

基于新闻大数据的北极地区地缘关系研究

李 萌^{1,2}, 袁 文¹, 袁 武³, 牛方曲¹, 李汉青⁴, 胡段牧^{1,2}

(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 中国科学院大学, 北京 100049;
3. 北京理工大学计算机学院, 北京 100081; 4. 公安部第一研究所, 北京 100048)

摘要: 随着全球气候变暖, 北极地区海冰大面积消融, 引发了严重环境问题, 同时使得北极航道成为可能, 北极地区的战略地位显著提升。作为近北极国家, 北极地区国际关系变化对中国的北极政策有直接影响。全面实时地分析北极地区地缘关系及其变化特征, 对中国制定北极地区的政治、经济、外交政策具有重要指导作用。海量全球实时开放数据库的出现以及大数据技术的发展, 如GDELT新闻事件数据库提供了覆盖全世界的源于各国主要媒体服务平台的新闻数据, 为地缘关系实时监控及分析提供了可能性。本文利用GDELT数据库, 引入labeled-LDA主题分析的理论与方法, 挖掘了2013—2019年北极圈内8个国家和地区社会发展关键要素, 构建了国家(地区)交互网络, 发现了北极地区国家(地区)之间关系的演变格局。主要结论为: ① 北极地区热点新闻主题聚焦于气候变化/冰盖融化, 冰盖融化是北极地缘关系变化的主要驱动因子; ② 冰盖融化新闻热度与海冰监测数据变化存在极强的相关性; ③ 随着冰盖融化, 北极地区的社会经济军事活动激增, 其主导权的争夺日趋激烈, 总体呈现出俄罗斯、加拿大主导的格局。

关键词: 北极地区; 地缘关系; 时空数据挖掘; 主题模型; 交互网络; 大数据

DOI: 10.11821/dlxb202105004

1 引言

早在1942年, 乔治·雷纳就指出了北极地区的重要性^[1]。近年来, 随着全球气候变暖, 北极地区海冰面积逐渐减少, 在带来环境问题的同时也在一定程度上提高了北极资源的可利用性, 北极战略地位显著提升^[2], 而中国作为近北极国家, 北极地区地缘关系变化对中国的北极政策有重大影响^[3]。全面实时的分析北极地区地缘关系及其变化特征, 对中国的政治、经济、外交具有重要参考价值^[4-5]。

国内外学者对北极地区地缘关系研究主要集中在以下3个方面: ① 通过对重大事件的案例分析研究北极地区地缘关系走向^[6]; ② 通过对各国北极政策的分析研究北极地区地缘关系变化^[7-12]; ③ 通过分析各国在北极热点问题中的参与度研究北极地缘关系格局^[13-14]。然而国际关系错综复杂、瞬息万变, 涉及到政治、经济、文化等多个方面, 数据的全面性和实时性直接影响国际关系分析的准确性。仅通过重大事件案例分析、政策研究、单

收稿日期: 2019-07-20; 修订日期: 2020-04-28

基金项目: 中国科学院战略性先导科技专项(XDA23100103); 中国科学院重点部署项目(ZDRW-ZS-2017-4) [Foundation: The Strategic Priority Research Program of the Chinese Academy of Sciences, No.XDA23100103; The Key Program of the Chinese Academy of Sciences, No.ZDRW-ZS-2017-4]

作者简介: 李萌(1995-), 女, 山西阳泉人, 硕士生, 主要从事基于时空语义普遍关联的大数据分析。

E-mail: limeng18@mailsucas.ac.cn

通讯作者: 袁文(1974-), 男, 四川岳池人, 博士, 副研究员, 主要从事自然语言时空处理及检索技术、基于时空语义普遍关联的大数据计算等研究。E-mail: yuanw@lreis.ac.cn

1090-1104 页

一主题分析等静态分析方法无法准确而全面的分析北极地缘关系变化。随着大数据时代的到来,新技术的涌现和迅速发展使得海量数据的获取、存储和计算成为可能,为研究地缘关系时空演化格局提供了新机遇^[15]。新闻大数据因内容覆盖面广、获取容易、更新频率高等优点得到了众多学者的青睐。Su等基于新闻大数据分析尼泊尔地震后相关新闻覆盖量和不同国家援助情况的变化,为了解国际和地方组织及政府在震后救援方面的工作提供了新思路^[16];马明清等运用主题模型提取了“一带一路”沿线25个国家和地区相应的新闻主题词,进而对各国的稳定度和社会发展态势进行了量化分析^[17];陈小强等在构建中国与周边国家“合作—冲突”事件网络的基础上对中国及周边国家的地缘关系进行了时间和空间上的分析^[18]。但是,目前的研究大多是基于新闻数据库自身的事件类别编码,通过对骚乱、抗议、冲突、外交等固定类别的新闻数据进行分析,主题选择单一且灵活性差,无法反映国际关系变化的驱动因子。

本文基于GDELT(The Global Database of Events, Language, and Tone)数据库,根据从中抽取的新闻原始网页链接抓取新闻全文,构建北极地区新闻数据库,在此基础上利用labeled LDA主题分析方法对北极圈内8个国家和地区2013年4月—2019年12月新闻中包含的社会发展关键要素进行分析,并根据新闻事件中国与国之间的交互行为构建了国家(地区)交互网络。同时将新闻主题分析结果与北极海冰范围监测数据进行相关分析,验证新闻数据代替真实监测数据研究国家(地区)间交互网络的可行性。最终通过对国家(地区)交互网络的特征分析,从时间和空间2个维度探索北极地区国际交往主题和地缘关系发展变化。

2 数据来源与研究方法

2.1 数据来源

本文涉及北极圈内的8个国家和地区,分别是加拿大、俄罗斯、挪威、瑞典、丹麦、芬兰、冰岛和格陵兰岛。其中格陵兰岛已于2009年6月21日正式自治。作为大西洋与北冰洋的战略要冲,格陵兰岛是唯一连接亚欧、欧美环北冰洋航道的枢纽地区,对北极地区地缘关系有重要影响^[19],因此本文将其列入研究范围。此外,美国阿拉斯加州位于北极圈内,作为世界唯一的超级强权国家,奉行全球战略,深入介入北极事务,但近年其北极战略发生重大战略转向,认为北极已进入“战略竞争的新时代”,是与中国和俄罗斯进行战略竞争的“潜在走廊”^[19]。北极地缘中的“美国因素”、本国对该区域的影响及其与周边国家的关系变化研究,单一篇文章难以概全。为控制研究范围,不将问题复杂化,故未将美国列入研究范围。

本文所用的新闻数据来源于GDELT全球知识图(Global Knowledge Graph, GKG)数据库。GDELT是一个实时、开源的全球新闻事件数据库,它每15 min更新一次,实时监控100多种语言的全球各国广播、印刷和网络新闻数据,对其进行文本分析并提取出人物、地点、事件和信息来源等关键信息。GDELT事件主要类型如表1所示。

表1 GDELT事件类型

Tab. 1 GDELT event type

事件类型				
发表公开声明	上诉	表示合作意向	咨询	进行外交合作
需求	提供援助	产出	调查	进行物质合作
不被通过	拒绝	威胁	抗议	展示武力姿态
减少关系	胁迫	攻击	战斗	使用非常规的处理方式

此外, GDELT 全球知识图数据库还包含了对新闻内容的情感分析, 由 TONE 字段表示, 其得分范围为-100 (负向情感) ~100 (正向情感)。这为文章以情感导向分析北极地区国际关系提供了数据支持。

由于 GDELT 数据库自 2013 年 4 月 1 日起才实现全球新闻的每日更新, 2013 年之前新闻数据量很少, 不足以支撑基于文本主题挖掘的区域国际关系分析, 因此本文选择 2013 年作为研究起点。另一方面, GDELT 事件类别以政治为主, 冰盖融化、动物保护、航运交通等北极地区热点问题相关类别较少, 难以应用于本文。因此, 将 GDELT 数据库作为数据来源, 抽取出新闻事件的网页链接, 进而利用爬虫技术获取新闻事件网页全文, 用于挖掘新闻主题信息, 并结合事件时间、事件地点等信息, 构建北极地区新闻数据库。

2.2 研究方法

2.2.1 LDA 与 Labeled LDA 主题模型 LDA (Latent Dirichlet Allocation) 是一种无监督的文档主题生成模型, 包含词、主题和文档三层结构, 可以用来识别大规模文档集 (document collection) 或语料库 (corpus) 中潜藏的主题信息^[20]。它认为一篇文章的每个词都是通过“以一定概率选择某个主题, 并从这个主题中以一定概率选择某个词语”产生的, 假定新闻文档中词出现的概率是:

$$p(\text{word}|\text{document}) = \sum_{\text{topic}} p(\text{word}|\text{topic})p(\text{topic}|\text{document}) \quad (1)$$

主题向量 $\vec{\theta}_m$ 、词向量 $\vec{\varphi}_k$ 分别服从参数为 $\vec{\alpha}$ 、 $\vec{\beta}$ 的 Dirichlet 先验分布 $\Delta(\vec{\alpha})$ 、 $\Delta(\vec{\beta})$, w 表示词汇, z 表示主题, 则式 (1) 的联合概率分布可表示为:

$$p(\vec{w}, \vec{z}, \vec{\alpha}, \vec{\beta}) = p(\vec{w}|\vec{z}, \vec{\beta})p(\vec{z}, \vec{\alpha}) \quad (2)$$

通过 Gibbs 抽样得到形如式 (3) 的概率分布:

$$p(z_i = k|w, z) = \left(\frac{n_{m, -i}^k + \alpha_k}{\sum_{k=1}^K n_{m, -i}^{(k)} + \alpha_k} \right) \times \left(\frac{n_{k, -i}^t + \beta_t}{\sum_{t=1}^V n_{k, -i}^{(t)} + \beta} \right) \quad (3)$$

式中: $n_{m, -i}^k$ 表示新闻文档 m 中未分配到主题 k 的特征词个数; $n_{k, -i}^t$ 表示特征词 t 未分配给主题 k 的次数; $\sum_{k=1}^K n_{m, -i}^{(k)}$ 表示新闻文档中每个主题未被分配的特征词个数; $\sum_{t=1}^V n_{k, -i}^{(t)}$ 表示主题 k 的每个特征词未被分配的次数。通过已知数据及文档生成规则, 可以实现参数估计, 求出公式 (3) 中最大超参数 α 和 β 的值, 从而挖掘出新闻时空数据库中的隐含主题。LDA 主题模型使用简单、执行效率高, 广泛应用于主题识别、文本分类等领域, 但也存在主题向量难以解释、不便于调参、精度受限等问题^[21]。

Labeled LDA 模型在 LDA 的基础上引入标注数据, 采用监督学习的方法建立标签和主题之间的映射关系, 避免了主题向量难以解释的问题, 显著提高分类精度^[21]。但标注数据的获取需要耗费大量人力物力, 对数据涵盖范围广、样本规模大的情形不适用。

2.2.2 国家(地区)交互网络 国家(地区)交互网络是以国家(地区)为节点, 以交互行为做边的用来描述国家(地区)间交互模式的网络结构图。自新闻媒体兴盛以来, 大大小小的事件都会体现在新闻报道中, 国与国之间的交往亦是如此。当 2 个国家(地区)同时出现在一篇新闻报道中, 即可认为这 2 个国家(地区)参与了该新闻事件, 即他们在在该类事件中发生了交互。本文将任意 2 个国家(地区)在新闻中同时出现的关系定义为“交互”, 新闻的主题即为该次交互的主题, 基于北极地区新闻数据库构建了一个无向有权的国家(地区)交互网络 $G(V, E, W)$:

$$W=\{w_{L_1}, w_{L_2}, \cdots, w_{L_N}\} \tag{4}$$

$$w_{L_i}=\frac{\sum_t\sum_{n_i}Numarts_j}{\sum_t\sum_N\sum_{n_i}Numarts_j} \tag{5}$$

式中：点集 V 表示北极圈内的 8 个国家（地区）；边集 E 表示在时间 t 内国家（地区）间的交互集合； W 表示边的权重，即国家（地区）间交互的强弱。通过 Labeled LDA 主题分析将国家（地区）之间的交互按照新闻内容分为 N 类， t 时间内每类主题新闻量占比即为该主题的权重 w_{L_i} 。每类主题新闻量由该类新闻的源文档总数累加得到。

3 北极地区新闻主题分析

本文获取了 2013—2019 年北极圈内 8 个国家和地区的新闻数据（表 2），采用 LDA 主题模型对新闻主题进行探索性分析，通过对各主题主题词的人工解释，并结合北极文献分析结果，从中选出北极地区最受关注的 8 类主题作为目标主题。随后采用人工标注的方法对新闻数据进行标注，得到更加精细的目标主题训练数据，使用 labeled LDA 方法训练得到北极地区新闻主题分析模型，每个主题对应的频次排名前 8 的主题词如表 3 所示。

其中冰盖融化类新闻主题词以“ice”“climate”“glacier”“sheet”“cap”为主，新闻内容与北极地区气候变暖引起的冰盖融化事件和领域科学研究进展相关。而气候变化类新闻则主要描述了北极气候变暖、环境变化等内容。“climate”一词亦出现在动物保护类新闻的主题词中，与“bear”“wildlife”“animal”等词一起，构成了北极地区动物保护的

表 2 2013—2019 年北极地区新闻频次分主题统计(条)
Tab. 2 Statistics on the number of news pieces in the Arctic by topic, 2013-2019

年份	动物保护	航运交通	气候变化	经济活动	军事活动	区域治理	资源开发	冰盖融化
2013	324	398	766	625	1803	712	773	499
2014	671	575	977	1336	5422	1120	953	580
2015	956	2381	1625	1045	3616	1342	1434	1199
2016	1447	3915	1377	1195	3923	1476	1028	1716
2017	1029	1603	1326	832	4080	1460	943	1109
2018	1280	1193	2147	933	3244	1270	930	1338
2019	1589	1537	3591	991	4605	3162	1220	1940

表 3 2013—2019 年北极地区各新闻主题对应的高频词
Tab. 3 High frequency words corresponding to each news topic in the Arctic from 2013 to 2019

事件	top1	top 2	top 3	top 4	top 5	top 6	top 7	top 8
冰盖融化	ice	climate	glacier	sheet	cap	record	researcher	melting
动物保护	bear	polar	wildlife	climate	animal	ice	conservation	study
航运交通	ship	cruise	ice	polar	route	coast	expedition	Norwegian
区域治理	policy	development	region	oil	country	paper	cooperation	state
气候变化	climate	temperature	cold	weather	record	degree	polar	wind
资源开发	energy	oil	fuel	drilling	fossil	gas	company	climate
经济活动	oil	market	price	natural	industry	company	car	gas
军事活动	military	defense	force	state	missile	Nordic	war	logistics

主要内容。细观主题词分布,“oil”一词在区域治理、资源开发、经济活动类新闻中都占有较大比重。区域治理类新闻以北极地区区域间政治交流合作为主,极地丰富的石油、天然气资源使得资源开发成为区域交流合作的重要主题,也使得北极地区经济活动围绕资源开展开。除此之外,出于领土扩张、抢占资源、维护稳定等目的,军事活动和航运交通也是北极新闻的热点问题。为研究这8类目标主题间的内在联系,本文统计了各新闻主题在不同年份的新闻数,以新闻数代表其热度做相关分析,结果如表4所示。

表 4 2013—2019 年北极地区各新闻主题相关系数
Tab. 4 The correlation coefficient of various news topics in the Arctic from 2013 to 2019

	动物保护	航运交通	气候变化	经济活动	军事活动	区域治理	资源开发	冰盖融化
动物保护	1	0.621688	0.774918	0.323984	0.398627	0.758592	0.443681	0.970136
航运交通	0.621688	1	0.11132	0.37161	0.128924	0.205274	0.459066	0.668356
气候变化	0.774918	0.11132	1	0.042325	0.284297	0.924086	0.491428	0.797022
经济活动	0.323984	0.37161	0.042325	1	0.815929	0.150709	0.353416	0.206339
军事活动	0.398627	0.128924	0.284297	0.815929	1	0.462412	0.333612	0.261695
区域治理	0.758592	0.205274	0.924086	0.150709	0.462412	1	0.500057	0.794655
资源开发	0.443681	0.459066	0.491428	0.353416	0.333612	0.500057	1	0.513771
冰盖融化	0.970136	0.668356	0.797022	0.206339	0.261695	0.794655	0.513771	1

可以发现气候变化与区域治理、冰盖融化与动物保护2组主题有极强的正相关,而冰盖融化、动物保护分别与气候变化、区域治理的相关性也很强。此外航运交通与冰盖融化有着较强的正相关,经济活动与军事活动相关性较强。这8类主题间互为正相关,互相影响,但又以气候变化、冰盖融化为主,这表明气候变暖导致的冰盖融化是影响北极国际关系变化的核心动力。气候变暖带来的一系列环境问题为北极各国发展提出了新挑战,同时冰盖融化也为北极资源开发、航道开辟、经济发展带来了新机遇,机遇与挑战并存,北极地区进入以气候变化为驱动的地缘政治时代。

4 基于新闻主题的北极地区国家(地区)交互网络时空格局演化

新闻的数量在一定程度上反映了某种话题的热度,而一类新闻热度的增加与降低往往与某些实际观测数据的变化有关。为了证明用新闻数据代替真实监测数据研究国家(地区)间交互网络的可行性,本文从影响北极国际关系格局的驱动因子——冰盖融化出发,从美国国家冰雪数据中心(National Snow and Ice Data Center)获取了2013年9月—2019年9月北极地区海冰范围最小值的监测数据,将其与2013—2019年冰盖融化新闻热度数据做相关分析(图1)。

北极地区冰盖融化新闻热度与海冰最小范围监测数据呈现极强的负相关,其相关系数为-0.97 ($P < 0.001$),具有极其显著的统计学差异。从图1中可以发现,海冰范围减少的年份,冰盖融化类新闻热度

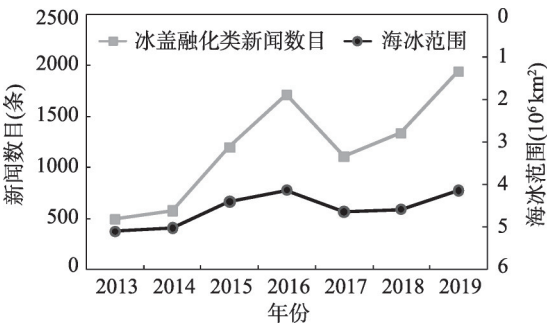


图1 2013—2019年北极地区“冰盖融化”新闻热度与海冰面积相关分析
Fig. 1 The correlation coefficient between the number of ice sheet melting news pieces and sea ice area in the Arctic from 2013 to 2019

均有所增加,二者的增减性完全对称。由此证明了用新闻数据代替真实监测数据研究国家(地区)间交互网络是可行的,新闻热度与真实监测数据极强的相关性是用新闻数据代替真实监测数据研究国家(地区)间交互网络的必要条件。本文将基于北极地区新闻主题分析数据,从时间和空间两个维度研究国家(地区)交互网络特征,进而探讨北极地区地缘关系演化格局。

4.1 北极地区国家(地区)交互网络时序演化特征

随着冰盖融化,北极地区的社会经济军事活动激增,其主导权的争夺日趋激烈,总体来看,北极地区国家(地区)交互网络朝着各国参与度更高、交互更加密切的方向发展,但又因不同交互主题的内在特征而在细节上有所不同。

区域治理主题的国家(地区)交互网络完美体现了其作为北极地区最直接热点问题的特点。如图2所示,2013—2019年北极地区这8个国家(地区)一直非常积极地参与北极区域治理,各国间均有交互,且交互的密切程度在逐年增加,至2019年,已经达到非常密切的程度。然而2018年短暂出现了俄罗斯—加拿大交互极为密切而其他国家(地区)交互较为均衡的大国主导的特征,但随着北极气候变化、资源争夺的加剧,北极圈内国家(地区)绝对不会因为国力问题放弃对北极事务的参与权,反而会根据自身特点逐步扩大参与的广度与深度。

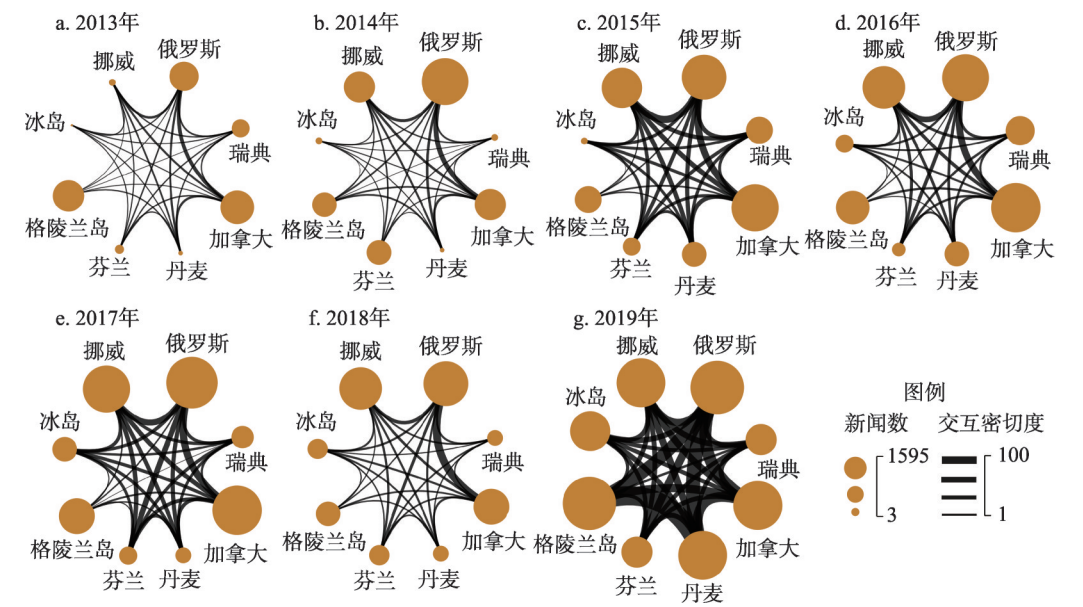


图2 2013—2019年北极地区“区域治理”国家(地区)交互网络时序变化
Fig. 2 Changes of national interaction network of regional governance in the Arctic from 2013 to 2019

大部分主题的国家(地区)交互网络都经历了由交互稀疏到交互稠密的过程,军事活动、航运交通、气候变化、冰盖融化、动物保护皆是如此,由最开始的国家(地区)参与度低、交互稀疏到后来的各国积极参与、交互变密。但这一变化过程的长短在这些主题中又有所不同。

对于与北极圈地运动、航线控制权争夺、资源开发等关系各国切身利益的主题,如军事活动、航运交通,其交互网络图分别在2014年与2015年便完成了由稀疏到稠密的转变(图3、图4)。2013—2019年北极地区军事交互逐渐加剧,最初军事交互主要发生在

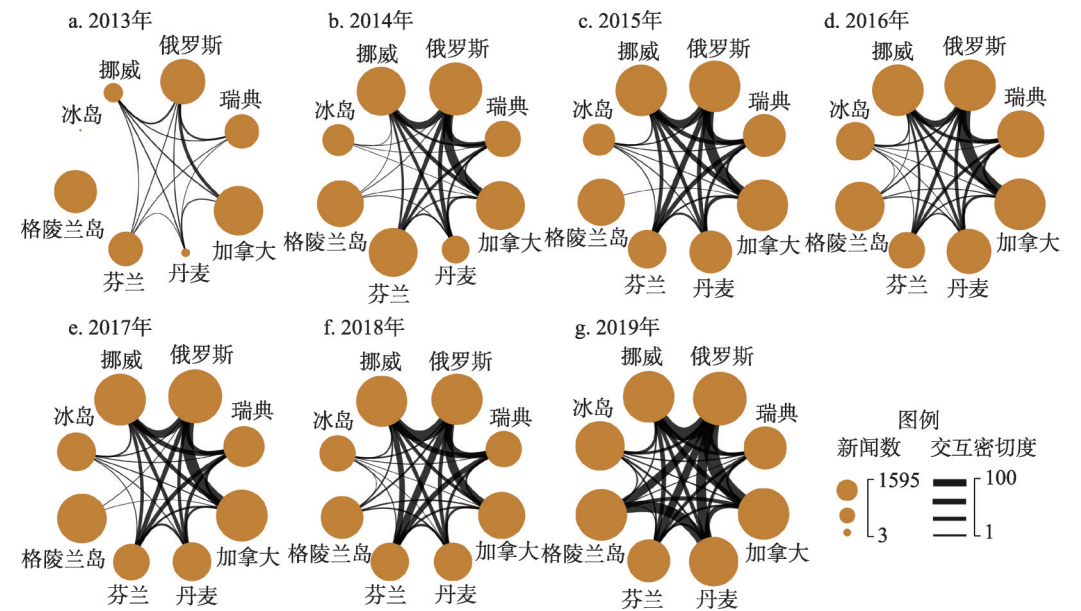


图3 2013—2019年北极地区“军事活动”国家(地区)交互网络时序变化
Fig. 3 Changes of national interaction network of military activity in the Arctic from 2013 to 2019

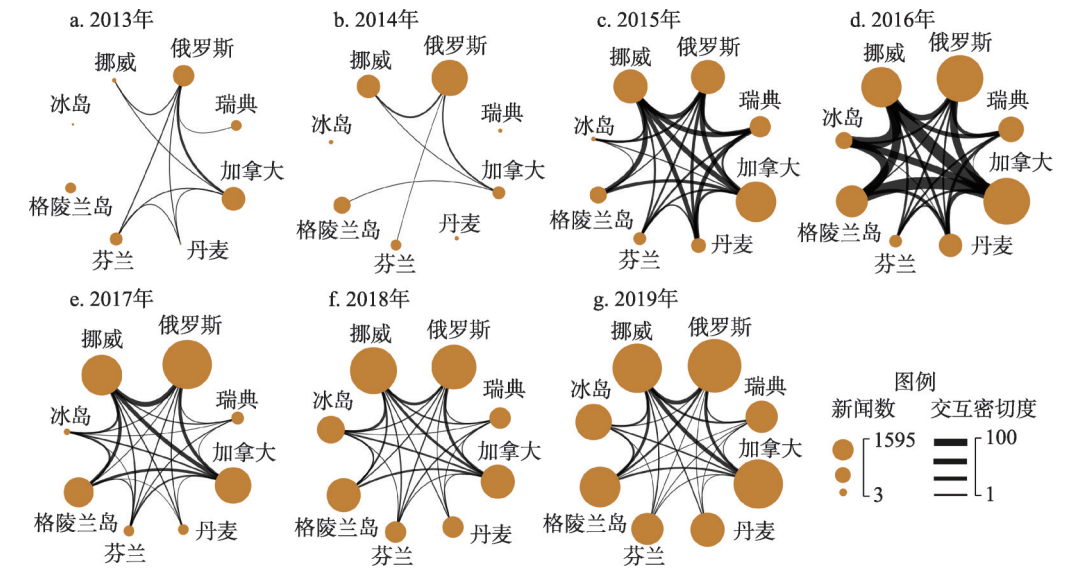


图4 2013—2019年北极地区“航运交通”国家(地区)交互网络时序变化
Fig. 4 Changes of national interaction network of channel traffic in the Arctic from 2013 to 2019

俄罗斯与加拿大之间,随后俄罗斯与挪威的军事交互也逐渐密切,至2019年俄罗斯与北极圈内大部分国家(地区)都有密切的军事交互。文献资料显示,俄罗斯在北极地区的军事活动带有明显的对抗性^[22],这表明北极地区未来将形成以俄罗斯为主的一对多的军事交互格局。航运交通主题国家(地区)交互网络时序变化与军事活动不同,航运交通网络在2015年趋于稠密,2016年是航运交通类交互最为密切的一年,在此之后虽然北极各国间仍保持有一定交互,但国家(地区)间的交互密切程度开始下降至普通水平。事

实上, 2014年以来海冰范围逐年减少, 至2016年降至7年最低水平。这为北极通航提供了绝佳的自然条件, 北极各国纷纷抓住这一机遇发展航运经济。2016年后, 北极海冰范围有所回升, 航运开发的热潮也逐渐退去, 导致其交互密切度开始下降。但航运交通在全球气候变暖的大背景下始终是北极地区待争夺资源的一种, 在很长一段时间内, 各国都将在航运交通方面保持密切的联系^[23], 当下一轮海冰范围骤降来临时, 交互密切程度将会与2016年一样急剧增加。

对于气候变化、冰盖融化、动物保护等与北极利益争夺没有直接关系的主题, 其交互网络图由稀疏走向稠密的时间跨度更长, 交互密切度较之前的主题也更低。如图5和6所示, 直至2018年, 北极各国在气候变化主题的交互才趋于密切, 而冰盖融化交互网络在2019年才达到较为稠密的结构。在这两类主题中, 格陵兰岛成为主要参与方, 积极与北极地区其他国家就气候变化、冰盖融化问题进行交互。这与它独特的地理环境有关, 格陵兰岛有84%的国土面积被冰雪覆盖, 冰盖面积达180 km², 其冰层平均厚度达2300 m, 气候变暖、冰盖融化对它有直接的影响^[10]。图7为动物保护主题国家(地区)交互网络时序变化, 虽然它在2016年便达到了各国交互较为均衡的网络结构, 但其交互密切程度整体水平偏低, 北极动物保护依旧面临着严峻的挑战。

资源开发与经济活动主题国家(地区)交互网络时序变化具有与众不同的特点(图8)。2013—2015年其交互密切程度有所增加, 但2016年突然降低, 此后再次随着时间变化逐年增加。2016年是近7年北极海冰范围最小的年份, 随着北极冰盖融化加剧, 环境威胁使得资源开发节奏变慢, 经济活动也受此影响, 交互减弱。而北极地区资源开发难度大、经济活动不发达则是这2类主题交互密切程度远低于其他主题的主要原因(图9)。

虽然国家间不同交互主题的交互密切程度随时间变化有所不同, 但总体来看与北极气候变化有极大关系, 而因气候变化引起的冰盖融化更是影响北极地区这8个国家(地区)在航运交通、资源开发、经济活动、动物保护等方面交互的重要驱动因子, 北极地区已经进入以气候变化为驱动的地缘政治时代。近年来, 随着全球气候变暖, 北极地区

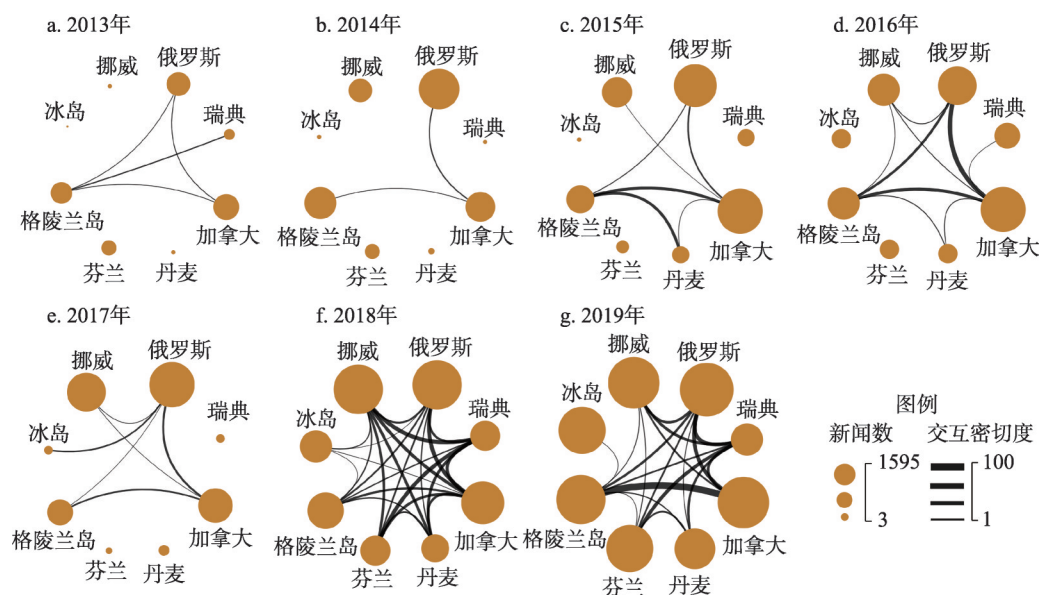


图5 2013—2019年北极地区“气候变化”国家(地区)交互网络时序变化

Fig. 5 Changes of national interaction network of climate change in the Arctic from 2013 to 2019

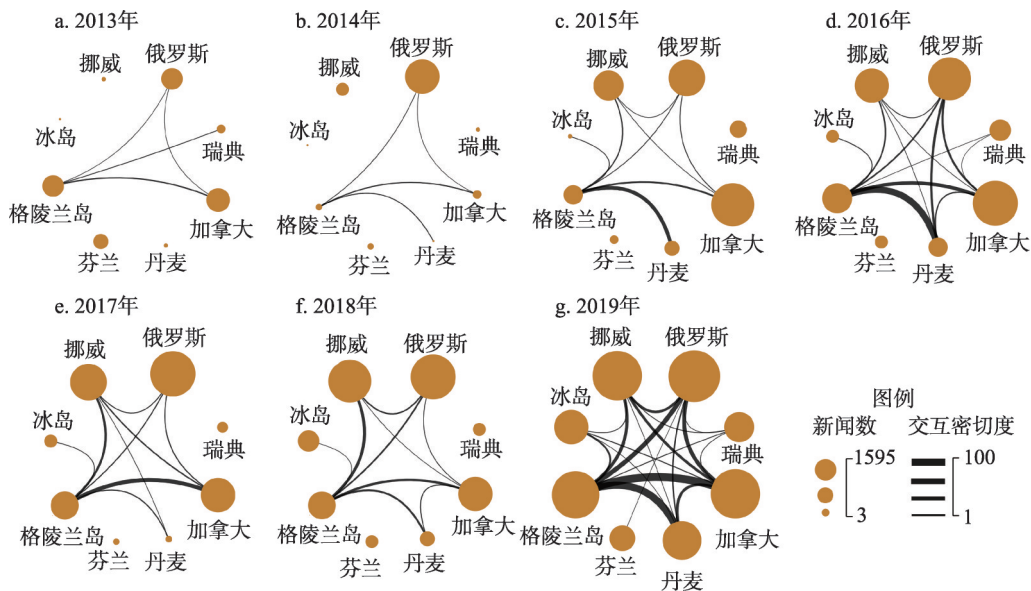


图6 2013—2019年北极地区“冰盖融化”国家(地区)交互网络时序变化
Fig. 6 Changes of national interaction network of shrunken ice sheet theme from 2013 to 2019

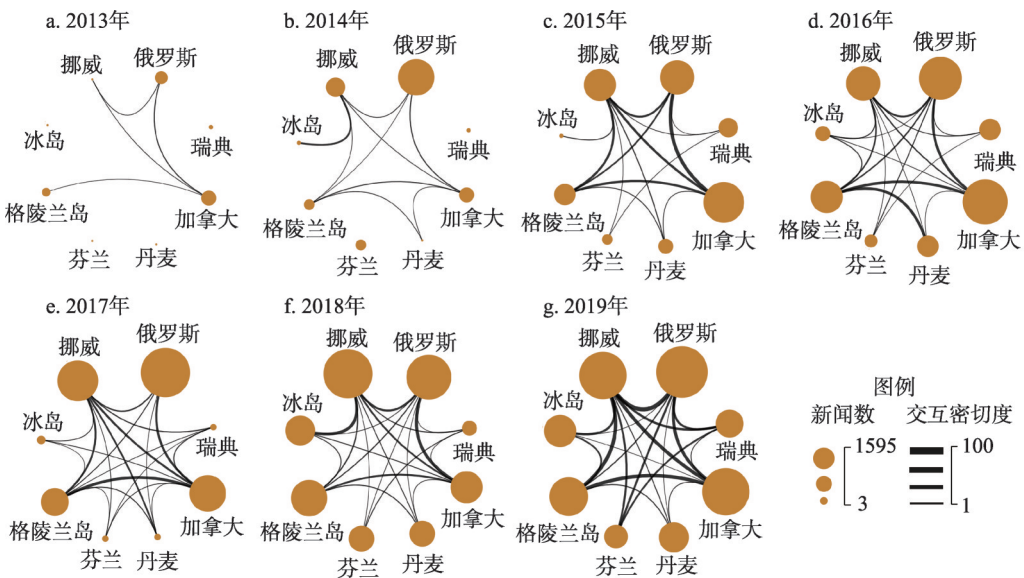


图7 2013—2019年北极地区“动物保护”国家(地区)交互网络时序变化
Fig. 7 Changes of national interaction network of animal protection in the Arctic from 2013 to 2019

海冰面积逐渐减少。这为北极航道建设、资源开发提供了极大便利, 北极利益相关各国纷纷投入其中。与此同时, 气候变暖带来的一系列环境风险也牵动着北极各国的神经, 使得北极地区国家(地区)间的交互行为合作与竞争并存、机遇与挑战同在。不过目前北极地区国家(地区)间的竞争主要建立在北极发展的角度上, 具体表现为治理框架内的有限政策和行动博弈, 北极地区地缘关系依然将向着合作共赢的方向发展。

4.2 北极地区国家(地区)交互网络空间分布格局

北极各国在综合国力、地理区位优势、利益诉求等方面的不同导致了北极地区国家

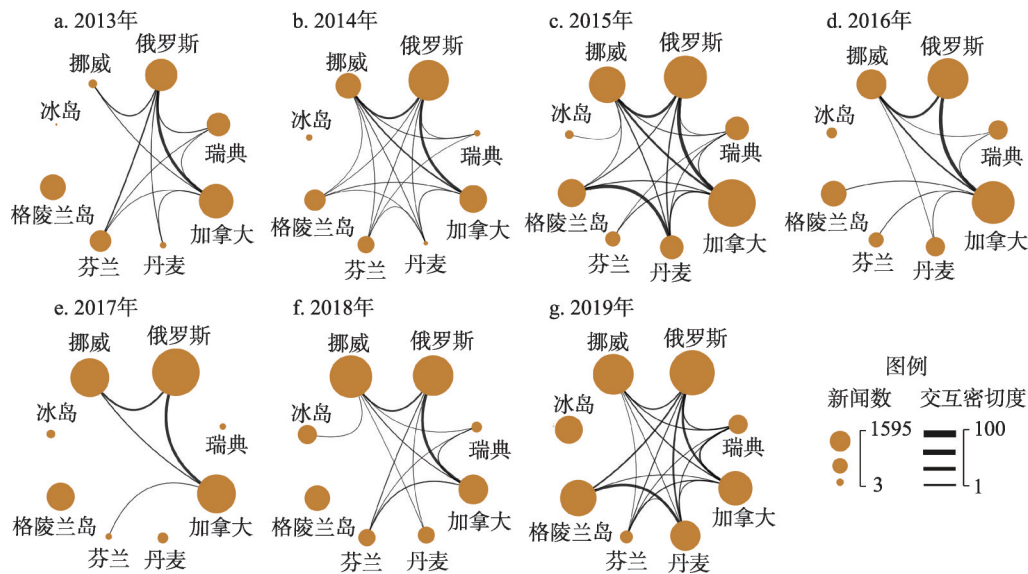


图8 2013—2019年北极地区“资源开发”国家(地区)交互网络时序变化
Fig. 8 Changes of national interaction network of resource development in the Arctic from 2013 to 2019

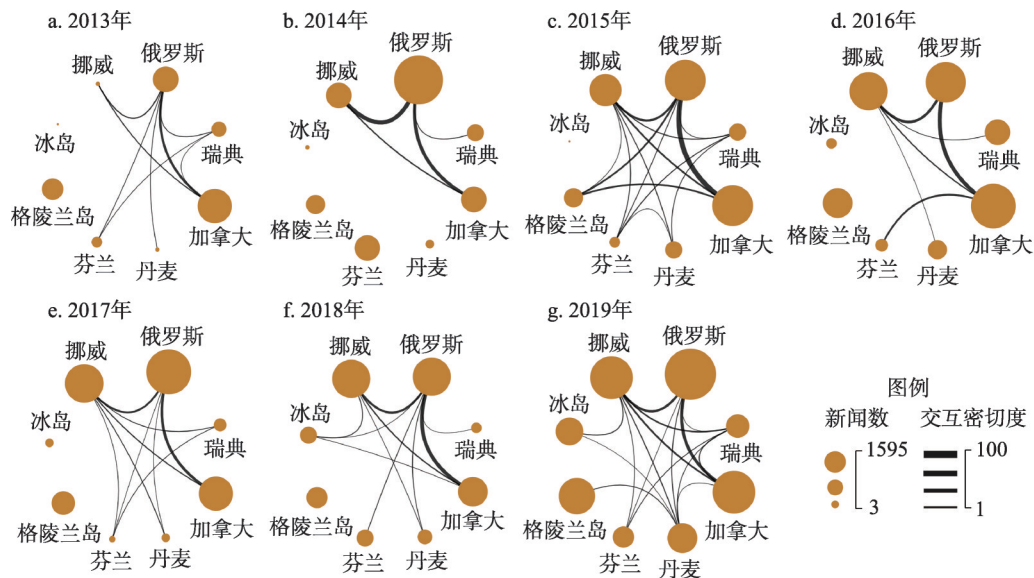


图9 2013—2019年北极地区“经济活动”国家(地区)交互网络时序变化
Fig. 9 Changes of national interaction network of economic activity in the Arctic from 2013 to 2019

(地区)交互网络独特的格局。统计2019年北极地区国家(地区)交互网络中不同情感类别的新闻数,得到表5所示的结果。计算两国交互中“积极”与“非常积极”新闻数在除“中性”外其他新闻总数中的比例,以此表示国与国交互中的正向情感占比。计算两国交互中各情感类别的新闻总数,以此表示国与国之间的交互密切程度。最后用颜色渐变表示正向情感占比,用线宽表示交互密切度,绘制2019年北极地区国家(地区)交互网络空间分布图(图10)。

从图10可以看到,该图8个节点的度均为7,北极地区这8个国家(地区)中任意2

表 5 2019 年北极地区(地区)国家交互网络情感分析结果(条)

Tab. 5 Sentiment analysis of national interaction networks in the Arctic in 2019 (piece)

交互	非常消极	消极	中性	积极	非常积极	交互	非常消极	消极	中性	积极	非常积极
加拿大—俄罗斯	1	125	2351	21	0	芬兰—瑞典	0	10	893	10	0
挪威—俄罗斯	0	54	2107	20	0	加拿大—瑞典	0	18	857	1	0
丹麦—格陵兰岛	0	10	1848	7	0	格陵兰岛—挪威	0	4	818	19	1
加拿大—格陵兰岛	1	27	1568	18	0	冰岛—俄罗斯	0	5	815	12	0
格陵兰岛—俄罗斯	0	41	1485	11	0	丹麦—瑞典	0	10	754	12	0
丹麦—俄罗斯	0	9	1435	8	0	格陵兰岛—冰岛	1	25	686	20	0
加拿大—挪威	0	12	1394	17	0	加拿大—芬兰	0	45	647	8	2
芬兰—俄罗斯	0	72	1301	4	0	加拿大—冰岛	1	8	632	16	0
挪威—瑞典	0	18	1136	14	0	丹麦—芬兰	0	4	637	8	1
俄罗斯—瑞典	0	29	1046	3	0	丹麦—冰岛	0	3	624	12	0
丹麦—挪威	0	7	957	19	0	芬兰—冰岛	0	3	523	14	0
加拿大—丹麦	0	5	953	10	0	冰岛—瑞典	0	7	487	8	0
冰岛—挪威	0	7	916	28	0	格陵兰岛—瑞典	2	15	412	0	0
芬兰—挪威	1	23	903	16	0	芬兰—格陵兰岛	0	8	380	1	0

注：数值为国家(地区)交互网络中不同情感类别新闻数;情感得分属于[-100, -10)为“非常消极”,属于[-10, -5)为“消极”,属于[-5, 5)为“中性”,属于[5, 10)为“积极”,属于[10, 100]为“非常积极”;表格按新闻总数降序排列。

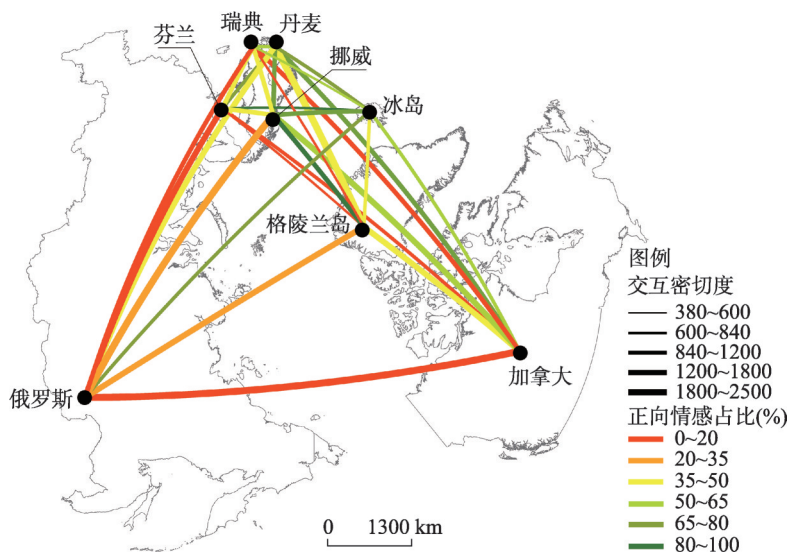


图 10 2019 年北极地区国家(地区)交互网络空间分布

Fig. 10 Spatial distribution of national interactive networks in the Arctic in 2019

个国家(地区)间均有较为密切的交互,可见其交互的频繁,各国都在积极参与北极事务。从交互的情感导向来看,2019 年北极各国的交互中非常消极和非常积极的新闻数都很少,且中性新闻占比在 92.17%~99.09%之间,占绝大多数,这表明 2019 年北极地区国际关系总体较为稳定。但依旧有 15 组交互正向情感占比不足 50%,其中俄罗斯与加拿大的负向情感交互最为明显,在此影响下,北极地区的局势依旧紧张。

俄罗斯与加拿大作为北极地区综合实力较强的 2 个国家,它们与各国交互时的情感

导向对北极地区局势有重要影响。2019年俄罗斯参与的交互占8国总交互次数的18.56%,位列第一。在北极事务上,俄罗斯一直秉持以科研为先导、以军事为后盾、以合作为手段、以先占为依托的理念^[24]。图10显示俄罗斯除了与冰岛的交互以正向情感为主外,与其他国家(地区)的交互均以负向情感为主。这与近年来以美国为首的北约国家对俄罗斯展开持续性的战略挤压有很大关系,为了寻找新的战略支点,俄罗斯愈发积极地参与北极事务,并展现出更强硬的姿态。加拿大在北极事务上的参与度仅次于俄罗斯,占14.8%。它与北极各国交互时的情感导向有极强的派系特征,具体表现在与同为北约成员国的丹麦、挪威、冰岛交互时正向情感占主导,与其他国家(地区)则负向情感占主导。

北极地区国家(地区)交互网络中,冰岛与各国的交互几乎均以正向情感为主,即使与格陵兰岛的交互正向情感占比低于50%,但依旧有43.48%。冰岛地狭人稀,是北极地区体量最小的国家。由于其综合国力有限,在北极事务中参与度最低。同时它的经济发展又高度依赖北极地区的自然资源。这一系列因素导致了它与各国交互中正向情感为主的特点。

此外,格陵兰岛因其重要的地理位置、丰富的矿产资源以及欲实现独立的政治诉求引起了各国的高度关注,其中加拿大、俄罗斯和丹麦是格陵兰岛交互网络中占比最大的3个国家,且负向情感占比较高。格陵兰岛地处北极东北航线、西北航线和中央航线的交汇处,战略地位极其重要^[25]。在各方势力纷纷争夺北极航道控制权的大背景下,加拿大因其北部北极群岛位于北极西北航道重要位置而与格陵兰岛的交互格外密切^[26]。格陵兰岛还是美俄战略较量的重要竞技场,著名的图勒空军基地就位于格陵兰岛西北海岸,在美俄关系紧张的今天,俄罗斯和格陵兰岛的交互也十分密切。而丹麦与格陵兰岛交互非常密切的原因在于格陵兰岛曾是丹麦属地,丹麦为维护其北极利益加强了对格陵兰岛的军事与经济管控力度。

综合来看,北极地区国家(地区)交互网络在空间分布上总体呈现以俄罗斯、加拿大为主导的格局。其他北极国家(地区)也根据自身特点,在热点问题、本国特色主题方面发挥作用。随着北极气候变化、资源争夺的加剧,北极圈内国家(地区)绝对不会因国力问题放弃对北极事务的参与权,反而会根据自身特点逐步扩大参与的广度与深度。

中国作为北极理事会正式观察员国也在积极参与北极事务,这不仅是国家利益的需要,也是国际法赋予中国的合法权益和时代赋予中国的国际责任。在以俄罗斯、加拿大主导的北极国际关系交互网络下,中国应以“冰上丝绸之路”建设为抓手,稳步推进与北极圈内国家(地区)的经贸合作,积极参与北极地区航道开辟、生态环境保护和资源养护与利用等事务,为北极地区社会经济繁荣和可持续发展做出积极贡献。

5 结论与讨论

5.1 结论

本文基于北极地区新闻数据库研究北极地区地缘关系演化格局,运用LDA与labeled LDA模型挖掘北极地区新闻热门主题,构建国家(地区)交互网络模型,通过对模型的分析揭示北极地区国家(地区)交互网络时空演化特征,为中国积极参与北极事务提供建议。主要结论为:

(1) 北极地区最受关注的热点问题包括军事活动、气候变化、航运交通、区域治理、冰盖融化、动物保护、资源开发与经济活动。这8类主题间互为正相关,互相影

响,但又以气候变化、冰盖融化为主。随着全球气候变暖,冰盖融化成为影响北极地缘关系变化的主要驱动因子,北极地区进入以气候变化为驱动的地缘政治时代。

(2) 冰盖融化新闻热度与海冰监测数据变化存在极强的相关性,用新闻数据代替真实检测数据研究国家(地区)间交互网络是可行的。

(3) 随着冰盖融化,北极地区的社会经济军事活动激增,其主导权的争夺日趋激烈,近7年来,北极地区国家(地区)交互网络朝着各国参与度更高、交互更加密切的方向发展,但又因不同交互主题的内在特征而在细节上有所不同。区域治理主题交互密切程度一直很高,加拿大、俄罗斯、挪威、瑞典、丹麦、芬兰、冰岛和格陵兰岛对北极区域治理均表现出十分高昂的积极性。军事活动、航运交通、气候变化、冰盖融化、动物保护主题的国家(地区)交互网络经历了由稀疏到稠密的过程,但又因其与各国切身利益密切程度不同而导致网络结构趋于稠密的速度不同。资源开发与经济活动主题网络则受气候变化、冰盖融化等自然环境变化影响经历了交互密切程度先增加、后降低,而后逐步增加的过程。

(4) 北极地区国家(地区)交互网络总体呈现以俄罗斯、加拿大为主导的空间分布格局。俄罗斯与加拿大作为北极地区中综合实力较强的大国,一直积极参与北极事务,就热点问题与周边各国进行密切交互,尤其关注区域治理、航运交通、军事活动等主题。而其他国家(地区)也根据自身特点,在热点问题、本国特色主题方面发挥作用。随着北极气候变化、资源争夺的加剧,北极圈内国家(地区)绝对不会因为国力问题放弃对北极事务的参与权,反而会根据自身特点逐步扩大参与的广度与深度。

5.2 讨论

本文从北极地区新闻热门主题出发研究北极地区地缘关系演化特征。随着全球气候变暖、冰盖融化,北极地区的社会经济军事活动激增,北极战略地位显著提升。越来越多的国家开始积极参与北极事务,使得北极地区地缘关系变得更加复杂多变。但本文侧重于研究北极圈内国家(地区)在北极事务上的交互关系,其他北极圈外国家虽然也对北极地区的发展和地缘关系走向有一定影响,但其政治诉求、利益模式、参与形式与北极圈内国家(地区)都有不同,下一步研究拟将构建更加完整的北极事务国际关系分析网络。

参考文献(References)

- [1] Ye Binhong, Cheng Yang, Wang Li, et al. A systematic review on geo-relations research in the Arctic. *Progress in Geography*, 2019, 38(4): 489-505. [叶滨鸿, 程杨, 王利, 等. 北极地区地缘关系研究综述. *地理科学进展*, 2019, 38(4): 489-505.]
- [2] Du Debin, Qin Dahe, Ma Yahua, et al. The age of cryopolitics is coming. *Bulletin of Chinese Academy of Sciences*, 2020, 35(4): 514-522. [杜德斌, 秦大河, 马亚华, 等. 冰冻圈地缘政治时代的到来. *中国科学院院刊*, 2020, 35(4): 514-522.]
- [3] He Qisong. Climate change and Arctic geopolitical game. *Foreign Affairs Review (Journal of Foreign Affairs College)*, 2010, 27(5): 113-122. [何奇松. 气候变化与北极地缘政治博弈. *外交评论(外交学院学报)*, 2010, 27(5): 113-122.]
- [4] Guan Xiaoguang, Li Zhenfu. The economic effect of the "Polar Silk Road" on China and the Arctic countries along the route. *Journal of Dalian University of Technology (Social Sciences)*, 2021, 42(1): 45-57. [关晓光, 李振福. "冰上丝绸之路"对中国与沿线北极国家的经济效应研究. *大连理工大学学报(社会科学版)*, 2021, 42(1): 45-57.]
- [5] Xiong Chenran, Wang Limao, Liang Maolin. The construction of the "Polar Silk Road" and its geo- strategic significance on China. *Academic Exploration*, 2018, 12: 26-32. [熊琛然, 王礼茂, 梁茂林. "冰上丝绸之路"建设与中国的地缘战略意义. *学术探索*, 2018, 12: 26-32.]
- [6] Byers M. Crises and international cooperation: An Arctic case study. *International Relations*, 2017, 31(4): 375-402.
- [7] Ma X M. China's Arctic policy on the basis of international law: Identification, goals, principles and positions. *Marine*

- Policy, 2019, 100: 265-276.
- [8] Li Zhenfu. Arctic countries network and China's Arctic strategy. *Northeast Asia Forum*, 2015, 24(2): 31-44, 127. [李振福. 大北极国家网络及中国的大北极战略研究. *东北亚论坛*, 2015, 24(2): 31-44, 127.]
- [9] Xie Xiaoguang, Cheng Xinbo. The construction of China-Russia "Ice Silk Road" under the influence of Russian Arctic Policy's adjustment. *Journal of Liaoning University (Philosophy and Social Sciences)*, 2019, 47(1): 184-192. [谢晓光, 程新波. 俄罗斯北极政策调整背景下的“冰上丝绸之路”建设. *辽宁大学学报(哲学社会科学版)*, 2019, 47(1): 184-192.]
- [10] Guo Peiqing, Wang Junjie. The geopolitical influence of Greenland independence. *Contemporary International Relations*, 2017(8): 58-64. [郭培清, 王俊杰. 格陵兰独立问题的地缘政治影响. *现代国际关系*, 2017(8): 58-64.]
- [11] Xiao Yang. Greenland: The anchor of Danish Arctic strategy's transformation? *Pacific Journal*, 2018, 26(6): 78-86. [肖洋. 格陵兰: 丹麦北极战略转型中的锚点? *太平洋学报*, 2018, 26(6): 78-86.]
- [12] Zhu Baolin. Interpretation of Canada's Arctic strategy from the perspective of medium sized countries. *Forum of World Economics and Politics*, 2016(4): 141-155. [朱宝林. 解读加拿大的北极战略: 基于中等国家视角. *世界经济与政治论坛*, 2016(4): 141-155.]
- [13] Ye Binhong, Cheng Yang, Wang Li, et al. The geo-economics relationship in the Arctic Region. *World Regional Studies*, 2021, 30(2): 234-244. [叶滨鸿, 程杨, 王利, 等. 北极地区地缘经济关系演变研究. *世界地理研究*, 2021, 30(2): 234-244.]
- [14] Huang Jixia, Zhang Tianyuan, Wang Li, et al. Spatial distribution patterns and the accessibility evaluation of oil and gas resources in Russia. *Acta Geographica Sinica*, 2020, 75(9): 2009-2024. [黄季夏, 张天媛, 王利, 等. 俄罗斯油气资源空间分布格局及可达性评估. *地理学报*, 2020, 75(9): 2009-2024.]
- [15] Cheng Changxiu, Shi Peijun, Song Changqing, et al. Geographic big-data: A new opportunity for geography complexity study. *Acta Geographica Sinica*, 2018, 73(8): 1397-1406. [程昌秀, 史培军, 宋长青, 等. 地理大数据为地理复杂性研究提供新机遇. *地理学报*, 2018, 73(8): 1397-1406.]
- [16] Su Y, Lan Z Y, Lin Y R, et al. Tracking disaster response and relief following the 2015 Nepal Earthquake. 2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC), 2016, 1: 495-499.
- [17] Ma Mingqing, Yuan Wu, Ge Quansheng, et al. Big data analysis of social development situation in several regions along the "Belt and Road". *Progress in Geography*, 2019, 38(7): 1009-1020. [马明清, 袁武, 葛全胜, 等. “一带一路”若干区域社会发展态势大数据分析. *地理科学进展*, 2019, 38(7): 1009-1020.]
- [18] Chen Xiaoqiang, Yuan Lihua, Shen Shi, et al. Analysis of the geo-relationships between China and its neighboring countries. *Acta Geographica Sinica*, 2019, 74(8): 1534-1547. [陈小强, 袁丽华, 沈石, 等. 中国及其周边国家间地缘关系解析. *地理学报*, 2019, 74(8): 1534-1547.]
- [19] Xu Guangmiao. Arctic order in a changing world: Generation mechanism and change logic. *Russian, East European & Central Asian Studies*, 2021, 1: 106-124, 157-158. [徐广森. 变动世界中的北极秩序: 生成机制与变迁逻辑. *俄罗斯东欧中亚研究*, 2021, 1: 106-124, 157-158.]
- [20] Wang Peng, Gao Cheng, Chen Xiaomei. Research on LDA model based on text clustering. *Information Science*, 2015, 33(1): 63-68. [王鹏, 高铨, 陈晓美. 基于LDA模型的文本聚类研究. *情报科学*, 2015, 33(1): 63-68.]
- [21] Li Wenbo, Sun Le, Zhang Dakun. Text classification based on Labeled-LDA model. *Chinese Journal of Computers*, 2008, 31(4): 620-627. [李文波, 孙乐, 张大鲲. 基于Labeled-LDA模型的文本分类新算法. *计算机学报*, 2008, 31(4): 620-627.]
- [22] Sun Qianjie, Ma Janguang. Deterrence strategy of Russia in geopolitical game in the Arctic region. *Journal of Shanghai Jiao Tong University (Philosophy and Social Sciences)*, 2017, 25(1): 14-21. [孙迁杰, 马建光. 论北极地缘政治博弈中俄罗斯的威慑战略. *上海交通大学学报(哲学社会科学版)*, 2017, 25(1): 14-21.]
- [23] Liu Dahai, Ma Yunrui, Wang Chunjuan, et al. Developments of Arctic passage resources under global climate change. *China Population, Resources and Environment*, 2015, 25(Suppl.1): 6-9. [刘大海, 马云瑞, 王春娟, 等. 全球气候变化环境下北极航道资源发展趋势研究. *中国人口·资源与环境*, 2015, 25(Suppl.1): 6-9.]
- [24] Luo Yingjie, Li Fei. The great powers' Arctic game and China's Arctic energy security: Probing into the advancement path of the "Ice Silk Road" framework. *Journal of International Security Studies*, 2020, 38(2): 91-115, 159. [罗英杰, 李飞. 大国北极博弈与中国北极能源安全: 兼论“冰上丝绸之路”推进路径. *国际安全研究*, 2020, 38(2): 91-115, 159.]
- [25] Jacobsen M. Greenland's Arctic advantage: Articulations, acts and appearances of sovereignty games. *Cooperation and Conflict*, 2020, 55(2): 170-192.
- [26] Shen Xinyi, Zhang Yu, Chen Changsheng, et al. Long-term spatial and temporal variations of sea ice in the northwest passage of the Canadian Arctic Archipelago. *Chinese Journal of Polar Research*, 2021, 33(1): 71-87. [沈心仪, 张瑜, 陈长胜, 等. 加拿大北极群岛区域西北航道海冰冰情长期时空变化特征研究. *极地研究*, 2021, 33(1): 71-87.]

Big data analysis on geographical relationship of the Arctic based on news reports

LI Meng^{1,2}, YUAN Wen¹, YUAN Wu³, NIU Fangqu¹, LI Hanqin⁴, HU Duanmu^{1,2}

(1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China; 3. School of Computer Science &

Technology, BIT, Beijing 100081; 4. First Research Institute of the Ministry of Public Security of PRC, Beijing 100048, China)

Abstract: Global warming has caused a shrinkage of the Arctic sea ice cover. It endangered the environment, but made the Arctic channel possible. Therefore, the strategic position of the Arctic has been improved significantly. As a near-Arctic country, China has formulated relevant Arctic policies, which will be directly affected by the changes in the international relations between the eight Arctic countries (region). A comprehensive and real-time analysis on the variation characteristics of the Arctic geographical relationship is required in China, which can provide instructions for formulating political, economic and diplomatic countermeasures in the Arctic. Massive global real-time open databases, such as GDELT (The Global Database of Events, Language, and Tone) global news event database, provide news data from major media in various countries. And this makes it possible to monitor the geographical relationship in real time. Based on the GDELT database and the method of labeled-LDA (Latent Dirichlet Allocation), this paper explored the key elements in social development of eight Arctic countries (region) over the past seven years. This paper also constructed the national interaction network, and identified the evolution pattern of the relations between the Arctic countries. The following conclusions were drawn: (1) Arctic news focused on climate change and ice sheet melting, which had become the main driving factor for the change of international relations in the Arctic; (2) There was a strong correlation between the number of ice sheet melting news pieces and the sea ice area; (3) With the melting of the ice sheet, the social, economic and military activities in the Arctic were booming, and the competition for dominance was becoming increasingly fierce. Generally, a pattern dominated by Russia and Canada was formed.

Keywords: Arctic; geographical relationship; spatiotemporal data mining; topic model; interactive network; big data