

# 知识网络视角下科学、技术、产业间 创新驱动关系识别方法研究

王超<sup>1</sup>, 许海云<sup>1</sup>, 齐砚翠<sup>2</sup>, 武华维<sup>3</sup>

(1. 山东理工大学管理学院, 淄博 255000; 2. 山东师范大学图书馆, 济南 250014;

3. 西北师范大学档案馆, 兰州 730070)

**摘要** 开展科学、技术、产业之间创新驱动关系研究, 对于实现科学发现、技术创新与产业的深度有效融合与发展具有重要意义。本文从知识网络视角展开科学、技术、产业间创新驱动关系识别方法研究。首先, 通过构建时效知识网络, 利用社区网络结构熵来表征知识网络结构特征; 其次, 利用网络结构特征信息间的传递熵探究创新过程中科学、技术、产业三个系统相互驱动的关系; 再其次, 采用互信息探测主要驱动关系下知识传递路径的变化特点; 最后, 以再生医学(干细胞)领域作为实证对象, 实证研究发现不同时间阶段, 科学、技术、产业之间的相互驱动关系会发生转变, 三者知识传递路径上知识内容前后继承关系也不同, 科学驱动技术过程中的知识传递路径继承性特征持续增强, 而技术驱动科学过程中的知识传递路径继承性呈现不稳定、非连续性特征。此外, 通过多种方式论证了本文方法的科学有效性, 可以用于其他创新领域科学、技术、产业间创新驱动关系的识别。

**关键词** 知识网络; 科学-技术-产业关联; 驱动关系; 结构熵

## Identification Method of Innovation Driving Relationships among Science, Technology, and Industry from the Perspective of the Knowledge Network

Wang Chao<sup>1</sup>, Xu Haiyun<sup>1</sup>, Qi Yancui<sup>2</sup> and Wu Huawei<sup>3</sup>

(1. Business School, Shandong University of Technology, Zibo 255000; 2. Shandong Normal University Library, Jinan 250014;

3. Archives of Northwestern Normal University, Lanzhou 730070)

**Abstract:** Research on the relationship between innovation drivers in science, technology, and industry is crucial for achieving deep and effective integration and development of scientific discovery, technological innovation, and industry. This study investigates innovation-driven relationships between science, technology, and industry through the lens of knowledge networks. First, a temporal knowledge network is constructed to capture the structural characteristics of the knowledge network using the concept of structural entropy in the community network. Subsequently, this study explores the relationship between the three systems of science, technology, and industry in the innovation process by analyzing the transfer entropy between the characteristic information of the network structure. Furthermore, the changing characteristics of the knowledge transfer path under the main driving relationship are examined using mutual information. Finally, empiri-

收稿日期: 2023-01-11; 修回日期: 2023-06-01

基金项目: 国家自然科学基金项目“基于弱信号时效网络演化分析的变革性科技创新主题早期识别方法研究”(72274113); 山东省自然科学基金项目“基于弱信号分析的变革性创新主题早期识别方法研究”(ZR202111130115)。

作者简介: 王超, 男, 1988年生, 博士, 副教授, 硕士生导师, 主要研究领域为科技与产业情报分析; 许海云, 通信作者, 女, 1982年生, 博士, 教授, 博士生导师, 主要研究领域为科学计量, E-mail: xuhaiyunnemo@gmail.com; 齐砚翠, 女, 1996年生, 硕士研究生, 主要研究领域为科学计量; 武华维, 男, 1985年生, 博士, 副研究员, 主要研究领域为知识图谱与知识计算。

cal research focuses on the field of regenerative medicine (stem cells) to investigate how the mutual driving relationships among science, technology, and industry vary at different stages. This study also reveals different inheritance patterns of knowledge content on the knowledge transmission path among the three systems. In science-driven technology processes, the inheritance characteristics of the knowledge transfer path are consistently strengthened. However, in technology-driven science, the inheritance of the knowledge-transfer path is unstable and exhibits discontinuous characteristics. Moreover, this study demonstrates the scientific validity of the proposed method through various means and suggests its applicability in identifying innovation-driven relationships among science, technology, and industry in other innovation fields.

**Keywords:** knowledge network; science-technology-industry linkage; driving relationship; structural entropy

## 0 引言

当前国际环境发生巨大变化,产业发展的技术来源、技术路线均面临重大变革,传统产业转型升级遭遇技术障碍,科学理论和前沿技术短缺成为突出矛盾。科学、技术与产业间的协同发展科学理论的突破、技术的变革创新、新兴及未来产业的发展提供了重要突破口。在此背景下,亟待弄清科学、技术、产业间的创新双向互促关系的特点,即三者间的创新驱动关系如何。开展科学、技术、产业之间创新驱动关系研究对于实现科学发现、技术创新与产业的深度有效融合和发展具有重要意义。现有研究更多强调的是科学、技术、产业三者之间一种线性或非线性的相关关系,鲜有研究涉及科学、技术、产业之间的驱动关系,尚未有效刻画科学、技术、产业间创新驱动关系的特征及规律。本文尝试从知识网络视角展开科学、技术、产业间创新驱动关系识别研究。具体来说,通过构建时效知识网络,以社区网络结构熵来表征知识网络结构特征,利用网络结构间的耦合关系与结构特征信息的传递熵探究创新过程中科学、技术、产业三个系统的相互驱动关系,最后利用互信息来探测主要驱动关系下知识传递路径的变化特点,以期通过探索科学、技术、产业之间相互推动、相互反馈的特点,拓展现有科技与产业创新理论,深入揭示科学、技术、产业创新要素的流动和扩散规律,促进知识在科学-技术-产业间的有序流动。

## 1 相关研究

### 1.1 科学、技术、产业的概念与关系

科学是客观规律及知识的发现结果<sup>[1]</sup>,主要解决客观世界中“是什么”“为什么”的理论问题。技术是改造客观世界的各种手段、方式、方法的总和,主要解决客观世界中“做什么”“怎么做”的

实践问题<sup>[2]</sup>。科学和技术是构成科技的辩证统一体,两者的发展体系与路径不同,但两者有机融合、相互促进,共同决定了科技的质量、效益和走向<sup>[3]</sup>。产业在微观经济和宏观经济的层面具有不同的含义,前者是指一组特定的生成性或盈利性的企业,后者是指能够生成一组密切相关的原料、商品或服务的经济部门<sup>[4]</sup>。国内普遍认可的产业概念是介于微观和宏观经济层面之间的集合概念,是以生成经营服务单位为基础,能够提供紧密相关产品或服务的社会经济行业<sup>[5]</sup>。

关于科学、技术、产业三者间的关系,主要有线性关系和非线性关系两类认识。其中,线性关系观点认为三者之间存在线性转化关系,现有学者常选用创新链模型来描述科学、技术、产业之间线性转化过程<sup>[6-7]</sup>,即科学产生的理论知识会转化为实际应用,实际应用中出现的各种新的手段及方法会推动技术创新,技术创新能够促进产业的迭代与升级。从知识流动的角度来看,这种线性关系反映了知识从科学到产业活动中单向、逐次渐进的知识流动过程,但是在一定程度上对知识从产业到科学技术的反馈机制以及知识活动的系统性关注不足,不利于反映科学技术与产业的真实互动场景。

非线性关系观点认为三者之间存在非线性的互动关系,即科学技术与产业的互动过程由诸多主体、要素参与,相互作用关系并不是简单线性的关系。起初,有学者聚焦于科学与技术之间存在的非线性互补<sup>[8-9]</sup>,即科学与技术的二元关联互动关系。之后,有研究将这种非线性互补的对象延伸到产业层面,即科学与产业<sup>[10]</sup>、技术与产业<sup>[11]</sup>的二元互联互动关系,强调三者之间既有“正向作用”,也有“反向作用”<sup>[12]</sup>。近年来,有学者尝试将科学技术与产业看作一个有机运行的整体<sup>[13]</sup>,即科学-技术-产业间三元互联互动关系。较为典型的是螺旋互动模式,该视角关注科技与产业在时间轴上的交互、融合及其反馈,以螺旋状的创新演化形式表述三者

的关联互动模式<sup>[14]</sup>。这种非线性关系可以反映出科学、技术、产业之间存在着相互融合、驱动的复杂互动关系。在这种非线性互动关系基础上, 本文进一步开展具有方向性、可测度的驱动关系研究。

## 1.2 科学、技术与产业的关系测度研究

科学、技术、产业的内涵及三者之间的关系辨析为定量化揭示三者之间的联系奠定了理论基础, 有学者尝试对三者之间的关系进行定量化测度。例如, 彭帅等<sup>[13]</sup>利用专利引文关系来建立关联测度模型, 但是研究依赖于单一的专利数据源, 无法更全面地反映三者之间的关联关系。针对单一数据源无法全面反映三者之间关系的局限性, Kenta等<sup>[15]</sup>利用科学论文、专利数据和产业经济普查数据, 围绕专利引用关系以及联合申请专利关系等信息, 构建了测度三者之间关系的指标。武华维等<sup>[16]</sup>采用类似于Kenta等<sup>[15]</sup>研究的三种数据源, 以科学、技术、产业三者的知识元素作为研究对象, 利用知识元素的耦合来测度三者的互联关系。部分学者尝试利用科学、技术、产业的分类标准之间的映射关系来测度三者之间的关系, 如Han<sup>[17]</sup>、Park等<sup>[18]</sup>利用科技论文分类标准和专利分类标准, 将科学、技术与产业联系起来, 以此构建知识流动的模型来测度三者间的关系。Wong等<sup>[19]</sup>利用构建科技论文、专利分类标准、产业三种分类标准之间的映射关系, 在此基础上实现对科学、技术、产业间关系的测度。

科学、技术、产业之间存在着复杂的互动关系, 而且随着时间的推移而发生变化。但是, 现有研究通常侧重于某个时期内科学、技术、产业之间的关系, 分析时间静态视角下三个对象彼此之间的关系特点。另外, 现有方法更多强调的是三者之间一种线性或非线性的相关关系, 鲜有研究涉及不同阶段下科学、技术、产业之间具有互动方向性的驱动关系。鉴于此, 本文尝试在知识网络的视角下, 利用网络结构间的耦合关系与结构特征信息的传递熵来探究创新过程中科学、技术、产业三个系统的前后相互驱动关系, 以期拓展现有科技创新与产业创新理论。

## 2 研究方法

知识网络视角下的科学、技术、产业间创新驱动关系的识别方法, 旨在发现三者之间相互驱动关系的模式与特征, 主要思路是利用三者之间知识网络结构的耦合关系与结构特征信息的传递熵来探究

相互驱动关系。本文的识别方法可以分为3个模块7个步骤, 如图1所示。

模块1: 数据获取及处理。从多来源数据库中获取科学文本、技术文本和产品文本数据, 利用自然语言处理技术对获得的文本数据进行结构化处理, 获取具有实体概念意义的主题词(步骤1)。

模块2: 知识网络构建及网络社区划分。根据主题词之间的共现关系构建知识网络, 并利用主题词的时间属性构建时效知识网络(步骤2); 利用网络社区发现算法对时效知识网络中各时间片上的快照网络进行网络社区划分(步骤3)。

模块3: 基于网络结构熵的驱动关系识别。计算时效知识网络的网络社区结构熵, 用于刻画科学、技术、产业知识网络结构的变化趋势(步骤4); 利用贪婪高斯分割算法对三类知识网络结构的变化趋势进行分段, 以定量方式把握知识网络结构变化的各个阶段特点(步骤5); 依据网络结构变化趋势的定量分段, 分别计算各时间段科学、技术、产业知识网络间的传递熵, 以此判断各时间阶段三者之间的创新驱动关系, 并识别时间段内最显著的驱动关系(步骤6); 计算具有显著驱动关系的两类知识网络时间前后各个网络社区间的互信息, 判断知识主题(网络社区)间的继承性; 识别最大互信息, 发现具有显著继承关系的知识主题; 获取具有显著继承关系的知识主题间的知识交集, 得到具有显著继承性的知识内容; 计算相邻时间段继承性知识内容间的相似性, 探析科学、技术、产业间驱动关系下主要知识传递路径(步骤7)。

具体的步骤及方法详述如下。

### 2.1 获取及处理数据

本文方法的数据是可以表征科学、技术、产品知识内容的文本描述信息, 具有广泛性。科技文本主要来源于期刊论文、会议论文、基金项目等形式的科学研究成果, 技术文本主要来源于专利文献、技术标准、技术手册等形式的技术研究成果, 产品文本主要来源于产品说明、产品注册、产品研发等形式的产业活动成果。

同时, 考虑到摘要是对信息主要内容的高度概括, 将摘要数据作为构建知识网络的具体数据对象, 以此为基础, 展开科学(S)、技术(T)、产业(I)间的互动关系识别研究。摘要数据通常是非结构形式存在的文本数据, 需要利用自然语言处理技术进行相应的分词、词性标注、同义词/近义



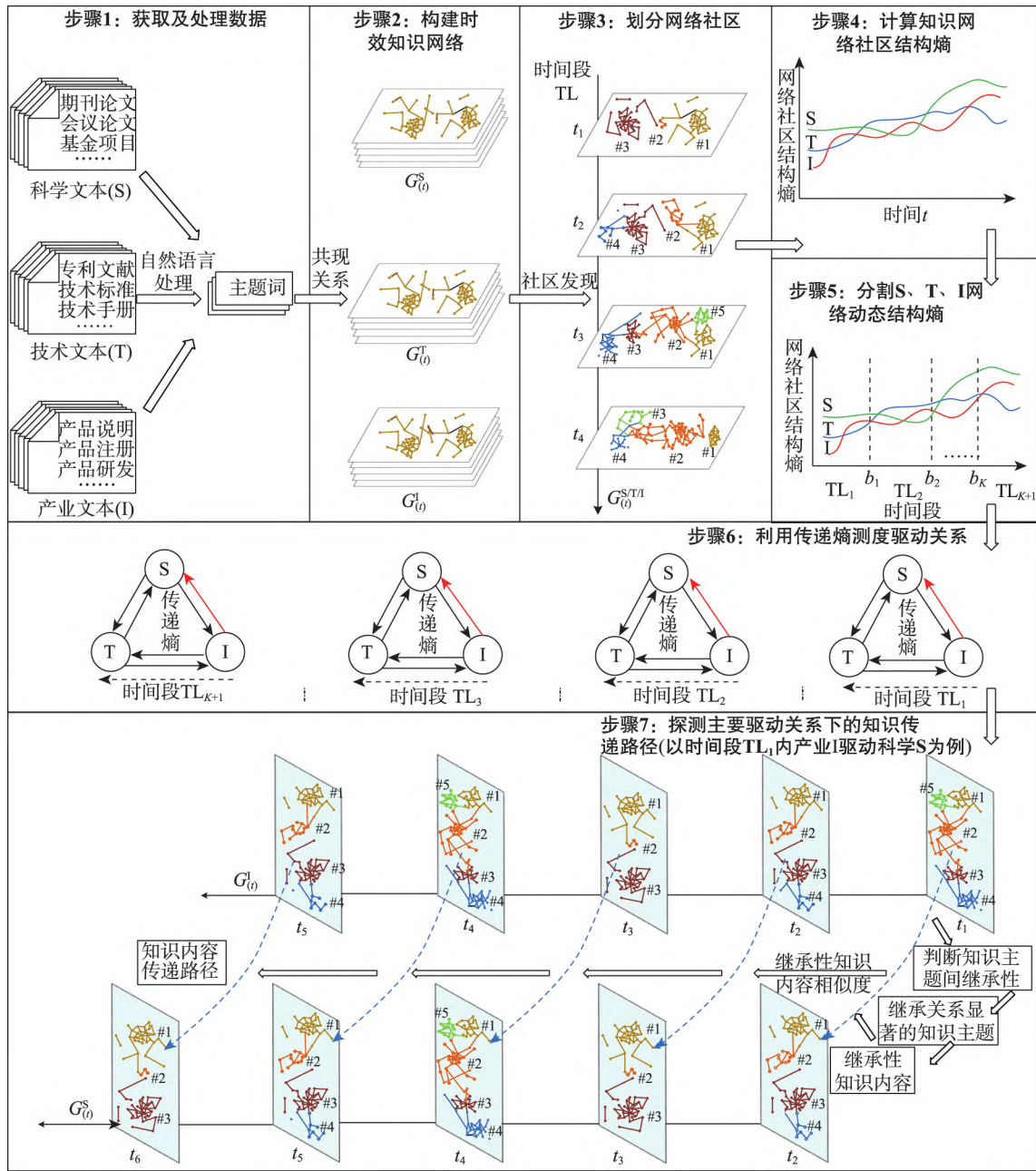


图1 知识网络视角下科学、技术、产业间创新驱动关系识别研究框架

词合并等结构化处理,提取具有实体概念意义的主  
题词。

## 2.2 构建时效知识网络

采用网络形式表示一个系统(对象)的优势在  
于网络结构可以较好地描述系统的动力机制,可以  
评估不同网络之间的相互影响程度<sup>[20]</sup>。因此,本文  
在知识网络的形式下探析科学、技术、产业间创新  
驱动关系。具体而言,利用知识网络结构间的耦合  
关系探究创新过程中科学、技术、产业三个系统间

前后驱动关系,更多体现的是三个系统间具有方向  
性的动态互动关系。

相较于静态知识网络,具有时间属性的时效知  
识网络可以更好地反映知识网络间动态关系的变化  
以及知识扩散活动的动力机制<sup>[21]</sup>,因此,可以将创  
新驱动关系的识别研究立足于时效知识网络中。时  
效知识网络  $G_{(t)}^O$  是由一段时间  $t = \{t_1, t_2, \dots, t_n\}$  内时间  
片上对象  $O$  的静态知识网络(快照网络)  $G_{(t)}^O =$   
 $\{G_{t_1}^O, G_{t_2}^O, \dots, G_{t_n}^O\}$  组成的,其中,  $O = \{S, T, I\}$ ,  $S$ 、 $T$ 、  
 $I$  分别为科学、技术、产业;  $t$  时刻静态知识网络记

为  $G_t^o = (V_t, E_t)$ ,  $V_t$  和  $E_t$  分别为  $t$  时刻网络的顶点及边。

### 2.3 划分知识网络社区

知识网络社区体现了知识网络空间内部知识单元团簇性与集聚性<sup>[22]</sup>。具有团簇性的知识单元之间紧密联系,不同团簇之间的知识单元联系稀疏。基于知识单元的这种团簇性可以将知识网络空间划分为若干具有相同知识主题在社区结构,可以认为知识网络社区是知识网络中观视角下的知识主题表现形式,知识网络社区的发展变化可以揭示领域研究主题的演化情况<sup>[23]</sup>。

本文选用由 Newman<sup>[24]</sup> 提出的基于模块度的 Fast GN (Girvan and Newman) 网络社区发现算法识别网络社区结构,该方法具有运行速度快、准确性高的优点。 $t$  时刻静态知识网络  $G_t^o$  的知识网络社区可以表示为  $G_{m,t}^o = \{C_{1,t}^o, C_{2,t}^o, \dots, C_{m,t}^o\}$ , 其中,  $m$  为网络社区数量。

### 2.4 计算网络社区结构熵

网络结构熵是从拓扑结构的角度发现网络结构相关特征,并能够量化反映知识网络全局特征信息的结构度量<sup>[25]</sup>,对于理解知识网络结构及知识网络演化动力学具有重要作用。在复杂网络中,结构熵与网络结构分布的集中度、网络各部分的差异性、网络结构异构性成反比,与网络结构稳定性、网络结构均匀程度成正比<sup>[26-27]</sup>。

本文采用网络社区结构熵<sup>[28]</sup>量化科学、技术、产业知识网络的结构特征,不仅可以知识网络社区数量及规模信息融合,以表征全局网络特征,同时结合时间序列可以形成网络社区动态结构熵,能够更好地反映科学、技术、产业知识网络结构特征的动态变化及耦合关系。

网络社区结构熵主要利用 Fast GN 算法获得的网络社区及每个社区中的节点数量来计算,其计算公式为

$$E_{\text{com}}(t) = - \sum_{i=1}^{M(t)} \left[ \frac{c_{i,t}}{N_t} \log \frac{c_{i,t}}{N_t} \right] \quad (1)$$

其中,  $M(t)$  为  $t$  时刻网络社区的数量;  $N_t$  为  $t$  时刻网络规模;  $c_{i,t}$  为  $t$  时刻网络社区  $i$  中节点的数量。熵值越小,表示网络社区结构分布越集中,各社区结构的差异性越大,网络社区结构越不稳定;熵值越大,表示网络社区结构分布越分散,各社区结构的差异性越小,网络结构社区趋于均衡。

### 2.5 分割 S、T、I 知识网络动态结构熵

考虑到 S、T、I 三个知识网络之间互动关系动态变化的差异性,需要以三类知识网络社区的动态结构熵作为观测值构建多元时间序列,并将其分割为若干具有统计特征的子序列,确保这些子时间序列的统计特征具有段内同质性、段间异质的特点<sup>[29]</sup>,以更好地探究 S、T、I 知识网络结构在不同时间阶段的互动关系特点。

常见的分割方法主要面向一元(即单一观测值)时间序列<sup>[30]</sup>,这些方法无法适用多变量的多元时间序列,如果简单地将每个变量按照一元时间序列分段,那么无疑会割裂各个变量之间的关系。为此,本文采用贪婪高斯分割算法(greedy Gaussian segmentation, GGS)<sup>[31]</sup>来实现对 S、T、I 知识网络动态结构熵进行量化的多元时间序列分割,该方法使用自上而下的贪婪算法来检测时间序列的分割点,具有较强的扩展性和适用性。其主要思路是将多元时间序列  $[E_1, E_2, \dots, E_t] \in R^{3 \times \text{TL}}$  (其中, TL 表示时间序列的长度,时间  $t = 1, \dots, \text{TL}$ , 3 表示科学、技术、产业三个网络结构熵的时间序列)划分为若干具有不同均值和协方差且符合多变量高斯分布的子时间序列,然后使用自上而下的迭代算法进行分割点搜索或边界组合优化,获取  $K-1$  个分割点  $b_1, b_2, \dots, b_{K-1}$ ,最终将时间序列划分为  $K$  个子时间序列。

### 2.6 利用传递熵测度驱动关系

传递熵(transfer entropy)是一种基于香农熵来检测两个系统之间定向信息传递的非参数信息度量方法,通过对系统的动力学及其耦合性质的最小假设,来量化两个系统之间信息传输动态性及方向性<sup>[32]</sup>,尤其是对于高度非线性或非高斯分布的数据,传递熵更适合量化系统间的因果关系<sup>[33]</sup>。在本文中,科学、技术、产业间的创新驱动关系是建立在科学、技术、产业知识网络结构间的耦合关系之上,这种驱动关系具有明显的方向性。因此,可以利用基于传递熵构建创新驱动关系的测度模型,以此获得具体领域的科学、技术、产业之间的驱动关系。

设  $x$ 、 $y$  为某一时间段内科学、技术、产业知识网络中的任意两个动态结构熵变量,那么该时间段中网络结构  $y$  到网络结构  $x$  的信息传递熵定义为

$$E_{y \rightarrow x} = \sum_{x,y} p(x_{n+1}, x_n, y_n) \log_2 \frac{p(x_{n+1}|x_n, y_n)}{p(x_{n+1}|x_n)} \quad (2)$$



传递熵  $E_{y \rightarrow x}$  反映了网络结构  $y$  对  $x$  的不确定性大小的影响，即网络结构  $y$  传递给  $x$  的信息量，也可以理解为网络结构  $y$  的变化对于  $x$  变化的影响程度。因此，依据这个信息量可以判断某个时间段内科学、技术、产业知识网络中任意两个对象之间的驱动关系，在此基础上，可以构建该时间段内科学、技术、产业间驱动关系强度指数  $DI_K$ ，即

$$DI_K = \begin{cases} E_{y \rightarrow x} - E_{x \rightarrow y} > 0 \\ E_{x \rightarrow y} - E_{y \rightarrow x} < 0 \\ 0, & E_{y \rightarrow x} - E_{x \rightarrow y} = 0 \end{cases} \quad (3)$$

若  $E_{y \rightarrow x} - E_{x \rightarrow y} > 0$ ，则网络结构  $y$  与  $x$  存在驱动关系， $y$  的变化会影响  $x$  的变化，即  $y$  为因， $x$  为果；若  $E_{y \rightarrow x} - E_{x \rightarrow y} < 0$ ，则网络结构  $x$  与  $y$  存在驱动关系， $x$  的变化会影响  $y$  的变化，即  $x$  为因， $y$  为果；若  $E_{y \rightarrow x} - E_{x \rightarrow y} = 0$ ，则网络结构  $y$  与  $x$  互不影响，彼此之间不存在驱动关系<sup>[34]</sup>。利用驱动关系强度指数  $DI_K$  既可以比较科学、技术、产业之间是否存在驱动关系，又可以判断哪两个对象在这段时间存在最为显著的驱动关系（主要驱动关系），如果两个对象的驱动关系最显著，那么意味着这两个对象在某个时间段内的知识创新活动前后推动的耦合关系明显。

## 2.7 探测主要驱动关系下的知识传递路径

利用科学、技术、产业之间的主要驱动关系，可以进一步从微观层面探析科学、技术、产业间知识传递的主要路径。考虑到互信息不仅可以度量两个随机变量相互共享的信息量<sup>[35]</sup>，也可以用于反映知识传递过程中的继承关系<sup>[36]</sup>。本文利用知识内容互信息，获得某时间段内驱动关系最显著的两个对象之间知识耦合所形成的知识创新内容的主要继承性传递路径，本文将此称为主要知识内容驱动路径，以此探析科学、技术、产业间驱动发展过程中的知识内容变化特点。

设具有  $m$  个社区的知识网络  $x_m$  到  $y_m$  的具有最大的驱动关系强度指数  $DI_m = E_{x_m \rightarrow y_m}$ ，主要分析步骤如图 2 所示。

Step1. 计算社区间互信息，判断知识主题间继承性。

计算  $t$  时间点知识网络  $x_m$  的各个网络社区  $C_{i,t}^{x_m}$  与  $t+1$  时间内知识网络  $y_m$  的各个网络社区  $C_{i,t+1}^{y_m}$  之间的互信息，计算公式为

$$I(C_{i,t}^{x_m}, C_{i,t+1}^{y_m}) = \sum_{C_{i,t}^{x_m}, C_{i,t+1}^{y_m}} p(N_{x,y}) \log \frac{p(N_{x,y})}{p(N_x)p(N_y)} \quad (4)$$

其中， $p(N_{x,y})$  为网络社区  $C_{i,t}^{x_m}$  和  $C_{i,t+1}^{y_m}$  中各个主题词（节点）的联合概率分布， $p(N_x)$ 、 $p(N_y)$  是边缘概率分布。

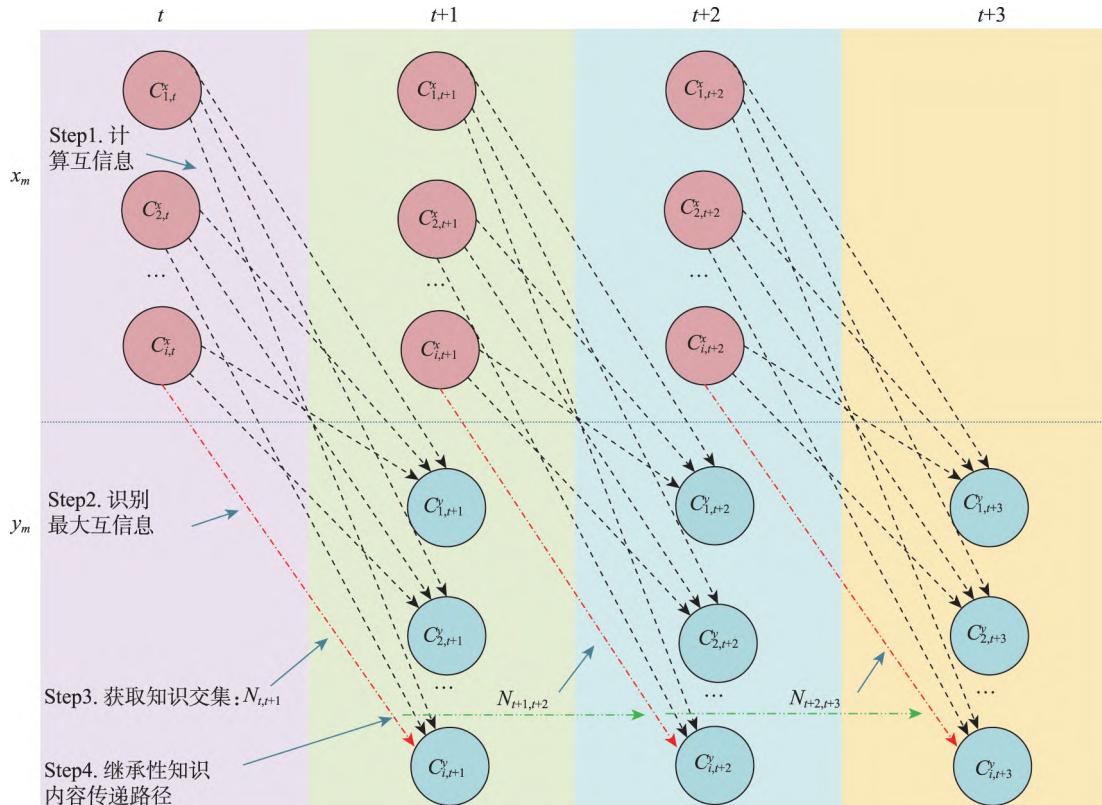


图 2 基于互信息探测驱动路径示意图

具有显著驱动关系的两个知识网络  $x_m$ 、 $y_m$  各社区间的互信息反映了这两个领域知识主题之间的信息传递量。互信息越大, 表明  $C_i^{x_m}$  传递给  $C_{i+1}^{y_m}$  的信息量越多,  $C_{i+1}^{y_m}$  所代表的知识主题对  $C_i^{x_m}$  所代表的知识主题继承性越强; 互信息越小, 表明  $C_i^{x_m}$  传递给  $C_{i+1}^{y_m}$  的信息量越少,  $C_{i+1}^{y_m}$  所代表的知识主题对  $C_i^{x_m}$  所代表的知识主题继承性越弱<sup>[36]</sup>。

Step2. 识别最大互信息, 发现具有显著继承关系的知识主题。

具有驱动关系的两个知识网络  $x_m$ 、 $y_m$  各社区间的互信息值大小不一, 这意味着主题间的继承性具有差异性。最大互信息表示主题间具有最显著的继承性, 利用这种显著的继承关系, 有助于发现某时间段内主要驱动关系中具有显著继承性关系的知识创新内容的主要传递路径。通过互信息值的对比分析, 可以获得时间段内  $t$  时间具有最大互信息的两个网络社区  $C_i^{x_m}$  与  $C_{i+1}^{y_m}$  的集合  $\{(C_{i,1}^{x_m}, C_{i,2}^{y_m}), (C_{i,2}^{x_m}, C_{i,3}^{y_m}), \dots, (C_{i,t-1}^{x_m}, C_{i,t}^{y_m})\}$ , 利用该集合可以形成  $t$  时间内知识网络  $x_m$ 、 $y_m$  之间主要继承性知识传递路径。

Step3. 获取具有显著继承关系的知识主题间的知识交集, 得到具有显著继承性的知识内容。

网络社区间的知识交集是信息共享的主要内容, 具有最大互信息网络社区间的知识交集是具有显著继承关系的知识内容, 在此基础上, 可以获得主要继承性知识内容的传递路径。具体来说, 分别获取  $t$  时间具有最大互信息的网络社区  $C_i^{x_m}$  与  $C_{i+1}^{y_m}$  之间的主题词交集  $N_{i,t+1} = N_{C_i^{x_m}} \cap N_{C_{i+1}^{y_m}}$ , 这些主题词表示了继承性知识内容。最终获得主题词交集集合  $\{N_{1,2}, N_{2,3}, \dots, N_{t-1,t}\}$ , 由该集合形成了主要继承性知识内容传递路径。

Step4. 计算继承性知识内容间的相似性, 分析知识内容的传递特点。

计算主要继承性知识内容传递路径中相邻继承性知识内容的相似性, 可以分析继承性知识内容的变化特点。如果继承性知识内容的相似度逐渐变大, 那么说明知识传递路径上知识内容前后继承关系逐渐增强, 体现了两个领域驱动过程中知识呈现会聚性<sup>[36]</sup>; 如果继承性知识内容相似度逐渐变小, 那么说明知识传递路径上知识内容前后继承关系逐渐减弱, 体现了两个领域驱动过程中知识内容发生转折, 新的知识内容可能出现。具体地,

本文利用余弦相似度算法计算继承性知识内容的相似性。

### 3 实证研究

再生医学是面向受损 (患病) 细胞、器官、组织的再生、修复或替换的一门新兴医学, 其中干细胞是再生医学的核心内容, 再生医学 (干细胞) 逐渐成为现代临床医学中具有广阔应用前景的研究领域, 能够促进医学领域向崭新的高度发展<sup>[37]</sup>。鉴于此, 本文选用再生医学 (干细胞) 领域作为科学、技术、产业间创新驱动关系识别方法的实证对象。

#### 3.1 数据获取及处理

针对再生医学 (干细胞) 领域, 本文选用 PubMed、Derwent Innovation 数据库和 Cortellis 数据库分别作为获取论文、专利、产品信息的数据来源。检索时间为 1997 年 1 月 1 日至 2020 年 12 月 31 日, 分别获取 392683、44454、1347 条数据。如图 3 所示, 再生医学 (干细胞) 论文和专利的数量一直处于增长状态, 尤其是专利数量自 2014 年以来发展迅速, 而产品的数量变化波动较大。

使用 SemRep 文本处理工具和 Python 语言中的 SpaCy 工具包对非结构化的摘要数据进行语料清洗、分词、词性标注、去停用词等结构化处理, 提取具有实体概念意义的主题词, 从科技论文、专利数据和产品数据中提取主题词的数量分别为 84473、55664、2896 个。

#### 3.2 网络社区分析

利用上述已获得的数据, 分别构建了从 1997 年至 2020 年共 24 年的再生医学 (干细胞) 科学、技术、产业知识网络。图 4 展示了再生医学 (干细胞) 产业知识网络从 1997 年至 2020 年的变化。进一步分析 S、T、I 知识网络规模的变化 (图 5) 可知, S、T 知识网络规模呈增长趋势, 说明该领域

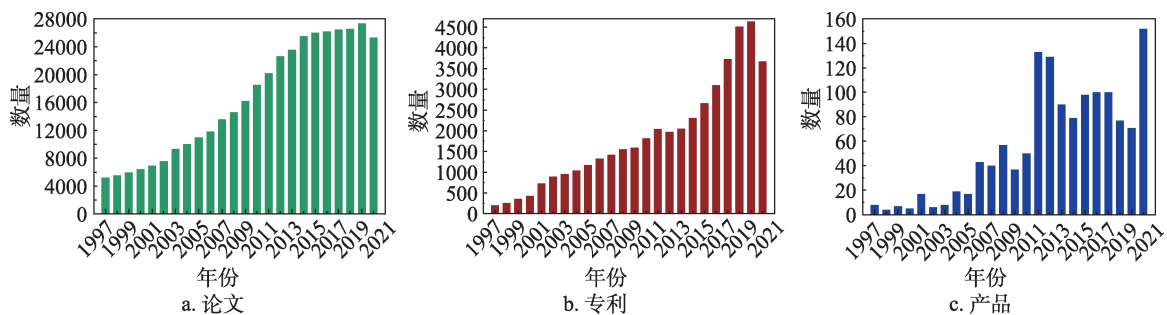


图3 再生医学(干细胞)论文、专利、产品数量变化



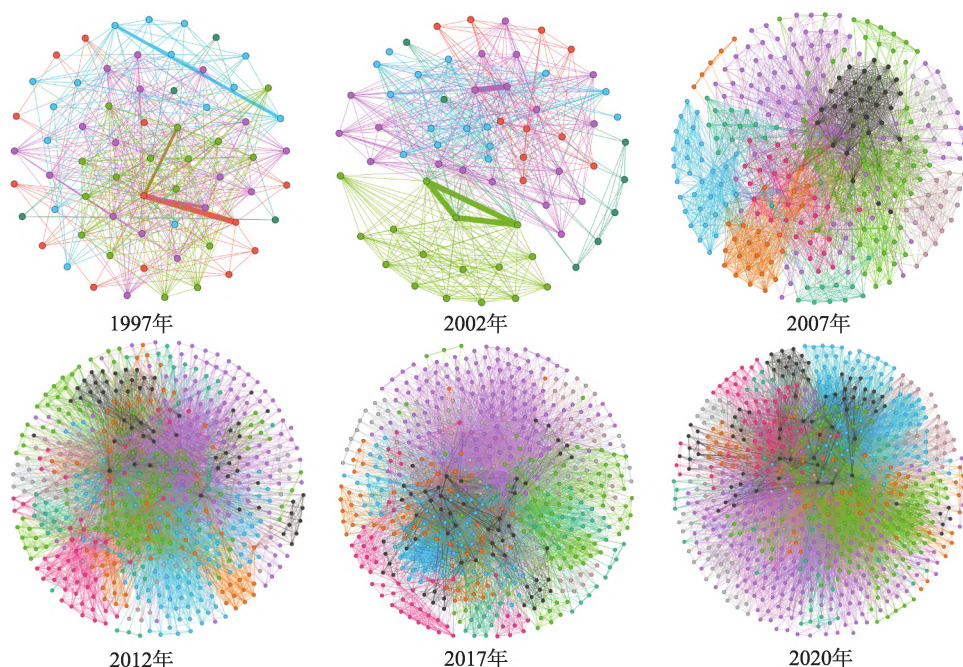


图4 再生医学(干细胞)产业知识网络演化(1997—2020年)

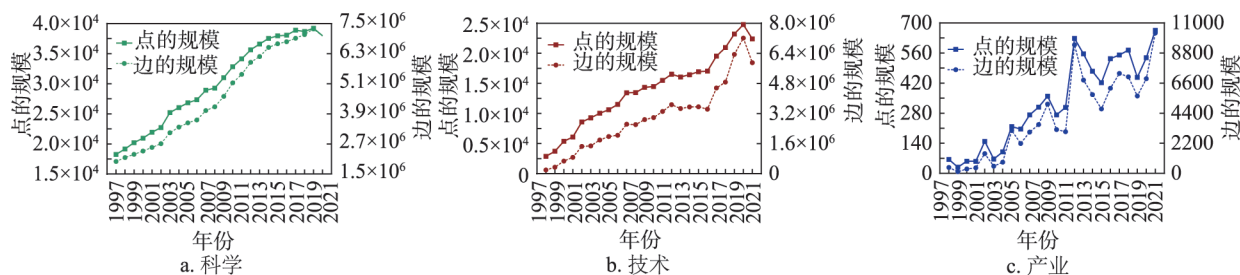


图5 再生医学(干细胞)科学、技术、产业知识网络规模变化

中科学研究和技术创新内容持续丰富；I知识网络规模发展呈现一定的波动，且知识网络规模相对最小，说明该领域产业创新和开发的内容丰富度相对集中、有限。

通过 Fast GN 网络社区发现算法的计算发现(图6)，科学知识网络的社区数量最多，表明再生医学(干细胞)领域的科学研究主题最为丰富；产业知识网络的社区数量维持在相对较低的水平，说明该领域产业研究内容相对较为集中。从总体来看，科学研究知识网络中的社区数量变化以M形曲线波动，且呈降低趋势，说明科学研究知识网络的研究主题在扩张与收敛的状态中交替演化，研究主题内容整体上逐渐收敛，表明该领域科学研究的方向逐渐趋向集中。技术知识网络中的社区数量也以M形曲线波动，但是波动频率和幅度相对科学研究知识网络较低，说明该领域技术创新内容的扩张和收敛维持在相对稳定的范围。产业知识网络的社区

数量变化波动相对较少，整体趋势呈现水平方向波动，这再次表明该领域的产业创新和开发内容相对集中、有限。

### 3.3 动态结构熵分析

图7是S、T、I知识网络社区动态结构熵分布图，产业的知识网络社区结构熵总体上高于科学、技术知识的结构熵，技术知识网络次之，科学知识网络结构熵最小，这说明再生医学领域产业知识网络中的网络社区分散、各社区结构的差异性较小、知识社区网络结构较为均衡。科学知识网络的网络社区较为集中，各社区的差异性大，网络社区结构不均衡，这进一步说明，再生医学的产业研究分散，尚未形成特色鲜明的研究方向，而科学和技术的研究方向相对较为集中，各研究方向的内容差异性大。

从图7可以看出，1997—2020年的知识网络结构



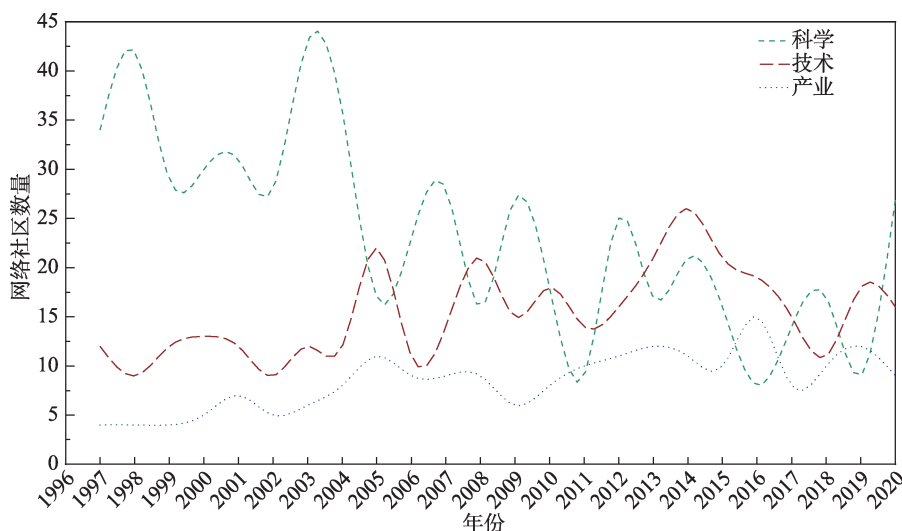


图6 再生医学(干细胞)科学、技术、产业知识网络社区变化

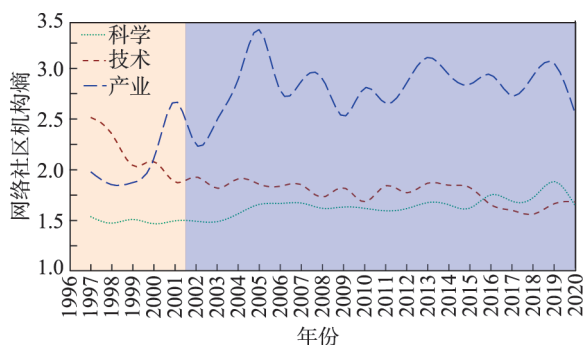


图7 科学、技术、产业知识网络社区动态结构熵及分段

变化可以分为两个阶段，第一阶段是1997—2001年，第二阶段是2002—2020年。在第一阶段（1997—2001年），产业知识网络的社区结构熵呈增长趋势，技术知识网络的社区结构熵呈降低趋势，科学技术知识网络结构熵呈水平发展趋势，说明该阶段产业研究方向分散、技术研究逐渐集中、科学研究较为稳定。在第二阶段（2002—2020年），产业知识网络的社区结构熵在2002—2005年呈现快速增长趋势，2006年之后的结构熵值在2.5~3.5的范围内上下波动，波动幅度较大，产业研究方向分散，研究特色不明显；技术知识网络的社区结构熵整体上呈现下降趋势，波动幅度相对较小，科学知识网络的社区结构熵整体上呈现上升趋势、波动幅度最为平缓，科学和技术的研究方向集中且研究特色明显。

### 3.4 驱动关系测度

(1) 第一阶段（1997—2001年）的驱动关系测度。①科学与技术之间存在驱动关系。科学知识网络结构对技术知识网络结构影响明显，传递熵值为

0.8113（表1）；而技术知识网络结构变化对科学知识网络结构变化无影响，传递熵值为0。因此，科学对技术驱动作用明显。②科学与产业之间不存在驱动关系。两者之间的传递熵相同，科学知识网络结构变化对产业知识网络结构变化产生的影响与产业知识网络结构变化对科学知识网络结构变化产生的影响程度相同，双方之间的驱动关系不明显。③技术与产业之间存在驱动关系。尤其是产业知识网络结构变化对技术知识网络结构变化的影响相对明显，传递熵值为0.1226；但是技术知识网络结构变化对产业的知识网络结构变化无影响，传递熵值为0。因此，产业对技术略有驱动作用。

表1 各时间段内科学(S)、技术(T)、产业(I)间的传递熵值

传递熵	1997—2001年		2002—2020年	
	熵值	DI值	熵值	DI值
$E_{S \rightarrow T}$	0.8113		0.0111	
$E_{T \rightarrow S}$	0.0000	0.8113	0.5943	0.5943
$E_{S \rightarrow I}$	0.1887		0.0111	
$E_{I \rightarrow S}$	0.1887	0.0000	0.3055	0.3055
$E_{T \rightarrow I}$	0.0000		0.0053	
$E_{I \rightarrow T}$	0.1226	0.1226	0.0053	0.0000

综上，该时间段内科学知识网络对技术知识网络结构的驱动关系强度指数DI最大，科学对于技术的驱动作用最为明显。

(2) 第二阶段（2002—2020年）的驱动关系测度。①科学与技术之间存在驱动关系。技术知识网络结构变化对科学知识网络结构变化影响明显，传递熵为0.5943；而科学知识网络结构变化对技术知识网络结构的变化影响不显著，传递熵仅为

0.0111。因此，技术对科学的驱动作用明显。②科学与产业之间存在驱动关系。产业知识网络结构变化对科学知识网络结构变化影响较为明显，传递熵值为0.3055；而科学知识网络结构变化对产业的知识网络结构变化影响不显著，传递熵仅为0.0111。因此，产业对科学具有一定驱动作用。③技术与产业之间不存在驱动关系。两者之间的传递熵相同，技术知识网络结构变化对产业知识网络结构变化产生的影响与产业知识网络结构变化对技术知识网络结构变化产生的影响程度相同，双方之间驱动关系不明显。

综上，该时间段内技术知识网络对科学知识网络结构的驱动关系强度指数DI最大，技术对于科学的驱动作用最为明显。

综合来看，在再生医学创新过程的第一阶段，技术创新是主要的被驱动对象，科学对技术具有关键的驱动作用，产业对技术略有驱动作用；在第二阶段，科学研究是主要的被驱动对象，其中技术对科学具有重要驱动作用，产业对科学具有一定驱动作用。

### 3.5 知识传递路径分析

(1) 第一阶段（1997—2001年）的知识传递路径分析。该阶段以科学驱动技术为主，通过计算该时间段内 $t$ 年科学S的各网络社区与 $t+1$ 年技术T的各网络社区之间的互信息及交集主题词数量（图8）发现，S与T的最大互信息值呈现逐渐增大的趋势，说明技术继承自科学的知识内容的信息量逐渐增多；S与T之间具有最大互信息社区间的知识交集主题词数量呈现逐渐增长的趋势，说明科学向技术的主要知识传递过程中继承性知识内容不断丰富。

进一步分析科学驱动技术过程中继承性知识内容的传递路径（图9）发现，时间相邻的继承性知

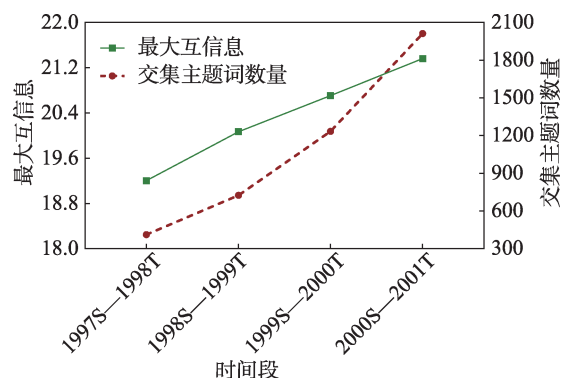


图8 第一阶段 $t$ 年科学(S)与 $t+1$ 年技术(T)在各社区间的最大互信息与交集主题词数量

识内容间的相似度由0.1655逐渐变大为0.6957，说明科学驱动技术过程中知识传递路径上知识内容前后继承关系逐渐增强，知识会聚性趋势明显，也说明科学驱动技术的创新需要持续发挥。

(2) 第二阶段（2002—2020年）的知识传递路径分析。该阶段以技术驱动科学为主，通过计算该时间段内 $t$ 年技术(T)的各社区与 $t+1$ 年科学(S)各社区之间的互信息及交集主题词数量（图10）发现，T与S的最大互信息值在21~23的范围内小幅波动，说明科学继承自技术的知识内容的信息量维持在相对较为稳定的水平。

T与S之间最大互信息网络社区的知识交集主题词数量呈现波动变化的趋势，说明在技术向科学的主要知识传递过程中继承性知识内容范围是不断变化的，尤其是2004T—2007T时间段继承性知识内容范围缩减明显，2008T—2010T时间段呈现扩大趋势，2011T—2013T时间段又呈现缩减趋势，2015T—2018T时间段呈扩大趋势。从继承性知识内容的范围变化特点可以看出，技术驱动科学发展的过程中，技术创新对于科学研究的反馈存在更高的不确定性；继承性知识内容范围越扩大，反映出技术能够反馈给科学研究的技术支持和需求导向更加

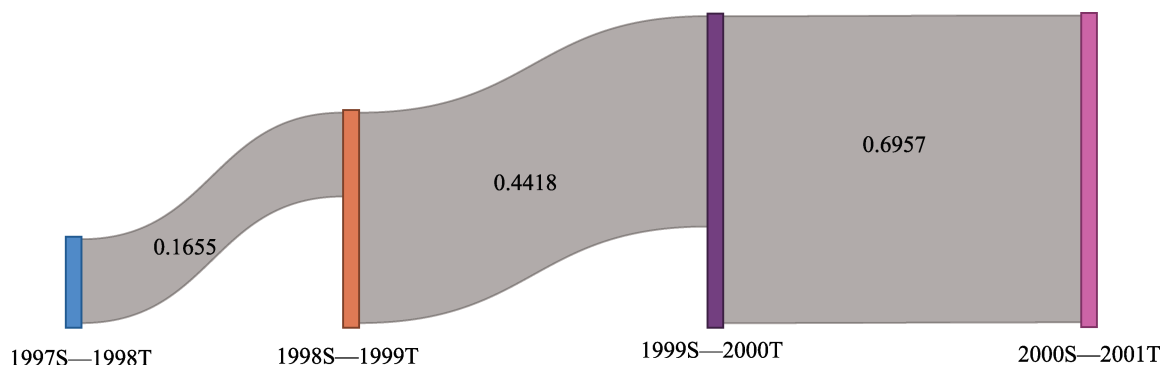


图9 第一阶段科学驱动技术过程中继承性知识内容的传递路径

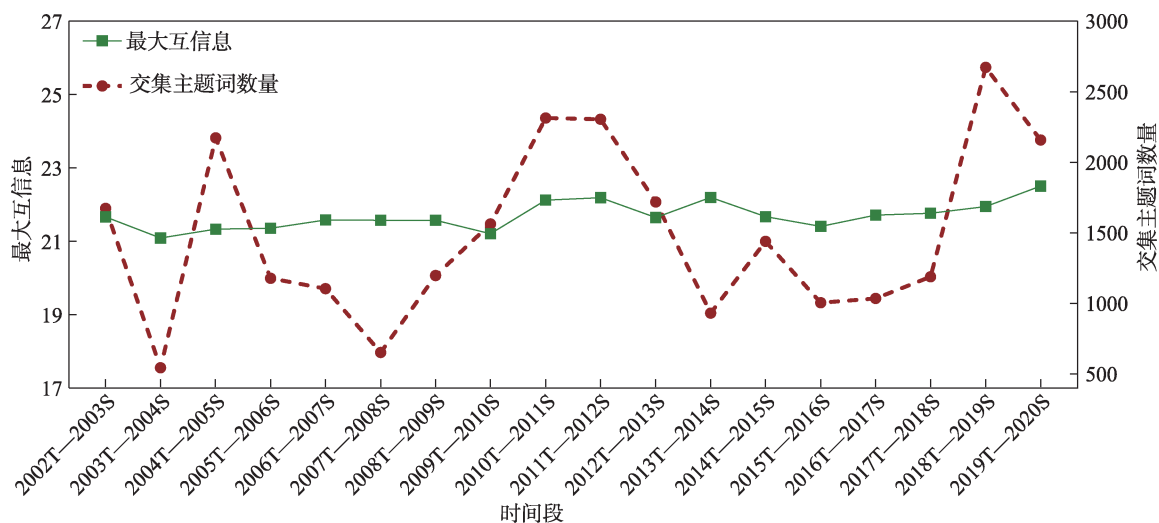


图 10 第二阶段  $t$  年技术(T)与  $t+1$  年科学(S)在各社区间的最大互信息及交集主题词数量

广泛，而继承性知识内容范围的缩减趋势，反映出随着技术的逐渐成熟，技术提供给科学研究的技术支持和导向逐渐汇聚。

进一步分析技术驱动科学过程中继承性知识内容的传递路径（图 11）发现，相邻时间的继承性知识内容间的相似度也呈现波动变化，但是具有波动幅度大、波峰高的特点，说明技术驱动科学过程中知识传递路径上知识内容前后继承关系的差异性变化明显。其中，2003T—2006T 时间段继承性知识内容相似度陡增，达到 0.7523，说明该时间段内知识内容前后继承关系逐渐增强、知识会聚性趋势明显；2006T—2008T 时间段继承性知识内容相似度陡降，降至 0.0024，说明该段时间知识内容前后继承关系逐渐减弱，意味着技术驱动科学过程中知识内容开始发生转折、新的知识内容可能出现；2008T—2011T 时间段继承性知识内容相似度陡增，达到 0.7379，说明该时间段内知识内容前后继承关系逐渐增强、知识会聚性趋势明显；除 2015T—2016T 时间段外，2011T—2019T 时间段继承性知识内容间

相似度略有波动，但总体上维持在降低水平，说明在这段时间内，除了 2015T—2016T 时间段继承关系明显外，知识传递过程中的前后时间的知识内容转折明显。

总体上看，技术驱动科学过程中知识传递路径上知识内容前后继承关系的差异性变化明显，表明技术向科学传递的知识内容是非连续的、动态的，这也说明此过程中技术对科学问题的技术支持和导向是动态变化的，预示着潜在颠覆性技术产生的可能。

### 3.6 实证结果的验证

本文采用文献调研与数据统计分析的方式对实证结果进行验证。

(1) 2001 年是再生医学（干细胞）领域发展的重要时间节点，考虑到该领域面临着严峻的人类伦理问题，时任美国总统 George W. Bush 于 2001 年出台了限制干细胞研究的联邦法案<sup>[38]</sup>，这对该领域的科学研究带来巨大影响<sup>[39]</sup>。这一重要时间节点与本

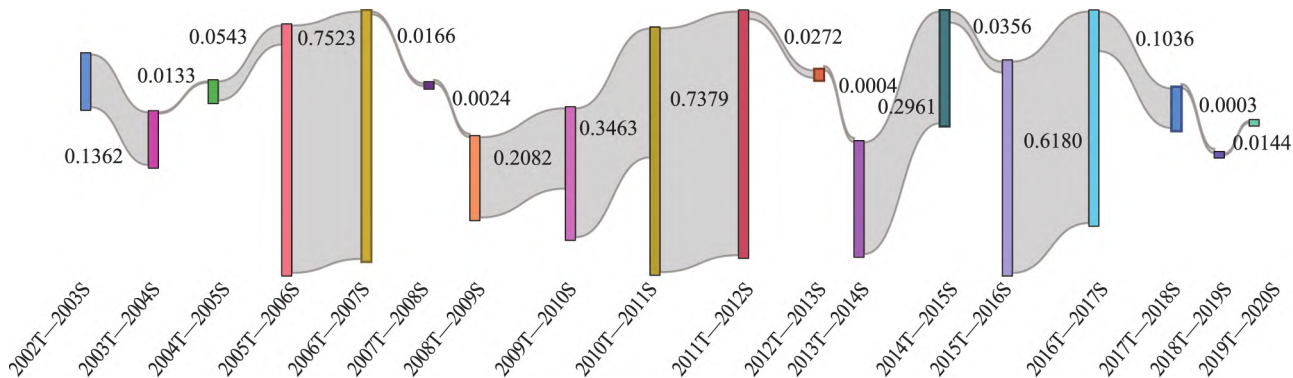


图 11 第二阶段技术驱动科学过程中继承性知识内容的传递路径



文实证研究中知识网络社区动态结构熵分割的时间节点相吻合。

(2) 通过对现有文献资料的梳理发现, 2002 年以前再生医学(干细胞)领域取得了许多重要科学突破, 尤其是 1997—2001 年干细胞领域以科学研究为主。例如, 1997 年加拿大的 Dominique Bonnet 和 John Dick 研究发现癌症是从偏离轨道的干细胞中生长出来的<sup>[40]</sup>; 1998 年生物学家 James Alexander Thomson 以“胚胎干细胞”为题发表了发现人类胚胎干细胞的研究, 2001 年第一个关于成人脐带干细胞移植的研究被发表<sup>[41]</sup>。进一步利用 Lens 数据开放平台<sup>[42]</sup>检索再生医学(干细胞)的论文及专利数据分析发现(图 12), 2002 年以前被专利引用的论文数量及施引专利数量增长率总体呈上升趋势, 这种引用关系也在一定程度上说明 2002 年以前技术受科学研究影响越来越大; 2002 年以后, 被专利引用的论文数量及施引专利数量增长率总体呈下降趋势, 尤其是 2006 年施引专利数量增长率首次进入负增长, 从 2011 年开始施引专利数量增长率持续以负增长下降, 说明该领域科学研究对于技术创新的影响逐渐减弱。这一特征与两个阶段的驱动关系特点基本吻合。

(3) Bergman 等<sup>[43]</sup>研究发现, 2001 年前后是干细胞技术发展的重要时间节点, 尤其是自 2000 年开始, 干细胞技术的专利数量申请急剧增长<sup>[44]</sup>。现有的干细胞专利申请内容及技术组合促使科学研究朝着特定的方向发展, 特别是使科学研究更接近于临床<sup>[45]</sup>, 这与实证分析中驱动关系测度相关研究发现基本吻合。

(4) Hu 等<sup>[46]</sup>从科学与技术相结合的视角探索再生医学(干细胞)知识前沿的过程中发现, 在

2006 年、2011 年这两个时间节点前后的知识内容发生了差异性变化, 这与本文 2002—2020 年时间段知识传递路径的实证分析结果基本吻合。

### 3.7 与现有方法的对比

为了对本文方法的相对优势进行验证, 采用基于网络属性 Copula 熵的知识网络间依赖关系非参数估计法<sup>[47]</sup>以及由 Ghawi 等<sup>[48]</sup>提出的知识网络间相似度测度法作为对照方法, 分别测度科学、技术、产业之间的互动关系。第一种方法中, 两个网络的 Copula 熵值与 0 值的距离越远, 两者的相互依赖的关系越大。从图 13 可以看出, 在点的规模、边的规模和社区数量三类网络属性下 S 与 T 之间存在明显的相关性, S 或 T 与 I 的相关性相对较低。利用第二种方法可以计算出  $t$  时间与  $t+1$  相邻时间下 S、T、I 知识网络间相似度(图 14), S 与 T 之间的相似度较高, 而 S 或 T 与 I 的相似度较低。与本文方法对比发现, 这些方法多从科学、技术、产业之间的关联强度来测度三者之间的关系, 强调更多的是三者之间一种线性或非线性的整体相关关系, 较少考虑三者的时序关联特征, 忽视了三者之间互动关系的方向性的转变特征。本文提出的识别方法通过分阶段测度三者关系的强度及方向性, 实现了对三者动态视角下关联关系的特征分析。

## 4 总结

本文提出一套知识网络视角下科学、技术、产业间创新驱动关系识别方法。该方法从知识网络视角下, 利用知识网络结构间的耦合关系探究创新过程中科学、技术、产业三个系统间前后相互推动的

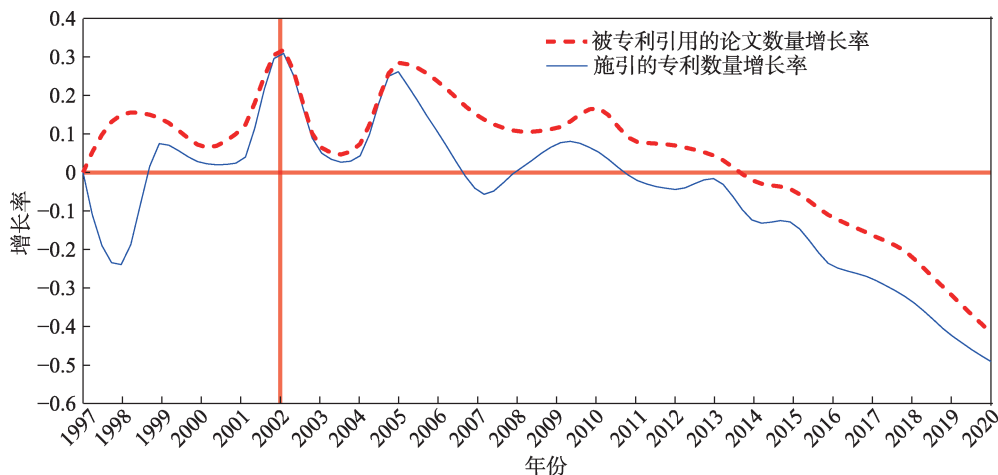


图 12 1997—2020 年再生医学(干细胞)领域论文被引及施引专利增长率变化

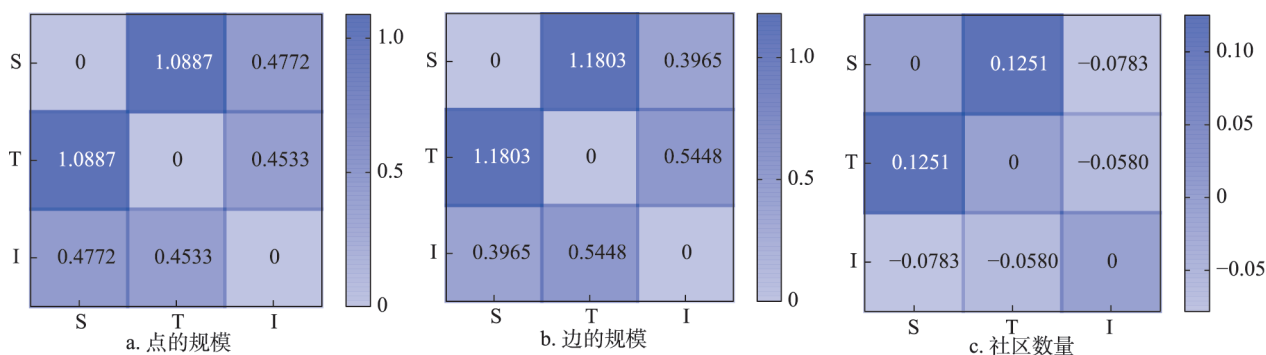


图 13 基于网络属性 Copula 熵的科学(S)、技术(T)、产业(I)知识网络关系测度

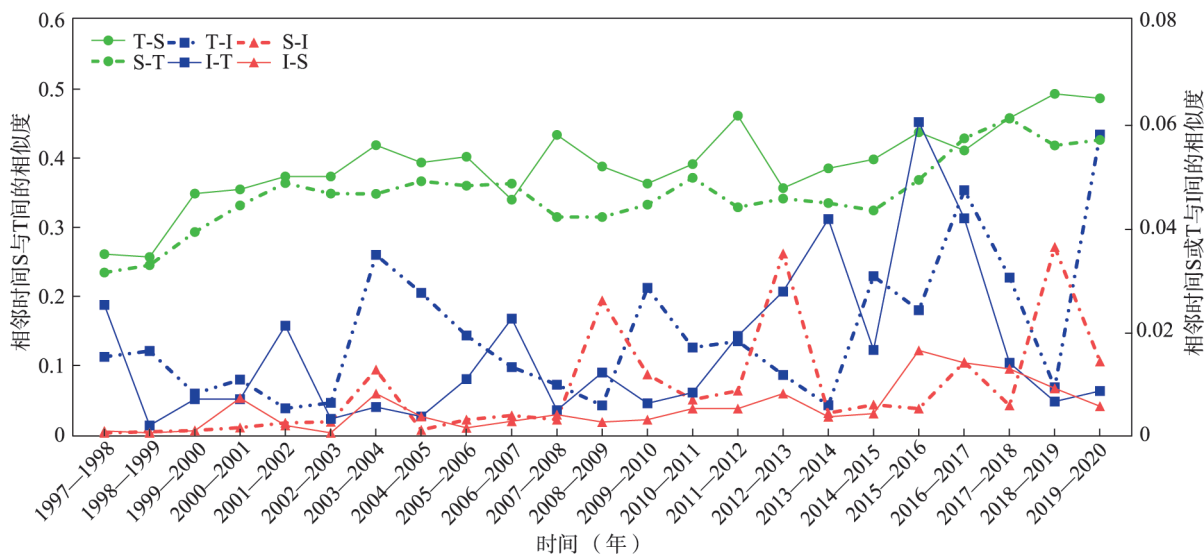


图 14 相邻时间下科学(S)、技术(T)、产业(I)知识网络间相似度

驱动关系；利用网络社区结构熵量化科学、技术、产业知识网络的结构特征，结合时间序列形成的网络社区动态结构熵，以反映知识网络结构特征的动态变化；采用贪婪高斯分割算法，实现对科学、技术、产业知识网络动态结构熵的定量化分割，从整体上把握科学、技术、产业三个知识网络之间互动关系动态变化的不同阶段特点；利用基于传递熵构建创新驱动关系的测度模型，实现了科学、技术、产业之间的驱动关系的测度，可以发现时间段内最为显著的驱动关系，最后利用这种最显著的驱动关系，可以进一步从微观层面探析科学、技术、产业间驱动关系下知识传递的主要路径。

本文选取再生医学（干细胞）领域作为实证对象，验证了本文方法的可行性，对实证结果的多维度验证论证了方法的有效性。相较于已有研究，本文将对科学、技术、产业之间关系测度由线性或非线性的相关关系测度拓展到三者之间的具有方向性的驱动关系测度，这种驱动关系的探究推进了对科

学、技术、产业间创新驱动的动力机制的认知。

此外，本文尚存在一定的局限性。首先，本文采用社区网络结构熵来表征知识网络结构特征，反映了介观层面的知识网络结构特点。未来，更全面的网络结构特征表征需要兼顾微观层面和宏观层面的网络结构熵；其次，知识网络视角下科学、技术、产业间的驱动关系反映的是三者知识内容方面的继承性耦合关系，但是三者之间驱动关系的形成，除了科学、技术、产业本身，还受到多种因素的影响，如科学、技术、产业创新主体之间的空间分布及合作特点。未来研究需要拓展到更多维度的创新生态视角。

## 参 考 文 献

- [1] Feynman R. What is science?[J]. Resonance, 2011, 16(9): 860-873.
- [2] 钱时惕. 什么是技术及科学与技术的关系——科学与人文漫话之二[J]. 物理通报, 2009(11): 55-56, 64.
- [3] Xu H Y, Winnink J, Yue Z H, et al. Topic-linked innovation paths

- in science and technology[J]. *Journal of Informetrics*, 2020, 14(2): 101014.
- [4] Wikipedia. Industry[EB/OL]. (2022-04-29). <https://en.wikipedia.org/w/index.php?title=Industry&oldid=1085329087>.
- [5] 芮明杰. 产业经济学[M]. 2版. 上海: 上海财经大学出版社, 2012.
- [6] Turkenburg W C. The innovation chain: policies to promote energy innovations[M]. New York: The UN Publication, 2002.
- [7] Sen N. Innovation chain and CSIR[J]. *Current Science*, 2003, 85(5): 570-574.
- [8] Mowery D C, Nelson R R. Sources of industrial leadership: studies of seven industries[M]. Cambridge: Cambridge University Press, 1999.
- [9] Pavitt K. Technology, management and systems of innovation [M]. Cheltenham: Edward Elgar Publishing, 1999.
- [10] 张鹏, 雷家骕. 基于科学的产业发展模式研究——以心电图和石墨烯产业为例[J]. *科学学与科学技术管理*, 2015, 36(9): 40-53.
- [11] Xu H Y, Yue Z H, Pang H S, et al. Integrative model for discovering linked topics in science and technology[J]. *Journal of Informetrics*, 2022, 16(2): 101265.
- [12] 江泓. 科学革命、技术革命、产业革命的关系和发展阶段[J]. *历史教学*, 1992(6): 8-13.
- [13] 彭帅, 张春博, 杨阳, 等. 科学-技术-产业关联视角下石墨烯发展国际比较——基于专利的计量研究[J]. *中国科技论坛*, 2019(4): 181-188.
- [14] 李天柱, 刘小琴, 马佳, 等. 从科学预言到新兴产业: 纳米技术创新的启示[J]. *中国科技论坛*, 2015(6): 22-25.
- [15] Kenta I, Kazuyuki M, Ryuichi T, et al. Measuring science intensity of industry using linked dataset of science, technology and industry[R]. Tokyo: Research Institute of Economy, Trade and Industry, 2017.
- [16] 武华维, 王超, 许海云, 等. 知识耦合视角下区域科学—技术—产业协同创新水平的评价方法研究[J]. *情报理论与实践*, 2020, 43(5): 91-98, 8.
- [17] Han Y J. Linking science and technology with industry[J]. *Journal of Technology Management & Innovation*, 2007, 2(2): 7-17.
- [18] Park H W, Suh S H, Lee J T. Scientific and technological knowledge flow and technological innovation: quantitative approach using patent citation[C]// 2011 Proceedings of PICMET'11: Technology Management in the Energy Smart World. Piscataway: IEEE, 2011: 1-13.
- [19] Wong C Y, Fung H N. Science-technology-industry correlative indicators for policy targeting on emerging technologies: exploring the core competencies and promising industries of aspirant economies[J]. *Scientometrics*, 2017, 111(2): 841-867.
- [20] Holme P, Saramäki J. Temporal networks[J]. *Physics Reports*, 2012, 519(3): 97-125.
- [21] 楼凤丹, 周银座, 庄晓丹, 等. 时效网络结构及动力学研究进展综述[J]. *电子科技大学学报*, 2017, 46(1): 109-125.
- [22] 刘自强, 许海云, 罗瑞, 等. 基于主题关联分析的科技互动模式识别方法研究[J]. *情报学报*, 2019, 38(10): 997-1011.
- [23] 白如江, 冷伏海. k-clique社区知识创新演化方法研究[J]. *图书情报工作*, 2013, 57(17): 86-94.
- [24] Newman M E J. Fast algorithm for detecting community structure in networks[J]. *Physical Review E*, 2004, 69(6): 066133.
- [25] Dehmer M, Mowshowitz A. A history of graph entropy measures [J]. *Information Sciences*, 2011, 181(1): 57-78.
- [26] Cai M, Cui Y, Stanley H E. Analysis and evaluation of the entropy indices of a static network structure[J]. *Scientific Reports*, 2017, 7: Article No.9340.
- [27] 黄丽亚, 霍宥良, 王青, 等. 基于K-阶结构熵的网络异构性研究[J]. *物理学报*, 2019, 68(1): 325-336.
- [28] Almog A, Shmueli E. Structural entropy: monitoring correlation-based networks over time with application to financial markets [J]. *Scientific Reports*, 2019, 9: Article No.10832.
- [29] Gedikli A, Aksoy H, Unal N E. Segmentation algorithm for long time series analysis[J]. *Stochastic Environmental Research and Risk Assessment*, 2008, 22(3): 291-302.
- [30] Huang L, Chen X, Zhang Y, et al. Identification of topic evolution: network analytics with piecewise linear representation and word embedding[J]. *Scientometrics*, 2022, 127(9): 5353-5383.
- [31] Hallac D, Nystrup P, Boyd S. Greedy Gaussian segmentation of multivariate time series[J]. *Advances in Data Analysis and Classification*, 2019, 13(3): 727-751.
- [32] Schreiber T. Measuring information transfer[J]. *Physical Review Letters*, 2000, 85(2): 461-464.
- [33] Choi I, Kim W C. Detecting and analyzing politically-themed stocks using text mining techniques and transfer entropy—focus on the republic of Korea's case[J]. *Entropy*, 2021, 23(6): 734.
- [34] 蔡爽, 胡瑾秋, 张来斌. 基于传递熵与KELM的炼油化工过程风险传播路径分析方法[J]. *中国安全生产科学技术*, 2018, 14(3): 19-26.
- [35] Khan F H, Qamar U, Bashir S. SentiMI: introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection[J]. *Applied Soft Computing*, 2016, 39: 140-153.
- [36] 李健, 韩毅. 基于主路径的领域知识扩散演变研究——以太阳能领域为例[J]. *情报学报*, 2014, 33(8): 883-891.
- [37] Rajabzadeh N, Fathi E, Farahzadi R. Stem cell-based regenerative medicine[J]. *Stem Cell Investigation*, 2019, 6: 19.
- [38] The Bush Ranch. President discusses stem cell research[EB/OL]. (2001-08-09) [2023-01-01]. <https://georgewbush-whitehouse.archives.gov/news/releases/2001/08/20010809-2.html>.
- [39] Levine A D. Research policy and the mobility of US stem cell scientists[J]. *Nature Biotechnology*, 2006, 24(7): 865-866.
- [40] Bonnet D, Dick J E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell



- [J]. Nature Medicine, 1997, 3(7): 730-737.
- [41] Charitos I A, Ballini A, Cantore S, et al. Stem cells: a historical review about biological, religious, and ethical issues[J]. Stem Cells International, 2021, 2021: Article ID 9978837.
- [42] The Lens - free & open patent and scholarly search[EB/OL]. [2023-01-01]. <https://www.lens.org/lens>.
- [43] Bergman K, Graff G D. The global stem cell patent landscape: implications for efficient technology transfer and commercial development[J]. Nature Biotechnology, 2007, 25(4): 419-424.
- [44] Mathews D J H, Graff G D, Saha K, et al. Access to stem cells and data: persons, property rights, and scientific progress[J]. Science, 2011, 331(6018): 725-727.
- [45] Hernandez-Melchor D, Lopez-Bayghen E, Padilla-Viveros A. The patent landscape in the field of stem cell therapy: closing the gap between research and clinic[J]. F1000Research, 2022, 11: 997.
- [46] Hu Z Y, Pang H, Wei L, et al. Identifying research fronts based on scientific papers and patents using topic model: a case study on regenerative medicine[C]// Proceedings of the 7th Global Tech-Mining Conference, Atlanta, US, 2017.
- [47] 黄朝君, 贾建伟, 秦赫, 等. 基于 Copula 熵-随机森林的中长期径流预报研究[J]. 人民长江, 2021, 52(11): 81-85.
- [48] Ghawi R, Pfeffer J. A community matching based approach to measuring layer similarity in multilayer networks[J]. Social Networks, 2022, 68: 1-14.

(责任编辑 冯家琪)