从而获取该设备在空间上的位置移动信息[15].相比于路网,手机数据所反映的交通信息更客观地描述了实际的交通量;相比于基于公路流、铁路流的研究方法,手机数据可以更加全面地获取到一定范围内的人口移动信息,并且拥有更高的时空精度[16-17].手机数据

收稿日期:2020-02-25

*通信作者:zhangjing5946@sina.com

可以同时涵盖多种交通方式所产生的位置移动信息,更加全面和直观地反映不同城市之间的关联强度.但目前利用手机移动数据对京津冀地区的研究尚少.

土地利用-交通相互作用模型(land use-transport interaction model, LUTI)是模拟城市发展过程的常用模型,主要用来研究土地利用决策对社会活动和交通发展的影响,或交通决策对社会活动和城市空间分布的影响.LUTI 模型侧重城市内部结构的变化,对于空间公平性和社会群体的空间聚集分析较为薄弱[18].信息点(point of information,POI),即描述信息的点,其包含该点的坐标位置及其他所需要的信息,可以充分反映城市内部的功能区差异,且城市基础设施 POI 的分布模式、分布密度在城市空间分析中具有重要意义[19].而手机数据则可以提供必要的交通信息.基于这些考虑,本文提出设想,使用 POI 数据作为城市土地利用信息,使用手机数据作为交通信息.在这个层面上,通过研究城市间的土地利用与交通,研究城市之间的城市体系.

京津冀包括北京、天津、石家庄、唐山、秦皇岛、邯郸、邢台、保定、张家口、承德、沧州、廊坊和衡水,共计 13 个城市.2014 年以前,京津冀地区的发展缺乏统一规划,区域一体化进程缓慢、结构趋同现象严重[20-22].随着雄安新区进入国家规划,京津冀一体化发展的需求空前盛大.对京津冀地区的城市体系现状进行研究,有利于把握当前京津冀都市圈内各城市的发展特征,对未来规划具有重要的参考价值.本研究以土地利用-交通相互作用模型为基本思路,利用京津冀地区 6 类城市基础设施 POI 数据作为城市土地利用属性,手机 LBS 数据作为交通属性.本研究旨在通过 POI 和人口数据研究当前京津冀地区以医疗、学校、银行、超市商场、餐饮和宾馆酒店为代表的城市基础设施的空间分布特征及其空间公平性;通过手机数据研究各城市在网络结构下的空间交互强度以及依附关系;通过对比 POI 和手机数据,挖掘京津冀地区资源较为丰富、空间交互较弱,具有发展潜力的地区,为京津冀一体化发展提供参考与建议.

1 数据来源与研究方法

1.1 数据来源

本文的 POI 数据来自百度地图开放平台(https://lbsyun.baidu.com/),获取了截至 2015 年 6 月的 POI 数据,医疗、学校、银行、超市商场、餐饮和宾馆酒店分别为 14 651、23 150、25 129、51 896、39 312 和 10 187 条.每个城市基础设施 POI 包含 4 方面的信息:设施名称、设施类别、经度和纬度.利用这些数据点的空间分布特征,可以反映出区域内不同城市之间基础设施的空间分布关系.

手机数据是来自百度迁徙所属的人口移动大数据项目.本文采集了 2015 年 4 月 11—26 日期间的 16 个自然日(10 个工作日和 6 个休息日)京津冀地区 13 个城市之间逐小时人口流动数据. 16 个自然日采集的样本总量超过 200 万条,日均超过 12 万条,从而保证了本文推断的有效性,其中工作日累计样本量约 123 万条,休息日累计样本量约 82 万条.

1.2 研究思路与方法

1.2.1 均衡性评价

本文所采取的评价方法为变异系数(coefficient of variation, C_V)[23]和赫芬达尔指数[24](herfindahl-hirschman index, HHI). C_V 也称离散系数,反映了地区均衡度,其值越小,说明该区域越均衡,离散程度越小,公式为

$$C_V = \frac{\sigma}{\mu},$$

式中 σ 表示标准差,μ 表示平均数. HHI 经常用来衡量产业集中度,公式为

$$I_{HH} = \sum_{i=1}^{N} (X_i/X)^2 = \sum_{i=1}^{N} S_i^2,$$

式中 X_i 表示某地区某类 POI 数量,X 表示一定空间范围内的同类 POI 总数,$S_i = X_i/X$ 表示某地区的某类 POI 占有率,N 表示地区总数. HHI 反映区域规模指标的聚集程度,数值越大说明越集中,当区域属于完全垄断状态时,则 HHI = 1.

另外,为了直观地反映城市内资源与人口的关系,本文设计了城市基础设施资源总数与常住人口总数的比值,对城市的 POI 分布及供需情况进行评价,简称资源-人口比例,具体方法为

$$R = \frac{M'/M}{P'/P},$$

式中 R 表示某区域内某类资源的资源-人口比例,M'表示该区域某类 POI 数量,M 表示京津冀同类 POI 总数,P'表示该区域常住人口数量,P 表示京津冀常住人口总数.

1.2.2 聚类算法

聚类是将具有相似属性的样本聚到一个类中

的数据挖掘方法,旨在识别多变量数据中的离散模式[25].该方法通过样本之间的相似性进行分组,通过树状图反映各要素之间的结构与关系.

1.2.3 优势流方法

优势流方法是一种网络简化方法,该方法参照城市规模,依据某一个城市主要要素流的流向,将这个城市归属到对其最具控制力的中心[26].一般在一定空间范围内某城市所处层级水平越高,表明其中心性、主导性越强,对周边城市或地区的控制力和影响力越大,服务范围越广.

1.2.4 社区发现算法

社区发现算法用于发现网络中的社区结构[27].在社交网络中,每位用户都相当于一个点,用户之间通过互相的关注关系构成了整个网络的结构,在这样的网络中,有的用户之间连接较为紧密,有的用户之间连接关系较为稀疏,连接较为紧密的部分可以被看成是一个社区,其内部的节点之间有较为紧密的连接,而隶属不同社区的节点间的连接则较为稀疏.

社区发现是一个复杂而有意义的过程,用于探测并揭示不同类型复杂网络中固有的社区结构[27].其中,Fast Unfolding 算法是基于模块度对社区进行划分的算法,通过不断划分社区,使划分后整个网络的模块度不断增大.划分后的网络模块度越大,说明社区划分的效果越好.Fast Unfolding 算法通过反复运算使模块度增大,并根据运算结果对网络进行重构,直到网络结构不再改变[28],最终聚合成围绕少数中心点形成的网络,这与城市体系中的单中心、双中心、多中心城市结构很相似.因此该算法非常适用于挖掘城市之间的潜在关系. Fast Unfolding 算法公式为

$$Q = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{k_i \times k_j}{2m}\right]\delta(C_i, C_j),$$

式中:Q 表示整个社区的模块度;m 表示该网络中的权重总和,A_{ij}表示节点 i 和节点 j 之间的权重;k 表示与顶点连接的边的权重;C_i 和 C_j 分别代表节点 i 和节点 j 被分配到的社区;δ 用于判断顶节点 i 和节点 j 是否被分配的同一社区.

1.2.5 反距离权重法

反距离权重法假定每个测量点都有一种局部影响,而这种影响会随着距离增大而减弱[29],这很符合人们在对城市基础设施进行出行选择时的就近原则.因此本文使用反距离权重法对 POI 和手机数据所反映的城市连通度进行插值.

2 基于公共服务的京津冀地区城市体系分析

2.1 各地区 6 类城市基础设施分布情况及均衡性分析

为了能够直观地查看各地区城市基础设施的分布情况,本文将各地区 6 类 POI 的数量进行了类型内对比,以获得各地区 6 类 POI 在京津冀都市圈内所占该类 POI 总量百分比,如表 1 所示.同时,计算了 6 类城市基础设施 POI 在不同地区的资源-人口比,如表 2 所示.

表 1　6 类城市基础设施 POI 在不同地区的比例和不同地区常住人口量比例　　单位:%

城市	医疗	学校	银行	超市商场	餐饮	宾馆酒店	总量	常住人口
北京	32.4	26.7	30.2	34.9	50.4	44.9	37.1	19.8
天津	15.3	13.1	19.8	17.4	14.6	9.9	15.8	14.1
石家庄	10.3	11.0	8.9	9.8	6.5	6.5	8.9	9.2
唐山	6.7	6.3	6.4	6.3	5.6	5.5	6.1	7.1
秦皇岛	3.7	3.5	3.1	3.9	4.0	8.2	4.0	2.8
邯郸	4.2	5.9	4.6	2.8	1.7	2.1	3.3	8.6
邢台	2.4	3.7	3.5	1.7	1.2	1.2	2.2	6.7
保定	9.7	8.7	6.4	10.2	6.1	7.3	8.2	9.5
张家口	3.1	4.7	3.5	2.0	1.6	3.4	2.7	4.0
承德	2.9	4.6	3.0	1.5	1.5	3.8	2.5	3.2
沧州	2.6	4.0	4.4	2.5	1.4	2.1	2.7	6.8
廊坊	4.8	4.5	3.8	5.6	4.4	4.2	4.7	4.2
衡水	1.9	3.3	2.4	1.4	1.0	0.9	1.8	4.0

表 2 6 类城市基础设施 POI 在不同地区的资源-人口比及总量排名

城市	医疗	学校	银行	超市商场	餐饮	宾馆酒店	总量	总量排名
北京	1.64	1.35	1.53	1.76	2.55	2.27	1.87	1
天津	1.09	0.93	1.40	1.23	1.04	0.70	1.12	3
石家庄	1.12	1.20	0.97	1.07	0.71	0.71	0.97	5
唐山	0.94	0.89	0.90	0.89	0.79	0.77	0.86	7
秦皇岛	1.32	1.25	1.11	1.39	1.43	2.93	1.43	2
邯郸	0.49	0.69	0.53	0.33	0.20	0.24	0.38	12
邢台	0.36	0.55	0.52	0.25	0.18	0.18	0.33	13
保定	1.02	0.92	0.67	1.07	0.64	0.77	0.86	6
张家口	0.78	1.18	0.88	0.50	0.40	0.85	0.68	9
承德	0.91	1.44	0.94	0.47	0.47	1.19	0.78	8
沧州	0.38	0.59	0.65	0.37	0.21	0.31	0.40	11
廊坊	1.14	1.07	0.90	1.33	1.05	1.00	1.12	4
衡水	0.48	0.83	0.60	0.35	0.25	0.23	0.45	10

从表 1 可知,在 6 类城市基础设施中,北京均处于优势地位,其中医疗 POI 占京津冀地区总量的 32.4%,餐饮 POI 占 50.4%,宾馆酒店 POI 占 44.9%,其 6 类 POI 总量占京津冀地区 POI 总量 37.1%. 天津市 6 类 POI 均在京津冀地区排第 2 位,北京和天津两地 6 类 POI 相加可以发现,银行、超市商场、餐饮和宾馆酒店 4 类 POI 之和均已超过京津冀地区 6 类 POI 总量的 50%,而常住人口之和为 33.9%. 从表 2 可知,6 类 POI 在北京、天津的资源-人口比远高于其他地区,其 POI 总量的资源-人口比例分别位于第 1 位和第 3 位,以城市基础设施提供者的身份向周边城市提供 6 类城市基础设施服务. 秦皇岛的资源-人口比例虽然位居第 2 位,但这主要是因为其旅游城市的特性,流动人口高、常住人口少,并不能认为其具有向周边城市提供城市基础设施服务的能力. 位居第 4 位的廊坊同样存在大量流动人口. 由此可见,大多数城市基础设施均分布在北京和天津,其城市基础设施公平性远高于其他城市,并向周边城市提供城市基础设施服务. 根据表 1 的数据,可以计算出 6 类 POI 在京津冀地区的 C_V 和 HHI. 如表 3 所示.

表 3 6 类 POI 在京津冀地区的变异系数和赫芬达尔指数

基础设施	变异系数	赫芬达尔指数
医疗	1.093	0.162
学校	0.841	0.127
银行	1.061	0.157
超市商场	1.224	0.183
餐饮	1.735	0.291
宾馆酒店	1.498	0.236
全类型	1.257	0.189

通过 C_V 与 HHI 之间的变化规律可以看出 6 类城市基础设施均处于过度集中或过度离散的状态. 餐饮与宾馆酒店这类生活服务类型 POI 在京津冀地区集中度较高,而医疗、学校、银行这类基础保障类型 POI 则相对较为平均.

本文还将各地区 6 类 POI 的数量进行了区域内对比,以获得各地区 6 类 POI 在京津冀都市圈内所占本地区 6 类 POI 总量百分比,如表 4 所示. 从 POI 总量中可以发现,邯郸、邢台、沧州、衡水所包含的 POI 总量均低于 4.0%,总比为 12.1%,但 4 市所包含的常住人口总量达到 26.1%,4 市资源-人口比例

表 4 不同地区 6 类城市基础设施 POI 在本地区的比例 单位:%

基础设施	北京	天津	石家庄	唐山	秦皇岛	邯郸	邢台	保定	张家口	承德	沧州	廊坊	衡水
医疗	7.6	8.5	10.1	9.6	8.0	11.2	9.5	10.4	10.1	10.4	8.3	8.8	9.6
学校	10.2	11.7	17.4	14.5	12.4	25.2	24.0	14.9	25.0	26.3	20.4	13.5	26.6
银行	12.4	19.1	15.3	16.0	12.0	21.3	24.3	12.0	19.9	18.8	24.7	12.4	21.2
超市商场	29.7	34.7	35.0	32.4	30.9	26.4	25.1	39.4	23.1	19.5	29.4	37.4	25.6
餐饮	32.6	22.1	17.6	21.9	23.9	12.0	13.6	17.8	14.0	15.3	12.5	22.4	13.8
宾馆酒店	7.5	3.9	4.6	5.6	12.8	3.9	3.5	5.5	7.9	9.7	4.7	5.5	3.2

为京津冀都市圈内的倒数 4 名,属于重度城市基础设施需求地,具有较低的城市基础设施公平性,因此应加强城市基础设施建设,或加强与周边基础设施丰富的城市之间的互动. 而张家口与承德 2 地虽然 POI 总量 <4.0%,但由于其常住人口量较少,城市基础设施需求量低. 在河北省的各城市中,医疗和学校 POI 在石家庄以及保定比较集中,且石家庄略高于保定.

从表 4 可知,超市商场 POI 在京津冀地区均占较高比例. 北京、秦皇岛、廊坊、天津和唐山的餐饮占据较高比例. 由于秦皇岛是一个旅游城市,宾馆酒店的占比为京津冀地区最高,而承德、衡水、邯郸和张家口则以学校所占比例较高.

2.2　6 类城市基础设施 POI 的空间分布密度

为了能够更直观地反映出 6 类城市基础设施的分布情况,本文计算了 6 类 POI 在各城市的点密度. 点密度体现了每个输出栅格像元周围的点要素的密度. 从概念上讲,每个栅格像元中心的周围都定义了一个邻域,将邻域内点的数量相加,然后除以邻域面积,即得到点要素的密度(见图 1).

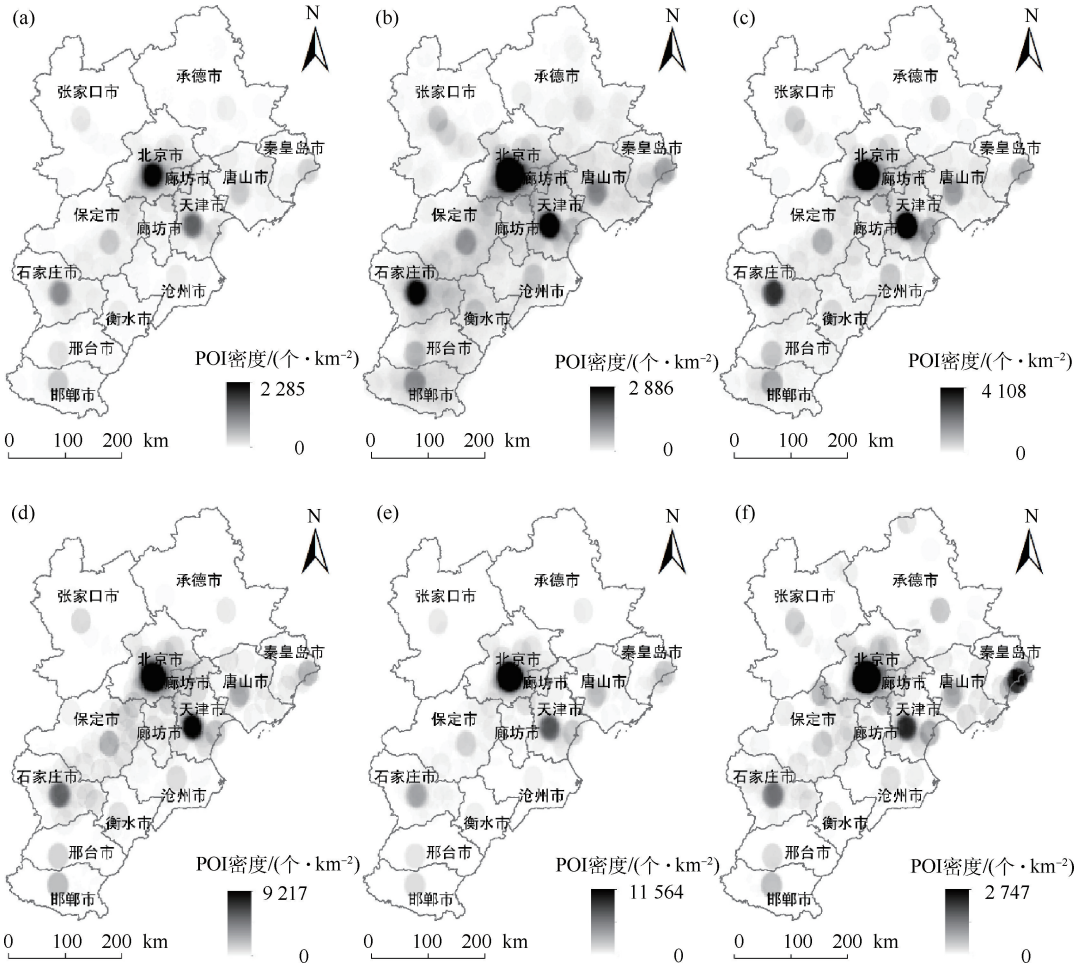


图 1　京津冀地区 6 类 POI 密度分布

(a)医疗;(b)学校;(c)银行;(d)超市商场;(e)餐饮;(f)宾馆酒店

从图 1 可知,6 类 POI 在北京均呈现较为集中的形态,以医疗和餐饮最为密集. 学校、银行和超市商场这 3 类 POI 则在北京、天津和石家庄均有集中态势,且在各地均有较为密集的分布,其中学校 POI 相对最为平均,说明教育资源在各地区分布较为平衡. 宾馆酒店类则是除北京和天津外,在秦皇岛市呈现密集形态.

2.3　各城市 POI 总量聚类

根据各城市的 POI 数量,对京津冀都市圈内的城市进行了聚类(图 2). 根据 POI 的性质,聚类距离越近的地区其 POI 分布特征越相似,距离越远则表示其 POI 分布特征差异较大。图 2 反映了北京、天津两地在整个京津冀都市圈中具有较高的城市基础设施占有率. 值得注意的是,对于河北省内的城

市,石家庄、保定与其他城市的差异,体现了石家庄和保定在河北省内,相对于其他城市有较高的 POI 占有率,即具有更大的发展潜力以及带动周边城市发展的能力.

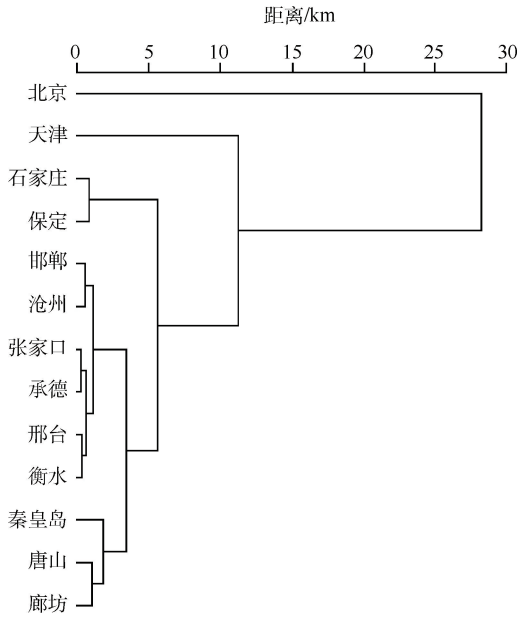


图 2　京津冀地区城市 POI 聚类

整体对比可知,北京、天津在 6 类 POI 的空间公平性上全面处于明显优势地位,对周边城市具有一定的领导作用. 位于第 2 梯度的则是石家庄和保定,这两地在河北省内的基础设施空间公平性最高. 而南部的邢台和邯郸不仅 POI 分布稀少,从距离上也难以利用到北京、天津的资源.

3　基于人口流量的京津冀地区城市体系分析

3.1　基于人口流量的京津冀社区划分

对京津冀都市圈各城市之间的 LBS 数据进行统计,得到共计 156 条不对称网络联系(图 3),通过优势流方法获得其中占主导地位的23 条主要人口流动线路. 对占主导地位的 23 条线进行二次分类,使用自然断裂法分为 5 级,断点类别基于数据中固有的自然分组. 将分类间隔加以识别,可对相似值进行最恰当地分组,并可使各个类之间的差异最大化. 从图 3 可知,北京-廊坊两地之间的人口交互强度最高,日均人口流量达到 36 839 人次. 北京-保定和北京-天津之间的人口交互强度次之,分别为 17 646 与 13 660 人次. 天津-廊坊、天津-唐山也以 10 188 与 7 207 人次保持在第 3 梯度. 由此可见,廊坊与北京和天津两地的人口流量都处于较高的水平,说明三地之间的空间交互程度较高. 整体上,除上述地区外,承德、张家口、石家庄、沧州与北京和天津两地的人口流量较大. 相对而言,河北省 11 个城市之间人口流量较小,均在日均434 人次以下.

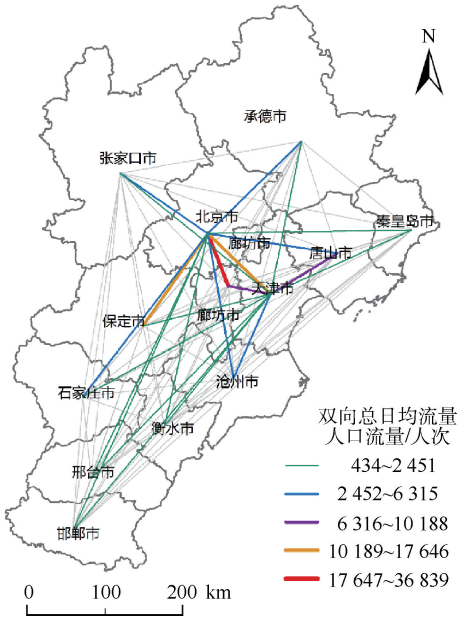


图 3　京津冀地区日均人口流量图(双向)

使用 Fast Unfolding 算法进行社区发现京津冀地区呈现双中心的结构(图 4). 其中,石家庄、保定、邢台、承德、邯郸、廊坊和张家口以北京为中心;而唐山、秦皇岛、衡水和沧州则以天津为中心.

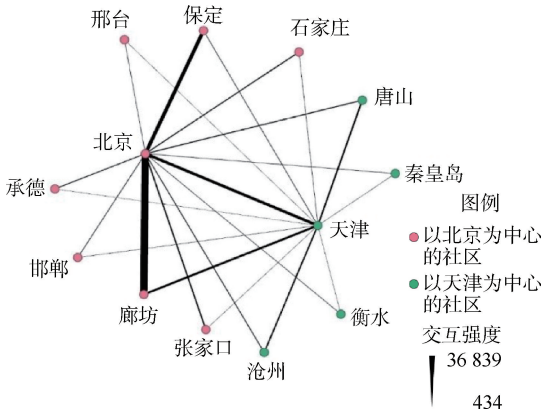


图 4　基于 Fast Unfolding 算法的网络结构下京津冀地区社区发现结果

3.2　网络结构下的中心城市差异

为了更直观地观察北京、天津 2 个中心城市的人口流量,制作了双向人口流量图,并使用对数函数对数据进行拉伸,使其更加显著地反映进出关系

(图 5),从图 5 可知,以北京和天津为中心的人口移动整体趋势为进京津方向人口流量大于出京津方向,且北京更为显著. 以北京为中心的城市体系中,除了廊坊,均明显呈现进大于出的趋势,仅秦皇岛差距微弱. 而天津为中心的城市体系中,沧州、唐山和保定进津方向明显多于出津方向,其他城市并无明显差异. 值得注意的是,廊坊市在以北京为中心的城市关系中,不同于整体进京大于出京的趋势,虽然差距微弱,出京却大于进京. 总体来说,北京市对周边城市的影响力大于天津市.

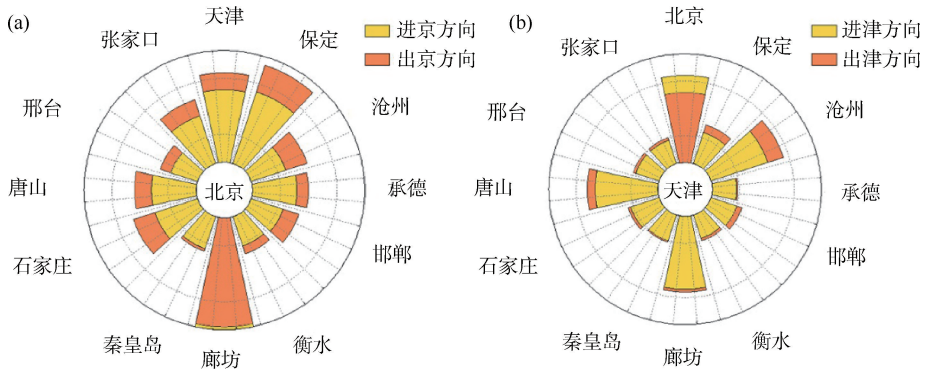


图 5　北京、天津进出双向日均人口流量示意图

(a)北京;(b)天津

4　POI 与人口流量所反映的城市热度差异

使用反距离权重法,计算各地区 POI 以及人口流量热度(图 6). 从图 6 可知,京津冀都市圈各城市 POI 热度与人口流量热度整体上很相似,但石家庄的 POI 热度与人口流量热度差异却很大. 石家庄虽然拥有丰富的基础设施资源,但其作为省会城市对周边城市的影响能力不足,在网络结构中的核心地位与层次性不明显. 实际交通交互量远低于其基础设施的承载能力,基础设施无法得到充分利用,其周边城市也难以利用到石家庄的城市基础设施资源. 再加上邢台、邯郸和北京、天津的距离较远,联系较弱,使京津冀南部地区在整个网络中处于薄弱状态. 石家庄丰富的城市基础设施资源和薄弱的交通交互量,使其具有成为京津冀都市圈内另一个中心城市的可能性.

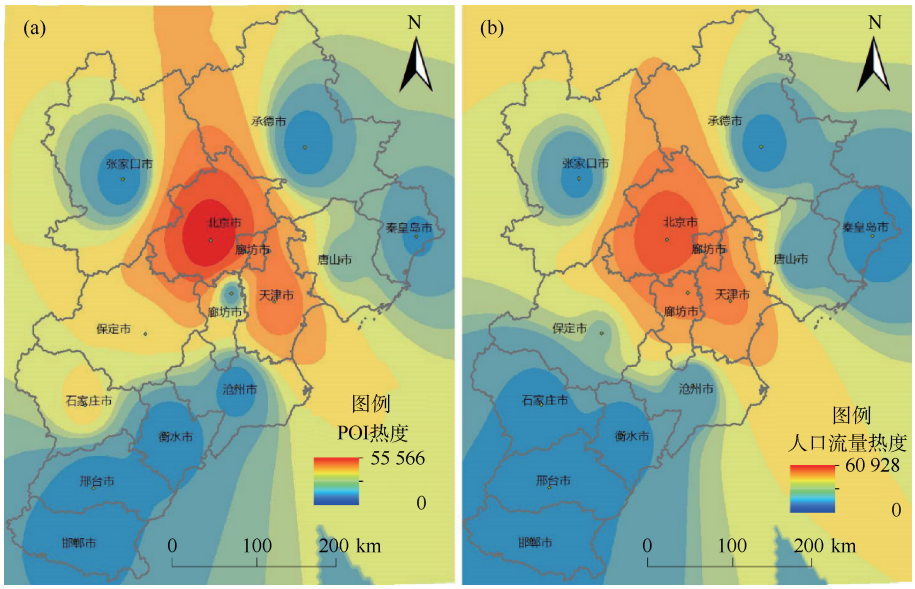


图 6　京津冀地区各城市 POI 热度与人口流量热度

(a)POI 热度;(b)人口流量热度

5 结论和讨论

本文证明了使用 POI 数据与手机数据替代 LUTI 模型中原本的土地利用和交通信息进行城市体系研究具有可行性,且二者结论具有一致性. 北京、天津和石家庄的城市基础设施资源远远比京津冀都市圈内的其他城市丰富,具有向周边城市提供城市基础设施服务的能力;围绕北京和天津形成的双中心网络结构可以有效带动北部区域发展,但是南部地区当前并没有中心城市,并且与现有中心城市的交互较弱;位于南部的石家庄虽然现在与周边城市的交互较弱,但具有丰富的城市基础设施资源,有较高的发展潜力. 伴随雄安新区的设立,京津冀地区将从双中心变为多中心的形态,配合保定的城市基础设施资源,可以很好地带动包括沧州和衡水在内的京津冀中部地区发展,而南部城市群依然处于薄弱的地位. 基于本研究的结果,在京津冀后续发展计划中,加强发展石家庄作为省会城市的优势,可以更好地带动南部城市群发展,为京津冀一体化发展提供助力. 当然,城市体系的研究还有很多方面,本文只是从公共服务和交通联系 2 个角度对京津冀都市圈的城市体系进行了探究,未来还可从更多角度进行研究,更好地为京津冀一体化发展提供参考和依据.

参 考 文 献

[1] HALL P, MARSHALL S, LOWE M. The changing urban hierarchy in England and Wales, 1913—1998[J]. Regional Studies, 2001,35(9):775 - 807.

[2] IOANNIDS Y M, OVERMAN H G. Spatial evolution of the US urban system[J]. Journal of Economic Geography, 2004,4(2):131 - 156.

[3] CAPELLO R, CAMAGNI R. Beyond optimal city size: an evaluation of alternative urban growth patterns[J]. Urban Studies, 2000,37(9):1479 - 1496.

[4] 王维, 于洁. 长江经济带社会-经济-资源-环境协调机理及其时空格局[J]. 首都师范大学学报(自然科学版), 2019,40(5): 73 - 83.

[5] 张秉忱, 石楠,金晓春. 城市基础设施是城市发展的支柱[J]. 城市规划, 1989(1):20 - 24.

[6] 杨胜利, 段世江. 京津冀医疗卫生资源配置的公平性[J]. 河北大学学报(哲学社会科学版), 2016,41(2): 63 - 69.

[7] 董香书. 京津冀医疗资源协同发展的实现路径[J]. 首都经济贸易大学学报, 2016,18(4):20 - 27.

[8] 赵梓渝, 杨冉, 魏冶, 等. 中国人口流动重力模型回归系数的定量估算及其空间异质性研究[J]. 地理科学, 2018,38(9): 1439 - 1448.

[9] 孟斌. 北京城市居民职住分离的空间组织特征[J]. 地理学报, 2009,64(12):1457 - 1466.

[10] 沈颖洁, 韩宝睿, 马健霄. 大都市近郊通勤交通需求管理策略[J]. 重庆交通大学学报(自然科学版), 2014,33(6):118 - 121 + 141.

[11] 申悦, 柴彦威. 基于 GPS 数据的北京市郊区巨型社区居民日常活动空间[J]. 地理学报, 2013,68(4):506 - 516.

[12] 刘小丽. 区域统筹视角下城市基础设施规划的编制与思考:以广东省城市基础设施建设"十三五"规划为例[J]. 城市地理, 2018(2):50.

[13] 吴威, 曹有挥, 曹卫东, 等. 长三角地区交通优势度的空间格局[J]. 地理研究, 2011(12):2199 - 2208.

[14] 方可, 王磊, 陈佳. 基于 GIS 平台的城市人口数据应用模式探索:以武汉市为例[J]. 华中建筑, 2015,33(12): 83 - 88.

[15] 周晓津, 姚阳. 基于大数据的京沪人口流动流量、流向新变化[J]. 大数据, 2016,2(3):49 - 59.

[16] 朱桃杏, 吴殿廷, 马继刚, 等. 京津冀区域铁路交通网络结构评价[J]. 经济地理, 2011,31(4):561 - 565 + 572.

[17] 傅毅明, 赵彦云. 基于公路交通流的城市群关联网络研究:以京津冀城市群为例[J]. 河北大学学报(哲学社会科学版), 2016,41(4):91 - 100.

[18] 牛方曲. LUTI 模型的概念结构、实现方法及发展趋势[J]. 地理科学, 2017,37(1):46 - 54.

[19] 禹文豪, 艾廷华. 核密度估计法支持下的网络空间 POI 点可视化与分析[J]. 测绘学报, 2015,44(1):82 - 90.

[20] 李兴国, 刘艳芳. 京津冀地区城市层级结构及协同发展战略[J]. 国土与自然资源研究, 2014(6):1 - 2.

[21] 王海龙, 朱洪兴. 京津冀地区的城市引力研究[J]. 企业经济, 2010(3):97 - 100.

[22] 孙铁山, 李国平, 卢明华. 京津冀都市圈人口集聚与扩散及其影响因素:基于区域密度函数的实证研究[J]. 地理学报, 2009,64(8):956 – 966.
[23] ADBI H. Coefficient of variation[J]. Encyclopedia of Research Design, 2010, 1: 169 – 171.
[24] RHOADES S A. The Herfindahl-Hirschman index[J]. Federal Reserve Bulletin, 1993, 79: 188.
[25] 孙吉贵, 刘杰, 赵连宇. 聚类算法研究[J]. 软件学报, 2008, 19(1):48 – 61.
[26] 魏冶, 修春亮, 刘志敏, 等. 春运人口流动透视的转型期中国城市网络结构[J]. 地理科学, 2016,36(11): 1654 – 1660.
[27] 刘大有, 金弟, 何东晓,等. 复杂网络社区挖掘综述[J]. 计算机研究与发展, 2013, 50(10):2140 – 2154.
[28] BLONDEL V D, GUILLAUME J L, LAMBIOTTE R, et al. Fast unfolding of communities in large networks[J]. Journal of Statistical Mechanics:Theory and Experiment, 2008(10): P10008.
[29] BARTIER P M, KELLER C P. Multivariate interpolation to incorporate thematic surface data using inverse distance weighting(IDW)[J]. Computers & Geosciences,1996, 22(7): 795 – 799.

Actuality of Urban Hierarchy in Beijing-Tianjin-Hebei Region from the Perspective of Big Data

——An Example of POI and Smartphone Data

WANG Kunyang[1]　QIAO Bowen[2]　ZHANG Jing[2]

(1. Graduate School of Integrated Arts and Sciences, Hiroshima University, Hiroshima 7398521;
2. 3D Information Collection and Application Key Lab of Education Ministry, Capital Normal University, Beijing 100048)

Abstract: Beijing-Tianjin-Hebei integration is a goal to build a new capital economic circle for the future. It is a major national strategy to realize the complementarity of the internal advantages of Beijing-Tianjin-Hebei and promote the development of the Circum-Bohai Sea Economic Zone. To this end, it is necessary to fully study the current status of the Beijing-Tianjin-Hebei urban hierarchy. For this reason, it is necessary to deeply study the current status of urban hierarchy in Beijing-Tianjin-Hebei metropolitan region. This study selected six types of point of urban infrastructure information data and smartphone location service traffic movement data among the Beijing-Tianjin-Hebei metropolitan region. This study analyzed the current status of the study area from the aspects of both resources and interaction. Adopted the land use-transport interaction model as the theoretical background, used varieties of methods including commensurability evaluation, clustering, and community discovery, to study the urban system structure situation. It is clearly showing urban infrastructure resources are highly concentrated in Beijing and Tianjin and seriously inadequate in Xingtai and Handan. The aspect of population interaction presents a dual-center structure of Beijing and Tianjin. Shijiazhuang is weakly at the transportation interaction, and its core and hierarchy relationship in the network structure is not obvious. The southern region of the Beijing-Tianjin-Hebei metropolitan region is weak in the entire network. Development of influence of Shijiazhuang in the southern region can effectively provide assistance for "Beijing-Tianjin-Hebei integration".

Keywords: POI data; location based services(LBS); Beijing-Tianjin-Hebei region; urban hierarchy; traffic