



基于专利数据的技术主题扩散量化研究与实现*

王 丽 刘细文

(中国科学院文献情报中心 北京 100190)

(中国科学院大学经济与管理学院图书情报与档案管理系 北京 100190)

摘要:【目的】基于专利数据对技术主题扩散进行量化研究,为提前发现或预判技术扩散提供线索。【方法】以技术主题为研究单元,基于专利引用关系构建技术扩散关系;融合技术扩散的强度、速度、广度等三个维度构建技术扩散综合测度指标;在此基础上,实现技术主题扩散量化测度模型。【结果】从石墨烯领域 100 个技术主题的扩散分析来看,模型可以快速遴选综合扩散度高的技术主题,如多种石墨烯制备方法、“石墨烯用于晶体管”等;模型的生成结果还包括扩散的具体方向,如“石墨烯用于散热”主题向“腔壳设备”等主题强力扩散,向“含锂复合氧化物”等主题快速扩散。【局限】对三个维度的技术扩散测度指标进行了线性归一化,未深入研究各指标权重。未来可以根据情报应用目的优化各指标的权重值,提升综合测度指标的鲁棒性。【结论】综合测度优化了单一扩散指标测度的不均衡性,技术主题扩散测度模型可以快速生成有效的情报信息。

关键词: 技术扩散 专利数据 主题模型 测度指标 石墨烯

分类号: TP393

DOI: 10.11925/infotech.2096-3467.2021.0915

引用本文: 王丽,刘细文.基于专利数据的技术主题扩散量化研究与实现[J].数据分析与知识发现,2022,6(6):1-10.(Wang Li, Liu Xiwen. Measuring Diffusion of Technology Topics with Patent Data[J]. Data Analysis and Knowledge Discovery, 2022, 6(6): 1-10.)

1 引言

技术的进步往往伴随着创新思想和技术的扩散,因此提前预判技术扩散有助于发现技术机会。Rogers 于 1962 年提出创新扩散理论^[1],试图解释新思想和新技术如何传播、为什么传播。在扩散研究中,大部分创新都是技术创新,所以多数情况下技术扩散成为创新扩散的同义词。专利文献作为创新活动的完整记录,被视为技术、产品、工艺创新的重要信息载体和再创新基础,专利引文反映一项技术自

诞生之日起不同时期的开发活动状况,提供了技术扩散的线索。此外,在创新发展过程中,发明人需要了解 and 获取已有的相关知识和技术,专利也成为发明人之间信息交换的重要手段。发明人为获得专利权,必须向专利局证明其发明的创造性以及相比现有技术的新颖性。专利文献中的引文即是这种证明的举证,一定程度上体现了在先技术的发展和扩散。

利用专利引用来描述技术扩散已成为创新扩散研究的一个重要方向,基于专利引用的扩散模式通常表现为专利公开后被引用的水平和分布^[2-3]。经

通讯作者(Corresponding author): 刘细文(Liu Xiwen), ORCID: 0000-0003-0820-3622, E-mail: liuxw@mail.las.ac.cn.

*本文系国家科技图书文献中心基金项目(项目编号: 2021XM59)的研究成果之一。

The work is supported by National Science and Technology Library (Grant No.2021XM59).

济合作与发展组织(Organization for Economic Cooperation and Development, OECD)的统计平台OECD.STAT认为专利可以跟踪技术扩散^[4],发现特定发明对后续发明的影响。Jaffe等^[5]在关于知识溢出的开创性论文中写道,“无迹可寻的知识流有时会留下书面记录,以专利引用的形式”。此后,专利引用被广泛用于衡量知识或技术在不同维度的扩散:地理空间^[6-9]、时间^[10]、技术领域^[11]、组织边界^[12-13]、联盟伙伴关系^[14]和社会网络^[15]等。

技术扩散包括技术产生、扩散渠道、扩散时间、技术采纳者4个要素。从技术扩散的4个要素来看,专利文献本身是技术信息的一种载体,引用途径构建了扩散渠道,专利的各种时间成为研究扩散时间的基础,这三个要素在专利视角的扩散研究中映射对象基本一致。但从技术扩散采纳者的角度来看,有国家、企业、行业/技术领域等不同的研究对象。本文以技术领域为研究对象,基于专利引用关系构建技术扩散关系,通过设计测度指标实现对技术扩散的量化研究。

在情报研究实践中,情报人员需要通过实证分析确定技术产生与发展的轨迹,预判哪些技术领域具有发展潜力。基于专利数据的技术扩散量化研究,为实证研究提供了快捷、有效的情报技术手段,可以有效发现技术领域之间的相互作用(依赖、相辅等);可以产生许多情报信息,如哪些领域存在组合效应、哪些领域存在知识溢出等;并帮助相关决策者遴选预计会对研发或产业产生关键影响的技术领域,协助规划和实施一套最佳的科技与产业政策。

2 研究现状

基于专利引用的技术扩散测度方面,常用指标有引用数量、引用时滞(Citation Lag)以及特定研究目标的相关指标,如地理空间扩散。Jaffe等^[5]通过比较施引专利与被引专利之间的地理差异衡量知识扩散,通过两者之间的引用时滞衡量知识扩散的速度;在其后续的研究中^[2,16],通过建立引用生成的指数随机模型来估算研究对象(国家、机构、领域等)的引用强度和扩散速度,其中引用强度被描述为测度专利潜能的指标,例如所含不可测技术进步的潜能,以及推动未来技术发展的“丰富性”或“重要性”等。

Han等^[17]基于国民经济行业分类(Standard Industrial Classification, SIC),通过相对引用量($c_{ij}/\sum c_{ij}$,其中 c_{ij} 是行业 i 专利被行业 j 专利引用的频次)测度传统行业与新兴行业间的无实体技术流动。Park等^[18]在比较不同领域科学知识利用和技术知识扩散差异时,分别对比引用数量及引用时滞,测度技术知识扩散的体量和速度。Verdolini等^[19]基于引用时滞结合国家技术创新特征建立指数模型来估算国家/区域间的技术扩散参数,从而评估一个国家/区域创新思想扩散的影响因素,如地理距离、技术距离、跨语言等因素。Nemet^[20]基于专利分类代码的三级技术分类结构(如国际专利分类代码的部、大类、小类三级)测度能源领域所属技术的内部扩散和外部扩散,其中不同属“部”级的“大类”被认为是对外扩散,通过“小类”来测度同属“大类”的技术间的远近,从而实现向外扩散的测度。Hohberger^[21]利用专利构建技术扩散模式,研究科学知识对扩散速度和倾向的影响,其中,以引用时滞测度扩散速度,以IPC分类代码测量一项专利的技术知识范围,以引用国家研究扩散的地理分布。贵淑婷等^[10]构建引文网络测度技术扩散速度,以IPC技术分类的第三级小类作为技术领域单元,在引用时滞基础上引入网络聚类数量和跨领域数量来计算技术领域的扩散速度,随着时间增长领域的技术扩散速度持续大幅度增加,论文随后通过网络主路径进一步对比领域的扩散速度变化。孙轶楠^[9]在国际技术流动测度研究中,以国家为研究对象,通过施引专利的国家范围、施引频次、引用时滞综合评估某个国家的国际技术扩散情况。Hur^[15]基于Hall技术分类体系(HJT Classification System)^[22]构建领域间的专利引用强度($C_{ij}/N_i \times N_j$,其中 N_i 、 N_j 为领域 i 、 j 的专利数量, C_{ij} 为领域 i 专利被领域 j 专利引用的频次)加权网络,通过领域节点的出度入度研究跨领域技术流动。Cho等^[23]以IPC大类为节点构建引文网络,利用出度中心性来测度可持续水资源领域的技术扩散影响力。罗晓梅等^[24]依据IPC分类号确定技术领域,在相对引用量基础上构建最短路径树模型,进而研究技术的跨领域扩散速度。Britto等^[25]构建30个技术领域的引用矩阵来研究技术流动,基于引用量计算技术领域赫希曼指数(Herfindahl-Hirschman Index, HHI)测度技术吸收

广度。

基于专利引用的技术扩散测度的相关研究,部分通过引用数量、引用时滞绝对量进行测度^[18];部分在引用数量、引用时滞基础上引入相对量进行扩散测度^[15,17,24];相对复杂的研究通过构建模型估算引用数量、引用时滞的函数变量以进行参数测度^[2,5,16,19]。引用数量或其变体一般用来测度扩散的强度或潜能,引用时滞或其变体用来测度扩散的速度,在技术对外扩散范围方面可以通过技术向不同分类领域扩散来测度。目前,技术扩散测度的研究对象多样化(国家、机构、行业/领域等),因而研究目的不一,多数研究为引用数量、引用时滞、对外扩散范围的两两组合测度,以技术领域为单元的扩散研究常以现有专利分类作为研究对象^[10,20-21,23,25]。本文以技术主题为研究对象,设计技术扩散强度(S)、速度(V)、广度(D)三个维度的测度指标并构建综合测度指标,最终实现对技术扩散的量化研究,该方法可以辅助快速发现具有发展潜力的技术主题。

3 技术主题扩散量化研究方案

创新发展追踪相关研究发现^[26],重大技术进步往往来自一系列技术创新,或称作技术创新集群(Innovation Cluster)。例如心脏起搏器便是一个创新集群,它是一系列技术创新的结果,包括晶体管、紧凑型电池以及其他在先发明。一项技术创新包含多个技术要素,还可能与另一项技术创新或想法密切相关。例如信息技术产业的软硬件进步使流体动力学计算机模型的发展成为可能,这对涡轮机的设计至关重要。创新集群形成的创新组合或创新方案,扩大了技术扩散效应。技术创新存在相互依赖性,从组合或者集群的角度探索创新集群的扩散效应具有理论和实证支撑。映射到专利上,一项技术创新可能由一组专利组成,技术集群可能由多组专利组合构成,因而存在不同层面的组合。技术集群中的技术创新存在技术关联性,一组专利之间存在更深层次的技术相关性或相似性,这些关联性或相似性可以通过技术主题挖掘进行不同程度的揭示。本文采用隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)模型构建技术主题,以专利集合构成的技术主题为单元进行技术扩散研究。结合名词

短语自动抽取以及文本数据通用规则构建 LDA 主题模型预处理模块^[27-28],以“技术词”(词组)替代“词”(单个词)作为训练语料,进而优化 LDA 主题模型的生成结果,提升技术主题的可读性。

专利引用关系,将技术从一个发明连接到另一个发明,在这种引用关系中,产生了技术知识的流动,技术被传递和扩散。以技术主题为研究对象的量化扩散测度可以转化为以专利集合为单元的前向引用量化和测度研究。一项专利 C 引用另一项专利 P, C 为施引专利,为区分说明,本文将 P 称为基础专利。分别对基础专利数据集和施引专利数据集进行主题挖掘,生成的主题分别称为技术主题和扩散主题,每个主题包含多件专利。基于专利引用,以技术主题为对象,形成技术扩散关系,如图 1 所示。有的主题扩散强势、扩散范围广泛,如何通过量化测度发现这些主题是本文的研究重点。

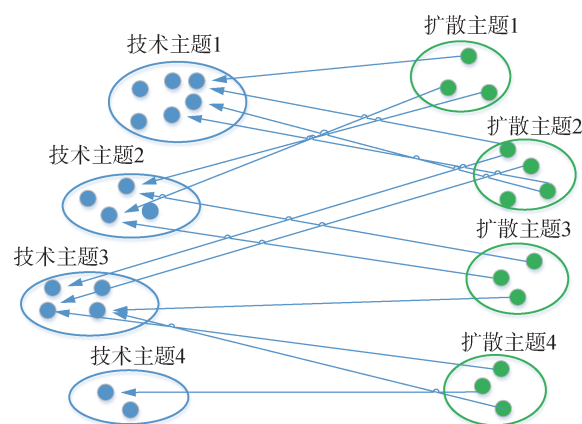


图 1 技术扩散关系示意图

Fig.1 Diagram of Technology Diffusion Relationship

围绕技术产生、扩散渠道、扩散时间、技术采纳等 4 个扩散要素^[1],构建基于专利集合的量化扩散指标,进而实现以技术主题为单元的量化扩散测度。研究的总体思路框架如图 2 所示。基于某技术领域(如石墨烯技术领域)专利数据及其施引专利数据,分别形成该领域的基础专利数据集和施引专利数据集;基于 LDA 主题模型分别对两个数据集进行主题挖掘,生成技术主题和扩散主题;然后基于专利引用关系构建技术主题扩散关系;以技术主题为对象,设计扩散测度指标,建模实现扩散的量化研究。

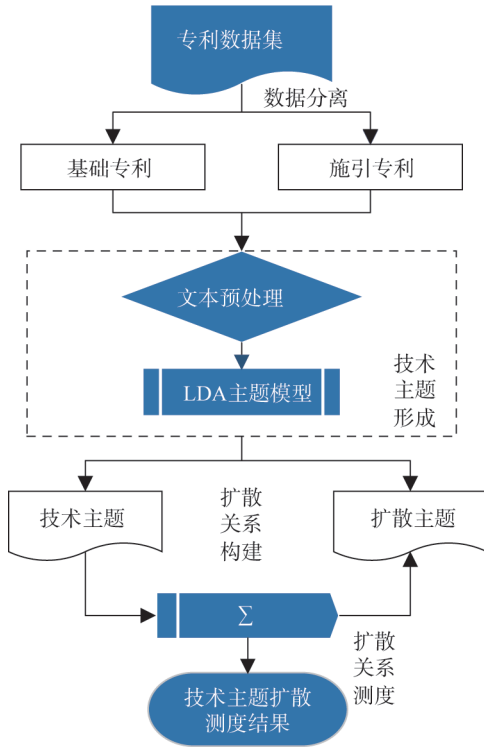


图2 研究框架

Fig.2 Research Framework

4 技术主题扩散测度指标的设计

从扩散的角度看,一个技术主题扩散到其他不同的主题,扩散的领域越多其组合创新机会越大,如人工智能技术中涉及的众多主题或子领域;反过来,一个技术主题由不同的主题扩散而来,来源越多其涵盖的技术知识越广,知识溢出的可能性越大,如合成生物的众多主题或子领域。研究技术扩散的广度和强度有助于发现这些特性。

技术扩散的速度受技术感知性的影响^[1],即技术的优势、兼容性、复杂性、可试用性以及可观察性,反之,技术扩散度可以体现技术的这些特性。渐进式技术创新,优势可感知、兼容性好、复杂性低、可试用、易观察,在短期内会带来预期结果,扩散速度较快;颠覆式技术创新,优势突出、与现有技术不兼容、复杂性高、不可试用、不易观察,不会在短期内带来直观结果,扩散速度很慢。研究技术扩散的速度有助于洞悉这些特性。

本文基于专利数据,考虑技术扩散4个要素,以技术主题为单元,研究技术主题之间的扩散关系,通

过构建技术扩散的测度指标,对扩散关系进行量化研究。为此,本文提出综合性指标来测度技术扩散能力,称其为技术扩散度(TD)。技术扩散度综合了技术扩散的强度(S)、速度(V)、广度(D)三个测度指标,如公式(1)-公式(3)所示。

$$TD_k = S + V + D = Norm\left(\sum_k C_{p,i} / \sum_T C_{p,i}\right) + Norm\left((T_{HL} - V_d) / T_{HL}\right) + Norm\left(\sum_k d_i\right) \quad (1)$$

$$V_d = \left(\sum_k C_{p,i} \times f_{p,i} / \sum_k C_{p,i} - mean(f_{p,i})\right) \quad (2)$$

$$\left(\sum_T C_{p,i}\right) / 2 = \left(\sum_{j=int(Y_s - T_{HL})+1}^{Y_s} C_{p,j}\right) + (T_{HL} - int(Y_s - T_{HL})) \times C_{p,int(Y_s - T_{HL})} \quad (3)$$

其中, TD_k 表示技术主题 k 的扩散度,扩散度越大该技术主题的扩散能力越强,通过技术主题扩散的强度(S)、速度(V)、广度(D)归一化求和来计算 TD_k 值, $Norm$ 采用线性函数归一化,即利用最大值和最小值进行标准化处理。

以技术主题为单元采用相对引用量^[17]计算扩散的强度 S , 即 $Norm\left(\sum_k C_{p,i} / \sum_T C_{p,i}\right)$, 强度越大说明该技术主题被采用的越多。 $C_{p,i}$ 表示技术主题 k 中专利 p_i 的被引数量, 通过计算技术主题 k 的被引总量占总被引量的百分比来揭示技术主题的扩散强度。以主题之间的引用情况表征扩散的广度 D , 即 $Norm\left(\sum_i d_i\right)$, 其值越大说明该技术主题被采用的广泛性越强。 d_i 表示技术主题 k 与扩散主题 i 的扩散关系, $d_i = \begin{cases} 1 & \text{存在扩散关系} \\ 0 & \text{不存在扩散关系} \end{cases}$, 扩散主题数量揭示技术主题 k 的扩散广度。

$Norm\left((T_{HL} - V_d) / T_{HL}\right)$ 表征技术扩散的速度 V , 其值越大表示该主题扩散越快。 V_d 表示技术主题 k 的扩散速度, $f_{p,i}$ 表示技术主题 k 中专利 p_i 的申请时间, 采用平均申请时间表征技术主题 k 的产生时间。 $C_{p,i}$ 表示技术主题 k 的前向施引专利 $p_d i$ 在技术主题 k 中的施引数量, $f_{p,i}$ 表示技术主题 k 的前向施引专利 $p_d i$ 的申请时间, 通过计算 $p_d i$ 申请时间在技术主题 k 中的引用平均表征技术主题 k 的被采用时间, 即 $\sum_i C_{p,i} \times f_{p,i} / \sum_i C_{p,i}$ 项。本文采用引用时滞^[17]表

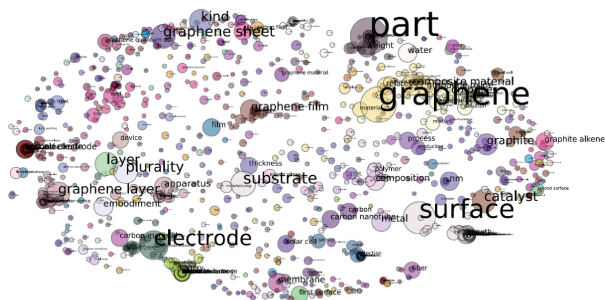


图4 石墨烯技术主题分布图

Fig.4 Topic Landscape Based on Patents Relating to Graphene

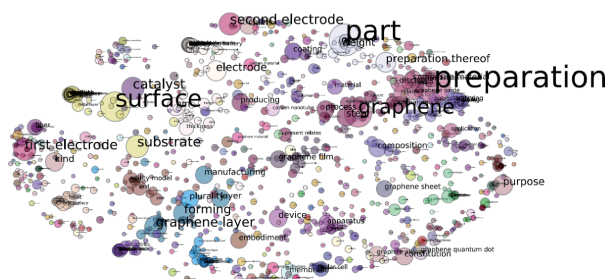


图5 石墨烯技术扩散主题分布图

Fig.5 Diffusion Topic Landscape Based on Forward Citation Patents of Graphene Patents

5.2 扩散分析

基于技术扩散测度模型生成石墨烯领域扩散指标表。扩散度排名前10的技术主题及其扩散指标如表1所示。其中,技术产生年(平均)为技术主题所含专利的平均申请年,可以披露技术主题的研究持久性和新颖性,此外,分别按扩散强度和扩散速度遴选TOP3扩散主题,以发现强力或快速扩散主题。例如,扩散度排名第一的主题O_37(石墨烯微片制备)领域,基于本文的技术主题扩散指标,该主题的平均产生时间为2013年左右,扩散强度0.50,扩散速度0.99,扩散广度0.61,综合扩散度为2.11;该主题向D_5(聚合物复合材料工艺)、D_48(电池电极)、D_34(碳纳米管)等主题强力扩散,向D_95(生物牙传感器)、D_92(偏光板的粘合膜)、D_93(保温杯)等主题快速扩散。不难发现采纳速度非常快的主题技术难度较低、实用性较强,多属于渐进式技术创新,在一定程度上佐证了本扩散模型的可用性。

作为新兴材料,高品质石墨烯制备技术是该领

域的核心主题。通过扩散分析得到的TOP10技术主题中O_37(石墨烯微片制备)、O_45(微波还原法制备石墨烯)、O_28(氧化还原法制备石墨烯)和O_79(金属/铜基体生长石墨烯薄膜)涉及石墨烯制备,O_37较为综合,O_45、O_28、O_79为具体制备方法。氧化还原法和化学气相沉积法是主流的规模化石墨烯制备方法^[30],氧化还原法(O_28)成本较低、工艺简单,但非碳杂质多、厚度不易控制、存在环境污染等问题,改进方法(O_45)致力于改善这些缺陷;化学沉积法(O_79)可以规模化制备高质量石墨烯,但剥离和转移还存在一定挑战,成本较高,基体改进为其研究重点。

从情报方法应用角度来讲,本研究目的旨在为技术机会发现提供一条可量化途径,案例研究的实验数据检索时间为2018年10月,根据2019年后的石墨烯研发和市场反观案例分析结论,以验证该方法的有效性。扩散度作为综合测度指标,可以有效发现被引用次数不高的主题。例如,在TOP10技术主题中,O_13(石墨烯层基器件)、O_57(石墨烯核壳结构材料)、O_3(石墨烯用于散热/热管理)等技术产生年较新,尽管扩散强度不高,但扩散速度非常快、扩散领域广泛,从100个技术主题中脱颖而出。从石墨烯研发和市场来看:2020年欧洲石墨烯旗舰计划启建二维材料实验性试产线(2D-EPL)^[31],旨在将石墨烯及相关层状材料集成到半导体平台中,建立电子、光子和传感器等器件的原型生产;2020年,石墨烯在散热领域产业化应用成果丰硕,华为、小米等头部企业纷纷发布石墨烯导热膜、石墨烯导热片等相关产品^[32];此外,2019年后石墨烯核壳结构材料的重大研发进展或成果不多,但石墨烯良好的性能使得石墨烯核壳结构材料在催化^[33]、生物^[34]、光电^[35]和能源存储^[36]等众多领域展现了广阔的应用前景。通过综合测度技术主题扩散情况,可以快速找出一些被引次数不高但极具潜力的技术主题,为进一步情报研究提供线索和证据。

基于扩散指标值生成多维度的可视化图表,有利于直观地观察技术扩散关系、强度、广度等。基于引用关系数据的石墨烯领域技术主题扩散如图6所示。图6(a)直观地显示了扩散度排名前20的技术主题向哪些主题进行了扩散,以及向这些领域扩散

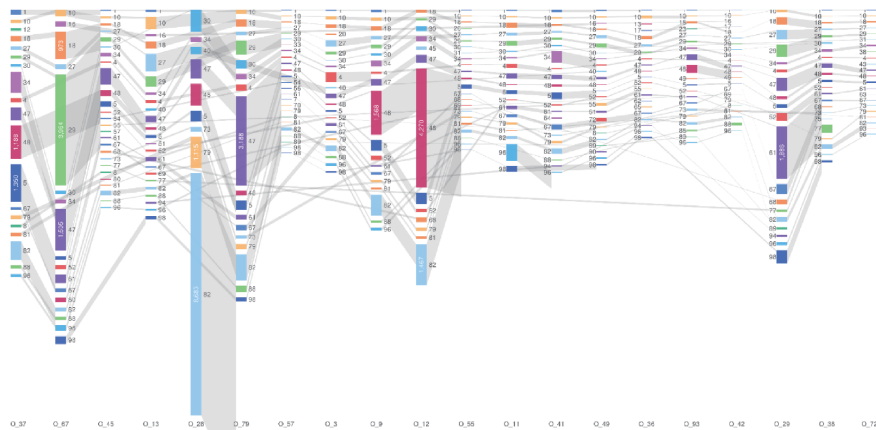
表1 石墨烯领域的技术主题扩散分析(TOP 10)
Table 1 Diffusion Analysis of Technology Topics in the Graphene Field

排名	石墨烯技术主题	技术产生年 (平均)	技术 扩散度	扩散 强度	扩散 速度	扩散 广度	扩散主题(按强度)TOP 3	扩散主题(按速度)TOP 3
1	O_37:石墨烯 微片制备	2013	2.11	0.50	0.99	0.61	D_5:聚合物复合材料工艺 D_48:电池电极 D_34:碳纳米管	D_95:生物牙传感器 D_92:偏光板的粘合膜 D_93:保温杯
2	O_67:石墨烯 用于晶体管	2013	2.03	0.72	0.81	0.50	D_29:半导体器件(晶体管) D_47:石墨烯薄膜 D_18:衬底结构	D_0:3d打印增强丝 D_21:发光元件 D_69:能量吸收装置
3	O_45:微波还原法 制备石墨烯	2013	2.02	0.18	0.85	1.00	D_47:石墨烯薄膜 D_48:电池电极 D_5:聚合物复合材料工艺	D_98:LED器件 D_35:生产生物燃料 D_32:(藻类)蛋白提取
4	O_13:石墨烯层基器件 (传感器、探测器等)	2014	1.99	0.27	0.88	0.83	D_27:传感器 D_10:纳米孔传感/探测器 D_29:半导体器件(晶体管)	D_92:偏光板的粘合膜 D_74:纳米管织物层 D_93:保温杯
5	O_28:氧化还原法 制备石墨烯	2015	1.98	1.00	0.92	0.06	D_82:石墨烯复合材料 D_79:石墨烯原材料配比 D_30:石墨烯成本问题	D_97:细胞培养 D_94:信息存储单元 D_95:生物牙传感器
6	O_79:金属/铜基体 生长石墨烯薄膜	2013	1.92	0.59	0.83	0.50	D_47:石墨烯薄膜 D_82:石墨烯复合材料 D_29:半导体器件(晶体管)	D_92:偏光板的粘合膜 D_52:热致发声装置 D_94:信息存储单元
7	O_57:石墨烯核壳 结构材料	2014	1.90	0.04	0.92	0.94	D_82:石墨烯复合材料 D_34:碳纳米管 D_47:石墨烯薄膜	D_96:PCB/LED散热 D_93:保温杯 D_95:生物牙传感器
8	O_3:石墨烯用于 散热/热管理	2015	1.88	0.17	0.99	0.72	D_4:腔壳设备 D_27:传感器 D_82:石墨烯复合材料	D_49:含锂复合氧化物 D_88:二维材料复合膜制备 D_87:加热元件
9	O_9:超级电容 电极材料	2014	1.88	0.34	0.92	0.61	D_48:电池电极 D_82:石墨烯复合材料 D_5:聚合物复合材料工艺	D_97:细胞培养 D_0:3d打印增强丝 D_98:LED器件
10	O_12:电池 负极材料	2014	1.80	0.59	0.93	0.28	D_48:电池电极 D_82:石墨烯复合材料 D_5:聚合物复合材料工艺	D_74:纳米管织物层 D_95:生物牙传感器 D_52:热致发声装置

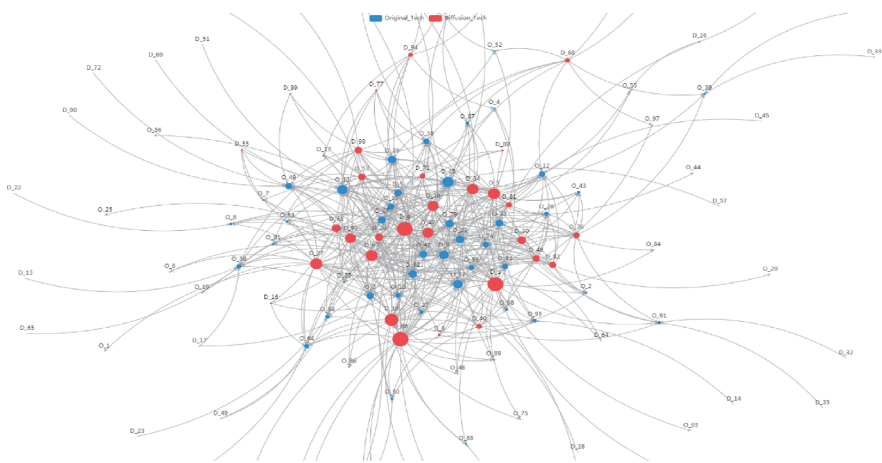
的强度如何。例如技术主题O_67(石墨烯用于晶体管),向D_29(半导体器件)、D_47(石墨烯薄膜)、D_18(衬底结构)等主题进行强力扩散。图6(b)直观地显示了各主题扩散和采用的广度。例如技术主题O_45(微波还原法制备石墨烯)、O_13(石墨烯层基器件/传感器、探测器等)、O_9(超级电容电极材料)等的扩散广度较高,表明这些领域组合创新的机会较大。再如扩散主题D_88(二维材料复合膜制备)、D_4(腔壳设备)、D_1(纳米颗粒合成方法)等采纳了较多的技术主题,表明这些领域涵盖的技术知识广泛,有知识溢出的可能。

6 结 语

本文基于专利数据从技术扩散的角度探索和研究技术价值、发展方向、再创新机会以及组合创新机会等,为发现技术机会的情报方法提供了一条有效路径。构建的技术扩散测度模型可以快速生成有效的情报信息,通过扩散测度可遴选扩散强劲的技术主题,同时发现其扩散的具体技术方向;通过反向观察技术采用,可发现知识范围广泛的扩散主题,有助于发现技术组合。本文综合技术扩散强度、速度、广度进行技术扩散量化研究,综合测度优化了单一扩散指标测度的不均衡性,例如可以通过综合测度发



(a)



(b)

图6 基于引用关系数据的技术主题扩散

Fig.6 Technology Topics Diffusion Based on Patent Citation Data

现扩散强度不明显但扩散快速、扩散领域广泛的技术主题,未来可以进一步优化各指标的权重值,提升综合测度指标的稳定性。

成功的技术扩散形成了技术和采纳者之间的扩散关系。本文基于专利引用关系对已经形成的扩散关系进行量化研究,而那些尚未形成扩散关系的技术是否在一定范围或某些集群内存在潜在的扩散性?提前预判一项技术创新的潜在扩散方向,有助于促进创新发展、最大化创新功效,同时,通过预测多项创新技术的潜在扩散方向,有助于反向发现更有价值的创新技术。结合推荐系统、深度学习、链路预测等技术或方法,未来可以进一步研究专利之间的潜在引用关系以实现技术潜在扩散的预判,这对于技术机会发现有着现实意义。本文构建的技术扩

散测度模型为进一步预测潜在技术扩散提供了量化研究的基础。

参考文献:

- [1] Rogers E M. Diffusion of Innovations [M]. The Fifth Edition. New York: The Free Press, 2003.
- [2] Jaffe A B, Trajtenberg M. Flows of Knowledge from Universities and Federal Laboratories: Modeling the Flow of Patent Citations over Time and Across Institutional and Geographic Boundaries [J]. PNAS, 1996, 93(23): 12671-12677.
- [3] Bacchiocchi E, Montobbio F. Knowledge Diffusion from University and Public Research. A Comparison Between US, Japan and Europe Using Patent Citations[J]. The Journal of Technology Transfer, 2009, 34(2): 169-181.
- [4] OECD. STAT. Technology Diffusion[OL]. [2021-04-01]. <https://>

- stats.oecd.org/Index.aspx?DataSetCode=PAT_DIFF.
- [5] Jaffe A B, Trajtenberg M, Henderson R. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations[J]. *The Quarterly Journal of Economics*, 1993, 108(3): 577-598.
- [6] Almeida P, Kogut B. Localization of Knowledge and the Mobility of Engineers in Regional Networks[J]. *Management Science*, 1999, 45(7): 905-917.
- [7] Alcácer J, Gittelman M. Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations[J]. *Review of Economics and Statistics*, 2006, 88(4): 774-779.
- [8] Fischer M M, Scherngell T, Jansenberger E. Geographic Localisation of Knowledge Spillovers: Evidence from High-Tech Patent Citations in Europe[J]. *The Annals of Regional Science*, 2009, 43(4): 839-858.
- [9] 孙轶楠. 基于专利引用的国际化技术流动测度研究——以玉米育种技术领域为例[J]. *情报杂志*, 2016, 35(1): 51-56. (Sun Yinan. Research on the Measurement of International Technology Flow Based on Patent Citation—A Case Study in the Field of Maize Breeding[J]. *Journal of Intelligence*, 2016, 35(1): 51-56.)
- [10] 贵淑婷, 彭爱东. 基于专利引文网络的技术扩散速度研究[J]. *情报理论与实践*, 2016, 39(5): 40-45. (Gui Shuting, Peng Aidong. Research on the Speed of Technology Diffusion Based on Patent Citation Network[J]. *Information Studies: Theory & Application*, 2016, 39(5): 40-45.)
- [11] 黄鲁成, 刘玉敏, 吴菲菲, 等. 基于专利全引用信息的技术知识扩散特征研究——以石墨烯技术为例[J]. *科学学与科学技术管理*, 2017, 38(4): 149-161. (Huang Lucheng, Liu Yumin, Wu Feifei, et al. Research Framework of Technical Knowledge Diffusion Characteristics Based on Patent Full Citation Information: The Example of Graphene[J]. *Science of Science and Management of S.&T.*, 2017, 38(4): 149-161.)
- [12] Peri G. Determinants of Knowledge Flows and Their Effect on Innovation[J]. *Review of Economics and Statistics*, 2005, 87(2): 308-322.
- [13] 孙冰, 徐晓菲, 苏晓. 技术扩散主路径及核心企业的识别研究——以手机芯片专利引文网络为例[J]. *情报学报*, 2019, 38(2): 201-208. (Sun Bing, Xu Xiaofei, Su Xiao. Identification of the Main Path of Technology Diffusion and Core Enterprises: A Case Study of Patent Citation Network in Mobile Phone Chips[J]. *Journal of the China Society for Scientific and Technical Information*, 2019, 38(2): 201-208.)
- [14] Gomes-Casseres B, Hagedoorn J, Jaffe A B. Do Alliances Promote Knowledge Flows?[J]. *Journal of Financial Economics*, 2006, 80(1): 5-33.
- [15] Hur W. The Patterns of Knowledge Spillovers Across Technology Sectors Evidenced in Patent Citation Networks[J]. *Scientometrics*, 2017, 111(2): 595-619.
- [16] Jaffe A B, Trajtenberg M. International Knowledge Flows: Evidence from Patent Citations[J]. *Economics of Innovation and New Technology*, 1999, 8(1-2): 105-136.
- [17] Han Y J, Park Y. Patent Network Analysis of Inter-Industrial Knowledge Flows: The Case of Korea Between Traditional and Emerging Industries[J]. *World Patent Information*, 2006, 28(3): 235-247.
- [18] Park H W, Kang J. Patterns of Scientific and Technological Knowledge Flows Based on Scientific Papers and Patents[J]. *Scientometrics*, 2009, 81(3): 811-820.
- [19] Verdolini E, Galeotti M. At Home and Abroad: An Empirical Analysis of Innovation and Diffusion in Energy Technologies[J]. *Journal of Environmental Economics and Management*, 2011, 61(2): 119-134.
- [20] Nemet G F. Inter-Technology Knowledge Spillovers for Energy Technologies[J]. *Energy Economics*, 2012, 34(5): 1259-1270.
- [21] Hohberger J. Diffusion of Science-Based Inventions[J]. *Technological Forecasting and Social Change*, 2016, 104: 66-77.
- [22] Hall B, Jaffe A, Trajtenberg M. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools[R]. National Bureau of Economic Research, 2001.
- [23] Cho I, Park J, Heo E. Measuring Knowledge Diffusion in Water Resources Research and Development: The Case of Korea[J]. *Sustainability*, 2018, 10(8): 2944.
- [24] 罗晓梅, 王亢抗, 乔铮. 基于有效距离的新兴技术跨领域扩散模型[J]. *技术经济*, 2019, 38(4): 10-14. (Luo Xiaomei, Wang Kangkang, Qiao Zheng. Construction of Interdisciplinary Diffusion Model of Emerging Technology Based on Effective Distance[J]. *Technology Economics*, 2019, 38(4): 10-14.)
- [25] Britto J N D P, Ribeiro L C, Araújo L T, et al. Patent Citations, Knowledge Flows, and Catching-Up: Evidences of Different National Experiences for the Period 1982-2006[J]. *Science and Public Policy*, 2020, 47(6): 788-802.
- [26] Globe S, Levy G W, Schwartz C M. Key Factors and Events in the Innovation Process[J]. *Research Management*, 1973, 16(4): 8-15.
- [27] 王丽, 沈湘. 文本预处理后的LDA模型主题发现与技术演进研究[J]. *农业图书情报*, 2019, 31(4): 19-28. (Wang Li, Shen Xiang. Research of Topics Discovery and Tech Evolution Based on Text Preprocessed LDA Model[J]. *Agricultural Library and Information*, 2019, 31(4): 19-28.)
- [28] 王丽, 邹丽雪, 刘细文. 基于LDA主题模型的文献关联分析及可视化研究[J]. *数据分析与知识发现*, 2018, 2(3): 98-106. (Wang Li, Zou Lixue, Liu Xiwen. Visualizing Document Correlation Based on LDA Model[J]. *Data Analysis and Knowledge Discovery*, 2018, 2(3): 98-106.)
- [29] 万秀明, 周斌. 测算期刊“半衰期”与期刊科学管理[J]. *大学图书馆情报学刊*, 1996, 14(1): 42-44. (Wan Xiuming, Zhou Bin. Calculation of Half-Life Period of Scientific Journal and Scientific Administration of Periodicals[J]. *Journal of Academic*

- Library and Information Science, 1996, 14(1): 42-44.)
- [30] Lin L, Peng H, Liu Z. Synthesis Challenges for Graphene Industry[J]. Nature Materials, 2019, 18(6): 520-524.
- [31] Graphene Flagship. 2D-Experimental Pilot Line[OL]. [2021-09-22]. <https://graphene-flagship.eu/innovation/pilot-line/>.
- [32] 赛迪工业和信息化研究院. 2021年中国石墨烯产业发展趋势[OL]. [2021-09-22]. <https://www.ccidgroup.com/info/1105/32697.htm>. (CCID. Trend of Graphene Industry Development in China in 2021[OL]. [2021-09-22]. <https://www.ccidgroup.com/info/1105/32697.htm>.)
- [33] 中国科学院重大科技任务局. 铜纳米颗粒/石墨烯核壳结构材料催化研究获进展[OL]. [2021-09-22]. http://www.bmrpd.cas.cn/gzky/kyjz/clny/201602/t20160225_4536698.html. (Bureau of Major R&D Programs, Chinese Academy of Sciences. Progress in Catalysis of Copper Nanoparticle@graphene Core-Shell Materials[OL]. [2021-09-22]. http://www.bmrpd.cas.cn/gzky/kyjz/clny/201602/t20160225_4536698.html.)
- [34] Liu Z K, Li S K, Xia X, et al. Recent Advances in Multifunctional Graphitic Nanocapsules for Raman Detection, Imaging, and Therapy[J]. Small Methods, 2020, 4(4): 1900440.
- [35] Naseri T, Pour-Khavari F. Bimetallic Core-Shell with Graphene Coating Nanoparticles: Enhanced Optical Properties and Slow Light Propagation[J]. Plasmonics, 2020, 15(4): 907-914.
- [36] Sarigamala K K, Shukla S, Struck A, et al. Rationally Engineered

3D-Dendritic Cell-Like Morphologies of LDH Nanostructures Using Graphene-Based Core-Shell Structures[J]. Microsystems & Nanoengineering, 2019, 5: 65.

作者贡献声明:

王丽:设计研究方案,模型建立,代码实现,案例分析,论文撰写;
刘细文:提出研究思路,论文最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: wangli@mail.las.ac.cn。

[1] 王丽. Tech Diffusion.ipynb. 技术主题扩散测度模型。

[2] 王丽. Graphene.csv. 石墨烯专利数据。

[3] 王丽. Graphene_TTopic.csv. 石墨烯技术主题。

[4] 王丽. Graphene_DTopic.csv. 石墨烯扩散主题。

[5] 王丽. Graphene_Topic_tech_diff_Original.csv. 石墨烯技术主题扩散指标总表。

收稿日期:2021-08-27

收修改稿日期:2021-11-30

Measuring Diffusion of Technology Topics with Patent Data

Wang Li Liu Xiwen

(National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

(Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China)

Abstract: [Objective] This paper tries to quantify the diffusion of technology topics based on patent data, aiming to predict their dissemination in advance. [Methods] First, we constructed the technology diffusion relationship with the patent citation data. Then, we constructed a comprehensive measuring index for technology diffusion from their strength, speed and breadth. Finally, we built the model measuring technology topic diffusion. [Results] We examined our model with 100 topics in the graphene field, which quickly identified topics with high comprehensive diffusion scores. We also found the diffusion directions of graphene patents. [Limitations] We only normalized three measuring indices for technology diffusion with Min-Max Scaling, while their weights not being optimized for the applications. [Conclusions] The proposed model could help us find intelligence effectively with the help of multiple measurements.

Keywords: Technology Diffusion Patent Data Topic Model Measuring Index Graphene