

全球科研论文合作网络的结构异质性及其邻近性机理

刘承良^{1,2}, 桂钦昌¹, 段德忠^{1,2}, 殷美元¹

(1. 华东师范大学城市与区域科学学院, 上海 200241;

2. 华东师范大学科技创新与发展战略研究中心, 上海 200062)

摘要: 以科研论文为媒介的知识合作网络已成为知识溢出的重要通道,但目前学术界对全球科研合作网络结构的复杂性涌现机制缺乏深入的探讨。基于2014年Web of Science核心合集所收录的科研论文合著数据,借助大数据挖掘技术、复杂网络、空间统计和重力模型分析,刻画了全球科研论文合作网络的拓扑结构、空间格局及其邻近性机理。结果发现:① 拓扑结构上,形成了以美国为核心的层级网络,具有小世界性和等级层次性,发育出典型的等级“核心—边缘”结构。② 空间格局上,以美国、西欧、中国和澳大利亚为顶点的“四边形”成为全球科研论文合作网络的骨架;三大中心性指标值的空间分异明显,强度中心性形成以美国为极核,加拿大、澳大利亚、中国及西欧诸国为次中心的“一超多强”格局,与之类似的介数中心性呈现北美、西欧和东亚“三足鼎立”的形态,度中心性分布则相对均匀,表现出“大分散、小集中”的“多中心—边缘集散”格局。③ 重力回归分析发现,地理距离抑制了国际科研论文合作,不过其影响力较弱;社会与经济邻近性对全球科研论文合作具有明显的促进作用,语言差异不是国际科研合作交流的障碍。

关键词: 科研论文合作网络;结构异质性;复杂网络分析;重力模型;邻近性机制

DOI: 10.11821/dlxb201704014

1 引言

21世纪以来,科技全球化的深度和广度不断加强,科学正在走向全球,成为“全球化的科学”。国际合作发表论文的比例已从1995年的约25%飙升至2011年的35%左右,世界范围内的知识流动日益频繁,国际科研合作成为大势所趋^[1]。当前的科学研究已经从个人、机构和国家进入合作时代^[2]。国际合作已成为前沿科学发现的主导力量,催生了众多高质量的科研成果^[3]。与此同时,悄然兴起的全球科学网络正在改变二战后形成的大西洋轴心格局,中国、南非、印度和巴西等新兴科技力量迅速崛起,中东、东南亚和北非国家的科学强劲发展,加速推动世界向多极化发展,全球科技创新版图正在加速重构^[1,4]。科技创新成为各国打造核心竞争力的关键所在,深刻影响和改变了国家力量的对比,正在重塑世界经济格局。国家是参与全球竞争的重要主体,大力提升本国科技创新能力,抢占全球创新网络的核心位置,主导或引领未来的科学发展成为许多国家发展的出发点和落脚点。中国也概莫能外,在《“十三五”国家科技创新规划》发布之际,探讨国际科研合作网络具有重要的现实意义。

收稿日期: 2016-06-14; 修订日期: 2016-11-25

基金项目: 国家自然科学基金项目(41571123, 41471108) [Foundation: National Natural Science Foundation of China, No.41571123, No.41471108]

作者简介: 刘承良(1979-), 男, 博士, 副教授, 硕士生导师, 中国地理学会会员(S110009837M), 主要从事经济地理复杂性研究。E-mail: clliu@re.ecnu.edu.cn

通讯作者: 桂钦昌(1991-), 男, 硕士生, 主要从事交通与创新网络复杂性研究。E-mail: 970995302@qq.com

当今,知识流动日益频繁,区域性和全球性的科研合作网络正在涌现,其关联及演化的复杂性引起西方经济地理学者广泛关注。研究范畴论及专利引用^[5]、专利联合申请^[6]、合著科研论文^[7]、共同参与研发项目^[8-9]等,研究尺度涵盖个人^[10]、组织^[11]、区域^[12]、国家甚至全球^[13-14],研究内容则聚焦于科研合作网络的拓扑结构^[15-17]、空间格局^[18-19]、演化过程与影响机制^[20-26]等领域。

近十年,社会网络或复杂网络分析方法的兴起,为洞察科研合作网络的结构洞、中心性、集聚性、小世界性、无标度性、分形性等结构复杂性提供强大支撑。学者普遍发现科研合作网络具有尺度依赖性,全球尺度网络多具小世界性,具有高的聚集系数和短的平均路径长度;国家尺度多为无标度网络,网络节点度分布服从幂律函数;区域创新网络则更多地体现出随机网络特征^[15-17]。与此同时,知识合作网络在空间形态上发育出显著的“核心—边缘”结构和等级层次性^[18-19],呈现出集聚与分散的双重特征。一方面,由于信息通信和交通技术的变革,地理空间约束不断减弱(distance is dead),全球知识合作网络向均衡化演进(world is flat)^[20-21],显性知识通过图书、期刊、会议、科研人员流动等渠道进行传播,横跨城市、区域、国家等多种地理空间,具有全球扩散的趋势^[22]。另一方面,隐性知识的传播仍然受到距离衰减的影响,总是高度聚集于某些区域(如西欧、北美和亚太地区)、国家(如美国、英国、德国、加拿大、中国等)、城市(如纽约、旧金山、伦敦等),呈现钉子状的空间集聚特性,具有地方依赖性(place-dependence)^[23-25]。受到世界城市研究小组(GaWCgroup)的鼓励,一些学者采用SCI数据库来研究全球主要城市间的知识合作网络联系的不均衡性^[26-27]。研究发现,科研论文产出上以西欧和北美的城市为主导,但近年亚太地区的城市增速较快,开始挑战前者的统治地位;就国际合作而言,欧洲城市占据网络的主体,伦敦、巴黎和阿姆斯特丹扮演全球知识合作网络枢纽的角色^[26-27]。

此外,这种结构异质性背后的邻近性影响机制也被广泛审视。他们强调邻近性在创新过程中的作用,主要构建地理邻近性、认知邻近性、组织邻近性、制度邻近性和社会邻近性^[28]来分析知识创新活动的本地蜂鸣与全球通道^[29-31],以及邻近性与创新互动、创新绩效的关系^[32]。研究发现,地理邻近性越大,越有利于创新主体交流互动和隐性知识溢出,促进研发合作^[33-35];在知识合作网络中主要发挥间接作用,但有助于其他邻近性的形成与发展^[36]。认知邻近性反映了主体共享知识基础的程度^[37],是合作过程发生的前提条件,也是创新主体选择合作伙伴的决定因素^[38],但过度的认知邻近性不利于知识溢出,因为主体间知识过多重叠往往会导致“认知锁定”^[28]。相似地,组织邻近性也对知识合作具有促进作用,同属某一企业集团的公司之间更有可能发生合作关系,从而减少不确定性^[8, 28]。制度邻近性是指主体受到非正式约束和正式规则制约的相似性^[39-40]。研究发现,制度距离阻碍着科研合作^[9, 12, 35]。而社会邻近性,随着主体之间的测地线距离衰减,有利于形成闭合的知识合作网络^[38]。英国的专利合作研究表明,社会邻近性促进共同申请专利^[10]。也有研究发现,邻近的社会关系对科研合作的作用并不显著^[8]。

近年,国内学者对创新体系、创新联系、创新网络和知识网络展开了深入的研究,取得了丰硕的研究成果^[41-43],主要包括:①采用专利合作数据,探讨了特定产业(生物技术、装备制造业等)的技术合作特征^[44-45];②以论文合作为视角,分析了城市间的知识流动特征^[46-48];③以跨国公司R&D投资为例,着重揭示了全球研发网络的空间特征和创新溢出效应^[49-50];④基于空间作用视角,采用系列测度指标和引力模型,分析了城市间的创新联系强度与空间格局^[51]。成果普遍聚焦于中国实践,研究发现中国已形成以上海、北京为顶层的五级塔型创新城市体系,整体呈现“东强西弱”的创新联系格局^[45-46, 51]。

与此同时,知识合作网络或创新联系网络发育背后的邻近性机制也有所论及。对装备制造业技术合作机制研究,发现地理邻近性是驱动技术创新网络形成的基础和演化的首要因子,社会邻近性的作用逐渐增强,丰富了主体间的知识流动渠道^[52]。对生物技术领域的论文合著机制研究表明,地理邻近性和组织邻近性的相互作用共同推动着知识网络空间结构的演化^[53]。普遍认识到,随着知识网络发育日益成熟,地理距离的影响力逐渐下降,社会距离和组织距离的约束性得以加强^[17]。比较专利合作和论文合作,邻近性机制对后者具有更强的解释力,社会和组织邻近性的作用超过地理和认知距离^[16]。也有学者得出相反的结论,如Ma等指出地理邻近性在中国城市间科学合作网络中的作用并没有下降,而是逐渐增加^[54]。

综上,以科研论文和专利技术为主体的知识合作网络研究已成为经济地理学研究的热点和前缘。国内外的知识合作网络研究大多集中于特定产业、特定区域、特定学术索引,鲜有全球尺度的知识合作网络研究;基于复杂网络分析的网络拓扑结构异质性研究成为研究重心,但缺乏知识产出及合作规模的权重考虑,空间结构的等级层次性及其邻近性机制研究也有待加强。因此,本文基于2014年的全球科研论文合作数据,通过大数据挖掘科研论文的国际合作规模,构建基于论文发表和合作量加权的国际知识合作网络,采用复杂网络和空间统计方法,绘制科技全球化背景下的国际科研论文合作图谱,应用负二项式回归和重力模型验证科研合作网络与多维邻近性的作用机制,以尝试做出有益的补充。

2 数据来源和研究方法

2.1 数据来源

合著科研论文是科研合作最直接的体现形式,成为研究全球科研合作的主要途径之一。2015年自然出版集团发布《Nature Index 2015 Collaborations》报告,基于旗下68本期刊2014年刊文数据对全球94个国家间的科研论文合作网络首次进行了图示化分析。鉴于其论文数据范围和国家数量的局限性,本文进一步将数据源拓展至汤森路透(Thomson Reuters)的Web of Science(WOS)核心合集,囊括了211个国家(或地区)^①、所有学科领域的作者合著论文数据;为了便于比较分析,选择其2014年收录的2037461篇论文作为数据源。首先,通过Python数据爬虫,将论文作者的地址信息汇总至国家尺度,提取汇总出211个国家(或地区)的科研论文合作数据,保存为txt文本。然后,利用C++编程语言将各个国家(或地区)的合作文本生成关联矩阵,对角线单元格赋值该国的科研论文数量,其余单元格赋值两两国家作者论文合作的总数量。接着,在第一步基础上,剔除具多个国家(或地区)作者联合署名论文数据,获得各国国内独立论文发表量(表1,仅列论文发表量前15)。

其中,在洲际尺度,欧洲(7个)和亚太地区(5个)成为主要贡献者。在国家层面,美国“一家独大”,无论论文发表总数量、国内独立发表数量,还是合作国家数量和国际合作发表数量,均遥遥领先于其他国家,具有较高的首位度,其发文总量是第15位瑞士的15.9倍,是第2位中国的2.11倍。就国际合作比例而言,发达国家明显高于新兴经济体。其中,欧洲国家国际合作比例最高,普遍接近50%,远高于世界平均水平。特别是瑞士,尽管论文规模不大,但其论文国际合作比例高达68.26%。北美次之,美国与加

① 在Web of Science核心合集中,英国被划分为英格兰、威尔士、苏格兰、北爱尔兰;中国的数据包括中国大陆、香港和澳门;本文中的“国家”均应为“国家(或地区)”。

表 1 论文发表数量排名前 15 国

Tab. 1 Top 15 countries by published papers

排名	国家	合作国家 数量(个)	论文发表 数量(篇)	国内独立发表 数量(篇)	国际合作发表 数量(篇)	国际合作 比例(%)
1	美国	205	575379	400184	175195	30.45
2	中国	177	272116	204579	67537	24.82
3	英格兰	194	140114	68117	71997	51.38
4	德国	182	132970	65055	67915	51.08
5	日本	169	97161	70595	26566	27.34
6	法国	190	89868	40903	48965	54.49
7	加拿大	191	86567	44798	41769	48.25
8	意大利	181	82868	44645	38223	46.13
9	澳大利亚	189	74522	38661	35861	48.12
10	西班牙	180	71530	38330	33200	46.41
11	印度	170	66167	51142	15025	22.71
12	韩国	165	63831	46436	17395	27.25
13	荷兰	180	50038	21099	28939	57.83
14	巴西	169	47940	32468	15472	32.27
15	瑞士	177	36196	11490	24706	68.26

拿大论文合作比例介于 30%~50%。亚太地区（中国、日本与韩国）位居第 3，均低于 30%，国际化程度较低，以国内独立刊文为主。

2.2 研究方法

借鉴图论构造全球科研论文合作网络连通图 $G=(V, E)$, $N=|V|$ 为网络节点（国家或地区）数, $M=|E|$ 为网络边（国家间的论文合作关系）数。运用原始法, 将相同国家不同称谓进行整合, 每一个国家仅保留一个名称, 删除不与他国产生科研联系的节点, 以国家论文发表总量为节点权重, 以国家间的论文合作发表量作为边的权重, 构建加权无向科研论文合作网络。

2.2.1 网络中心性模型

(1) 度中心性 (degree centrality, C_D) 指与该节点直接相连的其他节点的个数, 表征节点连接程度。在科研论文合作网络中, 节点度表示与该节点产生论文合作关系的国家数量:

$$C_D(i) = \sum_{j=1}^N a_{ij} \quad (1)$$

式中: a_{ij} 表示国家科研合作邻接矩阵, 有科研合作则赋值为 1, 无则赋值为 0。

(2) 接近度中心性 (closeness centrality, C_C) 表示节点到其他所有节点最短路径之和的倒数乘其他节点个数。节点的接近度越大, 表明节点越居于网络的中心。在科研论文合作网络中, 节点接近度表示该国与其他国家之间的科研论文合作的欧式距离:

$$C_C(i) = \frac{N-1}{\sum_{j=1, j \neq i}^N d_{ij}} \quad (2)$$

式中: d_{ij} 表示节点 i 和 j 之间的最短路径数; N 表示节点个数。

(3) 介数中心性 (betweenness centrality, C_B) 是测量网络中所有最短路径中经过该点的数量比例。节点的介数值越大, 表明节点控制网络的能力越强。在科研论文合作网络中, 节点介数表示该国承担“中介”或“中转站”的能力:

$$C_B(i) = \sum_{j=1, k=1, j \neq k \neq i}^N \frac{N_{jk}(i)}{N_{jk}} \quad (3)$$

式中: N_{jk} 表示节点 v_j 和 v_k 之间的最短路径条数; $N_{jk}(i)$ 表示节点 v_j 和 v_k 之间的最短路径经

过节点 v_i 的条数。

(4) 强度中心性 (strength centrality, C_s) 是无权网络中节点度的自然推广。节点 v_i 的点权 s_i 定义为与它关联的边权重的总和。在科研论文合作网络中, 节点强度表示该国对外科研合作中, 两两科研论文合作数量的总和:

$$C_s(i) = \sum_{j \in N_i} w_{ij} \quad (4)$$

式中: N_i 表示节点 v_i 的相邻节点集合; w_{ij} 表示连接节点 v_i 和节点 v_j 的合著发表论文数量。

2.2.2 重力模型 鉴于国家间的论文合著数据为非负整数, 且被解释变量的方差明显大于期望, 存在“过度分散”。因此, 负二项式回归 (negative binomial regression) 方法被引入研究多维邻近性与科研论文合作网络的作用机制:

$$I_{ij} = \alpha + \beta_1 Mass_i + \beta_2 Mass_j + \beta_3 Geodistance_{ij} + \beta_4 Socproximity_{ij} + \beta_5 Ecoproximity_{ij} + \beta_6 Lanproximity_{ij} + \varepsilon_{ij} \quad (5)$$

式中: α 为常数项; ε_{ij} 为随机误差项; I_{ij} 表示国家 i 和国家 j 之间合作发表的论文数量, 也是本文的被解释变量; $Mass_i$ 、 $Mass_j$ 分别表示国家 i 和国家 j 发表的论文数量, 作为重力模型中国家质量的代理变量, 以上 3 个变量均来自 WOS 数据库; $Geodistance_{ij}$ 表示国家 i 和国家 j 之间的地理邻近性, 通过计算各个国家首都之间的球面距离而获得, 提取于法国 CEPII-GeoDist 数据库; $Ecoproximity_{ij}$ 是国家经济邻近性的虚拟变量, 如果两国同属世界银行的收入群组分类则赋值为 1, 否则为 0; $Lanproximity_{ij}$ 是国家间语言邻近性的虚拟变量, 如果两国使用相同的官方语言则赋值为 1, 否则为 0, 该变量来源于法国 CEPII-Language 数据库。社会邻近性本是测度合作主体之间的社会关系, 其值的大小反映了社会关系的亲疏。本文基于国家之间的相对合作强度, 构建杰卡德指数 (Jaccard index) 方法来计算^[7, 55], 其含义是两个集合 A 和 B 的交集元素在 A 与 B 的并集中所占的比例 (表 2), 其计算公式如下:

$$Socproximity_{ij} = \frac{I_{ij}}{C_s(i) + C_s(j) - I_{ij}} \quad (6)$$

式中: $C_s(i)$ 、 $C_s(j)$ 分别为国家 i 和 j 的强度中心性, 即国家 i 与所有国家的两两合作的总和。

3 拓扑结构异质性

3.1 集聚性

将国家间论文合作网络 (.net) 和国家分区 (不同颜色) 数据导入 VOSviewer 生成国

表 2 变量和数据来源

Tab. 2 Variable descriptions and data sources

变量名	变量描述	数据来源
I_{ij}	国家 i 和国家 j 之间合作发表的论文数量, 单位篇。	WOS
$Mass_i$	国家 i 发表的论文数量, 单位千篇。	WOS
$Mass_j$	国家 j 发表的论文数量, 单位千篇。	WOS
$Geodistance_{ij}$	国家 i 和国家 j 之间的地理邻近性, 采用各个国家首都之间的距离表示, 单位 1000 km。	法国 CEPII 数据库
$Socproximity_{ij}$	国家 i 和国家 j 之间的社会邻近性, 通过计算国家 ij 之间的杰卡德相似系数而获得。	-
$Ecoproximity_{ij}$	国家 i 和国家 j 之间的经济邻近性, 国家 ij 同为世界银行划分的收入群组, 计为 1, 否则为 0。	世界银行国家收入群组分类
$Lanproximity_{ij}$	国家 i 和国家 j 之间的语言邻近性, 国家 ij 使用共同的官方语言, 计为 1, 否则为 0。	法国 CEPII 数据库

际科研论文合作网络图(图1),图中节点大小与国家论文数量成比例,节点连线大小与国家之间的论文合作规模成正比。从节点分布来看,欧洲国家是全球科研论文合作网络的主体;亚洲区域次之,主要包括中国、日本、韩国、印度、台湾、新加坡等国家或地区;而美国处在全球科研论文合作网络的中心,与图中97%的节点产生科研论文合作联系,是全球科研合作的中枢。相比而言,南美洲、非洲和大洋洲则地处网络边缘,聚集成群,形成独立的社团,其区域核心分别是巴西、南非和澳大利亚(图1)。这些发现与《Nature Index 2015 Collaborations》的报告结论基本一致^[56]。

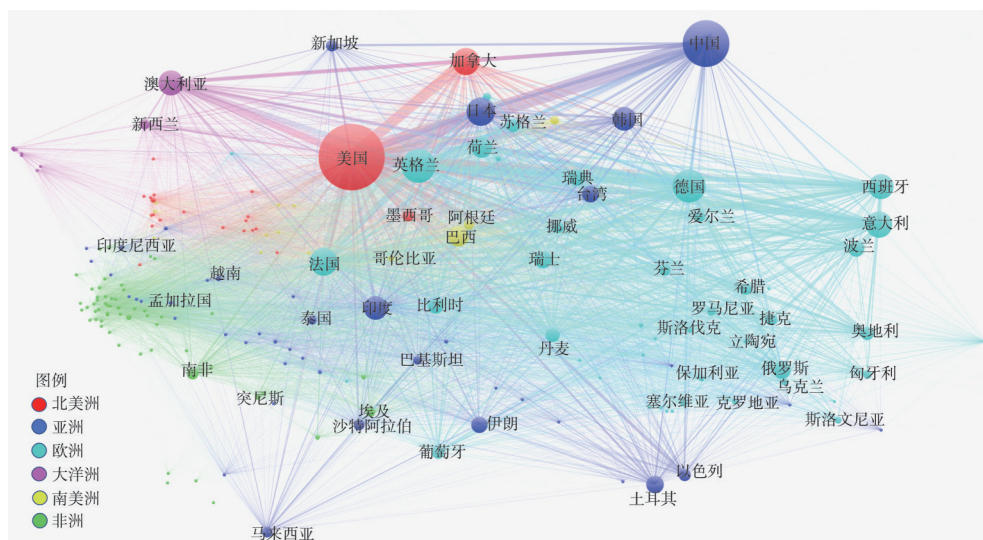


图1 全球科研论文合作网络的拓扑结构(据Pajek和VOSviewer绘制)

Fig. 1 The topological structure of global scientific collaboration network

从统计特征来看,与2005年的SCI索引网络相比^[57],2014年全球科研论文合作网络的节点数、边数、密度、平均度中心性都有大幅增加,表明十年间参与全球科研合作网络的国家日益增多,各国间的知识流动日益频繁。与此同时,全球科研论文合作网络密度为0.45,节点平均度值为94.10,表明整个网络中各国家之间知识溢出频繁,合作联系密切,具有较高的凝聚性。此外,整个网络中心性分布不均衡,尤其是强度中心性和介数中心性差异悬殊,基尼系数Gini大于0.7,变异系数CV也超过1.5,表明介数和强度高度集中于少数节点,而度中心性分布(变异系数为0.5498,基尼系数为0.3154)则相对均衡(表3)。

3.2 小世界性

与同等规模随机网络相比,全球科研论文合作网络具有较大的集聚系数,较小的平均路径长度,发育典型的小世界性。该网络的平均集聚系数为0.73,大于同等规模的随机网络($C_{ER} \approx 0.44$),表现出较强的集聚性,国家间科研论文合作以短距离的联系为主,具有明显的“轴—辐式”网络组织特征。它的平均路径长度为1.56,略低于同等规模的随机网络($L_{ER} \approx 1.78$),意味着国家间的论文合作联系紧密,具有较好的网络通达性。除此之外,累积度分布函数遵循良好的指数分布律($y = 1.9406e^{-0.018x}$, $R^2 = 0.7853$),具有较高的置信水平。Leydesdorff^[57]基于SCI数据库的国际科研论文合作网络和Haunschild等^[58]基于文献管理软件Mendeley的读者使用网络,均发现全球尺度的合作网络具有显著的小世界性(表3)。

表3 全球科研论文合作网络的统计特征
Tab. 3 Statistical characteristics of global scientific collaboration network

指标	规模				小世界性		度中心性			介数中心性		强度中心性	
	节点数	边数	密度	直径	平均路长	平均集聚系数	平均度	基尼系数	变异系数	基尼系数	变异系数	基尼系数	变异系数
本文(2014)	211	9928	0.45	3	1.56	0.73	94.10	0.32	0.55	0.71	1.85	0.83	2.63
SCI(2005) ^[55]	194	9400	0.25	3	1.80	0.79	48.60	-	-	-	-	-	-
Mendeley(2012) ^[56]	178	6774	0.43	3	1.58	-	76.02	-	-	-	-	-	-

3.3 等级层次性

中心势是衡量整个网络中心化程度的重要指标, 全球科研论文合作网络的度中心势为0.53, 接近度中心势为0.64, 表明整个网络具有比较明显的向某个节点或某些节点集中的趋势, 存在“中心—外围”结构。

采用Pajek块模型分析中的层次聚类算法(hierarchical clustering), 依据度中心性值, 获取层级文件, 并转换为分区文件, 将全球科研论文合作网络划分为5个等级(表4)。第一层级: 美国, 处于科研合作网络的中心。美国是全球最大的科研论文产出国, 拥有最多的科研合作伙伴, 国际合作论文发表数量位居第一, 四大中心性指标皆排在首位, 成为全球科学研究的领导者, 处于全球知识网络金字塔的顶端。西欧7国(英格兰、法国、意大利、德国、荷兰、瑞士、西班牙)、北美的加拿大以及亚太地区4国(中国、日本、韩国、澳大利亚), 构成全球科研论文合作网络的第二梯队。他们的各项指标大致相当, 网络性质相似, 科研论文合作联系密切, 形成高连通网络。第三层级由39个国家或地区组成, 包括南非、土耳其、新西兰、墨西哥、巴西、俄罗斯、以色列、印度等。他们都是所在区域的科研论文产出大国, 甚至是区域的科技创新中心(如巴西和南非), 各项指标均高于网络平均值, 相互之间有着稠密的科研合作联系。肯尼亚、越南、古巴、秘鲁、阿拉伯联合酋长国、印度尼西亚、菲律宾等46个节点构成第四层级。相比而言, 他们之间科研论文合作程度不高, 但平均度中心性、接近度中心性和密度皆高于网络的平均值。最后剩余的113个国家或地区组成第五层级, 主要来自非洲、太平洋岛国、西印度洋地区、西亚、东南欧等地区。这些国家或经济发展水平低, 或国土面积小、人口密度低, 或深陷地缘政治的泥潭, 科学研究事业发展缓慢, 处在全球知识合作网络金字塔的底部。

从Pajek生成的分区文件以2D格式输出到VOSviewer, 颜色的冷暖随层级的高低设置, 获得全球科研论文合作网络的等级层次结构图(图2)。图2中节点大小与该节点的强度成正比, 边大小与论文合作数量正相关。从图2可以发现, 全球科研论文合作网络已经发育形成典型的“核心—边缘”等级渐进式形态, 可以划分为核心区、半边缘带、边缘带三大国家集团。其中, 核心国家集团由第一和第二层级组成, 共13个国家, 是全

表4 全球科研论文合作网络的等级层次结构特征
Tab. 4 Statistical characteristics on the hierarchical structure of global scientific collaboration network

等级	节点数量	平均度中心性	平均接近度中心性	平均介数中心性	平均强度中心性	密度
第一层级	1	205	0.977	0.040	286062	1
第二层级	12	179.41	0.875	0.015	90281.66	1
第三层级	39	148.41	0.775	0.005	22226.28	1
第四层级	46	117.46	0.697	0.002	3289.13	0.76
第五层级	113	55.81	0.579	0.001	268.22	0.13
整个网络	211	94.10	0.660	0.003	11459.13	0.45

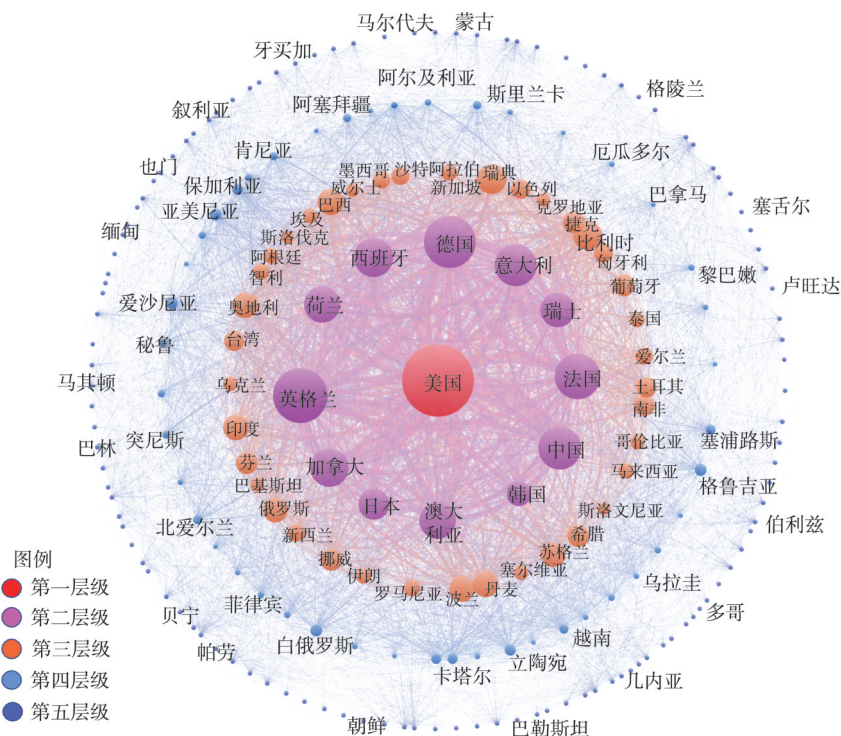


图2 全球科研论文合作网络的等级层次结构(据Pajek和VOSviewer绘制)

Fig. 2 The hierarchical structure of global scientific collaboration network

球科研论文合作网络的hub节点和枢纽,以洲际之间科研论文合作为主。主要位于西欧(英格兰、法国、意大利、德国、荷兰、瑞士、西班牙)、北美(美国和加拿大)和亚太(中国、日本、韩国、澳大利亚)三大区域,内部之间的连接数量多、密度大(网络密度为1)、强度高(最大联系值为32291),形成完全连通网络,“富人俱乐部”特征显著。第三和第四层级构成半边缘国家集团,共85个国家,是区域科研论文合作网路的重要门户,以洲内尺度的科研论文合作为主导。半边缘地带国家内部联系紧密(网络密度为0.9165),联系强度中等(最大联系值为2470);与核心国家之间的合作联系数量较多且强度较大,但与边缘国家之间的纽带较少且强度低。第五层级(剩余113个国家)组成边缘国家集团,处在全球科研论文合作网络的边缘,扮演附属角色,彼此之间联系数量少(网络密度仅为0.13)且强度非常低(最大联系值仅为27)(表4)。

4 空间结构异质性

4.1 网络联系:以四边形为骨架

选取发文量不低于1000篇的节点,合作量不低于500的边,绘制全球主要国家科研论文合作网络图谱(图3)。研究发现,世界各国的科研论文产出具有明显的空间不均衡性,非洲、中东、中亚、中美地峡与西印度洋群岛、西太平洋群岛等人口稀疏或经济欠发达区域构成全球科研论文产出的低谷区,欧洲、北美洲、亚太地区则是科研活动的热点区域,成为科研论文产出的高值区。整个合作网络呈现以北美(美国与加拿大)、欧盟、东亚(中国、日本、韩国)和澳大利亚等科研产出高值区为顶点的四边形主骨架,

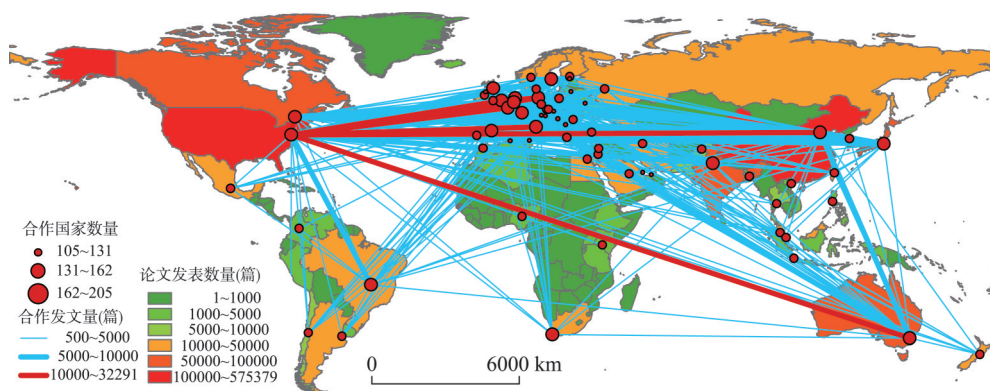


图3 全球主要国家科研论文合作网络图谱

Fig. 3 Spatial distribution of global scientific collaboration network

南美洲和非洲成为其外围区域,通过次一级的联系流与核心区域连接。这表明科研论文合作网络并没有受到明显的地理距离约束,主要合作流发生在这些相距甚远的核心节点之间^[59]。其中,欧洲国家间论文合作联系紧密,对外联系以亚太和北美为主,英格兰、德国、意大利、法国、西班牙和瑞士是欧洲对外论文合作联系的核心枢纽。美国、加拿大和墨西哥构成北美洲科研合作网络的重要节点,内部联系以美国和加拿大为主,对外与欧洲联系最为频繁,与亚太地区(主要是中国、澳大利亚和日本)联系强度最大。随着亚太地区经济的崛起,中国、韩国、日本和澳大利亚成为全球科研论文合作的积极参与者,域内以中国与澳大利亚、日本、韩国间科研联系为主,对外联系则高度汇聚于西欧和北美。

4.2 度中心性:大分散小集聚

整体上,无论是Moran's I 指数(0.1581),还是Gini系数和变异系数CV(表3),度中心性的国际分异程度均明显较介数和强度中心性小,表明全球科研论文合作网络呈弱集聚性和分散化的空间态势(图4)。大部分国家或地区科研论文合作程度较小且差距不大,小部分具较高论文合作程度的国家高度集中于北美、西欧和亚太三大组团。度中心性排名前20的国家中,有90%国家或地区位于以上三大区域,呈现出显著的区域集聚性特征。究其原因,北美和西欧是前三次科技革命的发源地和众多国际科研合作组织的所在地,成为全球科研产出的高地和科研合作流的“集线器”;而近年亚太相关国家经济的腾飞,创新驱动战略的实施,留学人才的回流,也为其国际科研论文产出和合作提供源动力。

4.3 介数中心性:三足鼎立

从洲域尺度来看,高介数中心性的国家或地区仍然集中于北美、西欧和亚太三大增长极,呈现出“三足鼎立”的非均衡格局(图5)。具体而言,除南非以外,所有的非洲国家介数中心性值均低于0.005,是全球介数分布的低值区。整个亚欧大陆,东边是亚太极值区,西边是欧洲极值区,东西之间形成“塌陷地带”,整体表现出“四周高、中间低”的盆地形态。而美洲则形成“北高南低”的空间差异,美国与加拿大构成北美极值区,中美洲和南美洲大部分国家介数值较小,科研合作的桥梁作用有待加强。其原因在于:①这三大增长极,主要包括美国、加拿大、中国、澳大利亚、英国、法国、德国、意大利、西班牙等,拥有垄断性的科研资源,成为全球科学研究的领导者。从高校数量和质量来看,以绝对的优势远超其他区域,拥有347所世界500强高校,有75所高校进

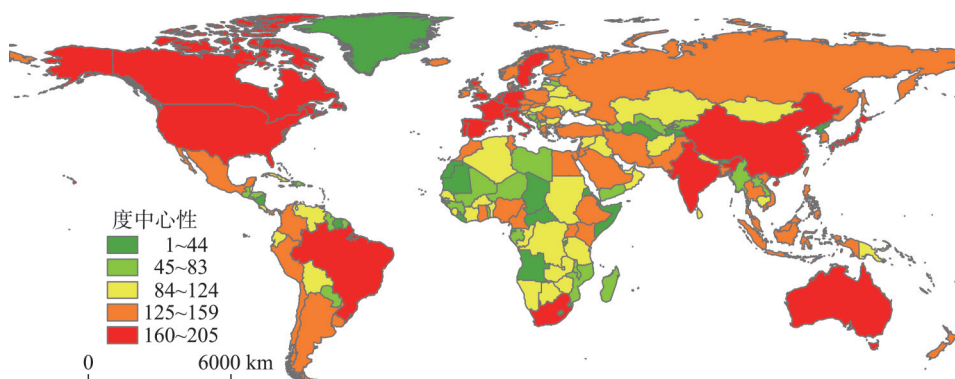


图4 度中心性的全球空间分布

Fig. 4 Global spatial distribution spatial distribution of nodal degree centrality

入世界大学100强^①。② 英国、法国、意大利、德国与西班牙等曾为世界上最主要的殖民主义国家，殖民地广泛分布在世界各地，至今仍有许多海外领地。研究发现殖民地国家与宗主国保持密切的科研合作关系^[60]。这些国家仍是很多非洲国家、太平洋岛国首要的科研合作对象。③ 共同文化有利于维持紧密的合作关系^[61-62]。英语是目前使用最为广泛的语言，欧美文化是当今全球的主流，其在学术交流上具有天然的文化亲近性和语言便利性。以2014年所收录的文献为例，英语论文高达1960161篇，占总体的96.21%，占据绝对的主导地位，其中大多数论文由欧洲、美国、加拿大和澳大利亚等国家的学者完成。

4.4 强度中心性：一超多强

相比前两大中心性指标，强度中心性的空间分布更加不均匀，表现出“一超多强”的空间格局（图6）。其中，在科研论文产出和合作规模上，美国成为独一无二的超级大国。其科研合作强度中心性值达到286062，合作论文数量为175195篇，均位居世界第一。此外，强度中心性约为第二位的2倍，国际合作论文数量约为第二位的3倍，具有较高的首位度。英格兰、德国、法国、中国、意大利、西班牙、加拿大、荷兰、澳大利亚、瑞士、日本、瑞典等国，与美国差距较明显，但均突破5万论文合作规模，且远高于其它国家或地区，进入全球科研论文合作的强国序列。而非洲、中亚、东南亚、加勒比海地区等构成全球科研论文合作强度的低值区和外围区。

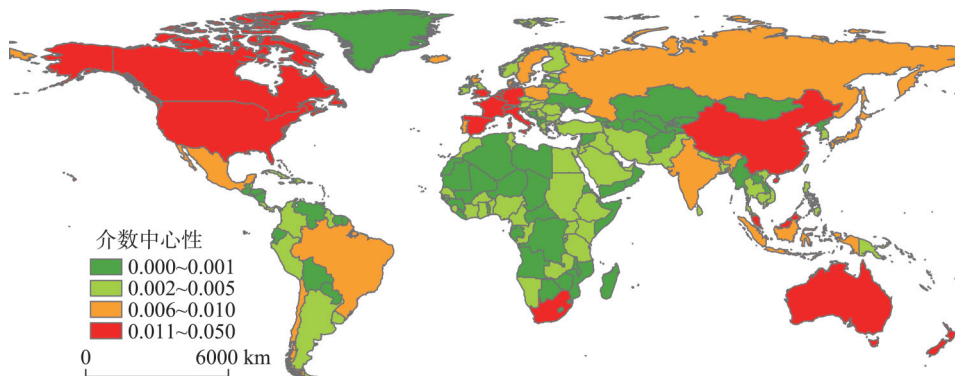


图5 介数中心性的全球空间分布

Fig. 5 Global spatial distribution spatial distribution of nodal betweenness centrality

① 新华出国. 解析2015年USNews世界大学前500名榜单. 新华网, 2015-02-24.

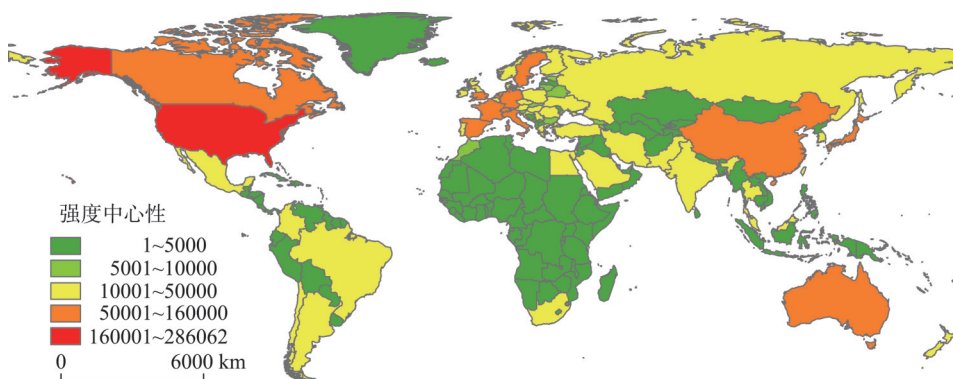


图6 强度中心性的全球空间分布

Fig. 6 Global spatial distribution spatial distribution of strength centrality

5 邻近性机理

采用重力模型和负二项式回归分析方法,检验了国家间地理邻近性、社会邻近性、经济邻近性、语言邻近性及发表论文量与科研论文合作量的估计结果(表5)。从模型拟合程度来看,Alpha参数均不等于0,各因变量均具1%的显著性水平,模型拟合度较高,具较好解释力。

模型1揭示出,国家论文发表量和地理距离是影响科研论文合作的重要因素。一方面,国际论文合作规模与各自的科研论文产出成正比,即发表的科研论文越多,两国之间存在合作的可能性及合作量越大,表明全球科研论文合作网络存在显著的强强联合、合作共赢的网络演化态势。与此同时,该发现与已有研究结论相似,如Hoekma等^[12]对欧洲的知识合作研究和Cassi等^[34]对全球葡萄酒产业的科研合作研究均发现国家的科研能力积极地促进科研合作。另一方面,科研论文合作量与两国的地理距离成反比,即科研工作者之间的地理距离与合著的论文数量呈显著的负相关,这也与许多研究发现相一致^[7, 9, 12]。原因在于,研究者更愿意寻找邻近的科研合作伙伴,通过面对面的互动促进隐性知识的传播和非正式的交流,从而提高科研产出的效率。与其他集中于国家尺度(中国、美国)、区域尺度(欧盟)的研究相比^[12, 63-64],全球尺度的地理距离对科研论文合作的阻碍作用较小。可能缘由是,许多前沿或重大的研究领域,如应对全球性挑战(气候变化),已非某个人、机构或国家能够独立完成,需要跨区域或全球性的科研联合攻关;同时交通变革、通信技术发展也进一步减少了“距离的严苛管制(the tyranny of distance)”^[65]或“磨擦(friction of distance)”作用。

在模型2中,发现国家间的社会邻近性对科研论文合作具有显著的正向作用,意味着两国的社会关系越密切,越有利于科研合作的展开和深化。当主体之间基于信任、友谊和频繁的互动所建立的合作关系,能有效地减少繁琐的流程,便于非正式的知识交换和增加双边合作的可能性^[38]。正如Plotnikova等发现,社会邻近积极而又显著地促进国际制药合作研究^[35]。原因是科技人才的全球性流动,往往会加速其社会关系的全球化,从而明显地促进国际学术合作和交流。如大批留学人员归国,极大地加强了中国与北美(美国与加拿大),以及亚太(新加坡、日本、澳大利亚)与欧洲(英国、德国、法国)的科研合作^[66]。值得一提的是,尽管许多R&D机构呈现全球化转移,但大量的R&D合作面向母国总部^[49],因“技术封闭”很少向东道国企业溢出,密切的国际R&D转移并未带来显著的母国—东道国间技术创新合作。

表5 重力模型的负二项式回归估计结果
Tab. 5 Estimation results of the negative binomial spatial interaction models

	模型1	模型2	模型3	模型4
国家1发文量	0.0304*** (0.0010)	0.0168*** (0.0005)	0.0168*** (0.0005)	0.0172*** (0.0005)
国家2发文量	0.0129*** (0.0010)	0.0072*** (0.0003)	0.0072*** (0.0003)	0.0071*** (0.0003)
地理邻近性	-0.0934*** (0.0050)	-0.0195*** (0.0031)	-0.0198*** (0.0030)	-0.0231*** (0.0030)
社会邻近性	-	0.4950*** (0.0114)	0.4843*** (0.0118)	0.4777*** (0.0117)
经济邻近性	-	-	0.1934*** (0.0304)	0.1637*** (0.0301)
语言邻近性	-	-	-	-0.6055*** (0.0437)
常数	3.5188*** (0.0483)	1.7361*** (0.0373)	1.6978*** (0.0371)	1.8197*** (0.0383)
样本量	9315	9315	9315	9315
Alpha	2.4574	1.4363	1.4307	1.3952
Wald chi2	1549.80	3630.23	3928.45	4241.28
Prob > chi2	0.0000	0.0000	0.0000	0.0000
Log pseudolikelihood	-41131.909	-37827.278	-37804.322	-37667.596

注：*： $p < 0.10$ ；**： $p < 0.05$ ；***： $p < 0.01$ 。

从模型3可以得出，人均国民收入也是影响科研论文合作的重要因素。国家间的经济发展水平越接近，相应的科学技术需求越一致，其科研人员开展科研论文合作的可能性和规模就越大，即经济邻近性会促进科研论文合作^[9, 35]。如Scherngell等对中国各省域之间的论文合作机制研究发现，省际经济水平差距越大，科研合作量越少^[7]。必须指出的是，全球科研论文合作与经济邻近性的相关程度较低。主要原因是，低收入国家，虽经济相似性高，但无法提供较大的研发投入来维持科研论文合作关系；而部分高收入国家“国小人少”，缺乏足够的研究型教育科研机构，造成对外科研合作水平不高。

已有研究表明，语言邻近性对国际科研合作缺乏促进作用。如，Cassi等的全球葡萄酒业研发合作和Plotnikova等的国家层面的制药研究合作，均发现语言差异并不阻碍国际科研合作^[34-35]。模型4进一步证实，语言已经不是阻碍科研论文合作全球化的影响因素。究其原因，WOS核心合集中绝大多数科研论文用英语撰写，英语已经成为使用最为广泛的语言和首选的第二语言，通过英语纽带，可以促进不同语系国家科研人员的论文合撰。但是，Scherngell等^[9]研发合作网络（欧盟25国的255个NUTS-2区域）和Hoekman等^[33]研究项目合作网络（欧盟33国的313个NUTS-2区域）的实证分析发现，由于语言接近或文化趋同，便于项目的交流、学习、管理和创新，从而对科研合作具有促进作用。

6 结论与讨论

6.1 结论

本文基于WOS科研论文数据挖掘，采用空间统计方法、复杂网络分析和重力模型系统描绘了全球科研论文合作网络的拓扑结构、空间格局及其邻近性机制：

（1）拓扑结构上，全球科研合作网络密度较大，平均度值较高，具较高的凝聚性。

与同等随机网络相比,国际科研合作网络具有较大的集聚系数,较小的平均路径长度,累积度分布服从指数律,发育典型的小世界性。三大中心性分布上符合帕累托法则,呈现明显的结构异质性。美国成为拓扑网络的极核,欧洲国家构成网络的主体,亚太国家正发展成网络的新增长点。这种不均衡性表现出一定的等级层次性,可以划分为5个层级,形成明显的“核心—边缘”网络结构,一二级组成网络核心,其余层级构成边缘地带。

(2) 空间格局上,科研合作网络联系强度分布不均匀,形成北美、欧洲、东亚和澳大利亚为顶点的四边形主骨架。三大中心性指标值空间差异显著,呈现各具特色的空间形态。度中心性表现为全球分散与区域集聚的双重特征,整体上以低度中心性国家为主,分布较均匀,小部分高度中心性国家或地区主要集聚于北美、欧洲、亚太的三大区域。介数中心性呈现“三足鼎立”的非均衡格局,亚太、欧洲和北美构成三大极值区。强度中心性则展现出“一超多强”的层级格局,美国是独一无二的超级大国,英格兰、德国、法国、中国等国是紧随其后的科研产出和合作强国。

(3) 影响机制上,通过重力模型和负二项式回归分析发现,国际科研论文合作与两国的科研规模成正比,受到地理距离的阻抗作用,地理邻近性具显著的正向效应。社会和经济邻近性也是影响国际科研论文合作的重要因素,对各国家科研工作者的合作具积极的促进作用。语言不再是国际科研合作的障碍因素,不同语言之间的知识合作更加频繁。

6.2 讨论

① 全球科研合作网络是一个复合的、动态的网络,有必要在论文合著基础上增加专利、R&D、科研项目合作等数据,开展较长时间尺度的知识合作网络演化研究,厘清其时空结构的形成过程、演化趋势和影响机制。② 除了邻近性之外,国际经济形势、殖民关系或双边关系,以及国内科教水平、人力资源和科研投入等社会经济因素的影响有待开展计量分析。此外,网络位置、关系强度、网络密度、网络派系、节点中心性、结构洞等网络动力学机制也值得深究。③ 本文所构建的国家科研论文合作矩阵,采用的是全计数方式(full count),其前提假设是合作双方同等重要。忽略了论文合著者的贡献差异性,未来需要采用加权分式计数(weighted fractional count)方式进行优化。

参考文献(References)

- [1] Royal Society. Knowledge, Networks and Nations: Global Scientific Collaboration in the 21st Century. Amsterdam: Elsevier, 2011: 5-6.
- [2] Adams J. Collaborations: The fourth age of research. *Nature*, 2013, 497(7451): 557-560.
- [3] Adams J, Loach T. Comment: A well-connected world. *Nature*, 2015, 527(7577): S58-S59.
- [4] Adams J. Collaborations: The rise of research networks. *Nature*, 2012, 490(7420): 335-336.
- [5] Breschi S, Lissoni F. Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, 2009, 9(4): 439-468.
- [6] TerWal A L J. Cluster emergence and network evolution: A longitudinal analysis of the inventor network in Sophia-Antipolis. *Regional Studies*, 2013, 47(5): 651-668.
- [7] Scherngell T, Hu Y. Collaborative knowledge production in China: Regional evidence from a gravity model approach. *Regional Studies*, 2011, 45(6): 755-772.
- [8] Balland P A. Proximity and the evolution of collaboration networks: Evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies*, 2012, 46(6): 741-756.
- [9] Scherngell T, Barber M J. Distinct spatial characteristics of industrial and public research collaborations: Evidence from the Fifth EU Framework Programme. *The Annals of Regional Science*, 2011, 46(2): 247-266.
- [10] Crescenzi R, Nathan M, Rodríguez-Pose A. Do inventors talk to strangers? On proximity and collaborative knowledge creation. *Research Policy*, 2016, 45(1): 177-194.
- [11] Broekel T, Boschma R. Knowledge networks in the Dutch aviation industry: The proximity paradox. *Journal of Economic Geography*, 2012, 12(2): 409-433.
- [12] Hoekman J, Frenken K, Van Oort F. The geography of collaborative knowledge production in Europe. *The Annals of*

- Regional Science, 2009, 43(3): 721-738.
- [13] D'Este P, Guy F, Iammarino S. Shaping the formation of university-industry research collaborations: What type of proximity does really matter? *Journal of Economic Geography*, 2013, 13(4): 537-558.
- [14] Balland P A, De Vaan M, Boschma R. The dynamics of interfirm networks along the industry life cycle: The case of the global video game industry, 1987-2007. *Journal of Economic Geography*, 2013, 13(5): 741-765.
- [15] De Prato G, Nepelski D. Global technological collaboration network: Network analysis of international co-inventions. *The Journal of Technology Transfer*, 2014, 39(3): 358-375.
- [16] Li D, Wei Y D, Wang T. Spatial and temporal evolution of urban innovation network in China. *Habitat International*, 2015, 49: 484-496.
- [17] Li Dandan, Wang Tao, Zhou Hui. The structural characteristics of knowledge spillover networks based on different spatial and temporal scales. *Scientia Geographica Sinica*, 2013, 33(10): 1180-1187. [李丹丹, 汪涛, 周辉. 基于不同时空尺度的知识溢出网络结构特征研究. *地理科学*, 2013, 33(10): 1180-1187.]
- [18] Leydesdorff L, Wagner C S, Park H W, et al. International collaboration in science: The global map and the network. *El Profesional de la Información*, 2013, 22(1): 87-95.
- [19] Huallacháin B Ó, Lee D S. Urban centers and networks of co-invention in American biotechnology. *Annals of Regional Science*, 2014, 52(3): 799-823.
- [20] Cairncross F. The death of distance: How the communications revolution is changing our lives. Massachusetts: Harvard Business Press, 2001: 24-35.
- [21] Friedman T. The World Is Flat: A Brief History of the Globalized World In the 21st Century. London: Penguin, 2005: 393-395.
- [22] David P A, Foray D. An introduction to the economy of the knowledge society. *International Social Science Journal*, 2002, 54(171): 9-23.
- [23] McCann P. Globalization and economic geography: The world is curved, not flat. *Cambridge Journal of Regions, Economy and Society*, 2008, 1(3): 351-370.
- [24] Heimeriks G, Boschma R. The path-and place-dependent nature of scientific knowledge production in biotech 1986-2008. *Journal of Economic Geography*, 2014, 14(2): 339-364.
- [25] Gertler M S. Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, 2003, 3(1): 75-99.
- [26] Matthiessen C W, Schwarz A W. The top-level global research system, 1997-99: Centres, networks and nodality. An analysis based on bibliometric indicators. *Urban Studies*, 2002, 39(5/6): 903-927.
- [27] Matthiessen C W, Schwarz A W, Find S. World cities of scientific knowledge: Systems, networks and potential dynamics: An analysis based on bibliometric indicators. *Urban Studies*, 2010, 47(9): 1879-1897.
- [28] Boschma R. Proximity and innovation: A critical assessment. *Regional Studies*, 2005, 39(1): 61-74.
- [29] Torre A, Gilly J P. On the analytical dimension of proximity dynamics. *Regional Studies*, 2000, 34(2): 169-180.
- [30] Bathelt H, Malmberg A, Maskell P. Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 2004, 28(1): 31-56.
- [31] Ibert O, Hautala J, Jauhiainen J S. From cluster to process: New economic geographic perspectives on practices of knowledge creation. *Geoforum*, 2015, 65: 323-327.
- [32] Bouba-Olga O, Carrincazeaux C, Coris M, et al. Proximity dynamics, social networks and innovation. *Regional Studies*, 2015, 49(6): 901-906.
- [33] Hoekman J, Frenken K, Tijssen R J W. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 2010, 39(5): 662-673.
- [34] Cassi L, Morrison A, Rabellotti R. Proximity and scientific collaboration: Evidence from the global wine industry. *Tijdschrift voor Economische en Sociale Geografie*, 2015, 106(2): 205-219.
- [35] Plotnikova T, Rake B. Collaboration in pharmaceutical research: Exploration of country- level determinants. *Scientometrics*, 2014, 98(2): 1173-1202.
- [36] Balland P A, Boschma R, Frenken K. Proximity and innovation: From statics to dynamics. *Regional Studies*, 2015, 49(6): 907-920.
- [37] Nooteboom B. Innovation, learning and industrial organisation. *Cambridge Journal of Economics*, 1999, 23(2): 127-150.
- [38] Boschma R, Frenken K. The spatial evolution of innovation networks. A proximity perspective//Boschma R, Martin R. *The Handbook of Evolutionary Economic Geography*. Cheltenham: Edward Elgar Publishing, 2010: 120-135.
- [39] North D C. Institutions, Institutional Change and Economic Performance. Cambridge: Cambridge University Press,

- 1990: 68-79.
- [40] Ponds R, Van Oort F, Frenken K. The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 2007, 86(3): 423-443.
- [41] Lv Guoqing, Zeng Gang, Gu Nana. Literature Review of regional innovation network: An economic geographical perspective. *Economic Geography*, 2014, 34(2): 1-8. [吕国庆, 曾刚, 顾娜娜. 经济地理学视角下区域创新网络的研究综述. *经济地理*, 2014, 34(2): 1-8.]
- [42] Lu Lachang, Huang Ru, Liao Qian. Several theoretical issues on innovation geography. *Scientia Geographica Sinica*, 2016, 36(5): 653-661. [吕拉昌, 黄茹, 廖倩. 创新地理学研究的几个理论问题. *地理科学*, 2016, 36(5): 653-661.]
- [43] Si Yuefang, Zeng Gang, Cao Xianzhong, et al. Research progress of glocal innovation networks. *Progress in Geography*, 2016, 35(5): 600-609. [司月芳, 曾刚, 曹贤忠, 等. 基于全球—地方视角的创新网络研究进展. *地理科学进展*, 2016, 35(5): 600-609.]
- [44] Li Dandan, Wang Tao, Wei Yehua, et al. Spatial and temporal complexity of scientific knowledge network and technological knowledge network on China's urban scale. *Geographical Research*, 2015, 34(3): 525-540. [李丹丹, 汪涛, 魏也华, 等. 中国城市尺度科学知识网络与技术知识网络结构的时空复杂性. *地理研究*, 2015, 34(3): 525-540.]
- [45] Wang Qiuyu, Zeng Gang, Lu Guoqing. Structural evolution of innovation networks of China's equipment manufacturing industry. *Acta Geographica Sinica*, 2016, 71(2): 251-264. [王秋玉, 曾刚, 吕国庆. 中国装备制造业产学研合作创新网络初探. *地理学报*, 2016, 71(2): 251-264.]
- [46] Lu Lachang, Li Yong. A research on Chinese renovation urban system: Based on urban renovation function. *Acta Geographica Sinica*, 2010, 65(2): 177-190. [吕拉昌, 李勇. 基于城市创新职能的中国创新城市空间体系. *地理学报*, 2010, 65(2): 177-190.]
- [47] Ma H, Fang C, Pang B, et al. Structure of Chinese city network as driven by technological knowledge flows. *Chinese Geographical Science*, 2015, 25(4): 498-510.
- [48] Chen W, Xiu C, Liu W, et al. Visualizing intercity scientific collaboration networks in China. *Environment and Planning A*, 2015, 47(11): 2229-2231.
- [49] Zhu Ying, Du Debin. The spatial organization of R&D globalization by multinational corporations. *Economic Geography*, 2005, 25(5): 620-623. [祝影, 杜德斌. 跨国公司研发全球化的空间组织研究. *经济地理*, 2005, 25(5): 620-623.]
- [50] Du Debin, Zhou Tianyu, Wang Yong, et al. The situation and trends of world's R&D industry. *World Regional Studies*, 2007, 16(1): 1-6. [杜德斌, 周天瑜, 王勇. 世界R&D产业的发展现状及趋势. *世界地理研究*, 2007, 16(1): 1-6.]
- [51] Lu Lachang, Liang Zhengji, Huang Ru. The innovation linkage among Chinese major cities. *Scientia Geographica Sinica*, 2015, 35(1): 30-37. [吕拉昌, 梁政骥, 黄茹. 中国主要城市间的创新联系研究. *地理科学*, 2015, 35(1): 30-37.]
- [52] Lv Guoqing, Zeng Gang, Gu Nana. Dynamic evolution of innovation network in China's equipment manufacturing industry: Geographic proximity versus social proximity. *China Soft Science*, 2014(5): 97-106. [吕国庆, 曾刚, 顾娜娜. 基于地理邻近与社会邻近的创新网络动态演化分析: 以我国装备制造业为例. *中国软科学*, 2014(5): 97-106.]
- [53] Wang Tao, Li Dandan. Spatial structure evolution of knowledge network and its impact on the NIS: Case study of biotechnology in China. *Geographical Research*, 2011, 30(10): 1861-1872. [汪涛, 李丹丹. 知识网络空间结构演化及对NIS建设的启示: 以我国生物技术知识为例. *地理研究*, 2011, 30(10): 1861-1872.]
- [54] Ma H, Fang C, Pang B, et al. The effect of geographical proximity on scientific cooperation among Chinese cities from 1990 to 2010. *PloS One*, 2014, 9(11): e111705.
- [55] Leydesdorff L. On the normalization and visualization of author co-citation data: Salton's Cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, 2008, 59(1): 77-85.
- [56] Kogleck L, Multiples S. Strength in numbers. *Nature*, 2015, 527(7577): S50-S51.
- [57] Leydesdorff L, Wagner C S. International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2008, 2(4): 317-325.
- [58] Haunschild R, Bornmann L, Leydesdorff L. Networks of reader and country status: An analysis of Mendeley reader statistics. *Sociological Review*, 2015, 25(25): 867-876.
- [59] Wu Kang, Fang Chuanglin, Zhao Miaoqi. The spatial organization and structure complexity of Chinese intercity networks. *Geographical Research*, 2015, 34(4): 711-728. [吴康, 方创琳, 赵渺希. 中国城市网络的空间组织及其复杂性结构特征. *地理研究*, 2015, 34(4): 711-728.]
- [60] Pouris A. A scientometric assessment of the Southern Africa development community: Science in the tip of Africa. *Scientometrics*, 2010, 85(1): 145-154.
- [61] Anonymous. Developing partnerships. *Nature*, 2015, 527(7577): S60-S63.
- [62] Larivière V, Gingras Y, Archambault É. Canadian collaboration networks: A comparative analysis of the natural

- sciences, social sciences and the humanities. *Scientometrics*, 2006, 68(3): 519-533.
- [63] Andersson D E, Gunessee S, Matthiessen C W, et al. The geography of Chinese science. *Environment and Planning A*, 2014, 46(12): 2950-2971.
- [64] Singh J, Marx M. Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science*, 2013, 59(9): 2056-2078.
- [65] Blainey G. *The Tyranny of Distance: How Distance Shaped Australia's History*. Melbourne: Sun Books, 1966.
- [66] Pincock S. China's diaspora brings it home. *Nature*, 2015, 527(7577): S68-S71.

Structural heterogeneity and proximity mechanism of global scientific collaboration network based on co-authored papers

LIU Chengliang^{1,2}, GUI Qinchang¹, DUAN Dezhong^{1,2}, YIN Meiyuan¹

(1. School of Urban and Regional Science, East China Normal University, Shanghai 200241, China;

2. Institute for Innovation and Strategic Studies, East China Normal University, Shanghai 20062, China)

Abstract: Despite increasing importance of academic papers in global knowledge flows, the structural disparities and proximity mechanism related to international scientific collaboration network attracted little attention. To fill this gap, based on data mining from Thomson Reuters' Web of Science database in 2014, its heterogeneities in topology and space were portrayed using visualizing tools such as Pajek, Gephi, VOSviewer, and ArcGIS. Topologically, 211 countries and 9928 ties are involved in global scientific collaboration network, but the international network of co-authored relations is mono-centric and dominated by the United States. It exhibits some features of a "small-world" network with the smaller average path length of 1.56 and the extremely large cluster coefficient of 0.73 compared to its counterpart, as well as the better-fitting exponential distribution accumulative nodal degree. In addition, the entire network presents a core-periphery structure with hierarchies, which is composed of 13 core countries and the periphery of 198 countries. Spatially, densely-tied and high-output areas are mainly distributed in four regions: West Europe, North America, East Asia and Australia. Moreover, the spatial heterogeneity is also observed in the distributions of three centralities. Amongst these, the countries with greater strength centrality are mainly concentrated in North America (i.e. the US and Canada), Western Europe (i.e. the UK, France, Germany, Italy and Spain), and China, noticeably in the US, which forms the polarizing pattern with one superpower of the US and great powers such as China and the UK. Similarly, the big three regions consisting of West Europe, North America and Asian-Pacific region have the peak betweenness centrality as well. Slightly different from the two above, the distribution of nodal degree centrality is uneven in the world, although regional agglomeration of high-degree countries is still observed. Last but not least, the proximity factors of its structural inequalities were also verified by correlational analysis, negative binomial regression approach and gravity model of STATA. The findings further confirm that geographical distance has weakened cross-country scientific collaboration. Meanwhile, socio-economic proximity has a positive impact on cross-country scientific collaboration, while language proximity plays a negative role.

Keywords: scientific collaboration network; structural heterogeneity; complex network