

长江中游城市群社团结构时空演化数据集 (2000–2014)

高 鹏¹, 何 丹^{2*}, 宁越敏³

1. 上海师范大学全球城市研究院, 上海 200234;
2. 华东师范大学城市与区域科学学院, 上海 200241;
3. 华东师范大学中国现代城市研究中心, 上海 200062

摘 要: 基于 11315 企业征信系统和国家企业信用信息公示系统挖掘长江中游城市群生产性服务业企业总部-分支数据, 由此构建长江中游城市群空间关联网, 并通过模块度、PageRank 值和冲积图等方法, 研究了 2000–2014 年长江中游城市群社团结构的时空演化特征, 并形成相应的数据集。数据集采用.xlsx 表格存储, 数据量为 311 KB, 包括了过程数据和结果数据。结果表明: (1) 长江中游城市群空间关联网逐渐形成分别以武汉、长沙和南昌为中心向外辐射的轮轴形态; (2) 长江中游城市群空间关联网形成武汉社团、长沙社团和南昌社团 3 个内部联系紧密的城市社团; (3) 各城市社团的网络地位在动态中不断调整; (4) 各城市社团之间的关系存在非均衡性和非对称性特征; (5) 行政分割特征明显, 各社团内部形成“核心-次核心-边缘”的拓扑结构。

关键词: 空间关联网; 城市社团; 时空演化; 长江中游城市群

DOI: <https://doi.org/10.3974/geodp.2022.01.15>

CSTR: <https://cstr.science.org.cn/CSTR:20146.14.2022.01.15>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志(中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2021.08.10.V1> 或 <https://cstr.science.org.cn/CSTR:20146.11.2021.08.10.V1>.

1 前言

城市群空间结构演化大致包括单中心城市主导、多中心城市间竞争以及网络化依赖与竞争三个发展阶段^[1]。其中, 城市群网络化是区域内各类资源要素动态流动的最高表现形式, 也是城市群形成发育过程中理想的城市化模式^[2,3]。基于网络和联系视角的城市群, 与经济地理学中“功能区”的概念类似^[4], 其空间组织形式强调地域边界以内的城市相互作用强度比边界外的更为紧密, 其空间形态表现为特定区域内以中心城市为核心向周围辐射

收稿日期: 2021-09-09; 修订日期: 2021-11-13; 出版日期: 2021-03-25

基金项目: 教育部人文社会科学重点研究基地重大项目(17JJD790007); 上海市哲学社会科学项目(2021WQQ001)

*通讯作者: 何丹, 华东师范大学城市与区域科学学院, dhe@re.ecnu.edu.cn

数据引用方式: [1] 高鹏, 何丹, 宁越敏. 长江中游城市群社团结构时空演化数据集(2000–2014)[J]. 全球变化数据学报, 2022, 6(1): 111–117. <https://doi.org/10.3974/geodp.2022.01.15>. <https://cstr.science.org.cn/CSTR:20146.14.2022.01.15>.

[2] 高鹏, 何丹, 宁越敏. 长江中游城市群社团结构时空演化数据集(2000–2014)[J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.08.10.V1>. <https://cstr.science.org.cn/CSTR:20146.11.2021.08.10.V1>.

构成的多个城市的集合体。城市之间的竞争在很大程度上转化为城市群之间的竞争^[5]；而在城市群内部，联盟与竞争关系也日趋复杂化。社会网络、复杂网络分析给地理学带来新的概念和分析范式，为洞察城市群内部空间组织提供有力支撑。学者借助子群、派系等凝聚子群算法或社团发现算法，普遍发现城市群内部具有与其整体相似的“内部联系紧密、外部联系稀疏”的若干个城市集合体或城市社团（City Community）^[6-8]。2015年4月国务院批复通过的《长江中游城市群发展规划》明确指出，要将长江中游城市群打造成为中国经济新的增长极，并推动形成多中心、网络化发展格局。长江中游城市群作为一个由多个城市子群（包括武汉都市圈、环长株潭城市群、环鄱阳湖城市群）构成的跨省域巨型城市群，城市群社团化空间组织格局及其动态演化趋势是决定其能否持续健康发展的关键所在。因此，本数据集基于生产性服务业企业总部-分支数据，从城市网络的角度构建了长江中游城市群社团结构时空演化的数据集。该数据集可为研判和优化区域发展格局提供数据支持。

2 数据集元数据简介

《长江中游城市群社团结构时空演化数据集（2000-2014）》^[9]的名称、作者、研究区域、数据年代、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表1。

表1 《长江中游城市群社团结构时空演化数据集（2000-2014）》元数据简表

| 条 目 | 描 述 |
|-----------|---|
| 数据集名称 | 长江中游城市群社团结构时空演化数据集（2000-2014） |
| 数据集短名 | CommunityStructure_MRYR |
| 作者信息 | 高鹏, 上海师范大学全球城市研究院, geogaopeng@163.com 何丹, 华东师范大学城市与区域科学学院, dhe@re.ecnu.edu.cn 宁越敏, 华东师范大学城市与区域科学学院, ymning@re.ecnu.edu.cn |
| 研究区域 | 长江中游城市群面积为 31.7 万 km ² , 包括湖北、湖南、江西的 1 个副省级城市、27 个地级市和 3 个县级市, 共计 178 个县级地理单元 |
| 数据年代 | 2000-2014 |
| 数据格式 | .xlsx 数据量 311 KB |
| 数据集组成 | 2000-2014 年长江中游城市群的空间关联网络矩阵数据、模块度数据、PageRank 值数据、城市社团划分数据、各城市社团间联系强度数据等 |
| 基金项目 | 教育部人文社会科学重点研究基地重大项目 (17JJD790007); 上海市哲学社会科学规划项目 (2021WQQ001) |
| 出版与共享服务平台 | 全球变化科学研究数据出版系统 http://www.geodoi.ac.cn |
| 地址 | 北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所 |
| 数据共享政策 | 全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据集和通过《全球变化数据学报（中英文）》发表的数据论文。其共享政策如下：（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 ^[10] |
| 数据和论文检索系统 | DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS |

3 数据研发方法

3.1 数据来源

图 1 反映了数据库构建流程。首先，本研究利用 11315 企业征信系统¹提供按地区关键词查询功能，在地域研究单元范围内输入“子公司”“分公司”“办事处”等关键词，获取研究区内分支机构名称，同时得到其总部的企业名录。其次，登录国家工商行政管理总局的国家企业信用信息公示系统²，进行二次查询逐条确认并补充所需信息。最后，依据《国民经济行业分类标准（GB/T 4754—2011）》，对采集的样本企业经营范围进行归类处理，筛选出交通运输、仓储和邮政业、信息传输、软件和信息技术服务业、金融业、房地产业、租赁和商务服务业以及科学研究和技术服务业等六个行业的企业；根据得到的企业总部和分支机构的地理位置，仅保留企业总部和分支机构在异地的样本。事实上，正是在最近 20 年中，生产性服务业经历了快速增长，并成为区域网络形成的关键贡献者。因此，本研究按时间顺序对样本进行分类，分别筛选出 2000、2007 和 2014 年的样本企业。最终总共获得 11,564 个有效样本。

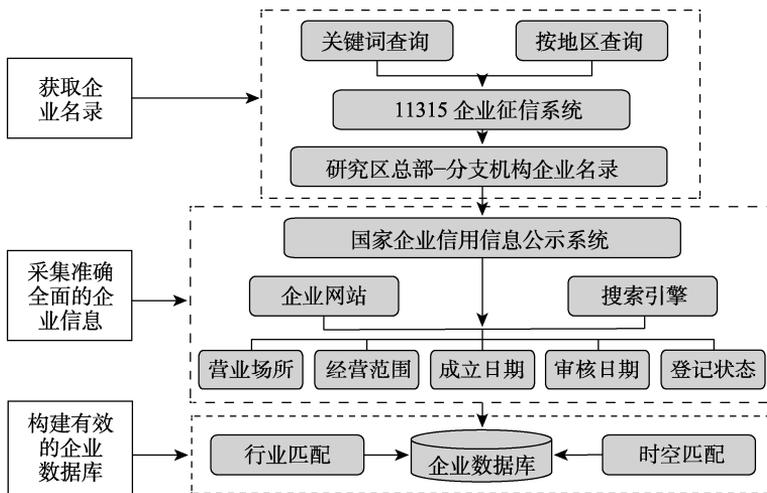


图 1 生产性服务企业数据库构建流程

3.2 技术路线与研究方法

首先，以有联系的空间单元为网络节点，提取生产性服务企业在节点间有总部-分支机构联系的边，分别构建 2000、2007 和 2014 年长江中游城市群关联网络；其次，利用社团发现算法中的模块度指标对长江中游城市群关联网络进行社团划分；再次，计算每个社团中各节点的 PageRank 值并据此绘制城市社团动态演化冲积图，随后考察社团之间的非均衡性和非对称性特征；最后，深入分析城市社团内部的结构特征和动态演化规律。

（1）模块度：城市网络中部分城市根据彼此间的连接性和联系强度会形成若干社团，

¹ 11315 企业征信系统. <https://www.11315.com/>.

² 国家企业信用信息公示系统. <http://www.gsxt.gov.cn/index.html>.

社团内节点之间联系相对紧密,而社团之间联系相对稀疏。Newman 等人定义了模块度^[11],用来定量描述网络中的社团,衡量网络社团结构划分的质量。

$$Q = \sum_{m=1}^n \left[\frac{l_m}{L} - \left(\frac{d_m}{2L} \right)^2 \right], m \leq n \quad (1)$$

式中, Q 为模块度,最大值和最小值分别为 1 和 0,该值越接近 1 表示对社团结构的划分质量越好; n 为社团数量; L 为网络中城市联系值; l_m 为第 m 社团内部的联系值; d_m 为社团 m 中与各节点相关联的联系值总和。

(2) PageRank 算法: PageRank 算法是一种用来对网络中节点的重要性排序的算法^[12],节点的 PageRank 值计算如下:

$$PR_i = (1-d)/n + d \sum_{i \in M_i} [w_{ij} \times PR_j / D_j] \quad (2)$$

式中, PR_i 为 PageRank 值; n 为网络中节点的数量; M_i 为与节点 i 相连的节点数量; w_{ij} 为节点 i 和节点 j 之间的联系; D_j 为节点 j 的中心度; d 为衰减因子,通常取 0.85。

(3) 冲积图: 社团结构的演化过程既包括各社团内部节点、关系和结构的变化,又包括社团间关系、位置的变化。Rosvall 等提出的冲积图 (alluvial diagram) 方法^[13]能够直观清晰地呈现出社团结构的演化过程。在冲积图中,各社团的名称由社团内部 PageRank 值最大的节点命名; 社团的位置代表其在网络中的地位,越靠近冲积图底部的社团其地位则越高。

4 数据结果

4.1 数据集组成

《长江中游城市群空间关联网络时空演化数据集 (2000–2014)》组成包括: (1) 长江中游城市群空间关联网络矩阵数据; (2) 长江中游城市群空间关联网络模块度数据; (3) 长江中游城市群空间关联网络城市社团划分数据; (4) 长江中游城市群空间关联网络 PageRank 值数据; (5) 长江中游城市群各城市社团间联系强度。该数据集以.xlsx 格式存储,数据量为 311 KB。

4.2 数据结果

(1) 长江中游城市群空间关联网络逐渐形成分别以武汉、长沙和南昌为中心向外辐射的轮轴状形态。伴随着武汉、长沙和南昌生产性服务企业总部的增加,与周边城市的联系逐渐增强,考察各城市的首位联系城市可以更清楚地看到,首位联系包括三个城市的边数由 2000 年的 45 条猛增到了 2014 年的 117 条,且基本形成了以三个城市为中心向外辐射的轮轴状结构特征。但这从侧面反映出,除三个城市以外,其余众多城市之间的联系非常微弱,尤其是跨省级行政边界的联系更弱,这将在后文中做进一步探讨。

(2) 长江中游城市群空间关联网络形成 3 个内部联系紧密的城市社团。一般而言,如果模块度达到 0.3 以上,表明网络中社团结构较为明显。研究结果显示,当把 2000、2007 和 2014 年长江中游城市群空间关联网络的社团分裂次数设定为 3 次时,模块度达到相应年份的最大值并均超过 0.3,分别为 0.575、0.488 和 0.45,社团划分效果最佳,这意味着研究时段内长江中游城市群空间关联网络形成 3 个内部联系紧密的城市社团。此外,随着时间

的推移，模块度呈递减趋势，表明各城市社团内部结构略有松动，各城市社团之间的联系相对来说提升明显。

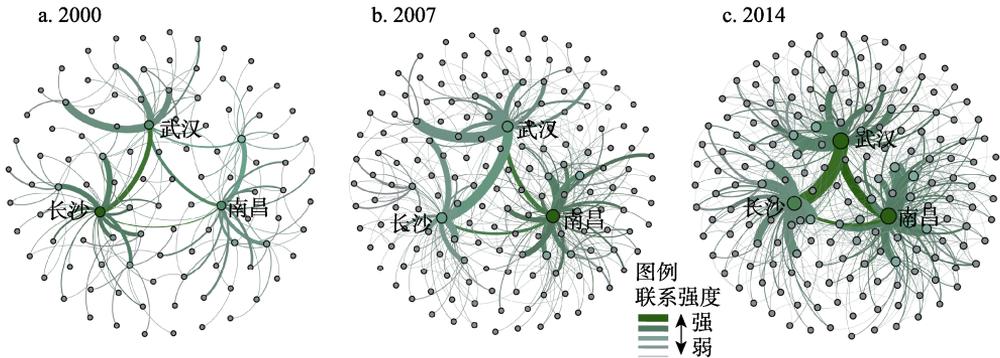


图 2 长江中游城市群空间关联网

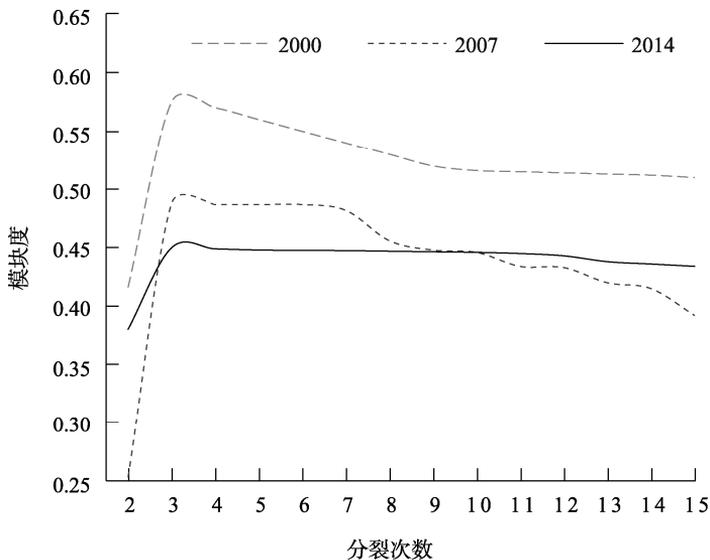


图 3 长江中游城市群空间关联网模块度

(3) 各城市社团的网络地位处于动态变化之中。以各城市社团中 PageRank 值最高的城市来命名社团名称，据此得到各年份的城市社团包括武汉社团、长沙社团以及南昌社团。紧接着，依托 Mapequation 平台绘制长江中游城市群社团结构时空演化冲积图^[4]，越靠近冲积图底部的城市社团，其网络地位就越高。2000–2007 年，除长沙社团中的少量成员分别成为另外两个社团的成员外，社团之间没有成员上的变动；南昌社团的对外连接性最高且在网络中地位愈趋重要，武汉社团反超长沙社团成为网络中对外连接性第二大社团。2007–2014 年，社团间的成员变动幅度较前期明显增加，其中南昌社团中有相当数量的成员演化为武汉社团的成员；武汉社团再次反超南昌社团成为对外连接性最高的社团。

(4) 各城市社团之间的关系存在非均衡性和非对称性特征。如图 4 所示，2000 年武汉社团–长沙社团的联系最为紧密，武汉社团–南昌社团次之，长沙社团–南昌社团的联系最弱；2007 年武汉社团–南昌社团成为联系最为紧密的一对社团，并且联系强度远超第二位

的武汉社团-长沙社团; 2014年武汉社团-长沙社团和武汉社团-南昌社团的联系强度相当, 社团间互动较为活跃, 且与长沙社团-南昌社团的差距进一步拉开。可见, 三个社团之间的关系存在显著的非均衡性特征。如果考虑到社团间的主导联系流向, 社团间的关系并非是对称的, 三个年份中武汉社团均为净流出社团, 而长沙社团和南昌社团始终是净流入社团, 表明武汉社团的对外经济辐射能力较强, 对其他两个城市社团产生经济外部性作用。

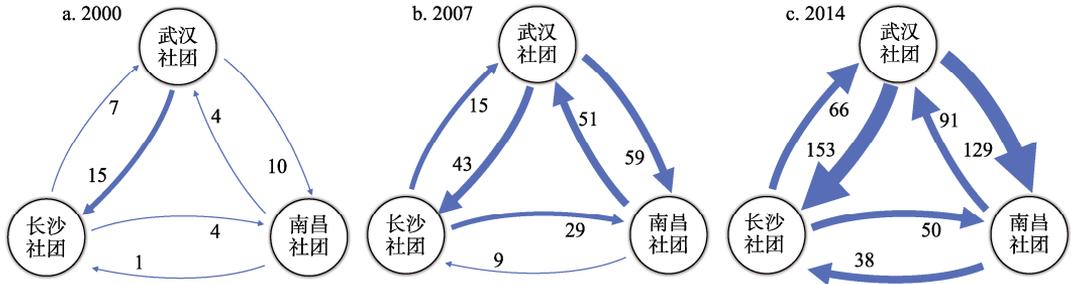


图4 长江中游城市群各城市社团之间的有向联系

(注: 图中的数字代表生产性服务企业的分支机构数量, 即城市社团之间的单向联系强度)

(5) 行政分割特征明显, 各社团内部形成“核心-次核心-边缘”的拓扑结构。进一步对各城市社团的内部结构进行研究发现^[14]: 一方面, 省级行政边界是影响社团结构演化的主导因素, 跨省域“组团”的现象越来越少见; 另一方面, 武汉、长沙、南昌分别是其所属社团中的核心城市, 在社团内部的空间组织中发挥主导作用, 紧密围绕核心城市分布的是市级节点, 属于各社团的次级核心, 是社团内部网络流通的重要媒介, 而各社团的边缘则分布着大量地理位置偏远、交通条件欠佳、经济基础薄弱的县级节点。

5 结论

本研究针对长江中游城市群社团化空间组织现象, 从企业网络的视角构建了长江中游城市群空间关联网络, 并通过模块度、PageRank 值和冲积图等方法, 分析了长江中游城市群社团结构的时空演化特征, 据此研发了《长江中游城市群社团结构时空演化数据集 (2000-2014)》, 该数据集不仅有助于从城市网络的角度形成对城市群空间结构和组织新的认识, 也可为区域协调发展政策的制定等提供基础数据和参考。数据结果显示: 研究时段内, 长江中游城市群空间关联网络联系日益密切, 形成分别以武汉、长沙和南昌为中心向外辐射的轮轴状形态。各年份长江中游城市群空间关联网络被分成武汉社团、长沙社团和南昌社团这3大城市社团, 且各城市社团在网络中的地位不断调整, 不同社团之间的关系存在非均衡性和非对称性特征。此外, 随着时间的推移, 跨省域“组团”的现象消失, 城市社团明显受到行政边界的分割, 各社团内部形成圈层式的等级结构。

本研究利用网络大数据挖掘企业总部-分支企业名录, 利用国家企业信用信息公示系统验证数据的准确性并进一步补充有效信息, 是一种企业网络大数据与官方数据相结合的有益尝试, 即体现了数据的时效性, 又保证了数据的可靠性。此外, 本研究还将社会网络分析与空间分析相结合, 从城市网络的角度揭示了城市群社团化空间组织现象, 可以为相关研究提供一条较为科学的实证分析路径。需要注意的是, 本数据集中企业总部-分支

数据是投资的企业数量数据，后续工作需要挖掘投资额度数据，以更加精准地构建空间关联网络。

作者分工：何丹和宁越敏对数据集的开发做了总体设计；高鹏和何丹采集、处理并核实了数据；高鹏进行了数据运算并撰写了数据论文。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

- [1] Scott, A. J. *Global City-Regions: Trends, Theory, Policy* [M]. New York: Oxford University Press, 2001.
- [2] 年福华, 姚士谋, 陈振光. 试论城市群区域内的网络化组织[J]. *地理科学*, 2002, 22(5): 568–573.
- [3] 潘峰华, 方成, 李仙德. 中国城市网络研究评述与展望[J]. *地理科学*, 2019, 39(7): 1093–1101.
- [4] Anderson, W. *Economic geography* [M]. New York: Routledge, 2012.
- [5] 张凡, 宁越敏, 娄曦阳. 中国城市群的竞争力及对区域差异的影响[J]. *地理研究*, 2019, 38(7): 1664–1677.
- [6] 方大春, 孙明月. 高铁时代下长三角城市群空间结构重构——基于社会网络分析[J]. *经济地理*, 2015, 35(10): 50–56.
- [7] Zhang, W. Y., Derudder, B., Wang, J. E., *et al.* Regionalization in the Yangtze River Delta, China, from the perspective of inter-city daily mobility [J]. *Regional Studies*, 2018, 52(4): 528–541.
- [8] Gao, P., He, D., Sun, Z. J., *et al.* Characterizing functionally integrated regions in Central Yangtze River Megaregion from a city-network perspective [J]. *Growth and Change*, 2020, 51: 1357–1379.
- [9] 高鹏, 何丹, 宁越敏. 长江中游城市群社团结构时空演化数据集(2000–2014)[J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.08.10.V1>. <https://cstr.science.org.cn/CSTR:20146.11.2021.08.10.V1>.
- [10] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/10.3974/dp.policy>. 2014.05(2017年更新).
- [11] Newman, M. E. Finding community structure in networks using the eigenvectors of matrices [J]. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, 2006, 74(3): 1–22.
- [12] Brin, S., Page, L. Reprint of: The anatomy of a large-scale hypertextual web search engine [J]. *Computer Networks*, 2012, 56(18): 3825–3833.
- [13] Rosvall, M., Bergstrom, C. T. Mapping change in large networks [J]. *PlosOne*, 2010, 5(1): 1–7.
- [14] 高鹏, 何丹, 宁越敏等. 长江中游城市群社团结构演化及其邻近机制——基于生产性服务企业网络分析[J]. *地理科学*, 2019, 39(4): 578–586.