doi:10. 3969/j.issn. 1672-4623.2023.01.005

全球人口迁移网络结构探测与时空分析

高谢庆[1]，秦　昆[1*]，喻雪松[1]，张　凯[1]
（1. 武汉大学 遥感信息工程学院，湖北 武汉 430079）

摘　要：人口迁移已成为各个国家或地区之间频繁发生的交互现象，人口迁移网络研究是社会网络与复杂网络的重要应用领域。首先以全球人口迁移数据为基础，以国家或地区为节点，以国家或地区之间的人口迁移为边构建网络；然后研究全球人口移动网络的结构探测与时空分析方法，探索全球人口迁移网络的结构特点和时空变化规律。研究结果表明，全球人口迁移网络具有无标度、小世界与异配等特性，其网络结构的变化与地理位置、国际关系、国际事件紧密相关。研究结果对于全球尺度的人口问题研究具有一定的参考价值。

关键词：全球人口迁移；复杂网络；幂律分布；社区提取；时序变化
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2023）01-0026-06

Structure Detection and Spatio-temporal Analysis of Global Population Migration Network

GAO Xieqing[1], QIN Kun[1], YU Xuesong[1], ZHANG Kai[1]
(1. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China)

Abstract: Population migration has become a frequent interaction phenomenon among countries or regions, and the study of population migration network is an important application field of social network and complex network. Firstly, based on the global population migration data, we constructed a network with countries or regions as nodes and population migration between countries or regions as edges. Then, we studied the structure detection and spatio-temporal analysis methods of global population migration network, and explored the structural characteristics and spatio-temporal change rules of global population migration network. The results show that the global migration network has the characteristics of scale-free, small-world and heteropogy, and the changes of network structure are closely related to geographical location, international relations and international events. This study has certain reference value for the study of population problems at global scale.

Key words: global population migration, complex network, power law distribution, community extraction, time-series change

随着科技与经济的发展，各个国家或地区之间的联系和交往越来越密切，人口迁移已成为各个国家或地区之间频繁发生的交互现象。数据处理技术的进步为人口迁移数据提供了更好的获取方式，提高了数据的真实性与可靠性。人口迁移作为世界发展的物质流、信息流、资金流、技术流的重要载体，推动着各个国家或地区的物质、资金、信息等生产要素的加速流动和优化，全球人口迁移网络化正逐步形成。

对于人口迁移的研究，目前大部分学者主要试图解释或模拟人口迁移趋势，或侧重于应用，如利用人口流动网络揭示社会结构[1-2]，将人口迁移与经济、社会事件等要素进行联系以研究其相关关系[3-4]。人口迁移研究主要关注城市[5]、国家[6]或全球[7-9]范围内的人口流动，多数学者以某个或几个城市以及国家之间的人口迁移数据为基础开展研究，只有少数学者针对全球人口迁移问题进行研究。探索全球尺度的人口迁移，能深刻反映全球人口迁移的特征，更加充分认识世界人口迁移的变化趋势。目前，国内外关于全球人口迁移的分析研究主要是通过统计方法和统计数据进行全球人口迁移数据估算以及全球人口迁移趋势预测，部分利用到复杂网络理论的研究也大多是同一时间内的对比性研究或为其他数据分析起辅助作用[10]。本文以全球人口迁移数据为基础，基于复杂网络和社会网络分析理论和方法[11-12]，对全球人口迁移网络的结构特征和基本性质进行探索，并分析了其时空变化规律，拟为全球人口问题的研究提供基础。

1　研究数据与方法

1.1　研究数据

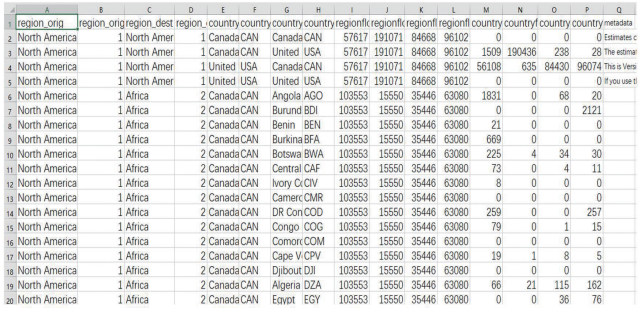
本文采用国家人力资源和社会保障部公开的国家

收稿日期：2022-07-06。
项目来源：国家自然科学基金资助项目（42171448）。
（*为通信作者）

编码数据和全球人口迁移数据两个数据集。其中，国家编码数据用于提取各个国家的坐标信息，如图1a所示，包含国家编码、国家名称和国家经纬度字段；全球人口迁移数据为1990—2010年20 a间196个国家每5 a的人口迁移量，该数据根据参考文献[13]的方法估算得到，根据连续的人口存量表，利用统计缺失数据方法估算得到1990—2010年以5 a为一个时间段196个国家之间的单边流量，每5 a约有3.8万条单边人口迁移数据，数据下载网址为：https://download.gsb.bund.de/BIB/global_flow/，数据形式如图1b所示，共包含16个字段。通过原文的数据质量评估，该数据的具体国家或地区的净移民流量与联合国公布的净移民估计数非常接近，且该数据包含196个国家或地区，占国家人力资源和社会保障部公开信息中244个国家或地区的80%，说明其具有较高的准确性和较好的覆盖量，具有分析的可行性。本文采用数据涉及的字段包括country_orig_id（人口迁出国编码）、country_dest_id（人口迁入国编码）、countryflow_1990（1990—1995年人口迁移量）、countryflow_1995（1995—2000年人口迁移量）、countryflow_2000（2000—2005年人口迁移量）、countryflow_2005（2005—2010年人口迁移量）等。

	A	B	C	D	E	F
1		code2	code	en_name	x	y
2	0	AW	ABW	Aruba	-69.9669	12.50292
3	1	AF	AFG	Afghanista	66.02712	33.82696
4	2	AO	AGO	Angola	17.54667	-12.292
5	3	AI	AIA	Anguilla	-63.0655	18.23533
6	4	AX	ALA	Aland Islan	20.06898	60.21442
7	5	AL	ALB	Albania	20.06397	41.13468
8	6	AD	AND	Andorra	1.571298	42.5518
9	7	AE	ARE	United Ara	54.33079	23.91806
10	8	AR	ARG	Argentina	-65.1223	-35.0014

a 国家编码数据



b 全球人口迁移数据

图1　研究源数据展示图

1.2　研究方法

本文的研究方法主要以复杂网络理论为基础，对基于全球人口迁移数据构建的复杂网络进行基础测度统计分析、结构探测和时空分析；从而对全球人口迁移网络的无标度和小世界特性以及基础测度和社区结构的特征展开研究，技术路线如图2所示。

作为对全球人口迁移的整体探索，本文主要考虑国家或地区之间的整体迁移关系，不考虑其中的迁入迁出，所构建的网络为无向网络，通过数据收集与处理对数据中的单边流量进行合并处理，获得整体的人口迁移量。根据数据集的4个时间段，本文以国家或地区为节点分别建立了4个无权无向网络与4个加权无向网络，无权网络的边主要体现各个国家或地区之间是否发生迁移关系，加权网络的边则表示国家或地区之间的人口迁移量。

在完成网络的构建后，对网络的测度指标进行统计，根据两种网络的不同含义，本文先对无权网络的整体属性进行测度统计，主要包括网络节点数 N 、连边数 M 、平均度<k>、图密度 D 、平均聚类系数 C 、平均路径长度 L 以及度同配系数 DA；然后通过对加权网络的部分测度进行幂律分布分析和社区检测分析来探索网络结构特征。本文采用基于模块度的Louvain算法进行社区检测，并以此为基础研究全球人口迁移网络的集群分布，揭示网络深层次的结构特征。为了直观探测网络的社区结构，本文为全球人口迁移网络节点添加空间位置信息，在地理空间中表示网络的社区划分结果。

在对网络结构有了一定认识后，再通过对网络测度的时序变化分析以及对社区检测结果的时序和空间变化分析，完成对全球人口迁移网络的时空分析，并得出相应结论。

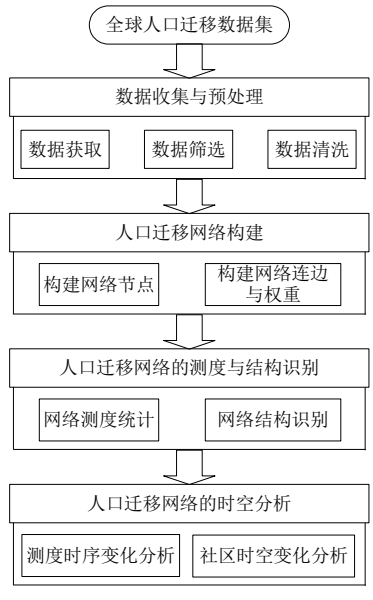


图2　技术路线图

2　研究结果与分析

2.1　网络测度统计分析

2.1.1　网络基础测度统计

网络基础测度是根据复杂网络理论中对于不同测度的定义公式计算而来，测度统计结果如表1所示，N 为该时段内发生过人口迁移现象的国家数量，M 为该时段内发生过迁移关系的国家对数量，<k>为该时段内所有国家发生迁移关系的平均数，D 为迁移关系发生的紧密程度，可以看出，1990—2010年每5 a内全球至少有80%的国家或地区参与了人口迁移，且网络有较大的边数和平均度数；由<k>可知，每个国家均与其他一半及以上的国家存在迁移关系，全球各个国家/地区之间的迁移关系非常紧密。

对于平均聚类系数和平均路径长度的分析，需要与相似随机网络的测度进行对比。因此，构建了一个节点数为196，每个节点与其他节点连接概率为50%的随机网络，该网络的测度指标统计如表2所示。通过对比可以看出，该网络具有较大的平均聚类系数与较小的平均路径长度，这也反映了该网络具有一定的小世界特性，说明各个国家之间人口迁移的连通程度较高。网络的度同配系数为负值，说明该迁移网络是异配的，具有迁移关系的国家数量相差过大的国家之间更倾向于发生迁移活动，而具有迁移关系的国家数量相近的国家则更少发生迁移活动。

表1　无权迁移网络测度指标统计结果

时间段	N	M	<k>	D	C	L	DA
1990—1995	196	9 856	100.57	0.51	0.75	1.48	−0.25
1995—2000	196	10 150	103.57	0.53	0.76	1.47	−0.24
2000—2005	196	9 677	98.74	0.50	0.75	1.49	−0.26
2005—2010	196	9 895	100.97	0.52	0.76	1.48	−0.26

表2　随机网络的测度指标统计结果

N	M	<k>	D	C	L	DA
196	9 525	97.19	0.50	0.49	1.50	−0.02

2.1.2　网络测度的幂律分布分析

在真实世界的复杂网络中，与随机网络或规则网络不同的是大多数节点倾向于与少量核心节点进行连接，这种机制使得复杂网络具有无标度性质。本文探究了网络中的资源分布规律，对加权网络节点强度与边强度进行幂律分布分析。4个时段的加权网络节点

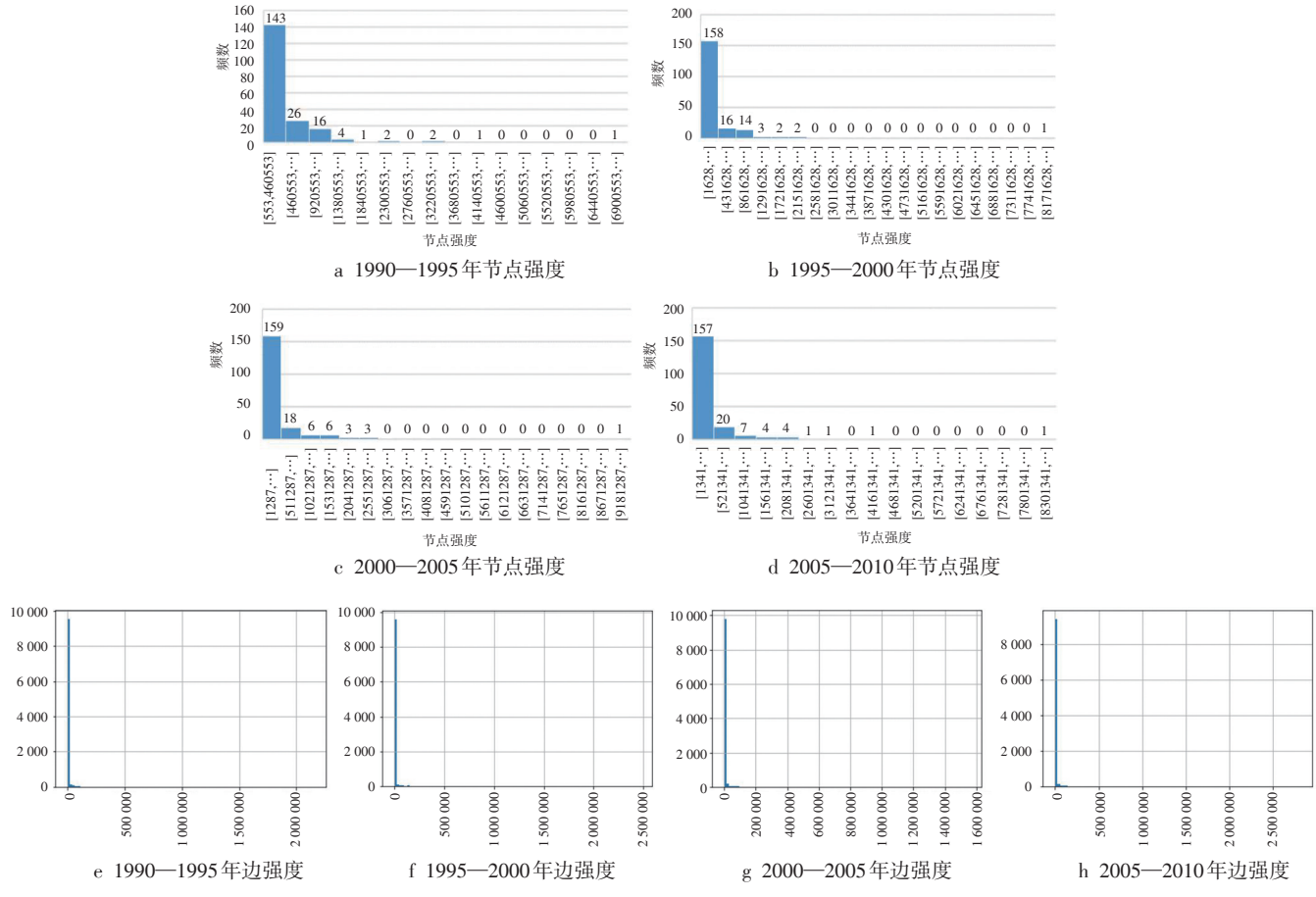


图3　加权网络节点强度与边强度分布

强度与边强度分布如图3所示，由节点强度分布图可知，不论哪个时段内，一个国家与其他国家发生迁移的次数均存在非常大的差异，只有极少数国家发生了大量的迁移关系，而绝大多数国家发生的人口迁移次数都很少，非常符合幂律分布的基本原则；由边强度分布图可知，边强度也满足幂律分布规律，与节点强度分布不同的是，边强度分布的差异更显著，只有极端少数连边的边强度值较大，绝大多数连边的边强度值都较小。

2.1.3　网络测度时序变化分析

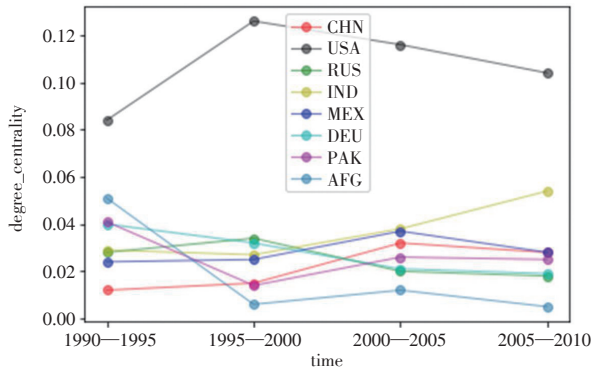
本文选择了部分度中心性较高且具有一定代表性的国家节点进行基于度中心性的测度时序变化分析，对全球人口迁移网络的时空分析方法进行深入探索，为全球人口迁移分析提供新思路、新视角。

部分国家节点度中心性时序变化如图4所示，图4a中表示了中国（CHN）、美国（USA）、俄罗斯（RUS）、印度（IND）、墨西哥（MEX）、德国（DEU）、巴基斯坦（PAK）和阿富汗（AFG）的节点度中心性变化，可以看出，美国的度中心性一直遥遥领先于其他各国，但在1995—2000年增加后，后期一直在减少；除美国外，其他各国的度中心性均在0~0.06左右变化。

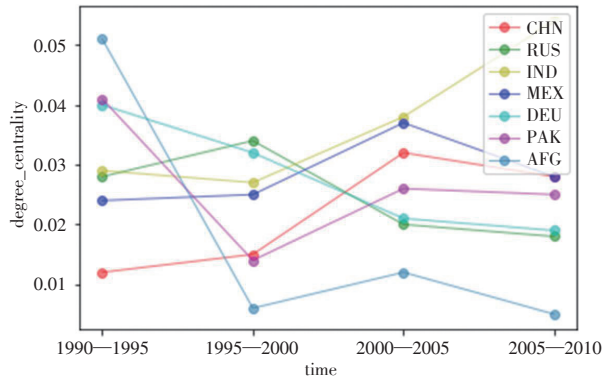
为了更加明显地分析除美国外其他国家的度中心性变化，不添加美国的其他国家节点的度中心性变化情况如图4b所示，可以看出，处于中东地区的巴基斯坦与阿富汗的度中心性变化跨度较大，这与1990—2000年发生的中东战争相对应；处于亚洲的中国（大陆）与印度的度中心性整体上在增加；处于北欧地区的俄罗斯与德国的度中心性整体上在减少，这些变化正好对应了1990—2010年中国（大陆）的经济发展、印度的人口增长以及苏联解体后俄罗斯的发展状况与两德统一后德国的发展状况。

2.2　社区探测与时空变化分析

本文利用基于模块度的Louvain算法进行社区探测。该算法能在聚类过程中遍历所有节点，相较于其他算法，对于本实验数据的聚类效果更加明显，社区探测与划分结果如图5所示，图中圆心代表网络的国家节点的地理位置中心，圆形的颜色代表节点所属社区。通过分析4个时段的网络社区划分图可知，地理位置较近的国家节点更容易被划分到同一社区中，具有相同邻海的国家虽然地理位置不相近也有较多处于同一社区的现象，在一定程度上体现了地理位置对人口迁移的影响。



a 度中心性时序变化图

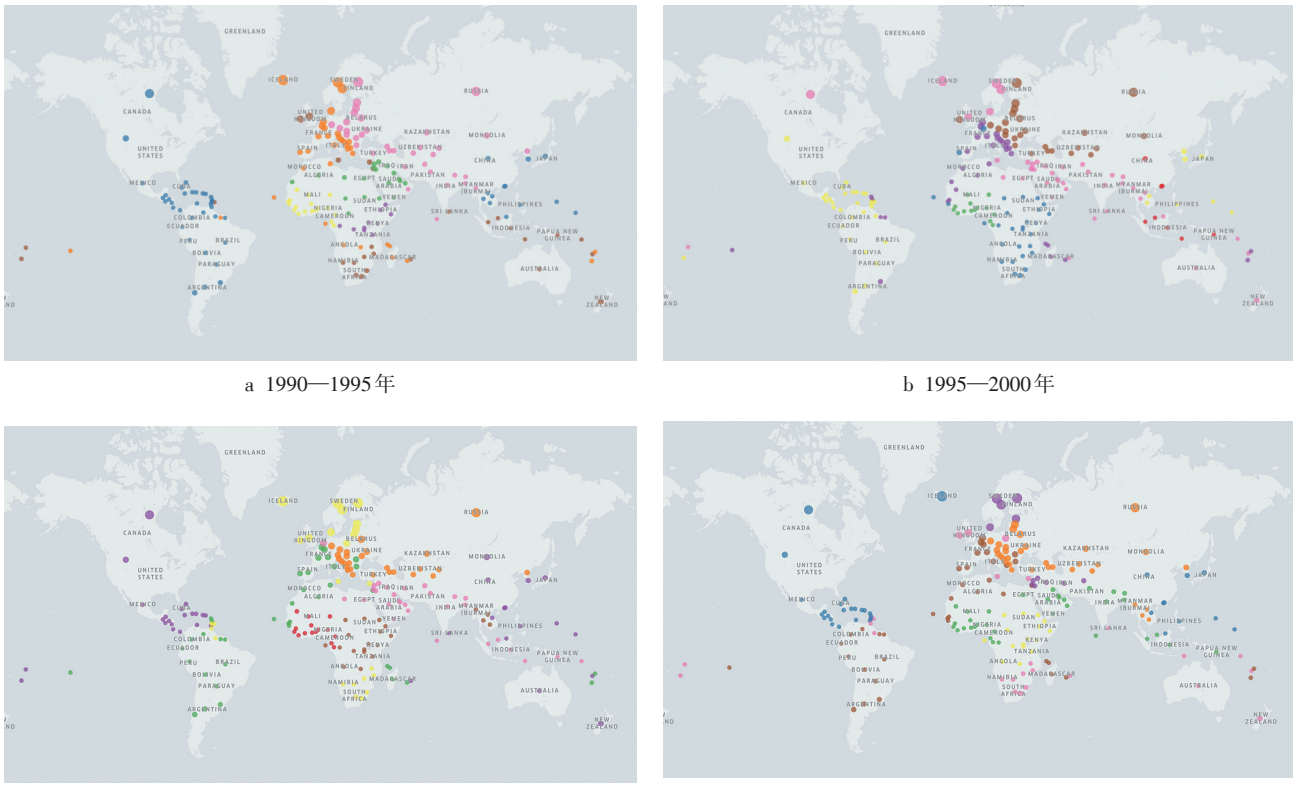


b 度中心性时序变化图（部分）

图4　部分国家节点度中心性时序变化图

为了对社区结构的时空变化进行更具体的分析，本文选择4个时段内中国（大陆）所在的社区进行变化探测。4个时段内中国（大陆）所在社区的社区划分如图6所示，可以看出，变化较大的是1995—2000年中国（大陆）所在的社区，1995—2000年中国（大陆）只与东南亚部分国家存在于同一社区，且社区内节点数量远小于另外3个时段，这与1995年中国援外工作的改革相对应，中国在1995年后加强了与东南亚部分国家的援助与合作工作，从而提高了人口迁移规模。

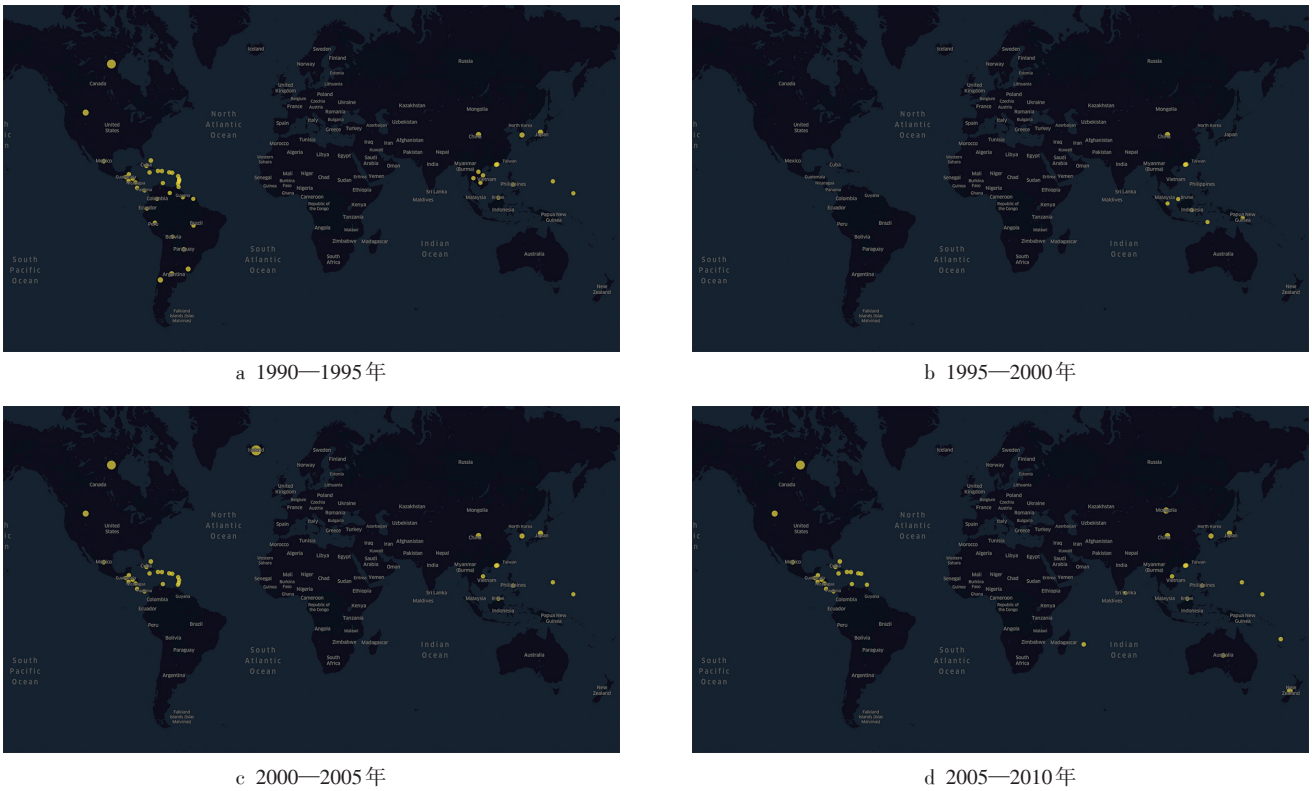
除1995—2000年的社区变化较大以外，对比其他3个时段的社区划分图可知，有许多共同的国家和地区节点，因此可通过提取社区划分结果的共同节点进行进一步分析。通过统计发现，只有中国香港与中国澳门在4次划分中一直与中国（大陆）在同一个社区，说明中国国内3个地区的人口迁移关系一直都非常密切；除了中国香港与中国澳门，还有26个国家在3个时段内均与中国（大陆）划分在同一个社区，如图7所示，可以看出，这26个国家大多出现在东南亚地区和北美地区，包括美国、加拿大、日本、韩国等。结合上文对边强度变化的分析可知，1995—2000年中国与加拿大、日本、韩国等国家之间的人口迁移量



a 1990—1995年　　b 1995—2000年

c 2000—2005年　　d 2005—2010年

图5 网络社区划分地理位置可视化展示



a 1990—1995年　　b 1995—2000年

c 2000—2005年　　d 2005—2010年

图6 中国（大陆）所在社区可视化展示

减少，正好对应了1995—2000年社区划分的明显变化；而中国与美国之间人口迁移量激增、美国与美洲其他国家人口迁移关系密切，也对应了中国与一些美洲国家多次划分在同一社区的现象。



图7　社区划分共同节点可视化展示

3　结　语

本文以全球人口迁移数据为基础，构建了全球人口迁移网络，并运用复杂网络和社会网络分析的理论与方法对人口迁移网络的测度与结构特征进行挖掘，探索分析了人口迁移网络的时空变化特征，为全球人口迁移和全球人口问题研究奠定了基础。同时，本文为全球人口迁移活动研究提供了新思路，即从复杂网络理论的基本原理出发，结合时空分析等其他理论方法对全球人口迁移特征进行深入探索与研究。

本文研究发现全球人口迁移网络具有无标度、小世界和异配等特性，即全球人口迁移在各个国家或地区之间发生得非常密集，但人口迁移总量较大的国家或地区与具有较大双边迁移量的国家对在整体中只有极小的占比，且人口迁移量差距较大的国家之间更倾向发生迁移关系。通过对社区结构进行探测与变化分析发现，全球人口迁移变化与地理位置、国际关系、国际事件具有一定联系，对于全球人口迁移的探测、分析与预测具有重要的指示意义。后续研究可考虑结合各国人口政策、国际关系以及全球航空交通等，进行多因素的综合分析，从而为全球人口问题研究提供更好的决策支持。本文对于全球人口迁移网络的特性大多给出了定性的证明，而对于其特性的强弱并没有一个定量的分析，还需在后续对于该网络的学习和研究中进行补充和完善。

参考文献

[1] Sobolevsky S, Szell M, Campari R, et al. Delineating Geographical Regions with Networks of Human Interactions in an Extensive Set of Countries[J]. Plos One, 2013, 8(12):1-10

[2] Hawelka B , Sitko I , Beinat E , et al. Geo-located Twitter as Proxy for Global Mobility Pattern[J]. Cartography and Geographic Information Science , 2014 , 41(3): 260-271

[3] Fagiolo G, Mastrorillo M. Does Human Migration Affect International Trade? A Complex Network Perspective[J]. Plos One, 2014, 9(5):97-101

[4] Sgrignoli P, Metulini R, Schiavo S, et al. The Relation Between Global Migration and Trade Networks[J]. Physica A: Statistical Mechanics and Its Applications, 2015, 417: 245-260

[5] González M, Hidalgo C, Barabási A L. Understanding Individual Human Mobility Patterns[J]. Nature, 2008, 453: 779-782

[6] Amini A, Kung K, KANG C G, et al. The Impact of Social Segregation on Human Mobility in Developing and Industrialized Regions[J]. EPJ Data Science, 2014, 3(1): 1-20

[7] Greenwood M J. Human Migration: Theory, Models, and Empirical Studies[J]. Journal of Regional Science, 1985, 25(4):521-544

[8] Fagiolo G, Mastrorillo M. International Migration Network: Topology and Modeling[J]. Physical Review E, 2013, 88(1):12 812

[9] Tranos E, Gheasi M, Nijkamp P. International Migration: A Global Complex Network[J]. Environment and Planning B: Planning and Design, 2015, 42(1):4-22

[10] Belyi A, Bojic I, Sobolevsky S, et al. Global Multi-layer Network of Human Mobility[J]. International Journal of Geographical Information Science: IJGIS , 2017 , 31(7): 1 381-1 402

[11] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用[M]. 北京:清华大学出版社, 2006

[12] Newman M E J. The Structure and Function of Complex Networks[J]. Siam Review, 2003, 45(2):167-256

[13] Abel G J, Sander N. Quantifying Global International Migration Flows[J]. Science, 2014, 343(6 178):1 520-1 522

第一作者简介：高谢庆（1999—），硕士研究生，主要研究方向为时空大数据分析、人口移动网络分析、地理多元流。